UGZY/BY-07

BLOCK - 1

HEREDITY AND PHENOTYPE

Uttar Pradesh
Rajarshi Tandon Open University

UGZY/BY-07
Genetics

Block
1
HEREDITY AND PHENOTYPE

GENETICS

Introduction

We welcome you to the study of this course on genetics. Genetics is the science of heredity that relates to the study of the structure and function of genes and the mechanisms of their transmission from one generation to the next. The differences in many a characteristics among organisms are a direct result of differences in the genes they carry. These differences have resulted due to the evolutionary processes. *Mutations* – heritable changes in the genes, and *recombination* – the shuffling of genes and chromosomes at meiosis are the raw materials on which *selection* acts favouring a particular combination of genes for a given environment. The principles of heredity were first recognised by Gregor Mendel in 1860s but the systematic study of this discipline began in about 1900. And eversince the science of genetics has never stopped expanding. Its growth, however, has been very rapid and phenomenal in the past four decades eversince the discovery of the structure of the genetic material by Watson and Crick. One has to be really alert and hardworking to keep pace with advances made in this branch of science. It is generally said that scientific knowledge doubles every ten years, but as far as genetics is concerned the doubling time is just about two years.

Modern genetics is divided into three main branches: i) **transmission genetics,** ii) **molecular genetics,** and iii) **population genetics.**

i) **Transmission genetics** deals with the study of transmission of genes from one generation to the next. This first branch relies mainly on the same kind of experimental approach used by Gregor Mendel in the middle of the nineteenth century. Organisms with different traits (different phenotypes) are mated and transmission of these traits to the next generation is observed. Similarly, these progeny organisms can be mated with others of the same or different phenotypes, and again the transmission of traits can be observed. Since Mendel pioneered this approach of genetics, it is also frequently called **Mendelian** or **Classical genetics.** And as it deals with the transmission of genes to successive generations we also call it **Transmission genetics.**

ii) **Molecular genetics** is the study of the molecular structure of genes as well as the nature, expression and regulation of these molecules. Whereas transmission genetic studies are mainly carried out with eukaryotic organisms, viruses and bacteria are the source of study in molecular genetics. This branch of genetics paved the way for the development of genetic engineering and recombinant DNA studies which have far-reaching implications in the field of agriculture and medicine.

iii) **Population genetics** deals with the study of behaviour of genes in populations. More importantly population genetics is concerned with changes in the gene frequencies of populations over a period of time. Processes such as mutation, recombination, selection etc. tend to affect gene frequencies of populations. In fact population geneticists define evolution as changes in gene frequencies.

These branches are not exclusive of one another, rather, they reinforce each other. In this course you would find an integration of these three branches.

General aims of the course

This course aims at presenting the subject matter of genetics in the broader context of study of heredity. It examines not only the inheritance of genes which affect the characters of organisms but also the developmental processes whereby the characters are produced. The major questions that have concerned geneticists, in the past and the present, are highlighted. Some of the important questions that this course would address are: What is the genetic material? How it is packaged, where is it located in a cell? How does it express itself in a coordinated and organised manner? What are the sources of similarity and variability among individuals of a species? What is the behaviour of genes in populations? And in the present context, what is the relevance of genetic studies to humans and so on and so forth.

What you should already know : Genetics (LSE-03) is a second level course that you are going to study after completing the prerequisite Cell Biology (LSE-01) course. While developing this course, we have kept the CLASS XII genetics syllabus of NCERT as the baseline. We assume, that after the study of the above two courses, you are now well versed with the biological terminology related to genetics and cell biology. You may refer to the above two sources for your clarification of any basic concepts, if you think so.

General guidance for the study material

To make an effective use of the study material you may find the following points useful:

1) At the beginning of most units there is a study guide. This is designed to draw your attention to the major features and any special requirement for the text that follows.

2) All objectives for each unit are placed in the beginning of the unit, just after the introduction. It is important to understand that the objectives are a statement of what we hope that you can absorb. They are a guide to the topics that may be assessed in the assignments or final examination. Therefore, we suggest that you go through the objectives carefully when you start working on a unit.

3) Read the text slowly and attentively spending sufficient time on illustrations, photographs, the data in the tables, flowcharts etc.

4) You will find that all the units of the course are not of equal length. Some require more study time and others less. We suggest that you break up the longer units into three or four convenient portions to facilitate better learning.

5) Attempt all Self-assessment Questions (SAQs) as this will enable you to check your progress and see whether you have achieved the objectives set for the section to a satisfactory level. You may write your answers in the space provided in the unit itself.

6) Do not skip the terminal questions (TQs) particularly the problems given at the end of each unit. They would help you to get the concepts clarified and to diagnose any areas of weakness in your study. Make a separate note-book to write answers of the TQs that would help you in future to revise them.

7) Each unit contains a number of important ideas, concepts and terms and these are italicised or made bold when they are defined or introduced for the first time in the text. You should give special attention to them.

8) There are some Box items in the units, that provide information of special interest or some additional information related to that particular topic. You are however, not expected to memorise them for your examination.

9) Similarly, we do not expect you to remember and reproduce the results and data of all the genetic experiments included in different units. But nevertheless, examine them critically so that you can understand the concept better.

10) The glossary given at the end of each block explains the technical terms appearing in the units of that particular block.

11) 'Further Reading' at the end of the block gives you a list of books, some of which are available in your Study Centre library and in the market too. You may refer to these books in case you are interested in knowing more about a specific topic.

12) You are required to do two assignments (for the purpose of continuous assessment and feedback) at the end of this course – one TMA and one CMA. Each assignment has questions based on the units of all the four blocks. Do not wait till the time you complete your study of the entire course. It would be easier for you to write answers on the topics on which you have completed your study.

13) At the end of each block of this course you would find a detachable feedback form. Your feedback on each unit/block would enable us to know your difficulties and suggestions for improving the course material. Therefore, don't forget to send them back to us as soon as you complete reading each block.

We hope that you will enjoy studying this course.

BLOCK 1 HEREDITY AND PHENOTYPE

This is the first block of the course that contains six units dealing with the basic concepts of transmission of genetic material. Unit 1 describes the classic work of Gregor Mendel and the basic principles of Mendelian genetics that resulted from his work. A thorough understanding of these principles is necessary for dealing with more complex phenomenon presented later. Also introduced in this unit are problem solving and data analysis as standard methods in genetic study.

Mendelian analysis cannot be applied to all situations. The exceptions to, and the extensions of Mendelian analysis, such as the existence of multiple alleles, lethal genes, modification of dominance relationships, epistasis, sex linkage and many others are discussed in Unit 2.

Units 3 to 5 deal with various aspects of genetics of sex. Unit 3 elaborates the importance of sexual reproduction and the genetic basis of sex determination. Unit 4 explores further one of the major exceptions to Mendelian genetics – sex linkage. This is discussed by considering several examples in inheritance of genes located on the X and Y – the sex chromosomes. The dosage compensation mechanism that equalises the X-linked gene activity in both the sexes is also discussed. The genetic basis of development with respect to the differentiation of sexes is described in Unit 5. It explains the role of genetic information in early development in normal instances, and also reviews the occurrence of developmental mishaps leading to a path of anomalous development.

Unit 6 is about the classical genetic topic of linkage, crossing-over and genetic mapping. We deal with questions such as what is linkage and recombination? What are the principles of gene mapping? And how the order and the distance between genes in eukaryotic chromosomes is determined.

Objectives of the Block

After studying this block you should be able to:

- interpret and explain the results of mono-, di- and tri-hybrid crosses based on Mendel's laws of inheritance;
- recognise and describe giving examples the extensions and modifications of Mendelian genetic analysis;
- elaborate the mechanisms of sex determination and discuss the importance of sexual reproduction in evolution;
- describe sex linkage in terms of X-, and Y- linked genes, their transmission, and the mechanism of regulating dosage of sex-linked genes in both the sexes;
- discuss the genetic basis of development of sexes;
- illustrate and explain the concepts of linkage, crossing-over and genetic mapping.

Acknowledgements

I thank Dr. Amrita Nigam, Dr. Masood Parveez and especially Dr. V.V. Raghavan for their valuable comments.

UNIT 1 MENDEL'S LAWS OF INHERITANCE

Structure

1.1 INTRODUCTION

Genetics is the study of heredity. It is an ancient discipline. Atleast 4000 years ago, in Sumeria, Egypt and other parts of the world, farmers recognised that they could improve their crops and their animals by selective breeding. Their knowledge was based on experience and they recognised that many features of plants and animals were passed on from generation to generation. Furthermore, they were aware that desirable traits – such as size, speed and weight of animals could sometimes be combined by controlling mating, and that in plants, crop yield and resistance to arid conditions could be combined by cross pollination. The ancient breeding programmes were not based on experimental studies because nothing was known of genes or any of the principles of heredity.

In this first unit of the course, we start with a brief outline of the history of development of modern genetics and then we would discuss Mendel's laws : law of segregation and independent assortment.

Objectives

After studying this unit you would be able to:

- describe the historical developments and the landmarks in genetics (Sections 1.2 to 1.7);
- outline the contributions of Mendel to the science of genetics (Section 1.2);
- define and use correctly the terms gene, allele, locus, genotype, phenotype, dominance, recessiveness, homozygous, heterozygous and testcross (Section 1.3);
- describe Mendel's breeding experiments (Sections 1.4 to 1.6);
- state and explain Mendel's laws of inheritance (Sections 1.5 and 1.6);
- use various notations to represent and explain genetic crosses (Sections 1.5 to 1.7);
- apply Mendel's laws to solve problems in genetics involving monohybrid, dihybrid and trihybrid crosses (Sections 1.5 to 1.7).

Fig. 1.1: Gregor Johann Mendel – the father of the Science of Modern Genetics.

Born – 22 July, 1822 in Czechoslovakia.

Died – 6 January, 1884 in Brno (Czechoslovakia).

He was born in a small village and was the only son of a peasant family. Mendel had an insatiable curiosity about the natural and the physical world and was keenly interested in the diversity of living beings. His parents could not afford his higher education. He joined a monastery of St. Thomas, which was at that time in Austria and now in Czechoslovakia. His interest in Botany began early in life, as farming and the development of new varieties of apples were his family's chief occupation. His early interest was further stimulated by his formal education which centred round mathematics, physics, botany and zoology. The monastery provided him a stimulating environment, as it was a center of cultural, intellectual and religious life, and its members and visitors included many notable scholars and scientists of that period. In 1851, he joined the University of Vienna and upon completing his course he returned to his teaching responsibilities at Brno. His experiments in plant hybridization were

(Cont.)

Study Guide

As you go through this course, you will be introduced to several exercises on problem solving in genetics, some of which are worked out examples to explain the principles and concepts, and others are devices to reinforce or check your comprehension of the theory as well as its application to more practical, genetical situations. We hope you would find problem solving a challenging and rewarding job. Here are some tips that you may find useful in problem solving:

Firstly, analyse the problem carefully;
Secondly, translate words into symbols, defining each one;
and thirdly, solve the problem logically.

The first two steps are critical and if done carefully make the third step easier.

Table 1.1 in the subsection 1.2.2 lists the important breakthroughs in the field of genetics, compiled in a chronological order. You are not expected to memorise it at this stage as you may not be familiar with some of the terms used, but you must go through it carefully, more than once, to appreciate how this science developed over the years.

1.2 DEVELOPMENT OF GENETICS – A HISTORICAL PERSPECTIVE

As systematic study of genetics began, answers were sought to questions such as i) why are we like our parents? ii) why do organisms belonging to a group have an astonishing similarity to each other? The *first geneticists* were the persons who observed that "like begets like" and then used this basic observation to improve their strains. Domestication of animals like dogs, cats and cattle, was started when humans learned how to control the matings of these species. Once animals were tamed, people could breed more of those which they found most useful and the control of lineage remained with humans. Cultivation of crops required that seeds of one generation were saved to produce the next generation. Therefore, the seeds that were saved were from the largest and the healthiest plants.

Next, scientists started seeking answers to the question as to why we are like our parents but are also different from them in many respects. It was only after the work of Mendel was published in 1860s that answers to these questions started to emerge.

1.2.1 Birth of Genetics

Modern genetics originated with Gregor Mendel's (Fig. 1.1) work. It is based on this paper entitled *"Experiments in Plant Hybridisation"* published in 1866 in the Proceedings of the Society of Natural History in Brno. Mendel carried out detailed investigations of inheritance in garden pea. He performed elaborate plant hybridisation experiments and kept accurate pedigree records of his results. With the data obtained, he was able to formulate the basic principles of inheritance.

Mendel proposed the concept of hereditary units. According to him equal number of these units (factors) inherited from each parent determined the observable characters of the offspring. This was the first conceptualisation of what is now referred to *particulate* inheritance. Characteristics themselves are not inherited but the particles, units or factors that determine or control the observable traits are transmitted from parents to offspring. The appearance of the character in the offspring is determined by the particular combination of factors inherited from the two parents. This was the beginning of the concept of a gene, which is the modern term for the hereditary units or particles originally described by Mendel.

Mendel's work was not appreciated by the rest of the scientific community until 1900, when three botanists Carl Correns in Germany, Hugo de Varies in the Netherlands and Erich von Tschermak in Austria, rediscovered his work after each had independently reached similar conclusions. They all found Mendel's report while scanning the literature for related work and cited it in their own publications. William

Bateson (Fig. 1.3), an English scientist, coined the term *"genetics"* in 1905 for this developing science. The term was derived from Greek word which means *"to generate"*. Many consider Bateson as the real founder of genetics as he was the first to have Mendel's papers translated into English and the first one to show that Mendel's theory was also applicable to animals.

1.2.2 Growth of Genetics : From Mendel to Genetic Engineering

Genetics has come a long way and at present is a mature and dynamic science. The science of genetics was built on the foundation laid by Mendel but is owes its present stature to the contributions of a large number of scientists. In fact, the history and development of genetics is a subject worthy of study. You will be studying about the important contributions of various researchers in different units of this course. Nevertheless, a list of salient developments/contributions in the field has been compiled in Table 1.1 to give you some idea of their chronology.

Table 1.1 : The Genetics Time Line (modified from Gardner et al., 1991).

Year	Scientists (S)	Contribution
1865	Mendel	read his paper to the Brünn Society for Natural History
1866	Mendel	his paper published in the Proceedings of the Brünn Society for Natural History
1868	Miescher	first study of DNA
1900	De Vries, Correns, and Tschermak	Mendel's work discovered
1902	Boveri & Sutton	demonstrated the presence of paired chromosomes (homologs) in diploid species
1905	Bateson	named the science genetics
1908	Hardy & Weinberg	formulated the "Hardy Weinberg Law" relating genotypic frequencies to gene frequencies in randomly mating populations
1909	Johannsen	introduced the term gene
1909	Garrod	book on Inborn Errors of Metabolism
1910	Morgan	established the sex-linked inheritance of white eyes in *Drosophila melanogaster (Nobel Prize 1933)*
1911	Morgan	postulated the chromosomal basis of linkage
1913	Sturtevant	construction of a genetic map
1927	Muller	reported the use of the *CIB* technique to demostrate that X-rays are mutagenic *(Nobel Prize, 1946)*
1928	Griffith	discovery of transformation in *Diplococcus pneumoniae*
1931	Creighton & McClintock	demonstrated that genetic recombination is correlated with the exchange of morphological markers on chromosomes.
1940	Oliver	demonstration of recombination within the lozenge functional unit in *Drosophila*
1941	Beadle and Tatum	one gene one enzyme concept *(Nobel Prize 1958)*
1944	Avery, MacLeod & McCarty	demonstrated that the pneumococcal "transforming principle" is DNA
1946	Lederberg & Tatum	discovered conjugation in bacteria *(Nobel Prize 1958)*
1950	McClintock	first to present a paper on "Transposable elements" in Maize *(Nobel Prize 1983)*
1952	Hershey & Chase	demostrated that the genetic material of bacteriophage T2 is DNA *(Nobel Prize 1969)*
1952	Zinder & Lederberg	discovered the phage-mediated transduction in bacteria
1953	Watson & Crick	worked out the double-helix structure of DNA using the X-ray diffraction data of Wilkins and the base composition data of Chargaff *(Nobel Prize 1962)*

carried out in the monastery garden (see Fig. 1.2) for several years beginning from 1826. Mendel combined his talents, background and interests in a series of experiments with garden peas. His experiments are now recognised as classic example of carefully planned and executed scientific research.

Fig. 1.2 : Monastery garden where Mendel's experiments on peas were conducted.

Fig. 1.3 : William Bateson. (1861-1926)

1955	Benzer	first to present a paper on the fine structure of the Phage T4 rII locus
1956	Tjio & Levan	established that the normal diploid chromosome number in human is 46
1957	Fraenkel-Conrat & Singer	demonstrated that the genetic information of tobacco mosaic virus is stored in RNA
1958	Meselson & Stahl	demonstrated that DNA replication is semi-conservative
1958	Kornberg	isolated DNA polymerase I from *E.Coli (Nobel Prize 1959)*
1959	Ochoa	discovered RNA polymerase *(Nobel Prize 1959)*
1961	Jacob & Monod	proposed the "Operon Model" for regulating gene expression *(Nobel Prize 1965)*
1964	Yanofsky & colleagues; Brenner & colleagues	established colinearity between polypeptide products
1964	Temin	proposed the DNA provirus from RNA tumour viruses *(Nobel Prize 1975)*
1965	Holley	worked out the first complete nucleotide sequence of a tRNA *(Nobel Prize 1968)*
1966	Nirenberg, Khorana & coworkers	established the complete genetic code *(Nobel Prize 1968)*
1970	Nathans & Smith	isolated the first restriction endonuclease *(Nobel Prize 1978)*
1970	Baltimore	identified reverse transcriptase of RNA tumour viruses *(Nobel Prize 1975)*
1972	Berg	produced first recombinant DNA *(Nobel Prize 1980)*
1976	Bishop & Varmus	demonstrated the protooncogene to oncogene relationship *prize 1989)*
1976	Hozumi & Tonigawa	demonstrated somatic rearrangements of gene encoding antibodies
1977	Breathnach, Mandel & Chambon; Jeffreys & Flavell	demonstrated the presence of introns in eukaryotic genes
1977	Maxam & Gilbert; Sanger, Nicklen & Coulson	description of the DNA sequencing techniques
1977	Sanger & colleagues	worked out the complete 5387 nucleotide sequence of phage ϕX174
1978	Three different laboratories	discovered "splicing" of adenovirus RNAs
1982	Sanger & colleagues	worked out the complete 48,502 nucleotide-pair sequence of phage lambda
1983	Cech & Altman	established the existence of catalytic RNAs *(Nobel Prize 1989)*
1985	Jeffery	DNA finger-printing
1988	Watson	coordinated the "human genome-project"
1989	NIH Recombinant DNA Advisory Committee	recommended approval of first-human "gene transplant" experiment
1989	Tsui, Collins & colleagues	colored the "cystic fibrosis gene"
1990	Saiki	Polymerase chain reaction

The above table, however, is not complete as we intended to make in concise. Presently, new information is accumulating at an unprecedented rate. Most of the important contributions were made between 1970s and 1980s. The next few years are going to be exciting as many scientists are working out details of the molecular mechanisms responsible for differentiation in eukaryotes that includes humans as well. Several challenging areas in the field are open for research, and perhaps one day we would be able to understand the molecular basis of complex phenomena like learning, memory, cancer and aging.

1.3 SOME BASIC GENETIC TERMINOLOGY

Before we move on to discuss in detail Mendel's experiments with pea and the laws of inheritance formulated by him, let us first understand some basic genetic terminology.

1.3.1 Gene

In subsection 1.2.1 you have studied that Mendel proposed the concept of hereditary units, which were later on identified as genes. We shall elaborate on it further. Let us consider different traits such as flower colour, seed shape or height in plants, each of these characters is controlled by a different gene.

1.3.2 Genes and Alleles

The inheritance of any character can be studied only when there are two contrasting conditions, such as yellow versus green seed colour (as observed by Mendel in peas), normal pigmentation versus absence of pigmentation *(albinism)* in humans and other animals, and brown versus black coat colour in guinea pigs. An individual expresses one or the other, but not both contrasting conditions at the same time. *Genes that govern variations of the same characteristic and that occupy corresponding loci on homologous chromosomes are termed* **alleles** (see Fig. 1.4).

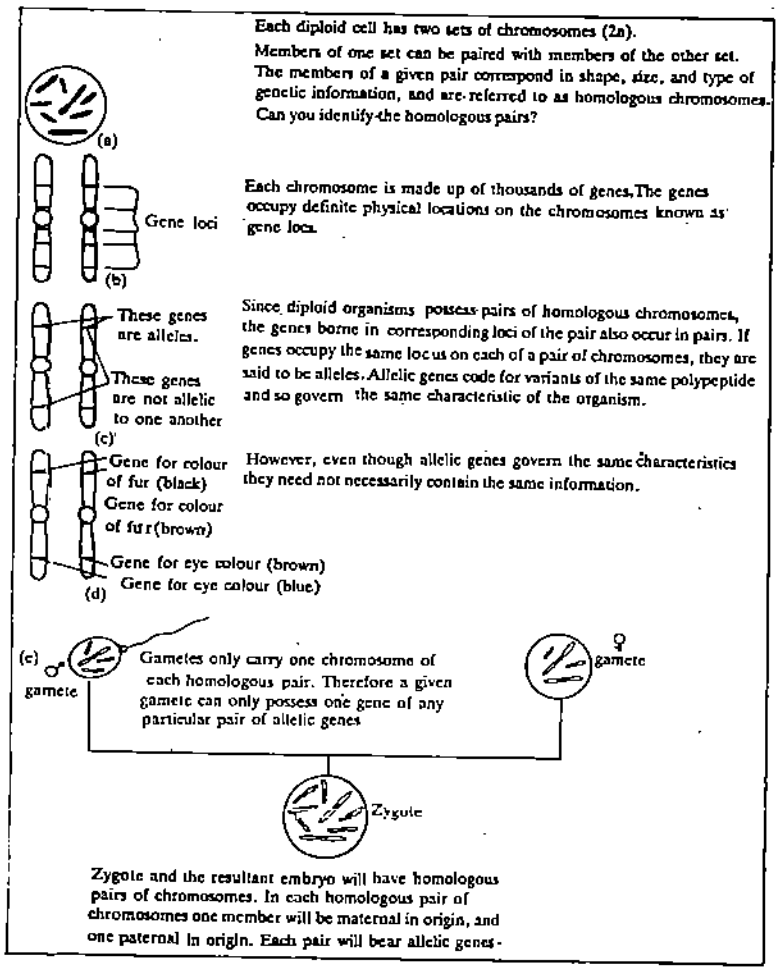


Fig. 1.4 : Homologous chromosomes and alleles.

Geneticists use the term *allele* to emphasise that there are two or more alternative forms of a gene that can occur at corresponding specific loci. The possible variants of a gene at any given locus are known as alleles, each of which is assigned a single letter (or a group of letters) as its symbol. In the example below, we shall consider the trait – height of pea plant. Tall forms are denoted by T and the short ones are denoted by t. T and t are alleles of the same gene. Since they are present in pairs they are represented as TT, tt or Tt.

It is customary to indicate a *dominant* allele with a capital letter and a *recessive* allele with a lower case letter. The choice of the letter or letters themselves is generally determined by the first allelic variant in a particular locus. For example, Mendel studied the alleles that govern seed colour – yellow versus green. The allele that is responsible for the yellow colour of the seed is designated as Y, and the allele for the green colour y. Because discovery of the yellow allele made identification of this locus possible, we commonly refer to the locus as the *yellow locus*, although pea seeds are most commonly green. The term locus is used to designate not only a location on a chromosome but also a kind of "generic" gene controlling a particular *kind* of characteristic; thus Y (yellow) and y (green) represent a specific pair of alleles of a locus involved in determining seed colour in peas. To prevent confusion, the dominant allele is always designated first and the recessive allele second (Yy, never yY). Geneticists use the term *gene* sometimes to specify a locus and sometimes to specify one of the allelic variants at that locus. Usually the meaning is clear from the context.

Mendel's particulate hereditary elements or factors are called genes.

There are many genes within each chromosome, each generally different from the other and each controlling the inheritance of one or more characteristic(s). The *members of a homologous pair of chromosome have similar set of alleles arranged in the same order.* The regularity of the mitotic process ensures that each diploid daughter cell receives a pair of each chromosome and therefore a pair of each gene. As the chromosomes separate during meiosis, and become associated with new partners at the time of fertilisation the alleles also separate and associate with new partners. The entire concept of Mendelian Genetics depends on these simple facts.

1.3.3 Homozygous and Heterozygous

Homozygous: *Same* factor.
Heterozygous: *Different* factors.

In an individual two identical alleles may exist for a given character and, hence, the individual is referred to as *Homozygous* (e.g. AA and aa). If there are two non-identical or different alleles for a given character, the individual is referred to as *heterozygous* (e.g. Aa).

Let us examine a situation where both the parents are homozygous. The male is homozygous recessive aa, and female is homozygous dominant AA. During meiosis in the male the two 'a' alleles separate from each other so that each sperm has only a single 'a' allele. Similarly, in the female parent each egg has one 'A' allele. The fertilisation of the 'A' egg by 'a' sperm results in a heterozygous animal with 'Aa'.

The form of gene which occurs in an individual in nature is called the **'wild type'** while a **'mutant type'** is the one in which the genetic material is somewhat altered. Mutants arise due to various reasons. You will study the topic of mutation in details in Unit 16.

The alleles which express themselves in both homozygous and heterozygous conditions are known as the **dominant alleles**. For example, TT represents tallness (homozygous). The individuals having the alleles TT, and Tt would be tall as T is a **dominant** allele and it can express itself in both homozygous and heterozygous conditions.

Some alleles express themselves only in homozygous condition and are referred to as **recessive alleles**. In the height characteristic, dwarfness can be seen in individuals that have the alleles 'tt', i.e., **recessive alleles** are expressed in homozygous condition only.

1.3.4 Phenotype and Genotype

From the above subsection (1.3.3) it is clear that some alleles are dominant while others are recessive. This indicates that we *cannot always* deduce the alleles present

in an organism by just looking at it. The **phenotype** is the term used to specify the appearance of an individual in a given environment with respect to a certain inherited trait. The *genetic constitution* of that organism, most often expressed in symbols, is its **genotype**.

Physical appearance : Phenotype.

Genetic constitution : Genotype

Using the same example as above let us understand the concepts of phenotype and genotype. If the female parent has the genotype TT (homozygous, dominant) can you guess the phenotype? Yes! It is Tall. If the male parent is homozygous recessive 'tt' can you tell what is its phenotype? The answer is dwarf. If these two parents are crossed then what would be the genotype and phenotype of the offspring? The genotype would be 'Tt' and phenotype would be tall.

Similarly, an individual has alleles that govern hair colour. These specific alleles constitute the individual's genotype for hair colour. The actual hair colour that the individual exhibits is its phenotype (black, brown, red, blonde etc.). The crux of the above discussion is that the *genetic constitution* of an individual is referred to as its **genotype** and the expression of these genes as its **phenotype.**

Phenotype should be used in the broadest possible sense. Phenotype commonly refers to visible characteristics such as size, shape, colour of the parts of an organism, height, weight, length of limbs and so on. But other less obvious characteristics also comprise the phenotype: behavioural traits (twitches, yawns, posture, smiles per hour), physiology (heart rate, blood pressure, level of basal metabolism) and biochemical characteristics (cholesterol level, presence or absence of particular enzymes, blood type). In fact, any trait of an organism which can be described and/or measured is the phenotype of that character.

We hope that you have clearly understood the above basic terms. You would learn about many more terms at relevant places in the following units.

SAQ 1

From the list of the items given below, select all those which are : i) phenotype, ii) genotype, iii) homozygous, dominant iv) homozygous, recessive, and v) heterozygous. Write your answers in the space provided.

a) Rr _____
b) TT _____
c) Tt _____
d) tall _____
e) RR _____
f) Yy _____
g) rr _____
h) terminal flower _____
i) tt _____

SAQ 2

Make appropriate matches of the terms mentioned in column A with the explanation in column B. Write your answers in the boxes.

A		B
a) Genes	☐	i) any character of an organism that can be measured or described
b) Phenotype	☐	ii) different alleles for specific traits on homologous chromosomes
c) Chromosomes	☐	iii) contain hereditary units
d) Heterozygote	☐	iv) once called the inheritance "factors" by Mendel

1.4 MENDEL'S CLASSICAL EXPERIMENTS WITH PEA

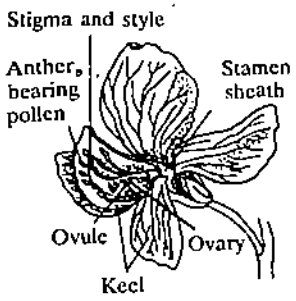


Fig. 1.5 : A pea flower that is partially cut to show the reproductive parts.

Mendel's studies provide an outstanding example of good scientific technique. He chose research material well suited for the study of the problem at hand, designed his experiments carefully, collected large amount of data and used mathematical analysis to show that the results were consistent with his explanatory hypothesis.

1.4.1 Choice of Material

Before Mendel, several investigators carried out research work to understand the principles of inheritance, but they failed to reach meaningful conclusions because of the unsuitability of the system they were studying. The garden pea with which Mendel worked has several suitable features. The foremost were that these plants were cheap, easy to obtain, required little space, had shorter generation time, produced many offspring and could be crossed easily. The pea flowers are bisexual and are usually self-fertilised, that is, the ovule (female gamete) is fertilised by pollen (male gamete) from the same flower as both the male and female parts of the flower are closed in a petal box or keel (see Fig. 1.5). Pollen from another plant can be experimentally introduced to the stigma of a flower to bring about cross-pollination.

Cross-pollination can be encouraged experimentally either by removing stamens from female parent and placing pollen from another variety on the stigma of its flowers (Fig. 1.6a), or by placing the pollen from the male parent on the stigma of a different plant, whose stamens have been removed (Fig. 1.6b). Therefore, with the common pea plant, the geneticists can perform crosses in the way they choose and can easily establish lineage or pedigree of each plant in a particular cross.

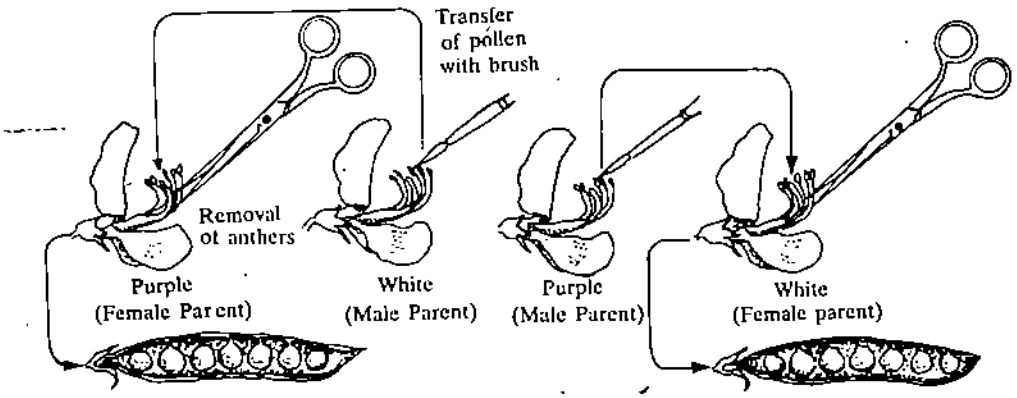


Fig. 1.6 (a,b) : Cross pollination in pea. In a) the anthers are removed in the purple flowers to prevent self-pollination. Similarly, in b) anthers are removed in the white flowers and it is usd for cross-pollination.

The pea plants varied with respect to a number of characteristics such as plant height, seed texture, used colour and flower colour. Such a variation is essential if anything at all is to be learned about the inheritance of any character. If, for example, all pea plants were of the same height and had the same flower colour, generation after generation, no information would be gained from following plant height and flower colour in genetic studies. A characteristic must have alternative traits or variant forms that can be followed if insight is to be gained regarding the inheritance pattern.

We have just discussed Mendel's choice of material for experimentation. You will find an explanation of Mendel's experiments and how he deduced the laws of inheritance from them in the following sections.

1.5 MONOHYBRID CROSSES AND MENDEL'S LAW OF SEGREGATION

Before we discuss Mendel's experiment let us clarify the terminology encountered in breeding experiments. The Parental generation is called the **P** or **P_1 generation.** The

progeny produced from the mating of individuals in the parental generation is called the **first filial generation** or F_1. The subsequent generation produced by breeding of the F_1 offspring is termed the **F_2 generation**. Selfing the offspring of each generation results in F_3, F_4, F_5 generations and so on.

1.5.1 Mendel's Experiments and Results

Mendel undertook the study of seven characters of pea plants and their variant forms as shown in Fig. 1.7. Each traits was studied individually such as tall versus dwarf, red flower colour versus white, yellow seeds versus green and so on. He performed **monohybrid crosses** —*crosses in which only one pair of contrasting or alternative trait is involved.* A cross is a mating between two individuals leading to the fusion of gametes. For example, to study the inheritance pattern of height, he crossed true breeding tall plants with dwarf plants. True breeding lines are those plants which produce progeny exactly like themselves when allowed to self pollinate normally. Thus, he obtained true breeding tall (TT) and dwarf (tt) plants. Using these true breeding lines which differed in only one pair of contrasting traits Mendel made monohybrid crosses, homozygous tall and dwarf plants (P_1 generation) were crossed to give offspring (F_1) which were all tall (see Fig. 1.8). The same results were obtained whether he used the tall plant as male or female parent, i.e.,

Tall female (♀) × dwarf male (♂) ------ All offspring tall
Dwarf female (♀) × tall male (♂) ------ All offspring tall

Such a cross is also called a **reciprocal cross.**

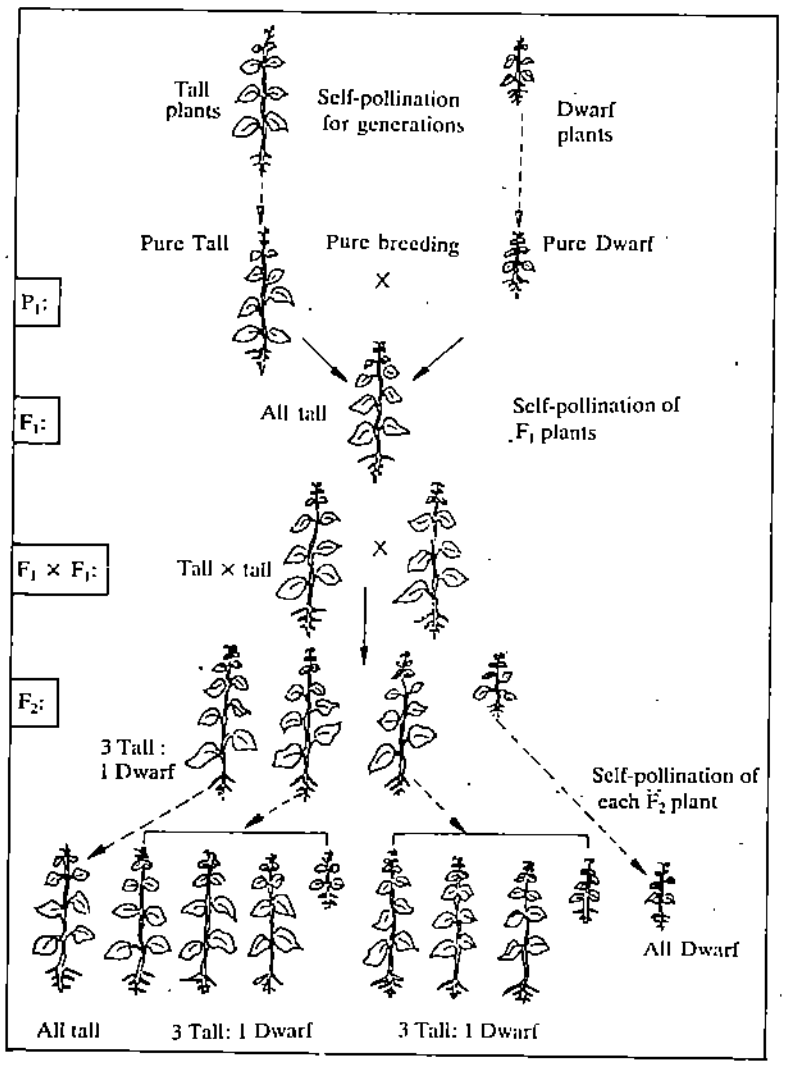


Fig. 1.8 : Monohybrid cross in the pea plant followed through three generations. The results of the reciprocal cross, i.e., dwarf × tall are identical.

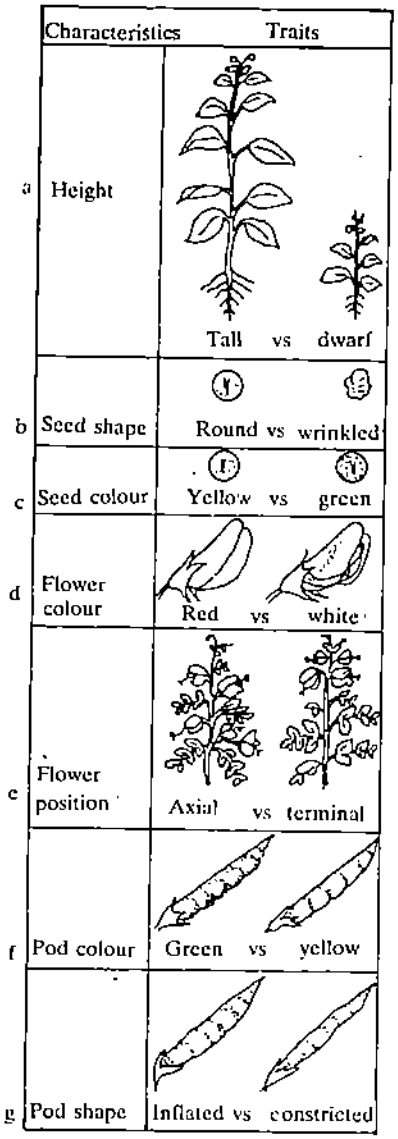


Fig. 1.7: The seven characteristics (a-g) in pea plant studied by Mendel. Each character has two well defined phenotypes that are easily recognised.

Since a cross was made between a pure breeding tall and a pure breeding dwarf parent, Mendel surmised that these F_1 tall individuals were not the same as the P_1 tall parent because they had alleles from the dwarf parent as well. He allowed these

individuals to self pollinate and studied the F_2 generation (see Fig. 1.8 again). In one cross he actually obtained 787 tall and 277 short offspring. The ratio comes out to be 3:1. He further self pollinated (selfed) the F_2 offspring and found that the short plants continued to produce only short one (see Fig. 1.8). On the other hand the tall plants were of two types: approximately one third of them produced only tall offspring upon self fertilisation; the remaining two thirds gave rise to both tall and short, again a ratio of approximately 3:1.

Using the same quantitative approach Mendel analysed the behaviour of the six other pairs of traits. The results of such experiments were very similar to the ones obtained with reference to the height of the plant (see Table 1.2).

Table 1.2 : A Summary of Mendel's Results from the Seven Different Monohybrid crosses (modified from Ayala, F.J. and Kiger, J.A., 1984).

Character	The Phenotype of F_1 Individuals	Number of F_2 Individuals		
		Dominant	Recessive	Total
Seeds: smooth versus wrinkled	All smooth	5474	1850	7324
Seeds: yellow versus green	All yellow	6022	2001	8023
Flowers: purple versus white	All purple	705	224	929
Flowers: axial versus terminal	All axial	651	207	858
Pods: inflated versus pinched	All inflated	882	299	1181
Pods: green versus yellow	All green	428	152	580
Stem: tall versus short	All tall	787	277	1064
Total or average		14,949	5,010	19,959

1.5.2 Mendel's Explanation and Derivation of the Law of Segregation

Mendel questioned as to how a trait present in the P_1 generation could disappear in the F_1 and then reappear with full expression in the F_2 generation. Further, the F_1 progeny resembled only one of the parents in their phenotype and did not breed true. The F_1 progeny possessed the potential to produce F_2 progeny that resembled the parents of P_1 generation. Mendel concluded that each one of the alternate traits, the tallness and the dwarfness, was determined by a particulate factor. He reasoned that these factors which were transmitted from parents to progeny through the gemetes, carried hereditary information. At present these factors are known as **Genes.** Since Mendel was examining pairs of contrasting traits each factor was considered to exist in alternate forms (which are known as alleles). Each of these alternate factors or alleles specifies a trait. For example, for the gene that controls the height of pea plant, there is one form or allele that results in the production of a tall plant and another allele that results in dwarf plant.

Mendel reasoned further that for pure breeding strains of the peas both egg and pollen must have carried or transmitted forms of the factor. Since both traits were seen in the F_2 whereas only one appeared in the F_1, each individual must have contained both factors, one for each of the alternate traits. Furthermore, since only one of the characters was seen in the F_1's, the expression of the other trait must somehow have been masked by the other factor, a feature called *dominance*. For the tall × dwarf cross the F_1 plants were all tall. Thus, the allele for tallness is **dominant** to the allele for dwarfness. Conversely dwarfness is said to be **recessive** to tallness. Similar conclusions can be made for the other six pairs of characters. The dominant and recessive forms for each pair of traits are already indicated in Table 1.2.

We can understand the crosses more easily if we use symbols instead of pictures or Figures (as in Fig. 1.8) for the alleles. In a cross between tall and dwarf plants, symbol T denotes the factor for tallness and the symbol t for dwarfness. You already know the convention that the dominant allele is denoted by the capital letters and the recessive allele is denoted by lowercase or small letters. Using these symbols we can diagram the cross as shown in Fig. 1.9.

Since each parent is true breeding and each contains two copies of the same allele, thus, the genotype of the parental plant grown from the tall seeds is 'TT' and that of the dwarf plant is 'tt'. You may recall that the true-breeding individuals always carry a pair of same alleles, that is, they are homozygous genotypes.

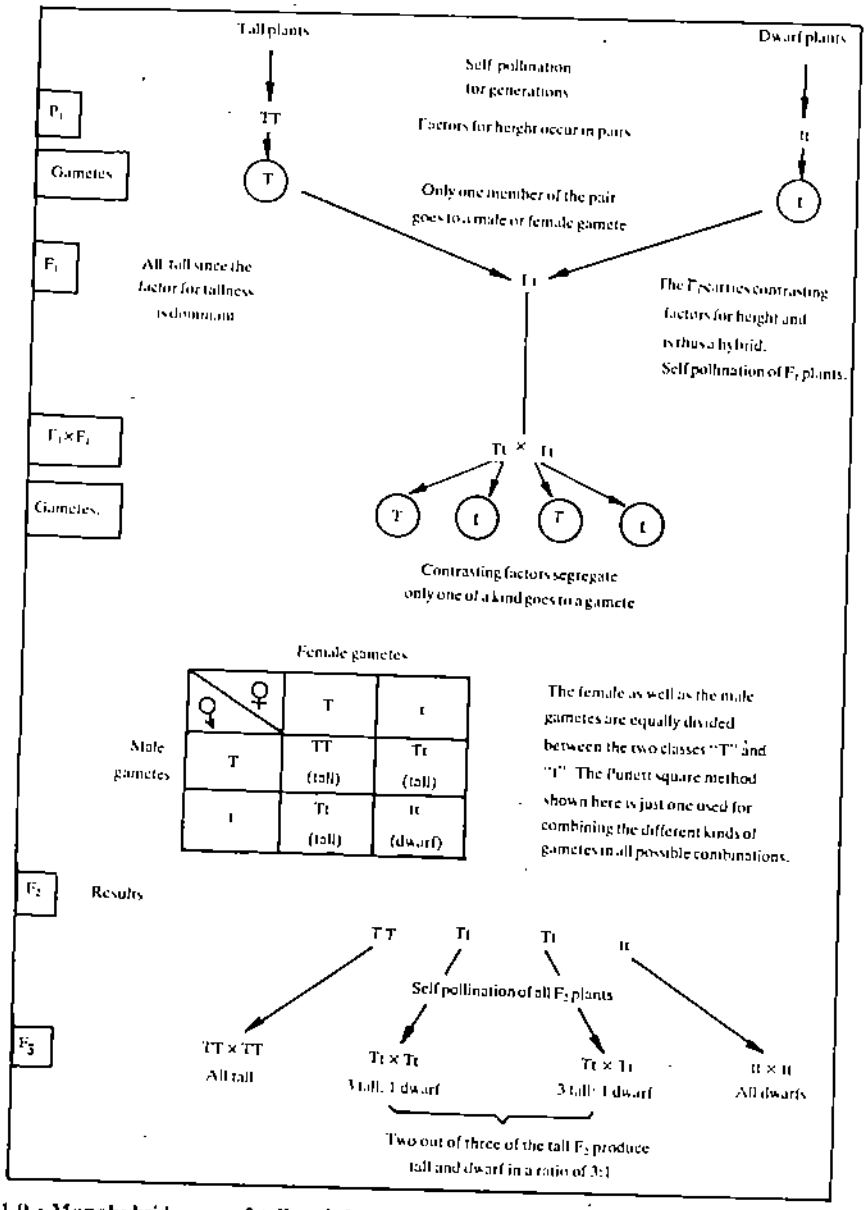


Fig. 1.9 : Monohybrid cross of tall and dwarf pea plants. The results of the reciprocal cross are the same. Compare with Fig. 1.8.

When these plants produce gametes by meiosis, each gamete contains only one copy of the gene (i.e., one allele), the plants from tall parents produce T-bearing gametes, and the dwarf plants produce t-bearing gametes. When these gametes fuse during the process of fertilisation, the resulting zygote has both T and t factors, i.e., a genotype 'Tt' (see Fig. 1.9). What are such plants that have two different alleles for a specific trait known as? Yes, it is heterozygous. Because of the dominance of the 'T' allele all the plants developed from the F_1 zygotes are phenotypically tall.

The plants derived from the F_1 generation differ from the tall parent in that they produce two types of gametes in equal numbers: T– and t– bearing. All the possible F_1 gamete combinations are represented in the matrix in Fig. 1.9 called as a **Punnett square** after its originator R. Punnett. These gametic fusions/combinations give rise to the zygotes that produce the F_2 generation.

Three types of zygotes are produced : TT, Tt and tt. The relative proportion of these zygotes is 1:2:1 respectively, and the F_2 generation exhibits a phenotypic ratio of 3:1 (3 tall, 1 dwarf).

The results are the same for crosses involving the other six character pairs.

1.5.3 The Law of Segregation

Based on the results of his monohybrid cross, Mendel proposed his *first law*, known as the *law of segregation*. It states that the members of a gene pair (alleles) segregate

(separate) from each other during the formation of gametes. As a result, half the gametes carry one allele and the other half carry the alternate allele. The progeny are produced by the random combination of gametes from the two parents. In proposing the law of segregation, Mendel clearly differentiated between the factors (genes) that determined the traits (the genotype) and the traits themselves (the phenotype). From modern perspective, this means that at the gene level the members of a pair of alleles of a gene on a chromosome, segregate during meiosis, so that any offspring receives only one member of the pair from each parent. This gene segregation parallels the separation of members of pairs of homologous chromosomes at anaphase I in meiosis.

To sum up, Mendel's first law states that the members of a gene pair (allele) segregate (separate) from each other while forming gametes; half the gametes carry one allele, and the oth.. - half carry the other allele.

Representing Crosses Using a Branch Diagram or Forked-line Method

The use of Punnett square to consider the pairing of all possible gametic types from the two parents (shown in Fig. 1.9) is an acceptable way of predicting the ratios of genotypes in the next generation. There is, however, another method — known as the branch diagram or the forked-line method for representing the crosses. We require the application of our mathematical knowledge here.

Do you remember the basic principles of probability? If you are not very sure about it, then you may refer to the course Mathematical Methods (MTE-03), Block 3, Unit 12. Nevertheless, for your ready reference we give below some basics of probability in Box 1.1.

Coming back to the branch diagram method of representing a cross, we have illustrated one such cross in Fig. 1.10. This diagram shows the selfing of $F_1 \times F_1$ of tall plants.

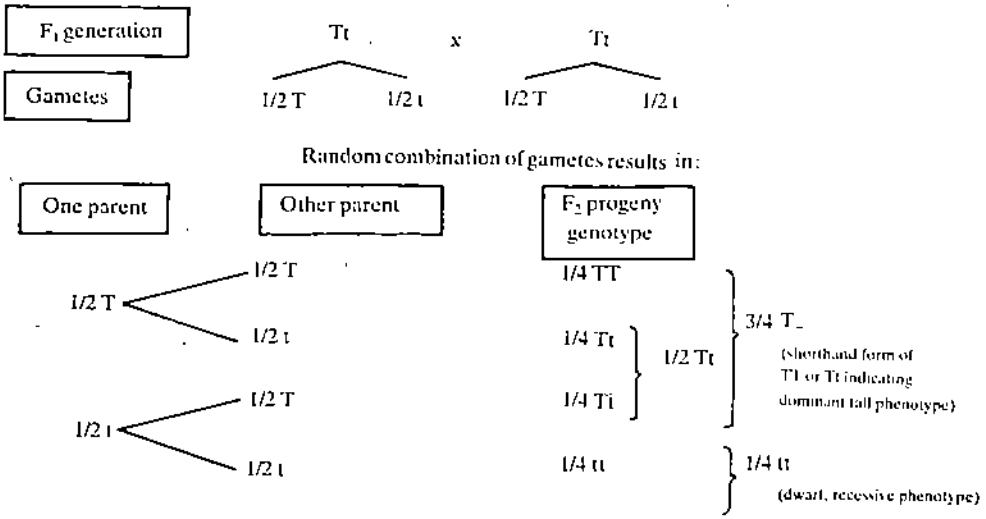


Fig. 1.10 : Representing the cross as shown in Fig. 1.8 and 1.9 by a branch diagram method. This method helps in calculating the ratios of phenotypes of the F_2 generation in an easy manner.

The F_1 plants have the genotype 'Tt'. Both eggs and pollen are produced in the flowers of the plants grown from these seeds. In the meiotic process an equal number of T and t gametes are expected to be produced so we can say that half the gametes are 'T' and the other half are 't'. Thus, ½ is the predicted frequency of these two types but just as tossing a coin many times does not always give exactly half heads and half tails, the gametic frequency may not be exactly realised. However, more the number of chances (e.g., tosses) the more likely you will obtain the true frequency.

From the rules of probability the relative expected frequencies of the three possible genotypes in the F_2 generation can be predicted. To produce a TT plant, an egg with T must fuse with a pollen grain carrying T. The frequency of T carrying eggs is ½ and the same holds good for the pollen grains also. Therefore, the relative expected proportion of 'TT' plants in the F_2 is $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. A similar proportion is expected for dwarf or 'tt' progeny.

What about the Tt progeny? Again the frequency of 'T' in one gametic type is ½ and the frequency of t in the other gametic type is also ½. There are, however, two ways in which Tt progeny can be obtained. The first involves the fusion of an egg with T and a pollen with t, and the second involves the fusion of an egg with t and a pollen with T. Applying the product rule (see Box 1.1) the probability of one or the other occurring is the sum of the individual probabilities, or ¼ + ¼ = ½.

From the above discussion, prediction is made that ¼ of the F_2 progeny would be TT, ½ would be Tt and ¼ will be tt. These figures are exactly similar to those found with the method shown in Fig. 1.10. Therefore, you can apply either of these methods for any monohybrid cross.

Box 1.1 : Some Basics of Probability

A probability is the ratio of the number of times a particular event occurs to the number of trials during which the event could have happened. For example, the probability of picking a heart (from 13 hearts) from a deck of cards (52 cards) is p (heart) = 13/52 = ¼. That is, we would expect, on the average to pick a heart from a deck of cards once in every four trials.

Probability and the *laws of chance* are involved in the transmission of genes. As a simple example, let's consider a couple and the chance that their child will be a boy or girl. Let us assume that an equal number of boys and girls are born. The probability that the child be a boy is ½ or 0.5. Similarly, the probability that the child will be a girl is ½.

Let us see some rules of probability. Firstly, we take up the **product rule**. This rule states that the probability of two independent events occurring simultaneously is the product of each of their probabilities. Thus, the probability that a family will have two girls in a row is ¼. That is, the probability of the first child being a girl is ½, the probability of the second being a girl is also ½, and the product rule of the probability of the first and the second being girl is ½ × ½ or ¼. Similarly, the probability of having three boys in a row is ½ × ½ × ½ = ⅛.

The **sum rule** of probability states that the probability, that one of two mutually exclusive events will occur is the sum of their individual probabilities. For example, if two dice are thrown, what is the probability of getting two sixes or two ones? The probability of getting two sixes is found by using the product rule. The probability of getting one six, p (one six) is ⅙, since there are six faces to a dice. Therefore, the probability of getting two sixes, p (two sixes), when two dice are thrown is ⅙ × ⅙ = 1/36. Similarly, p (two ones) = 1/36. To roll two sixes or two ones involves mutually exclusive events, so the sum rule is used to find the probability. The answer is 1/36 + 1/36 = 1/18.

1.5.4 Confirmation of the Law of Segregation : The Use of Test Crosses

Mendel did a number of tests to ensure the validity of his results when formulating his law of segregation. He continued the self-fertilisations to the F_6 generation and found that in every generation both the dominant and recessive characters appeared. He concluded that the law of segregation was valid no matter for how many generations it was carried out.

In another experiment, instead of allowing the F_1 heterozygous individuals to self-fertilise, he crossed them back to the recessive parent. Such a cross to a recessive is called a **test cross** or some times a **back cross** (see the adjacent Margin Remark) because it is used to test whether a dominant individual is heterozygous or homozygous. For example, let us consider the case of smooth and wrinkled seeds to understand it more clearly. The smooth (S) character is dominant over the wrinkled (s). In the Fig. 1.11 see the two crosses (a & b).

When a plant is crossed with any of its parents, it is called a backcross. And when a plant is specifically crossed with its homozygous, recessive parent, then it would be called a testcross or test backcross

In the Fig. 1.11 see cross a. Here the gametes produced by parent 1 are of two types S and s, whereas the second parent produces only one kind of gametes, i.e., s. As a result of random fusion of gametes, half the progeny of the test cross are Ss

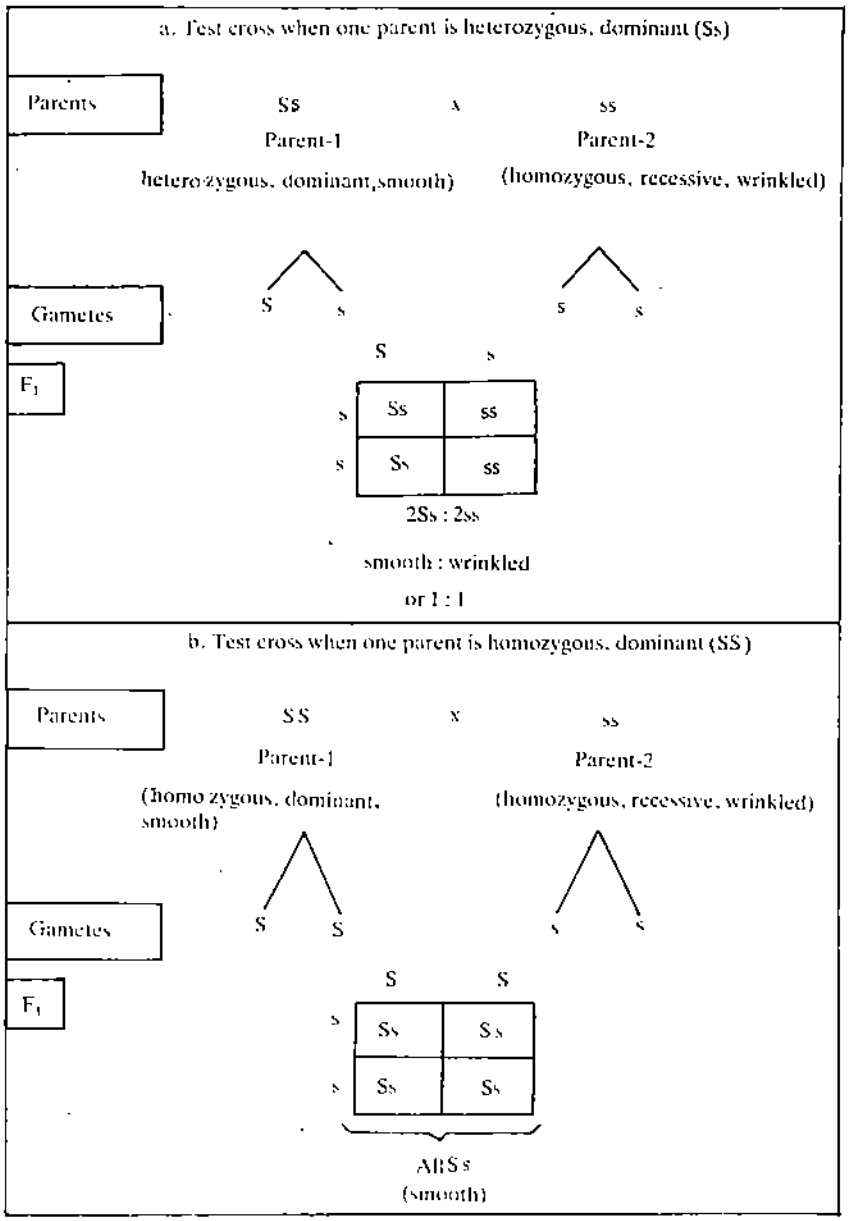


Fig. 1.11 : (a & b) Cross with a recessive parent and to confirm the law of segregation.

heterozygotes having a smooth phenotype. Owing to the dominance of the S allele, and the other half are ss homozygotes and have a wrinkled phenotype. Test crosses are useful in determining the genotype of F_1 individuals as well as their parents.

The ratio 1:1 of dominant to recessive phenotypes would indicate that the F_1 smooth seeds used in the test cross are heterozygous. We can similarly predict the outcome of a test cross of a homozygous dominant (see Fig 1.11, cross b). The first parent produces only one kind of gametes, that is, S, and so does the other, producing only s gametes. Thus, all the progeny are Ss and have smooth seed phenotype.

Therefore, we can conclude that if an organism in a test cross gives rise to progeny with the dominant phenotype, then the test-crossed plant must have been homozygous dominant (as shown in cross b above). For the Ss plant a 1:1 ratio of dominant to recessive phenotypes is expected as is seen in test cross 'a'.

The above discussion highlights the fact that the phenotypes of the progeny of the test cross indicates the genotype of the individual tested. If all the progeny show dominant phenotype, the individual tested is homozygous dominant. If there is an approximately 1:1 ratio of progeny with dominant and recessive phenotypes, then the individual is heterozygous.

1.5.5 Reasons for Mendel's Success

Mendel's success with the hybridisation experiments was mainly due to the following reasons :

First, careful choice and selection of experimental material, that is, garden pea. Besides several advantages it offers as an experimental material as described in subsection 1.4.1, it must be mentioned that it is a diploid organism having two sets of chromosomes. Through many generations of natural self fertilisation, garden peas had developed into pure lines, each of which could be differentiated by its own characteristic features. Furthermore, in the seven pairs of contrasting traits, Mendel chose to study the two parental plants exhibiting well-defined, visible, contrasting morphological traits. The factors responsible for the expression of the seven pairs chosen for study are located on seven different chromosomes which was a fortunate coincidence and this avoided the phenomenon of linkage.

Second, much of Mendel's success in his experiments may be attributed to his good judgement in making crosses between the parents that differed in only one trait. In other words, he considered only one trait at a time when analysing the progeny of a cross.

Thirdly, he methodically documented the results of his experiments. Fourthly, by the application of mathematics he could analyse the segregating populations and make proper interpretations.

SAQ 3

Pollen from a tall pea plant was dusted on the stigma of a flower from a dwarf plant. What would be the reciprocal cross?

.....

.....

SAQ 4

Put a tick mark (✓) on the correct choice:

a) The offspring of mating between two pure strains are called:
   i) hybrids
   ii) mutants
   iii) the P_1 generation
   iv) the F_2 generation

b) An organism with two copies of the same alleles is:
   i) homozygous for the trait
   ii) homologous for the allele
   iii) heterozygous for the trait
   iv) heterologous for the allele

c) A woman without dimples marries a man who has dimples and who is known to be heterozygous for the trait. What is the chance their first child will have dimples?
   i) one in four
   ii) one in two
   iii) three out of four
   iv) it is certain

d) A test cross is done to find out:
   i) the genotype of an individual by examining the phenotypes of its offspring from a particular mating
   ii) the genotype of an individual by testing for its DNA content
   iii) whether a mating is fertile
   iv) whether two species can inter-breed

e) A test cross distinguishes between:
   i) two homozygous forms
   ii) a homogygous dominant and the heterozygous form
   iii) two heterozygous forms
   iv) a homozygous recessive and a heterozygous form

f) The allele for black hair colour (B) is dominant over the allele for white hair colour (b) in guinea pigs. A test cross between a black male and a white female produced a litter of five black and one white guinea pigs. The genotype of the father is:
   i) unknown, due to the small sample size
   ii) BB
   iii) Bb
   iv) bb

1.6 DIHYBRID CROSSES AND MENDEL'S LAW OF INDEPENDENT ASSORTMENT

The experiments of Mendel that we have discussed so far have dealt with crosses between plants differing in only one character or alleles at one locus. The most logical extension of Mendel's experiments is to determine what happens when more than one pair of characters is simultaneously involved in the cross. He performed a number of dihybrid crosses in which the parents made different contributions for two traits. From these experiments he proposed his **second law, the law of independent assortment.** *This law states that the factors or genes for different traits assort independently of one another.* In other words, the genes on different chromosomes behave independently for the production of gametes.

Let us examine the results of Mendel's experiments in which he crossed plants that differed in two characters. An example is the mating of plants involving smooth or round (S)/wrinkled (s) and yellow (Y)/green (y) seed traits. As you already know, yellow is dominant to green and smooth is dominant to wrinkled. Mendel made crosses between true breeding smooth, yellow plants (SSYY) and wrinkled, green plants (ssyy). The results are shown in the Fig. 1.12.

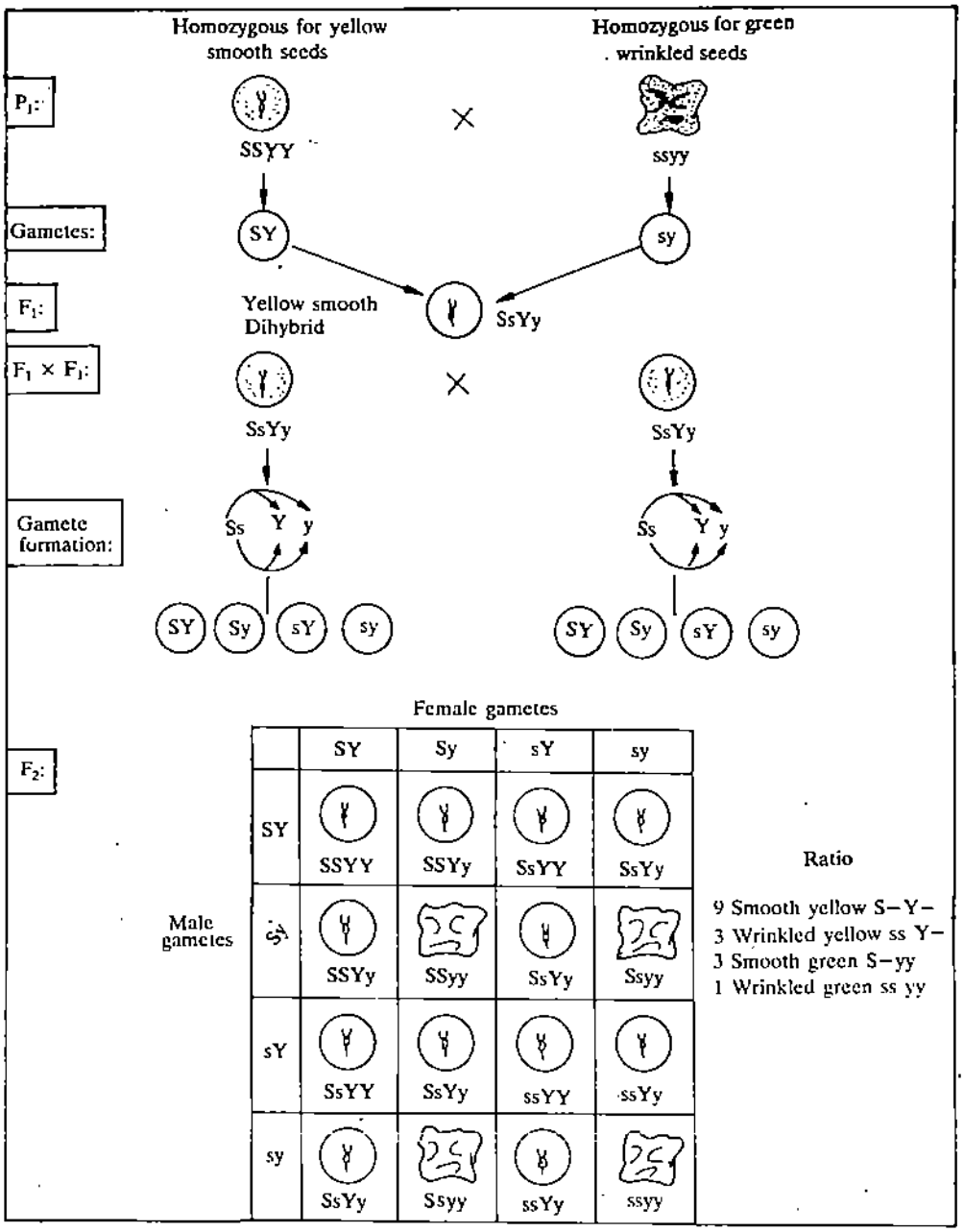


Fig. 1.12 : A dihybrid cross followed through the second generation. On inspecting the Punnett square one can find nine different genotypes in the F_2 generation. Because of dominance in one member of each pair of alleles, only four phenotypes are expressed.

All the F_1 seeds from this cross were smooth and yellow as predicted from the results of monohybrid crosses. The smooth yellow parents produce SY gametes which give rise to SsYy zygotes when fused with the sy gametes from the other parent. Because of the dominance of the smooth and yellow traits, all F_1 seeds are smooth and yellow. As in the monohybrid crosses, the F_1 were selfed to give rise to the F_2 generation.

Mendel thought of two possible outcomes. One, that the genes for the traits from the original (P_1) parent would be transmitted to the progeny together. If this was the case then a phenotypic ratio of 3:1 smooth-yellow: wrinkled-green would be obtained. The second possibility was that each of the traits would be inherited independently of the other.

Let us examine the Fig. 1.12 carefully. The dihybrid plants grown from the F_1 seeds produce four types of gametes: SY, Sy, sY and sy. Because of the independence of the two pairs of genes there is an equal frequency of each gametic type. When the F_1 plants are selfed, these four types of gametes fuse randomly in all possible combinations. All the possible 16 gametic fusions are represented on the Punnett square in the above figure. In the dihybrid cross, 9 different genotypes are obtained but because of dominance only 4 phenotypes are observed.

1 2	SSYY, SSYy,	2 SsYY 4 SsYy	=	9 smooth, yellow
1	SSyy,	2 Ssyy	=	3 smooth, green
1	ssYY,	2 ssYy	=	3 wrinkled, yellow
1	ssyy,		=	1 wrinkled, green

Mendel's actual observations are shown in the Table 1.3.

Table 1.3 : Results of Mendel's Dihybrid Cross

Generation	Seed Type	Number	F_2 Ratio
Parents	Round, green vs wrinkled, yellow		
F_1	Round, yellow		
F_2	Round, yellow	315	9.84
	Wrinkled, yellow	101	3.16
	Round, green	108	3.38
	Wrinkled, green	32	1.00
		556	

From this example, it is clear that if two pairs of characters are inherited independently then the selfing of F_1 individuals will yield a 9:3:3:1 ratio in the F_2 generation. This ratio arises as a result of independent assortment of two gene pairs into gametes and of the random fusion of these gametes. The 9:3:3:1 ratio is the product of two 3:1 ratios, that is, $(3:1)^2 = 9:3:3:1$. This prediction was met in all dihybrid crosses that Mendel performed. In every case the F_2 ratio was close to 9:3:3:1. Based on the results, Mendel concluded that the factors (genes) determining different traits were transmitted independently.

Let us now sum up as to what we have learnt in this section so far. *The principle of independent assortment states that genes on one chromosome segregate independently of those on another chromosome in the production of gametes.*

Branch Diagram of a Dihybrid Cross

As is already apparent, it is quite a tedious, time and space consuming job to construct a Punnett square of gametic combinations, then to count the numbers of each phenotypic class from all the genotypes produced. This exercise is not too difficult for dihybrid crosses, but for more than two gene pairs it becomes rather complex. So, it is better that we get into the habit of calculating the expected ratio of phenotypic classes by using a branch diagram.

We shall use the same example as the one used in Fig. 1.12 where two gene pairs assort independently into gametes. As you have seen earlier that on selfing F_1 of an Ss heterozygote gave rise to progeny of which three fourths were smooth and one fourth were wrinkled. Genotypically, the former class had atleast one dominant S allele, that is, its members were SS or Ss. A convenient way to signify this situation is to use a dash to indicate an allele that has no effect on the phenotype. *Thus, S— means that phenotypically the seeds are smooth and genotypically they are either SS or Ss.*

Now consider the F_2 produced from a selfing of Yy heterozygotes; again a 3:1 ratio is seen with three-fourth of the seeds being yellow and one-fourth of the seeds being green. The same holds good for the smooth vs. wrinkled seeds in the F_2 ratio. Let us now find out the expected proportion of F_2 seeds that are smooth and yellow. This ratio can be obtained by the product of the probability that an F_2 seed will be smooth and the probability that it will be yellow, or $\frac{3}{4} \times \frac{3}{4} = \frac{9}{16}$ (see Fig. 1.13, 1st ratio). Similarly, the expected proportion of the F_2 individuals that are wrinkled and yellow is $\frac{3}{4} \times \frac{1}{4} = \frac{3}{16}$ (see Fig. 1.13, 2nd ratio). Extending this calculation to all possible phenotypes, as shown in the Figure, we obtain the ratio 9:3:3:1.

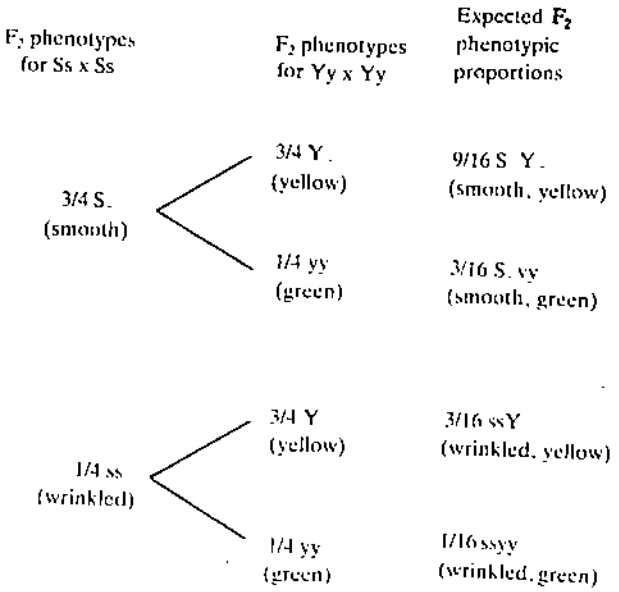


Fig. 1.13 : The example used here is the same as that of Fig. 1.12. The above figure shows the calculation of the F_2 phenotypic ratio by the branch diagram method.

If one wishes to know the genotype of the F_1's and the F_2's from a dihybrid cross, a test cross may be used. The example that we shall use to explain this point is already familiar to you that is of smooth/wrinkled, and yellow/green characters of pea seeds. The F_1 here is a double heterozygote (see Fig. 1.14) SsYy. As you know this F_1 produces four types of gametes in equal proportions and these are SY, Sy, sY and sy (see again Fig. 1.14). You may recall that in a test cross, we use a doubly homozygous recessive plant and in this case it is ssyy. A test cross like this one, shows a 1:1:1:1 ratio in the offspring of SsYy:Ssyy:ssYy:ssyy genotypes which means a 1:1:1:1 ratio of smooth-yellow:smooth-green:wrinkled-yellow:wrinkled-green phenotypes. This 1:1:1:1 phenotypic ratio is diagnostic of test crosses in which the "unknown" parent is a double heterozygote.

In the F_2 of a dihybrid cross there are nine different genotypic classes but only four phenotypic classes. The genotypes can be ascertained by test crossing. Use Table 1.4 for obtaining the genotypes. Substitute A, a, B and b in the table with the respective character involved, i.e., with S, s, Y or y. The test crosses in the table (from top to bottom) lists the 9 genotypic classes and the top column (from left to right) indicates the four phenotypic classes.

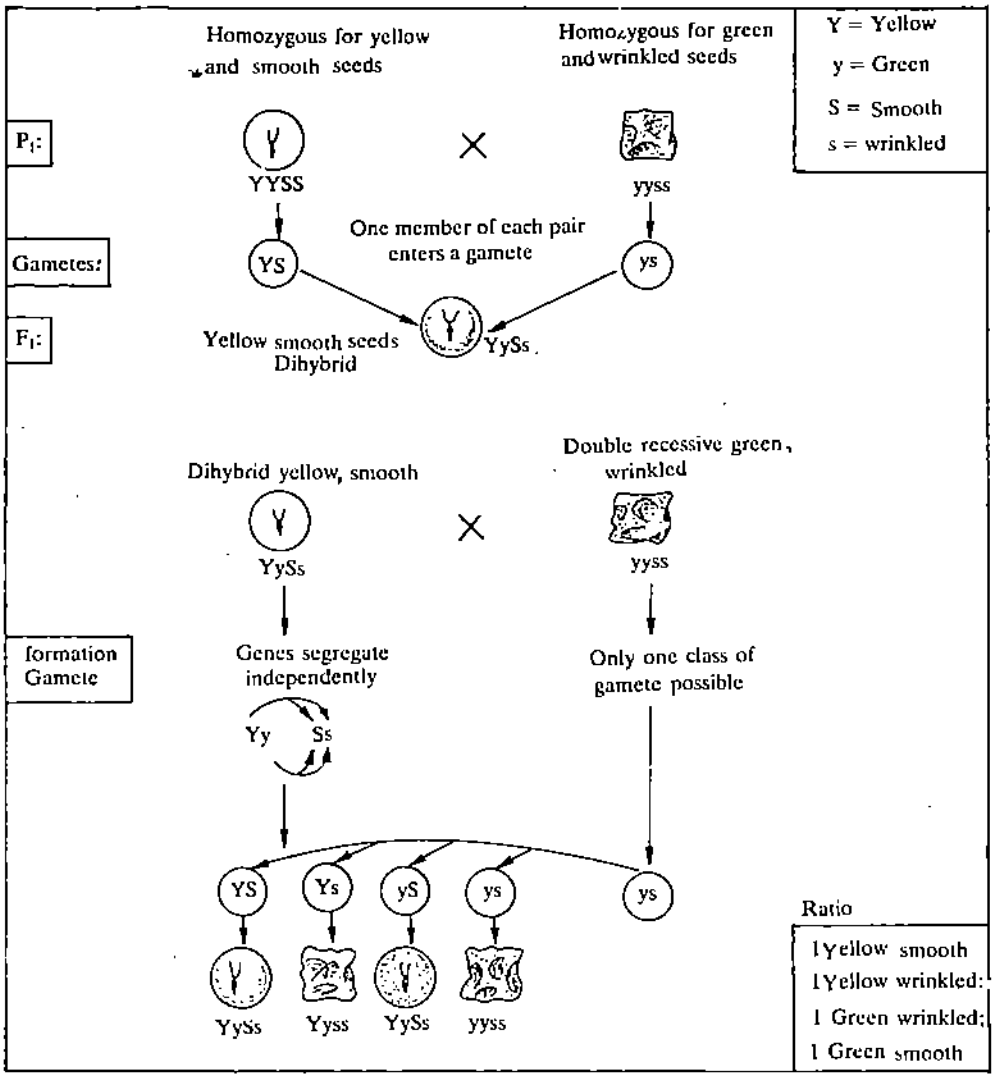


Fig. 1.14 : A dihybrid (YySs) being test-crossed by double recessive, green wrinkled parent (yyss).

Table 1.4 : Proportions of Phenotypic classes Expected From test crosses of strains with various genotypes for two gene pairs.

Test crosses	Proportion of Phenotypic Classes			
	A- B-	A- bb	aa B-	aa bb
AABB × aabb	1	0	0	0
AaBB × aabb	½	0	½	0
AABb × aabb	½	½	0	0
AaBb × aabb	¼	¼	¼	¼
AAbb × aabb	0	1	0	0
Aabb × aabb	0	½	0	½
aaBB × aabb	0	0	1	0
aaBb × aabb	0	0	½	½
aabb × aabb	0	0	0	1

From the above discussion it can be concluded that a test cross is a truly diagnostic approach to confirm the genotype.

SAQ 5

Given below is a Punnett square, go through it carefully and then answer the following questions.

		Female			
		AB	Ab	AB	Ab
Male	AB	AABB	AABb	AABB	AABb
	Ab	AABb	1	AABb	AAbb
	aB	AaBB	AaBb	2	AaBb
	ab	AaBb	Aabb	AaBb	aabb

Key : A and B are dominant
a and b are recessive

a) The genotype of the male parent is _____.
b) The genotype of the organism in box 1 is _____.
c) The organism in box 2 is _____ for trait 'A'.
d) The organism in box 2 exhibits _____ dominant genetic traits.

1.7 TRIHYBRID CROSSES

Mendel also applied his laws to the inheritance of three characters in other garden pea crosses. He found that both these laws hold good in this case also. Such crosses involving three characters are called trihybrid crosses. In such crosses too, the proportions of F_2 genotypes and phenotypes are predicted precisely with the same logic used before, with each character pair considered independently. You can try drawing a Punnett square to find out the genotypes as well as the phenotypic ratios, but here we give a branch diagram to show the derivation the relative frequencies of the phenotypic classes. In the example illustrated in the Fig. 1.15, the independently assorting character pairs in the cross are: smooth versus wrinkled seeds; yellow versus green seeds, and purple versus white flowers, there are 64 combinations of 8 maternal and 8 paternal gametes. Combination of these gametes give rise to 27 different genotypes and 8 different phenotypes.

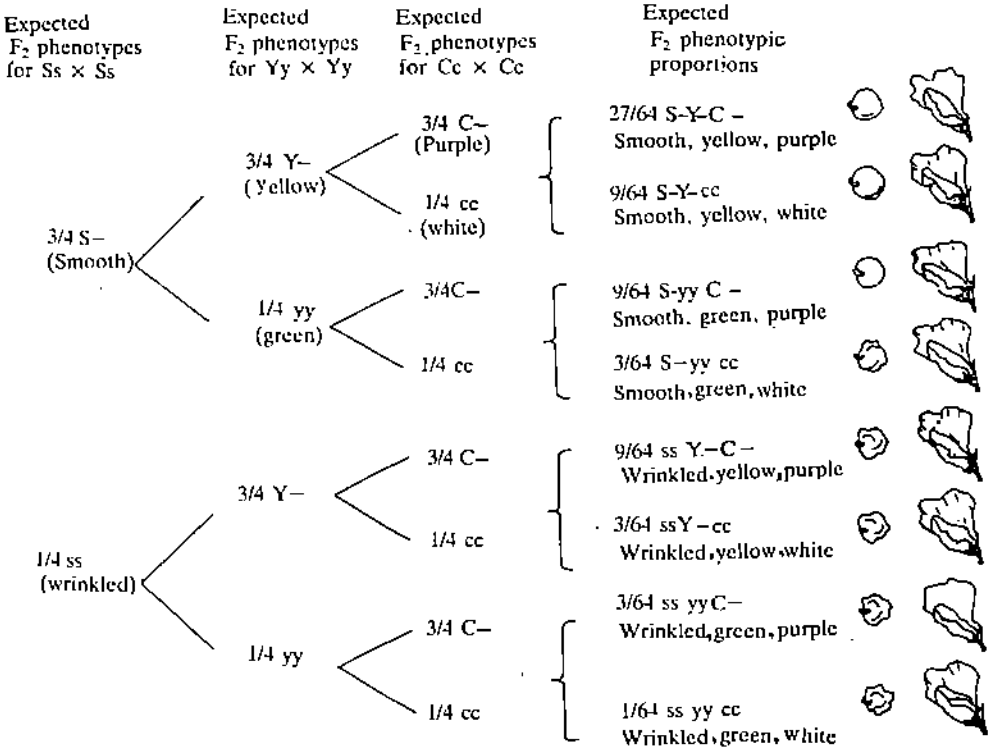


Fig. 1.15 : Branch diagram of a trihybrid cross. The ratio of the eight phenotypic classes above are calculated in the same manner as you have learnt in Figs. 1.10 and 1.13.

1.8 SOME GENERALISATIONS

We hope you have understood clearly, Mendel's laws and the methods used for finding out genotypes and phenotypes from different types of crosses. In each of the

previous examples, the F_1 is heterozygous for the gene(s) involved in the cross, and the F_2 is generated by selfing the F_1's wherever possible. You should remember that selfing is possible only in plants and in animal situations only cross fertilisation is possible.

In monohybrid crosses there were 2 phenotypic classes in the F_2, in dihybrid crosses there were 4, and in trihybrid crosses there were 8. The general rule is that there will be 2^n phenotypic classes in the F_2, where n is the number of independently assorting, heterozygous gene pairs (see Table 1.5).

Table 1.5 : Number of Phenotypic and Genotypic classes expected from self crosses of heterozygotes in which all genes show complete dominance.

Number of Segregating Genes	Number of Phenotypic Classes	Number of Genotypic Classes
1*	2	3
2	4	9
3	8	27
4	16	81
n	2^n	3^n

* For example, from Aa × Aa, two phenotypic classes are expected, with genotypic classes AA, Aa and aa.

This rule holds **only** *when a dominant recessive relation holds for each of the gene pairs.* Furthermore, there were 3 genotypic classes in the F_2 of monohybrid crosses, 9 in the dihybrid crosses, and 27 in trihybrid crosses. A simple rule is that the number of genotypic classes will be 3^n, where n is the number of heterozygous gene pairs.

The phenotypic rule (2^n) can also be used to predict the number of classes that will come from a multiple heterozygous F_1 used in a test cross. Here, the number of genotypes in the next generation will be the same as the number of phenotypes.

The examples given so far have dealt with situations in which each locus was represented by a maximum of two allelic variants and in most of these examples, one of the alleles has been dominant and the other recessive. It is true that a single diploid individual will have only two alleles for a particular locus and that a haploid gamete will only have a single allele for each locus. However, if we survey a population, we may find additional alleles for the locus in question. If three or more alleles for a given locus exist within the population, we can say that the locus has *multiple alleles.* A great many loci are known to have multiple alleles. Usually each identifiable allele can produce a distinct phenotype, and certain patterns of dominance and recessivity can be discerned when any of the two alleles are combined in various permutations and combinations. You would study the multiple alleles in detail in section 2.3 of Unit 2.

SAQ 6

a) How many different kinds of gametes can an organism of genotype AaBBCc produce?
   i) 3
   ii) 4
   iii) 9
   iv) 16

b) A trihybrid cross is made in yeast, both parents have a genotypes of AaBbCc. What proportion of the offspring will be of genotype aabbcc?
   i) 0
   ii) $\frac{1}{4}$
   iii) $\frac{1}{16}$
   iv) $\frac{1}{64}$

c) In peas, Tall (T) is dominant over dwarf (t), Yellow (Y) is dominant over green (y), and smooth (S) is dominant over wrinkled (s). What fraction of the offspring in the following cross would be homozygous recessive for all gene pairs in the cross : YyTtss × YyttSs?

1.9 SUMMARY

In this unit you have learnt that:

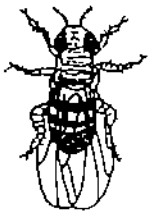
- Genetics is a branch of biology concerned with the structure, transmission and expression of hereditary information. More than a century ago, an Austrian Monk, Gregor Mendel worked on garden peas, with seven contrasting characters, and kept detailed data of his experiments. He analysed the data mathematically and laid foundation for the science of modern genetics.
- Genes are present on chromosomes. The term locus refers to the site a gene occupies in a chromosome.
- Different alternate forms of a particular gene are called alleles; they occupy corresponding loci on homologous chromosomes.
- An individual that carries two identical alleles for a given locus is said to be homozygous for that locus. If the two alleles are different then the individual is said to be heterozygous.
- According to Mendel's law of dominance, one allele (the dominant allele) masks the expression of the other allele (the recessive allele) in a heterozygous individual. For this reason an individuals's appearance (phenotype) may be different from the genetic constitution (genotype).
- According to Mendel's law of segregation, during meiosis, the alleles for each locus separate or segregate, from each other as the homologous chromosomes separate. When haploid gametes are formed, each will contain one and only one allele for each locus.
- According to Mendel's law of independent assortment, during meiosis, each pair of alleles will segregate independently of the other pairs. Alleles at different loci, therefore, assort randomly into the gametes.
- A cross between homozygous parents (P generation) that differ from each other with respect to their alleles at one locus is called a monohybrid cross; if they differ at two loci, it is a dihybrid cross. The first generation of offspring is heterozygous and is called the first filial or F_1, generation; the generation produced by a cross of two F_1 individuals is the second filial or F_2, generation. If a F_1 individual is crossed with a homozygous recessive individual the cross is known as a test cross.
- Mendel first demonstrated that inheritance was based on the segregation of alleles at a locus. Through self-fertilisation of F_1 garden pea plants, obtained from crosses between plants with different characteristics, he observed a 3:1 ratio of dominant to recessive phenotypes in the F_2 progeny. The law of segregation developed from these experiments states that heterozygous individuals produce equal proportions of gametes containing members of a pair of alleles. Mendel's second principle of independent assortment states that alleles at different loci assort independently of each other resulting in a 9:3:3:1 ratio in F_2 progeny from a dihybrid cross, and in a 27:9:9:9:3:3:3:1 ratio in a trihybrid cross.

1.10 TERMINAL QUESTIONS

1) Determine whether the following statements are True or False.
   a) Mendel crossed different types of pea plants to blend their characteristics.
   b) Homozygotes have identical alleles for a given trait on homologous chromosomes.
   c) An organism having the genotype Tt makes only one type of gametes.
   d) All organisms with identical phenotypes have identical genotypes.
   e) The phenotype determines whether an organism is heterozygous for that trait.
   f) Phenotype determine whether an organism is heterozygous or homozygous.
   g) Any organism displaying a recessive trait must be homozygous for that trait.
   h) A cross between a homozygous dominant tall pea plant and a heterozygous tall pea plant yields 100% tall pea plants.
   i) The inability for a person to roll his tongue represents a complete absence of the dominant gene (assuming tongue rolling is the dominant trait).
   j) A cross between a homozygous dominant male and a heterozygous dominant female would not have any homozygous recessive.

2) Two long-winged flies were mated. The offspring consisted of 77 flies with long wings and 24 flies with short wings (see the given figure). Is the short-winged condition-dominant or recessive? What are the genotypes of the parents?

Fruit flies



Long-winged (male)

Short-winged (female)

3) The ability to roll the tongue into almost circle is conferred by a dominant gene in humans, while its recessive allele fails to confer this ability. A man and a woman are heterozygous for tongue rolling, and have three sons. These three sons marry women who are not tongue rollers. Assuming that each of the three sons has a different genotype, show by diagram what proportion of their children might have the ability to roll their tongues.

4) Normal length of fur in rabbit is controlled by the dominant allele R, and a short type of fur called "rex" is determined by the recessive allele r. The dominant allele B is responsible for black fur colour; while the recessive allele b determine brown.
   a) Diagram a dihybrid cross between a homozygous rabbit with normal-length black fur and rex rabbit with brown fur. What are the phenotypic ratios resulting from this cross?
   b) What proportion of the normal, black rabbits in the F_2 generation of this cross can be expected to be homozygous for both pairs of genes?
   c) What would be the expected phenotypic and genotypic results of a back cross between a member of the F_1 generation and a fully recessive rex, brown parent?

5) Consider three gene pairs Aa, Bb and Cc, each of which affects a different character. (In each case the upper case letter signifies the dominant allele and the lower case letter the recessive allele). These three gene pairs assort independently of each other. Calculate the probability of obtaining.
   a) An AaBBCc zygote from a cross AaBbCc × AaBbCc;
   b) An AaBBcc zygote from a cross aaBBcc × AAbbCC;
   c) An ABC phenotype from a cross AaBbCC × AaBbcc;
   d) An abc phenotype from a cross AaBbCc × aaBbcc;

1.11 ANSWERS

Self-assessment Questions

1) a) heterozygote, genotype
   b) homozygous dominant, genotype
   c) heterozygote, genotype
   d) phenotype
   e) homozygous dominant, genotype
   f) heterozygote, genotype
   g) homozygous recessive, genotype

h) phenotype
i) homozygous recessive, genotype

2) a) iv
b) i
c) iii
d) ii

3) Pollen from dwarf plant dusted on the stigma of a tall plant.

4) a) i
b) i
c) ii
d) i
e) ii
f) iii

5) a) AaBb
b) AAbb
c) heterozygous
d) 2

6) a) ii,
b) iv

Hint: The frequency of getting aabbcc separately is 1/4. Therefore, the probability of the offspring having all three of these homozygous pairs is $1/4 \times 1/4 \times 1/4 = 1/64$.

c) This problem can be solved using the rule of probability for independent events. The probability of independent events occurring together is the product of their individual probabilities.
First determine the expected frequency of homozygous recessive for each gene pair.

	Y	y
Y	YY	Yy
y	Yy	yy

	T	t
	Tt	tt
t	Tt	tt

	S	s
	Ss	ss
s	Ss	ss

Then multiply the expected homozygous recessive frequencies together to find the fraction of offspring with all the homozygous recessive gene pairs.

$1/4yy \times 1/2tt \times 1/2ss = 1/16$ yyttss

Terminal Questions

1) a) False
b) True
c) False
d) False
e) False
f) False
g) True
h) True
i) True
j) True

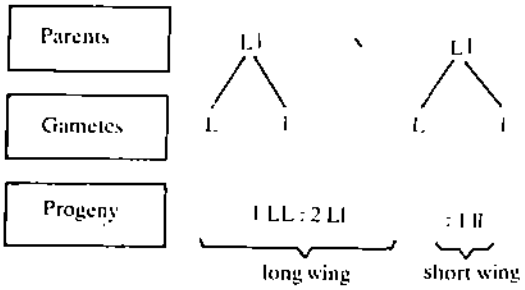
2) When we are not told which of the characteristics is dominant and which is recessive, we can deduce it from the ratio of phenotypes in the progeny. We know that 77 flies have long wings and 24 have short wings. This gives us an approximate ratio of 3 long-winged flies to every 1 short-winged fly.

$77/24 = 3/1$

As you know, the three-to-one ratio signifies that dominant and recessive characteristics are most likely involved. Since there are 3 long-winged ones, it suggests that short-wingedness is recessive.

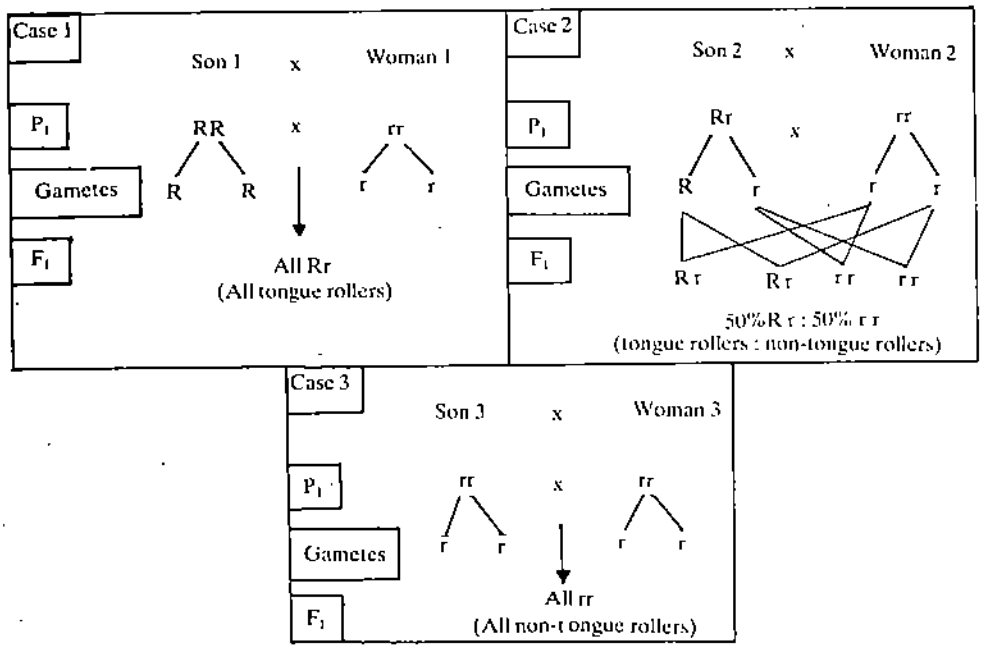
The presence of short-winged flies (homozygous recessive) in the progeny suggests that both parents carry the recessive gene and are thus heterozygotes.

Let L be the gene for long wings in flies and l be the gene for short wings. In the cross between two long-winged heterozygous parents:



Let us represent the dominant allele for the trait of tongue-rolling by R, and the recessive allele by r. Only three possible genotypes could result from the cross of the heterozygous parents. These alleles are RR, Rr and rr. Let one of the sons be RR, therefore, a tongue roller; the second son Rr and also a tongue roller; and the third son rr (homozygous recessive) and thus unable to roll his tongue. All three sons have married women who are not tongue rollers and who are, therefore, rr.

Let us now determine the proportion of the offspring who will be able to roll their tongues.



4) a) First draw the Punnett square of this cross. The genotype of the homozygous black rabbit with normal fur is BBRR and the genotype of the homozygous brown rex rabbit is bbrr. The result of the cross is shown in the figure given below.

Gametes	BR	Br	bR	br
BR	BBRR	BBRr	BbRR	BbRr
Br	BBRr	BBrr	BbRr	Bbrr
bR	BbRR	BbRr	bbRR	bbRr
br	BbRr	Bbrr	bbRr	bbrr

Summary of F_2 phenotypes:
9 Black, long : 3 Black, rex : 3 Brown, long : 1 Brown, rex. This is an example of a typical dihybrid cross. The ratio 9:3:3:1 is expected from all dihybrid crosses involving dominance.

b) By scanning the above diagram, one can count 9 normal-furred black F_2 rabbits (either BB or Bb or fur colour and either RR or Rr for fur length, since normal length of fur and black colour are dominant traits); of these only one is homozygous for both traits – BBRR
Total number of normal black F_2 rabbits expected = 9
Number of those expected to be homozygous for both traits = 1
Ratio = 1:9

The proportion of normal, black rabbits that can be expected to be double homozygous is thus 1/9.

c) Genotype of F_1 member – BbRr
Genotype of brown, rex parent – bbrr
The genotype and phenotypic results of this back cross are:

BbRr × bbrr
↓

Gametes	BR	Br	bR	br
br	BbRr	Bbrr	bbRr	bbrr

Summary of back cross results : 1 Black, long : 1 black, rex : 1 black, long : 1 brown, rex

5) First reduce the question into simple parts so that basic Mendelian principles can be applied. The key is that the genes assort independently so we must multiply the probability of the individual occurrence to obtain the final answer.

a) Consider the Aa gene pair. The cross is Aa × Aa, so the probability of the zygote being Aa is 2/4 since the expected distribution of genotypes is 1 AA : 2 Aa : 1 aa, as we have discussed. The probability of BB from Bb × Bb is 1/4, and that of Cc from Cc × Cc is 2/4, following the same sort of logic. The probability of an AaBBCc zygote is 1/2 × 1/4 × 1/2 = 1/16.

b) Similar logic is needed here, although we must be sure of the genotypes of the parental types since they differ from one gene to another. For the Aa pair the probability of getting Aa from AA × aa has to be 1. Next, the probability of getting Bb from BB × bb is 0; so on these grounds alone we cannot get the zygote, asked for from the given cross.

c) Since the question asks for the probability of getting a particular phenotype, we must think of dominance. Again we break up the question and consider each character pair in turn. The probability of an A phenotype from Aa × Aa is 3/4 from basic Mendelian Principles. Similarly, the probability of a B phenotype from Bb × Bb is 3/4. Lastly, the probability of a C phenotype from CC × cc is 1. Overall the probability of an ABC phenotype is 3/4 × 3/4 × 1 = 9/16.

d) An abc phenotype from AaBbCc × aaBbcc is 1/2 × 1/4 × 1/2 = 1/16.

UNIT 2 EXTENSIONS AND MODIFICATIONS OF MENDELIAN GENETIC ANALYSIS

Structure

2.1 INTRODUCTION

It was earlier thought that Mendel's law hold good for all situations and they form the basis for predicting the outcome of crosses in which segregation and independent assortment occurs. The research work conducted after the discovery of Mendel's work has confirmed this. And also, with the further advancement of knowledge it has come to light that the real world of genes and chromosomes is more complex than what was thought earlier. There are several exceptions and extensions to Mendel's laws though *all these do not invalidate Mendel's findings.* Rather, they show that there are situations other than the ones that can be explained by segregation and independent assortment of gene pairs. These situations must be accommodated into the fabric of genetic analysis, and we must adopt an expanded view of the Mendelian or Transmission Genetics. Therefore, some such extensions, e.g., multiple alleles, sex linkage are taken up in this unit. Also included are the other extensions such as the cases in which one allele may not be completely dominant to another allele at a gene locus; instances in which the products of different genes interact to produce modified mendelian ratios, lethal genes, pleiotropy, penetrance, expressivity, and the effects of environment on gene expression. We think, that a study of this unit would provide you greater insight into the elements of genetics and the mechanics of inheritance.

Objectives

After studying this unit you would be able to:

- distinguish between complete dominance, incomplete dominance, and codominance (Section 2.2);
- explain multiple alleles, and the mechanism of their inheritance (Section 2.3);
- recognise and explain the modified Mendelian ratios (Section 2.4);
- infer from the progeny ratios the operational lethal genes, and describe the inheritance pattern in parents carrying lethal genes (Section 2.5);

- explain the term pleiotropy and give examples of the effects caused by pleiotropic genes (Section 2.6);
- define what sex-linked characters are, and describe the mode of inheritance of sex-linked genes (Section 2.7);
- describe with the help of examples the degrees of gene expression — penetrance and expressivity, and differentiate between the two (Section 2.8);
- explain how environment affects the expression of genes in organisms (Section 2.9).

2.2 DOMINANCE

All the seven traits that Mendel selected had one allele dominant over the other, and when he examined the F_2 progeny there was a 3:1 dominant to recessive ratio of the phenotypes. In fact, Mendel suggested that dominance was a general characteristic of traits, similar to the principle of segregation and independent assortment. However, shortly after 1900, three different patterns of dominance were documented, and these are discussed below.

2.2.1 Complete Dominance

So far, we have only considered the phenomena of complete dominance. All the examples you have studied in Unit 1 have two alternative states : one, the phenotype produced by the dominant allele from its homozygous or heterozygous state (*AA* or *Aa*); and second, the phenotype resulting from the recessive allele in its homozygous state (*aa*). You may recall that in complete dominance, the heterozygotes (*Aa*) cannot be distinguished from the homozygotes (*AA*), that is, *AA* = *Aa* in phenotype. You may also recall that the completely recessive allele is phenotypically expressed only when it is homozygous (*aa*). **Complete dominance** and **complete recessiveness** are the two extremes of a range of dominance relationships. Many alleles, however, do not exhibit this dominance relationship.

2.2.2 Incomplete Dominance

In many cases one allele is not completely dominant to another allele, and this phenomenon is called **incomplete** or **partial dominance.** It was discovered by Carl Correns, one of the discoverers of Mendel's principles. He crossed red-flowered four-O'clock plants with white ones (see Fig. 2.1), the F_1 plants did not resemble either parent but were an intermediate pink. When the F_1 plants were self-fertilised, the phenotypic ratio of F_2 individuals was 1 red : 2 pink : 1 white or ¼, ½ and ¼ respectively. These experimental results indicate incomplete dominance. *In this all genotypes have different phenotypes, with the heterozygote being intermediate between the two homozygotes.* You might have noticed that the notations used in Fig. 2.1 are slightly different from those we used in the Unit 1. Here, R, stands for the trait flower colour, and since there is lack of complete dominance, we have used the subscripts 1 and 2 denote the different flower colours. That is, R_1 and R_2 denote red and white alleles respectively.

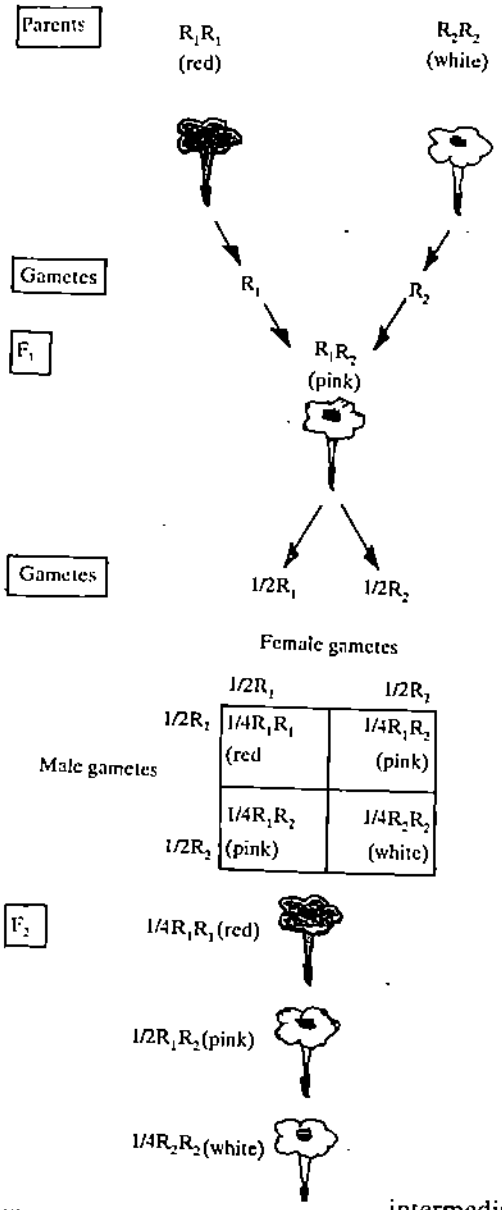


Fig. 2.1 : A cross between red and white four-O'clock flowers. The F_2 progeny shows incomplete dominance.

2.2.3 Codominance

When both alleles of a pair are **fully expressed** in a heterozygote, they are called **codominant alleles** and the phenomenon is known as **codominance**. Such alleles exhibit a unique pattern of expression with heterozygotes being phenotypically distinct from both the homozygotes and expressing both the alleles equally. In general, the two products are the same with respect to function but are different in exact amino acid sequence. In humans, the MN blood group antigens are a good example. The allele L^M for M-type blood is codominant with its allele L^N for N-type blood. The heterozygote L^ML^N expresses the characteristics of both M and N antigens (MN-type blood), that is, L^ML^N heterozygotes possess red blood cells that carry both

M and N antigens on their surfaces. Since the two alleles encode non-identical protein products, a mating between a homozygous M-type person (L^ML^M) and a homozygous N-type person (L^NL^N) would result in all heterozygous (L^ML^N) offspring. Mating between heterozygotes ($L^ML^N \times L^ML^N$) would result in a ratio of 1 M type (L^ML^M) : 2 MN-type (L^ML^N) : 1 N type (L^NL^N).

The phenotypic ratio of 1:2:1 has thus replaced the 3:1 ratio, because the alleles are both expressed in the heterozygote, that is, the alleles are codominant.

2.3 MULTIPLE ALLELES

So far you have learnt about just two alleles of a gene – for example, *A* versus *a*, *W* versus *w* or *G* versus *g*. Although each individual organism can contain no more than two alleles of any gene, some genes, however, have more than two alleles that are present among a group of organisms. For example, the human blood groups designated A, B, O or AB are determined by three types of alleles denoted by I^A, I^B and I^o. The blood group of any particular person is determined by the particular pair of alleles present in his or her genotype. In such a situation when more than two different forms of a gene exist in a species, they are referred to as multiple alleles.

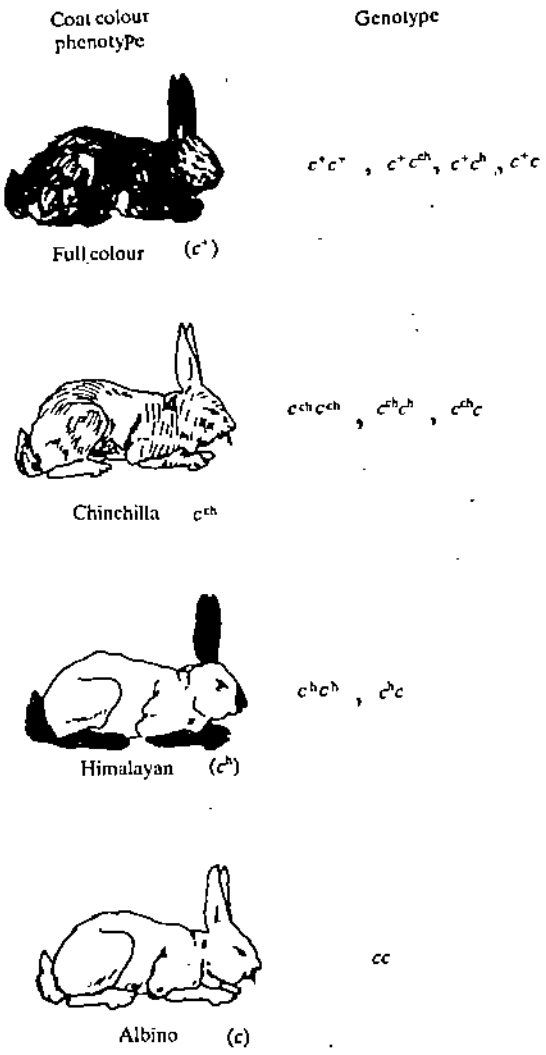


Fig. 2.2 : Multiple alleles. Coat colour in rabbit.

A classic example of multiple alleles involves coat colour in rabbits. Four alleles of the rabbit coat colour *(c)* gene have been seen (see Fig. 2.2) : c^+ – "wild type" or fully coloured coat; c^h – "himalayan", body white with tips of ears, nose, tail and legs black; c^{ch} – "chinchilla" entire body has light grey colour, and c – "albino". These show a gradation in dominance of $c^+ > c^{ch} > c^h > c$. That is, c^+ is dominant to each of the three remaining alleles, while c^{ch} is dominant to the remaining two in the series,

and so on. *Note that the members of the multiple allelic series are conventionally represented by the same letter or symbol, with appropriate superscripts to represent the different alleles.*

The above series show a gradation in dominance. In other series of multiple alleles, some may be codominant and others may be incompletely dominant so that the heterozygotes have a phenotype intermediate between those of their parents.

2.3.1 *ABO* Blood type Alleles in Humans

You have seen above that in humans, there are basically four blood groups that are governed by three alleles namely I^A, I^B and I^O. These alleles are of particular importance, as certain blood groups are incompatible. We would not be going into the how and why of the incompatibility of blood groups for the time being. But we shall examine, as to what are the phenotypes and genotypes of offspring obtained from the possible crosses between the males and females of all the three alleles. These are shown in Fig. 2.3, also see Box 2.1.

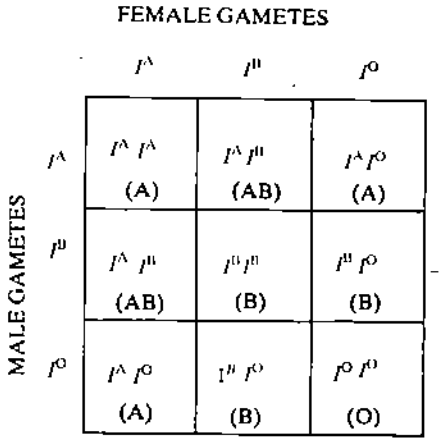


Fig. 2.3 : The six genotypes produced by the three alleles in the ABO blood group system. The corresponding phenotypes (given in parenthesis) are grouped into four phenotypes : A, B, AB and O.

From Fig. 2.3 you would find that in ABO blood system, *I^A and I^B are codominant* and *I^O is recessive* to both. I^AI^B heterozygotes have both A and B antigens on their red blood cells. I^OI^O homozygotes have no AB antigens on their red blood cells; I^AI^O and I^BI^O heterozygotes have A and B antigens respectively on their blood cells.

Box 2.1

The number of genotypes possible in a diploid organism with "n different alleles is found by adding 1 + 2 + 3 + ... + n, when n indicate all the integers between 3 ... and n, or by using the formula $\frac{n(n+1)}{2}$ Both of these expressions give the number of different ways objects (or alleles) can be combined two at a time. For the ABO blood system, with three alleles, the number of genotypes is either $1 + 2 + 3 = 6$ or $\frac{3(3+1)}{2} = \frac{(3)(4)}{2} = 6$. The same result is shown in Fig. 2.3. Supposing there were 4 alleles, there would be $\frac{(4)(4+1)}{2} = 10$ possible genotypes. Refer to the example of coat colour in rabbits where there are four alleles involved. Amongst these genotypes, the number of homozygotes is equal to the number of alleles, n, and the number of heterozygotes is equal to $\frac{(n)(n-1)}{2}$. For example, the six genotypes resulting for three alleles are composed of three homozygotes and $\frac{3(3-1)}{2} = \frac{(3)(2)}{2} = 3$ heterozygotes.

To summarise that has been said above, the numbers of homozygotes, heterozygotes and genotypes for the given multiple alleles can be calculated using the folloiwng table.

Alleles	Homozygotes	Heterozygotes	Genotypes
1	1	0	1
2	2	1	3
3	3	3	6
4	4	6	10
"	"	"	"
"	"	"	"
"	"	"	"
"	"	"	"
"	"	"	"
"	"	"	"
"	"	"	"
n	n	$\frac{n(n-1)}{2}$	$\frac{n(n+1)}{2}$

SAQ 1

Now that you have studied the Fig. 2.3 carefully. Can you write the genotype(s) for each blood group given below?

Blood Group	Genotypes
A	
B	
AB	
O	

You have just studied the genotypes produced by the three alleles in the blood group system. We shall now examine as to what makes these blood types different from one another. Differences among ABO blood group types occur because of the presence of different antigens and antibodies. *Antigens of the ABO blood system are molecules composed of a protein-sugar combination that occur on the surface of red blood cells. And antibodies as you know, are substances that are produced by the immune system in response to specific external agents.* Humans and other mammals produce antibodies and circulate them in the blood stream as a defence mechanism against foreign substances.

No antibodies are synthesised in normal individuals that react with antigens present on the individual's own cells. However, when type A blood and type B blood are mixed, the anti-A antibodies in the type B blood serum react with the antigens on the type A blood cells, and vice versa which produce **agglutination** or clumping of cells (see Fig. 2.4). Cross-matching blood types to determine compatibility is thus essential

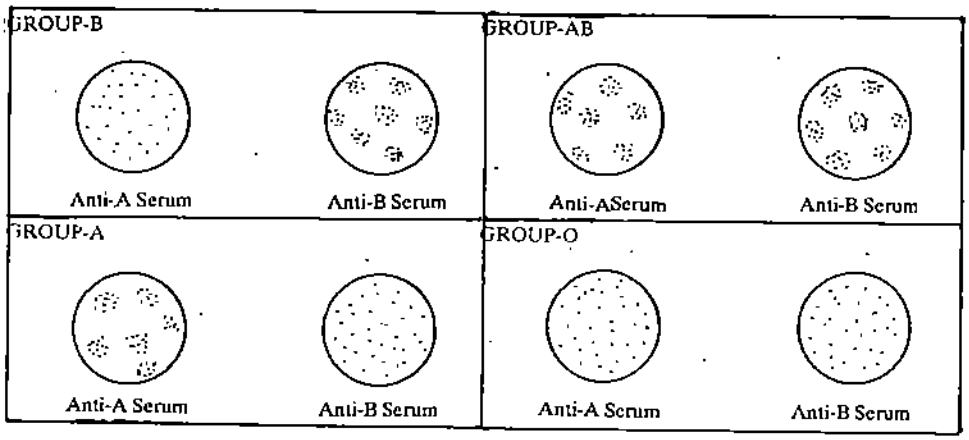


Fig. 2.4 : Agglutination reactions between the blood types in humans. Red blood cells of the type indicated at the top left corner of each box are mixed with serum of the type indicated below each reaction mixture (in circle). A clumped pattern of cells indicates that agglutination has occurred and the reaction is incompatible whereas the uniform distribution of cells indicates a compatible reaction.

in blood transfusions. In this process, blood donors and recipients are tested for the presence of antigens and antibodies that are compatible.

The conclusions that have emerged from the above discussions have been presented in Fig. 2.5.

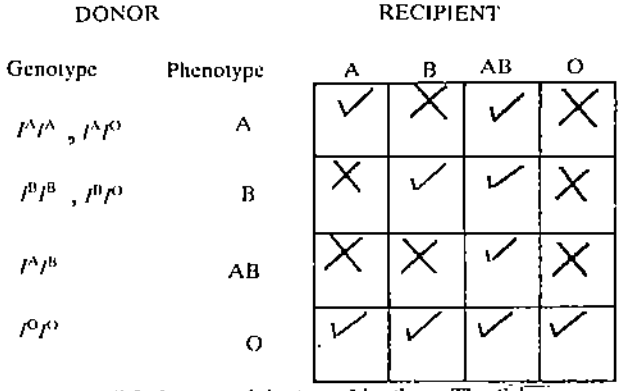


Fig. 2.5 : ABO donor recipient combinations. The tick mark (√) indicates compatibility in blood transfusion, and the cross (x) indicates incompatibility.

By now it should be clear to you why the type O individuals are **universal donors,** and type AB individuals **universal recipients.** Moreover you should be able to explain why type O recipients when given any blood other than O will suffer dire consequences, and type AB individuals can donate blood only to type AB individuals. You would learn more about the genetics of blood in Block 4, Unit 19 of this course.

2.3.2 *Rh* Factor Alleles in Humans

One major blood group that has been well publicised is the *Rh*. Its discovery goes back to 1940, when Landsteiner and Wiener injected red blood cells from Rhesus monkey (*Macaca rhesus*) into a rabbit, which in turn responded by producing antibodies against them. The antibodies thus produced were able to agglutinate the monkey's blood when mixed with it. When the rabbit serum, containing the antibodies against the monkey's blood was mixed with blood from humans there was clumping in 85% of the subjects tested. To check the specificity of this reaction, Landsteiner and Wiener had already removed from the serum all other antibodies known at that time that could possibly react with human antigens. Inspite of all this, clumping still took place. This meant that the rabbit serum contained an antibody against an antigen present in both human and monkey blood. Individuals whose blood cells react with the *Rh* antibody are termed *Rh*-positive, and those whose blood cells do not react are *Rh*-negative. The symbol *'Rh'* came from the first two letters of the species name of the monkey. A test for *Rh* incompatibility is accomplished by placing a drop of blood from the test organism on the slide and introducing anti-*Rh* serum. Agglutination of red blood cells indicates incompatibility, whereas an even distribution of red blood cells indicates no reaction.

The original antigen, now symbolised Rh^O is highly antigenic to humans. Thus, cross-matching of donor and recipient blood for the *Rh* factor, as well as of the *ABO* type is now used to avoid incompatibility following blood transfusions.

Rh-incompatibility also has a role to play in married life. In marriages between *Rh* positive men and *Rh*-negative women an interesting situation arises if the foetus is *Rh*-positive. It may be kept in mind that the allele for *Rh*-positive condition is dominant over that for *Rh*-negative condition. Due to the damage of small blood vessels blood is frequently exchange between the mother and the foetus during childbirth. Thus, *Rh*-negative mothers are often immunised by blood from *Rh*-positive foetuses. Usually no ill effects are associated with exposure of the mother to the *Rh*-positive antigen during the first childbirth unless the mother has been previously exposed to *Rh* antigen by transfusion. Subsequent *Rh*-positive children carried by the same mother, however, may be exposed to a higher titer of antibodies accumulated by the mother against the *Rh* antigen. Such children may develop symptoms of *haemolytic jaundice* and *anaemia*, a condition referred to as *erythroblastosis fetalis*. The symptoms may be mild, but in severe cases it may lead to the death of the foetus or newborn infant if appropriate steps are not taken by the physician.

Passive immunity for *Rh*-hemolytic anaemia can now be achieved by use of an *incomplete antibody* against the Rh^O antigen. This antibody does not agglutinate *Rh*-positive red blood corpuscles. Instead, these antibodies attach to antigen receptors on red cell surfaces and coat the cells. The coating of any *Rh* antigens from the *Rh*-positive foetus with incomplete antibodies inhibits the mother's capacity to form *Rh* antibodies. These incomplete antibodies may be injected into an *Rh*-negative mother immediately after she has given birth to an *Rh*-positive child. The injection of incomplete antibodies thus prevents the foetal antigens from eliciting the normal immune response of the mother. Injected antibodies dissipate within a few months and present no danger to the mother or foetus.

What is the genetic control of this *Rh* system? This must be the question arising in your mind. There are a few explanations for this, but the validity of these are yet to be established. Molecular studies, have revealed that multiple alleles exist either at a locus or loci that encode for the *Rh* antigen(s) and a single gene or several genes may be involved.

2.3.3 Incompatibility Alleles in Plants

As in humans and animals, the multiple alleles are known to play important role in plants too. It has been known for a long time that some plants just will not self-pollinate. A single plant may produce both male and female gametes, but both will not fertilise. The same plant, however, will cross with certain other plants. This phenomenon is called **self-incompatibility.** Of course, the pea plants used by Mendel were not self-incompatible, for he was able to self his plants with ease. What governs the compatibility in plants? This is the aspect that we shall discuss now. It is well established that incompatibility has genetic basis. And there are several different genetic systems acting in different self-incompatible species. These systems form very good examples of multiple allelism. One of the most common systems is found in many dicot and monocot plants, including sweet cherries, tobacco, petunias, and evening primroses. In each of these species, one gene determines compatibility/incompatibility relations, with many different allelic forms of the gene possible in different plants of any one species. In Fig. 2.6, three situations are

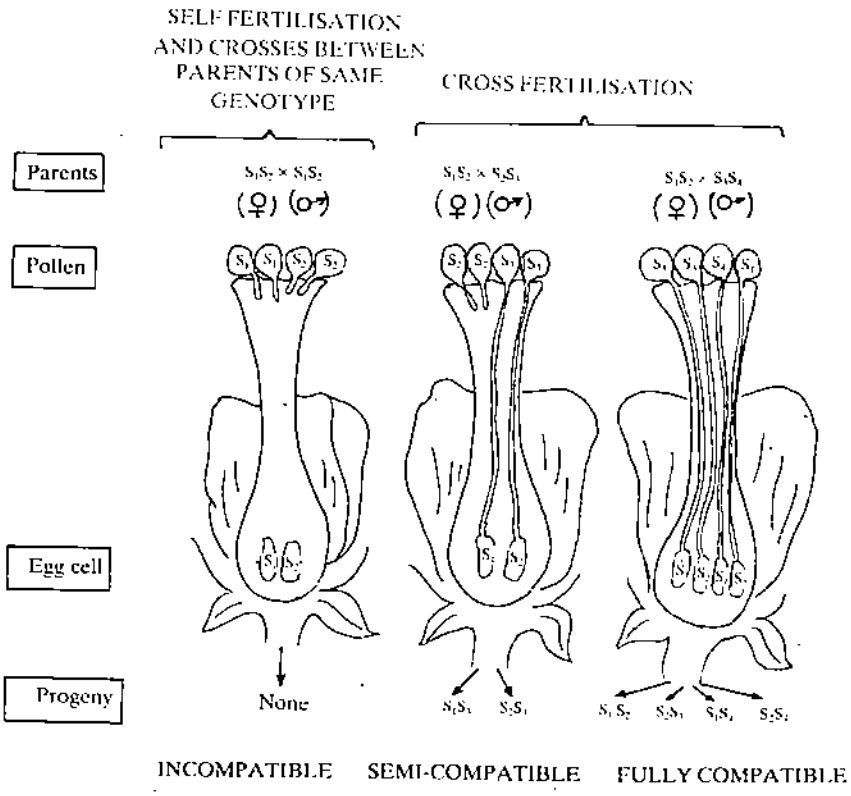


Fig. 2.6 : The role of multiple alleles in controlling incompatibility in plants. A pollen tube will not grow if the S allele that it contains is present in the female parent. In this diagram, for convenience only four alleles are shown, but as you know some systems have larger numbers of alleles. Such systems act to promote exchange of genes between plants by making selfing impossible and crosses between plants related by common ancestory unlikely.

depicted : first, fully incompatible reaction; second, semi-compatible reaction; and third, fully compatible reaction.

If a pollen grain bears an *S* allele that is also present in the maternal parent, then it will not grow; however, if that allele is not in the maternal tissue, then the pollen grain produces a pollen tube containing the male nucleus, and this tube effects fertilisation. The number of *S* allele in a series in some species may be very large—over 50 in the evening primrose and clover.

SAQ 2

Fill in the blanks with appropriate word(s).

a) Genes that exist in more than one form are called _____

b) A person with the blood type AB could not possibly be the parent of an offspring with blood group _____.

c) When red and white four-O'clocks are crossed, the colour of the offspring is pink. This is due to _____.

d) In all the Mendel's experiments, the two alternative alleles for a trait were _____

e) How many possible phenotypes are there for the ABO blood groups? _____

2.4 GENE INTERACTIONS AND MODIFIED MENDELIAN RATIOS

No gene acts by itself in determining the phenotype of an individual; the phenotype is the result of integrated pattern of molecular reactions that are under the direct control of genes. The examples that we have discussed so far and would subsequently take up in the following units, all have discrete biochemical basis. In a number of cases, the complex interactions between genes can be detected by genetic analysis. We shall discuss some such examples in this section.

SAQ 3

Recall Mendel's law of independent assortment that you studied in the previous unit. Let's assume that there are two independently assorting genes, each with two alleles *A* and *a*, and *B* and *b*. What type of genotype and phenotype will the progeny of a cross between doubly heterozygous individuals (i.e. *AaBb*) have and in what proportion?

.....

.....

.....

.....

If the phenotypes determined by the two allelic pairs are distinct – for example smooth versus wrinkled peas, tall versus short plants, then we get the familiar dihybrid ratio of 9:3:3:1. *Any alteration in this standard 9:3:3:1 ratio indicates that the phenotype is the product of the interaction of two or more genes.* These altered genotypic ratios are referred to as **modified mendelian ratios.**

For the 9:3:3:1 phenotypic ratio the genotype for these phenotypes can be represented in a short form as A–B–, A–bb aaB–, aabb respectively. The dash (–) indicates that the phenotype is the same whether the allele is dominant or recessive, e.g., *A*– means *AA* or *Aa*. This system, however, cannot be used when incomplete dominance or codominance is involved.

Let us now examine the interactions that result in modified mendelian ratios.

2.4.1 Gene Interactions that Produce New Phenotypes

All the previous examples of dihybrid crosses that we have discussed, were taken from Mendel's experiments. In these, he studied *two genes that affected different traits.* For example, the allelic pair for smooth/wrinkled peas had no effect on the allelic pair of yellow and green colour. But after the discovery of Mendel's laws, many scientists through their experimentation found that there are traits which are affected by several genes. It means that if **two allelic pairs affect the same characteristic**, there is a possibility of gene product interaction. A variety of phenotypic ratios can result depending on the type and the extent of interaction between the products of the nonallelic genes. We shall now discuss some examples to elaborate this point.

Comb Shape in Chickens : The comb shape in chickens is a classical example of such gene interaction. As a result of interaction between two allelic pairs, new comb shape phenotypes are produced. Figure 2.7 shows the four comb phenotypes that

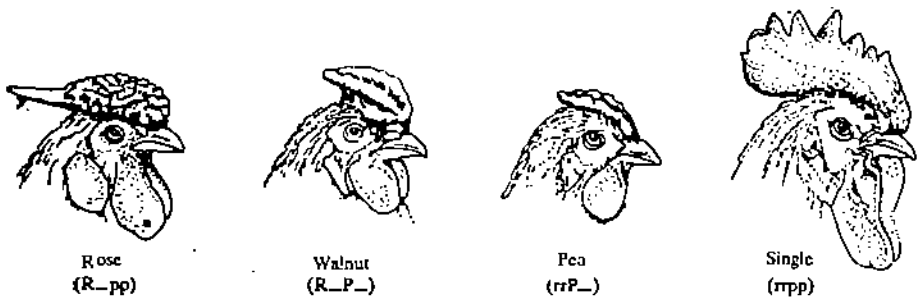


Fig. 2.7 : The genotypes and phenotypes of four comb shapes in chickens.

result from interaction of the alleles at two gene loci. Each of these types can be bred true. When we make crosses between the different comb-shaped birds, very interesting results are obtained. Carefully understand the three crosses a, b, and c in Figure 2.8. The 9:3:3:1 ratio in cross c matches the ratio of F_2 progeny resulting from a cross of parents differing in two genes. In cross c, this doubly dominant class is walnut, while the doubly recessive class is single.

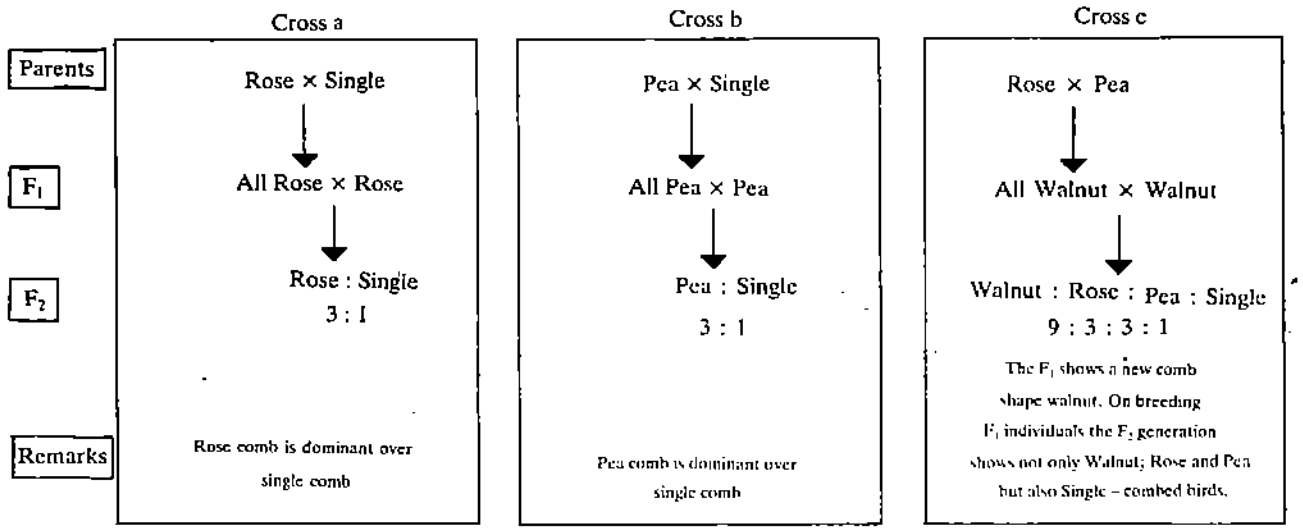


Fig. 2.8 : Crosses between true breeding : a) Rose- and single-; b) Pea- and Single-; and c) Rose- and pea-combed birds.

The overall explanation of the result (also see Fig. 2.9) is as follows: The *walnut comb depends* on the presence of *two dominant alleles R and P,* on two independently assorting gene loci. In the presence of atleast *one R allele* and with the *homozygosity* for the *recessive p allele* a *rose comb* results. Birds with atleast *one P allele* and *homozygosity* for the *recessive r allele* have a *pea comb.* The *doubly homozygous rrpp* birds have a single comb.

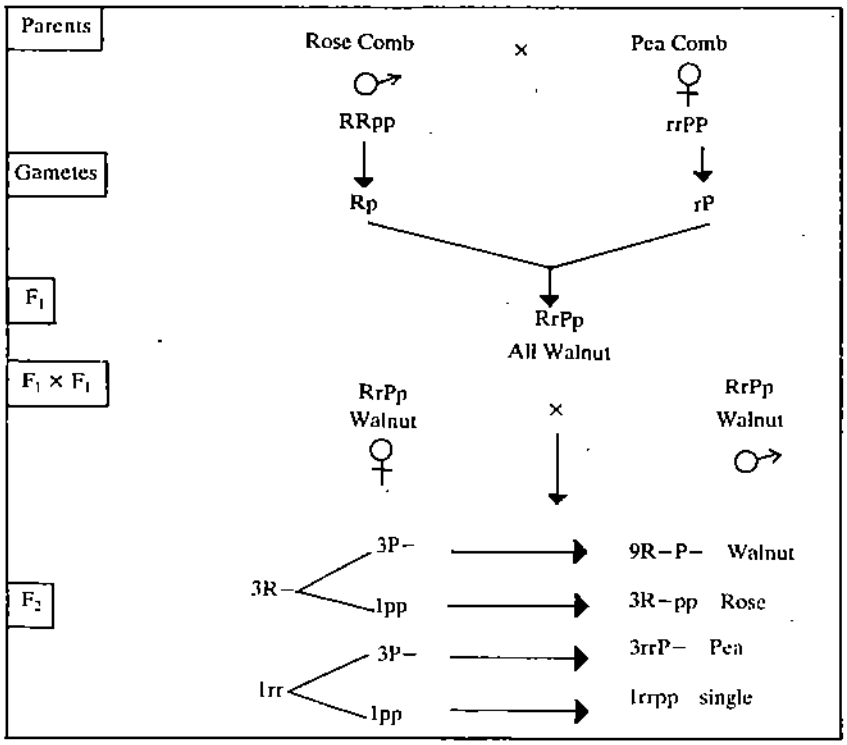


Fig. 2.9 : Results of crosses showing the interaction of genes for comb shape in fowl.

The biochemical basis for the four comb types is not clearly known. It is believed that the single-comb phenotype results from the activities of a number of genes other than the *R* and *P* genes. Thus *rrpp* birds do not produce any functional gene product that influences the comb phenotype beyond the basic — single comb. The dominant *R* allele might produce a gene product that interacts with the products of these genes, controlling the single-comb phenotype so that a rose-shaped comb results. Similarly, the dominant *P* allele might produce a gene product that interacts with the products of the single-comb genes to produce a pea-shaped comb. When the products of both the *R* and *P* alleles are present, they interact so that the comb develops differently resulting in the walnut comb.

Duplicate Interaction : Fruit shape in Summer Squash (9:6:1 ratio) : In summer squash, there are two varieties: spherical fruit and long-fruit varieties (see Fig. 2.10).

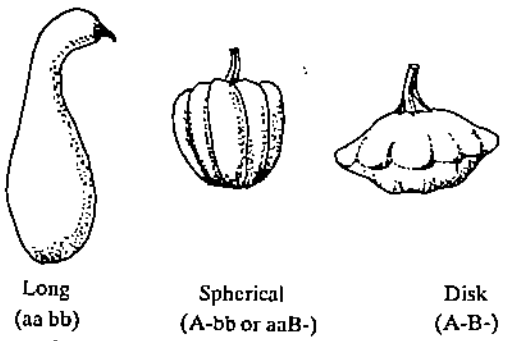


Fig. 2.10 : Three fruit shapes of summer squash.

The long-fruit varieties are always true-breeding. However, sometimes crosses between different true-breeding spherical varieties result in disk-shaped fruits (Fig. 2.10) in the F_1 generation. These F_1 individuals when crossed give rise in F_2 generation approximately 9/16 disk-shaped, 6/16 spherical and 1/16 long fruits. How does this happen? The explanation is as follows: the dominant allele alone (*A—bb* or *aaB—*) produce spherical fruit, while the two non-allele dominant alleles (*A—B—*) interact together to produce a new phenotype, that is, disk-shaped fruit. And the doubly homozygous recessive (aabb) gives rise to long fruits (see Fig. 2.11). Thus, in the cross between the sphere-shaped parentals *AAbb* and *aaBB*, the F_1's are disk-shaped and are doubly heterozygous *AaBb*. Thus an F_2 ratio of 9:6:1 is obtained.

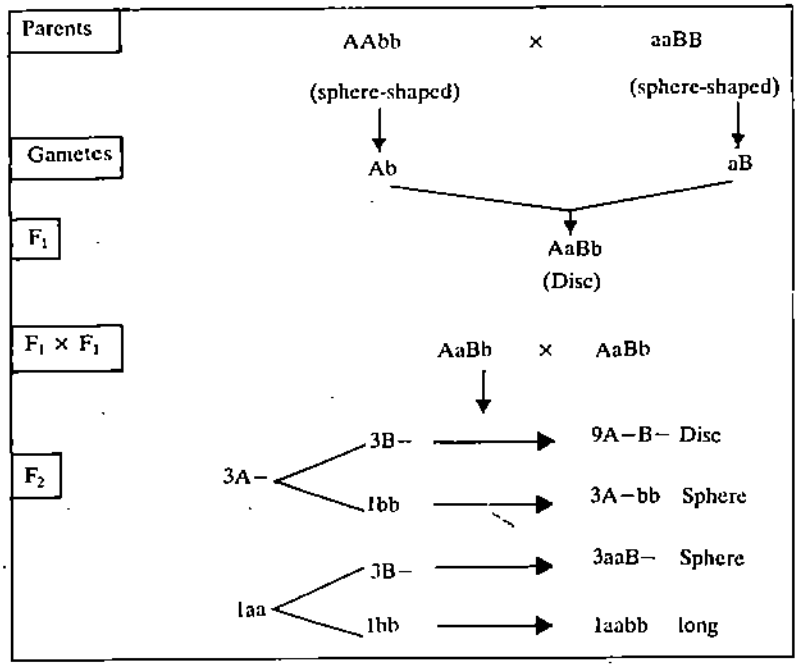


Fig. 2.11 : Results of crosses showing duplicate interaction of genes of fruit shape in summer squash.

The precise biochemical bases for the different shapes of squash fruit are not known fully. It is believed that in the absence of $A-$ and $B-$ allele products in the aabb squash, the functions of other genes determine the long fruits. If either the $A-$ or the $B-$ allele product, *but not both*, is present, the fruit is sphere-shaped. The disc fruit shape presumably results from the interaction of the $A-$ and $B-$ allele products. Since each of the dominant allele by itself results in the same phenotype but when both dominant alleles are present the phenotype is different, it is called duplicate interaction.

2.4.2 Epistasis

Epistasis is the interaction of nonallelic genes in which one gene masks the expression of another gene. Genes whose expression is masked by nonallelic genes are said to be *hypostatic. Epistasis may be caused by either the presence of homozygous recessives of one gene pair so that aa masks the effect of the B– gene. or may result from the presence of the dominant allele in a gene pair.* For example, *A* might mask the effect of the *B* gene. Moreover, the epistatic effect need not be just in one direction, as in the examples we just mentioned. *Epistasis can occur in both directions between two gene pairs.* With all these possibilities there can be quite a number of modifications of 9:3:3:1 ratio. Let us examine a few of them.

Recessive Epistasis : Coat Colour in rodents (9:3:4) : In one type of epistasis *aaB –* and *aabb* individuals have the same phenotype, so that overall phenotypic ratio is 9:3:4. An example is coat colour in rodents. Ancestral coat colour of mice, for example, is greyish as seen in ordinary wild mice. The other type of coat colouration is due to the presence of two pigments in the fur. Individual hairs are mostly black with narrow yellow bands near the tip. Such a colouration is known as agouti. It has a camouflage function and is found in many wild rodents including the wild rabbit, the guinea pig, the grey squirrel and wild mice. Another variation of coat colour is albino, in which the coat is white and the eyes are pink because of the absence of pigment in the iris. Albinos are true breeding and this variation behaves as a complete recessive to any other colour. Another variation is black coat colour which is produced as a result of the absence of yellow pigment from the agouti pattern. Black also breeds true and is recessive to agouti. When black mice are crossed with albinos, the F_1 progeny are all agouti. When these F_1 progeny consist of approximately 9/16 agouti animals, 3/16 black, and 4/16 albinos.

It is now known that three gene loci are involved in causing the colour phenotypes discussed above. At one locus the dominant C allele specifies a product that is necessary *for the production of pigment in the coat;* the *recessive c allele,* when *homozygous, prevents pigment formation* and *hence* the mice are *albino.* At a second locus the *dominant allele A specifies a product* that *determines* the *agouti factor.* Its recessive allele *a* is present in the homozygous state in all nonagouti mice, such as

blacks. The dominant allele *B* of the third locus specifies a product that governs the synthesis of black pigment. The recessive allele *b*, when homozygous, results in brown pigment. This latter locus is relevant to the example only when the basic hair colour is black, therefore, all mice involved must have atleast one *B* allele. The results can be interpreted as follows: we assume that the parental blacks are *aaCC* and the parental albinos are *AAcc*. The F_1 progeny obtained from them would be *AaCc*, which are agoutis since the genes *A* and *C* are responsible for the production of agouti pattern and colour respectively. In the F_2 animals the *A—C—* are agouti, the *aaC—* are black, and the *A—cc* and *aacc* are albino, giving a 9:3:4 ratio.

Duplicate Recessive Epistasis : Flower Colour in Sweet Peas (the 9:7 ratio) : The sweet pea occurs in a number of true-breeding varieties, most of which have descended from a purple-flowered wild sweet pea of Sicily. *Purple flower colour is dominant to white and gives a typical ratio of 3:1 in the F_2.* The white-flowered varieties breed true, and crosses between different white varieties usually produce white-flowered progeny. In some cases, however, crosses of two-breeding white varieties give only purple-flowered F_1 plants. When these F_1 hybrids are self-fertilised, they produce an F_2 generation consisting of about 9/16 purple-flowered sweet peas and 7/16 white-flowered peas.

Now, the next probable question is, how does this ratio occur? The *homozygosity of recessive alleles of any one gene results in white flowers. When both genes have one or more dominant alleles, the phenotype is purple.* Because nine of the sixteen categories in Fig. 2.12 are like this, there is a 9:7 ratio is the F_2 progeny.

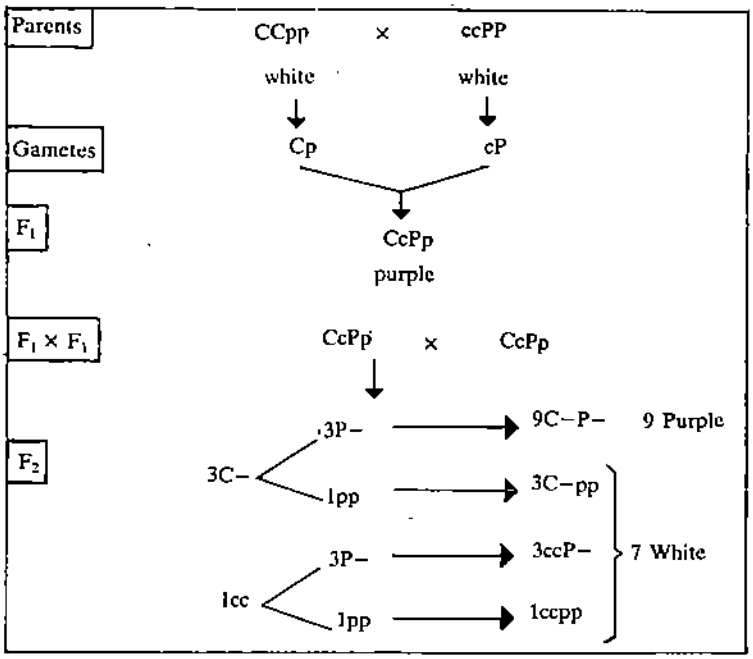


Fig. 2.12 : A cross between two types of sweet peas, showing the 9:7 ratio as seen in the F_2 generation.

Let us now understand the mechanistic basis of these phenotypes. Pigmentation in the flowers of sweet peas results from chemicals called anthocyanins. Production of anthocyanins occurs through a series of metabolic steps catalysed by enzymes that are themselves products of genes. Whenever a step in this synthetic process is blocked by the absence of a functional enzyme, pigmentation does not occur. A general pattern of pigment synthesis in sweet pea is illustrated in Fig. 2.13. From the figure it should be clear to you that *the enzyme products of gene C and gene P are both necessary for the production of anthocyanins.* When the genotype is *cc*, the first enzyme is not produced (see b_2 in Fig. 2.13); consequently, reaction to the intermediate product is stopped. Likewise, if the genotype is pp (see b_1, in Fig. 2.13), the lack of second enzyme blocks the second metabolic step. In other words, if the genotype is either *cc* or *pp*, the synthetic pathway is blocked and no pigments are produced resulting in white flower. Such an *interaction of genes to jointly produce a specific product is a type of epistasis called complementary gene interaction.*

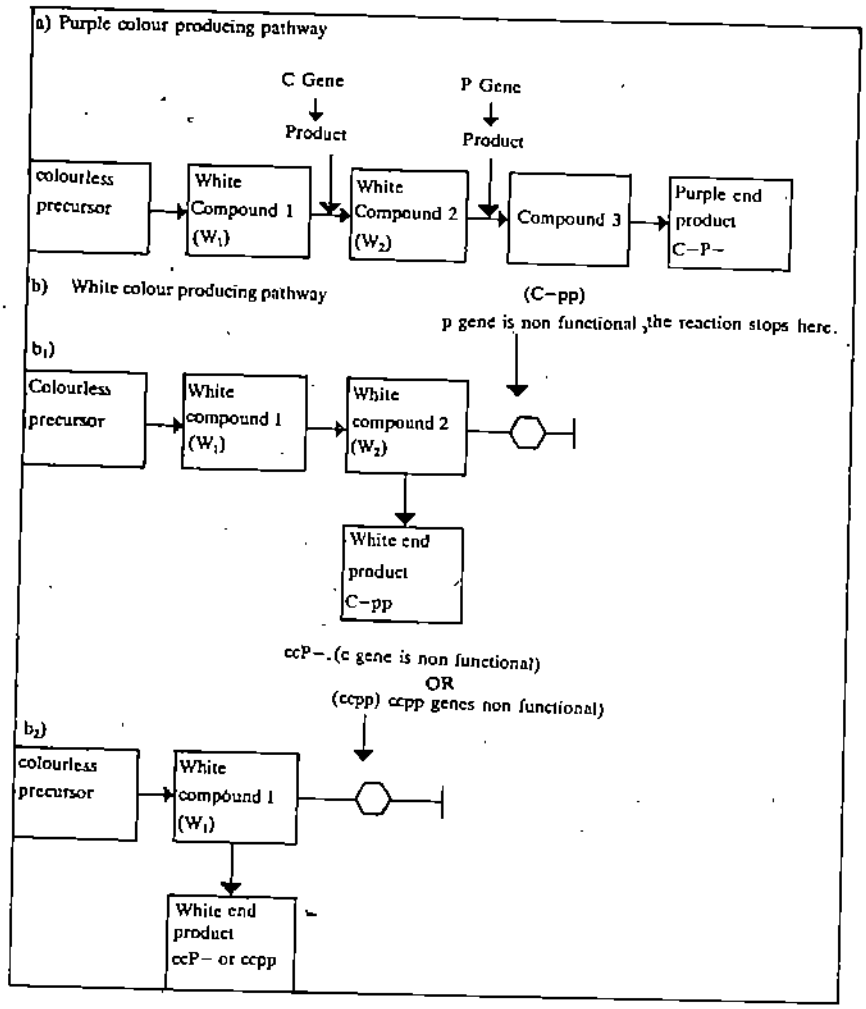


Fig. 2.13 : Pathways (a and b) for the production of purple and white colour respectively in the flowers of sweet pea. Note the relationship of genes *C* and *P* in the production of the pigment.

You would study later in the course that, many recessive diseases in humans are caused by metabolic errors. If the two individuals with metabolic errors for the same trait (but in different steps of a biochemical pathway) marry, their offspring should have a normal phenotype. For example, when two individuals with deaf-mutism caused by different recessive genes mate and have offspring. In the pedigree as shown in Fig. 2.14, II-3 and II-4 are the only two affected individuals because they are recessive for different genes. Note that the pedigree is different from that of a single recessive gene. Here, none of the progeny from a mating of two affected individuals are affected.

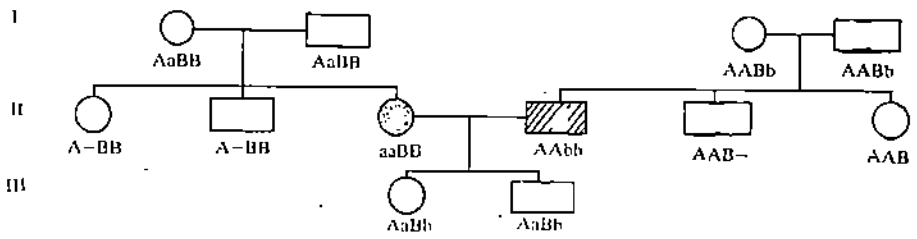


Fig. 2.14 : A pedigree illustrating a recessive disease caused by two genes. Only II-3 and II-4 are affected.

Duplicate Dominant Epistasis : Pod shape in Sheperd's Purse (15:1 ratio) : The 9:7 ratio is one of the types of the F_2 pattern caused by epistasis. In other situations the F_2 ratio may be 13:3, 9:3:4, 15:1 and so on. Notice that the different categories of a 9:3:3:1 ratio can be combined to obtain these ratios. The biochemical basis for the 15:1 ratio can be understood by examining one step of a pathway in which a dominant allele at either of two genes is enough to produce enzymes for catalysis of a given reaction (see Fig. 2.15).

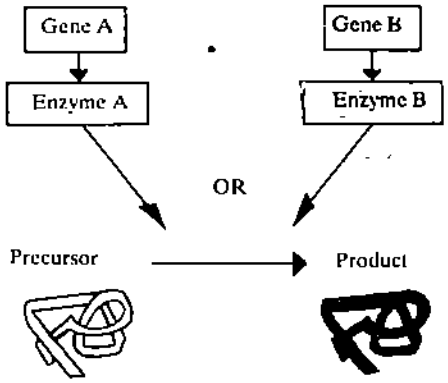


Fig. 2.15 : Duplicate gene action. Note, the enzyme from either A or B can catalyse the biochemical reaction.

Only when there is a double recessive, say *aabb*, is the pathway *blocked* and the aberrant phenotype expressed. *When either of the two genes can function to produce the dominant phenotype,* the gene action is called *duplicate gene action.* Sheperds' purse – a weed, offers a classical example of duplicate gene action. In this plant, *round seed pods* result *when dominant alleles are present at either of the two genes.* On the other hand *narrow seed pods* are *produced in the absence of dominant alleles at both loci. Hence it is an illustration of duplicate dominant epistasis.*

Dominant Epistasis : Fruit Colour in Summer Squash (12:3:1 ratio) : There are commonly three colours of summer squash fruits : *white, yellow* and *green.* In crosses between *white* and *yellow* and between *white* and *green, white* is always expressed. In crosses between *yellow* and *green, yellow* is expressed. *Yellow* thus acts as **recessive in relation** to *white,* but is *dominant in relation* to *green.* In other words, **there is an allele for** *white* **that is epistatic to those of** *yellow* **and** *green.*

a) *White* × *yellow*
All *white*

b) *White* × *green*
All *white*

c) *Yellow* × *green*
All *yellow*

White > *yellow* > *Green*
> = *dominant over*

Consider two gene pairs, *Ww* and *Yy.* In squashes the fruit with *W*– genotype is white no matter what the genotype is at the other locus. In *ww* plants, the fruit will be *yellow* if there is a dominant allele at the other locus, and *green* if it is not there.

SAQ 4

What will be the phenotype of :
W–*Y*– ....., *W*–*yy* ..... *ww*Y– and *wwyy* ..... plants?

The F_2 progeny, when an F_1 doubly heterozygous individual (*WwYy*) is self-fertilised shows a 12:3:1 ratio of *white : yellow : green* fruits (see Fig. 12.16). This example shows *complete dominance at both gene pairs, with one gene (the white one here) when dominant is epistatic to the other.*

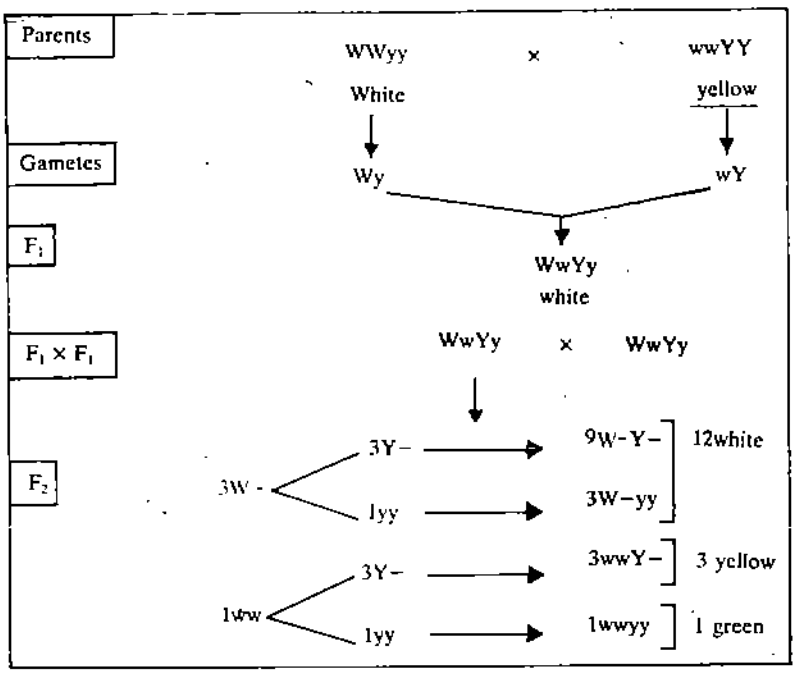


Fig. 2.16 : A cross between white and yellow fruits showing the 12:3:1 ratio in the F_2 generation.

The pathway to explain the 12:3:1 ratio is shown in Fig. 2.17. See the pathway b) in the figure first. The *white* substance *is converted to yellow* end product via a *green* intermediate. Both the steps are controlled by genes, the *green* to *yellow* step being controlled by the *Y* gene. The F_2 phenotypic ratio will be generated if the following conditions are present: i) the dominant allele *W* is nonfunctional thereby no inhibitor is produced, and hence the green intermediate is formed; ii) the dominant *Y* allele is functional that brings about the conversion of *green* to *yellow* pigment. Thus all the summer squash plants that have atleast one *W* allele will be *white*-fruited no matter which alleles are present at the *Y* locus, since no *green* substance can be made in this case (see Fig. 2.17 a). As a result, 12/16 of the F_2 squashes are *white*-fruited and have the genotypes *W*—*Y*— and *W*—*yy*. Also 3/16 of the F_2's are *yellow*-fruited and have the *ww Y*— genotype (see b). The *green* substance is made since there is no inhibition of the *white*-to-*green* step. As the functional *Y* allele product is present that catalyses the conversion of *green* substance to *yellow* substance. Lastly, 1/16 of the F_2's have *green* fruits and have a genotype *wwyy* (see the pathway *c* in the figure). Again, the *green* substance is produced because there is no inhibition of the *white*-to-*green* step, but due to nonfunctional *Y* gene, the green substance cannot be further converted to the *yellow* substance.

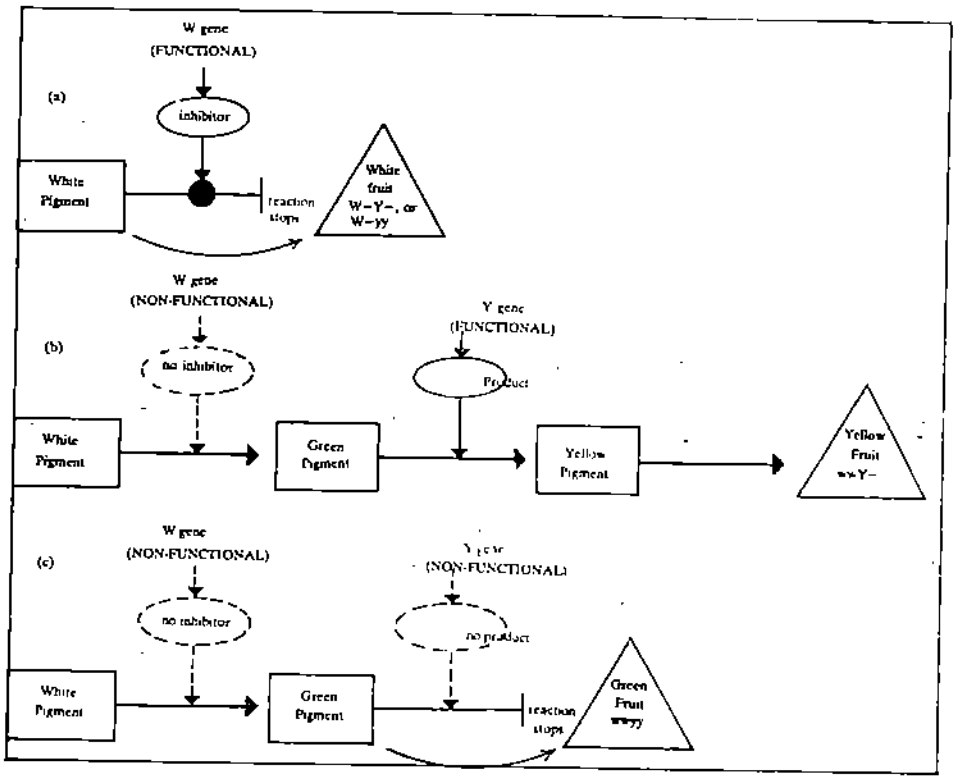


Fig. 2.17 : Pathways to show the production of : a) *white*, b) *yellow* and c) *green*-coloured fruits.

From the above examples, it should be clear to you that many types of interactions are possible between the products of gene pairs. Geneticists detect such interactions when they observe modifications of the expected phenotypic ratios in crosses. We have discussed some examples in which two nonallelic genes assort independently and in which complete dominance is exhibited in each gene pair. The ratios we discussed would necessarily be modified further if the genes did not assort independently or if incomplete dominance or codominance prevailed.

SAQ 5

Tick the correct answers.

a) Mice with the genotype *BB* and *Bb* are black and those with the genotype *bb* are brown. At another locus the genotypes *CC* and *cc* code for colour while the

homozygous recessive genotype *cc* codes for albinism. What phenotypic ratio would be expected from a cross between the mice with the genotype *BbCc*?

i) 9 black : 3 brown : 4 albino
ii) 12 black : 3 brown : 1 albino
iii) 9 black : 6 brown : 1 albino
iv) 9 black : 0 brown : 7 albino

b) What is the mode of inheritance in question (a) above?
i) codominance
ii) epistasis
iii) multiple allele
iv) none of the above

c) A 15:1 phenotypic ratio is observed in an F_2 generation. This ratio is characteristic of :
i) duplicate recessive epistasis
ii) recessive epistasis
iii) duplicate dominant epistasis
iv) both i) and ii)

2.5 LETHAL ALLELES

In the previous section, you studied the various gene interactions resulting in the modification of the ratio of F_2 individuals. The genetic ratios are also affected by several other factors. One of them is a class of genes – lethal genes, which you will study in this section.

Genes may affect viability as well as visible traits of an organism. The living beings carrying certain genes are disadvantaged as they have impaired structural as well as biochemical functioning. For example, the *Drosophila* flies having white eyes and vestigial wings have lower viability than their wild types. The detrimental physiological effects are apparently associated with the genes involved, that is, *w* and *vg* respectively. Some other genes have no effect on the appearance of a fly but do influence viability in some ways. Other genes have such serious effects that the organisms is unable to live. These are called **lethal genes** and the alleles involved in the situation are termed as **lethal alleles**.

If the lethal effect is dominant and immediate in expression, all individuals carrying the gene will die and the gene will be lost. Some **dominant lethals**, however, have a delayed effect so that the organism lives for some time. **Recessive lethals** present in the heterozygous condition have no effect but may come to expression when matings between carriers occurs.

We shall now take up an example for discussion, that clearly illustrates the functioning of these genes. In 1904, shortly after the rediscovery of Mendel's principles, a French geneticist, Lucien Cuénot, while carrying out experimental crosses on coat colour in mice, found that a gene was not consistent with the mendelian predictions. He observed from his experiments that the yellow body colour allele (*Y*) was dominant and agouti allele (*y*) was recessive. The crosses between two yellow mice (see Fig. 2.18) yielded approximately a 2:1 ratio of yellow to agouti mice rather than the expected ratio of 3:1. Further, when the yellow individuals (Yy) are crossed to the agouti (*yy*) Cuenot found that some agouti progeny are produced. He, therefore, concluded that yellow mice were heterozygous (*Yy*) and there were no yellow homozygotes (*YY*) in the progeny. Later, it was suggested that the yellow homozygotes were actually lethal, and they died while still in the uterus. It was found that approximately ¼ of the embryos from yellow by yellow crosses failed to develop. Therefore, the observed ratio of phenotypes differs from the expected ratio, as they die very young – much before reaching the reproductive age.

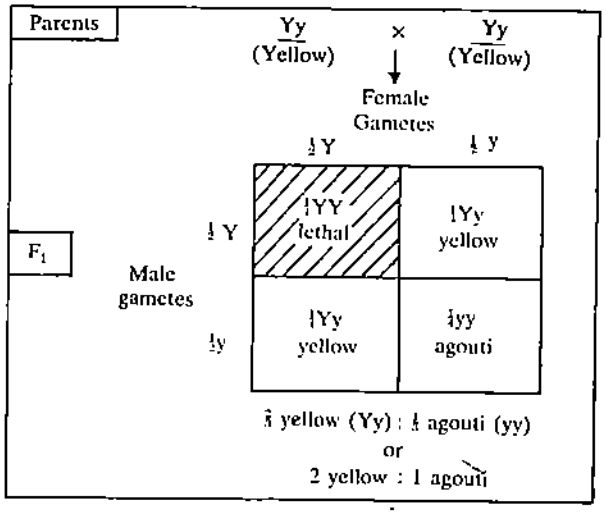


Fig. 2.18 : A cross between two yellow mice, yielding a 2:1 ratio in the offspring.

Such lethals are by no means exceptional and must always be considered in populations of plants and animals. Many lethals produce no pronounced effect at all on the phenotype, but they may make their presence known by a decrease in the life span or the very elimination of the carrier. It has been estimated that each human carries, on the average, about six lethal alleles.

How can an allele have a killing action? This perhaps may be the question arising in your mind. You may remember that metabolism is the result of many interlocked biochemical pathways. A defect in just one step can upset several others. Just one defective step can alter the entire chemistry of the body. A lethal, by blocking a critical reaction, can interfere with normal embryological development of any organ, say heart. The death of embryo may then follow. Lethals can thus decrease the chances of survival by causing various kinds of abnormalities in development and physiology.

Different lethals eliminate individuals at different stages of the life cycle. The **complete lethal** removes the carrier before the reproductive age so that those affected have no offspring, e.g., the allele for yellow coat colour in mice; in humans the recessive factor for **Tay-Sachs disease** which kills in infancy. In humans, the dominant factor for **Huntington's disease**, a fatal, deterioration of the nervous system, does not usually express itself before the age of 30. Such genetic determinants which can result in death but permit the carrier to live to reproductive age, are often grouped as **sublethals**. There is actually no sharp boundary during the life cycle at which lethals act.

Tay-Sachs disease (amaurotic idiocy) : A genetic recessive disorder that affects the central nervous system. Its clinical symptoms are : cherry-red spot in eye macula (the visible white portion of the eye); after 6-9 months rapid degeneration of vision and motor skills; death at about 2-4 years of age.

Huntington's disease (HD) : A fatal neurological disorder which is dominant. It is normally manifested after the age of thirty, but has been reported to occur at all ages. HD is characterised by mental and physical deterioration. There is progressive change in personality. As the disease progresses, the HD patients demonstrate twitching and uncontrollable muscle spasms.

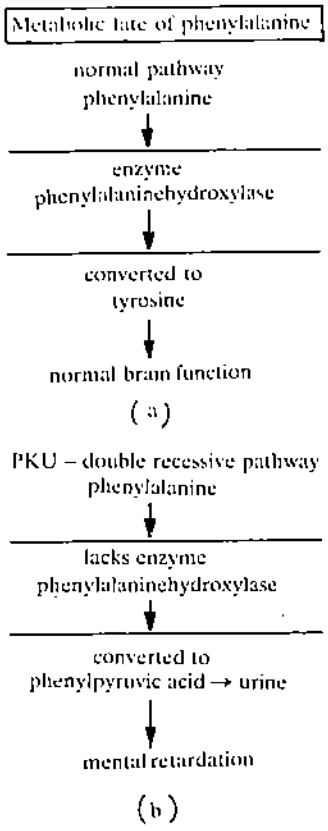
There is degeneration of Central Nervous System. (CNS) and loss of brain cells. This leads to fits of depression, insanity and suicide. At the time of death, the patients have lost about 25 per cent of their brain weight.

2.6 PLEIOTROPY

The action of a gene at the cellular level is unitary, that is, one gene one action. Sometimes the presence of a gene results in a broad spectrum of phenotypic changes, so that it appears that the **gene has multiple action.** This phenomenon is called **pleiotropy,** and is found primarily in higher organisms where complex and interrelated developmental events occur.

Many lethal alleles are pleiotropic. For example, the yellow coat colour in mice, just discussed, is an allele that affects more than one character, that is, it produces yellow colour of the coat in heterozygotes, and it also affects survival, causing lethality in homozygotes. Another example of multiple effects is the gene affecting seed shape in garden peas; this gene also affects starch grain morphology. In fact, many genes affect more than one trait. Mendel also noticed that genes causing the flower colours, like violet and white, also influenced seed colour and caused the presence or absence of coloured areas on the leaves. This is due to pleiotropy as a single gene affects more than one character.

Pleiotropic traits also occur in humans. One such disease is **phenylketonuria,**



abbreviated as **PKU**. This occurs in individuals that are homozygous for a defective, recessive allele. The diseased people lack the enzyme necessary for the metabolism of the amino acid phenylalanine. When normal and PKU individuals are compared, the level of phenylalanine is much higher in diseased group. In addition to this basic biochemical difference, a number of other features are seen in the untreated PKU patients, such as lower IQ, similar head size and lighter hair.

2.7 SEX-LINKED GENES

In the crosses we have discussed so far, it has not mattered which parent carried a particular allele in question. Reciprocal crosses gave identical results. For example, in all the Mendel's monohybrid crosses, F_1 and F_2 were the same regardless of which P_1 parent exhibited the recessive trait.

This, however, is not the case in crosses involving genes located on chromosomes involved in the determination of sex. Such genes are said to be **sex-linked.** The first thorough study of the sex-linked gene was conducted by T.H. Morgan in 1910, when he was studying the inheritance of eye colour in *Drosophila.* The wild-type flies have brick red eye colour. Morgan crossed mutant white-eyed flies with red-eyed flies. He found that the result of a cross between a white eyed male (*r/Y*) and red-eyed female (*R//R*) were different from those of a cross between white-eyed male. Both of these reciprocal crosses are illustrated in Fig. 2.19. (crosses a and b).

Compare the phenotypic ratios of F_2 generations of cross a with b. Is there any difference? Yes, there is a clear difference between the F_2 individuals of the two crosses. The phenotypic ratios thus depend on whether the P_1 white-eyed parent was male and female. Morgan was able to correlate these observations with a difference

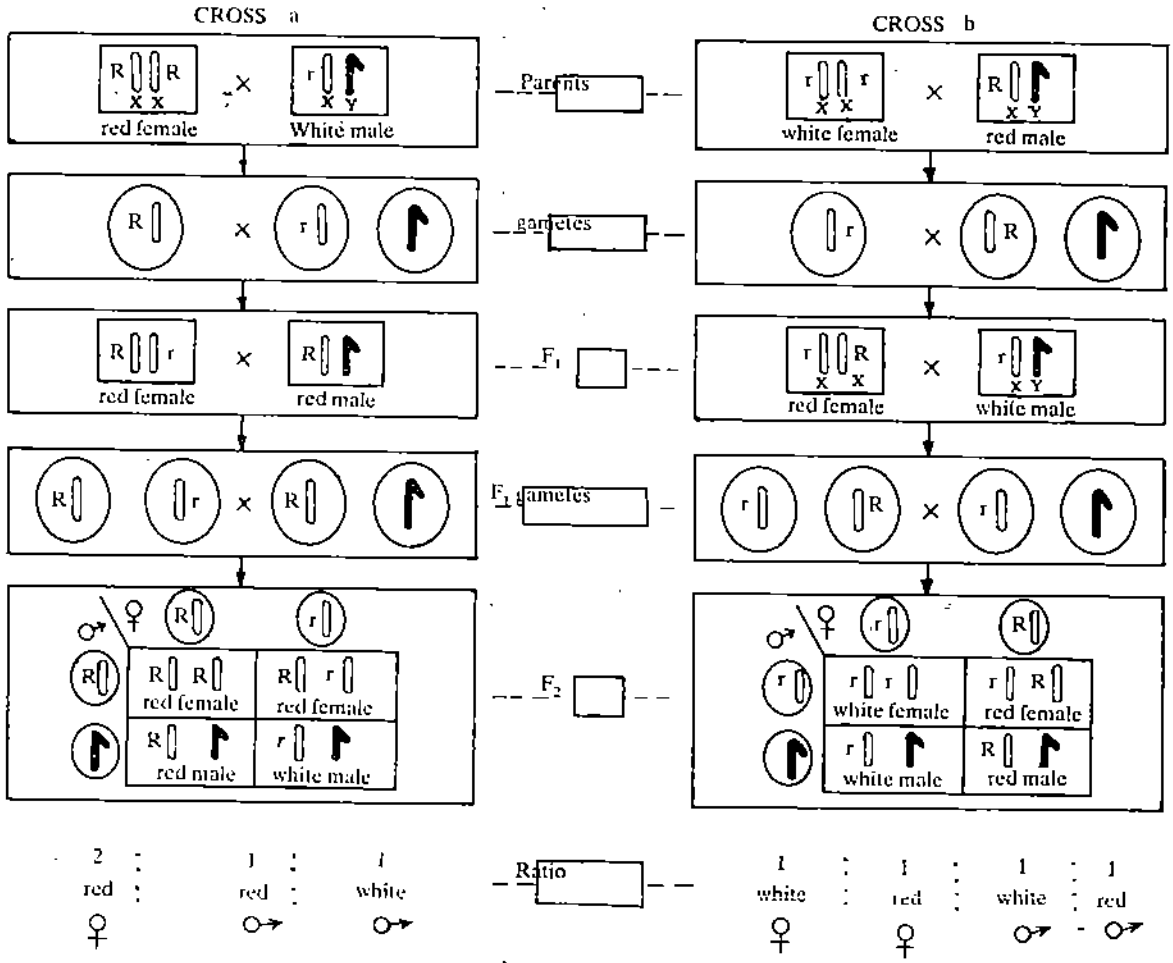
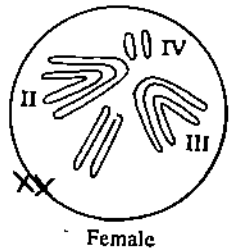


Fig. 2.19 : Crosses a and b illustrate inheritance of sex-linked genes in *Drosophila*. a) is a cross between red-eyed female and white-eyed male. And b) is the reciprocal cross of a). In the figure the long, rod-shaped structure (I) represents the x chromosome and the inverted J-shaped structure (ſ) stands for the Y chromosome.

found in chromosome composition between male and female *Drosophila*. It has *four* chromosomes and one of the chromosome varies between sexes (see Fig. 2.20). This chromosome pair (see XX or XY in the figure) is involved in sex determination mechanism and constitutes the **sex chromosomes.** The remaining chromosomes are called **autosomes.** The females possess two, rod-shaped homologs called the X chromosomes which are designated as XX. Males possess a single X chromosome and J-shaped Y chromosome, which are designated as XY.

On the basis of this correlation, Morgan postulated that the gene for white eye is present on the X chromosome but is absent from the Y chromosome. Females thus have two available genes, one on each X chromosome while males have only one available gene on their single X-chromosome. This explanation supposes that the Y chromosome lacks homologous loci to those on the X chromosome. However, the X and Y chromosomes still behave as homologues in that they do partially synapse with each other and segregate into gametes during meiosis.

Many species, including humans, possess an arrangement of sex chromosome as in *Drosophila*. Many sex-linked genes in humans have now been identified, e.g., the gene controlling one form of haemophilia and muscular dystrophy. We stop further discussion on this topic for the time being and shall resume it in an elaborated manner in Unit 4, that is, Sex Linkage and Dosage Compensation.



Female

Male

Fig. 2.20 : Chromosomes composition in *Drosophila melanogaster.*

2.8 DEGREES OF GENE EXPRESSION

In observing offsprings from a cross, we tend to think of a phenotype in terms of all or none of the phenomena. A trait is expressed or it is not — and we draw conclusions as to what are the types of genotypes based on that form of expression. While it is true that some phenotypes are always certain, such as, in pea plants the genotype *ss* will produce always the wrinkled seeds. Similarly, all of the Mendel's genes and all blood group types show an absolute and clear cut pattern. However, for some genes, the expression is variable.

2.8.1 Penetrance

Some individuals fail to show a particular trait, even though genetic analysis indicates that the controlling gene for that trait must be present in them. This aspect of gene action, that is *the frequency with which a genotype is expressed in the phenotype is called* **Penetrance.** For example, out of the eight individuals of a particular genotype, five express the diseased phenotype. The penetrance is $5/8 = 0.625$ or 62.5%.

A **completely penetrant gene** is always expressed; an incompletely penetrant gene may be expressed in some individuals, but not in others.

Incomplete penetrance of a gene may cause a trait to skip a generation; that is, a dominant trait present in a given generation may skip the first generation but appear again in the second generation. Such a case is illustrated in Figure 2.21. It shows a

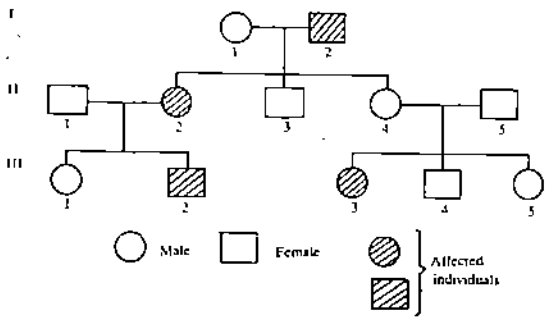


Fig. 2.21 : A pedigree illustrating the pattern of phenotypes with an incompletely penetrant dominant allele.

human pedigree where unaffected individual II-4 is the daughter as well as the mother of an affected individual each. This indicates she has the genotype *Aa*, but because of incomplete penetration, does not show the dominant phenotype.

2.8.2 Expressivity

Expressivity refers to the *intensity or range of expression of a trait in different parts of the same individual or different individuals.* This aspect of gene action differs from penetrance in that it describes the level of phenotypic expression, whereas penetrance refers to whether the phenotype is affected or not. Both penetrance and expressivity originate from variations in the degree to which a gene is manifested in the phenotype. These manifestations can be absent or be so slight as to pass unnoticed; in such individuals, a trait would be considered **non-penetrant**. However, when apparent in the phenotype, the same trait may vary in its effects from mild to severe, in which case the trait would be described as exhibiting not only **variable penetrance**, but **variable expressivity** as well.

Let us consider the example of *Drosophila*. In this the recessive allele eyeless (*e*) when present in homozygous condition causes a reduction in the size of the compound eye. However, the extent to which the eye is reduced varies considerably from individual to individual. In some, only slight decrease in eye size occurs, while in others the eye may be completely absent. There are also cases in which the fly has one normal eye and the other eye drastically reduced in size. Phenotypic variation can also be observed in humans. Polydactyly exhibits variable penetrance, it is also characterised by differences in expressivity. In different individuals showing this trait, extra digits may be present on the hands or on the feet or both hands and feet.

The source of this variation is partly genotypic and partly environmental. The genotype contains thousands of genes, and the actions of many of them are interrelated so as to modify one another's effects. For example, the level of expression of a trait is generally more similar among relatives than among unrelated individuals, provided that the relatives and unrelated individuals are raised in fairly similar environments. Such *genes that have secondary effect on a trait are called* **modifier genes**, and can sequentially influence the phenotype. The modifier genes can be seen in animals like house cat, where a '**dilute allele**' reduces the intensity of pigmentation from black to grey. Another example of modifier genes is seen in *D. melanogaster* mutants when these are kept in laboratory culture for many years, sometimes they do not have as extreme a phenotype as was first observed.

2.9 ENVIRONMENTAL MODIFICATION OF GENE EXPRESSION

You should not forget that activities of all the genes are taking place in a cellular environment, which in turn may be influenced by the external environment. In this section, we shall discuss the changes in phenotype that are induced by the environment.

The genotype establishes a fixed potential for the expression of any trait, but whether or not, and to what extent that potential is utilised, depends on the environment, both internal and external. For example, it is obvious that the potential normal growth cannot be achieved under conditions of deprivation. Factors such as climate, light, temperature, moisture, minerals, nutrients in diet, vitamins, hormones, and other aspects of the total environment play a major and permissive role in gene expression. The wild-type organism, be it a plant or an animal, presents the wild-type phenotype under congenial environmental conditions. However, the same organisms when placed in a harsh and restrictive environment must adapt its physiology for survival. As a consequence, the degree of expression of a given gene may change, leading to a corresponding change in some aspect of the phenotype. Thus, although the genotype is fixed, the response of the genotype to varying environmental conditions is flexible to the extent that component genes can vary in their range of expression.

Some genes, such as the blood group alleles, exhibit all-or-none expression regardless of the environment. Others, especially those with variable penetrance and expressivity, may or may not be expressed. A genotype with variable penetrance may not happen to encounter the environmental conditions necessary for the development of the trait. Hereditary fructose intolerance in humans is an example. Individual with this recessive trait are normal and completely healthy unless they happen to eat foods containing fructose or sucrose, or a disaccharide sugar composed of fructose and glucose. The defect is a deficiency of the enzyme fructose-1-phosphate aldolase, and ingestion of cane sugar or fruits and vegetables containing fructose results in nausea and vomiting in persons homozygous for the deficiency. A similar example is — sitosterolemia, a recessive inherited condition characterised by the accumulation of plant sterols, especially — sitosterol, in the blood and tissues. Phenotypic expression of this disease is associated exclusively with diet. This disorder results from the increased intestinal absorption of plant sterols which normally are absorbed only in trace amounts. If vegetable fats and foods with high fat content, such as nuts and chocolate, are eliminated from the diet and replaced with equivalent foods from animal sources (meat, eggs, milk, butter), recovery from this disease follows.

Lifestyle and physical activity can also be considered as agents contributing to gene expression. An example is a rare disorder, called *McArdle's disease*, which stems from a deficiency of the enzyme phosphorylase in muscles. Persons with this trait experience painful muscle cramps when they engage in strenuous exercise. Exercise requires the utilisation of glycogen as an energy source, and the absence of muscle phosphorylase blocks the first step in glycogen degradation in exercising muscle. A more sedentary life style and the avoidance of muscular stress generally eliminate the phenotypic expression of the trait.

Another interesting aspect of gene expression is the presence in many species of temperature sensitive alleles whose products are active only within a certain defined temperature range. One such temperature sensitive allele is found in rabbits. The coat colour of rabbit, as you know, is determined by a series of multiple alleles. One of these alleles, the recessive c^h determines a striking pigmentation pattern called Himalayan having white fur over most of its body, but black fur on the nose, ears, paws and tail (also see Fig. 2.2). The c^h is temperature-sensitive and, when homozygous, is capable of initiating melanin pigment formation at temperatures below 34°C. Since the extremities are ordinarily somewhat cooler than the rest of the body, black fur results. However, if the extremities are artificially warmed or the rabbit is kept at a higher temperature the new hair growth is white. Again, if some region of the body is artificially cooled the new growing hair are black. Other members of this allelic series, chinchilla, agouti and albino do not exhibit temperature sensitivity in their phenotypic expression. Instances of temperature-dependent gene expression can also be found in *Drosophila*. In this a genotype that produces blisters on wings is much more extreme in its expression at 19°C than at 25°C.

In the context of humans, monozygotic twins separated at birth provide a useful test situation. **Monozygotic** or **identical twins** are derived from the division and splitting of single zygote just after fertilisation. In this case both the derivatives of the egg are identical in their genetic composition. Although most identical twins are reared together and are exposed to very similar environments, some pairs are separated and raised in different settings. For any particular trait, average similarities or differences can be investigated. Such an analysis is particularly useful because characteristics that remain similar in different environments are believed to have a high genetic component. These data can then be compared with a similar analysis of **dizygotic** or **fraternal twins**, that originate from two different zygotes. Dizygotic twins are thus no more genetically similar than any two siblings.

Let us examine some characteristics of the twins that are reared together. See Table 2.2 for the various characters studies. Twins are said to be **concordant** for a given trait if both members express it or neither express it, and **disconcordant** if one shows the trait and the other does not. In the above table, the values of various traits in both types of twins are listed. Have a look at the table again and compare the values listed under MZ with those of DZ. If you see the concordance value that approaches 90 to 100 per cent in monozygotic twins, it indicates a large genetic contribution to the expression of the trait. This can be seen in traits like blood type and eye colour.

Table 2.2 : Comparison of concordance for various traits between monozygotic (MZ) and dizygotic (DZ) twins (From Klug & Cummings, 1986)

Trait	Concordance	
	MZ (%)	DZ (%)
Blood Types	100	66
Eye Colour	99	28
Mental retardation	97	37
Idiopathic Epilepsy	72	15
Schizophrenia	69	10
Diabetes	65	18
Allergy	59	5
Tuberculosis	57	23
Cleft Lip	42	5
Club Foot	32	3
Mammary Cancer	6	3

You have just seen how identical twins provide a good material for studying environmental modifications of gene expression. However, the best indications of environmental effects come from studies of species that can be raised in a large number of controlled environments.

A classic example of such a study was provided for several plant species in the 1940s by J. Clausen and his coworkers. To determine the effect of the environment they grew clones (genetically identical individuals) from cuttings at different altitudes. For example, they collected *Potentilla glandulosa*, a member of the rose family, from three locations : situated at heights of 100, 1,400 and 3,040 meters respectively. From these samples, they generated clones by vegetative propagation and grew them at three different locations. In other words, genotypes collected from the three locations were grown in each of the three environmental conditions (100m; 1,400m and 3,040m).

Table 2.3 shows the results of these experiments. Each row represents a genotype from the different areas, and each column indicates the environment in which these plants were grown. From this table it is apparent that both genetic and environmental factors are important in determining plant growth. For example, the plants from 100 m altitude grow well in both 100 m and 1,400 m altitudes but cannot even survive at high altitude that is, 3,040 m (see the top row, extreme right). This indicates the importance of environmental factors. On the other hand, the genotype B, grows much better than the other genotypes at the high altitude of 1,400 m indicating the importance of genetic factors.

Table 2.3 : Three genotypes of *Potentilla* plants each cloned and grown at three different altitudes.

Plants grown at / Plant specimens taken from	100m	1400m	3040m
Genotype A from 100m	 Optimum growth	 Lesser growth, forming a compact structure	Fail to survive
Genotype B from 1400 m	 Stunted growth few flowers	 Optimum growth	 Extremely reduced growth
Genotype C from 3040 m	 Reduced growth	 Fairly good growth fewer flowers	 Optimum growth.

The way in which the phenotype changes, for a given genotype, according to its environment is called the **norm of reaction.** Figure 2.22 gives norms of reaction similar to those of the low and middle altitude genotypes in Table 2.3. Genotype A, the low altitude clone, does best at the low end of the environmental spectrum and worst at the high end. The genotype B, the middle altitude clone, does best at intermediate altitude levels; falling off in phenotypic value in both low and high altitude environments.

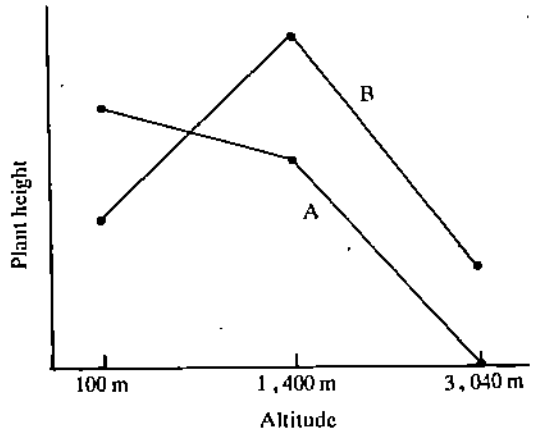


Fig. 2.22 : The norms of reaction for two different genotypes A and B.

Phenocopy : The term phenocopy designates the individuals whose phenotype has been altered by the environment in such a way that it simulates the phenotye usually associated with a particular genotype. Factors like temperature shocks, radiation, poisons, vitamin deficiencies, and a wide range of chemical agents, affect the normal organisms in many cases and result in production of new phenotypes. These phenotypes resemble some known genotypes such as the mutants. All these varied phenotypes produced are the phenocopies. A phenocopy can be produced only if the treatment is applied during a sensitive stage in development; application at other stages being ineffective. Presumably, such sensitive periods mark a turning point or critical period in embryonic determination or differentiation. We shall explain this with a phenotype — tetraptera in *Drosophila.* This fly has only one pair of wings, and the second pair is reduced to small knob-like structures known as the halters, that serve as the balancing organs. The recessive gene tetraptera in homozygous condition leads to the development of these small balancers into wing-like structures when a wild-type embryo is treated by heat shocks or exposure to ether, a phenocopy of the teraptera condition is produced (see Fig. 2.23). *The sensitive period during which treatment is effective occurs a few hours after fertilisation,* although the wing itself does not differentiate until much later, at the time of metamorphosis.

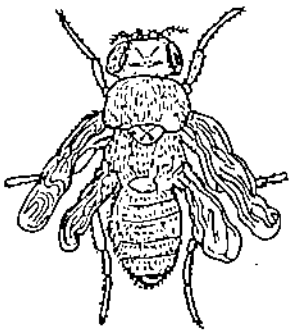


Fig. 2.23 : A tetraptera phenocopy in *Drosophila melanogaster.*

Besides heat shocks, many other agents too result in phenocopies. Even vitamin deficiencies are able to do so. For example, hens fed on a diet deficient in biotin produce chicks characterised by malformations of the long bones and beak. This phenotype is practically indentical with that which results from the homozygosity of the recessive allele for chondrodystrophy (see Fig. 2.24).

Chondrodystrophy — a hereditary deforming disease of skeletal system characterised by insufficient growth of long bones resulting in reducing length.

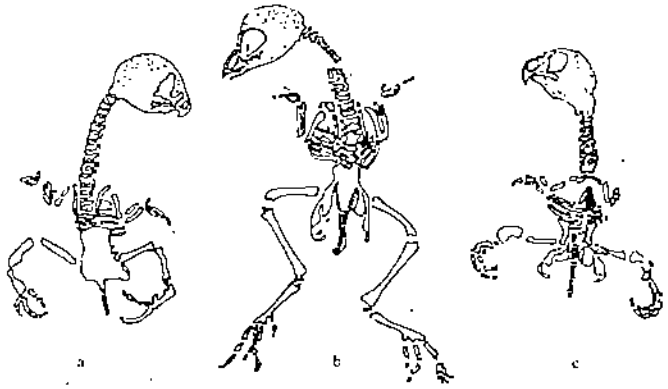


Fig. 2.24 : Skeleton of chick at the time of hatching. a) pattern of damage of the lethal factor chondrodystrophy; b) normal condition; c) effect of biotin deficiency.

In humans too phenocopies are seen. These arise due to malnutrition, injury, disease or by accidental exposure to toxic chemicals. A deficiency of the intestinal enzyme

lactase is an example. Absence of this enzyme may be inherited as a recessive condition, rare in infants, but much more common in adults, where it arises from development loss. However, lactase is an enzyme extraordinarily sensitive to any form of intestinal damage, and therefore, the disorder may also be acquired as a result of a variety of diseases that affect the intestines. Acquired cases of lactase deficiency can thus be considered phenocopies of the inherited trait.

One phenocopy caused by exposure to chemicals is a condition called **hyperoxaluria** in which calcium oxalate is deposited as crystals in tissues and as stone in the kidney. This phenotype is found in persons with an inherited recessive defect in the metabolism of glyoxylate. The individual phenotype also results from accidental poisoning with oxalate or ethylene glycol.

Other agents capable of inducing phenotypic changes are hormones, whose actions as chemical coordinators and growth stimulators are pervasive, affecting every organ system of the body. Changes in the level of various hormones, brought about by accident, disease or medical treatment, rather than by genotype, can cause profound alterations in metabolism, appearance, and behaviour.

To summarise, the expression of any gene can rarely be considered as an activity separated and distinct from the environment in which it occurs. Although gene action is usually compared or described under normal conditions, it must be remembered that different degrees of expression may be evoked by significant departures from the standard environment. It is the response of the genotype to such differing environments that make possible the adaptation and survival of the individual.

SAQ 6

Tick the correct answers.

a) Pleiotrophy occurs when a gene has:
   i) a complementary gene elsewhere
   ii) a small effect on the trait
   iii) reversible effects on the phenotype depending on age
   iv) many effects on the phenotype.

b) All children homozygous for an allele that causes Tay-Sachs disease develop symptoms of this disease. This allele has a penetrance that is:
   i) deep
   ii) shallow
   iii) 0%
   iv) 100%

c) A woman receives her X chromosomes from:
   i) her mother only
   ii) her father only
   iii) both her mother and father
   iv) DNA in the cytoplasm of her mother's egg

d) A man receives his X chromosomes from:
   i) his mother only
   ii) his father only
   iii) both his mother and father (part from each)
   iv) either his mother or father

e) How can a phenocopy and a mutation be differentiated?
   i) by crossing an unaffected individual with an affected individual
   ii) by crossing an affected individual with a carrier
   iii) by crossing two affected individuals
   iv) none of the above

2.10 SUMMARY

In this unit you have learnt that:

- Dominance occurs in different forms. It may be complete, as found by Mendel,

where the same phenotype results whether an adult is heterozygous or homozygous. The recessive allele is phenotypically expressed when it is homozygous. Dominance may be incomplete, with the heterozygote having a phenotype in between the two homozygotes, or two traits may be codominant, with both alleles expressed in the heterozygote.

- A gene can exist in several different states or forms—a situation called multiple allelism. The alleles are said to constitute an allelic series, and the members of a series can show any type of dominance relationship with one another.
- In many plants and animals, nonallelic genes do not function independently in determining the phenotypic characteristics. In some cases, interaction between gene products results in new phenotypes without modification of typical mendelian ratios. In other type of gene interaction, called epistasis, interaction between gene products causes modification of mendelian ratios because one gene product interferes with the phenotypic expression of another nonallelic gene or genes. The phenotype is controlled by the former gene and not the latter when both genes occur together in the genotype.
- A lethal gene is one that when expressed is fatal to the individual. Recessive lethal and dominant lethal genes exist, and they can be sex-linked or autosomal.
- Pleiotropy is seen when one gene affects several traits.
- Genes on the X and Y chromosome have different patterns of inheritance from those of autosomal genes. X-linked recessive traits occur almost entirely in males and generally skip generations in families.
- Individuals possessing the same genotype may show different degrees of phenotypic expression. Some genes exhibit variable penetrance in that not all genes possessing the gene express it. Even when completely penetrant, expressivity may be variable, and not all individuals or all structures of the same individual are affected to the same extent.
- Environment, in all the aspects, both external and internal, plays a major role in determining the degree to which genes are expressed. In addition the action of many environmental agents on developmental processes and other functions can produce phenocopies that mimic those produced by known genotypes.

2.11 TERMINAL QUESTIONS

1) In snapdragons, red flower colour (R) is incompletely dominant to white flower colour (r) and the heterozygote has pink flowers. Also, normal broad leaves (B) are incompletely dominant to narrow, grass-like leaves (b), the heterozygote has an intermediate leaf breadth. If a red-flowered, narrow-leaved snapdragon is crossed with a white-flowered, broad-leaved one, what will be the phenotypes of the F_1 and F_2 generations, and what will be the frequencies of the different classes?

2) This question refers to the inheritance of ABO blood groups in humans.
   a) Which alleles for blood group are codominant?
   b) What does the term codominant mean?
   c) What blood groups are possible in children of parents with type A and type B blood? Explain your answers.

3) When a white cow was mated with a red bull, all their offspring were a mottled red-and-white (roan) colour. If the two roan cattle were mated what offspring colour would be produced and in what ratios. Explain your answers.

4) In rabbits, a multiple allelic system controls coat colour in the following way: C is dominant to all other alleles and causes full colour. The Himalayan phenotype is caused by c^h and is dominant to all alleles except C. The chinchilla phenotype is caused by c^{ch} and is dominant to the recessive allele c, which results in albinism. Thus the order or dominance is $C > c^h > c^{ch} > c$. Shown below are a series of P_1 crosses, and from each the phenotype of one of the F_1 offspring. In each case, F_1 offspring from two separate crosses are mated. Determine the F_1 genotypes in each case. Then predict the phenotype ratio of offspring from the mating.

a) Himalayan × Himalayan → Himalayan ? → .....
Full colour × Himalayan → Full colour

b) Full colour × chinchilla → chinchilla ? → .....
Full colour × Himalayan → Himalayan

c) Full colour × chinchilla → Full colour ? → .....
Chinchilla × albino → albino

d) Full colour × chinchilla → Full colour ? → .....
Full colour × Himalayan → Full colour

e) Albino × albino → albino ? → .....
Chinchilla × chinchilla → chinchilla

5) Mrs Ram and Mrs Ganapathy had babies at the same hospital. Mrs Ram took home a girl, Anjali and Mrs Ganapathy a boy, Parag. However, Mrs Ganapathy was sure that she had a girl and she filed a suite against the hospital. Blood tests showed that Mr Ganapathy was type *O*, Mrs Ganapathy was type *AB*; and Mr and Mrs Ram were both type *B*. Anjali was type *A* and Parag was type *O*. Had an exchange occurred?

Anjali _____ _____ Mrs Ganapathy _____ _____
or Mr Ganapathy _____ _____
_____ _____ Mr Ram _____ _____ or _____ _____
Parag _____ _____ Mrs Ram _____ _____ or _____ _____
Draw arrows to indicate possible parents.

6) Which blood group type of the *ABO* system is known as the universal donor? The universal recipient? Explain your answers.

7) What is the difference between dominance and epistasis? Why didn't any genes used by Mendel show epistasis?

8) Whether the oat grains are red or non-red is determined by the dominant allele *R* and its recessive allele *r*. Non-red grain (*rr*) may be yellow or white, these characteristics being determined by a dominant allele *Y* (yellow and its recessive allele *y* (white). A cross between plants heterozygous for both pairs of genes produced 101 red, 23 yellow and 8 white offspring.
   a) What ratio do the data represent?
   b) What genotypes could be found among the red offspring?
   c) What genotypes would be found among the yellow offspring?
   d) What proportion of the red individuals will produce only red progeny if allowed to self fertilise?
   e) What proportion of yellow progeny will produce only yellow individuals if allowed to self fertilise?

9) How may a classic mendelian 9:3:3:1 ratio be converted into a) 9:7 or b) 15:1 ratio?.

10) In horses, suppose that there are three colours affecting genes with the following effects : *WW* is lethal, *Ww* is white, and ww allows colour; *BB* or *Bb* is black, and *bb* is chestnut; *OO* or *Oo* is solid and *oo* has white spots on colour. A white stallion and a white mare, each heterozygous for all these genes are mated. What are the expected frequencies of possible phenotypes in the live offspring? What frequencies of phenotypes would be expected if the stallion were WwBboo?

11) In foxes, two alleles of a single gene, *P* and *p* may result in different phenotypes in different combinations. *PP* results in lethality, *Pp* in platinum coat and *pp* in silver coat. What ratio is obtained when platinum foxes are interbred? Is the *P* allele behaving dominantly or recessively in causing lethality? In causing platinum coat colour?

12) In *Drosophila*, a sex-linked recessive mutation, scalloped (*sd*) causes irregular wing margins. Diagram the F_1 and F_2 results if: a) scalloped male is crossed with a normal female. b) Compare these results to those obtained if scalloped were not sex-linked.

13) In the snapdragon, flower colour may be red, white or pink. Flower shape may be personate or peloric. From the data of the following crosses, determine how flower colour and shape are inherited.

a) red peloric × white personate ⟶ F_1 = all pink personate
b) red personate × white peloric ⟶ F_1 = all pink personate
c) pink personate × red peloric ⟶ F_1 = ¼ red personate
¼ red peloric
¼ pink personate
¼ pink peloric
d) pink personate × white peloric ⟶ F_1 = ¼ white personate
¼ white peloric
¼ pink personate
¼ pink peloric

What phenotypic ratios would result from crossing the F_1 of a) to the F_1 of b) ?

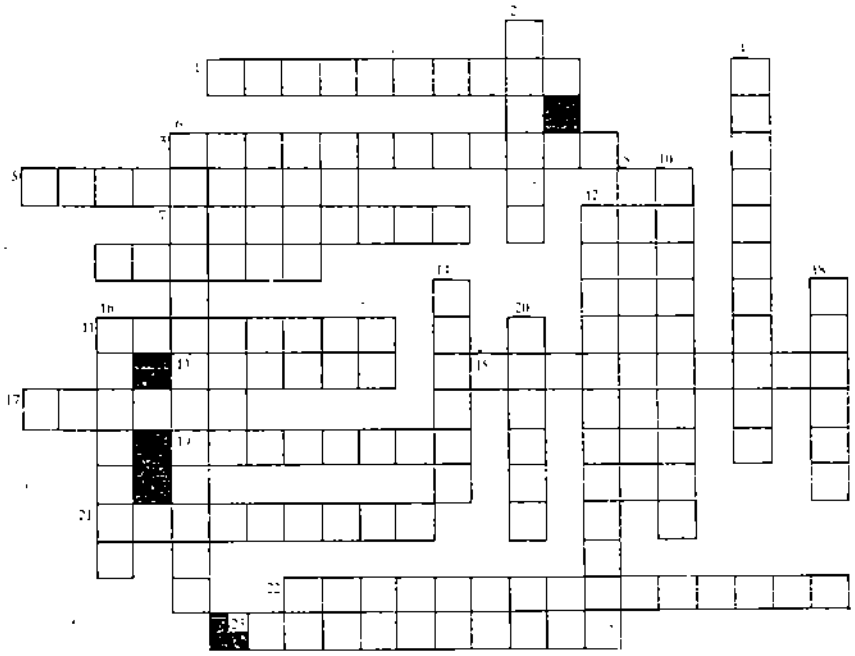
14) Discuss the difference between penetrance and expressivity of a trait. Draw a pedigree for a trait such as Huntington's disease, illustrating variable expressivity. Assuming *H* and *h* are the dominant and recessive alleles respectively, give the genotypes and phenotypes for three generations (one set of grandparents in the first generation, at least 3 progeny and at least five grand children).

15) a) How can phenocopies be used to study gene action?
b) What values and limitations do they have for investigations of this kind?
c) How can an investigator determine whether an altered phenotype is due to a mutation or phenocopy?

16) In i) through v), match the terms in column A to their appropriate description in column B.

A	B
i) Dominance	a) alleles of a gene causing suppression
ii) Phenocopy	b) alleles of different genes causing suppression
iii) Epistasis	c) expression of two or more genes at different loci
iv) Pleiotropism	d) environmentally produced phenotypes
v) Independent Assortment	e) multiple, unrelated effects produced by a single gene

17) Drawing the information and terminology presented in this unit, complete the following crossword puzzle.



Across

1) fruit fly genus
3) purple pea flowers
5) fraternal twins

7) study of inheritance
9) alternate form of gene
11) immunological response to foreign substances
13) defence system of body
15) percentage of individuals that show a particular phenotype among those capable of showing it
17) result is zygote
19) non-sex chromosome
21) genes masking genetic expression of other genes
22) defective metabolism of phenylalanine
23) multiple phenotypic effects

Down

2) coat colour in rodents
4) heterozygote exhibiting phenotype of both homozygotes
6) incompatibility of red blood cells
8) overt genetic expression
10) both twins exhibit or fail to exhibit a trait under investigation
12) degree to which a particular genotype is expressed in the phenotype
14) Gregor
16) immunological stimulus
18) monkey
20) death

2.12 ANSWERS

Self-assessment Questions

1) A – I^AI^A, I^Ai^O; B – I^BI^B, I^Bi^O;
AB – I^AI^B; O – i^Oi^O;

2) a) multiple alleles
b) O
c) incomplete dominance
d) dominant vs recessive
e) 4

3) 1/16 *AABB* : 2/16 *AABb* : 2/16 *AaBB* : 4/16 *AaBb* : 1/16 *AAbb* : 2/16 *Aabb* : 1/16 *aaBB* : 2/16 *aaBb* : 1/16 *aabb*

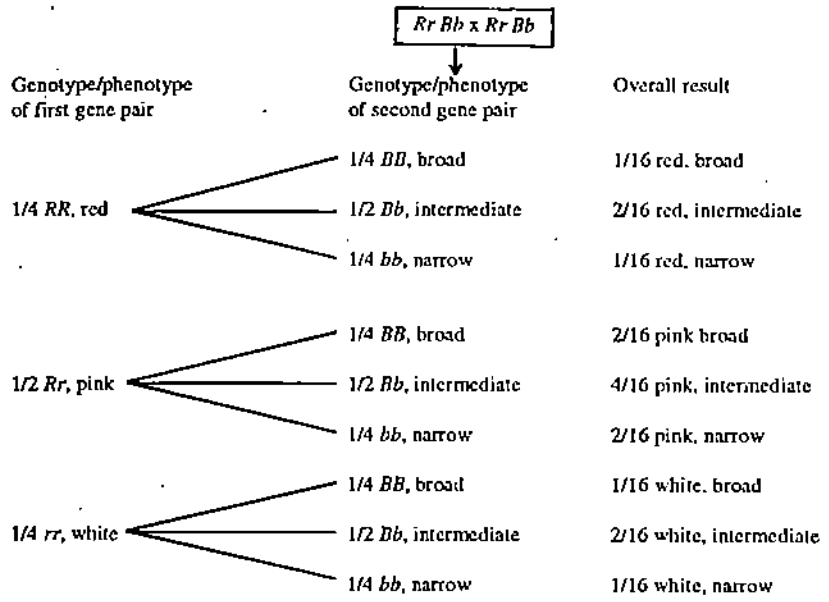
4) *W–Y–* = White
W–yy = White
wwY– = Yellow
wwyy = Green

5) a) i
b) ii
c) iii

6) a) iv
b) iv
c) iii
d) i
e) i

Terminal Questions

1) This is based on incomplete dominance. First assign genotypes to the parental phenotypes. Thus, the red, narrow plants are RRbb and white, broad-leaved plants rrBB. The F_1 of these two parents would be a double heterozygote RrBb. Because of incomplete dominance, these plants are pink-flowered and have leaves of intermediate breadth. Interbreeding F_1s give F_2 generation shown below. They do not have 9:3:3:1 ratio. Instead there is a different phenotype for each genotype as shown in the figure.



2) a) I^A and I^B (gene for A and B antigens)
   b) Codominance is a situation when both alleles for a given character are fully expressed in heterozygous individual.
   c) All the four blood groups (A, B, AB and O) are possible in the children. Type A parents (genotype I^AI^O) and type B parents (I^BI^O) produce the gametes containing I^O, I^A and I^O, I^B genes, which combine to give A (I^OI^A), B (I^OI^B), AB (I^AI^B) and O (I^OI^O) genotypes.

3) 1 white : 2 roan : 1 red
   The genes show incomplete dominance (see the figure below)

White		Red		Roan
WW	×	*RR*	⟶	WR
WR	×	WR	⟶	WR : WR : WR : RR
Roan		Roan		1 White 2 Roan 3 Red

4)

F_1 Genotypes	F_2 Phenotypes
a) $c^hc^h \times Cc^h$	½ colour : ½ himalayan
b) $c^{ch}c^{ch} \times c^hc^h$	All himalayan
c) $Cc^{ch} \times cc$	½ colour : ½ chinchilla
d) $Cc^{ch} \times Cc^h$	¾ colour : ¼ himalayan
e) $cc \times c^{ch}c^{ch}$	All chinchilla
or	or
$cc \times c^{ch}c$	½ chinchilla : ½ albino

5) Yes, Anjali is I^AI^A or I^AI^O and this matches with the Ganapathys because only they could have an I^AI^O child. Parag matches with Rams only.

6) Type O is the universal donor; type AB is the universal recipient. Type O has no antigens and therefore, is accepted by all types. Type AB has no antibodies and therefore, can accept all types of blood.

7) Dominance concerns the interaction to determine a phenotype of two alleles at a given locus, while epistasis involves interaction of genes at different loci. All of Mendel's genes affected different traits.

8) a) 12:3:1 phenotypic ratio
   b) red offspring : *RRYY, RRYy, RRyy, RrYY, RrYy, Rryy*
   c) yellow offspring : *rrYy, rrYY*
   d) Only those homozygous for red (*RR*) will produce red progeny exclusively upon selfing. These are *RRYY, RRYy,* and *RRyy*; they comprise $^4/_{12}$ or $^1/_3$ of the F_1 red individuals.
   e) Only those homozygous for the dominant *Y* will produce yellow progeny exclusively upon selfing. These are *rrYY*; they comprise $^1/_3$ of the F_1 yellow individuals.

9) When epistasis is operating between two loci the number of phenotypes in F_2 generation is less than four. When two phenotypes are seen in a 9:7 ratio, duplicate recessive genes are acting to produce identical phenotypes; two dominant genes are necessary to complement each other and express the trait. For example, if both *A* and *B* were needed for pigment formation, the following results would be expected:

P_1 *AAbb* (colourless) × *aaBB* (colourless)

↓

F_1 *AaBb* (coloured)

F_2 *9 A—B— : 3A—bb : 3aaB— : 1aabb*

9 coloured : 7 colourless

When a 15:1 ratio is seen, duplicate dominant genes are producing the same phenotype without a cumulative effect. When either of the dominant allele is present the trait is expressed. If the genes suppressed pigment formation, the following would result:

P_1 *AAbb* (colourless) × *aaBB* (colourless)

↓

F_1 *AaBb* (colourless)

F_2 *9 A—B— : 3A—bb : 3aaB— : 1aabb*

15 colourless : 1 coloured

10)

	Cross 1	Cross 2
White	32/48	16/24
Solid Black	9/48	3/24
Spotted Black	3/48	3/24
Solid Chestnut	3/48	1/24
Spotted Chestnut	1/48	1/24

11) 2/3 platinum, 1/3 silver. The *P* allele behaves as a recessive lethal, but as a dominant in its influence on coat colour.

12) a) F_1 — All ♂ scalloped : All ♀ normal
F_1 — ½ ♀ scalloped : ½ ♀ normal
½ ♂ scalloped : ½ ♂ normal

b) Assuming the female is homozygous, the F_1 consists of all normal ♂s and ♀s. In the F_2, 3 normal : 1 scalloped in both ♂♂ and ♀♀.

13) Here, white and red are the two homozygotes and pink is the heterozygote. In flower colour there is incomplete dominance. Personate is dominant to peloric. F_1 (a) × F_1 (b) cross : 3/16 red, personate: 1/16 red peloric; 6/16 pink personate; 2/16 pink peloric; 3/16 white personate; 1/16 white peloric.

14) Penetrance is the proportion of individuals of a given genotype that manifest a phenotype, while expressivity is the degree of expression of the trait in those who are supposed to have it.

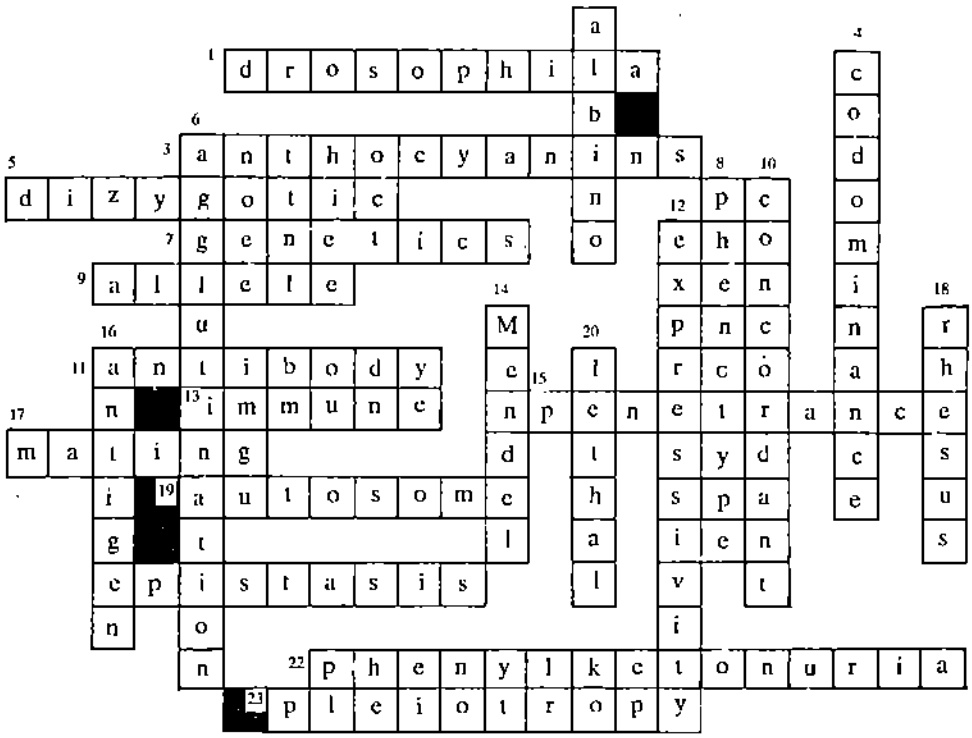
15) a) Environmental factors may influence developing phenotypes in the same way that genes do. Studies of phenocopies suggest mechanisms of gene action.

b) Phenocopies provide approach to the study of physiological genetics. Such studies, however, are indirect, and can only provide suggestions about the mechanisms of gene action.

c) Mutations are inherited; phenocopies are not inherited.

16) i) a
ii) d
iii) b
iv) c
v) c

17)



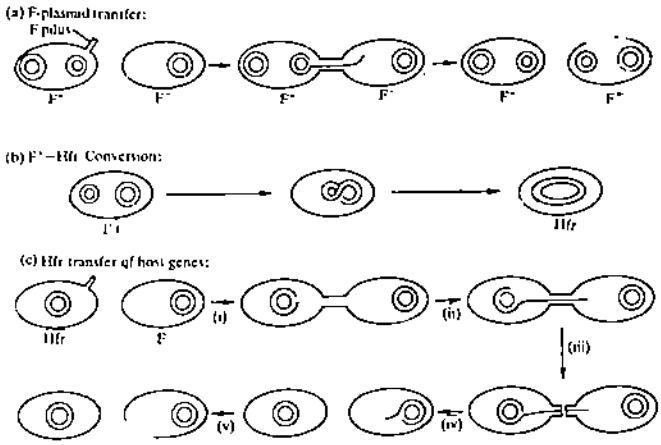


Fig. 3.1: Reproduction in bacteria. a) The F factor or plasmid passes from one cell to the other through F pilus, leaving a copy in the donor cell. b) F^+ to Hfr conversion. The F plasmid combines with the host chromosome. c) Transfer of Hfr genes into the recipient cell. i) First, the F pilus establishes a bridge between the conjugating cells. ii) Hfr chromosome begins replicating, with one daughter DNA molecule passing across the bridge to the recipient cell. iii) The bridge breaks during conjugation, leaving a piece of donor DNA in the recipient cell. iv) The donor DNA combines with the recipient chromosome. This may be a single-stranded event as shown here, or a double-stranded event. v) After the recipient cell replicates it contains a recombinant DNA, with part derived from the donor and part from the recipient.

3.2.2 In Eukaryotes

In most eukaryotes, especially higher animals, individuals normally exhibit one or two sex phenotypes; **female** or **male.** In such species, females produce the **female gametes — eggs, ovules** or **macrospores** and males produce the **male gametes — sperm, pollen** or **microsproes.** Species with separation of sexes in different individuals are called **dioecious** or **monosexual** organisms. All higher organisms and some higher plants are dioecious. Species in which both male and female gametes are produced by each individual are called **monoecious** or **bisexual** organisms. In lower animals, the production of both eggs and sperms by the same organisms is more commonly called **hermaphroditism,** and individual organisms producing both the types of gametes are termed **hermaphrodites.**

Although the two sex phenotypes are usually quite easily distinguished in humans and fruit flies, it is not universally the case. In lower or the 'primitive' eukaryotes, the two sexes are phenotypically indistinguishable except for their reproductive organs. Indeed, in lower eukaryotes the two genetically distinct types of gametes are sometimes morphologically indistinguishable. This is called **isogamy (iso** meaning 'same'). Isogamy occurs in several simple eukaryotes, such as the green alga *Chlamydomonas*, fungi–*Neurospora* and protozoa-*Paramecium*. They may however, be identified by their sexual reproduction pattern (see Figs. 3.2 and 3.3).

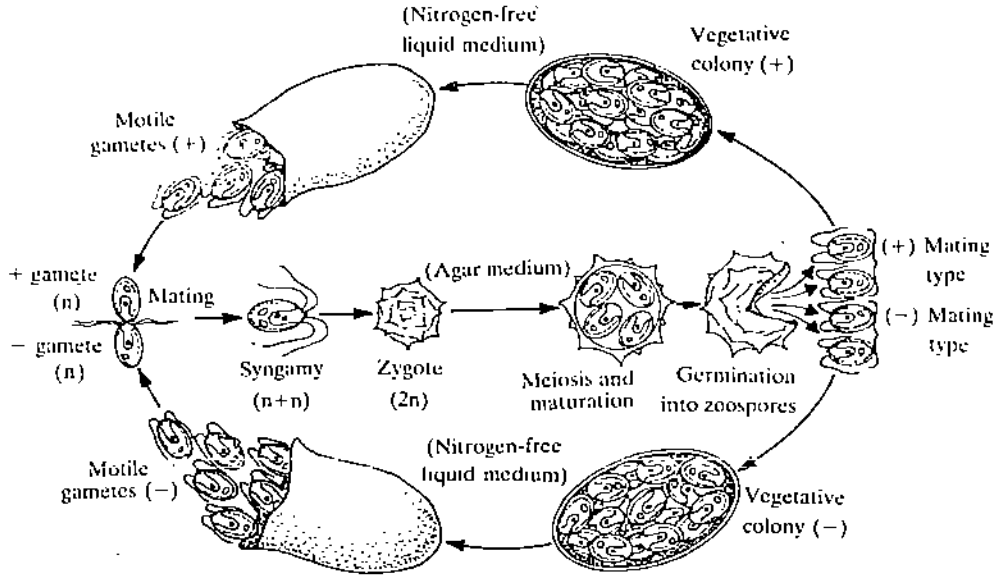


Fig. 3.2: Sexual reproduction in *Chlamydomonas reinhardtii* (After R. Sagar, 1972. *Cytoplasmic Genes and Organelles*, Academic Press, New Delhi).

An individual belonging to one mating type exchanges genetic material by fusing only with an individual of another mating type but never with its own mating type (see Fig. 3.4). Therefore, the similar looking male and female gametes, or isogametes are actually physiologically different, as is evident by their mode of reproduction.

Sexual differences between individuals probably originated first in their gametes. Most of the plants are hermaphrodite, producing both types of gametes, but have various adaptations to promote cross fertilisation. One such adaptation is the phenomenon of self-sterility. Its examples are cherry and tobacco plant. *Due to self sterility the plants have to undergo cross fertilisation,* and the *result is* the *recombination of genetic material.*

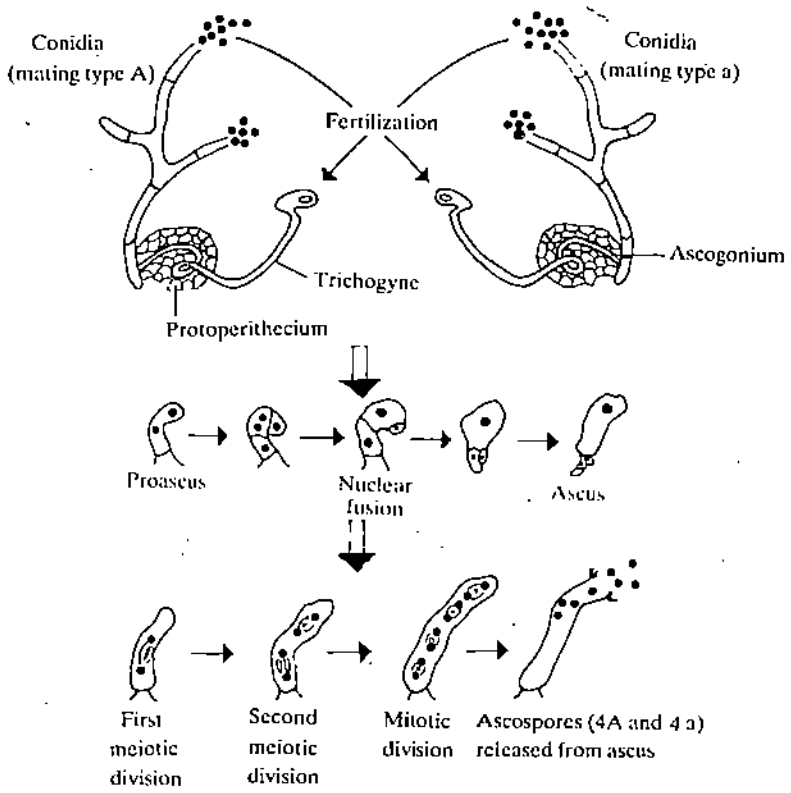


Fig. 3.3: Sexual reproduction in *Neurospora crassa* (Modified from Klug W.S. and Cummings, M.R. 1983. Concepts of Genetics. C.E. Merrill Publishing Co.

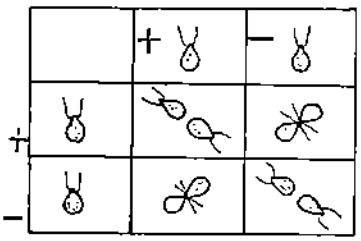


Fig. 3.4: Mating Types in *Chlamydomonas* sp.

3.2.3 Effect of Environmental Factors

In some lower organisms, though the genetic determiners of the sexes are present, the ambient environment plays a decisive role in the development of a particular sex type. And it appears that sex determination is non-genetic. The males and females have similar genotype, but stimuli from environmental factors initiate development towards one sex or the other. In *Equisetum* (horse tail) plant, for example, female characteristics develop when the plant is raised under favourable growth conditions, while in poor or unfavourable conditions males are formed.

Another salient example is of the marine worm *Bonellia* (see Fig. 3.5). These worms are very small. The males remain in a highly reduced form in the reproductive tract of female (Fig. 3.5, arrow). The female is many times larger than the male (compare their size in the figure). Any young worm reared from a single isolated egg becomes a female. If the newly hatched worms are released into water containing mature females, some young worms are attracted to females and become attached to the female proboscis. These are then transformed into males and these migrate to the female reproductive tract, where they become parasitic. Experimental studies have shown that even the extracts made from the female proboscis influence young worms towards maleness.

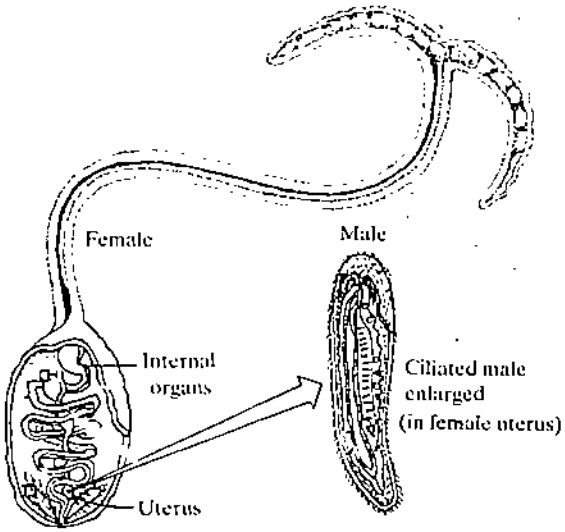


Fig. 3.5: The female and male of the marine worm *Bonellia viridis*. The male remains in highly reduced form in the uterus of the female (see arrow). Its enlarged version is drawn on the right (After Dobzhansky, T. 1955, *Evolution, Genetics and Man*., John Wiley and Sons, Inc.)

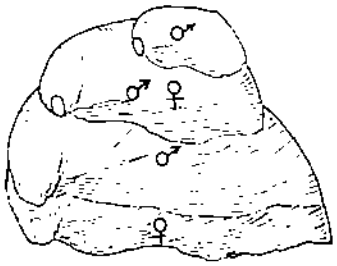


Fig. 3.6: *Crepidula* (slipper shells), a cluster. The second organism from the top is in the process of changing from male to female. When it undergoes the change and become ♀, it would be fertilised by the top organism which is a male.

In some reptiles, the temperature at which the fertilised eggs are incubated prior to hatching, plays a major role in determining the sex of the offspring. High temperatures during incubation have opposite effects on sex determination in different species. In the lizard *Agama agama*, some crocodiles and alligators, incubation at high temperature results in all male offspring while at low temperature it results in females. On the other hand, in the turtle, *Chrysema picta*, high egg incubation temperatures result in the production of mostly female progeny.

An interesting example is of the slipper shell *Crepidula*, where individuals tend to live stacked up on one another (Fig. 3.6). Young *Crepidula* are always male. However, as an individual ages, the male reproductive system degenerates. The reproductive system can then re-form as a male or become female, depending on the sexes of the other organisms in the cluster. If the organism is attached to a female, the reproductive system will redevelop as male. Isolation or the presence of a large number of males will induce a male to become a female. Once a female, the individual will no longer change.

These examples clearly illustrate the diversity of sex determintion mechanisms that exist in nature. Although the segregation on specific sex-determining genes and chromosomes is responsible for sex phenotype, but in most species such as *Bonellia, Chrysema picta* and *Agama agama*, some specific factor in the environment triggers the expression of either the genes producing a male phenotype or the genes producing a female phenotype. This knowledge has a great potential use in agriculture and conservation method (see Box 3.1).

Box 3.1: Sex Determination and Conservation

The information obtained from temperature-dependent sex determination is being used to conserve some species of sea turtles. Sea turtles spend virtually all their lives at sea, except for one night each year when the females laboriously haul themselves onto a beach to dig nests and lay eggs, before returning to the sea. Because turtles are so vulnerable at this time, many have been taken as food by humans at turtle nesting areas throughout the tropics and the turtle population has declined dramatically in the past century.

Conservationists began collecting eggs from sites, where endangered species nested. These eggs were incubated in captivity, but many of these early efforts were of limited success, because the eggs were incubated at only one temperature. All the artificially reared young were of one sex only.

Since the discovery of temperature-dependent sex determination, conservationists now incubate eggs at several temperatures or in thermally fluctuating environments that mimic the variations found at nesting sites in nature. Thus, young ones of both sexes are produced, and the future of these species is secure.

SAQ 1

Why is sexual reproduction favoured over asexual reproduction in nature?

.....

.....

SAQ 2

It is not uncommon to encounter self-sterile plants. What is the importance of this phenomenon?

.....

.....

3.3 GENETIC BASIS OF SEX DETERMINATION

In most organisms, the sex determining mechanism is under genetic control, free of environmental factors. In the genome of higher organisms, there are certain genes or chromosomes which regulates sex determination and thus sex is determined at the time of fertilisation.

3.3.1 Genic Type

In some organisms, certain **independent genes located on different chromosomes** are responsible for sex determination. Most species of fish illustrate this mode. The male may be represented as AA^{I} for the sex determining genes, and the female as AA (Fig. 3.7). According to this assumption, maleness is determined by the gene A^{I} which is dominant over A.

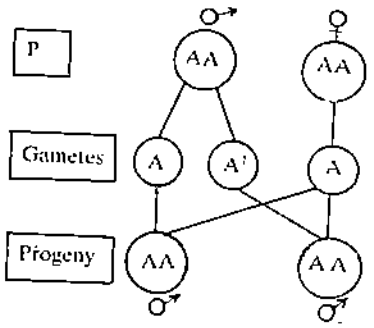


Fig. 3.7: Sex determination in fish at genic level.

In a parasitic wasp, called *Habrobracon*, genes are responsible for sex determination in a different way. The sex determining gene has multiple alleles (X^a, X^b, X^c..X^d). When two alleles in a zygote are different or heterozygous (X^aX^b, X^aX^c, X^bX^c, etc.) a normal, fertile female is formed. If the alleles in the zygote are in homozygous (X^aX^a, X^bX^b, etc.) or hemizygous (X^a,X^b, etc.) condition males are formed.

3.3.2 Chromosomal Type

It refers to the condition where the genes involved in sex determination are located on specific chromosomes known as the sex chromosomes.

First, we shall see how the genic type of sex determination evolved into a chromosomal type of sex determination mechanism. In the primitive forms, the only difference between the two sexes was in their gametes. Later in evolution, morphological or phenotypic difference developed in the two sexes of a species.

In primitive forms, sex determination was due to genes on autosomes (the genic type). In the process of evolution, gradually, the genes responsible for sex determination got localised on specific chromosomes – the *sex chromosomes*. These chromosomes were designated as 'X' and 'Y' or 'Z' and 'W', and they can usually be distinguished morphologically from each other. The remaining chromosomes of the complement are known as *'autosomes'* and are designated 'A'.

The X and Y chromosomes differ from each other in many respects. This is because, there is accumulation of sex determining genes on the respective sex chromosomes. Also, there is negligible crossing-over between the X and Y chromosomes. This helps to preserve gene combinations favouring distinct sexual differences. The consequence is that the Y chromosome bears mostly the genes essential for the male determination while all the other genes become inert. These regions got reduced in size in some species and are completely lost in others. This was how the heteromorphic sex chromosome evolved.

3.4 CHROMOSOMAL SEX DETERMINATION MECHANISMS

In most of the higher plants and animals, the chromosomal sex determination mechanisms are prevalent. Basically, five types of chromosomal mechanisms exist. These are XX-XY, ZZ-ZW, XX-XO, ploidy level, and compound chromosome system.

3.4.1 The XX-XY System

This is a common mode of sex determination in animals including man and some plants like the angiosperm genus *Lychnis* (see details in subsection 3.5.1). Both the sexes have equal number of chromosomes, of which one pair is of sex chromosomes. In females, the two sex chromosomes are similar and are called the X-chromosomes. Thus, the female is the **homogametic** sex ('homo' meaning same). Males in contrast usually possess an X chromosome, and one chromosome is dissimilar in morphology from all others and is known as the Y chromosome. Because the male sex chromosomes are different, the male is called the **heterogametic** sex ('hetero' meaning different).

In humans, the characteristic diploid chromosome number is 46 (Fig. 3.8). The females have 22 pairs of autosomes (AA) and a pair of X chromosomes (AA + XX). The males have 22 pairs of autosomes (AA) alongwith an XY pair (AA + XY). The sperms formed are either X-bearing or Y-bearing (Figs 3.8 and 3.9). This sex is

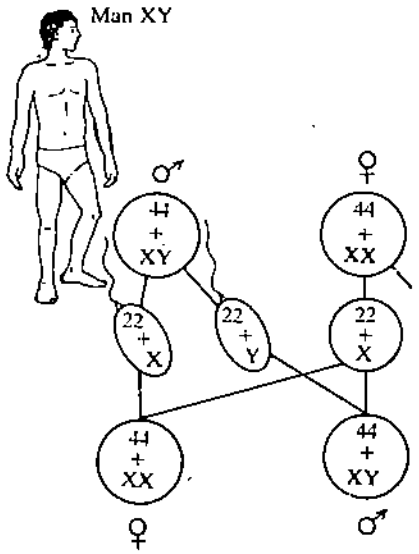


Fig. 3.8 : The XX-XY system of sex determination in man.

determined by the sperm because all eggs are similar and X-bearing. Studies on sex chromosomal abnormalities in man, helped in understanding the crucial role of Y-chromosome in determining maleness. It has been observed that a single Y chromosome, irrespective of the number of X-chromosomes present in the zygote, causes an individual to develop into a phenotypic male. And in the absence of a Y-chromosome, the zygote leads to femaleness. Genes on the Y chromosome direct

Fig. 3.9: Transmission of the sex chromosomes from parents to children.

differentiation of the embryonic gonad to a testis, whose hormonal products then induce a male phenotype. Genes on the Y chromosome also control spermatogenesis. Thus, the presence of the Y chromosome determines maleness, and without a Y chromosome, the female phenotype develops.

3.4.2 A Variation : The ZZ-ZW System

Under this system, the *male is the homogametic sex* and the *female is heterogametic.* In order not to confuse this type of sex determination with the XX-XY mechanisms, the male chromosomes are labelled ZZ, the female ZW. In this system the ovum determines the sex of the resultant offspring, because all sperms carry similar chromosomes. Other than the reversal of homo-, and heterogametic sexes, the ZZ-ZW system functions similar to the XX-XY system. This mode of sex determination has been observed in birds including domestic fowl (see Fig. 3.10), butterflies and moths, some fishes, reptiles and amphibians, and in a plant species *Frageria orientalis.*

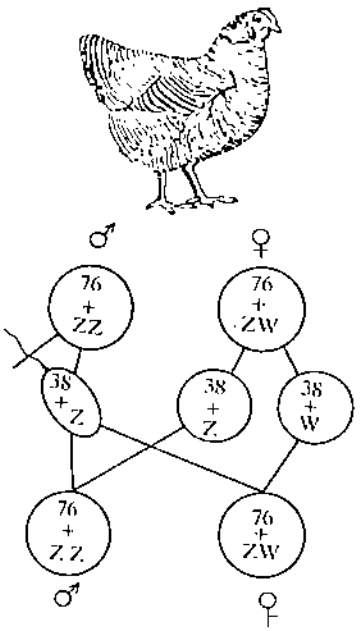


Fig. 3.10: The ZZ-ZW system of sex determination in domestic fowl.

3.4.3 The XX-XO System

In some species, the two sexes have different numbers of chromosomes. The difference often involves the sex determination mechanism. This phenomenon is called the XX—XO system (O indicates the absence of one sex chromosome). The female has two sex chromosomes just as in the XX—XY system, but the male has only one and is thus designated XO. In the species exhibiting this system, the diploid number of the chromosomes in male is one less than that of the female as a result of the absence of one sex chromosome. Consequently, the number of chromosomes is odd in males and even in females. The grasshopper (Fig. 3.11) is a good example of this mode of sex determination. The cricket and the beetle also exhibit a similar chromosomal basis of sex determination.

3.4.4 Sex Determination by Ploidy Level

In many species of hymenoptera (bees and wasps) and some mites and ticks, sex is determined by the number of sets of chromosomes or ploidy of the individual. Females are diploid, producing haploid gametes via meiosis. Most eggs are fertilised by haploid sperm from males, but a few are not. The fertilised eggs become females that show biparental inheritance. While the unfertilised eggs develop into haploid males (by parthenogenesis) that inherit their genes exclusively from their mother (Fig. 3.12). Haploid individuals, of course, cannot undergo normal meiosis, so males produce gametes via mitosis.

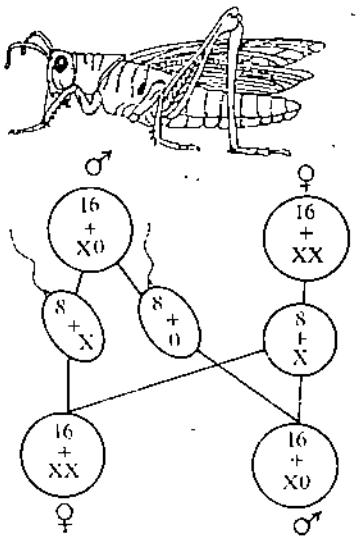


Fig. 3.11: The XX-XO system of sex determination in grasshopper.

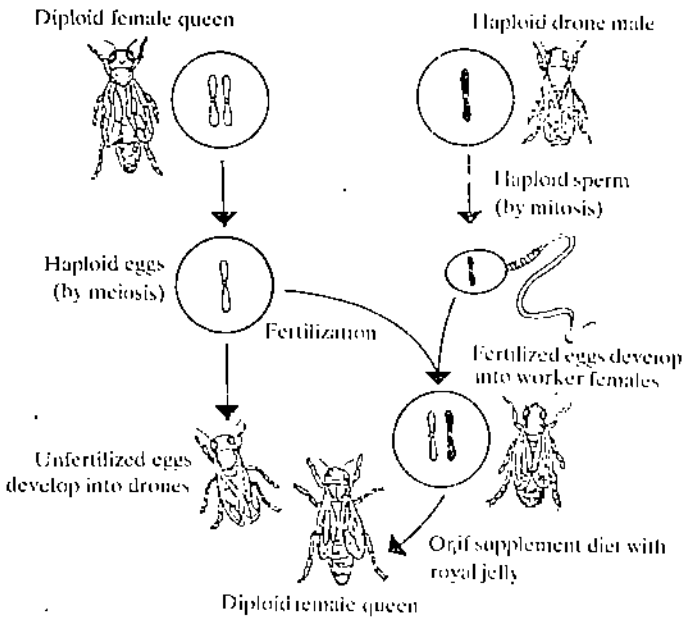


Fig. 3.12: Sex determination by ploidy level in honeybee.

Because males do not undergo meiosis for gamete production, all the sperm from one individual are genetically identical to each other and to the male parent. This has the interesting consequence of increasing the gametic relatedness of a male's

daughters. Remember in this system, a male produces no sons — only daughters. On an average, all daughters of one mated pair of bees share 75% of their genes, rather than the normal 50% relatedness of offspring of most species. The daughters are identical for all the 50% of the genes received from their father, plus one-half of the 50% of the genes from their mother, this makes a total of 75% genetic relatedness.

In some vertebrates also unusual degrees of ploidy are associated with a particular sex. Some species of lizards consist of mostly triploid (3N) females and few males. In fact, males are superfluous, because the females develop *parthenogenetically*. A triploid female produces a triploid egg through mitosis, which undergoes complete development without being fertilised. Haploid males too develop parthenogenetically in the similar manner.

3.4.5 The Compound Chromosome System

Although the X is most commonly found as a single chromosome or single homologous pair or chromosomes, some species have another variation — compound chromosomes. These are named so because a group of chromosomes (e.g. 8X, 12X, 6Y etc.), at the time of meiosis unite end to end and behave as single unit. In such species, there are large differences in the number of chromosomes in males and females. For example, in the nematode *Ascaris incurva* there are 8X chromosomes and 1Y. This species has 26 autosomes. The diploid number of chromosomes in males is 35 (i.e., 26A + 8X + Y), and in females is 42 (or 26A + 16X).

In the above example, X chromosomes exist as a compound chromosome. There are instances where both, the Y chromosomes and the X chromosomes form compound groups. One such example is of *Blaps polychresta*, where the male has 12X's and 6Y's in addition to 18 autosomes.

3.4.6 The Transfer Gene

One additional complicating factor in sex determination in *Drosophila* is worth examining briefly. This shows that the sex chromosomes (X and Y, Z and W) are not the only ones involved in sex determination, but in addition numerous autosomal genes also come into play. In *Drosophila* a recessive allele. *tra,* on the third chromosome (an autosome), when homozygous, "transforms" normal diploid females (AAXX) into sterile males. The *tra* gene has no effect in heterozygous condition. The XX *tra tra* flies have many sex characters of males (external genitalia, sex combs (see Fig. 3.13), and male type abdomen), but are sterile. So are XXY *tra tra* flies. The XY *tra tra* males, however, are normal and fertile.

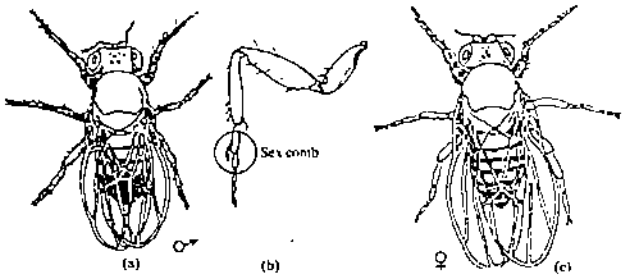


Fig. 3.13: *Drosophila melanogaster*, (a) male, (b) female. Two distinguishing features are: the merging of the posterior bands in the male versus their distinct separation in the female, and the longer more tapering abdomen of the female. (c) shows details of a male leg with sex comb (arrow).

3.5 SEX DETERMINATION : AN EXTENSION

This section pertains to some extensions of the XX-XY system discussed in the previous section. We shall discuss them with the help of two representative organisms: *Melandrium* and *Drosophila*.

3.5.1 *Melandrium*

Sex determination in a following plant named *Lychnis dioica* (formerly the genus *Melandrium*) of pink family (Caryopyllaceae) has been extensively investigated. This plant illustrates a *variation of XX-XY system*. The chromosome composition varies

between male and female plants of this species. This plant usually has four autosomal pairs of chromosomes and a pair of sex chromosomes. The male contains four pairs of autosomes plus an X and a Y chromosome. Female plants contain the same number of autosomes as the male but have two X chromosomes. In this plant abnormal chromosome composition is often seen. A correlation of abnormal chromosome composition in plants with the sex of the individual is depicted in Table 3.1.

Table 3.1: A comparison of chromosome and sex in *Melandrium*. (From: Klug W.S. and Cummings, M.R. 1983)

Chromosome Composition		Sex
Number of Sets of Autosomes (A)	Sex Chromosomes	
2A	XX	Normal Female
2A	XY	Normal male
2A	XXX	Female
3A	XX	Female
4A	XXXX	Female
2A	XYY	Male
2A	XXY	Male
3A	XY	Male
3A	XXXY	Male
4A	XXXY	Male

The Y chromosome has a strong masculinising influence, because when it is present, a male plant is always produced. The X chromosome has a feminising influence, but its influence is masked by the action of Y chromosomes. The XXY and XXXY plants are male (see Table 3.1) despite the presence of two and more X chromosomes.

Cytological studies have shown that the Y chromosomes is larger than the X chromosome, and any of the autosomes. Compare the X and Y chromosome in Fig. 3.14. Only a small portion of the X chromosome is homologous with a similar small bit of the Y chromosome, i.e., region IV. Regions I and IV of the Y chromosome suppress female development, counteracting the influence of region V of the X which promotes female development. If either region I or IV is missing, bisexual development occurs. Region II promotes the male development, and if male tissue develops, region III is essential for male fertility. In the absence of region III, male tissue develops, but the plant is sterile. A region of the X chromosome is also designated IV because it has been identified as the only part of the chromosome which synapses with the Y chromosome during meiosis.

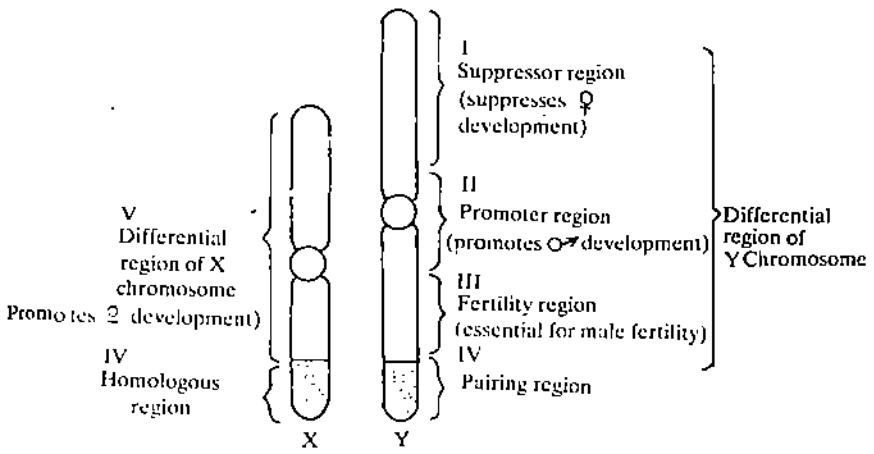


Fig. 3.14: X and Y chromosomes of *Lychnis*. Regions I, II and III bear holandric genes; and region V, X-linked genes. Genes in region IV are termed incompletely sex-linked. Note, the Y chromosome is larger than the X chromosome.

3.5.2 *Drosophila*

A special mention of *Drosophila* is necessary because it has played a key role in the study of inheritance. Let us now have a look at its sex determining mechanisms.

Fig. 3.15: Calvin B. Bridges (1884-1938). American geneticist who pioneered in research on chromosomes of *Drosophila melanogaster.*

The preliminary findings showed that a XX-XY system is operative in the determination of sex in *Drosophila.* It means that male *Drosophila* is heterogametic (XY). One might have assumed that the Y chromosome determines maleness in this species were it not for the investigations of Calvin B. Bridges (Fig. 3.15) in 1916. His investigations showed that the female determiners were located on the X chromosomes and male determiners were on the autosomes. No specific loci were identified at the time, but more recent evidence has demonstrated that many chromosome segments are involved. Specifically, *female-determining genes were shown to be carried on the X chromosomes, and male determining genes were* shown to be located *on the three autosomal chromosomes.*

It is apparent that Y chromosome is not needed at all for the life or even for maleness, as a fly can be male without a Y or can be female with one Y. Nevertheless, Y is essential for fertility, and is thus critical to the survival of the species and is not required for the determination of any sex.

The genic balance theory of sex determination was devised to explain the mechanism of sex determination in *D. melanogaster.* Bridges experimentally produced various combinations of X chromosomes and autosomes in this organism and deduced from comparisons that on X chromosome and two sets of autosomes (A) produced a normal male. The normal males had a ratio of X chromosomes to sets of autosomes of 0.5. Note that a set of autosomes consists of chromosomes II, III and IV (see Fig. 3.16). The combination of one X and two A's resulted in a normal diploid male; the combination of two X chromosomes and two sets of autosomes (2X + 2A, ratio of 2:2=1) produced a normal diploid female (see Table 3.2).

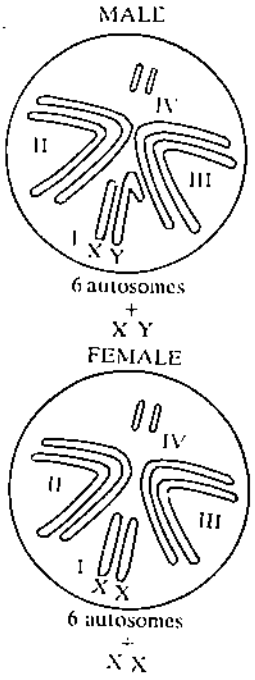


Fig. 3.16: The chromosomal composition of male and female *Drosophila melanogaster.*

To sum up, if the X/A ratio is 0.5 then the sex is male, and when 1.0 it is female. Any ratio between 0.5 and 1.0 results in intersexes. And those above 1.00 result in metafemale or superfemale, whereas those below 0.5 result in metamale (see Table 3.2).

Table 3.2: Ratio of X chromosomes to autosomes and corresponding sex types in *Drosophila melanogaster.* (After Bridges 1925, *American Naturalist, volume 59, pp 127-137.*)

X Chromosomes (X) and Sets of Autosomes (A)	Ratio X/A	Sex
1X 2A	0.5	Male
2X 2A	1.0	Female
3X 2A	1.5	Metafemale
4X 3A	1.33	Metafemale
4X 4A	1.0	Tetraploid female
3X 3A	1.0	Triploid female
3X 4A	0.75	Intersex
2X 3A	0.67	Intersex
2X 4A	0.5	Tetraploid male
1X 3A	0.33	Metamale

The first irregular chromosome arrangement from Bridges experiments resulted from **nondisjunction**—the failure of paired chromosome to disjoin or separate at anaphase. The X chromosomes, which ordinarily come together in pairs during the meiotic prophase of oogenesis and separate to opposite poles during anaphase remained together and migrated to the same pole. As a result some female gametes received two X chromosomes and some received no X chromosome (Fig. 3.17).

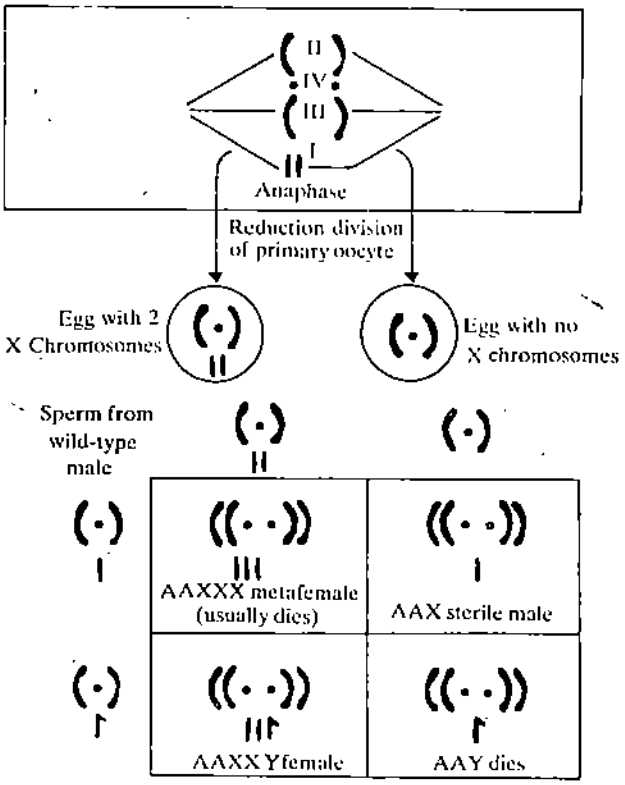


Fig. 3.17: Nondisjunction in *Drosophila*, and the zygotes that result from fertilisation by wild-type males (AAXY). The AAXXY females and the AAX males were exceptional flies in Bridge's experiment. In the primary oocycte, autosomes II and III are represented by pairs of bent rods: autosomes IV, which are small and take their place in the centre of the equatorial plate, are represented by a pair of dots. The X chromosomes (I) are symbolised as short rods; the Y chromosome introduced by the sperm is illustrated by an inverted 'J' shaped symbol.

Following fertilisation by sperm from wild-type males (AAXY), all zygotes had 2n autosomes (2A), but some received two X's from the mother and an X from the father (3X). The ratio of X chromosomes to sets of autosomes of 3:2 resulted in flies called **metafemales** that were inviable. The XXY flies (2X/2A) from the same mating were normal females in appearance. The XO (1X/2A) males are sterile, and those with a Y chromosome but no X did not survive. These results clearly show that, the *Y chromosome is not involved in sex determination.* It is, however, *required for male fertility.*

The flies with 4X/3A were also metafemales. Those with 4X/4A and also those with 3X/3A, both with an X/A ratio = 1 were females. The combinations 3X/4A = 0.75 and 2X/3A = 0.67 were intermediate in characteristics between males and females and were called "**intersexes**". Combination of 2X/4A = 0.5 were males and those of 1X/3A = 0.33 were metamales.

No other species has been as thoroughly investigated as *D. melanogaster.* The knowledge thus obtained has been used to understand the sex determination mechanisms operating in many other organisms.

SAQ 3

The diploid chromosome number of honey bee is 32. How many chromosomes will be found in the somatic cells of a) males, b) females, and in c) the sperm, and d) egg?

.....

.....

SAQ 4

What would be the phenotypic sex of following human beings having 22 autosomes as a haploid set and represented by 'A'?

4) Male Y and female X bearing sperm.

5) a) female (tetraploid)
   b) intersex
   c) intersex
   d) metamale
   e) female (diploid)
   f) male (sterile if no Y chromosome is present in *Drosophila*)

6) Female gametes would be (2X2A), (2XA), (X2A), and (XA). Zygotes and sex would be 3X3A female (triploid), 3X2A metafemale, 2X3A intersex, 2X2A female (diploid), 2XY3A intersex, 2XY2A female, XY3A metamale, XY2A male.

7) a) Male
   b) female
   c) bisexual
   d) female

8) a) male
   b) female
   c) male
   d) female

UNIT 4 SEX LINKAGE AND DOSAGE COMPENSATION

Structure

4.1 INTRODUCTION

One of the probable deductions from your study of Mendel's laws of inheritance and their extensions and modifications in Units 1 and 2 may be that contribution to inheritance is equal from both the parents. But sex linkage is a major exception to it. Sex linkage occurs when a gene controlling a trait is located on the sex chromosome. The sex chromosome bears several genes in addition to those directly concerned with sex determination. The inheritance of these genes follows a characteristic pattern which is different from that seen in the examples of monohybrid and dihybrid inheritance, that you have studied so far. The unique sex-linked pattern(s) of inheritance of any particular trait can be easily recognised and studied by pedigree analysis.

Sex linkage forms the main theme of this unit. But we begin by a brief discussion on mendelian factors that we now know are the genes and are located on the chromosomes. Then you will study examples of genes located on the sex chromosomes and their mode of transmission to the next generation. These would be explained with examples of X-, and Y-linked genes.

In Unit 3, you have seen that the sex chromosomes exhibit dimorphism. Due to the difference in chromosome complement there is inequality in dosage of genes present on the X chromosome in males and females. The Y chromosome bears genes mostly related to male differentiation. The X chromosome carries genes necessary for mediating some of the basic functions. Females with two X chromosome would then have the X linked genes in double the dose, whereas males would have the same genes in a single dose. A dosage compensation mechanism operates to equalise the X linked gene activity in both the sexes. You would study the operational details and importance of this mechanism in two systems – man and *Drosophila*, in the last section of this unit.

Objectives

After studying this unit you would be able to:

- relate the chromosome theory of inheritance to sex linkage (Section 4.2);
- distinguish the mode of inheritance between the X-linked genes and Y-linked genes (Section 4.3);

- explain with the help of examples the transmission of recessive and dominant X-linked traits (Section 4.3, Subsection 4.3.1);
- explain with examples the transmission of Y-linked genes (Section 4.3, Subsection 4.3.2);
- discriminate between sex-limited and sex-influenced genes, and enumerate their role in the control of secondary sexual characters (Section 4.4);
- explain the importance of dosage compensation mechanism (Section 4.5);
- differentiate between the type of dosage compensation in mammals and *Drosophila* (Section 4.5);
- describe the existence of female mosaics with respect to X-linked traits (Section 4.5).

4.2 THE CHROMOSOME THEORY OF INHERITANCE AND SEX-LINKAGE

Soon after the discovery of Mendel's work in 1900, most geneticists accepted the particulate nature of genes. Mendel had predicted that each gamete contains only one allele of each gene instead of two. This prediction was based on the fact that there is reduction in the number of chromosomes by one-half at the time of meiosis during gamete formation.

Fig. 4.1: Thomas Hunt Morgan b. 25 Sept. 1866, Kentucky, d. 4 Dec. 1945, California.

This notion, that chromosomes carry genes is the **Chromosome Theory of Inheritance.** The credit for the Chromosome Theory of Inheritance goes to Walter Sutton and Theodor Boveri. In 1902, these investigators independently recognised that the behaviour of Mendel's genes during production of gametes in peas precisely paralleled the behaviour of chromosomes at meiosis. The following parallels were drawn between the two: i) genes are in pairs, so are the chromosomes, ii) the members of gene pair segregate equally into gametes, so do the members of a pair of homologous chromosome, and iii) the different gene pairs act independently, so do the different chromosome pairs.

The proposition of the Chromosome Theory was a crucial new step in genetic thinking at that time. No longer were genes just disembodied factors, now they were a part of the observable entities in the cell nucleus. Some geneticists, particularly, Thomas Hunt Morgan (Fig. 4.1) remained skeptical of this idea. Ironically it was Morgan himself who in 1910 provided the first definitive evidence for the Chromosome Theory based on his studies on sex linkage.

Morgan worked with fruit fly *Drosophila melanogaster.* When he mated red-eyed flies R ‖ R (dominant) with white-eyed flies r ‖ r (recessive), the F_1 progeny were red-eyed. Furthermore, when Morgan mated red-eyed males of the F_1 generation with their red-eyed sisters, they produced about 1/4 white-eyed males, but no white-eyed females. In other words, the eye colour phenotype is X-linked. X-chromosome and eye colour are transmitted together because the genes governing this character are located on the X-chromosome. In a diploid individual, we know, the autosomes occur in pairs but as regards to X-chromosomes the female fly has two copies and the male has just one. However, Morgan was reluctant to draw this conclusion until he observed sex-linkage with two more characters – miniature wings and yellow body in the fruit fly. That was enough to convince him and other geneticists of the validity of the chromosome theory of inheritance.

4.3 SEX-LINKED INHERITANCE

The inheritance of genes located on the sex chromosomes follows a characteristic pattern which is different from those located on the autosomes. You are already familiar with this concept from your study of Section 2.7 of Unit 2, where we have discussed Morgan's discovery of sex linkage in *Drosophila.* You may revise that portion again before you begin your study of this section. Here, we are elaborating

on the XX-XY system as it pertains to man and most of the examples cited are also of human context.

4.3.1 X-Linked Traits in Humans

Let us first examine traits determined by genes on the X-chromosome, or the *X-linked traits.* The human X chromosome contains many genes that are required in both the sexes, whereas the Y chromosome contains only a few genes, principally the genes for maleness. More than two hundred traits have been found to be X-linked and only a few are known to be Y-linked. The traits controlled by genes located on the X-chromosome are also referred to as **sex-linked.** That is, the terms X-linked and sex-linked are used synonymously. It is more appropriate, however, to refer to these as X-linked traits, since they follow the pattern of transmission of the X-chromosome.

The X-linked traits have a unique mode of inheritance because females have two doses of X-linked genes, while males have only one. Males are thus hemizygous for X-linked traits.

X-Linked Dominant Genes

Dominant X-linked genes are always expressed in both the sexes just as in autosomal traits. One dose of X-linked dominant allele produces its effects in males as well as females. The hemizygous male transmits the gene to all its daughters but none to his sons. There is no father to son transmission. The heterozygous females transmit the trait to half their children, irrespective of their sex. On the other hand, females homozygous for the dominant allele produce all affected children. For example, a form of vitamin D-resistant rickets is inherited as an X-linked dominant trait (Fig. 4.2).

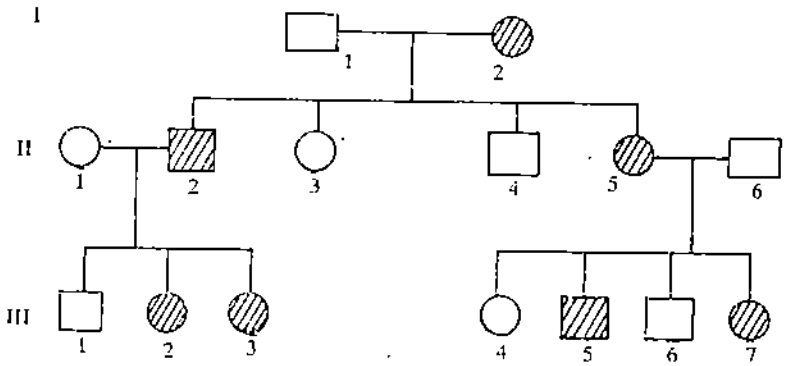


Fig. 4.2: Pedigree demonstrating X-linked dominant trait. The **affected progeny are shaded.**

X-Linked Recessive Genes

The opposite is true for recessive alleles. Males being hemizygous, always express the **recessive X-linked alleles.** Females, however, express recessive alleles only when they are homozygous. Thus, the **frequency of X-linked recessive traits is always lower in females than males.**

Most X-linked genes are recessive alleles, and they are discovered when their deleterious effects appear in males. Males transmit their X chromosome to every daughter, and their Y chromosome to every son. *Recessive X-linked traits thus show a pattern of inheritance, in which the phenotype is usually expressed only in males of alternate generations.* A male bearer transmits the recessive allele to daughters, who does not express the allele because it occurs in the heterozygous condition. However, each of her male offspring has a 50% chance of receiving that allele and expressing the phenotype. The trait should thus appear in 1/4 of her offspring (1/2 of her offspring are expected to be male and 1/2 of her sons receive the recessive allele: $1/2 \times 1/2 = 1/4$). The heterozygous female is a carrier of the allele. The X-linked allele is often said to show a criss cross pattern of inheritance. In this pattern of inheritance, the allele is transmitted from male to female, female to male, and the trait is expressed only in males in alternate generations (see Fig. 4.3). Well-known examples of X-linked recessive alleles include Red green colour blindness, Haemophilia, Glucose-6-phosphate dehydrogenase deficiency (G-6PD), Congenital hyperuricemia, Duchenne muscular dystrophy, and Ichthyosis.

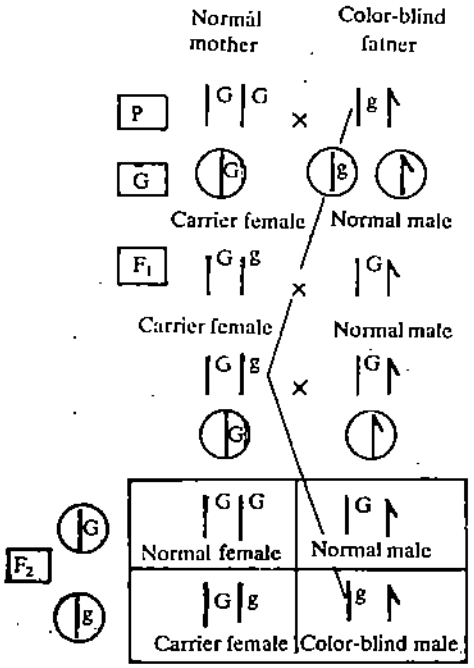


Fig. 4.3: Criss cross inheritance, i.e., the inheritance of a character from father to daughter to grandson. It is characteristic of a sex-linked gene. Genes are shown on the chromosomes illustrating a cross between a woman with normal vision and a green-colour defective man. The symbol *g* represents the sex-linked recessive gene for green colour defective vision, and *G* the normal condition.

SAQ 1

A husband and wife are normal although both their fathers have a trait which is X-linked recessive. What is the probability that their first child will be:

a) A normal son?
b) a normal daughter?
c) a son with the trait?
d) a daughter with the trait?

.....

.....

.....

.....

.....

.....

.....

.....

Red Green Colour Blindness: Many persons cannot perceive certain colours. The most common such defect is an inability to distinguish red from green. This condition is also called **partial colour blindness.** Colour perception is controlled by the cone-shaped cells in the retina of the eye. Three types of cone cells, each containing a specific light absorbing pigment (whose nature is protein), that perceives a specific portion of the visible spectrum (see Fig. 4.4) have been identified. These three types of cone cells are referred to as red-absorbing, green-absorbing and blue-absorbing cone cells.

By 1986, the genes that encode the above three light-absorbing pigments of the retina were isolated and their nucleotide sequence was determined. The sequences have been used to find out the amino acid sequences of the three light absorbing proteins. These light-absorbing proteins were found to have very similar structures. See Fig. 4.5, the red and green-receptor proteins differ at only a few of the amino

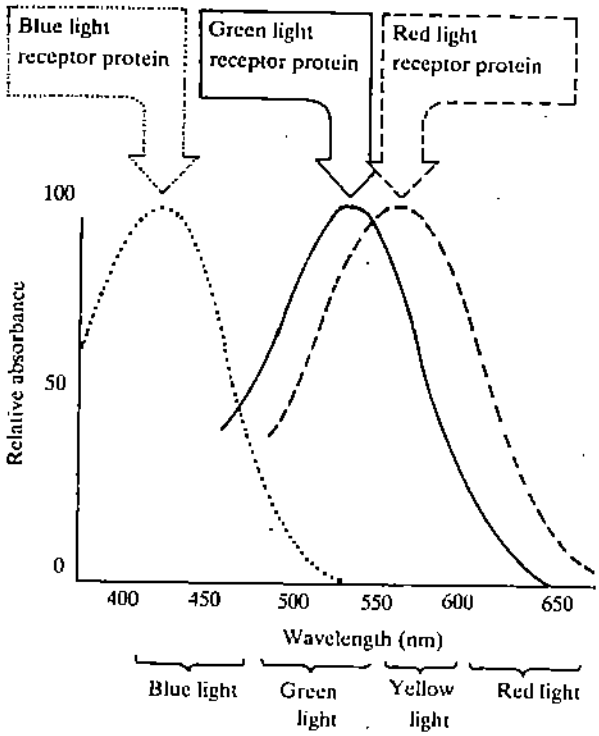


Fig. 4.4: Absorption spectra of the blue, green and red proteins present in the cone cells of the retina of the human eye. The ability of humans with normal colour vision to distinguish colours throughout the visible spectrum depends on the presence of all three proteins. Defective colour vision results from the absence of, or a defect in, one or more of these proteins. [After *Nathans, J. 1989. The Genes for colour vision, Sci. Amer. 260(2): 42-49].*

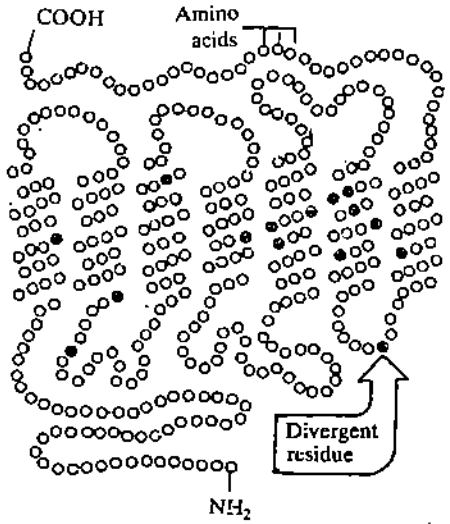


Fig. 4.5: The diagrammatic representation of the structure of red and green receptor proteins of humans. Each circle represents one subunit (amino acid) of each of the protein. Note the difference in the two proteins. [After *Nathans, J. 1989. The Genes for Colour Vision, Sci. Amer. 260(2) : 42-49].*

acid subunits. **The genes that encode the green-, and red-receptor proteins are located on the X-chromosomes,** thus the sex-linked patterns of inheritance are observed for defects in green and red colour vision. And the gene encoding the blue receptor protein was found to be located on chromosome-7, that is, an autosome. Let us now examine some crosses (see Fig. 4.6) and assess the pattern of inheritance of this defect. For the sake of simplicity we are only considering here one gene, that is, gene causing defects in the green receptor protein, as a single sex-linked recessive allele. Since the Y chromosomes carried no colour vision locus, the single allele is expressed causing colour blindness. Stop here for a minute and carefully study the five crosses in the figure. The symbol *G* denotes normal receptor pigment, and *g* is for the defective receptor pigment causing colour blindness.

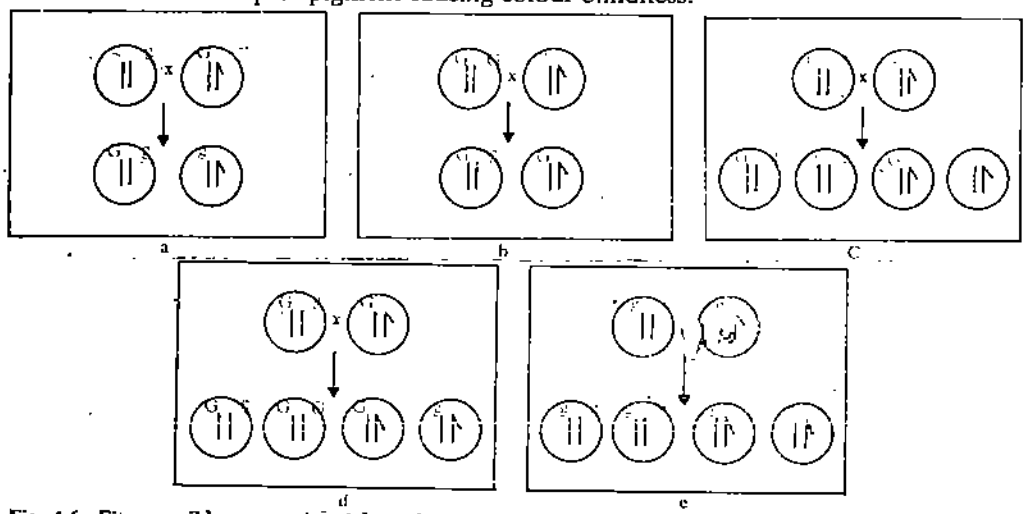


Fig. 4.6: Five possible crosses involving inheritance of sex-linked recessive trait — partial colour blindness: a) normal male × colour blind female; b) colour blind male × normal female; c) colour blind male × normal female who is a carrier; d) normal male × normal female who is a carrier; e) colour blind female × colour blind male.

In the above crosses have you noticed that sex-linked inheritance does not conform to the Mendel's laws of inheritance? Fig. 4.6 shows the results of reciprocal matings of affected and unaffected parents. The normal male and colour blind (recessive, homozygous) female produce normal but heterozygous daughters, but all sons have the disease (cross a). The reciprocal cross (b) demonstrates criss cross inheritance. A colour blind male (hemizygous) with a normal (homozygous) female produces no affected offspring, but the daughters are carriers (cross b). A colour blind male and a carrier female result in 50% colour blinds (see cross c). Another possibility, a normal male mated with a carrier female, produces all normal female offspring but 50% affected male offspring (see cross d). The mating of two colour blind individuals result in all colour blind offspring (see cross e) if they have the same colour blindness.

Haemophilia: Haemophilia is a disease in which one of the factors required for the normal clotting of blood (see Fig. 4.7) is deficient. As a result, the blood fails to clot or clots very slowly. Thus even minor injuries can cause profuse internal and external bleeding which can lead to death.

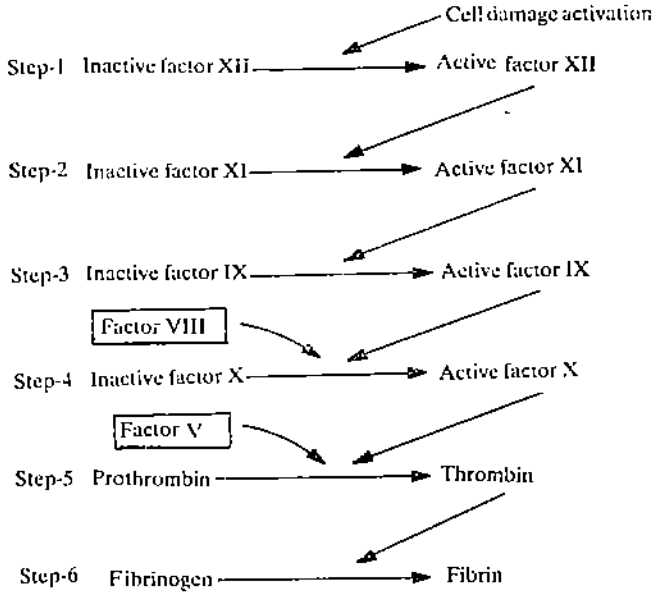


Fig. 4.7: Blood clotting is the end result of a series of reactions requiring various enzymes and cofactors, leading to the formation of insoluble fibrin. Many of the steps in the chain entail the conversion of a gene product to its active enzymatic form by the enzymatic action of the active product of a previous step. Cofactors (e.g., factor VIII and V) are required to work along with some of the active enzymes. Factor VIII, absent/defective in haemophilia A, is a cofactor required along with factor IX to activate factor X. If factor VII is absent as a result of a sex-linked recessive condition, the sequence is interrupted and the end result is defective clotting. In haemophilia B, factor IX is deficient.

Three forms of haemophilia controlled by three different gene loci are known. One of them is rare and is controlled by an autosomal recessive gene, while two forms, that is, haemophilia A and B result from recessive alleles at two X-linked loci. Haemophilia B also known as *Christmas disease,* comprises about 20% of all haemophilia and is caused by the deficiency of factor IX (see Fig. 4.7, step 3).

Haemophilia A, classical haemophilia is caused by an abnormality or deficiency of a protein cofactor known as factor VIII is located on X-chromosome. Factor VIII is needed for the activity of one of the enzymes — factor IX, in the series of events leading to the activation of thrombin. Absence of functional factor VIII interrupts the steps leading to the activation of thrombin, and consequently fibrin cannot form. Until recently, haemophilia A was untreatable and only about 25% of the affected males reached age of twenty five. Treatment with factor VIII now results in a longer life span.

The frequency of haemophiliacs is about one in ten thousand males, but is much lower in females, about one in one hundred million or less. A female haemophiliac can result from the mating of a heterozygous female with an affected male. Such a mating, is highly unlikely because very few male haemophiliacs survive long enough to reproduce. Haemophiliac females are also belived to die at the onset of menstruation.

Haemophilia is one of the earliest known diseases. According to Talmud, the Hebrew book of law, when excessive bleeding occurred during circumcision of two male infants of a mother, future male offspring were exempt. When sons of three sisters exhibited bleeding, sons of other sisters were also exempt. However, sons of brothers were not exempt, implying an understanding of criss cross pattern of inheritance.

Haemophilia A has been called the "Royal disease" because it affected males in the royal families of Europe. Queen Victoria, a carrier of the haemophilia allele had nine children (Fig. 4.8). Her eighth child, Leopold was a haemophiliac who died at the age of thirty one. Her other three sons were unaffected as they did not receive the allele. One daughter had no children, her status as a carrier cannot be assessed. Two daughters had children, none of whom were haemophiliac, indicating that mothers probably were not carriers. Two other daughters were carriers giving birth to haemophiliac sons.

The possible historical influence of haemophilia is tentalising. Victoria's third child was princess Alice, whose daughter Alexandra married Czar Nicholas II of Russia. The Czarina, Alexandra, had four daughters before giving birth to the long awaited son Alexis – the heir to the Russian throne. Unfortunately, Alexis had the haemophilia allele, a legacy from his great grand mother – Queen Victoria. Distressed over their son's condition, the Czar and Czarina turned to the monk Rasputin. While affairs of the state deteriorated, culminating in Russian revolution, the Czar was preoccupied with the health of his son.

Among Victoria's descendants, eight of twenty five males in four generations were haemophiliacs. Queen Victoria almost certainly received the gene for haemophilia A, as a result of mutation on the X chromosome which she received from her father Edward, Duke of Kent. He was fifty-two years old at the time of Victoria's birth and such mutations may occur more frequently in the germ cells of older males.

In the recent years, a serious threat to victims of haemophilia has arisen due to their continuing dependence upon blood transfusions. Such transfusions are one means of transferring **Acquired Immune Deficiency Syndrome (AIDS)** and some haemophiliacs have infact acquired AIDS in this way. Extensive surveillance of donor blood supplies is required to protect haemophiliacs and all others requiring transfusions.

Glucose 6-Phosphate Dehydrogenase (G-6PD) Deficiency: Another disease due to defective X-linked recessive allele is G-6PD. This is an important enzyme, for carbohydrate metabolism and maintaining stability of red blood cells.

Deficiency of enzyme G-6PD is a rare condition characterised by severe haemolytic anaemia (due to destruction of red blood cells) when exposed to environmental triggers such as fava beans, naphthalene and certain sulpha drugs.

Congential Hyperuricemia – Lesch-Nyhan Syndrome: This disease is characterised by the excess production of uric acid. A mother contributes the X-chromosome with defective gene to a male zygote. Half of the male children of carrier mothers may be expected to inherit the disease. They are deficient for the enzyme

hypoxanthine-guanine phosphoribosyl transferase (**HGPRT**). This enzyme is involved in nucleotide synthesis. Infants who receive the gene appear normal at birth. The initial symptom of the disease is the production of excessive uric acid in the urine as a result it appears as orange sand-coloured. By 10 months of age, the patients become abnormally irritable and lose motor control. Weak and flabby muscles prevent the child from sitting, walking and speaking normally. As the disease advances, there is deterioration of nervous system. Self-mutilation occurs, manifested by lip-biting, finger-chewing, teeth-grinding, and marked swinging of the arms. Eventually death occurs within a few years due to severe renal and neurological damage. Some of the patients live to their twenties.

Duchenne Muscular Dystrophy (DMD): It is another example of an X-linked recessive allele, that primarily affects young males. Half of the male progeny of a carrier female are expected to be affected. In the affected males deterioration begins between the ages of three and five years, but sometimes the affected individuals reach their teens. But they are confined to wheel chairs; and they die in their early twenties due to atrophy of their respiratory muscles. Only few affected males reproduce, so the condition is transmitted mainly by female carriers. This disease occurs in about one in every four hundred newborn males; and is several times more frequent then haemophilia. In 1986, the defective gene that causes DMD was isolated and studied. It was found that the absence or a protein - *dystrophin* caused DMD.

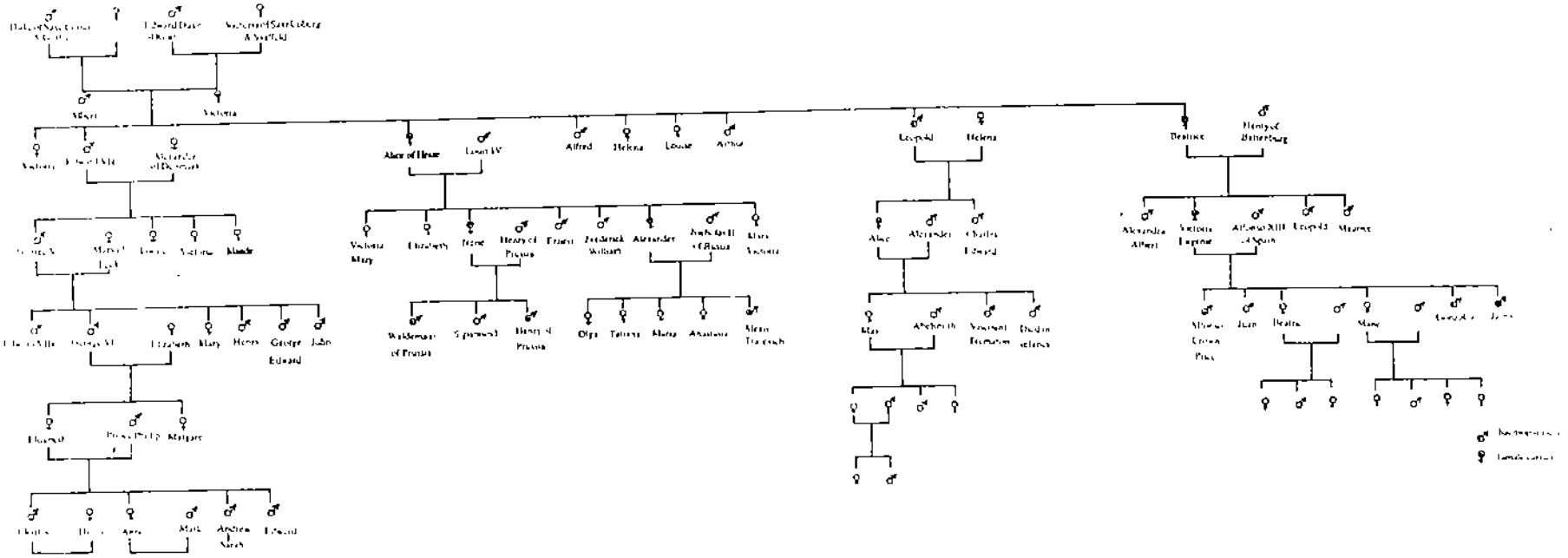


Fig. 4.8: Pedigree of Queen Victoria showing carriers and haemophiliacs. The present royal family of Great Britain is free of the haemophiliac allele.

Ichthyosis: So far you have seen the examples where the recessive X-linked are expressed in males. There are, however, instances where these are also expressed in females in certain situations. **Consanguineous matings** (see adjacent Margin Remark) can greatly increase the frequency of expression of X-linked traits in females. In consanguineous pedigrees containing X-linked recessive alleles, females have a high probability of carrying the X-linked allele, as they can receive the allele from either parent. In turn, matings of carrier females and affected males (Fig. 4.9) produce daughters and sons with an equal likelihood (i.e., 50%) of being affected.

Consanguineous means "of the same blood". The term means sharing genes derived from a common ancestor, related by descent.

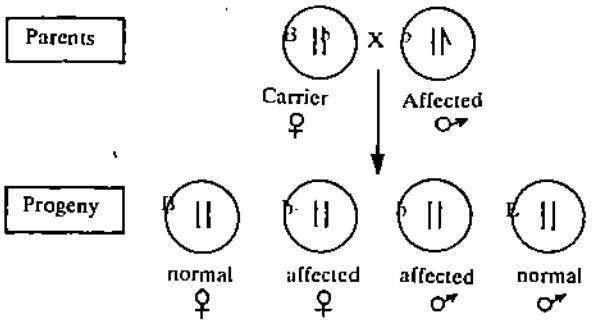


Fig. 4.9: Cross between ichthyosis carrier female and affected male showing 50% affected progeny.

Ichthyosis is a disorder characterised by extreme dryness, roughness and scaliness of the skin. The prefix 'ichthy' means fish-like. Children produced in situation as shown in Fig. 4.9 show ichthyosis at birth. Similarly, Fig. 4.10 shows a pedigrees of a family in which consanguineous mating resulted in expression of this X-linked trait in females. The allele for ichthyosis first appeared in male I-1. His first four daughters all transmitted the allele, without expressing the condition themselves. Male III-11 mated with his first cousin, once removed IV-3, who must have been a carrier, having received the ichthyosis allele through two generations of females. One of their three sons and two of their three daughters exhibited ichthyosis, a highly unlikely result without consanguineous mating.

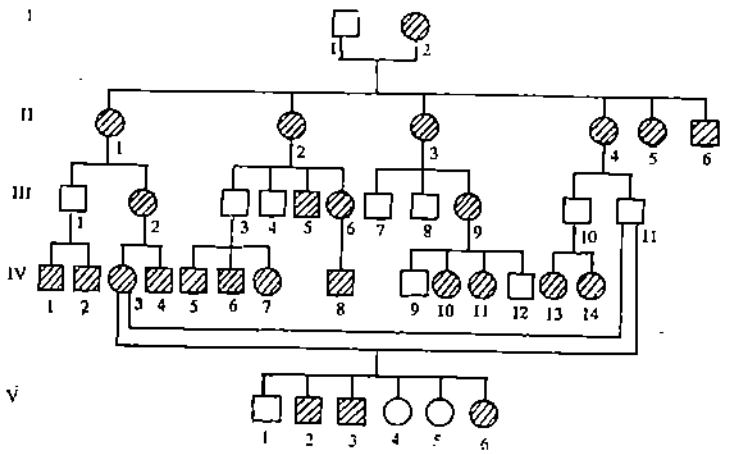


Fig. 4.10: A pedigree showing the occurrence of a X-linked recessive trait, ichthyosis in females as a result of consanguineous mating.

Before we go on to the next subsection, how about trying a couple of SAQs?

SAQ 2

A couple have a colour blind daughter and son with normal vision. What are the genotypes of the parent in this cross?

.....

.....

.....

.....

SAQ 3

Draw a three generation pedigree of a family starting from a couple, where male is a haemophiliac and the female is normal. They have 3 sons and a carrier daughter. The daughter marries a normal male and has 3 daughters and 2 sons. What is the probability of her children being carriers and haemophiliacs?

Fig. 4.11: Hairy ear in humans, a trait which is determined by a Y-linked gene.

4.3.2 Y-Linked Traits in Humans

Any gene that occurs exclusively on the Y chromosome is said to be **holandric** and it is not expressed in females. Such a Y-linked gene normally occurs in males and is transmitted only from father to son – **holandric inheritance.** Only a few Y-linked genes have been identified uptil now. One is the histocompatibility gene, known as the H-Y gene which is present on the short arm of the Y chromosome. Another important Y-linked gene is the TDF gene that codes for testis determining factor. This locus plays an important role in primary sex determination. The functional significance of TDF gene would be explained in Unit 5.

Another phenotype known to be associated with the Y chromosome is the condition **hypertrichosis.** The gene concerned with this condition leads to the development of hairy pinna (Fig. 4.11). This phenotype has been observed in the inhabitants of Australia, Ceylon, Israel, and India.

SAQ 4

A man has hypertrichosis of the ears, a condition due to a gene on the Y-chromosome. Show the types of male and female children he has.

.....

.....

.....

.....

.....

SAQ 5

Is a Y-chromosome linked gene supposed to be dominant or recessive in order to be recognised? How is a Y-linked gene transmitted to grand children?

.....

.....

.....

.....

.....

.....

.....

.....

4.4 SEX-LIMITED AND SEX-INFLUENCED TRAITS

Not all of the characters that differ in the two sexes are X-linked. There are certain traits that are determined by autosomal genes, but their expression is altered or influenced by the sex of the individual. Many a times these traits are confused with the sex-linked traits. Actually, they are entirely different in their mode of inheritance since their genetic determinants are not located on the sex chromosomes. There are two kinds of such traits: sex-limited, and sex-influenced traits.

4.4.1 Sex-Limited Traits

Sex-Limited Traits are traits expressed only in one sex, although the genes controlling them are present as well as transmitted to both the sexes. Therefore, males and females with the same genotype, with respect to a particular locus may have different phenotypes.

Sex-Limited Traits are determined by autosomal genes, whose phenotypic expression is determined by the presence or absence of one of the sex hormones. Since sex hormones are the limiting factors, the phenotypic expression of these genes is limited to one sex or the other. The most obvious examples are the secondary sex characteristics. Beard development in human beings is one such sex-limited character as men have beards, and woman normally do not. Yet studies indicate no significant differences between the sexes in number of hairs per unit area of skin surface except in their development. This appears to depend on sex hormone production. Any disturbance in these hormones in women may result in the development of beard. Similarly, the full development of breasts in females, and presence of prostate glands in males are the examples of sex-limited traits seen in human beings. Traits like egg laying in chickens, oviposition behaviour in insects are some other such examples. Milk production in mammals is limited to females, but certain bulls are in great demand among dairy breeders and artificial insemination associations because their mothers and daughters have increased milk production records.

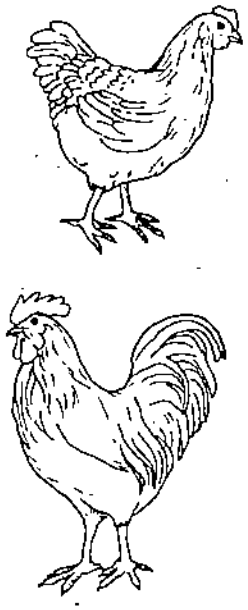


Fig. 4.12 : Hen-feathering (top), and cock-feathering (bottom) in domestic fowl.

Another classic example of a sex-limited trait is "cock feathering" in different birds. We consider here the example of domestic fowl, the males and females exhibit pronounced difference in their plumage. In the *leghorn* breed the males have long, pointed, curved, fringed feathers on tail and neck, but feathers on females are shorter, rounded, straighter, and without fringe (see Fig. 4.12). Thus males are cock-feathered and females are hen-feathered. In the breeds *Sebright bantam,* birds of both sexes are hen-feathered. However, in *Hamburg* and *Wyandotte,* both hen-, and cock-feathered males are seen, but all the females are hen-feathered. The feathering type depends on a single pair of alleles *H* and *h* in the following manner (Table 4.1).

Table 4.1: The Feathering Type in Domestic Fowl.

Genotype	Male	Female
HH	hen-feathered	hen-feathered
Hh	hen-feathered	hen-feathered
hh	hen-feathered	cock-feathered

Thus *Sebright bantams* are all HH. *Hamburgs* and *Wyandottes* may be *H–*, or *hh,* and *Leghorns* are all *hh.* Cock-feathering where it occurs is limited to the male sex.

4.4.2 Sex-Influenced Traits

Sex-Influenced or **Sex-Controlled Traits** appear in both sexes but occur in one sex more than the other.

The genes for **Sex-Influenced** Traits show differing patterns of expression in each sex-usually the trait behaves as dominant in one sex and a recessive in the other. Genes for **sex-influenced traits occur only on autosomes.**

Pattern baldness refers to a definite genetic pattern. In this condition hair usually thins on top ultimately leaving a fringe of hair low on the head (Fig. 4.13). Baldness may also arise due to various causes such as disease, radiation, thyroid defect.

The best documented example of sex-influenced inheritance is **pattern baldness** (Figs. 4.13 and 4.14). Individuals expressing pattern baldness begin to lose their hair on the front and the top of the head, relatively early in life, often in their twenties. Affected individuals are not totally bald: a distinct rim of hair surrounds their head in patterns varying from person to person. It has been proposed that a single pair of alleles is involved. The allele B_1 which is responsible for pattern baldness is dominant in males, and the heterozygous males therefore, express pattern baldness. In females, however, the gene is recessive. The allele for normal hair growth can be designated as B_2. Individuals with the genotype B_1 B_1 show pattern baldness, regardless of sex. In such situations, in females there is marked thinning, rather than total loss of hair on the top of the head. Persons with B_1B_2 genotype are bald if they are male but not bald if they are female. *The presence of male hormones, are strongly implicated in the expression of pattern baldness.*

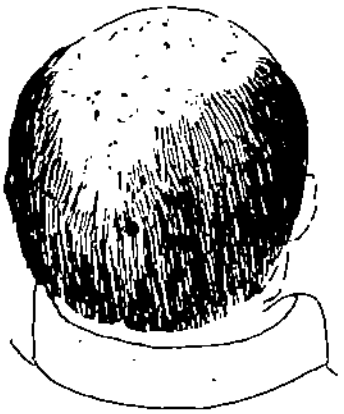


Fig. 4.13: Pattern-baldness in man.

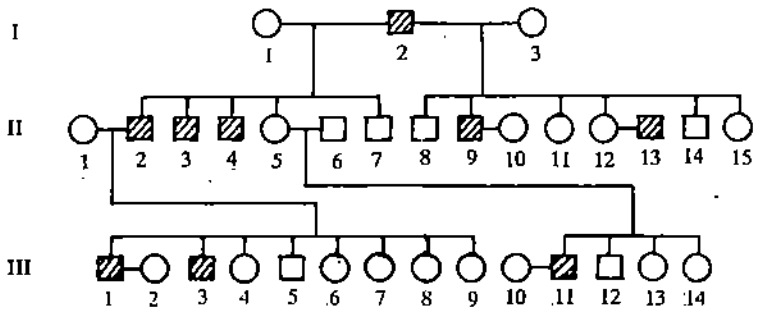


Fig. 4.14: Pedigree showing the incidence of pattern baldness in a family. The men represented by the dark squares became bald before they reached the age of 35. Those represented by light squares are over 35 and have thick hair. No woman in this family pedigree expressed the trait (After *Gardner et al. 1991, Principle of Genetics, John Wiley & Sons, Inc.*)

Some human traits, such as certain types of white forelock, absence of upper lateral incisor teeth, a particular type of enlargement of the terminal joints of the fingers, and cleft-lip, exhibit a pattern of inheritance characteristic of sex-influenced genes.

A few well-known examples of sex-influenced genes in animals are: spotting in cattle (mahogany and white dominant in males, red and white dominant in females), horned versus hornless condition in sheep (Fig. 4.15) where the autosomal gene involved is dominant in males and recessive in females.

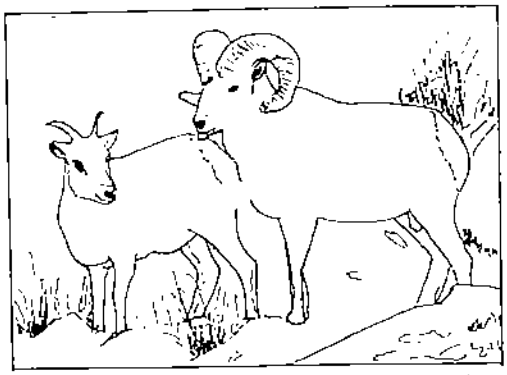


Fig. 4.15: Rocky mountain sheep showing sexual dimorphism in horn development. The male has large horns, whereas the female is devoid of them.

SAQ 6

Compare the inheritance pattern of the sex-limited traits with those of the sex-linked traits.

.....

.....

.....

.....

.....

4.5 DOSAGE COMPENSATION

Recall that in the XX-XY chromosome system, males have only one X chromosome (hemizygous) while the females have two. Thus the males have half the number of X-linked genes as females. In other words, the males have only one dose of X chromosomes and the females have double dose of X-chromosomes. We know that the amount of gene product in cells is related to the number of gene copies present, it would be expected that females would have double the amount of X-linked gene products as compared to males. Now the question arises, is there any compensation for this dosage difference between sexes? The answer is 'Yes', there is a mechanism which regulates the levels of gene products in such a way that both hemizygous and heterozygous/homozygous individuals that is, males and females, have the same amount of gene product. This mechanism is known as Dosage Compensation.

4.5.1 In Man

In human and other mammals, the necessary dosage compensation is accompanied by **inactivation** or **"turning off"** of one **of the X chromosomes** in females so that both males and females have only one functional X chromosome per cell. The inactive X chromosome, in females becomes tightly coiled into **'heterochromatin'**, a condensed form of chromatin visible as a dark spot – **X-chromatin** or **Barr body** (after its discoverer M.L. Barr) in the nucleus of female cell (Fig. 4.16.). Thus Barr body is the inactivated X chromosome. One X chromosome is necessary for normal development in both sexes, but if an individual (or either sex) has more than one X-chromosomes, all but one are inactivated and are visible in stained somatic cells as Barr bodies. Thus somatic cell nuclei of normal males have no Barr body, and those of normal females have one (also see Box 4.1).

Fig. 4.16: Barr Body in the nucleus of a cell of a normal female.

Box 4.1: Detection of Barr Body
A simple way to demonstrate the Barr Body in humans is by scraping epithelial cells from the buccal mucosa of females, and staining them with a specific dye. The nuclei in many cells would show a small, diamond-shaped structure about 1 μm in diameter, more deeply stained than the surrounding chromatin and usually located at the periphery of the nucleus. This body stains positively in the **Feulgen reaction** for DNA.

The hypothesis that *all but one X chromosome(s) are inactivated* in each cell was proposed by the geneticist Mary F. Lyon (Fig. 4.17) in 1961 and is known as **Lyon hypothesis.** Crucial evidence for this hypothesis was provided by sexually aneuploid individuals. **Aneuploidy** (meaning not the true number) refers to the possession of an abnormal number of chromosomes. Aneuploid individuals have the normal diploid number, plus or minus one or more chromosomes. Females lacking one X chromosome exhibit **Turner's syndrome,** designated 45, XO (45 chromosomes, with one X missing). Males with an extra X-chromosome have **Klinefelter's syndrome** designated 47 XXY. Cells from 45, XO females and 46 XY males have no Barr bodies (Fig. 4.18a); while those from 47, XXX females and 48 XXXY males have two (Fig. 4.8b); 46, XX females and 47, XXY males have one Barr body (Fig. 4.18c); and 48, XXXX females and 49, XXXXY males have three Barr bodies (Fig. 4.18d). Examination of the number of Barr bodies can be done easily to screen for sex-chromosomes abnormalities. You have already seen in Box 4.1 how to make a

Fig. 4.17.: Mary F. Lyon, (1925)

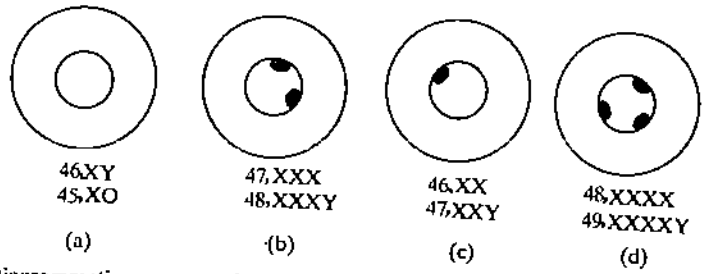


Fig. 4.18: Diagrammatic representation of varying number of Barr bodies in different genetic make ups.

preparation of the epithelial cells from the buccal cavity for examination. The mandatory "sex tests" that have been required for Olympic athletes includes a count of Barr bodies. Males disguised as females can be identified as they have no Barr bodies.

The inactivation of X chromosomes during development occurs at random. Early in development, the maternally derived X is inactivated or **lyonised** in some, while the paternally-derived X is inactivated in others. Thereafter, descendants of a particular cell have the same X inactivated (Fig. 4.19). If a female is heterozygous for an X-linked gene, she is **mosaic** for that trait. One of her X-chromosomes is active in roughly half of her cells while the second X is active in other cells. That is to say that some cells express X-linked genes inherited from the father while others express those passed on by the mother.

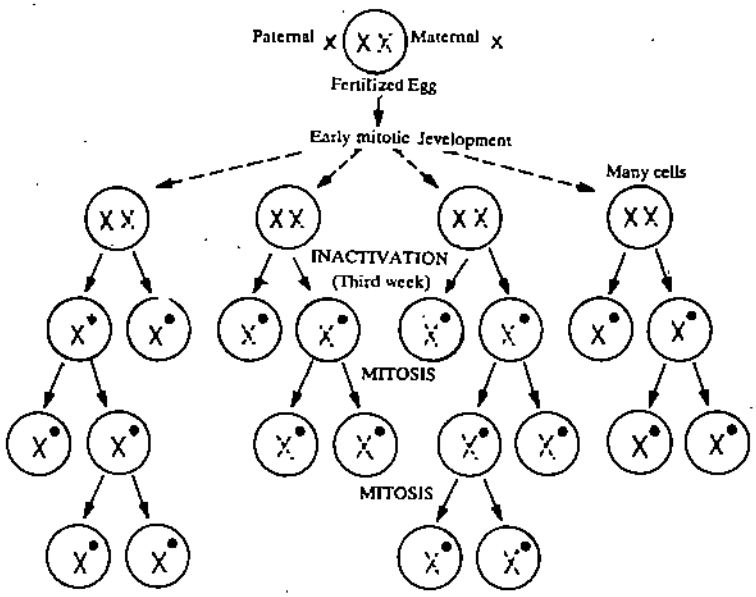


Fig. 4.19: Diagrammatic representation of the random inactivation of one of the two chromosomes in female cells. all progeny cells inactivate the same chromosome. All descendants of these cells have the same chromosome inactivated, so females are mosaics for their maternally-derived and paternally-derived X chromosomes.

Fig. 4.20: A calico cat with patches of colour resulting from random inactivation of X chromosomes bearing colour determining genes in cells giving rise to hair.

Calico cats (Fig. 4.20) exhibit mosaicism, due to dosage compensation. Several loci control coat colour in cats, but only one X-linked locus is involved in producing calico individuals. Two alleles occur at that locus, R and R'. In males (hemizygous), R produces rust coat colour and R' black. In females R inactivation produces clones of R-bearing rust fur intermixed with R'-bearing black fur – the calico cat. Thus almost all calico cats are females. Male calico cats, only result because of sex-chromosome aneuploidy. XXY males also undergo X inactivation, so an occasional calico male is seen.

When a female is heterozygous for a deleterious X-linked allele, the effects of the cell lines bearing the normal allele may compensate for the harmful effects of the cell lines bearing the deleterious allele. In females heterozygous for partial colour blindness, for examples, some cell clones in the retina are in fact colour blind, but the presence of other normal clones results in normal colour vision.

X inactivation in humans can sometimes be seen in females heterozygous for certain X-linked traits. Let us elaborate this point. *Ectodermal dysplasia* is a condition known to be X-linked and it involves the lack of some teeth and sweat glands in the affected individuals. Heterozygous females show a mosaic of areas of the jaw with and without teeth and patches of skin with and without sweat glands. In Fig. 4.21. the females in generation III are identical twins; they have developed from a single fertilised egg and are therefore genetically identical. However, due to random inactivation of different X chromosomes during development, they show considerable difference in the location of patches of skin lacking sweat glands. X-chromosome inactivation is one example of a development process that can produce phenotypic differences in genotypically identical individuals. Thus, even clones may differ significantly.

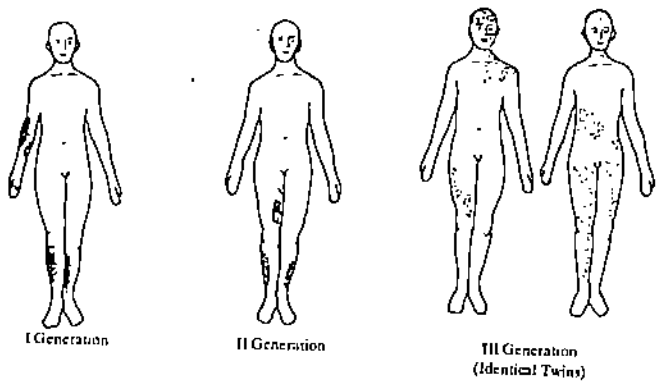


Fig. 4.21: Three generations of females, all showing mosaic phenotypes. All the females are heterozygous for the X-linked gene for *ectodermal dysplasia*. The dark regions in the figure indicate areas in which the sweat glands are missing. Note that the affected areas differ in each woman because the controlling gene is inactivated randomly even in the identical twins in generation III.

Another example that shows the genetic consequences of X-inactivation in females heterozygous for an X-linked gene is the enzyme G-6PD. Cell cultures from an individual heterozygous for the G-6PD gene, has alleles with two forms of the enzyme: G-6PD type A and G-6PD type B. In spite of the cells carrying both the alleles, half the cells express G-6PD type A enzyme, while the remaining express G-6PD type B enzyme. If one cell from this culture was made to grow in isolation then all cells arising from it express the same enzyme type as its parent cell. It confirms the hypothesis that X-inactivation is clonally transmitted, i.e., the same inactivated X-chromosome is passed to daughter cells throughout repeated cell divisions.

The phenomenon of random X-inactivation enables the detection of females heterozygous for a particular trait or enzyme. Some of the cells of this female will have gene expression like deficient male cells and some like normal female cells. Detection of carriers in this way may be of great help in genetic counselling especially in case of X-linked disorders. You would study some more examples of such disorders in Unit 10 of Block 2.

The inactivation of one of the two **X chromosomes** in females **must be reversible,** since females transmit both of their X chromosomes to their progeny in a functional state. This is especially clear in the case of hemizygous male progeny which receives either of the X chromosomes of the mother with equal probability, because the single X chromosome that each son receives must be fully active given that the X chromosomes contains many genes that are vital to the growth and development, indeed to the survival. The **reactivation-"turning on"** of the inactive heterochromatic X chromosomes of mammalian females occurs in germ cell lineages prior to oogenesis. Both X chromosomes of a female are active in the oogonial cells. The maintenance of the germ cells and ovarian structures requires the presence of two X-chromosomes. At this point you may wonder, whether normal reactivation ever fails to occur. There are considerable evidences that indicate abnormal reactivation of the heterochromatic X chromosomes. The most common form of inherited mental retardation in humans is its example.

4.5.2 In *Drosophila*

Dosage compensation occurs in fruit flies, but its mechanism is different from those of the mammals. No barr bodies are found in fruit flies. You have already learnt that in fruit flies, the X chromosome to autosome ratio is responsible for sex determination. Normal females have two X chromosomes and normal males have one X chromosome. *Dosage compensation in this case is achieved by increased transcriptional activity of genes on the single X chromosome in male cells relative to that of each of the X chromosome in female cells.* The *male has hyperactive X-chromosome,* approaching the level of activity of both of the females

This hyperactivity of X-chromosome can be cytologically seen as "puffed" bands in the salivary gland chromosomes (Fig. 4.22). *This is in variance to the inactive X-chromosome in mammals which appears condensed (sex chromation body).*

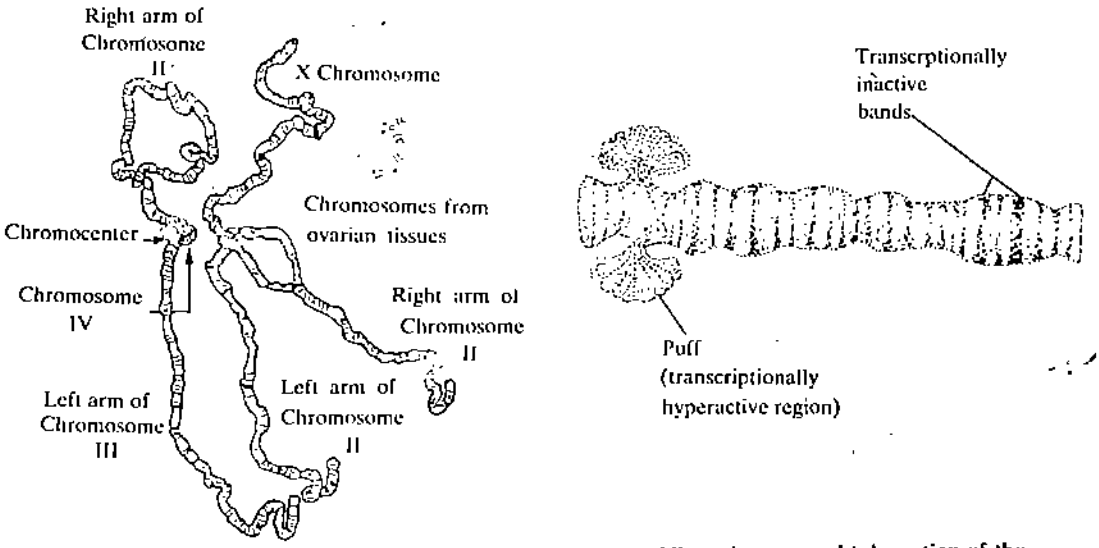


Fig. 4.22: a) The salivary gland chromosomes of *Drosophila melanogaster*. b) A portion of the puffed band enlarged.

SAQ 7

Indicate the expected number of Barr bodies in interphase cells of the following individuals: Klinefelter's Syndrome; Turner's syndrome; and Karyotypes 47 XYY, 47 XXX, and 48 XXXX.

.....

.....

SAQ 8

Cat breeders are aware that kittens with the calico coat pattern are invariably females. Why?

.....

.....

4.6 SUMMARY

In this unit you have learnt that:

- The chromosomes are the carriers of genes and the transmission of chromosomes from one generation to the next closely parallels that of the genes.
- In species with an XX-XY mechanism, genes on the sex chromosomes may be X-linked or Y-linked.
- The mode of transmission of sex-linked traits is different from that of the autosomal ones.
- The dominant X-linked traits are always expressed in both the sexes.
- Recessive X-linked traits show a criss-cross pattern of inheritance. Female heterozygotes are carriers who pass the trait to 50% of their male offspring. Recessive X-linked traits are expressed far less commonly in females than in males.
- Only a few Y-linked genes have been identified and amongst them TDF (testis determining factor) plays a role in male sex determination.
- Sex-limited and sex-influenced traits are the result of genes on the autosomes. The expression of these genes is sexually dimorphic.
- Dosage compensation regulates the level of gene products in such a way that both males and females have the same amount of gene products.

- In female mammals including man, dosage compensation occurs by the inactivation of all the X-chromosomes except one (forming Barr body or bodies) whereas in fruit flies it occurs by the hyperactivation of the single X-chromosome in males.

4.7 TERMINAL QUESTIONS

7 (oose the correct answer.

Which one of the following statements does not apply to human sex chromosomes?

a) carry allelic pairs
b) determine individual sex
c) are identical in women
d) are identical in man
e) both a and d

ii) Barr bodies result from:

a) inactivation of one X chromosome by the Y chromosome
b) a third X chromosome
c) inactivation of one X chromosome for dosage compensation
d) both b and c

iii) A man and a woman are both affected by vitamin D-resistant rickets which is a dominant sex-linked allele. All of the female offspring of these people are affected with rickets, but some of the males are not. What are the possible genotypes of the parents?

a) both are homozygous for the trait
b) the woman is heterozygous for the trait
c) the woman is homozygous and the man is heterozygous
d) this is not possible.

iv) The fly *Drosophila melanogaster* has a gene that codes for white eyes as recessive and X-linked. Red eyes result from the wild type allele at the same locus. A cross between a heterozygous red-eyed female and a white-eyed male would produce:

a) all red-eyed progeny
b) all white-eyed males and all red-eyed females
c) one red-eyed male and one white-eyed male
d) one red-eyed female and one white-eyed female
e) both c and d

v) The gene for pattern baldness is dominant in men, but exhibits recessiveness in women. The difference in expression results form:

a) the gene for baldness being X-linked
b) the gene for baldness being Y-linked
c) the expression of the gene depending upon the hormonal balance of the individual
d) both a and c
e) none of the above

8) Fill in the blanks:

i) Men have _____ pairs of autosomes and one_____pair of sex chromosomes.
ii) Women have _____ pairs of autosomes and one_____pair of sex chromosomes.
iii) The fertilisation of an egg by a Y sperm results in a _____ offspring.
iv) Genes which are Y-linked are called_____.
v) Sex-limited genes are those whose phenotypic expression is determined by the presence or absence of sex _____.
vi) Beard development in humans is generally limited to one sex (male), yet studies indicate that there is no real difference in the number of hairs per unit area of skin between men and women. This indicates that beard development is a _____ trait.

4) Match the terms in column A with their appropriate descriptions in column B:

A	B
i) Dosage compensation	a) X-chromatin
ii) Turner's syndrome	b) inactivation of one X chromosome
iii) Klinefelter's syndrome	c) inactivation of one X chromosome so as to reduce to half the allele
iv) Barr body	d) inactivation of all but one X chromosome
v) Lyon hypothesis	e) phenotypically female (XO)
	f) phenotypically male (XXY)

5) In sheep, the gene h^+ for the horned condition is dominant in males and recessive in females. If a hornless ram were mated to a horned ewe, what is the chance that:
   a) an F_1 male sheep will be horned or
   b) an F_1 female sheep will be horned?

6) In chicken, the gene h, which distinguishes hen-feathering from cock-feathering, is sex-limited. Males may be hen-feathered or cock-feathered, but females are always hen-feathered. If a cock-feathered male (hh) were mated to a homozygous (h^-h^+) hen-feathered female, what patterns of feathering might be expected among the (a) male F_2 and (b) female F_2 progeny?

7) In *Drosophila,* the gene for bobbed bristles (recessive allele bb, bobbed bristles; wild-type allele bb^+, normal bristles) is located on the X chromosome and on a homologous segment of the Y chromosome. Give the genotypes and phenotypes of the offspring from the following crosses: a) $X^{bb}X^{bb} \times X^{bb}Y^{bb+}$, b) $X^{bb}X^{bb} \times X^{bb}Y^{bb}$ c) $X^{bb+}X^{bb} \times X^{bb+}Y^{bb}$, d) $X^{bb+}X^{bb} \times X^{bb}Y^{bb+}$.

8) Make a diagram of a cross between a normal woman (whose father was defective in green colour vision) and a green colour-defective man. Summarise the expected results for sex and colour vision.

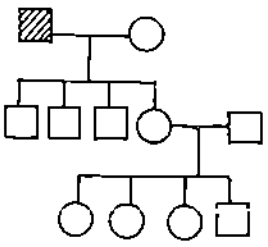
4.8 ANSWERS

Self-assessment Questions

1) Since both the husband and wife have fathers with the X-linked trait, the husband will not carry the trait as it receives its X-chromosome from his mother, but the wife carries the trait as she gets one X-chromosome from her father. a) The probability of having a normal son is fifty per cent. b) The chances of having a normal daughter is hundred per cent as a female is not affected with X-linked recessive trait unless she receives two genes for the trait. c) There is fifty per cent probability of having affected son. d) There is no chance of the daughter being affected, but there is a fifty per cent probability that they may be carrier.

2) Since the daughter is colour blind it can be assumed that both parents carry the genes for the trait. The father's genotype is therefore, hemizygous and that of the mother is heterozygous for the trait. The son must have received the normal X-chromosome from the mother and is therefore, normal.

3)



The probability is fifty per cent for the daughters being carriers and there is fifty per cent probability of the son being a haemophiliac.

4) Only the male children have hypertrichosis ear as the gene for it is Y-linked and the Y is passed from father to son. Females don't have the trait as they never possess a Y-chromosome.

5) Irrespective of whether a Y-linked gene is recessive or dominant it can be recognised as it is always present in hemizygous condition. The Y-linked gene is transmitted from grandfather to male grand-children through the father. The female grand-children are unaffected so are the daughters and the grand-children born from them.

6) Sex-linked inheritance patterns are quite different from those of sex-limited ones. The latter may be expressed in either sex, though with differential frequency. Genes for sex-limited traits express their effects in only one sex or the other and their action is clearly related to sex hormones. They are principally responsible for secondary sex characters.

7) Klinefelter-one; Turner-none; 47 XYY-none; 47 XXX-two; 48 XXXX-three.

8) Because the mosaic coat pattern is due to the expression of sex-linked heterozygous alleles according to the Lyon hypothesis.

Terminal Questions

1) i) The trait occurs more frequently in males than in females.
   ii) Traits are transmitted from an affected man through his carrier daughters to half his grandsons.
   iii) An X-linked allele is never transmitted directly from father to son.
   iv) All affected females have an affected father and a carrier or affected mother.

2) i) e
   ii) c
   iii) b
   iv) c
   v) c

3) i) 22, XY
   ii) 22, XX
   iii) male
   iv) holandric
   v) hormones
   vi) sex-limited

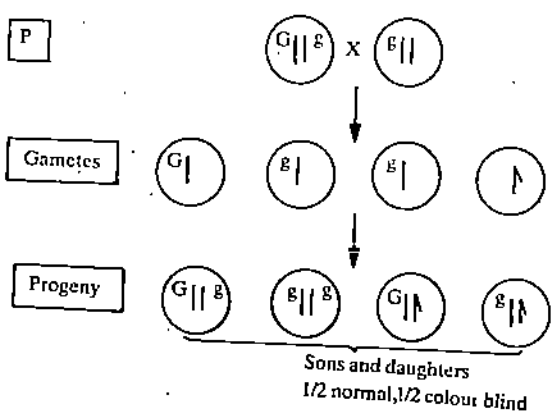
4) i) b
   ii) e
   iii) f
   iv) a
   v) d

5) a) 3/4
   b) 1/4

6) a) 3 hen-feathered : 1 cock-feathered
   b) All hen-feathered

7) a) 1/2 $X^{bb}X^{bb}$ bobbed females, 1/2 $X^{bb}Y^{bb+}$ wild males;
   b) 1/2 $X^{bb}X^{bb+}$ wild females, 1/2 $X^{bb}Y^{bb}$ bobbed males;
   c) 1/2 $X^{bb+}X^{bb+}$ and $X^{bb}X^{bb+}$ wild females, 1/4 $X^{bb+}Y^{bb}$ wild males, 1/4 $X^{bb}Y^{bb}$ bobbed males;
   d) 1/4 $X^{bb+}X^{bb}$ wild females, 1/4 $X^{bb}X^{bb}$ bobbed females, 1/2 $X^{bb+}Y^{bb+}$ and $X^{bb}Y^{bb+}$ wild males

8)



UNIT 5 DEVELOPMENTAL BASIS OF SEX

Structure

5.1 INTRODUCTION

The existence of two different sexes (mating types) is necessary to carry out sexual reproduction, which provides genetic variability. This is achieved in hermaphrodite plants by ensuring cross-fertilisation and in most animals by having two separate sexes. What determines the sex of an individual, must have been clear to you from your study of Units 3 and 4.

For human, it can be said that apart from the differences societies have created between the sexes, there is a basic biological difference between them which is undeniable. This difference is in the reproductive system and its functions. Most people are either 'normal' females or males, and they never question how or why they developed as they did.

Let us look into the genetic basis of development with respect to the differentiation of sexes. We will confine ourselves to the sexual differentiation in man to get a clear understanding of the phenomenon. The egg from female unites with the sperm from male to result in zygote or the first cell of the embryo, whose sexual development is predetermined by sex chromosome (X or Y) contributed by the sperm.

Since the role of genetic information in early development and formation of gametes is basically similar in most plant and animal species, thereby the general concepts can be applied to most of them. Information regarding intersexes and other ambiguous sex anomalies have been included in this unit. Some interesting facts about the sex ratio, factors affecting it and sex selection of offspring are also presented in this unit.

It would be useful to revise Units 3 and 4 as well as brush up your knowledge of reproductive system in man before beginning a study of this unit. You may refer to Unit 8, entitled Reproduction, of the course LSE-05 (Physiology).

Objectives

After studying this unit you would be able to:

- distinguish between sex determination and sexual differentiation (Section 5.2);
- evaluate the role of sex chromosomes in the differentiation of gonads (Section 5.2);
- describe the events, in chronological order, that lead to the differentiation and development of the male and the female gonads (Sub-sections 5.2.1 and 5.2.2);
- describe the causes of development of intersexes (Sub-section 5.3.1);
- contrast between true-, and pseudo-hermaphroditism (Sub-sections 5.3.2, 5.3.3 and 5.3.4);
- define free martins and explain their origin (Sub-section 5.3.5);
- contrast between hermaphrodites and mosaics (Section 5.4);
- describe the importance of sex reversal phenomenon in birds and animals (Section 5.5);
- define sex ratio and differentiate between primary, secondary and tertiary sex ratio (Section 5.6);
- describe the different post-, and pre-fertilisation techniques of sex selection of offspring (Section 5.7).

5.2 GONAD FORMATION

Early in development both XX and XY embryos form undifferentiated or all purpose gonads called **ovotestes** (Fig. 5.1 a). They are bipotential and can develop into either testes (Fig. 5.1 b) or ovaries (Fig. 5.1 c). Direction in which differentiation occurs depends on whether the X-, or Y-bearing sperm has fertilised the ovum. The Y-chromosome is required for production of **testes determining factor (TDF),** which stimulates the 'neutral' gonads to develop in the 6th week of pregnancy, at that time certain events occur that determine which sex the individual will be.

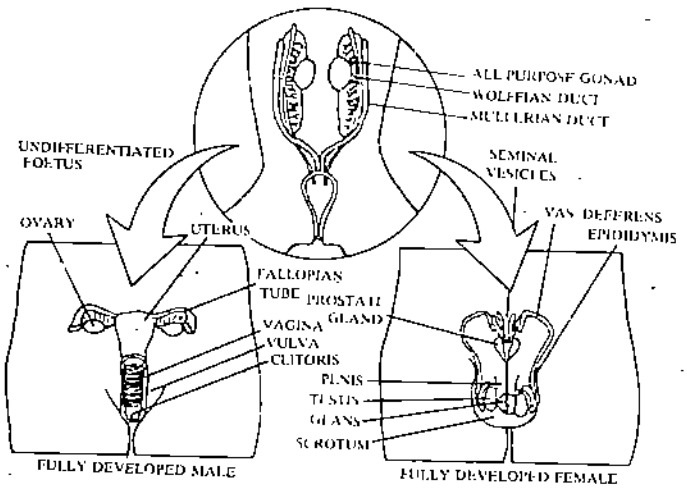


Fig. 5.1: In normal sexual differentiation, an all purpose gonad develops as either an ovary or a testis depending on the combination of chromosomes present. With an X chromosome from both the male and female, it develops as an ovary; with an X from the female and a Y from the male, it develops as testis. In males, the testes produce hormones, or androgens that convert certain embryonic structures into the appropriate male parts. Without the influence of these androgens, the same structures normally develop into the female counterparts.

The TDF gene located on Y chromosome is a master switch, that when turned on, activates an entire series of genes whose function is sex differentiation. No particular female inducing substance is known. So in general it can be said that the gonadal

development proceeds towards femaleness unless switched towards maleness via TDF (Fig. 5.2).

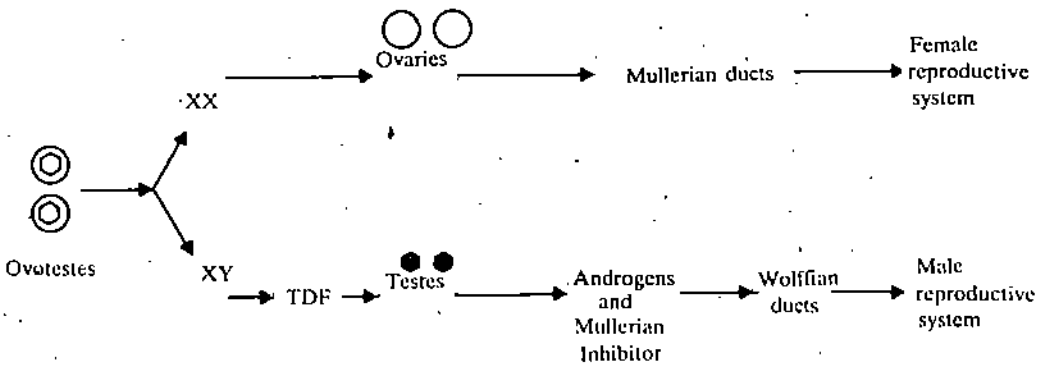


Fig. 5.2: Switch mechanisms that determine expression of human sex phenotype. Development proceeds towards maleness in two steps. First is the production of TDF under control of the Y chromosome. Second is the production of Müllerian inhibitor to inhibit the Müllerian ducts. Without TDF ovaries and Müllerian ducts develop, producing female structures.

5.2.1 Role of Hormones

In the eighth week of gestation, the testes formed earlier, begin to produce the hormone **testosterone,** some of which is converted to closely related substance **dihydrotestosterone,** or **DHT.** Such hormones are called **androgens.** The DHT goes on to convert the all purpose embryonic structures into glans penis, penis shaft, and scrotum. These structures would otherwise develop into their female equivalents: the clitoris, labia minora, and labia majora.

Embryos also start out with two sets of ducts, known as the Müllerian ducts and the Wölffian ducts. In the absence of testes, the Wölffian ducts degenerate, while the Müllerian ducts grow into uterus, fallopian tubes, and the inner part of the vagina. With testes present the opposite happens: androgens produced by the testes stimulate the Wölffian ducts to grow into seminal vesicles, vas deferens and epididymis. At the same time a testicular protein called **Müllerian inhibiting factor** does what its name implies: it prevents the Müllerian ducts from developing into the internal female organs (see also Fig. 5.2).

Thus, the genetic information on sex chromosomes is responsible for the *primary sex determination events.* So, sex is determined at the time when the baby is conceived (fertilisation) and no amount of listening to military marches or looking at pictures of athletes will alter its sex.

Under normal conditions once the gonads develop, further differentiation occurs under the influence of male or female hormones. Sex hormones play an important role in the *development of secondary sexual characteristics.* These include growth of beard, change of voice in males and breast development in females. The actual phenotypic differences between males and females are mediated by hormones, which in turn are produced according to genetic information programmed in the genome. The ability of the cells to respond to sex hormones is also under genetic control and it largely depends on their binding with specific receptors.

Hope now you can distinguish between the phenomena of sex determination and sex differentiation.

SAQ 1

At what stage of differentiation is testosterone produced? Comment on its role in the development of a particular sex.

.....

.....

.....

5.2.2 Role of Genes

Several genes and their products are involved in the process of sexual differentiation. Some of these have been well investigated and the knowledge thus obtained has enabled the geneticists to form a coherent picture of the process of sex differentiation and development of gonads.

H-Y Antigen: It has been known for a long time that sex determination in mammals is effected through the Y-chromosome which carries a factor that initiates male sexual development.

Gene for H-Y antigen is one of those genes which is considered to be important in sex differentiation. It was discovered in 1955 as a transplantation antigen in mice, and is found to occur in all mammalian species. The gene coding for the H-Y antigen is male specific (holandric). It is located on the short arm of the Y chromosome. The evidence for this is provided by the dose-related studies. If anti-H-Y antibodies are added to white blood cells, the H-Y antigen on the surface of the cell will bind to the antibody. And the XYY and XXYY individuals produce twice as much antigen.

The H-Y antigen *serves as an initial signal for primary sexual differentiation.* It is essential for the development of testis in mammals. It is believed that the H-Y antigen has long been involved in sex determination during evolution.

You may recall (from Unit 3) that in amphibians and birds, the female is the heterogametic sex. In such organisms, the female – not the male, expresses the H-Y antigen. It seems that the antigen has been preserved throughout evolution and just like in the mammals, it signals the primary sexual determination.

***Sxr* Gene:** Direct evidence for sex-determining genes on the Y chromosome has come from the studies of inheritance of a dominant sex-reversal *(Sxr) gene in mice. Sxr* causes zygotes with two X-chromosomes to develop as males with testes, but spermatogenesis is absent. Such males exhibit X-inactivation and are mosaic for X-linked genes. The use of recent techniques for DNA manipulation have suggested that when the *Sxr* containing segment of the Y-chromosomes is transferred to the X-chromosome during meiosis, the XX individuals formed develop as males with testes. But the adult XX *Sxr* males are sterile.

Based on the research works involving the *Sxr* trait in mice, it was concluded that whatever be the male-sex determining genes carried on the Y-chromosome, they are essential for instructing the undifferentiated embryonic gonad to develop as a testis, the first step in the male development pathway. In the absence of testis inducing functions, the undifferentiated gonad develops as an ovary.

***Tfm* Gene:** As indicated earlier, secondary sex development is a consequence of the sexual nature of the gonad that develops under the control of the sex chromosomal constitution. The developing testes secrete testosterone, a hormonal signal that induces male development. In the absence of this signal, female development occurs. Male development is controlled by a X-linked gene (Tfm^+) specifying a testosterone-binding protein that is present in the cytoplasm of all cells of male and female. This protein is a regulatory protein, activated by binding of testosterone (an effector molecule) (Fig. 5.3). The protein testosterone complex then enters the nucleus and activates the genes required for normal male differentiation. Mutations of the *Tfm* gene are known in several species including humans causing a syndrome called testicular feminisation (you would study about it in subsection 5.3.2). Cells of mutant *Tfm/Y* embryos are completely insensitive to the masculinising effect of testosterone; consequently the foetus develops all the external sexual characteristics of a female rather than those of a male. However, internally, testes develop rather than ovaries, and the testes suppress the development of fallopian tubes and uterus by secreting another male hormone known as factor chi x , resulting in blind vagina.

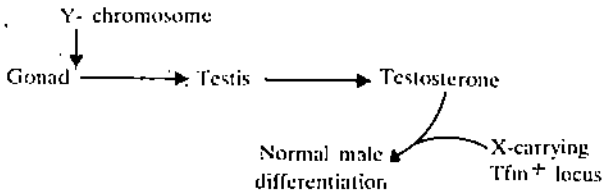


Fig. 5.3: Role of *Tfm*$^+$ locus in normal male development.

Thus a coherent picture of mammalian sexual development has emerged. The *Sxr* and *Tfm* mutations have helped to establish a framework for understanding the general features of the process of gonadal differentiation.

SAQ 2

Comment on the existence of XX males in mice.

.....

.....

.....

.....

.....

.....

.....

.....

5.3 INTERSEXES

This section particularly the subsection 5.3.1, is closely linked to the earlier section 5.2. In order to form a complete picture of what is happening at the time of normal as well as abnormal sexual differentiation, we suggest that you read these sections at a stretch.

5.3.1 Sex Chromosomes and Intersexuality

In humans, provided all developmental processes work correctly, the Y-chromosome confers maleness on the developing embryo and the lack of chromosome Y results in a female. Thus from the very beginning, each cell receives one type of non-reversible, sex labelling through the **sex chromosomes.** And as far as the sex chromosomes are concerned, there is no possibility of reversal. In females, the second X chromosome is required for the production of normal ovary, and some of *the genes on X chromosome are also essential for the production of normal testes in man.*

In Section 5.2, you have already learnt that the development of gonads of both the sexes begins from common, undifferentiated structures — the ovotestes. Gonadal sex differentiation begins earlier in the cells carrying the male XY sex chromosomes. The ovarian changes in a female occur later on in the developmental sequence.

If the male induction process does not begin according to schedule — early in embryonic development — abnormalities can be expected in gonadal genitalia at birth. If no induction occurs, the embryo will continue towards female development. The *correct sex chromosome complement and precise timing* are both necessary for induction processes of male development. Any disturbance in this process can create varying degrees of abnormal anatomical sexual development. Such errors or accidents in development can lead to the incomplete development of one sex or the partial display of both sexes — **intersex** in persons with normal sex chromosomes.

5.3.2 Male Pseudohermaphroditism

The most common cause of male Pseudohermaphroditism is **testicular feminisation,** which is an inherited sexual disorder. Those affected, display a normal feminine appearance and behaviour, though genetically they are male - XY. Such individuals develop female secondary sexual characters but are sterile. The vagina ends in a blind pouch and testes do not descend to their normal location in the scrotum but are located in the abdominal area. The testes produce female hormones — estrogens responsible for the secondary sexual characteristics. It is believed that *a defective gene alters the ability of Y chromosome to confer maleness on the embryo.* Some male Pseudohermaphrodites produce testosterone, but lack cytoplasmic testosterone receptors necessary for normal male masculinisation. Other male Pseudohermaphrodites have normal receptors but fail to produce testosterone. The gene for producing the testosterone receptor is located on the X chromosome.

In some cases of male Pseudohermaphroditism, an opening is present beneath the penis that simulates a vagina. The scrotum or scrotal sac is usually small and does not contain the (undescended) testes. Many of these individuals have normal-appearing female genitals, and may even undergo feminisation at puberty although they do not menstruate.

The male pseudohermaphrodites are sex chromatin negative. They do not show a Barr body and have the normal male sex complement of XY not XX.

5.3.3 Female Pseudohermaphroditism

This is a recessively inherited form of pseudohermaphroditism. The female pseudohermaphrodite has the normal-XX sex chromosome complement, but her genitalia display various degrees of phallic development. Female pseudohermaphroditism commonly results due to the **Congenital Androgenital Syndrome.** These females appear as intersexes because there is overproduction of testosterone due to hyperactive adrenal glands. Excessive testosterone inhibits the complete differentiation of the female duct system and stimulates the development of male sex organs and secondary sexual characters. There is variable expression of genital duct development. That is, the time when the male or female ducts depart from the normal depends on the levels of the hormones present.

If this recessive gene is present in an XY embryo then a **precocious male** develops, i.e., adult male characters appear at a very early age in the child. Cortisone, which inhibits testosterone activity, provides effective treatment in these conditions if administered early in embryogeny.

5.3.4 True Hermaphroditism

A true hermaphrodite possesses functional male and female reproductive systems or retains the bipotentiality of the embryonic gonad. This condition, however, is rare in humans. Such individuals show varying degrees of *intersexual development of genitals.* In many hermaphrodites, the chromosomal abnormalities in the form of XX/XY mosaicism are evident. You would study the sex mosaics in the following section.

5.3.5 Free Martins

Besides humans, sexual differentiation anomalies too occur in animals resulting in intersexes. In cattle, sheep, pigs and goats sometimes sterile intersex animals are born. This condition may occur when male and female foetuses are developing in the uterus simultaneously (i.e., twins) and there is a fusion of the placental membranes permitting mixing of foetal blood. As a result testosterone from the male induces male characteristics in the female twin. This kind of animals are called free martins. Free martins have XX-chromosome constitution, female internal genitalia; but male external genitalia and secondary sexual characters.

SAQ 3

If the diagnosis is testicular feminisation, what will be the genotype and phenotype of that individual?

.....

.....

.....

.....

.....

.....

.....

.....

SAQ 4

A human XX baby has both male and female characters. How could such an error in sexual development occur?

.....

.....

.....

.....

.....

.....

.....

5.4 SEX MOSAICS

In many species, some individuals are made up of several cell lines, each with different chromosome number. These individuals are referred to as *mosaics* or *chimeras*. Such conditions are the result of certain mishaps that take place at various stages of embryological development, and these affect the daughter cells. Three causes have been known that result in mosaics.

i) The first one is *mitotic nondisjunction of sex chromosomes* (Fig. 5.4). It happens in a manner similar to that found in meiosis. In a male zygote (XY), there are two possibilities (Fig. 5.4 a&b). One, there may be nondisjunction of X chromosome resulting in one cell line having XXY and other having YO chromosome (Fig. 5.4a).

Nondisjunction is the failure of chromosomes to properly separate into their respective nuclei during nuclear division.

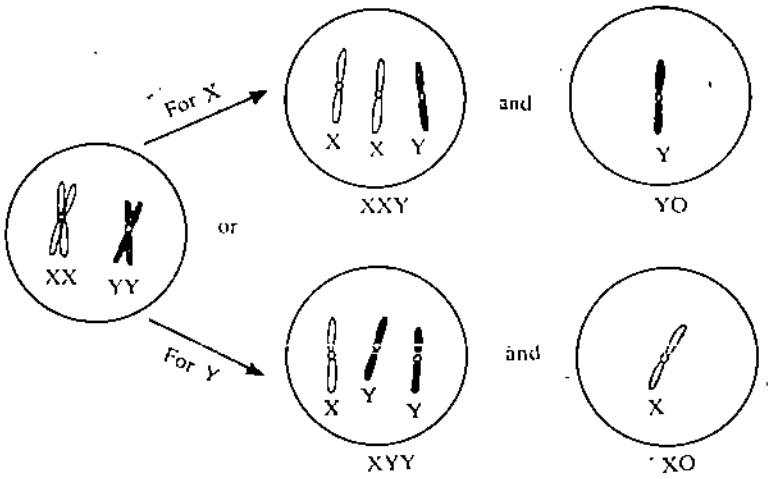


Fig. 5.4: Mosaicism due to nondisjunction of sex chromosomes in ♂ zygote. (a) the disjoining of X chromosome results in XXY and YO cell lines, (b) and the nondisjunction of the Y chromosome leads to the formation of XXY and XO cell lines.

The latter one is inviable. In the second possibility there is nondisjunction of Y chromosome resulting in cell lines: XYY and XO (Fig. 5.4 b). In a female zygote nondisjunction would result in XXX and XO cell lines (Fig. 5.5).

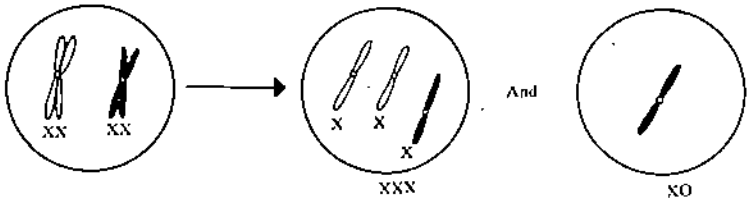


Fig. 5.5: Mosaicism due to nondisjunction of sex chromosomes in ♀ zygote leading to the formation of XXX and XO cell lines.

ii) The second cause is **chromosomal lagging.** The X chromosome may lag and fail to reach the pole at anaphase to be incorporated in the daughter nucleus (Fig. 5.6). It is left out in the cytoplasm where it disintegrates. As a result daughter cells of chromosomal complement AAXX and AAXO are formed. You may remember that in *Drosophila* AAXX results in females and AAXO in males. Therefore, the developing embryo having a mixture of AAXX - (female's) and AAXO - (male's) genotypes, have some parts that are like males and others like females. Flies with such chromosomal complement are seen in actual and are known as **gynandromorphs** or **gynanders** and also **sex mosaics.**

If the mishap, i.e., chromosomal lagging occurs at the one-called stage (Fig. 5.6) then **bilateral gynandromorphs** (Fig. 5.7.) are formed. If it occurs at a later stage of development then small patches of male tissue will be present among the female background.

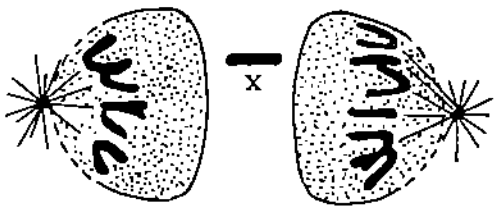


Fig. 5.6: *Drosophila melanogaster.* A lagging X chromosome in the first cleavage of zygote, illustrating the origin of bilateral gynandromorphs.

Sex chromosome mosaics are also known in humans. The karyotype of some of these mosaics are XX/X, XY/X, XX/XY, and XXX/X. In the case of XX/X persons, one line of cells have only one X chromosome whereas the other has two Xs. Similarly, in XXX/X individuals, one cell line has three Xs and the other has one X.

iii) The third cause of development of mosaics is by **dispermic fertilisation.** A known case of XX/XY chimera was formed by the above manner. In this two sperms, one X-bearing, and the other Y-bearing, unite with the nucleus of egg and one of the polar bodies respectively. The 'zygotes' thus formed, fuse and develop into one individual.

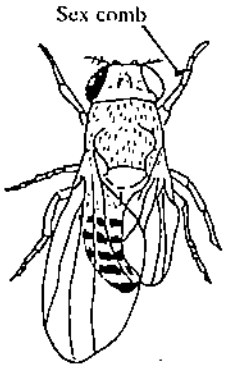


Fig. 5.7: *Drosophila*, gynandromorph. Left side is wild type XX female. Right side is XO male, hemizygous for white eye and miniature wings.

Because mosaicism can occur at any time producing a highly variable range of affected phenotypes, it is plausible that there would be many persons in the population that are mosaics to some degree. Such persons, although phenotypically normal, may have a reproductive risk. A portion of their gonadal cells could be chromosomally abnormal.

It must be noted that these *mosaic persons are not gynandromorphs.* No discrete patches of definite male or definite female tissues are seen in any kind of human mosaic. The mosaics may show a range of abnormal phenotypes, but hormonal activity in mammals will not allow the development in any one of them into two distinct classes of cells, i.e., those with typical female features intermingled with those that are obviously male.

Departures from normal number of chromosomes because of nondisjunction and other mishaps can affect autosomes as well as sex chromosomes. Any chromosome

may become altered in structure. These and other chromosomal anomalies would be discussed in detail in Units 9 and 10.

5.5 SEX REVERSAL

In the female birds normally only one gonad develops while the other side is suppressed. If ovariectomy is performed, the undifferentiated gonad develops into testis. In such a case, secondary sexual characters also appear and the resultant male can even father an offspring (i.e., it is fertile). A complete sex reversal can be seen in birds (unlike man) if hormones are injected in the embryo early in development.

In mammals early removal of gonads (castration, ovariectomy) or their destruction by disease results in an intersex appearance as it leads to partial or complete reversal of secondary sexual characters. Boys castrated before puberty (eunuchs) were employed in Middle East, North Africa and China as harem chamberlains. In Italy they were part of operatic stages and church choirs as the male hoarse voice does not develop in them.

From the above discussion it is evident that the potential hermaphrodite state can be more easily activated in birds than in mammals. In animal husbandry, castration is used to increase fat content and improve quality of meat by avoiding undesirable odours from meat, a characteristic of meat from adult males. It also contributes to docility of animals like oxen which are used for carrying loads.

A human intersex may opt for a sex reversal irrespective of the genetic constitution. Plastic surgery and hormonal treatment are given depending upon internal and external genitalia present. Such individuals, however, remain sterile.

5.6 SEX RATIO

Because sex is determined by the Y chromosomes, and because males produce X-, and Y-bearing gametes in approximately equal numbers, Mendel's law of Segregation predicts that the sexes should occur in equal proportions, or as commonly expressed in a male : female ratio of 1:1. In most species most of the time, the number of males and females is about equal, but this is not the case always. In humans, for example, different **sex ratios** occur for different age groups.

5.6.1 Primary Sex Ratio

It is the proportion of males and females conceived in a population. It is deduced from the male and female frequency in abortions, miscarriages and still-births, in addition to the live birth records. Although a 1:1 ratio is expected, as there is equal chance of the X-, and Y-bearing sperms to fertilise the egg, in actuality a deviation has been noted, which is more than 1.6 males : 1 female. This may be caused by one or more than one of the following three factors:

1) There may be selective fertilisation of the egg with a Y-bearing sperm due to its greater motility than the X-bearing sperm. The Y-bearing sperms may be more motile because of their lower chromosomal mass.

2) The environment of the female ducts may be more favourable for the survival of the Y-bearing than the X-bearing sperm.

3) The egg may react more preferentially to the approach of Y-bearing than to the X-bearing sperm.

5.6.2 Secondary Sex Ratio

It reflects the proportion of males to females at birth. It is easy to determine, but has disadvantage of not accounting for disproportionate embryonic or foetal mortality. World-wide data of secondary sex ratio indicates a value slightly different from the expected 1:1 and is about 1.06 males : 1 female. This unequal sex ratio at birth

appears to be an established phenomenon for which no clearly defined biological mechanism has been identified.

5.6.3 Tertiary Sex Ratio

The sex ratio at any particular time after birth is known as tertiary sex ratio. It has been found that in the age group around twenty years, the number of males exceed females. This trend gradually gets reversed and female population is higher after fifty years onwards due to increased male death rate in every age category (Table 5.1).

Table 5.1: Approximate sex ratios for the human species.

Time	Male:Female
Primary ratio (conception)	> 106:100
Secondary ratio (birth)	106:100
Tertiary ratio (post-natal)	
Second to fourth decades	100:100
Fifth decade	90:100
Sixth decade	70:100
Eighth decade	50:100
Tenth decade	20:100

Thus the excess of males at conception and birth progressively diminishes throughout life and the males get outnumbered more and more by females with time. The biological weakness of males resulting in higher mortality is partly understandable as they have only single X-chromosome. If this X-chromosome carries recessive alleles resulting in lower viability, sublethality or lethality, then a male carrying it would be affected. On the other hand, a female would escape the deleterious effects of these alleles as her other X-chromosome may carry a normal allele which compensates for the defective one. Interestingly a perfect sex ratio (1:1) is ensured during the prime reproductive years, irrespective of the primary and secondary sex ratios. This appears to be an evolutionary adaptation to facilitate mating and get the process of sexual reproduction going.

SAQ 5

What is the genetic basis of the approximately 1:1 sex ratio in majority of animals?

.....

.....

.....

.....

.....

.....

SAQ 6

In a population of 1440 individuals born, the number of males was 741 and of females 699. What is the sex ratio? How would you explain the excess of births of one sex as compared to the other?

.....

.....

.....

.....

.....

.....

.....

.....

.....

SAQ 7

In a small town the number of males under 21 years of age exceeded the number of females. From 22 years, the trend gradually reversed and the sex ratio was 100 males : 146 females for persons over 65 years. Account for this reversal.

5.7 SEX SELECTION OF THE OFFSPRING

It has been man's desire to choose and control the sex of his offspring. Early efforts in selecting sex of newborn involved changes in physical environment or in human behaviour around the time of conception. These methods were scientifically unsound. Currently three main approaches are employed for the selection of sex:

i) selective abortion of the foetus of unwanted sex;

ii) timing of fertilisation relative to ovulation;

iii) separation of sperm *in vitro* followed by artificial insemination. Of these three approaches the first is post-fertilisation selection, while the other two are known as pre-fertilisation techniques.

5.7.1 Prenatal Sex Determination

It has also been human curiosity to know the sex of the child before birth. In recent years, the technique of amniocentesis coupled with cell culture methods has opened the way to prenatal diagnosis of many genetic defects and the determination of sex. Amniocentesis involves withdrawing a sample of amniotic fluid surrounding the foetus, around 16 weeks of gestation. The foetal cells thus collected are stained and examined. They enable a highly accurate prediction of the sex of the foetus. Such a predetermination of sex is useful in case of families having a history of X-linked diseases. Similarly, chorionic villi can now be sampled much earlier in pregnancy (4 to 8 weeks of gestation) and cultured for prenatal detection of genetic defects and sex of the growing embryo. This early detection gives option to terminate the pregnancy, if the foetus is of the undesired sex or has any abnormality.

In a country like India, where we face the population explosion problem, predetermination of sex has been pleaded as a help in family planning. Many couples have large families for want of a child belonging to a particular sex, thereby increasing the population. Predetermination of sex with one of these techniques could be offered to such couples so that they are able to have a child of their choice. But preference of males in our society would lead to a very large number of females being aborted. It is difficult to predict the extent to which normal sex ratio would be disrupted, if people could choose the sex of their offspring. However, judging from past records (where male births have been favoured in most countries including India) human intervention is expected to alter the balance of nature. The result of this change is more likely to create new problems for society. Hence this approach cannot be recommended as a family planning method till the time the attitude towards both sexes becomes the same in our society.

5.7.2 Timing of Fertilisation

Earlier reports suggested that male conceptions are more likely to occur if fertilisation is around the time of ovulation and females from fertilisation after ovulation time. The reason for this may be the difference in maternal hormone levels during the fertile period. This hypothesis is likely to form the basis for manipulation of sex ratio in man as well as other species where artificial insemination can be carried out.

5.7.3 Separation of X and Y Chromosome Bearing Sperms

The slight difference in size, shape, weight and density between X and Y populations of spermatozoa have been taken into consideration for the physical separation of these two types.

Separation by Ultra-centrifugation: The X and Y-bearing sperms can be separated by ultra-centrifugation where the heavier X-bearing sperms settle down and the lighter Y-carrying ones float. This difference in weight of the two types of sperms is due to the extra chromatin of X-chromosomes as compared to the Y.

Electrophoretic Separation: This method is based on the assumption that X and Y-bearing sperms differ in their electric charge and can therefore be separated in an electric field.

Separation Based on Differential Motility: Columns of different types are used to separate the sperms on the basis of their differential motility. The semen is layered over albumin or sephadex columns. The sperms are then allowed to sediment or swim upwards in the columns, the Y-sperms being more motile move faster thereby getting separated from the X-sperms. So far this appears to be the most economic and extensively used method.

The control and manipulation of sex ratio in cattle is important in dairy cattle. Cows are needed in large numbers because of their milking ability and one bull is sufficient to mate with many cows. Therefore it is desirable to have more cows than bulls. This can be achieved by artificial insemination, using X-bearing sperms.

SAQ 8

When cows have twin calves of unlike sex (fraternal twins), the female twin is usually sterile and has masculinised reproductive organs. What is the calf in such a situation known as? Discuss how such an individual comes to exist?

.....

.....

.....

.....

.....

5.8 SUMMARY

In this unit you have learnt that:

- In mammals including man, sexual development proceeds from a stage where sex is not specified, as the embryological gonad is bipotential having the ability to develop into either sex.
- The genetic information on the sex chromosomes is responsible for the primary sex determination events.
- Y-chromosome plays a crucial role in sexual development of mammalian embryos as the presence of Y, determines male development and in its absence female development proceeds. Y-chromosomes carries genes for testis determining factor (TDF) which promotes testes (male gonad) development. There is no particular

substance which induces femaleness and so it can be deduced that in the absence of TDF, ovaries develop. In mammalian embryos the testes produces testosterone (male hormone) and Müllerian inhibitor. These help in the formation of male reproductive tract. In the absence of Mullerian inhibitor the female reproductive tract develops. Sex hormones are essential for normal male and female sexual development.

- Sex chromosomes are not the sole determinants of human sexual identity as there are males with an XX genotype. Clearly, there are genes on the autosomes that go into shaping the final sex phenotype of an individual. When these interrelated factors malfunction, human and animal intersexes may develop. Both male and female pseudohermaphrodites and the true hermaphrodites (contain both male and female reproductive organs) have been found
- Mosaic individuals contain cell lines with different genetic constitution. Such a condition may result due to mitotic nondisjunction, chromosome lag, or dispermic fusion.
- Sex ratio is the proportion of males to females in a population. Depending on when the sex ratio is assessed it is known as primary, secondary or tertiary.
- Man has always been curious to learn about the sex of unborn children, this is now possible by the techniques of amniocentesis and chorionic villi culture.
- It has been man's desire to produce sex of his choice and this is within his reach now as X and Y bearing sperms can be separated by various techniques.
- Since interference of man may disrupt the sex ratio and hence the natural balance, one should be very careful employing any of these techniques.

5.9 TERMINAL QUESTIONS

1) A sex-linked dominant mutation in the mouse, testicular feminisation (*Tfm*), eliminates the normal response to the testicular hormone – testosterone during sexual differentiation. An XY animal bearing the *Tfm* allele on the X chromosome develops testes, but no further male differentiation occurs. The external genitalia of such an animal are like female. From this information, what might you conclude about the role of the sex chromosomes in sex determination and differentiation in mammals?

2) a) In mice as in all mammals, the male is a heterogametic sex. Assume that sex-linked lethal gene is present in a strain of animals and that this causes the death of the embryo. How would this affect the sex ratio?

   b) Answer the same question if sex-linked lethal genes were present in a strain of chickens.

3) For humans, give the genetic sex of (a) true hermaphrodites, (b) masculinising male pseudohermaphrodites, and (c) feminising male pseudohermaphrodites, and (d) female pseudohermaphrodites?

4) Name the ways in which mosaicism can occur in *Drosophila*.

5) a) What is sex differentiation? Describe the role played by testis determining factor (TDF) in this.

   b) Distinguish between: (i) Hermaphrodite and Intersex; (ii) Androgenital syndrome and free martins; (iii) Primary and Secondary sex ratio.

   c) What are the advantages and disadvantages of prenatal sex detection? What social problems do you foresee if the secondary sex ratio is imbalanced?

6) Match the following:

i) Testosterone	a) Y-linked
ii) *Tfm*	b) Intersex
iii) Free martin	c) Sex mosaic
iv) Sxr	d) Produced by Testis
v) Gynandromorph	e) X-linked

7) Fill in the blanks:
   i) The genotype is ..... and phenotype is ..... in an individual who is suffering from testicular feminisation syndrome.
   ii) Primary sex ratio shows an increased number of ..... as compared to ..... at conception.
   iii) The two techniques primarily involved in early prenatal sex determination are ..... and ..... .
   iv) X and Y-bearing sperms can be separated on the basis of their difference in ..... and ..... by ultra-centrifugation and electrophoresis.
   v) To suppress the development of female reproductive tract the factor responsible is ..... .

5.10 ANSWERS

Self-assessment Questions

1) Once the testis is formed due to the presence of Y-chromosome it produces testosterone which stimulates the development of rest of the male reproductive system and suppresses the female development.

2) When a male has a XX genotype it can be assumed that one of the two X-chromosomes present carries the *Sxr* gene. This sex modifying gene is responsible for spermatogenesis and may have been transferred from the Y-chromosome to the X-chromosome. Presence of the *Sxr* gene modifies the sex to male in spite of the presence of two X-chromosomes.

3) The genotype of the individual is XY and the phenotype is that of a female in spite of the testis being present. The development of normal, functional male reproductive tract is prevented as the tissues are unable to respond to the male hormone testosterone. Since female development occurs in the absence of male differentiation, female genitalia and secondary sexual characters are present.

4) This may be a case of androgenital syndrome where because of the XX-chromosome constitution the individual is expected to develop into a female, while the excessive hormone adrenalin produced by the adrenal gland leads to masculinisation. The resultant individual has both male and female secondary sexual characters.

5) In majority of animals one of the two sexes is heterogametic (usually the male) and produces equal number of gametes of each type. These after fertilisation with gametes of the homogametic sex (the female) result in same proportion of males and females.

6) The ratio for males 741/1440 = 0.52 and for females is 690/1440 = 0.48. The sex ratio is slightly altered from the expected ratio, which is 1:1, with an excess of male births. The reason for this may be: (1) selective fertilisation of egg with Y-bearing sperms due to their faster motility; (2) the environment of the female reproductive tract may be more favourable for the survival of Y-bearing sperm; (3) the egg may react more preferentially to the approach of Y-sperm as compared to the X-sperm.

7) Excess of males at conception and birth progressively diminishes throughout life and the males get outnumbered by females gradually. This is due to the presence of a single X-chromosome in males unlike females. If this single X-chromosome carries deleterious gene, it can express itself in a single dose (hemizygosity) leading to differential mortality.

8) The male hormones of the male twin are produced earlier than the female hormones of the female twin. As a result, sexual differentiation in the female is abnormal because of the common placenta. The calf in such a situation is known as free martin.

Terminal Questions

1) The Y chromosome is essential for initial differentiation of testicular tissue. However, hormones subsequently produced by the testes are responsible for differentiation of the remainder of the male reproductive tract.

2) a) The females carrying the lethal gene would produce only half the expected number of male offspring since no male could be born with the allele. The 1:1 sex ratio would become distorted to 2:1 in favour of females.

   b) The males carrying the lethal gene would produce only half the expected number of female offspring. No hen would carry the allele. The sex ratio of 1:1 would change in favour of males to 2:1.

3) a) XX/XY
   b) XY
   c) XY
   d) XX

4) Mitotic nondisjunction and chromosome lag.

5) a) The differentiation and development of sex organs of either sex by the combined action of several genes. For the second part of the question refer to Section 5.2.

   b) i) See Section 5.3
   Hint: True hermaphrodites possess both male and female reproductive systems developed to various degrees. Intersexes have one sex or display some feature of both the sexes.
   ii) See Sub-sections 5.3.3 and 5.3.5.
   iii) See Sub-sections 5.6.1 and 5.6.2.

   c) Write your viewpoint on the issue.

6) i) d
   ii) e
   iii) b
   iv) a
   v) c

7) i) XY, female
   ii) males, females
   iii) amniocentesis, chrionic villi sampling
   iv) weight, electric charge
   v) mullerian inhibitor

UNIT 6 LINKAGE, CROSSING-OVER AND CHROMOSOME MAPPING

Structure

6.1 INTRODUCTION

The independent assortment of genes during meiosis at the time of gamete formation, is one of the fundamental processes of genetics. There are, however, situations where independent assortment of certain genes does not occur. These genes are transmitted together enblock because they are located on the same chromosome. Genes which are located on the same chromosome and transmitted together are called *linked genes* and are said to belong to a *linkage group*. Since it is the chromosome and not the gene which is the unit of transmission at meiosis, linked genes are not free to undergo independent assortment. How does one determine as to which genes on a chromosome are linked together? *Test crosses are useful for identifying linked genes.* This knowledge of linked genes and linkage group are used to construct a *linkage map* or *genetic map* of each chromosome.

The information obtained from genetic mapping is very useful in many aspects of genetic analysis. For example, chromosome maps can tell us whether certain genes that produce a particular phenotype are located together on the same chromosome. The knowledge thus obtained can be used in understanding; the regulation of gene expression in an organism; the DNA sequences in and around a particular set of genes, and it can also be applied to DNA recombinant research/technology.

In many instances, the linked genes, contrary to the expectations, may not be transmitted together. This is because, during the first meiotic prophase when homologous chromosomes undergo pairing, there is a reciprocal exchange of chromosome segments. This process is known as crossing-over. Crossing-over is responsible for the non-transmission of linked genes together and generation of recombinants.

In this unit you would learn about the concept of linkage, the evidence for the physical exchange of chromosome segments (crossing-over) resulting in recombination, and the construction of genetic map of a chromosome.

Before you begin your study of this Unit, we suggest that you once again look back and refresh your memory regarding the following aspects:

1) See Unit 17 of Cell Biology Course (LSE-01) and revise the chromosome behaviour during meiosis. Pay particular attention to the following points: i)When does chiasmata formation and crossing-over occur? ii) Between these structures is there an exchange of genes?
2) The dihybrid cross and Mendel's law of independent assortment in Unit 1 of this course, LSE-03 (Section 1.6).
3) Replication of DNA in chromosomes, see Unit 13 of LSE-01 (Section 13.4).

Objectives

After studying this unit you would be able to:

- explain the concept of genetic linkage and the basic terminology involved (Section 6.2);
- contrast between linked and unlinked genes giving examples (Section 6.2);
- distinguish between complete and incomplete linkage (Section 6.2);
- state the importance of testcross in understanding/detecting linkage (Section 6.2);
- describe the experiments of Bateson et al. and of Morgan, and relate how their analyses contributed to the development of the concept of linkage (Subsection 6.2.1);
- describe the cis-, and trans-configuration of linked genes, and show how these affect the phenotypic ratios (Subsection 6.2.1);
- identify from a given data, whether it is a case of linkage (Section 6.2);
- describe the concept of crossing-over highlighting its main features (Subsection 6.3.1 to 6.3.2);
- describe the cytological basis of crossing-over (Subsection 6.3.3);
- illustrate the molecular mechanisms of crossing-over (Subsection 6.3.4);
- explain how linkage maps are constructed, and their utility in genetic analysis (Section 6.4);
- comment on why Mendel did not find linkage in his experiments with pea (Section 6.5);
- solve genetic problems involving linkage, crossing-over and gene mapping (Section 6.2 to 6.4).

6.2 LINKAGE

All organisms have a large number of genes, their number being much more than the number of chromosomes. All genes on the same chromosome do not assort independently and therefore, provide another exception to Mendel's laws of inheritance. Sex linkage being one of the exceptions that you have already studied in Units 2 and 4. Genes whose patterns of inheritance deviate from that of independent assortment are often **linked.**

You have seen in Unit 1, Section 1.6 a dihybrid cross between a true-breeding strain of peas with round, yellow seeds (SSYY) and one with wrinkled, green seeds (ssyy). Independent assortment during gamete formation in the F_1 hybrid (SsYy) is expected to yield four types of gametes in equal proportions (SY, Sy, sY and sy), resulting in four phenotypes in the F_1 progeny in ratios of 9:3:3:1. But if parental association of alleles (i.e., S with Y, and s with y) are maintained during gamete formation in F_1 hybrid, only two types of gametes would be produced : SY and sy. In other words, the genes for seed shape and seed colour would be **completely linked** in the F_2 progeny would then be round, yellow seeds (SSYY and SsYy) and wrinkled, green seeds (ssyy) in a phenotypic ratio of 3:1. You should note a point here that the individuals of F_2 progeny are classified as round, yellow seeds or wrinkled, green seeds based on their phenotypes. **Complete linkage** of these genes makes it appear that round-yellow and wrinkled-green are inherited as **single trait.**

One of the most preferred methods to study linkage is by making test crosses in which one parent is homozygous for the recessive genes under study. The phenotypes of progeny in a test cross directly reflect the gametic types produced by the heterozygous parent. Figure 6.1 shows the use of test cross for studying independent assortment and complete linkage.

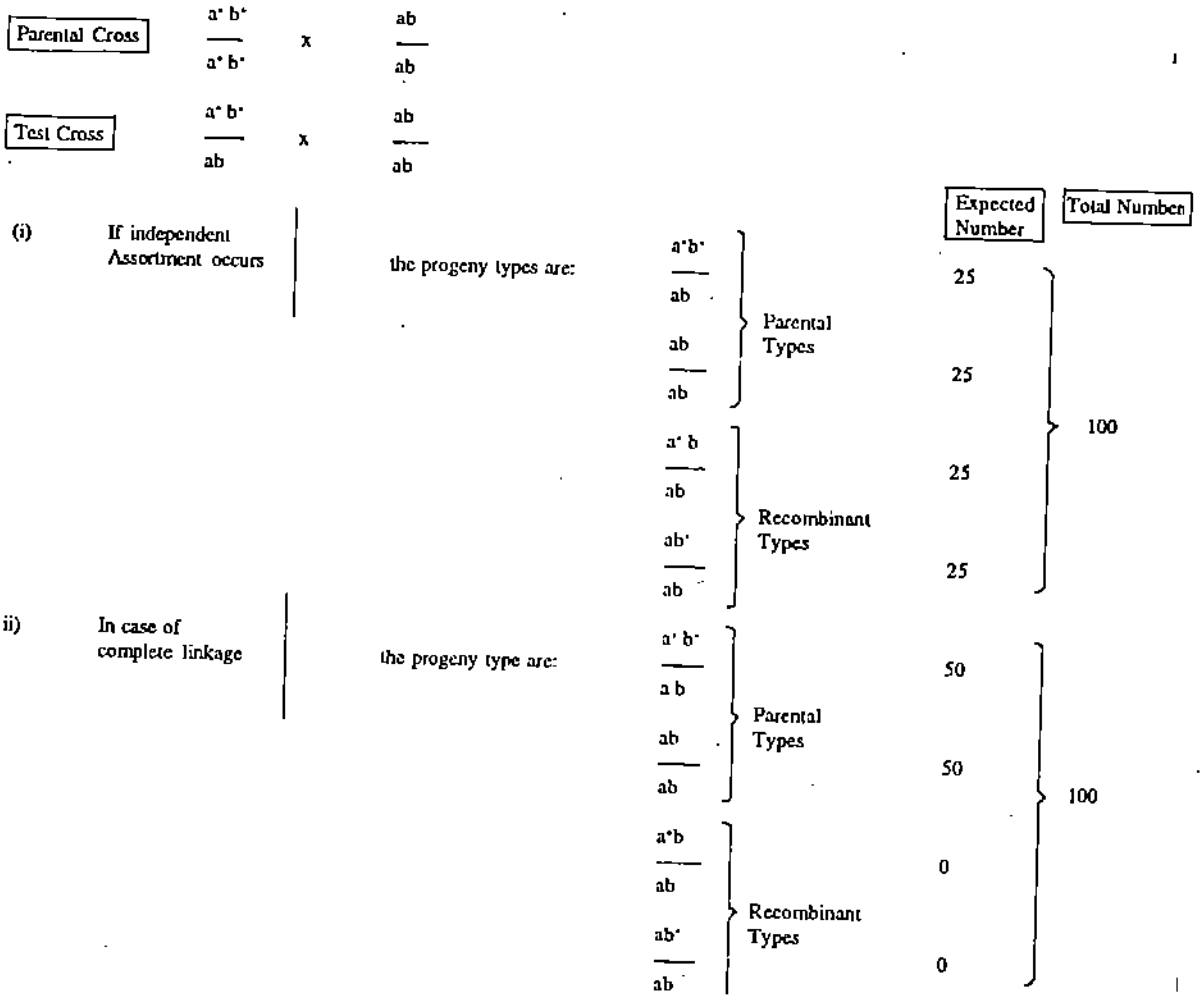


Fig. 6.1: Progeny genotypes from a test cross involving alleles of two genes a and b. The observed number of each progeny genotype are those expected among 100 progeny if (i) genes a and b assort independently, or (ii) they are completely linked. The bar $\left(\frac{ab}{a^+b^+}\right)$ separates alleles of one homolog from those of the other.

In the first case, where the independent assortment of alleles of two genes occurs, the four phenotypes are produced in equal numbers. Two classes of progeny exhibit the same association of alleles as seen in the parents (a^+b^+ and ab), and the other two classes of progeny exhibit new (recombinant) association of alleles (a^+b and ab^+). If parental-type and recombinant-type progeny occur in equal numbers, the genes *a* and *b* assort independently during meiosis in the heterozygous parent, and are said to be not **linked**. The **recombination frequency** between the two genes is defined as the summed frequency of recombinant types among the total progeny. In the example shown in Fig. 6.1 it is 50/100 = 0.50 or 50%). *Independent assortment is characterised by a recombination frequency of 50% and may indicate that the two genes possibly reside on different chromosomes.*

In the second case shown in Fig. 6.1, the complete association of alleles of two genes results in two genotypes among the progeny of a test cross. These progeny exhibit genotypes as those of the parents. Such complete linkage also indicates that the two genes may be located on the some chromosome.

In the above test cross (Fig. 6.1) we have not considered in a test cross whether the tester homozygous, recessive) is the ♀ parent or ♂ parent. We shall now examine two situations: one in which the ♀ parent is the tester and the ♂parent is doubly heterozygous (Fig. 6.2); and the second in which the ♂ parent is the tester and the ♀ parent is doubly heterozygous (Fig. 6.3). In these two crosses we shall see if there are any differences in the ratios of progenies of these two test crosses.

The first cross in *Drosophila* (Fig. 6.2) shows *complete linkage* of the recessive autosomal traits for black body (*b*) and purple eyes (*pr*). When the male is the doubly heterozygous parent, and there is complete linkage, then only sperms with the parental combinations of alleles are produced and transmitted to the progeny resulting in half black-body, purple-eyed flies, and half wild-type flies. Recombinations of parental alleles are not observed, so no black-body normal eye flies, or normal-body, purple eye flies are produced. The data shown in the Fig. 6.2 indicates that in a cross where ♀ is the tester and ♂ is doubly heterozygous, the progeny exhibits parental phenotypes which indicates the occurrence of complete linkage in this case.

Allele	Phenotype
b	black body (recessive)
b⁺	wild type (dominant)
pr	purple eye (recessive)
pr⁺	wild type

Parental cross: $\frac{b\ pr}{b\ pr}$ ♀ x $\frac{b^+\ pr^+}{b^+\ pr^+}$ ♂

Test Cross: $\frac{b\ pr}{b\ pr}$ ♀ x $\frac{b^+\ pr^+}{b\ pr}$ ♂

Progeny Type	Phenotype	Observed Number	Number expected from independent assortment
Parentals	$\frac{b\ pr}{b\ pr}$ Black body, purple eye	374	193
	$\frac{b^+\ pr^+}{b\ pr}$ wild type	398	193
Recombinants	$\frac{b\ pr^+}{b\ pr}$ Black body, normal eye	0	193
	$\frac{b^+\ pr}{b\ pr}$ normal body, purple eyes	0	193
	Total	772	772

Fig. 6.2: A test cross with F_1 progeny of *Drosophila* in which ♀ is the homozygous recessive (tester) and ♂ is the doubly heterozygous parent. Complete linkage is observed.

In the second cross is *Drosophila* (see Fig. 6.3) four types of progeny are observed: two parental types (black body, purple eyes; and wild type) and two recombinant types (black body, normal eyes; and normal body, purple eyes). These are the four types of progeny one would expect, if the genes for body colour and eye colour segregate independently. However, the observed ratio of recombinant to parental types $\frac{14+16}{210+240} = \frac{30}{450}$ is very different from that expected for independent

Allele	Phenotype
b	black body (recessive)
b⁺	wild type (dominant)
pr	purple eye (recessive)
pr⁺	wild type (dominant)

Parental cross: $\frac{b\ pr}{b\ pr}$ ♀ x $\frac{b^+\ pr^+}{b^+\ pr^+}$ ♂

Test Cross: $\frac{b^+\ pr^+}{b\ pr}$ ♀ x $\frac{b\ pr}{b\ pr}$ ♂

Progeny Type	Phenotype	Observed Number	Number expected from independent assortment
Parentals	$\frac{b\ pr}{b\ pr}$ black body, purple eye	210	120
	$\frac{b^+\ pr^+}{b\ pr}$ wild type	240	120
Recombinants	$\frac{b\ pr^+}{b\ pr}$ black body, normal eye	14	120
	$\frac{b^+\ pr}{b\ pr}$ normal body, purple eyes	16	120
	Total	480	480

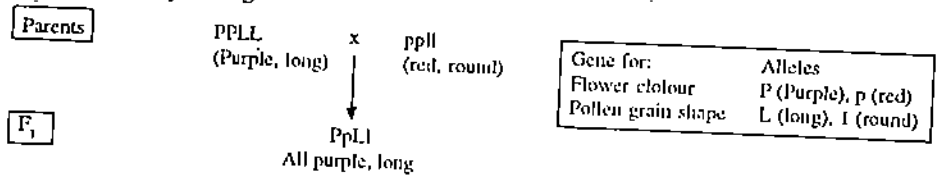
Fig. 6.3: A test cross with the F_1 progeny of *Drosophila* in which the ♂ parent is homozygous recessive and the ♀ parent is doubly heterozygous. Partial linkage is observed.

assortment ($\frac{120+120}{120+120} = \frac{240}{240}$).

Such a deviation from the ratio expected for independent assortment indicates there is linkage but is partial and not complete. The degree of linkage exhibited in such a cross (showing *partial or incomplete linkage*) is measured by the frequency of recombination. The recombination frequency in this case is 30/480 = 0.0625 or 6.25%. The unlinked genes, or the ones assorting independently exhibit a recombination frequencies of 50% (240/480 = 0.50 = 50%).

So far we have discussed the concept of linkage explaining the basic terminology involved. We would now review two pioneering works: one by Bateson, Saunders and Punnett and the other by Morgan, which demonstrated linkage experimentally.

Partial Linkage in Sweet Pea: The effects of linkage were first evident in the results of dihybrid cross in sweet peas (*Lathyrus odoratus*) that were reported by W Bateson, E. Saunders and R.C. Punnett in 1905. This cross is represented diagrammatically in Fig. 6.4.



Cross F_1 x F_1:

F_2

	Numbers obtained	Expected Numbers
Purple, long (P-L-)	4831	3915
Purple, round (P-ll)	390	1305
Red, long (ppL-)	393	1305
Red, round (ppll)	1338	435

Fig. 6.4: The Bateson, Saunders and Punnett cross. The expected numbers of the F_2 plants are calculated on the basis of the same total number of progeny (6952) and the 9:3:3:1 ratio obtained in a dihybrid cross by Mendel.

The sweet peas with purple flowers and long pollen grain (*PPLL*) were crossed with red-flowered plants with round pollen grains (*ppll*). Nothing was unusual about the F_1 progeny as all were purple and long (*PpLl*), showing these to be dominant traits. When the F_1 was inbred and each pair of alleles was examined separately, each one showed segregation like the Mendelian genes, that is, purple and red flowers were present in a 3:1 ratio, so were the long and round pollen traits.

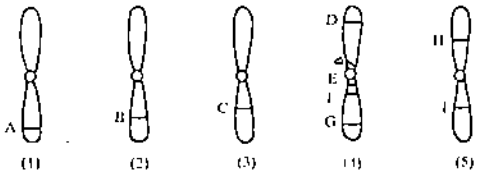
But when the traits were considered together, then partial linkage was seen. Among the 6952 F_2 plants, 4831 were purple long *(P—L—);* 390 were purple, round *(P—ll);* 393 were red, long *(ppL—);* and 1338 were red, round *(ppll)*. Compare in the Fig. 6.4 numbers of plants obtained with the expected numbers. You can see that the parental types (purple long *P—L—;* and red, round *ppll*) are present in excess of what was predicted about independently assorting genes. If these two traits had been located on different chromosomes, perhaps we would have got a 9:3:3:1 ratio of these F_2 progeny. In other words, 3915 purple, long; 1305 purple, round; 1305, red, long; and 435 red, round.

The data in Fig. 6.4 clearly shows that the two traits did not show complete linkage, had they done so a 3:1 ratio or 5214 purple, long and 1738, red, round progeny would have been expected. This clearly was a case of partial linkage. However, Bateson, Saunders and Punnett did not interpret their results in terms of behaviour of genes located on the same chromosome or linkage. Thomas Hunt Morgan was the first to relate linkage to the segregation of homologous chromosomes, and the occurrence of crossing-over between homologous chromosomes during meiosis. Morgan's interpretation of linkage was published in 1911 in a paper, where he reported the results of crosses involving linked genes in the fruit fly *Drosophila melanogaster*. Many of our current concepts of linkage, crossing-over, and chromosome mapping have evolved from the work of Morgan and his students C.B. Bridges; H.J. Muller and A.H. Sturtevant.

Morgan's Linkage Experiments with *Drosophila*: Morgan demonstrated the effects of linkage by two genes located on the second chromosome of *D. melanogaster* (see Figs 6.5 and 6.6). In both these crosses the recessive gene b in homozygous condition results in black colour of the body. The presence of its dominant allele b^+ results in

SAQ 1

Given below are five chromosomes (1-5) each having one or more genes. Observe them carefully and answer the following questions:



a) Fill in the blanks:
   i) The genes A,B, and C are ..... genes.
   ii) The genes D, E, F and G are ..... genes.
   iii) The genes H and I too are ..... genes.

b) How many linkage groups are seen in the above figure? Write their number in the box.

c) Write the linkage groups that are shown in the figure.

.....

.....

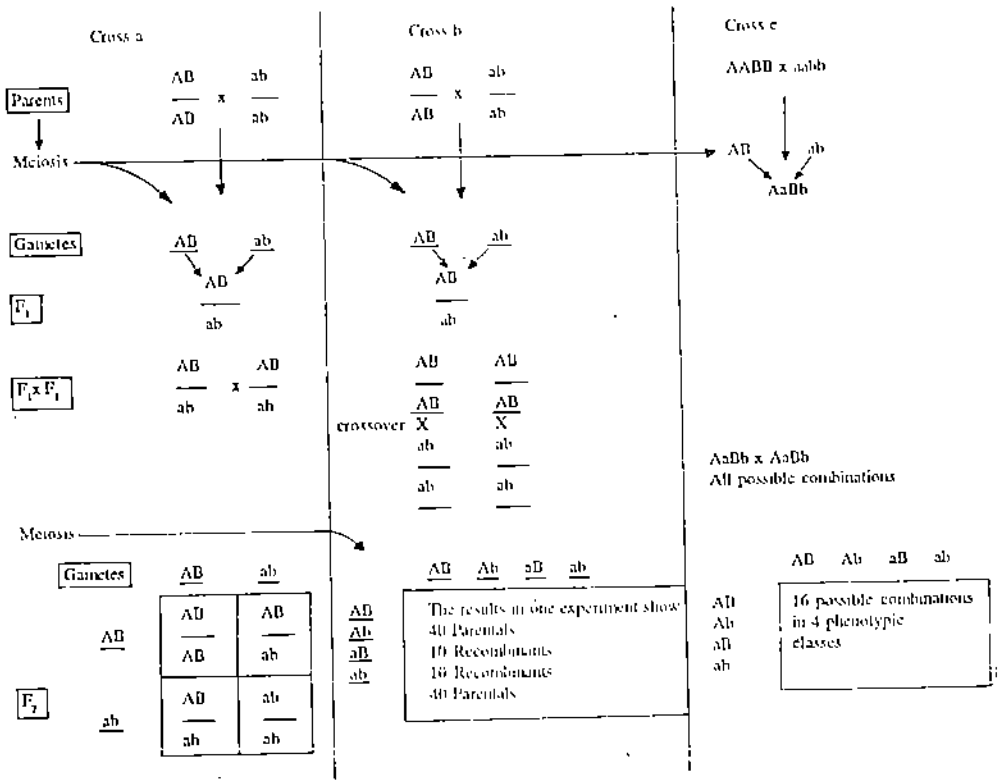
d) Write a statement that explains the situation in 1 to 5.

.....

.....

SAQ 2

Given below are three crosses – a, b and c. Study them carefully and answer the following questions.



a) **Cross a**
   i) The genes in meiotic gametes are ..... .
   ii) The phenotypic ratio of F_2 is ..... .
   iii) The genotypic ratio of F_2 is ..... .
   iv) This cross shows..... .

b) **Cross b**
   i) The genes in meiotic gametes are ..... .
   ii) F_1 inbreeding shows a special feature that is ..... .
   iii) The F_2 genotypes constitute basically ..... classes of individuals that are ..... and ..... .
   iv) This cross shows ..... .

c) **Cross c**
   i) The genes in meiotic gametes are ..... .
   ii) The F_1 progeny is ..... .
   iii) The phenotypic ratio of F_2 progeny is ..... .
   iv) This cross shows ..... .

SAQ 3

Which one of the following arrays represents the cis-configuration of the alleles A and B?

$\frac{ab}{AB}$ or $\frac{aB}{Ab}$

.....

6.3 CROSSING-OVER

Crossing-over is a physical exchange between chromatids in a pair of homologous chromosomes. It results in a new association of genes in the same chromosome. The role of crossing-over is important for evolution to take place. In fact, crossing-over and independent assortment are mechanisms that produce new combinations of genes. Natural selection can then act to preserve those combinations that produce organisms with maximum fitness, that is, maximum probability of perpetuation of the genotype.

6.3.1 The Concept of Crossing-Over

Following are the important features of crossing-over:

i) A gene is located on a chromosome at a particular site called a locus (plural - loci). The loci of the genes on a chromosome are arranged in a linear sequence.

ii) In a heterozygote, the two alleles of a gene occupy corresponding positions in the homologous chromosomes, that is, allele *A* occupies the same position in homolog 1 that allele *a* occupies in homolog 2 (see Fig. 6.8). The position of a gene in a given chromosome of a species is fixed or constant.

iii) Crossing-over involves the breakage and rejoining of two chromatids (of homologous chromosomes), resulting in reciprocal exchange of equal and corresponding segment between them (Fig. 6.8).

iv) Chromosomes with recombined or new combinations of genes are formed by the occurrence of crossing-over.

v) Crossing-over occurs more or less at random along the length of a chromosome pair. Thus, the probability of its occurrence between two genes increases with increasing physical separation of the genes along the chromosome.

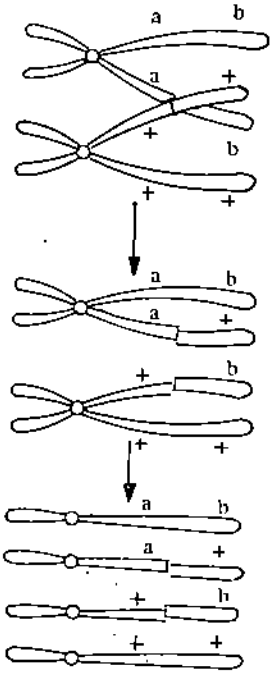


Fig. 6.8: Diagram illustrating the occurrence of crossing-over between two chromatids of homologous chromosomes. The exchange between two of the four chromatids results in two recombinants and two parental combinations at meiosis.

6.3.2 When does Crossing-Over Occur?

Crossing-over begins at pachytene stage, after the synapsis of the homologous chromosomes has occurred in zygotene stage of prophase-I of meiosis. Since chromosome replication occurs during interphase, meiotic crossing-over occurs in the post-replication four strand or tetrad stage. That is, after each chromosome has doubled such that four chromatids are present for each pair of homologous chromosomes.

6.3.3 Cytological Basis of Crossing-Over

Based on his results (Fig. 6.5 and 6.6), Morgan suggested that recombinations are formed as a result of pairing of homologous chromosomes during meiosis. A physical

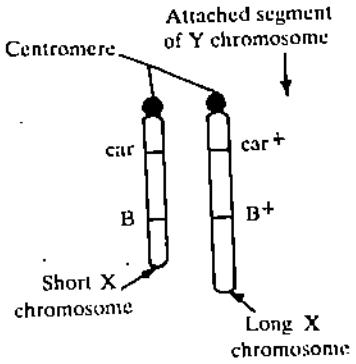


Fig. 6.9: Outline diagram showing the special features of the X chromosome, that were used in Stern's experiments.

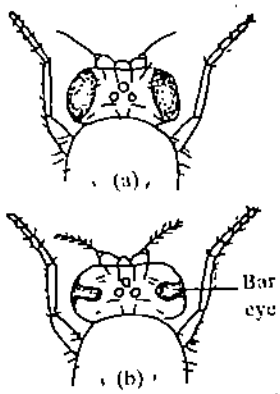


Fig. 6.10: Two phenotypes of eye shape in *Drosophila*.
a) normal, round eyes.
b) bar-shaped eye.

exchange of chromosome parts takes place by a process called crossing-over. In the germs cells of many organisms at the time of meiosis, one can actually see certain cross-shaped structures in which two of the four chromatids of homologous chromosome pairs appear to exchange parts (see Fig. 6.8). These cross-shaped structures are called *Chiasmata* (singular *Chiasma*).

The first direct evidence that relates the occurrence of physical exchange between chromosomes to the genetic recombination was provided by C. Stern (1931) in *Drosophila*, and by H.B.Creighton and B.McClintock in maize. Normally, the two chromosomes constituting a homologous pairs are morphologically, alike and are therefore, indistinguishable. But Stern, Creighton and McClintock, however, found 'homologs' that were not alike, and could be easily distinguished. Because at their ends they carried distinct morphological features that made them easily observable (see Figs. 6.9 and 6.12). These chromosomes are homologous along most of their length but differed in a small portion only. They, however, paired and segregated normally during meiosis. These structural alterations of the chromosomes permitted the microscopic recognition of parental and recombinant chromosomes. We shall now briefly discuss the experimental work carried out by Stern (in *Drosophila*) and McClintock (in maize using these genetic markers.

Stern discovered two X chromosomes in a *Drosophila* female strain that had undergone structural changes, that made them distinguishable from each other and from a normal X chromosome. One X chromosome had part of a Y chromosome attached to one end. The second X chromosome was shorter than normal, as its piece had broken off, and was translocated to one of the small, fourth chromosome (Fig. 6.9). These female flies were heterozygous for alleles of two genes located on the X chromosome. One gene affects eye shape, and other eye colour. The recessive allele *car* of the first gene results in carnation coloured eyes; its dominant allele car^+ yield red eyes. The dominant mutant allele *B* of the second gene results in bar-shaped eyes (Fig. 6.10) and the recessive allele B^+ produces round eyes (Fig. 6.10b) when homozygous. The females used in Stern's study carried the allelic pairs in **Cis-configuration** (see Fig. 6.9 and 6.11). Stern crossed such heterozygous females with males having carnation-coloured, round eyes, i.e., *car* B^+ males. The cross and results Stern obtained are shown in Fig. 6.11. Cross a shows a situation where

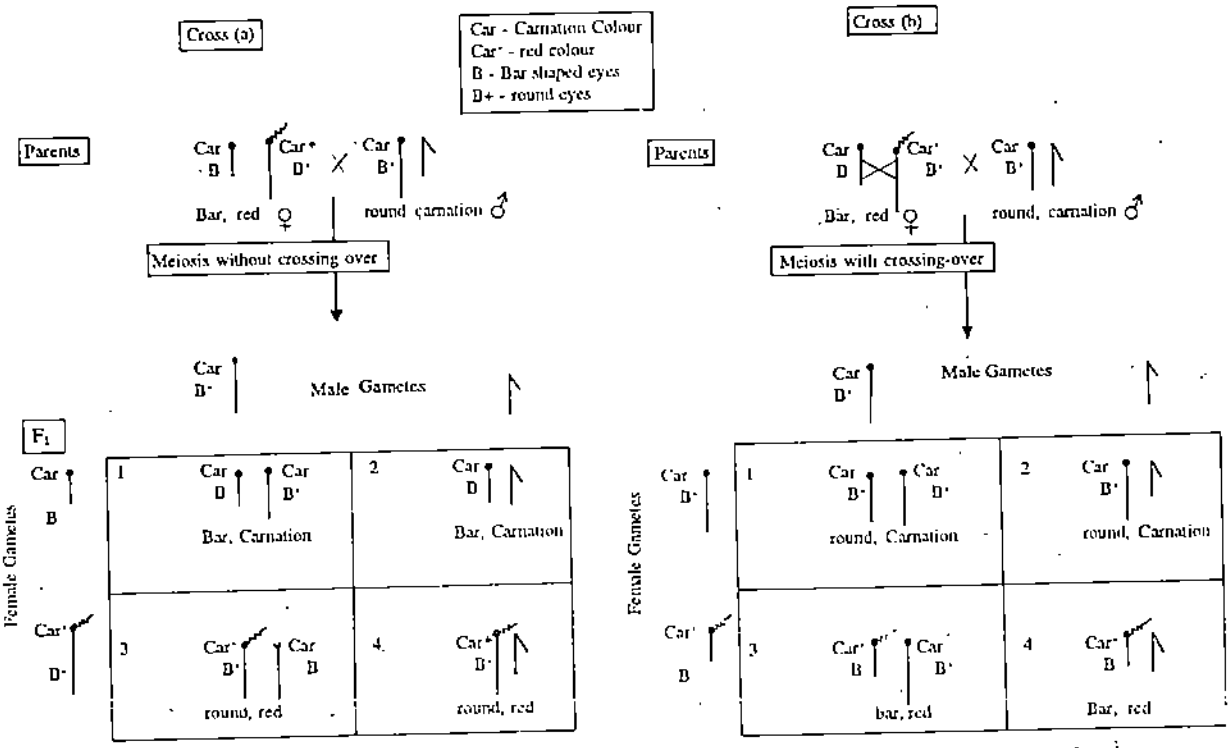


Fig. 6.11: The classic experiment in linkage, by Stern demonstrating that cross-over involves interchange of parts of homologous chromosomes. The cross b indicates the occurrence of crossing-over between the *car* and *B* loci. The knob-like structures on the chromosomes represent the centromere. You can see in the figure that the X chromosomes in the femal parent are morphologically distinguishable. The short one on the left is missing the distal end, its homolog (on the right) has an extra piece of Y chromosome attached to the centromere end, which is shown here as a horizontal zig-zagged line.

meiosis without crossing-over occurs during the gamete formation. And in cross b, there is meiosis with crossing-over between the *car* and *B* loci, during gamete formation in the female parent. The male *Drosophila* is unique as no crossing-over takes place. Now, look at the figure carefully.

Stern analysed the genotypes of the progeny to see whether there were only recombinations (see Cross b). Not only did he find the recombinant gene types but also found that the alleles on the X chromosomes of each recombinant progeny were precisely the same as predicted if crossing-over had occurred. For example, the recombinant male progeny with the bar-shaped red eyes ($car^{+}B/Y$) were found to carry the short X chromosomes, but now with the translocated piece of the Y chromosome at one end (see Cross b, 4). The male flies with normal-shaped, carnation eyes ($car\ B^{+}/Y$) contained long X chromosomes, but without the attached piece of the Y chromosome (see Cross b, 2).

You should note that the male flies were analysed and classified as a particular type on the basis of their phenotypes – the new combinations of the two phenotypic characteristics. The progeny had parental combinations of these traits. And also the progeny flies had X chromosomes that were produced by a cross-over event between the *car* and *B* loci.

Creighton and McClintock also experimentally demonstrated that there was a correlation between crossing-over and exchange of homologous chromosomes. They made crosses involving two loci on chromosome 9 and analysed their results. On this chromosome one gene controls aleurone colour (C - coloured; c - colourless) and the other controlled the type of carbohydrate in the endosperm (Wx-starchy; wx-waxy). They made use of a chromosome with densely staining knob at one end and a translocated piece from chromosome 8, at the other end (see Fig. 6.12). These two features made the chromosome visually identifiable.

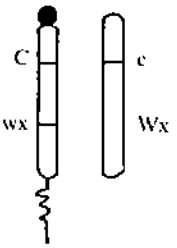


Fig. 6.12: Chromosome 9 of maize, showing a knob at one end and a translocated piece at the distal end. Note, the genes are arranged in trans-configuration.

They made a cross between heterozygous female – coloured, starchy and a male – colourless, starchy, and examined their progeny. The entire cross is illustrated in Fig. 6.13. Note the arrangement of alleles and cytological markers in the unique parent (Fig. 6.12). If cross-over occurs in this parent then the arrangement of alleles changes (see Fig. 6.14). Now examine the phenotypes of the recombinant (cross-over) offspring (cross b). This shows a new phenotype – colourless, waxy. Creighton and McClintock on observing this new phenotype found a rearrangement of the cytological markers.

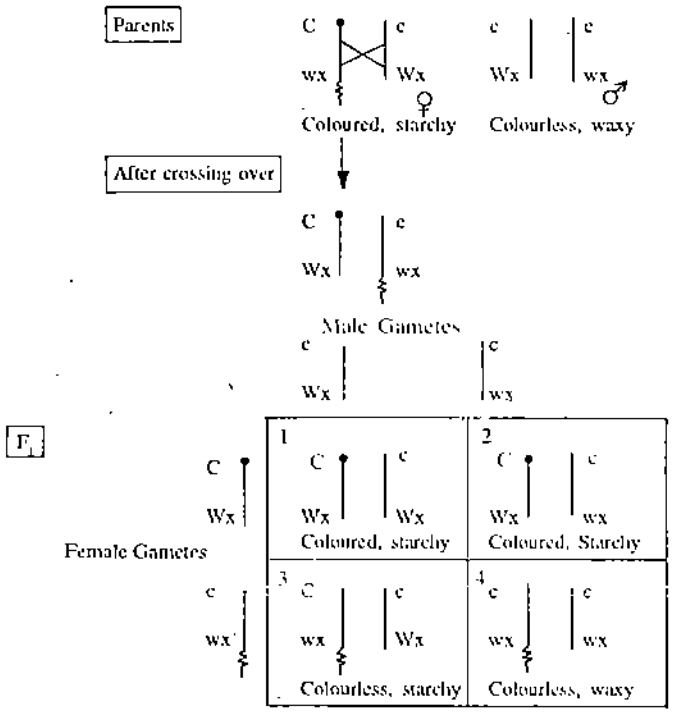


Fig. 6.13: Cross between heterozygote female and male with colourless aleurone, starchy endosperm in maize.

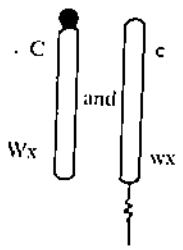


Fig. 6.14: Chromosomes of maize, after crossing over between the *C* and *Wx* loci. Now the genes are arranged in cis-configuration from the earlier trans-configuration.

Both the above experiments provide cytological proof of crossing-over and leave no doubt that an actual physical exchange between homologs occurs during crossing-over.

These two experiments by Stern and by Creighton and McClintock are classics in genetic studies. They confirmed Morgan's hypothesis that crossing-over involves the interchange of parts of homologous chromosomes. They also provided strong evidence that genes are located on chromosomes.

6.3.4 Molecular Mechanism of Crossing-Over

The classic experiments of Stern, and Creighton and McClintock confirmed the correlation between the formation of recombinants and crossing-over, but still several questions remained to be answered one of them was the nature of crossing-over.

A cross-over as you have seen above (Fig. 6.8) involves a reciprocal, physical exchange between homologous chromosomes. This suggests that a reciprocal exchange essentially occurs between the double helices of the DNA molecules found in each non-sister chromatids. For this process to take place, two homologus chromosomes come close to one another. In eukaryotes, crossing-over has been associated with the formation of a structure (or set of structures) called the synaptinemal complex (Fig. 6.15) which forms during prophase of the first meiotic division. This structure is composed primarily of proteins and RNA, and has been identified in a large number of eukaryotic species. Very little information is available

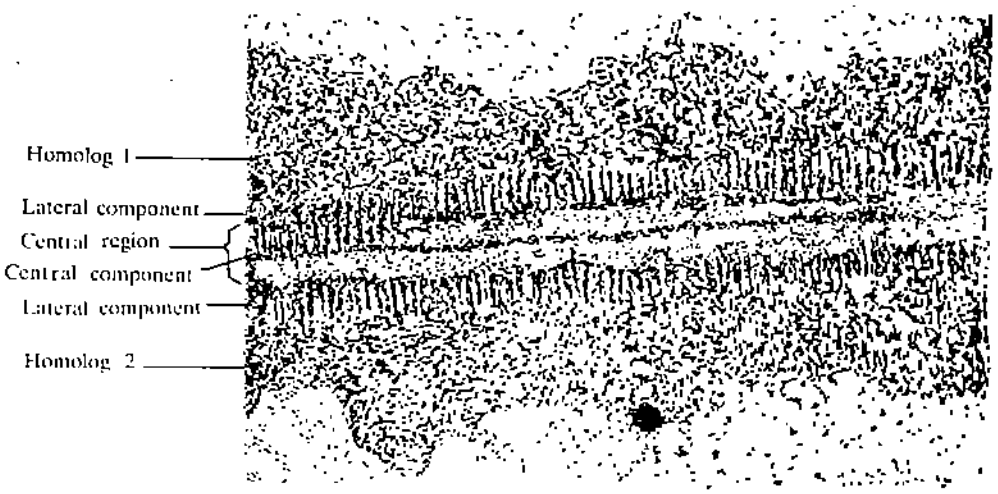


Fig. 6.15: Diagrammatic representation of the synaptinemal complex in *Neottiella rutilans* (Ascomycetes). A dense, central component is surrounded by a less dense space: together these make up the central region which is about 1000Å in width. The lateral components each about 500 Å in width are present on each side of the central region. The two homologous chromosomes are juxtaposed to the two lateral elements. (Redrawn from: M. Westergaard and D. Von Wett Stein, 1972 Ann. Rev. Genet. 6 : 71-110.

about the functions of the various components of the synaptinemal complex. It is known that small amount of DNA synthesis occurs at the time when the synaptinemal complex forms. This amount, however, is very small and is equivalent or even less than 1 per cent of the total DNA in the genome. It is believed that this DNA synthesis is involved in synapsis and/or crossing-over.

To explain the mechanism of crossing-over at molecular level, a number of models based on breakage and reunion of a DNA molecule have been proposed. The breakage of two homologous chromatids and the reunion of the parts in new arrangements (recombination) is the main feature of these models. One model for eukaryotic DNA recombination that survives was proposed by Robin Holliday in 1964. The important features of this model are illustrated in Fig. 6.16. This pathway begins when an *endonuclease* cleaves single strands of each of the two parental DNA molecules (*breakage*). Segments of the single strands on one side of each cut are displaced from their complementary strands. In this process *DNA-binding proteins, helix destabilisation proteins,* and *DNA unwinding proteins* or *DNA helicases* are also said to be involved.

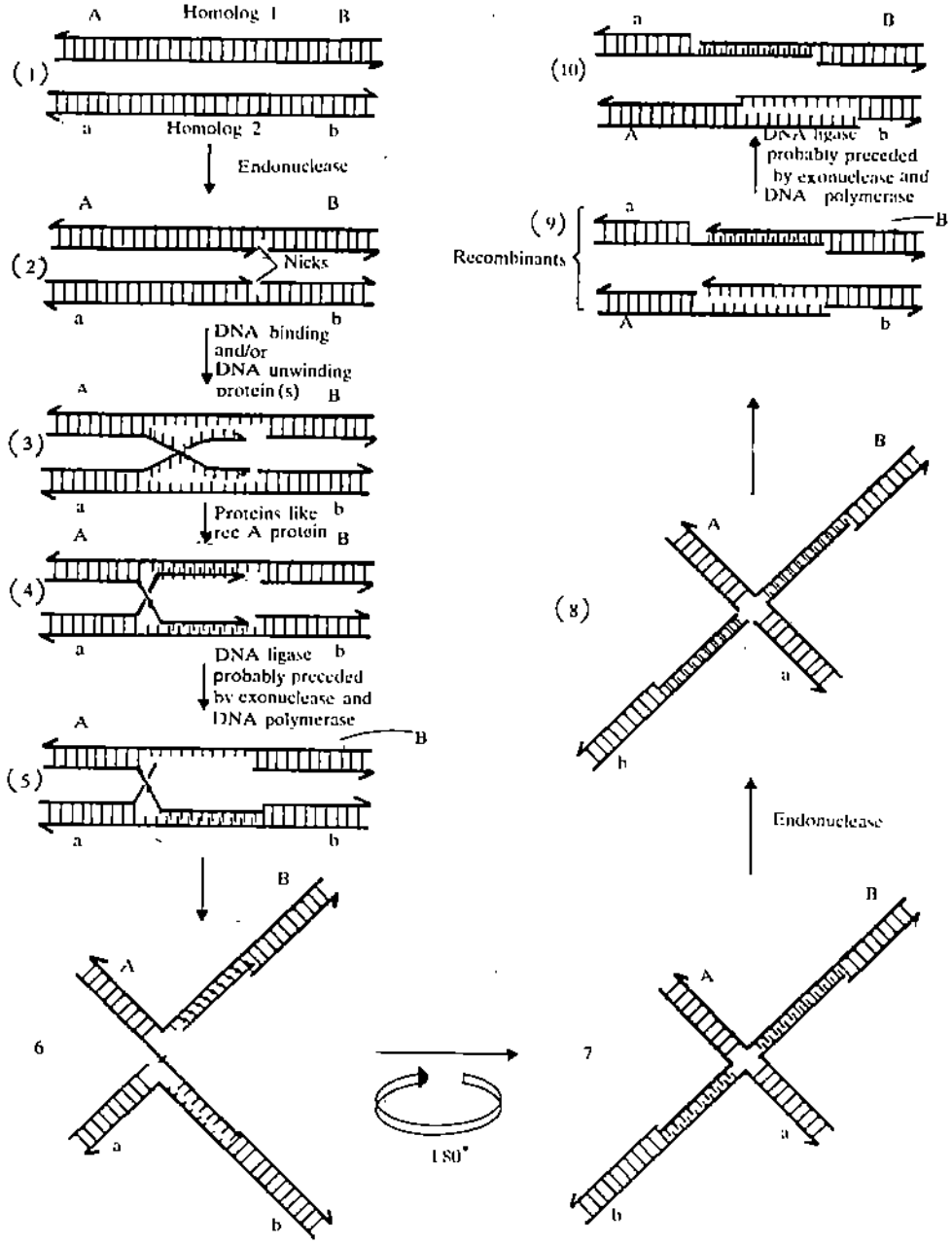


Fig.6.16: Pathway for the occurrence of crossing-over by breakage and reunion based on the model of Robin Holliday. 1) shows two homologs 1 and 2. homolog 1 carrying *A* and *B* genes and homolog 2 carrying *a* and *b* genes. 2) Initiation of cuts on single strands due to endonuclease 3) displacement of one side of the free ends of the cuts, with the help of DNA binding, helix-destabilisation, and/or DNA-unwinding proteins. 4) The cleaved single strand of homolog 1 base-pairs with the complementary intact strand of homolog 2 and vice versa. 5) The cleaved strands are then rejoined in recombinant combinations by single stranded bridge, by the enzyme DNA ligase. 6) and 7) X-shaped intermediates also known as Holliday intermediate or chi forms [after the Greek letter chi (x)] in different planar views. 8) The intact strands are cleaved at the intersection by endonuclease, and stage 9) is formed. Covalent closure of single stranded interruptions yields the intact recombinant chromosomes as shown in (10).

The displaced strands then exchange pairing partners, base-pairing with the intact complementary strands of the homologous chromosomes. This process is also aided by certain proteins such as **rec A protein.** The rec A protein mediates such a reaction by binding to the unpaired strand of DNA, searching for a homologous DNA sequence. A homologous double helix is found, promoting the displacement of a segment of one strand of the double helix by the unpaired strand.

The cleaved strands are then covalently joined in recombinant arrangements (*reunion*) by the enzyme DNA ligase. If the original breaks in the two strands do not occur at exactly the same site in the two homologs some "tailoring" will be required before DNA ligase catalyses the reunion process. This tailoring involves excision of a limited number of bases by an *exonuclease* and repair synthesis by a *DNA polymerase*. By this time an X-shaped recombination intermediate called a *"chi" form* or *Holliday intermediate* is formed. Such intermediates have been observed by electron microscopy (see Fig. 6.17) in several prokaryotic systems. A similar sequence of enzyme catalysed breakage and reunion events, involving the other two single strands, occurs to complete the process of crossing-over.

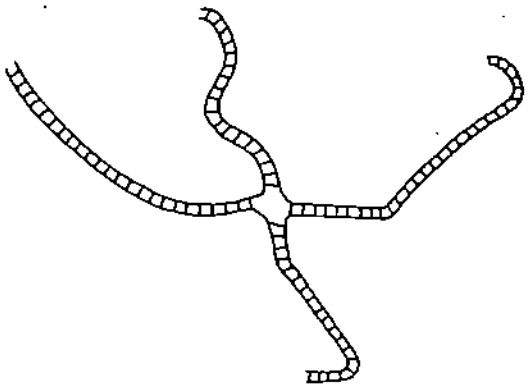


Fig. 6.17: Outline diagram of t. e x-shaped recombinant intermediate as seen under the electron microscope.

A current model that accounts for the formation of parentals as well as recombinants is shown in Fig. 6.18. It is based on the works in several laboratories including those of Meselson and Radding (1975); Potter and Dressler (1976); and Das Gupta et al. (1981). Up to stage 7, that is,till the formation of "chi form" or Holliday intermediate, it is the same as the Holliday model described above. The difference is from stage 8 onwards.

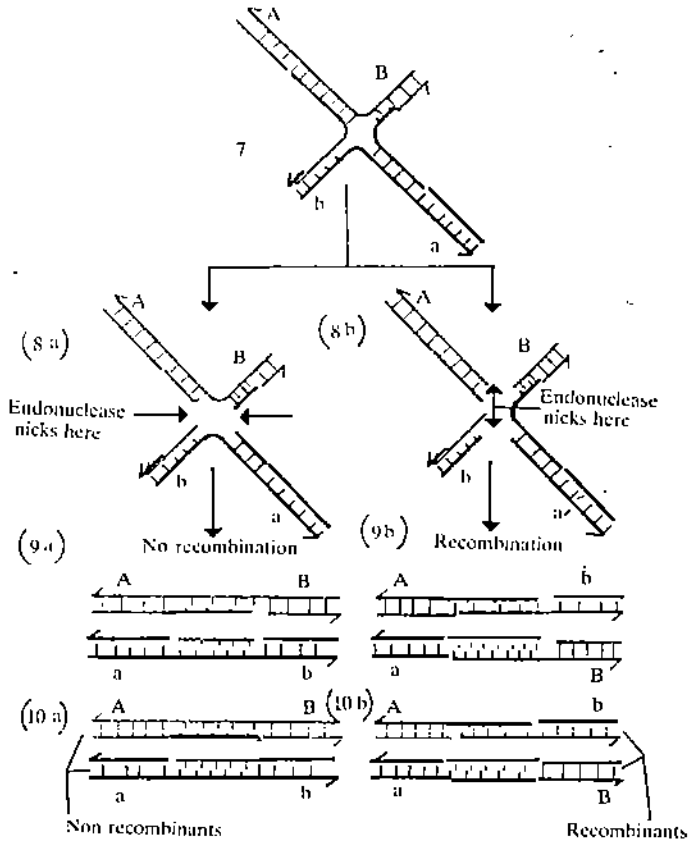


Fig. 6.18: Uptil stage (7) it is the same as Fig. 6.12 (Robin Holliday's model). In (8) there are two possibilities, first, endonuclease acts on sides (see 8a) resulting in nonrecombinants or parentals (9a and 10a). Second, if endonuclease nicks the opposite end then recombinants (9b and 10b) are formed.

The joined homologs are separated from each other by endonuclease attack on either pair of opposite strands (see 8a, b). Following the endonuclease cleavages, the strands separate. Each strand contains a small gap in one strand of DNA, that is, closed by the action of the ligase. Which pair of strands is cleaved determines whether the molecules are recombinants or not. If the two strands are cleaved on the sides then the parentals — *AB* and *ab* are formed (9a, 10a). And if the cleavage occurs on top and bottom, then recombinants — *Ab* and *aB* (9b, 10b) are formed.

Along with the modifications of the above model, evidences have accumulated that indicate homologous recombination occurs by more than one mechanism — very possibly by several different mechanisms. A double-strand break model was proposed by Szoztak et al. in (1983). The major difference between this model and the Holliday model is that, recombination is mediated by a double-strand break in one of the parental double helices, not by just single-strand breaks (see Fig. 6.19). These initial breaks in the two strands are enlarged further and gaps are formed in both the strands. The two single-stranded termini produced at the double stranded gap of the broken double helix invade the intact double helix, displacing the homologous strand in this region. Repair DNA synthesis then takes place filling the gaps using complementary strands of the other chromosome as templates. This process yields a recombination intermediate with the two double helices joined by two single-stranded bridges (chi structures). The bridges are resolved by endonuclease cleavages by the same process as for chi structures formed by the Holliday mechanism. This model, like the Holliday model, explains the production of chromosomes that are recombinant for genetic markers flanking the region in which cross-over occurs.

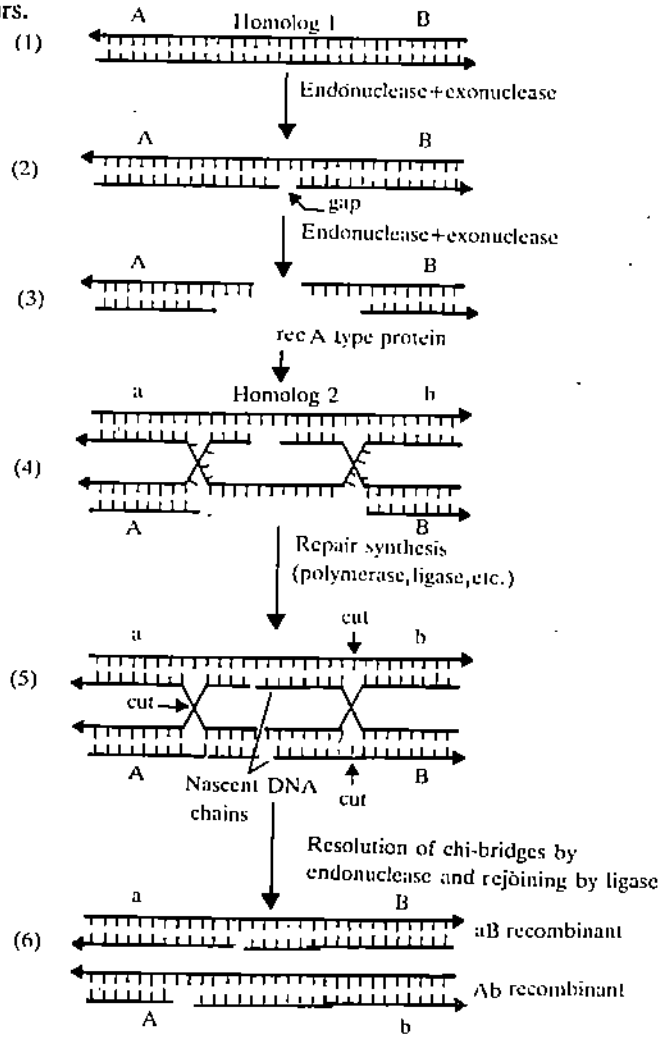


Fig. 6.19: A double-strand-break model of recombination (1-3) Gaps are produced in two strands of DNA of one chromosome (homolog 1) by sequential endonuclease and exonuclease activities. (4) the free single-stranded ends formed at the gaps invade the second DNA molecule (homolog 2) and with the help of rec A type proteins, and helicase displace the homologous strand of the intact double helix. (5) the gaps are filled by repair synthesis of DNA using the complementary strands of the intact, but locally unwound, double helix as templates. Ligation produces two double helices joined at two sites by single-stranded bridges (chi structures). (6) the bridges are resolved by the action of endonucleases, with subsequent rejoining catalysed by DNA ligase, and the recombinant chromosomes are produced (From Gardner et al. 1991 John Wiley and Sons Inc. New York.

SAQ 4

a) When does chiasmata formation and crossing-over occur?

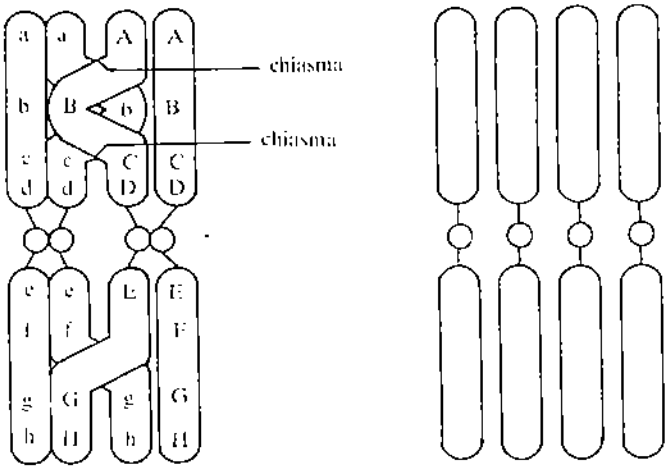
.....

b) Between which structure does exchange of genes occur?

.....

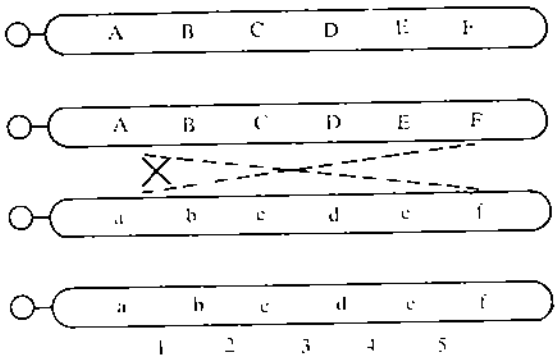
SAQ 5

The figure given below shows chiasma formation. The two chromosomes involved have the genes *a, b, ...h,* and *A,B,..... H.* Fill in the blank spaces as to what genes the resultant separated chromosomes have, and shade the areas where genetic exchange has taken place.



SAQ 6

In the cross-over given below, is there a greater chance of genes *A* and *B* being separated or genes *A* and *F*? Explain your answer.



.....

.....

.....

6.4 GENETIC MAPPING

The linkage of the genes in a chromosome can be represented in the form of a **genetic map.** It shows the linear order of the genes along the chromosome with the distances between adjacent genes proportional to the frequency of recombination between them. A genetic map is also called a **linkage map** or a **chromosome map.**

The unit of distance in a genetic map is called a **map unit (mu).** One map unit is equal to 1 per cent recombination or one centimorgan (cM), named in honour of T.H. Morgan. For example, two genes that recombine with a frequency of 4.5 per cent are said to be located 4.5 map units apart.

Alfred H. Sturtevant, a student of T.H. Morgan, was the first one to construct a genetic map based on his data on recombination frequencies that shows the location

of various genes on a chromosome. Sturtevant also pointed out that the map is linear or one dimensional. The first genetic map of the X chromosome of *Drosophila* constructed from his data in Table 6.1 is shown in Fig. 6.20.

Table 6.1: Recombination Frequencies for Some Sex-Linked Mutations in *Drosophila melanogaster*.

Genes	Recombination Frequency
yellow (y) and white (w)	0.010
yellow (y) and vermilion (v)	0.322
yellow (y) and miniature (m)	0.355
vermilion (v) and miniature (m)	0.030
white (w) and vermilion (v)	0.300
white (w) and miniature (m)	0.327
white (w) and rudimentary (r)	0.450
vermilion (v) and rudimentary (r)	0.269

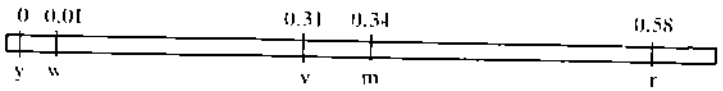


Fig. 6.20: The first genetic map of the X-chromosome of *Drosophila melanogaster*, showing the relative map positions of the genes *yellow*, *white*, *vermilion*, *miniature* and *rudimentary*. The *yellow* gene was arbitrarily chosen as zero on the genetic map.

Sturtevant reasoned that if genes were distributed along a chromosome, then the farther apart two genes lay the greater the likelihood that a cross-over would occur between them. Thus, the frequency of recombination between widely separated genes should be greater than that between closely located genes.

Let's have a look at the data in Table 6.1 once again. The *yellow* and *white* mutations exhibit a recombination frequency of 0.010 and therefore, must be fairly close to each other compared to *yellow* and *vermilion*, which exhibit a recombination frequency of 0.322. Moreover, *white* and *vermilion* show a recombination frequency of 0.300, indicating that *white* must be closer to *vermilion* than is *yellow*. Therefore, the data indicates that *white* must lie between *yellow* and *vermilion*. Similarly, it has been found out that the *vermilion* is located quite close to *miniature* (recombination frequency 0.030) and *vermilion* is more closely linked to *white* than is *miniature* (compare 0.300 to 0.327); therefore, *rudimentary* must be between *white* and *miniature*. The *rudimentary* appears to be almost unliked to *white*, as shown by its recombination frequency 0.450. This approaches the value of 0.50 expected for independent assortment. The location of *rudimentary* which might be on the left or the right of *white* on the map, was based on the observation that *vermilion* is closely linked to *rudimentary* than is *white* and therefore, must lie between the two.

The case of *white* and *rudimentary* (recombination frequency of 0.450) illustrates the point that two genes may reside on the same chromosome and yet assort independently, if they are very distantly located. For example, if a gene is located some distance right of *rudimentary* it would then assort independently of *white*. Such a gene would have a recombination frequency about 0.50 and yet could be shown to be on the same chromosome as *white* through common linkage to *rudimentary*.

In Fig. 6.20, the distance between adjacent genes is based on the observed recombination frequency. And you know that recombination frequency is defined as one map unit, or one centimorgan (cM). The location of each gene is designated by the map distance from one end of the set of linked genes. The *yellow* gene is arbitrarily chosen as the left end of the map with a position of 0.0, and the positions of the other genes are assigned by summing recombination frequencies between the nearest genes. Thus, *yellow* and *rudimentary* are separated by 58 cM on the map of the X chromosome.

6.4.1 Three-Point Crosses

Sturtevant presented his proof of the linearity of genetic map based on the analysis of data from crosses in which alleles of three different X chromosome genes were segregating – the three point crosses. We shall continue using the same example of the X-linked genes in *Drosophila* as above to describe these crosses.

You already know, that a test cross is very useful in studying linked genes, and it helped immensely in our understanding of linkage. Sturtevant test crossed a female

heterozygous for the mutations *yellow* body, *white* eye, *miniature* wings with a *ywm* male (see Fig. 6.21).

Test cross: $\frac{+\,+\,+}{y\ w\ m}\ ♀ \times \frac{y\ w\ m}{Y}\ ♂$

Progeny: Class	Genotype of maternal X chromosome present in sons or daughters	Observed Numbers
Parental	$\frac{+\,+\,+}{y\ w\ m}$	6972
m Recombined with y and w	$\frac{+\,+\,m}{y\ w\ +}$	3454
y Recombined with w and m	$\frac{y\ +\ +}{+\ w\ m}$	60
w Recombined with y and m	$\frac{y\ +\ m}{+\ w\ +}$	9
		10,495

Recombination frequency between:

w and m $\quad \frac{3454 + 9}{10,495} = 0.330$

y and w $\quad \frac{60 + 9}{10,495} = 0.007$

y and m $\quad \frac{3454 + 60}{10,495} = 0.335$

Fig. 6.21: A test cross involving three loci on the X chromosome of *Drosophila melanogaster*.

Consider the types of eggs produced by such a female. In addition to the parental combination of alleles, six other combinations are expected to occur because for each gene either of the two alleles may be present on one X chromosome. The total number of possible combinations is therefore, $2 \times 2 \times 2 = 8$ (also Fig. 6.21). Since these genes are linked, the different combinations do not occur with equal frequencies but are dictated by the parental association of alleles and by the frequency with which crossing-over occurs between each pair of genes. These recombination frequencies are shown in Fig. 6.21. If these genes are arranged in a linear order on the chromosome, then three different arrangements are possible (see Fig. 6.22).

You should note that the reciprocal recombinant types shown in the Fig. 6.18 cannot arise from the two cross-overs occurring during meiosis. In other words, if three genes are present linearly, the recombinant types cannot be produced independently of one another.

Only the other order I showing the *y-w-m* gene arrangement is consistent with the data in Fig. 6.21. The observed recombinantion frequency between *y* and *w* is 0.007 and that between *w* and *m* is 0.330. Therefore, the frequency of double cross-over recombinant class should be approximately $0.007 \times 0.330 = 0.00231$.

In Fig. 6.21, you can see that the most infrequent recombinant class is *w* recombined with *y* and *m*. Its recombination frequency being $9/10,495 = 0.00086$ and therefore, this is the class produced by double cross-over. Sturtevant showed that this type of relationship existed for alleles of any three genes on the X chromosome and only a linear map could account for the recombination data from many different crosses involving different sets of three genes.

Sturtevant's discovery of linear arrangement of genes on the chromosome is accorded an importance second only to Mendel's discovery of the gene. The linear nature of the chromosome provided a framework for all future work in genetics and presaged the discovery of the linear nature of the DNA molecule.

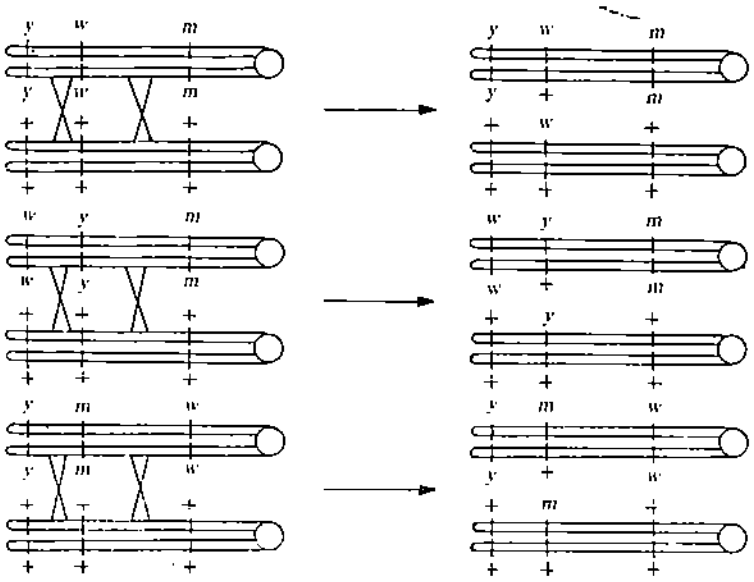


Fig. 6.22: The three possible orders of genes y and m with gametes produced from double recombination. The order of genes in I is with the gene w in the center is consistent with the data in Fig. 6.18.

Thus, *the three-point crosses provide information regarding both the order of the genes involved, and the recombination frequencies among them.* This type of analysis, first done by Sturtevant, forms the basis of all genetic mapping. The genetic map of the *Drosophila melanogaster* genome established by Sturtevant and other students in Morgan's laboratory is shown in Fig. 6.23.

6.4.2 Interference and Coincidence

The detection of double crossing-over has made it possible to determine whether exchanges in two different regions of a pair of chromosomes occur independently of each other. Using the information from the above example of *Drosophila* (data given in Fig. 6.21), we know from the recombination frequencies that the probability of exchange is 0.330 between *w* and *m*, and 0.007 between *y* and *w*. If crossing-over occurs independently in the two regions (that is, if the occurrence of one exchange does not alter the probability of the second exchange), the probability of the occurrence of an exchange in both regions is the product of these individual probabilities, that is, (0.330 × 0.007 = 0.00231) or 0.231 per cent. This means that in a sample of 10,495 the **expected** number of double cross-overs would be 10,495 × 0.00231 = 24.2 or 24. Whereas the number observed was only 9. Such deficiencies in the observed number of double cross-overs are common and are said to be due to **Interference** (*I*). That is, the occurrence of crossing-over in one region of a chromosome reduces the probability of a second cross-over in a nearby region.

The **coefficient of coincidence** (c) is the observed number of double recombinants divided by the expected number; its value provides a simple measure of the degree of interference, which is measured as:
Interference = I − coefficient of coincidence or I = I − c

From the data in our example, the coefficient of coincidence is 9/24 = 0.375, meaning that the number of double cross-overs that occurred was only 37 per cent of the number expected if crossing-over in the two regions were independent. It has been found experimentally that interference usually increases as the distance between the two markers becomes smaller, until a point is reached at which double crossing-over does not occur; that is, no double cross-overs are found and the coefficient of coincidence equals 0 (or the interference equals 1). The distance is about 10 map units in a variety of organisms. Conversely, when the total distance between the gene loci is greater than about 45 map units, interference disappears and the coefficient of coincidence becomes 1.

6.4.3 Linkage Analysis in Haploid Eukaryotes

Many of the single-celled eukaryotes are haploid during the vegetative stages in their life cycle. The alga *Chlamydomonas* and the mold *Neurospora* illustrate this

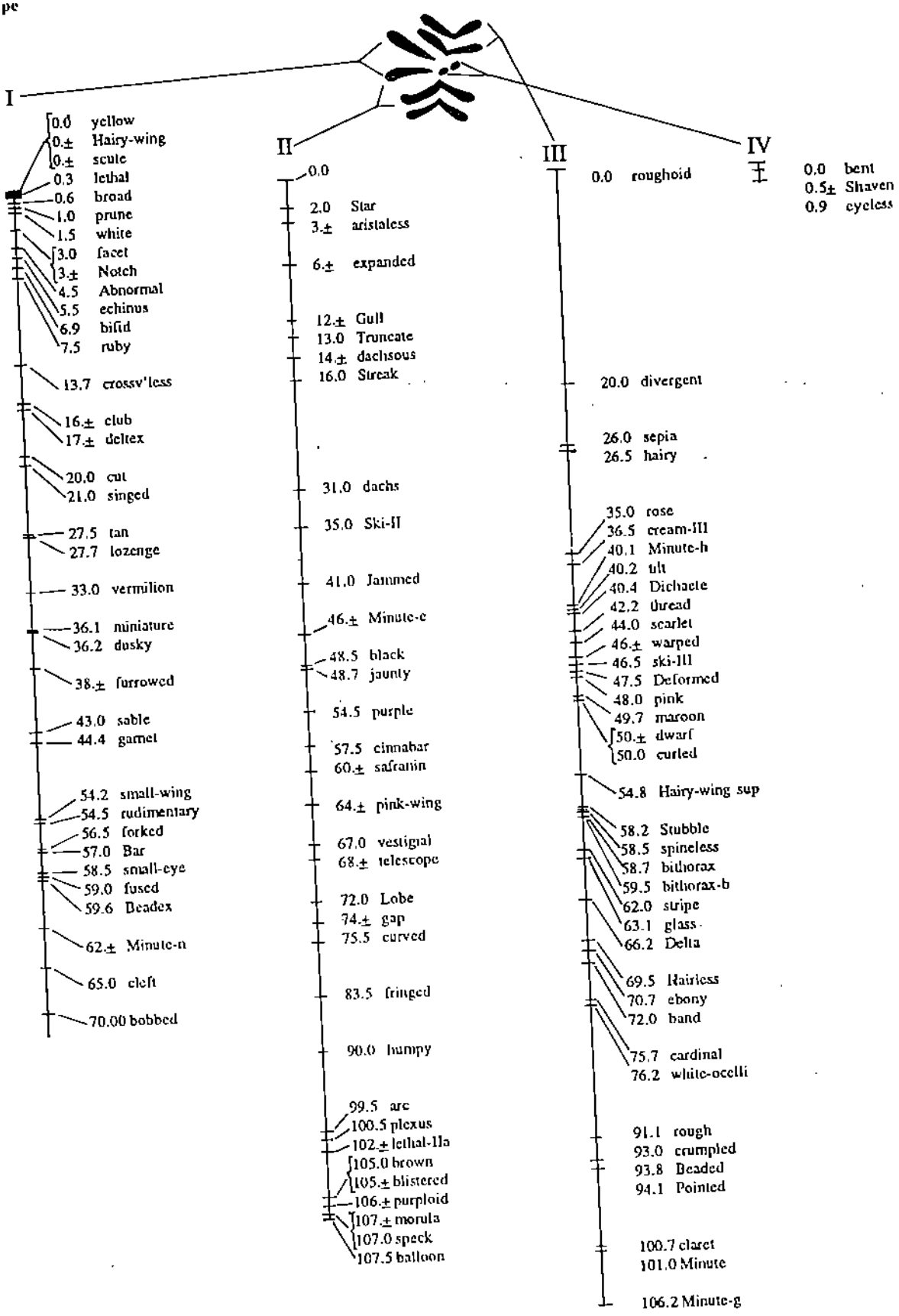


Fig. 6.23: An abbreviated genetic map of *Drosophila melanogaster*, showing the correspondence between each of the four linkage groups and a pair of chromosomes. The map positions of the genes in a chromosome are in map units from the gene closest to one end of the chromosome. Only a few of the many genes that have been mapped are shown. Capitalised symbols indicate a dominant mutation.

condition. These organisms form reproductive cells that fuse during fertilisation producing a diploid zygote. However, this structure soon undergoes meiosis, resulting in haploid vegetative cells which are then propagated by mitotic divisions.

Haploid organisms constitute favourite materials of geneticists. This is because these organisms have several important advantages in genetic studies compared with the complex diploid organisms. These haploids can be cultured and manipulated in genetic crosses much more easily. Above all, a haploid organism contains only a single allele of each gene which is expressed directly in the phenotype. This characteristic of haploids, greatly simplifies genetic analysis. Due to the presence of above features, the organisms such as *Chlamydomonas* and *Neurospora* have served as subjects of research investigations in linkage and mapping studies besides many other areas in genetics.

Tetrad Analysis of *Chlamydomonas*: It is one of the first haploid organisms used in genetic studies. A diagrammatic sketch of its life cycle is shown in Fig. 6.24. As you can see in the figure, the four haploid cells produced by meiosis are present together

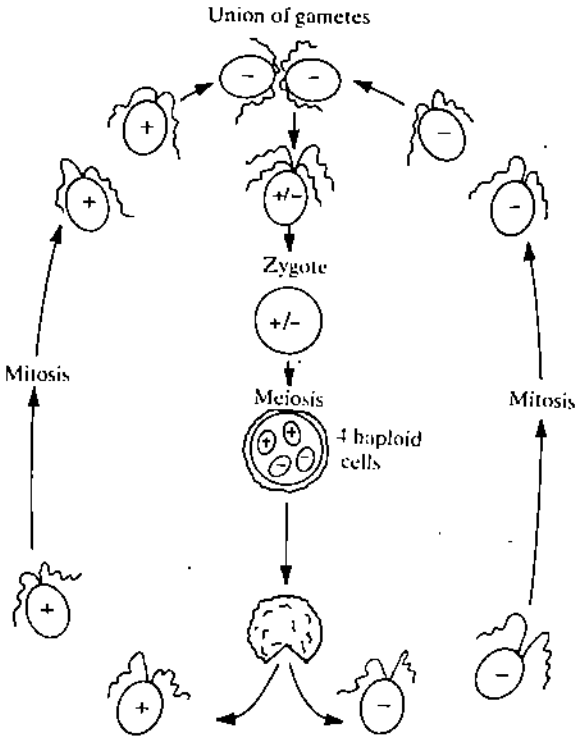


Fig. 6.24: The life cycle of *Chlamydomonas* sp.

in a cyst-like structure. They are, however, retained temporarily. This feature of their life cycle, thus permits the recovery and analysis of all the four products resulting from the meiotic division of one particular diploid cell. Each such group of haploid cells is called a cell tetrad or simply tetrad. And the genetic analysis of such cells which comprise all four products of meiosis is often termed as *tetrad analysis*. In performing an experiment, individual tetrads resulting from a cross are isolated and each cell is grown separately to determine their genotypes. Data on tetrads with like cells are then grouped and comparisons between the number of tetrads in each such group permit identification of independent assortment or of linkage between the genes used in the cross. The distinction between independent assortment and linkage is best illustrated by the following example. Let us assume that the mutant genes *a* and *b* are independent and are borne on separate chromosomes. A cross is made between the mutant and wild-type individual, that is, $ab \times + +$. For the sake of simplicity the wild types are written as $++$ and not $a+$ and $b+$. Now, suppose that 200 tetrads are derived from the above cross and the data is shown in Table 6.2.

Table 6.2: Tetrad Analysis in *Chlamydomonas*

Tetrad Type	Group 1 Parental (P)	Group 2 Nonparental (NP)	Group 3 Tetratype (T)
Genotypes	*ab* *ab* *++* *++*	*a+* *a+* *+b* *+b*	*++* *a+* *+b* *ab*
Number of Tetrads	45	46	9

Let us analyse the above data. In the Group 1 tetrads, two cells are *ab* and two cells are *++* in genotypes. Since these gene arrangements are the same as those found in the parents therefore the group 1 tetrads are designated as parental types (P). Group 2 tetrads are termed nonparental (NP) because of the presence of two cell types, that are *a+* and *+b*, and these represent a gene combination different from that of either parent. Comparing the number of P and NP tetrads, it can be seen that they are approximately equal in frequency, and thus the four types of cells: *ab, ++, a+* and *+b* are present in a 1:1:1:1 ratio. From this we can conclude that the genes *a* and *b* are assorting independently and are located on separate non-homologous chromosomes.

Next consider the Group 3 tetrads. All four types of gene combinations are present within each tetrad. Therefore, this group is designated as tetratype (T). The tetratypes indicate the occurrence of crossing-over between gene *a* and the centromere, or between *b* and the centromere (see Fig. 6.25). With either exchange (that is between *a* and centromere, or *b* and centromere), segregation of mutant gene from its wild-type allele will occur, thus resulting in tetratype.

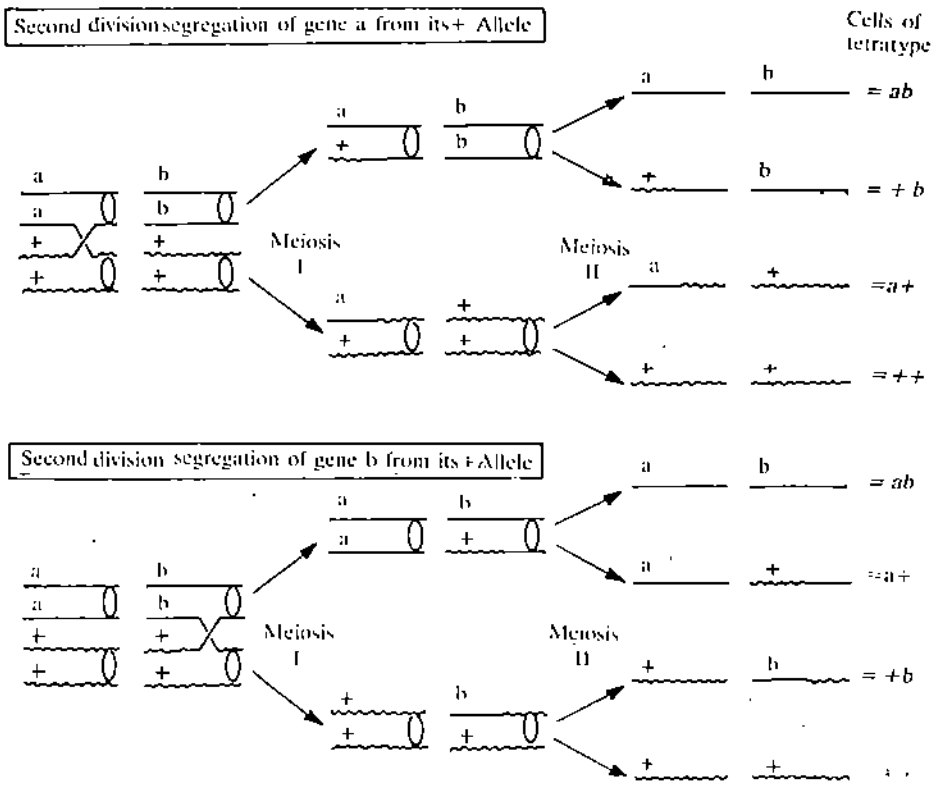


Fig. 6.25: Origin of tetratypes through second division segregation of independently assorting gene *a* and *b*. Second division segregation results from crossing-over between the gene locus and the centromere.

A tetratype thus contains two kinds of parental and two kinds of nonparental cells. You should note that the occurrence of second division segregation does not alter the ratio between parental and nonparental type cells and thus does not interfere with their independent assortment.

If the genes *a* and *b* were linked, the cross *ab* × *++*, would produce data quite different from those give in Table 6.2, provided of course that these gene loci were close enough together so that an exchange between them did not occur in every meiotic cell. In the sample data given below (see Table 6.3) compare the number of parental and nonparental tetrads, it is 73:3. This ratio is far from equal, and such a

Table 6.3: Tetrad Analysis in *Chlamydomonas*, when the Genes are Linked.

Tetrad Type	Group 1 Parental (P)	Group 2 Nonparental (NP)	Group 3 Tetratype (T)
Genotypes	*ab* *ab* *++* *++*	*a+* *a+* *+b* *+b*	*++* *a+* *+b* *ab*
Number of Tetrads	73	3	24

kind of ratio would occur if the genes were linked. On this basis we can conclude that genes *a* and *b* are linked. If it is so then how to account for the origin of nonparental tetrads (i.e., category 2 of Table 6.3). Nonparental tetrads can arise only from four-strand double crossing-over. As shown in Fig. 6.26, this also accounts for their rarity.

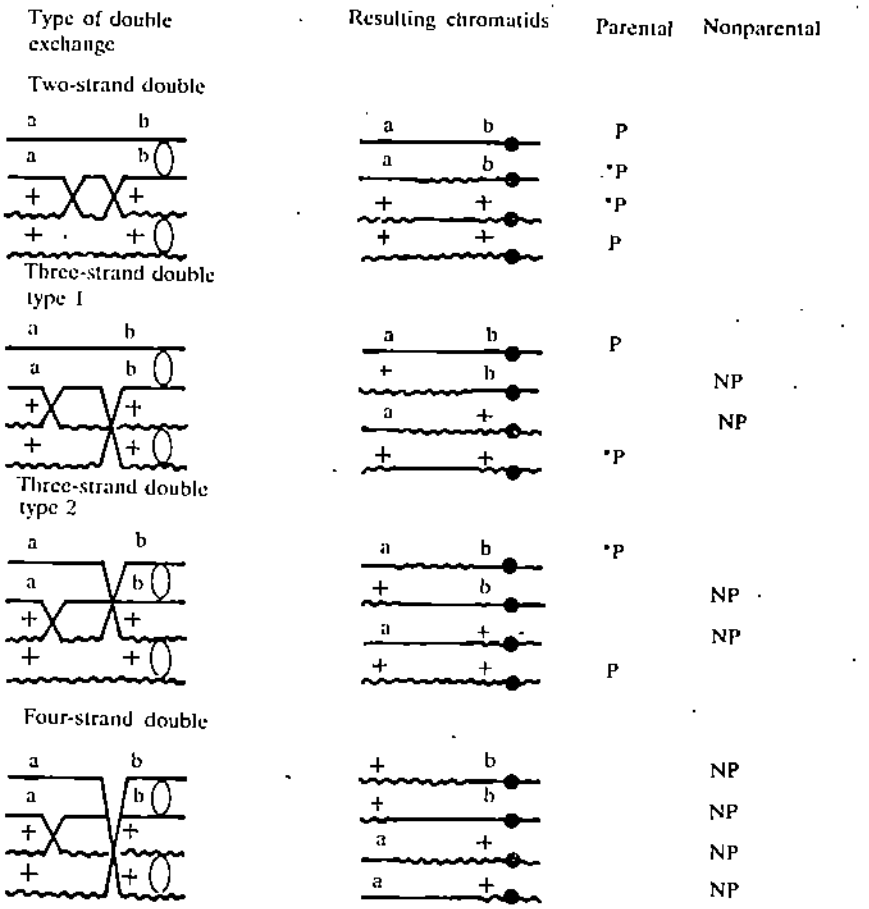


Fig. 6.26: The consequences of double crossing-over between linked genes. The possible kinds of double cross-overs are shown, along with the resulting chromatids. Progeny phenotypes are classified as parental (P) and nonparental (NP) or phenotypically parental originating from a double exchange (P'). Note that parentals and nonparentals occur with equal frequency.

Now consider the tetratypes in category 3, these also originate from crossing-over. Each such tetratype consists of two parental and two nonparental type cells. These originate (Study Figs 6.26 and 6.27 carefully) basically due to single cross-over and three-strand double cross-overs.

An estimate of the map distance between genes *a* and *b* can be calculated according to the following formula:

$$\frac{NP + \frac{1}{2}(T)}{\text{Total tetrads}} \times 100 = \text{Recombination Frequency (RF) or Map distance}$$

In this formula, the sum of the total number of nonparentals and one-half of the number of tetratypes is divided by the total tetrads observed in the experiment. Only half of the tetratypes are included because only half of the cells of this group are recombinant (*a*+ and +*b*). All nonparentals are included because in these tetrads, every cell is recombinant. Using the above formula and data in Table 6.3, the map distance between the genes *a* and *b* would be:

$$\frac{3 + \frac{1}{2}(24)}{100} \times 100 = 15\% \text{ or } 15 \text{ map units}$$

Now if three linked genes are considered in a cross, the gene order as well as recombination frequencies can be determined. For example, supposing a hypothetical

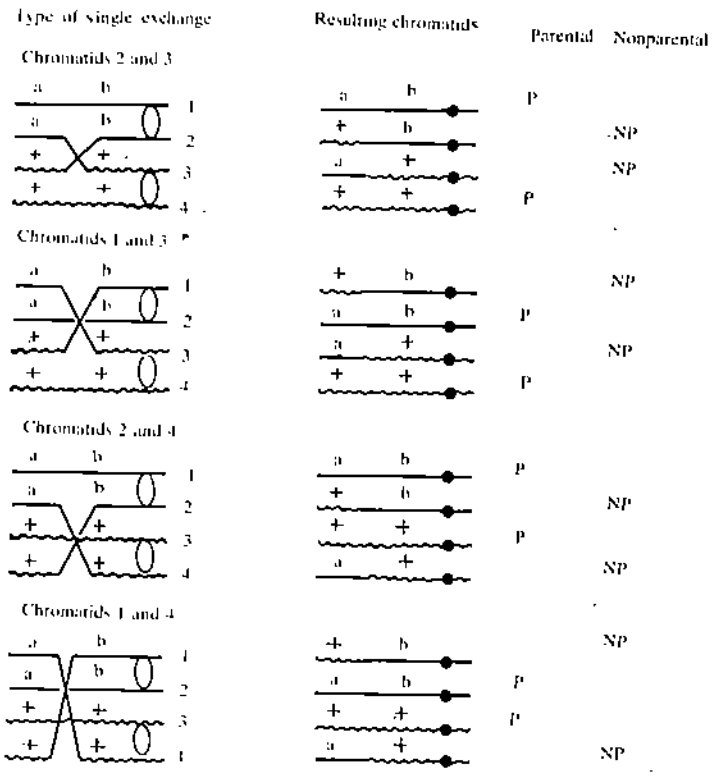


Fig. 6.27: Consequences of single exchanges between linked genes. Homologous chromosomes are represented in the four-strand stage of meiosis with sister chromatids attached to a single centromere. Crossing-over occurs between any two nonsister chromatids. These possibilities and the consequences in each case are shown. Progeny phenotypes are classified as parental (P) or nonparental (NP). Note that the parental and nonparental types occur with equal frequency.

cross *abc* × *+++* yields 100 tetrads which are grouped as follows (Table 6.4).

Table 6.4: Date Obtained from a cross abc × +++.

Category	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7
Genotype	*abc* *abc* *+++* *+++*	*a+c* *abc* *+b+* *+++*	*+bc* *+++* *abc* *a++*	*a+c* *a++* *+bc* *+b+*	*ab+* *abc* *+++* *++c*	*a+c* *a+c* *+b+* *+b+*	*ab+* *a++* *+bc* *++c*
No. of Tetrads	74	13	4	4	2	2	1

In order to determine the recombination frequencies for these genes, first we should examine the relationships between any two genes separate from the third. Each tetrad group can be categorised as P, NP or T with respect to the arrangement of *a* and *b* only, ignoring *c*. Similarly, genes *a* and *c* can be considered apart from *b*, and genes *b* and *c* apart from *a*.

Table 6.5: The Data in Table 6.4 Recyclassified and its Recombination Frequencies Determined.

Tetrad Group	*a* − *b*	*a* − *c*	*b* − *c*
1	74 P	74 P	74 P
2	13 T	13 P	13 T
3	4 T	4 T	4 P
4	4 NP	4 T	4 T
5	2 P	2 T	2 T
6	2 NP	2 P	2 NP
7	1 T	1 NP	1 T
Total	100	100	100
Total of all the Ps, NPs and Ts in the above tetrad groups	76 P 6 NP 18 T	89 P 1 NP 10 T	78 P 2 NP 20 T
Recombination Frequency	15%	6%	12%

Applying the formula for recombination frequency in each case, the map distance between each pair of genes can be determined. The gene which are the farthest apart will exhibit the highest frequency of exchange. From the data shown in Table 6.5 we can conclude that *a* and *b* are at opposite ends, while *c* is in the middle of this region of the linkage map, and the order *a-c-b*. Note that when genes *a* and *b* are considered apart from *c*, only 15 per cent recombination is evident. However, when gene *c* is included and the separate map distances are added a total recombination frequency of 18 per cent between *a* and *b* is obtained. The presence of gene *c* has permitted recognition of cross-overs that would otherwise go undetected.

Ordered Tetrads in *Neurospora*

Although all products of a single meiosis can be obtained in *Chlamydomonas*, the non-linear arrangement of the four haploid cells of a tetrad does not give any information as to which chromatids participate in particular exchanges during this process. In *Neurospora*, such inferences, however, can be made because the linear order of spores within the ascus sac directly reflects prior meiotic events. In *Neurospora* meiosis occurs immediately after fusion of nuclei of opposite mating types within a structure called ascus (see Fig. 6.28). The two meiotic divisions are

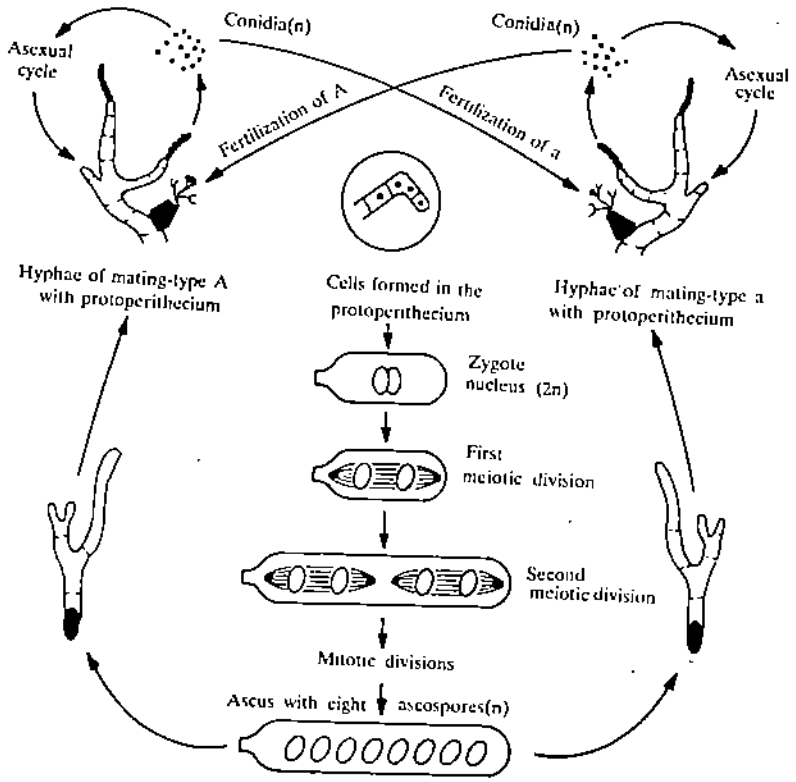


Fig. 6.28: Life Cycle of *Neurospora* sp. The vegetative body consists of partly segmented filaments called hyphae. Conidia are asexual spores that also function in the fertilisation of organisms of the opposite mating type. Note the alignment of nuclei in the ascus during meiosis and the subsequent division. The spindles in any of the divisions do not overlap, thus permitting the ordered segregation of the ascospores.

followed by a mitotic division to produce eight spores within the ascus. The spindles of these nuclear divisions do not overlap one another and so the order of spores in the sac reflects the order of segregation of centromeres and the first and the second meiotic divisions as shown in Fig. 6.28. The ordered development of spores from the tetrad makes it possible to detect crossing-over between the centromere and mutant genes as well as between mutant genes (Fig. 6.29). In the absence of crossing-over, alleles segregate from each other at the first meiotic division and are found in the spores at the opposite ends of the ascus. If a single cross-over occurs between the centromere and the different alleles of a gene, the different alleles do not segregate until the second meiotic division, producing an ascus in which parental genotypes are found at both ends of the ascus. A single cross-over between two genes produces an ascus in which alleles proximal to the cross-over segregate at the first division, and alleles distal to the cross-over segregate at the second division. Such asci contain half recombinant-type spores and half parental-type spores. The frequency of asci showing second division segregation is a measure of genetic distance between the centromere and the gene locus. Since only half of the chromatids in a tetrad are

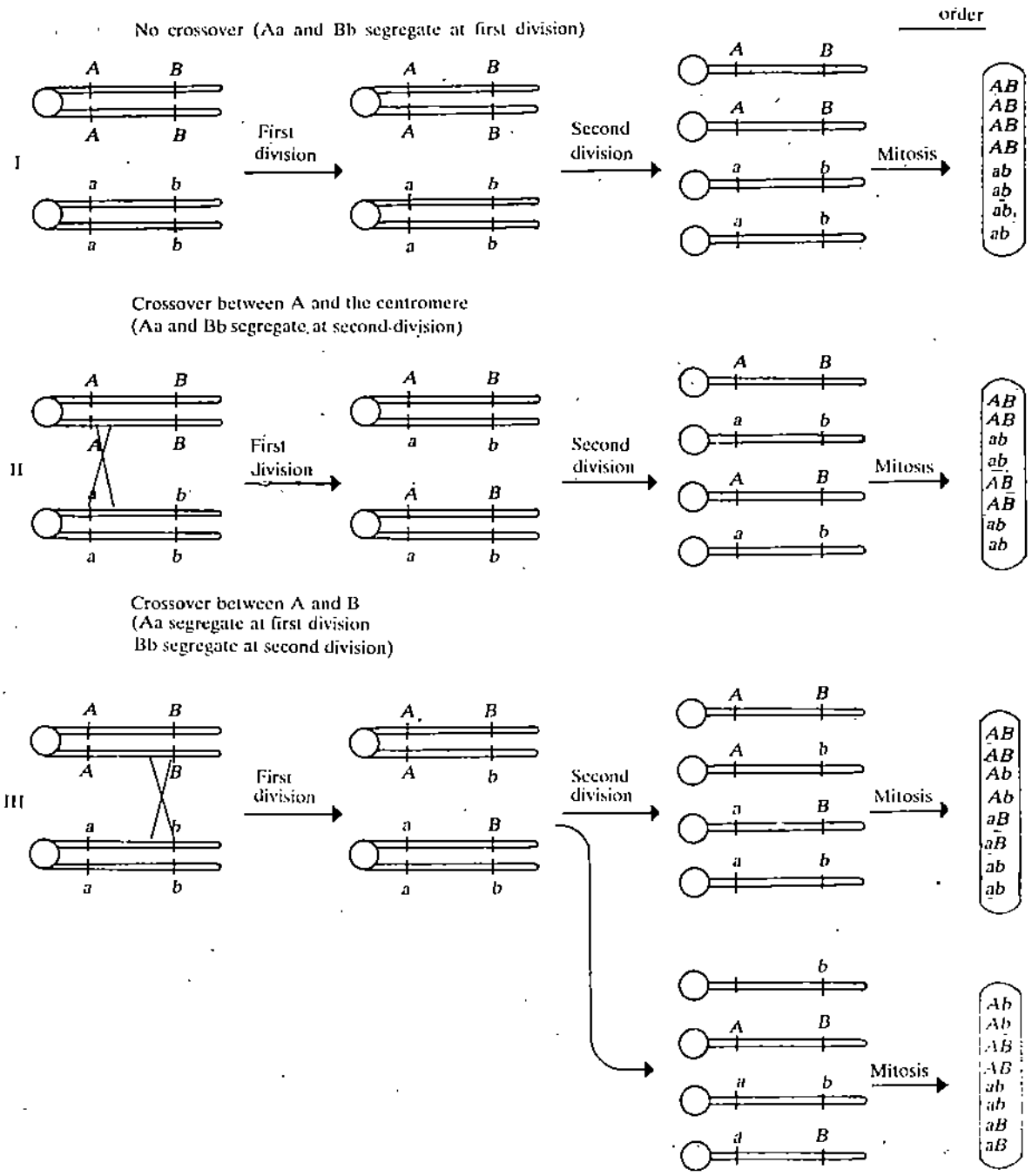


Fig. 6.29: The first and second-division segregation of alleles of each genotype based on their position in the ascus. Since the ascospores are haploid, therefore, a spore's phenotype reflects its genotype. A cross-over between the centromere and *Aa* results in second-division segregation of all markers distal (*Aa, Bb*) to the cross-over (II). A cross-over between *Aa* and *Bb* results in first-division segregation of markers proximal (*Aa*) to the cross-over and second division segregation of markers distal (*Bb*) to the cross-over (III). Depending on how the recombinant chromosomes align at the first division, four different orders are possible, but only two are illustrated here.

recombinant, the map distance between gene and centromere is half of the frequency of asci showing second-division segregation. Therefore, the map distance between a gene and its centromere is given by the equation

$$\frac{\frac{1}{2}\,(\text{Asci with second division segregation patterns})}{\text{Total number of Asci}} \times 100$$

The ordered arrangement of ascospores of *Neurospora* thus permits us to map the centromere with respect to mutant alleles of genes on the same chromosome.

Four types of double cross-over events can be observed in *Neurospora* tetrads. These are illustrated in Fig. 6.30. The occurrence of cross-over events involving three and four chromatids revealed by the ordered ascospores proves that crossing-over occurs at the tetrad stage, when four chromatids are present rather than at a two-strand

stage prior to replication of the DNA of the chromosomes. Such data also demonstrate a very important point, that is, the recombination events occurring during a single meiosis yield reciprocal recombinant types.

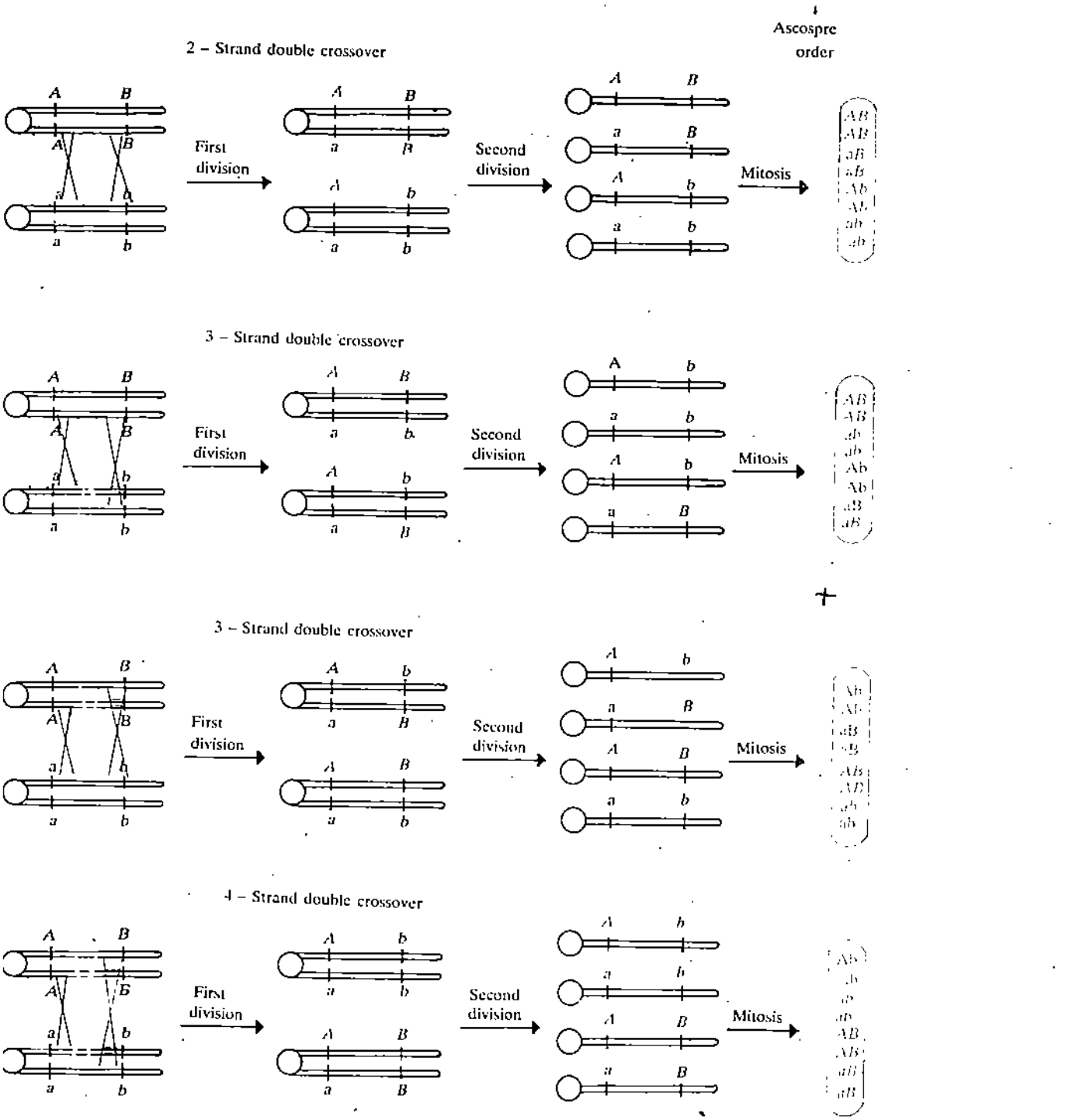


Fig. 6.30: The four types of cross-over events that can be detected in ordered asci.

6.5 WHY DIDN'T MENDEL FIND LINKAGE?

t is often said that Mendel had extremely good fortune in his experiments with arden pea as in none of his crosses did he come across any apparent linkage etween any of the seven traits he studied. Had Mendel obtained highly variable data haracteristic of linkage and crossing-over, these unorthodox ratios may have indered his successful analysis and interpretation of data.

)ne of the common and simplest explanation for the absence of linkage is that each f the seven genes was located on different linkage group or chromosome. As *Pisum ativum* has a haploid number of 7, so this speculation was widely accepted.

Stig Blixt, however, demonstrated the inadequacy of this hypothesis and explained why Mendel did not encounter ratios, characteristic of linkage and crossing-over (see Box 6.1).

Box 6.1: Why Didn't Gregor Mendel Find Linkage?

It is quite often said that Mendel was very fortunate not to run into the complication of linkage during his experiments. He used seven genes and the pea has only seven chromosomes. Some have said that had he taken just one more, he would have had problems. This, however is a gross oversimplification. The actual situation, most probably, is shown in Table 6.6. This shows that Mendel worked with three genes in chromosome 4, two genes in chromosome 1, and one gene in each of chromosome(s) 5 and 7. It seems at first glance that, out of the 21 dihybrid combinations Mendel theoretically could have studied, no less than four (that is, *a-i, v-fa, v-le, fa-le*) ought to have resulted in linkages. As found, however, in hundreds of crosses and shown by the genetic map of the pea[1], *a* and *i* in chromosome 1 are so distantly located on the chromosome that no linkage is normally detected. The same is true for *v* and *le* on the one hand and *fa* on the other, in chromosome 4. This leaves *v-le*, which ought to have shown linkage.

Table 6.6 : Relationship between Modern Genetic Terminology and Character Pairs used by Mendel

Character Pair Used by Mendel	Alleles in Modern Terminology	Located in Chromosome
Seed colour: yellow-green	*I-i*	1
Seed coat and flowers: coloured-white	*A-a*	1
Mature pods: smooth expanded-wrinkled indented	*V-v*	4
Inflorescences: from leaf axilsumbellate in top of plant	*Fa-fa*	4
Plant height: 1 m-around 0.5 m	*Le-le*	4
Unripe pods: green-yellow	*Gp-gp*	5
Mature seeds: smooth-wrinkled	*R-r*	7

Mendel, however, seems not to have published this particular combination and thus, presumably, never made the appropriate cross to obtain both genes segregating simultaneously. It is, therefore, not so astonishing that Mendel did not run into the complication of linkage, although he did not avoid it by choosing one gene from each chromosome.

Stig Blixt, 1975
From: Nature, Vol. 256, p.206.

SAQ 7

State whether the following statements are true or false.

i) Forty-five per cent crossing-over is equal to 45 map units.

ii) A genetic map is a good representation of the actual structure of a chromosome.

iii) Ten per cent crossing-over is equal to one map unit.

iv) In the formula:

$$X = \frac{\text{observed frequency of double cross-overs}}{\text{expected frequency of double cross-overs}}$$

X represents interference

v) In a three-point cross-over, the genes in the center position can be determined by comparing the parental type genes and the double cross-over type genes.

6.6 SUMMARY

In this unit you have learnt that:

- Certain non-allelic genes located on the same chromosome tend to remain together during meiosis rather than undergoing independent assortment. This phenomenon is called linkage. The indication of linkage is deviation from the

1:1:1:1 ratio of phenotypes in the progeny of a cross of the form *AaBb* × *aabb* when genes in a cross with two unlinked genes segregate, more than 50 per cent of gametes produced have parental combinations of the segregating alleles and less than 50 per cent have nonparental (recombinant) combination of alleles.

- The recombination of linked genes occurs by crossing-over, a process by which nonsister chromatids of the homologous chromosomes exchange corresponding segments. Crossing-over involves the breakage of individual chromatids and exchange of parts. This process of breakage and reunion is usually associated with a small amount of DNA repair synthesis. Crossing-over occurs after chromosome duplication, in the four-chromatid stage of meiosis. A given cross-over involves any two of the four chromatids.
- The frequencies of crossing-over between different genes can be used to determine the relative order and locations of the genes in chromosomes. This is called genetic mapping. Distance between adjacent genes in such a map (a genetic or linkage map) is defined to be proportional to the frequency of recombination between them. The unit of map distance (one map unit - mu or cM) is defined as 1 per cent recombination. One map unit corresponds to a physical length of the chromosome in which a cross-over event will occur. For short distances map units are additive.
- The four haploid products of individual meiotic divisions can be used to analyse linkage and recombination in some species of fungi and unicellular algae. The method is called tetrad analysis. In *Neurospora*, the meiotic tetrads are contained in a tubular sac or ascus, in a linear order making it possible to determine whether a pair of alleles segregated in the first or the second meiotic division. With such asci, it is possible to use the centromere as a genetic marker, and in fact centromere serves as a reference point from which all genes in the same chromosome can be mapped. Linkage analysis in unordered tetrads such as in *Chlamydomonas* uses the frequencies of parental (P), nonparental (NP) and tetratype (T) tetrads.

6.7 TERMINAL QUESTIONS

1) From the data given below answer i) and ii)

Genotype	Number of Progeny
QR	327
Qr	61
qR	56
qr	342

i) The above data indicates that genes:
   a) are completely linked
   b) are not linked
   c) are partially linked
   d) no conclusion can be made

ii) The per cent recombination that has occurred is:
   a) 10%
   b) 15%
   c) 30%
   d) 50%

Fill in the blank spaces with appropriate words:

i) Two events that take place during meiosis contribute to genetic variability: ..... and ..... . The general term for production of new gene combination is genetic ..... .

ii) The more chromosomes an organism has, the more genetic variability it gets from ..... .

iii) The longer the chromosomes of an organism, the more genetic variability it gets from ..... .

iv) Assume that the genes for hair colour and eye colour are closely linked, with the alleles for brown hair and brown eyes on one chromosome of a homologous pair; and the alleles for blonde hair and blue eyes on the other chromosome of the pair. A cross-over between these two loci would produce a child with ..... hair and ..... eyes, or ..... hair and ..... eyes.

v) Independent assortment recombines genes from ..... chromosomes, and crossing-over recombines genes from ..... chromosomes.

3) Tick mark the correct answer for the following:

i) Independent assortment during meiosis produces gametes with all possible combinations of maternal and paternal chromosomes. How many different kinds of haploid gametes does a diploid cell with four chromosomes produce?
   a) Two
   b) Four
   c) Eight
   d) Sixteen

ii) Assuming there is an independent assortment of chromosomes and no crossing-over in a cell with n pairs of chromosomes. How many different kinds of gametes is the organism expected to produce?
   a) n
   b) n^2
   c) 2^n
   d) $\frac{1}{2}$ n

iii) During which phase of meiosis does crossing-over take place?
   a) Prophase of meiosis I
   b) Prophase of meiosis II
   c) Anaphase of meiosis I
   d) Anaphase of meiosis II

iv) During crossing-over, genetic information is exchanged between two:
   a) chromatids of a chromosome
   b) long arm of a chromosome
   c) chromatids of two homologous chromosomes
   d) chromatids of two nonhomologous chromosomes

v) The probability of cross-over occurring between two gene loci is proportional to:
   a) the activity of the two loci
   b) the distance between the loci from the centromere
   c) the distance between the two loci
   d) the length of the chromosome

vi) The maximum frequency of a recombination at two loci is:
   a) 25%
   b) 50%
   c) 75%
   d) 100%

vii) The recombination frequency between gene *A* and *C* is 17 per cent and between *B* and *C* is 26 per cent. What is the sequence of genes in the chromosome?
   a) *ABC*
   b) *ACB*
   c) *BAC*
   d) cannot be determined from the information provided.

viii) In a test cross between two linked genes having 50% crossing-over between them, what phenotypic ratio of the progeny should be expected?
   a) 9:3:3:1
   b) 1:1:1:1
   c) 1:2:1
   d) 2:2

ix) In order to calculate map distances of genes on a chromosome, one must know the:
   a) number of mutant genes
   b) cross-over percentage
   c) recombination frequency of each locus
   d) a and c

x) In order to determine the gene sequence in a three-point cross you must know which of the progeny are:
   a) parental (non cross-over) types
   b) single cross-over types
   c) double cross-over types
   d) a, b and c

4) In a plant, long leaves (*S*) and green veins (*Y*) are dominant over short leaves (*s*) and yellow veins (*y*). The cross *SSYY* × *ssyy* produced an F_1 *SsYy*. When F_1 plants were inbred, the F_2 consisted of 570 long, green individuals and 190 short, yellow. Are the genes S and Y linked?

5) In *Drosophila*, the recessive, sex-linked genes abnormal eye facet (*fa*) and singed bristles (*sn*) show 18 per cent recombination.
   a) If a singed male is crossed to a $\frac{fa^+}{fa^+}$ female, what phenotypes are expected in the F_1?
   b) If the F_1 males and females are inbred what phenotypic proportions would be expected to occur in F_2 males and females?

6) Two recessive genes *ds* and *mp* are present in corn. These are linked and are 20 map units apart. From the cross:

$$\frac{ds\ mp}{+\ +} \times \frac{ds\ +}{+\ mp}$$

What percentage of the progeny would be expected to be both *ds* and *mp* in the phenotype?

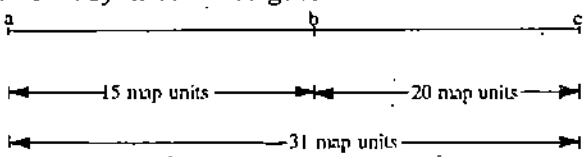
7) The following data was obtained in *Neurospora* by analysis of ordered spores from the cross $\frac{a\ +\ +}{+\ b\ c}$

The total number of tetrads was

(1)	(2)	(3)	(4)	(5)	(6)
a + + *a + +* *+ b c* *+ b c*	*a b c* *a b c* *+ + +* *+ + +*	*a b c* *+ b c* *a + +* *+ + +*	*a + +* *+ + +* *a b c* *+ b c*	*a + +* *a + c* *+ b +* *+ b c*	*+ + +* *+ + c* *a b +* *a b c*
40	40	5	5	5	5

   a) Which of the genes are linked?
   b) What is the map distance between them?
   c) What is the map distance between gene *a* and its centromere?

8) The following diagram summarises the recombination frequencies observed in an experiment to study three linked genes.



What can be the observed frequency of double crossing-over in this experiment? Calculate the interference.

6.8 ANSWERS

Self-assessment Questions

1) a) i) unlinked

ii) linked
iii) linked

b) 2

c) [D-E-F-G], [H-I]

d) Genes located on the same chromosome are called linked genes and belong to a linkage group.

2) a) i) linked
ii) 3:1
iii) 1:2:1
iv) complete linkage

b) i) linked
ii) exchange between chromatids
iii) 2, parentals, recombinants
iv) partial linkage

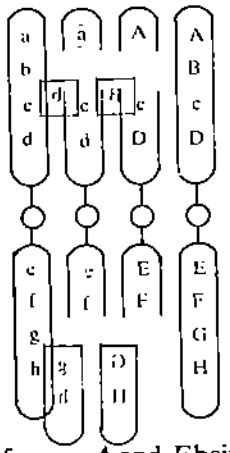
c) i) independent
ii) heterozygous
iii) 9:3:3:1
iv) nonlinkage

3) $\frac{ab}{AB}$

4) a) During the first prophase of meiosis.

b) Exchange occurs between the chromatids of homologous chromosomes.

5)



6) There is a greater chance of genes *A* and *F* being separated by crossing-over. For *A* and *B* to be separated, chiasma formation must occur at position 1. For *A* and *F* to be separated, the chiasma can form at any one of the five positions (1,2,3,4 or 5).

7) i) T
ii) T
iii) F
iv) F
v) T

Terminal Questions

1) i) c
ii) b

2) i) independent assortment, crossing-over, recombination
ii) independent assortment
iii) crossing-over
iv) brown, blue, blonde, brown
v) non-homologous, homologous

3) i) b
ii) c
iii) a
iv) c

v) c
vi) b
vii) c
viii) b
ix) c
x) d

4) The genes *S* and *Y* are linked; the ratio is 3:1 indicating complete linkage that no crossing-over. If not linked, a ratio of 9:3:3:1 would have been obtained.

5) a) F_1 females : $\frac{fa\ +}{+\ Sn}$ (wild type)

F_1 males $\frac{fa\ +}{Y}$ (facet)

b) From the data, the cross-over frequency between *fa* and *sn* is 18 per cent. Therefore, cross-over eggs will be *++* (9 per cent) and *fa sn* (9 per cent), and noncross-over eggs would be *fa +* (41 per cent) and *+ sn* (41 per cent). Union of eggs with *fa +*, X-bearing sperm will produce females in the phenotypic proportions of 50 per cent wild type and 50 per cent facet. Union of eggs with Y-bearing sperm will produce males in the phenotypic proportions of 41 per cent facet, 41 per cent singed, 9 per cent facet, singed and 9 per cent wild type.

6) Forty per cent of the gametes of $\frac{ds\ mp}{+\ +}$ parents will carry both *ds* and *mp;*

10 per cent of the gametes produced by the $\frac{ds\ +}{+\ mp}$ parent will carry both recessive genes. Therefore, $0.4 \times 0.1 = 0.04$ or 4 per cent of the progeny would be homozygous for these genes.

7) a) genes *b* and *c* are linked
b) 5 map units
c) gene *a* is 5 map units from its centromere

8) Let x be the observed frequency of double cross-overs. Then, single cross-overs in the $a - b$ interval occur with frequency $0.15 - x$ and single cross-overs in the $b - c$ internal occur with frequency $0.20 - x$. The observed recombination frequency between *a* and *c* equals the sum of the single cross-over frequencies (because double cross-overs are undetected), which implies that $0.15 - x + 0.20 - x = 0.31$ or $x = 0.02$. The expected frequency of double cross-overs equals $0.15 \times 0.20 = 0.03$ and the coincidence is therefore,

$\frac{0.02}{0.03} = 0.67$. The interference equals $1 - 0.67 = 33$ per cent.

GLOSSARY

agglutination: the clumping together of cells, especially red blood cells, caused by the reaction between an antigen on the cell surface and an agglutinin (antibody) in the serum.

albinism: a condition caused by lack of melanin production in the iris, hair and skin. In humans normally inherited as autosomal recessive.

allele: one of the two or more alternative forms of a gene, each possessing a unique nucleotide sequence. Different alleles of a gene are usually recognised by the phenotypes rather than by comparison of their nucleotide sequences.

antibody: a protein molecule that recognises and binds to a foreign substance introduced into the organism.

antigen: any large molecule, often a protein, which stimulates the production of specific antibodies or which bind specifically to an antibody.

artificial insemination: is the procedure by which semen from an acceptable donor is drawn into a syringe and deposited in the uterus at a time when fertilization is ensured.

autosomes: chromosomes other than the sex chromosomes.

back cross: a mating between a heterozygote and an individual homozygous for either the dominant or the recessive allele; differs from test cross in that it involves the crossing of an offspring with one of the parents.

barr body: is the darkly stained condensed X-chromosome observed in the interphase of mammalian female cells.

castration: removal of testes (male gonads) by surgery.

centromere: a chromosomal region that becomes associated with the spindle fibres during mitosis and meiosis.

chiasmata: the places at which pairs of homologous chromatids are in contact (from late prophase of meiosis to the beginning of the first anaphase) and at which exchanges of homologous parts between nonsister chromatids have taken place by crossing-over (singular: chiasma).

chromatid: in all duplicated chromosomes, each of the two longitudinal subunits that become visible during mitosis or meiosis.

chromosome: a thread like structure found in the nuclei of cells, composed of chromatin that stores and transmits genetic information.

chromosome set or chromosome complement: the normal gametic complement of chromosomes of a diploid individual.

chromosome puffing: a localised uncoiling and swelling in a polytene chromosome of e.g., *Drosophila*. It is regarded as a sign of active transcription.

clonal transmission: genetic constitution passed on from a single cell to all its daughter cells produced by mitosis.

codominance: a situation in which the heterozygote exhibits the phenotypes of both homozygotes.

complete dominance: the case in which one allele is dominant to the other so that at the phenotypic level the heterozygote is essentially indistinguishable from the homozygous dominant.

conception: fertilisation of the ovum by the sperm, it is the very beginning of pregnancy.

concordance: pairs or groups of individuals identical in their phenotype. In twin studies, a condition in which both twins exhibit or do not possess a trait under investigation.

consanguinity: the sharing of genes from at least one recent common ancestor.

cortisone: one of the hormones produced by the cortex of the adrenal gland.

crossing-over: the exchange of chromatid segments between homologous chromatids during meiosis; if different alleles are present on the chromatids, crossing-over can be detected by the formation of genetically recombinant chromatids.

deoxyribonucleic acid DNA: a polynucleotide in which the sugar residue is deoxyribose and which is the primary genetic material of all cells.

dihybrid cross: a cross in which two pairs of alleles are being followed.

diploid: having two complete haploid sets of chromosomes typical for the species or group.

discordance: in twin studies, a situation where one twin shows a trait but the other does not.

dizygotic twins: twins produced from separate fertilisation events; two ova fertilised independently. Also known as fraternal twins.

DNA ligase: an enzyme that creates a phosphodiester bond between the 5-PO_4 end of one polynucleotide and the 3-OH end of another, thereby producing a single, larger polynucleotide.

DNA polymerase: the enzyme responsible for synthesising DNA from deoxyribonucleoside triphosphates under the direction of a template DNA strand.

dominant: an allele or phenotype, that is expressed both in the homozygous or heterozygous state.

dosage compensation: a genetic mechanism which regulates the levels of gene products equalising their amount in males and females.

endonuclease: an enzyme that hydrolyses internal phosphodiester bonds in a polynucleotide.

endosperm: in flowering plants, a tissue specialised for nourishing the developing embryo.

epistasis: interaction of alleles at different loci; the control or masking of the expression of alleles at one locus by alleles at a second locus.

exonuclease: an enzyme that hydrolyses terminal phosphodiester bonds (at 3^1 or 5 ends) in a polynucleotide.

expressivity: the degree or range in which a phenotype for a given trait is expressed.

fertilisation: union of two special cells, the gametes.

F-factor: (fertility factor) bacteria have an independent circular DNA or episome which is responsible for fertility, and is also known as the sex factor.

first filial generation (F_1): the generaton of individuals produced by the first parental generation (P_1) or the first parents being considered.

fraternal twins: see dizygotic twins.

foetus: the newly developing individual within the mother's womb.

gene: classically a unit of inheritance occpying a specific site (locus) on a chromosome, that has one or more specific effects on an organism, and can both recombine with other such genetic units, and mutate independently to other allelic forms.

gene interaction: production of novel phenotypes by the interaction of alleles of different genes.

genome: the genetic content of cell or virus; in eukaryotes, it sometimes refers to only one complete haploid chromosome set.

genetic counselling: this includes analysis of risk for genetic defects in a family; imparting advice about the options available to avoid possible risks for future generations.

genotype: the genetic constitution of a cell or of an individual.

gonads: organ of animals which produce gametes.

gynandromorph: an individual exhibiting both male and female sexual differentiation.

haemophilia: a disease in which there is a blood-clotting defect. Males are usually affected as the trait is a X-linked recessive one.

haploid: having one complete set of chromosomes typical for that species.

hermaphrodite: (bisexual, monoecious): animals or plants producing both male and female gametes.

heterogametes: individuals which have a pair of dissimilar sex-chromosomes (one X- and one Y- chromosome or one Z and one W chromosome) and produce two different types of gametes.

heteromorphic: having a dissimilar morphology or appearance.

heterozygous: having alternative forms of a gene (alleles) at a given locus, one allelic form on each of the two homologues.

holandric inheritance: the pattern of transmission of genes linked to Y-chromosome. Only males are affected and the trait is passed on from affected males to all their sons but to none of their daughters.

homogametic: individuals which have a pair of similar sex chromosomes (X-chromosomes or Z-chromosomes) and produce only one type of gametes.

homologous: of chromosomes or chromosome segments that are identical with respect to their constituent genetic loci and their visible structure.

homozygous: having the same allele (gene form) present at a given locus on both homologous chromosomes.

hormone: organic substance produced in minute quantity in endocrine glands and transported to other parts where it exerts its effect.

HY-antigen: it is an antigen determined by Y-linked gene.

hybrid: an offspring of a cross between two genetically unlike individuals.

incomplete dominance: with reference to two alleles at a locus the expression of a heterozygous phenotype which is distinct from, and often intermediate to, that of either parent.

identical twins: see monozygotic twins.

immunoglobulin: the class of serum proteins having the properties of antibodies.

independent assortment: the segregation at meiosis of a pair of alleles independently of other allelic pairs whose loci are found on different chromosomes.

interference: a measure of the degree to which one cross-over by a chromatid affects the probability of a second cross-over by the same chromatid.

intersexes: individuals with characteristics of both males and females.

isogametes: morphologically identical gametes.

karyotype: the chromosome complement of a cell or an organism, characterised by the number, size and configuration of the chromosomes.

lethal gene: a gene whose expression results in death.

Lesch-Nyhan syndrome: the clinical condition resulting from deficiency of the enzyme Hypoxanthine guanine phosphoribosyl transferase (HGPRT) which is X-linked.

linkage: a measure of the degree to which alleles of two genes assort independently at meiosis or in genetic crosses.

linkage group: a set of gene loci that can be placed in a linear order representing the different degrees of linkage between loci.

linkage map: a chromosome map showing the linear order of the genes associated with the chromosome.

lyonisation: is the phenomenon of X-chromosome inactivation in female mammals to bring about dosage compensation.

marker: an allele whose inheritance is under observation in a cross.

metafemales: organisms having more than the usual dose of female determiners. Often sterile.

metamales: organisms having more than the usual dose of male determiners. Often sterile.

miscarriage: spontaneous detachment and expulsion of embryo or foetus.

monohybrid: an individual heterozygous at a locus under consideration, thus having a pair of alleles at that locus. Also, a cross in which a pair of alternative gene forms (alleles) is being followed.

monozygotic twins: twins derived from the splitting of a single zygote shortly after fertilisation; identical twins.

multiple alleles: three or more forms of a gene, any one of which can occur at a given locus on a specific chromosome.

nondisjunction: the failure of two sister chromatids or homologous chromosomes to separate during cell division, so that both go to the same pole, thus producing aneuploid nuclei.

norm of reaction: the range of all possible phenotypic responses of a particular gene or set of genes over all environments.

parental type: an association of genetic markers, found among the progeny of a cross that is identical to an association of markers present in a male gamete.

parthenogenesis: the production of an embryo from a female gamete without participation of a male gamete.

partial dominance: see incomplete dominance.

pedigree: a diagram showing the ancestral relationships among individuals of a family over two or more generations.

penetrance: the frequency (expressed as a percentage) with which individuals of a given genotype manifest at least some degree of a specific (mutant) phenotype associated with a trait.

penetrant: expressed; a gene showing any phenotypic manifestation is said to be penetrant; with respect to an individual, a gene is either penetrant or not; a qualitative measure of gene expression in an individual.

phenocopy: phenotype caused by an environmental factor that resembles a phenotype normally caused by a genetic factor; for example, sunbleached hair (environmentally caused) resembles blonde (genetically controlled) hair.

phenotype: the observable characteristics of an individual, resulting from the interaction between the genotype and the environment in which development occurs.

pleiotropic: the multiple phenotypic effects of a single allele.

recessive: any gene form that is not expressed in the presence of the other allele in the heterozygote. Also, any trait that is expressed phenotypically only when the responsible gene form is present in double dose in a homozygote.

reciprocal crosses: crosses in which the dominant allele is contributed by one parent and the recessive allele by the other parent and vice versa. Those in which each of the two strains provides the males in one cross and the females in the other, as in ♂ A × ♀ B and ♂ B × ♀ A.

recombinant type: an association of genetic markers, found among the progeny of a cross, that is different from any association of markers present in the parents (cf. parental type).

recombination: the creation of a new association of DNA molecules (chromosomes) or parts of DNA molecules (chromosomes).

Rh factor: an antigenic system first described in the rhesus monkey. Recessive rh individuals produce no Rh antigens and are Rh-negative, while RR and Rr individuals have Rh antigens on the surface of their red blood cells and are classified as Rh positive.

Rh incompatibility: where mother and foetus have different Rh blood types; incompatibility between an Rh-negative and Rh-positive foetus often causes *erythroblastosis fetalis.*

sex chromosome: a chromosome, such as X or Y in humans, which is involved in sex determination.

sex-linked: a trait which is expressed in only one sex in spite of not being sex chromosome linked.

sex-linkage: linkage of genes that are located on sex chromosomes.

sex-ratio: the number of males divided by the number of females (sometimes expressed in per cent) at fertilisation or conception (primary sex ratio), at birth (secondary sex ratio), or at sexual maturity (tertiary sex ratio).

synaptinemal complex: a structure found during meiosis, that mediates close pairing between homologous regions of chromatids.

test cross: a mating between a heterozygote and an individual homozygous for the recessive allele.

trihybrid: an individual heterozygous for three pairs of alleles.

transcription: the process by which genetic information from DNA in the chromosome is transmitted to m-RNA and finally results in the formation of gene product.

wild-type: the most commonly observed phenotype or genotype in a population, that is designated as the norm or standard.

X-linkage: genes found on the sex chromosome which is present as a pair in the homogametic sex. There are numerous genes linked on X-chromosome.

Y-linkage: genes present on the sex chromosome responsible for the heterogametic sex. There are few known Y-linked genes.

zygote: the diploid cell produced by the fusion of haploid gametic nuclei.

FURTHER READING

1) Ayala, F.J. & Kiger, J.A. Jr. 1984. Second Edition. *Modern Genetics.* The Benjamin Cummings Publishing Company Inc., California.

2) Carlson, E.A. 1985. *Human Genetics.* Tata McGraw Hill Publishing Co. Limited, New Delhi.

3) Gardner, E.J.; Simmons, M.J. & Snustad, D.P. 1991. Eighth Edition. *Principles of Genetics.* John Wiley & Sons, Inc. New York.

4) Gupta, P.K. 1987. Second Edition. *Genetics.* Rastogi Publications, Meerut.

5) Klug, W.S. & Cummings, M.R. 1986. Second Edition. *Concepts of Genetics.* Scott Foresman & Co., Illinois.

6) Lewin, B. 1983. *Genes.* Wiley Eastern Ltd., New Delhi.

7) Rao, K.R. 1986. Third Edition. *Text Book of Biochemistry.* Prentice Hall of India (Pvt) Ltd., New Delhi.

8) Strickberger. M.W. 1985. Third Edition. *Genetics.* Macmillan Publishing Company, New York.

9) Swanson, C.P; Merz, T. and Young W.J. 1990. Second Edition. *Cytogenetics.* Prentice Hall of India (Pvt) Ltd.; New Delhi.

Dear Student,

While studying these units you may have found certain portions of the text difficult to comprehend. We wish to know your difficulties and suggestions in order to improve the course. Therefore, we request you to fill and send us the following questionaire which pertains to this block.

QUESTIONAIRE

LSE-03
Block-1

Enrolment No. | | | | | | | | | | |

1) How many hours did you need for studying the units?

Unit Number	1	2	3	4	5	6
No. of hours						

2) How many hours (approximately) did you take to do the assignments pertaining to this block?

Assignment Number	1	2
No. of hours		

3) In the following table we have listed 4 kinds of difficulties that we thought you might have come across. Kindly tick (✓) the type of difficulty and give the relevant page number in the appropriate columns.

Page Number	Types of difficulties			
	Presentation is not clear	Language is difficult	Diagram is not clear	Terms are not explained

4) It is possible that you could not attempt some SAQ and TQs.
In the following table are listed the possible difficulties. Kindly tick (✓) the type of difficulty and the relevant unit and question numbers in the appropriate columns.

Unit No.	SAQ No.	TQ No.	Type of difficulty			
			Not clearly posed	Cannot answer on basis of information given	Answer given (at end of Unit) not clear	Answer given is not sufficient

5) Were all the difficult terms included in the glossary. If not, please list in the space given below.

6) Any Other Suggestion(s)

To,

The Course Coordinator (LSE-03; Genetics)
School of Sciences
Indira Gandhi National Open University
Maidan Garhi
New Delhi - 110 068

Uttar Pradesh
Rajarshi Tandon Open University

Block

2

THE PHYSICAL BASIS OF HEREDITY

BLOCK 2 THE PHYSICAL BASIS OF HEREDITY

In the first block you have studied the fundamental concepts of genetics. The main emphasis being on the principles and processes governing heredity. By now, we presume you are well equipped to study and comprehend more specific aspects of heredity that will be the subject matter of this and subsequent blocks.

This block consists of six units, Units 7 to 12, dealing mainly with the elements controlling the inheritance process. These have been discussed in the light of transmission and molecular genetics.

Thus far we have discussed the inheritance pattern as controlled by the genes located in the nucleus. In Unit 7 of this block, we explain the non-mendelian mode of inheritance or more precisely the extranuclear inheritance. We begin with the early experiments that indicated the presence of genetic determinants outside the nucleus, chloroplast and mitochondria. Thereafter, the characteristics of chloroplast and mitochondrial genome and associated protein synthesising machinery are discussed. These two extranuclear genomes have been mapped in a few species. In the last section of the unit, we have discussed the interaction between genomes – nuclear and chloroplast and nuclear and mitochondria, suggesting the possible origin of these organelles.

The focus of Units 8 to 10 is on human chromosomes. In Unit 8 you will learn the morphology and molecular organisation of human chromosomes. Further, the unit discusses the high resolution chromosomal banding studies and that their applications that led to an understanding of the phylogenic relationship between humans and the related taxa.

The chromosome aberrations, that is, abnormalities in chromosomal structure and number are described in Units 9 and 10 respectively. In recent years such studies have gained utmost importance as more and more evidence indicate their involvement in a number of human inherited disorders.

In Unit 11 we begin the study of a branch of genetics that has revolutionised the science of biology in its entirity – the molecular genetics. The chemical nature and structure of genetic material were proposed by James Watson and Francis Crick in 1953 in a two page paper they published in *Nature* under the title "Molecular structure of nucleic acids". The deduction of the structure and chemistry of DNA, along with the input of countless insights and discoveries of molecular nature, transformed the science of genetics.

The genetic study of microbes has played a significant role in the recent advances made in areas of molecular biology, recombinant DNA technology as well as in the production of biologically important substances such as insulin, somatotropin and blood clotting factors. Microbes, more specifically bacteria, have been quite useful for the combined genetic and biochemical approach to solve many a difficult problems in biology. The discovery and analysis of plasmid and bacteriophage systems revealed the finer details of genetic organisation which traditional genetic methods could not have unravelled. In Unit 12, we discuss the various aspects of bacterial and phage genetics with special reference to transfer and exchange of genetic material.

Study Guide

We advise you to go through again the 'Study Guide' provided in Block 1 of this course before you begin the study of this block. In this block there is the transition from transmission to molecular and microbial genetics. We would like to mention that the study of this block as the first one requires your full attention and concentration so that you can comprehend and assimilate the information as presented by us. While learning, if any of the fundamental concepts are not clear to you it would be useful to consult the NCERT Biology text books of X, XI, XII standards. And if you wish to explore the topics further, we invite your attention to the books listed under 'Further Reading' at the end of the block.

As in the previous block there is a feedback form at the end of this block too. We expect that you would fill it up and mail it to us. Your feedback on this block would help us know your opinion and suggestions, if any, relating to the study material.

Objectives

After studying this block you should be able to :

- discuss extra–nuclear genetic systems present in chloroplast and mitochondria of eukaryotic cell and their cooperation with nuclear genetic system;
- describe the techniques employed to study the various aspects of human chromosomes;
- describe the normal human chromosome at morphological as well as molecular level;
- recognise and explain the various kinds of structural chromosomal anomalies in humans and their resultant effects;
- recognise and explain the different forms of numerical chromosomal anomalies in humans and their resultant effects;
- discuss the nature and structure of genetic material;
- describe the mechanisms of recombination and complementation in a bacteriophage;
- list the various bacterial genetic transfer mechanisms and describe them briefly.

UNIT 7 EXTRA-NUCLEAR INHERITANCE

Structure

7.1 INTRODUCTION

In the earlier units on Mendel's laws of inheritance, you have learnt that genes obey the basic laws formulated by Mendel. Genes which follow such a pattern, are said to be governed by mendelian inheritance.

The notion that genes reside in the nucleus of eukaryotic organisms has become a firm part of general dogma of genetics. In fact, the mechanics of the nuclear processes of meiosis and mitosis, together with the understanding that genes are located on chromosomes, provide the basic set of operational rules for genetics. However, like most fundamental truths, this one does have exceptions and these exceptions form the basic frame-work of this unit. In eukaryotes, the special inheritance patterns of some genes reveal that these genes are located outside the nucleus. Such an inheritance called the extra-nuclear, extrachromosomal, cytoplasmic, uniparental of maternal inheritance, has been examplified in a variety of eukaryotic systems. We would mostly refer to this kind of inheritance in this unit as extra-nuclear inheritance.

The phenomenon of extra-chromosomal inheritance operates in all the eukaryotic organisms. Since there is invariably a preference in favour of the organisms which are easily amenable to experiments, certain plant species, the green alga, *Chlamydomonas* and fungi, *Neurospora* and *Saccharomyces* (yeast) have served good experimental systems for this type of inheritance.

In the first section of this unit we will discuss the basic nature of non-mendelian inheritance and mechanism of its transmission to the next generation. We would also see that not all maternal effects are due to non-mendelian inheritance. The non-mendelian inheritance is due to the genes residing on cytoplasmic organelles, chloroplasts and mitochondria. You will be introduced to certain well studied cases of such an inheritance controlled either by mitochondria or chloroplasts. You would see how inferences drawn from various experiments lead to certain concepts. Finally, we would also discuss the interdependence of nuclear and cytoplasmic genetic systems.

Objectives

After studying this unit you will be able to:

- distinguish between mendelian (nuclear) and non-mendelian (extra-nuclear) inheritance using the phenotypic outcome of reciprocal crosses,
- discuss the involvement of chloroplast and mitochondria in extra-nuclear inheritance,
- explain the nature of extra-nuclear inheritance,
- give evidences that genes are present in mitochondria and chloroplasts,
- appreciate the complexity of genetic systems residing in various cellular compartments of a cell and
- discuss the possible origin of organelles involved in extra-nuclear inheritance.

7.2 EXTRA-NUCLEAR INHERITANCE

Let's recall that genes in the case of diploid organisms exist in pairs and two members or alternative forms of a single gene are called alleles. The alleles are located at the same position (locus) on homologous chromosomes. The Law of segregation, as you have studied earlier, deals with the separation of the alleles at the time of gamete formation. In terms of chromosomes, it implies the separation of two homologues from each other during the first meiotic division. During the cross and in the formation of F_1 heterozygote, one allele is contributed by the female parent and the other comes from the male parent. The outcome of the cross in terms of phenotype of the F_1 is identical irrespective of which of the alleles, recessive or dominant comes from the female parent. Therefore, we say that in a typical mendelian inheritance, the phenotypic outcome of reciprocal crosses is identical as shown in Fig. 7.1.

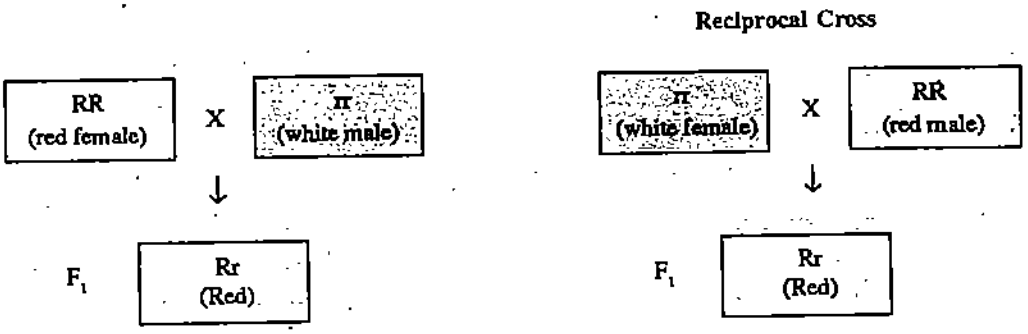


Fig. 7.1: Typical mendelian pattern of inheritance in reciprocal crosses.

The implication of this observation is that there are no differences whether gamete (R) is contributed by the female parent or by the male parent. The genetic contribution of the two parents is identical, resulting is the same phenotype in reciprocal crosses. However, there are exceptions to this general rule.

It has been observed in certain cases that the phenotype of F_1 in the reciprocal crosses is different. When phenotypes of the reciprocal crosses are not identical, it is the first indication that the inheritance of that particular character may not follow mendelian pattern. Such a phenotype is invariably influenced by the maternal parent. Broadly, this pattern of inheritance is termed non-mendelian inheritance.

7.2.1 Early Experiments

Soon after the rediscovery of Mendel's laws of inheritance in the beginning of this century, the results of various experiments were not in conformity with what was to be expected on the basis of nuclear inheritance. We will discuss some experiments which provided proof for the existence of extra-nuclear mode of inheritance.

Variegated Plants

One of the first convincing examples of extra-nuclear inheritance was found in higher plants. In 1909, a scientist named Carl Correns observed some strange variegation in Four O'clock (*Mirabilis jalapa*) plant. It had branches containing patches of both green and white tissues but some branches carried green leaves and others only white leaves. Fig. 7.2 shows such a variegated plant. In fact, you must have seen many such variegated plants in your surroundings.

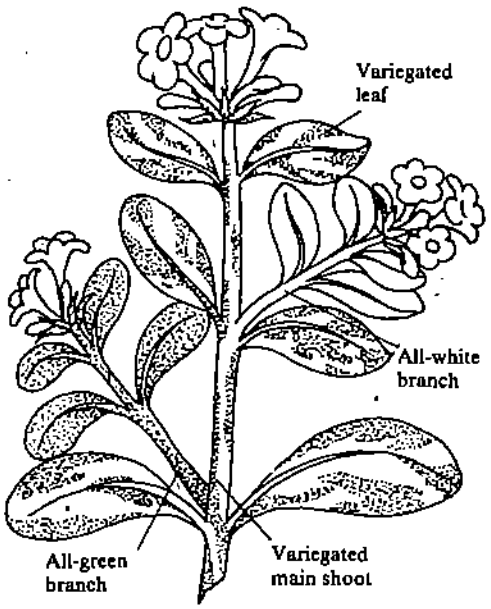


Fig. 7.2: Four O'clock (*Mirabilis jalapa*) plant.

Flowers appeared on all types of branches. He intercrossed different combinations by transferring pollen from one flower to another and obtained the following results (Table 7.1).

Table 7.1: Results of the crosses of different phenotypes in *Mirabilis jalapa*.

Phenotype of branch bearing egg parent (female)	Phenotype of branch bearing pollen parent (male)	Phenotype of progeny
White	White	White
White	Green	White
White	Variegated	White
Green	White	Green
Green	Variegated	Green
Green	Green	Green
Variegated	White	Variegated, green or white
Variegated	Green	Variegated, green or white
Variegated	Variegated	Variegated, green or white

Two basic features are obvious from these results.

(i) There is a difference between reciprocal crosses.

(ii) The phenotype of the maternal parent is solely responsible for determining the phenotype of all progeny. The contribution of the male parent appears to be nil.

The results are summarized in Fig 7.3.

Since there is no contribution of the male parent, this phenomenon has also been termed maternal inheritance. White plants lacking chlorophyll do not survive but other types i.e. variegated and green survive and can further be tested. This type of pattern is not transitional but keeps on persisting in the subsequent generations. How can such results be explained ? You know that the green leaf colour is due to the presence of chloroplasts. The colourless sectors or branches contain colourless plastids and

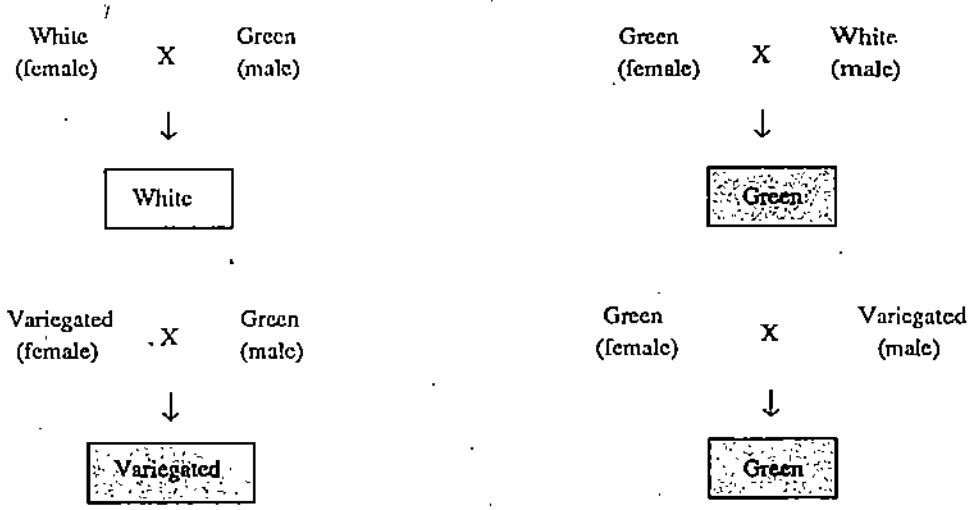


Fig .7.3: Non-mendelian pattern of inheritance in the reciprocal crosses of *Mirabilis jalapa.*

variegated sectors will have green as well as colourless plastids. The inheritance pattern can easily be explained with the diagram shown in Fig. 7.4, if chloroplasts contain

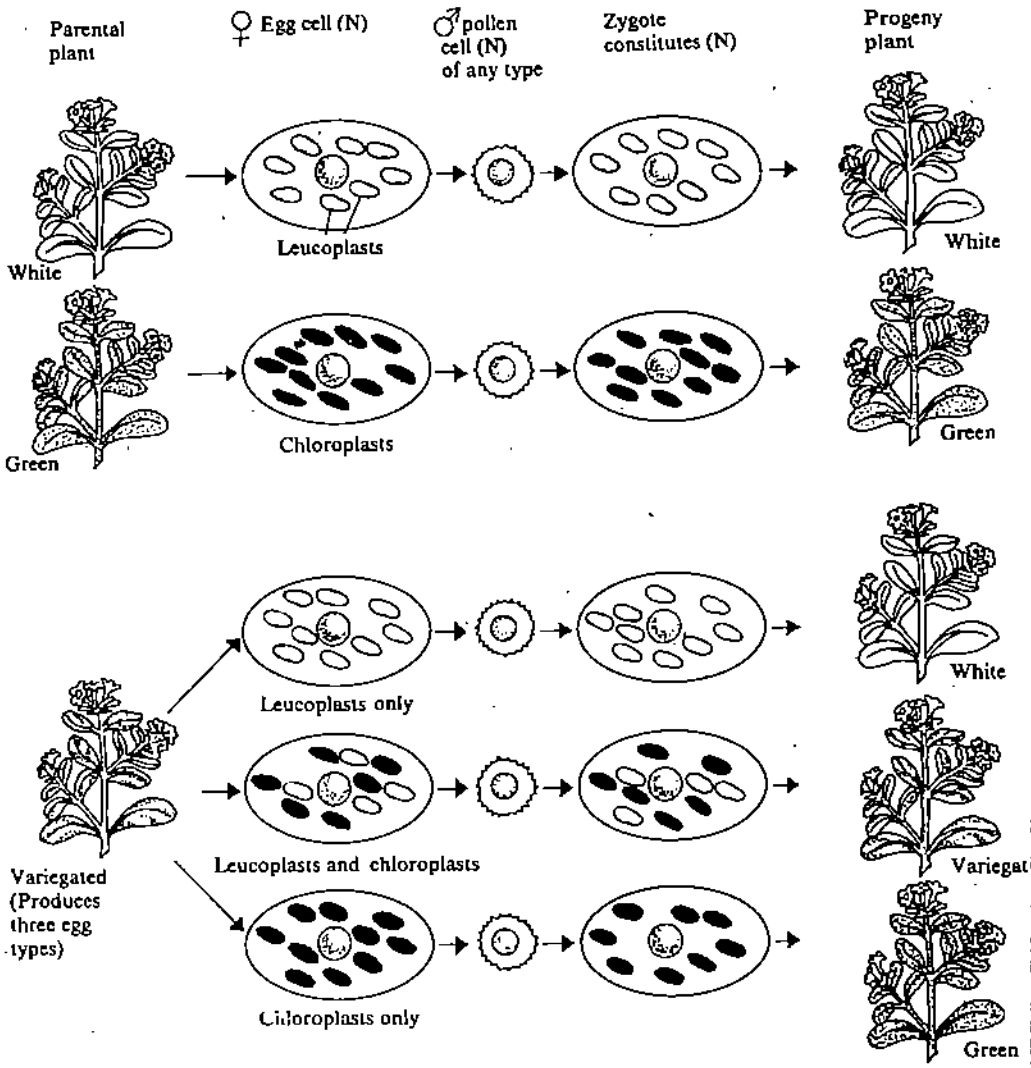
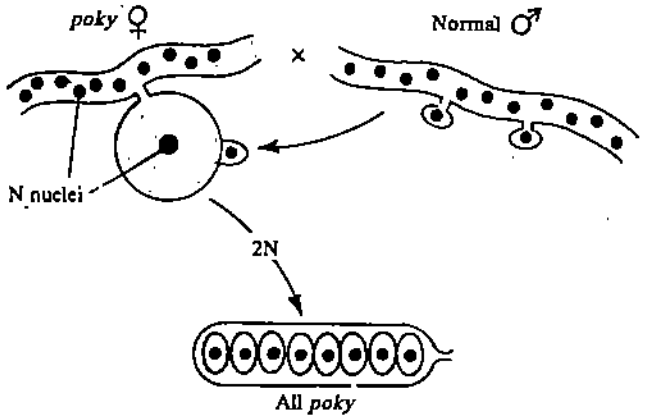


Fig. 7.4: A model explaining the results of *Mirabilis jalapa* crosses in terms of autonomous chloroplast inheritance. The first two crosses exhibit strict maternal inheritance. If the maternal branch is variegated, three types of zygotes can result, depending on whether the egg cell contains only white, only green, or both green and white chloroplasts. In the last case, the resulting zygote can produce both green and white tissue, so a variegated plant results.

genetic determinants and pollens do not contribute any chloroplast to the zygote. It is reasonable to assume that the pollen parent does not transmit any organelles since bulk of the cytoplasm of the zygote is known to come from the maternal parent via cytoplasm of the egg. The green branches produce eggs with only green chloroplasts and white branches produce eggs with colourless chloroplasts. Variegated branches apparently produce three kinds of eggs, some contain only white chloroplasts, some contain only green chloroplasts and some contain both kinds of chloroplasts. The egg containing both types of chloroplasts will produce a zygote containing both the types of chloroplasts. In subsequent mitotic divisions, chloroplast sorting occurs, whereby some cells contain only green chloroplasts and others colourless ones, resulting in variegated phenotype in the progeny individuals. Following these observations, various other cases of this type of inheritance have been observed in a variety of other plant species.

Poky Mutants

In 1952, a mutant strain of fungus, *Neurospora* called poky was identified. The poky mutant is slow growing and has abnormal amounts of cytochromes. You have already learnt (Unit 11, LSE-01) that cytochromes are mitochondrial electron transport proteins necessary for oxidative phosphorylation. In the poky strains, of the three main types of cytochromes, cytochrome *a* or *b* are absent but an excess of cytochrome *c* is present, *Neurospora* is a haploid fungus, but a eukaryote. It has two different mating types and the strains with differing mating types can be crossed. The crosses of poky to wild type produced the following results:



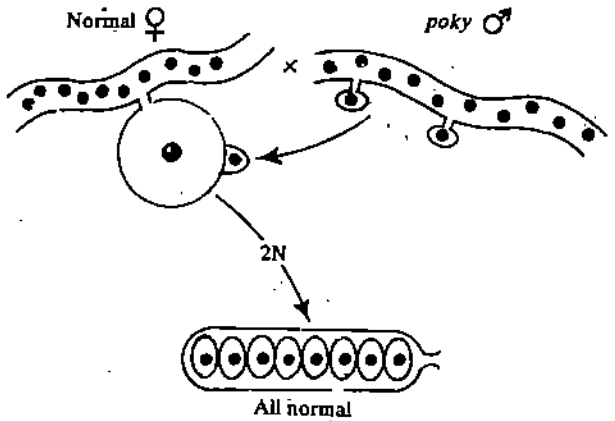


Fig. 7.5 : Results of reciprocal crosses of poky and wild *Neurospora*. (a) poky ♀ x wild type ♂ (b) wild type ♀ x poky ♂.

These results again show that there are differences in reciprocal crosses and the phenotype of the F_1 is solely controlled by the maternal parent.

SAQ 1

Which of the following statements are correct.

i) The mode of inheritance of extranuclear genetic system is distinct from mendelian inheritance. ☐

ii) The phenotypes of reciprocal crosses are identical in extra-nuclear inheritance. ☐

iii) In cytoplasmic inheritance the phenotype of reciprocal crosses are invariably influenced by the male parent. ☐

iv) Extra-chromosomal inheritance is recognised by uniparental (usually maternal) transmission through a cross. ☐

v) In cytoplasmic inheritance the egg cell would be expected to influence non-mendelian traits. ☐

7.3 MATERNAL VS. EXTRA-NUCLEAR INHERITANCE

Usually but not always, it is possible for a maternal inheritance pattern of reciprocal crosses to be due to nuclear genes or due to mendelian inheritance. One of the prominent example in this category is the direction of shell coiling in the water snail *Limnaea*. The coiling of the shell in snails, is of two types. Some snails coil to the right (dextral) while others coil to the left (sinistral coiling). All the F_1 progeny of the dextral female and sinistral male are dextral, but all the progeny of the cross between sinistral female and dextral male are sinistral (Fig. 7.6).

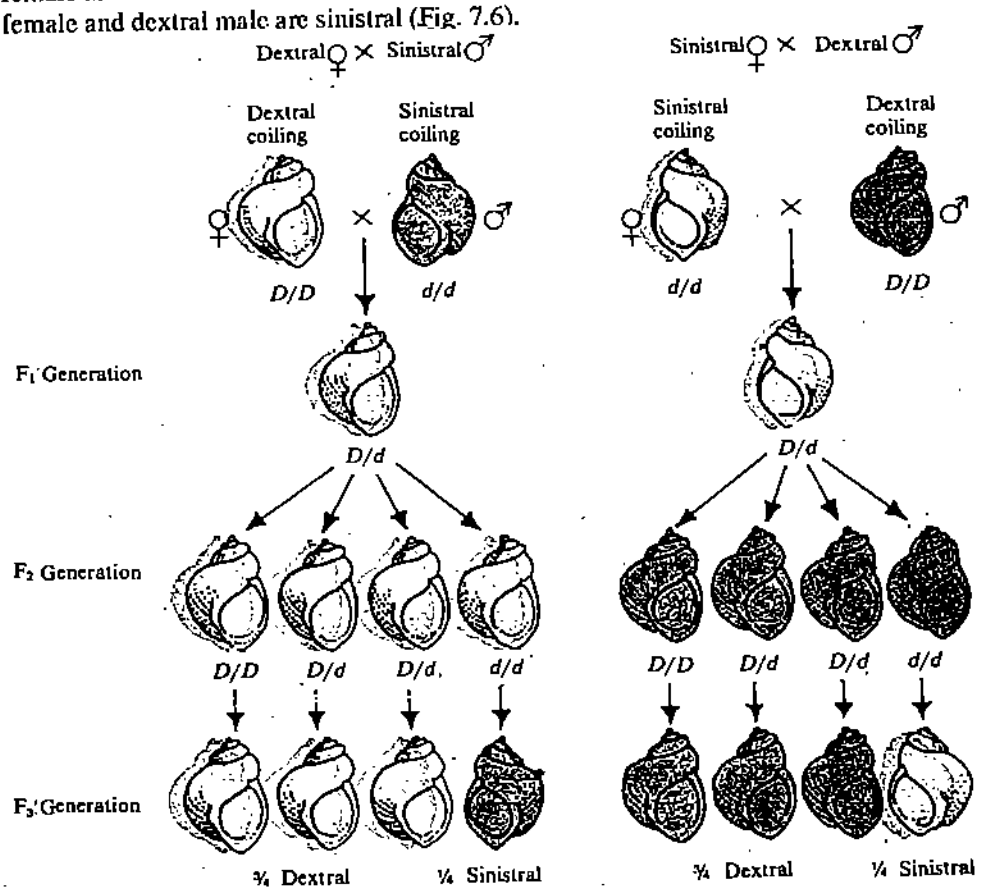


Fig. 7.6: The inheritance of dextral (D) and sinistral (d) nuclear genes for shell coiling in a species of water snail. The direction of coiling is determined by the nuclear genotype of the mother, not the genotype of the individual involved.

This appears similar to the observations made in the case of chloroplast inheritance or with the one observed with the poky mutant in *Neurospora*. However, in the F_2

generation the shell coiling is all dextral from both the pedigrees. The F_3 generation in this case reveals that inheritance of the direction of coiling is controlled by nuclear genes rather than extranuclear genes. It was found that 3/4th of F_3 are dextral and 1/4th are sinistral. The ratio reveals a mendelian segregation, 3:1 of the F_2 generation. Apparently dextral is dominant over sinistral. It was found that the coiling pattern is conditioned by a biochemical compound present in the maternal cytoplasm. Subsequently, this biochemical compound gets diluted due to somatic divisions, and the new compound is synthesised depending on the genes present in the female parent. It is apparent from the foregoing example that whether a maternal inheritance is controlled by extra-chromosomal genes, can only be confirmed if the inheritance pattern persists in subsequent generations.

7.4 SYSTEMS OF EXTRA-NUCLEAR INHERITANCE

We have seen earlier that genes residing in cytoplasm are responsible for extra-nuclear inheritance. The experiments with Four O'clock plants were conclusive enough to implicate chloroplasts in the inheritance of variegation. Similarly, the experiments with poky mutant in *Neurospora* show that the defect is in one of the components of mitochondria. When we talk about genes or genetic material we invariably imply DNA (deoxyribonucleic acid). You already know that DNA is the store-house of genetic information. Whatever information an organism requires to develop from a single-celled zygote to an adult is encoded in its DNA.

The experiments with plants and *Neurospora* and also with *Chlamydomonas* and yeast, which we will see in the later sections, clearly established the following:

(i) Certain characters are not governed by the principle of mendelian inheritance but follow extra-nuclear inheritance and

(ii) In the case of *Mirabilis*, chloroplasts also determine the inheritance pattern, and in the case of *Neurospora*, the inheritance is also controlled by mitochondria. Thus, the extra-nuclear inheritance implies the presence of genetic determinants in chloroplasts, mitochondria or both. Since the genetic determinants or genes invariably mean DNA (in certain cases RNA) the search continued for showing the presence of DNA in these organelles. Using both electron microscopic and biochemical studies it was established that both of these organelles contain DNA and also possess the machinery necessary to decode the information contained in their DNA.

7.4.1 Chloroplasts

In 1950's Ruth Sager isolated a mutant of *Chlamydomonas reinhardii* which cannot grow in the presence of an antibiotic, streptomycin. *Chlamydomonas* is a fresh water unicellular eukaryotic alga. It does not have different sexes but has mating types (mt^+ and mt^-). The haploid cells under conditions of nitrogen starvation behave like gametes and gametic fusion between dissimilar mating types occurs. The diploid cells after fusion, undergo meiosis and release four haploid cells (tetrads) as shown in Fig. 7.7.

When a strain, which can grow in the presence of streptomycin (sm^r), was crossed (Fig. 7.8) with a strain which cannot grow in the presence of streptomycin (sm^s), the following pattern of inheritance was observed:

$$sm^r\ mt^+ \times sm^s\ mt^- \rightarrow \text{all } sm^r$$

$$sm^s\ mt^+ \times sm^r\ mt^- \rightarrow \text{all } sm^s$$

We can draw two kinds of inferences from these crosses with respect to the inheritance of the *sm* gene:

(i) There are differences in reciprocal crosses and

(ii) the phenotype of F_1 is determined by mt^+ parent. It is similar to the pattern expected for a maternal inheritance and it is a case of uniparental inheritance. Using the analogy of higher system mt^+ is referred to as the female and mt^- as male. The mating type genes, mt^+ and mt^- segregate in 1:1 ratio, as expected for true mendelian inheritance.

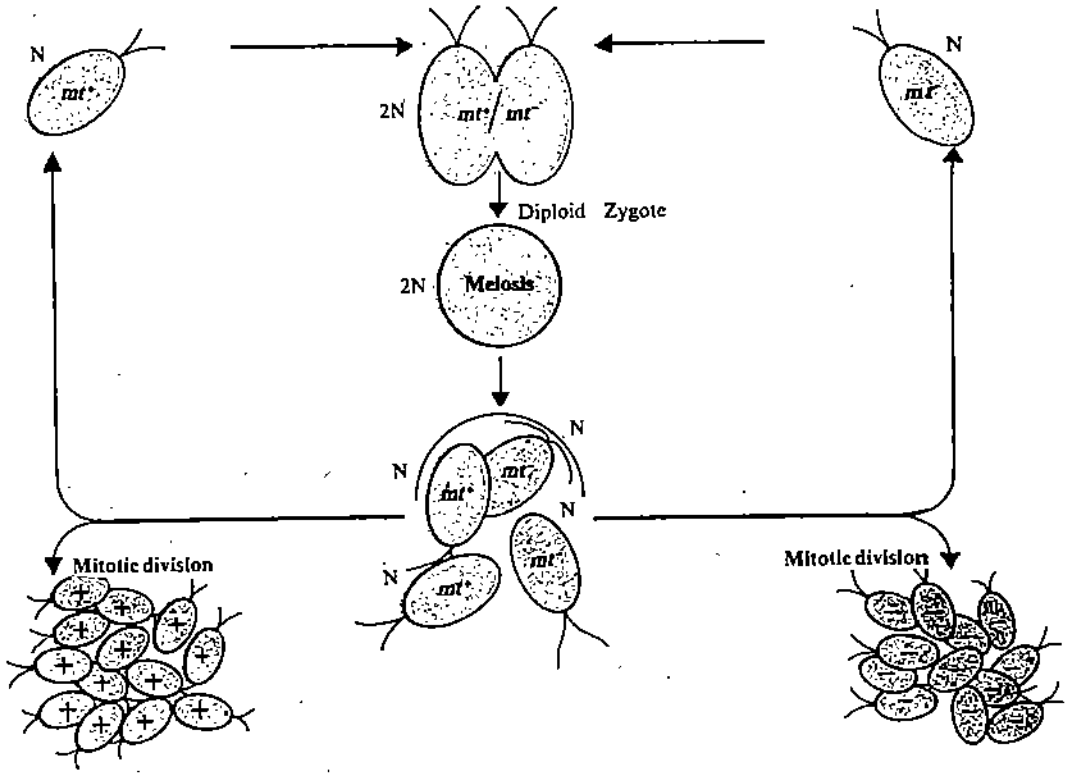


Fig. 7.7: Diagrammatic representation of the life cycle of *Chlamydomonas*. The nuclear mating-type alleles mt^+ and mt^- must be heterozygous in order for the sexual cycle to occur.

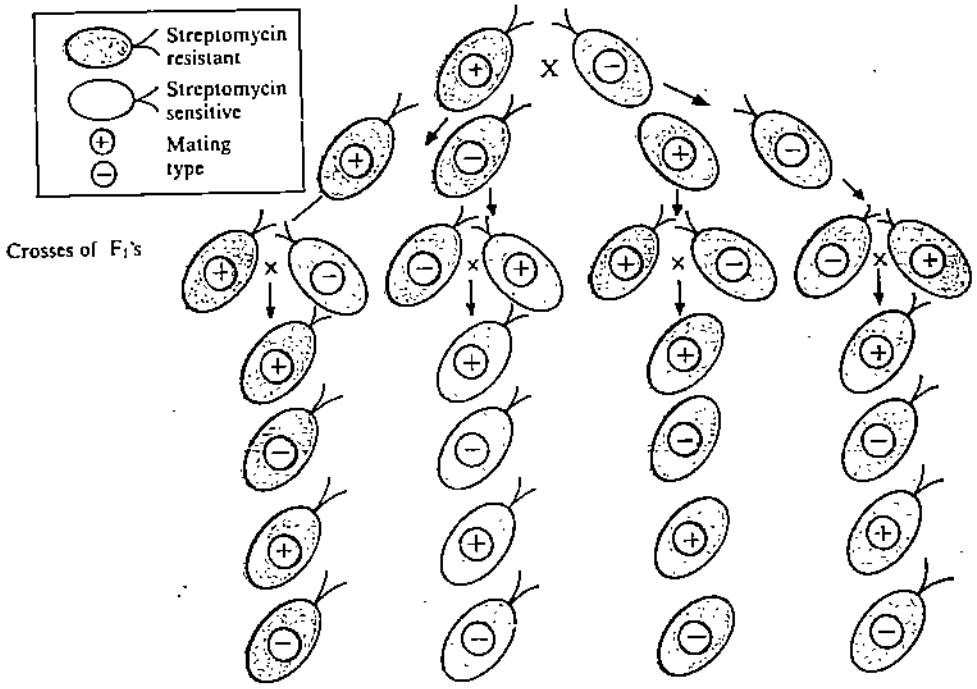


Fig. 7.8: Inheritance of streptomycin resistance in *Chlamydomonas*.

In the case of higher organisms, where a zygote is formed by fusion of egg with sperm, the cytoplasmic contribution of the male gamete is almost negligible. Under such conditions, it is easy to comprehend the mechanism of maternal inheritance. In case of *Chlamydomonas*, the cytoplasmic contribution of the two gametes, mt^+ and mt^- is almost identical. However, a uniparental inheritance is still observed and the cytoplasmic contribution of the mt^- parent, somehow has to be eliminated from the progeny. Therefore, a basic question arises, what happens to the cytoplasmic determinants of mt^- gametes? This puzzle was solved by Ruth Sager, who discovered that on gametic fusion, mt^- chloroplast DNA gets degraded. The mt^+ gene or a gene very closely linked to it, specifies a restriction-modification system. Restriction, in simplicistic terms, means degradation and modification implies protection. The system,

encoding the DNA modifying enzyme, modifies its own DNA which cannot be degraded by the restriction system. However, the *mt* $^-$ chloroplast which is not modified and thus not protected, is degraded by the restriction system specified by the *mt* $^+$ cells. It has been shown that an *mt* $^+$ linked gene encodes an endonuclease (enzyme which degrades DNA) which can differentiate DNA of its own cell from that of the *mt* $^-$ cell due to modification. Since the *mt* $^-$ DNA does not survive after fusion of *mt* $^-$ cell with *mt* $^+$ cells, no expression of *mt* $^-$ chloroplast DNA is observed, leading to the uniparental pattern of inheritance. Therefore, nature in this case has conspired to make *mt* $^+$ cells mimic the female gametic functions. However, in a fraction (1 in 1000) of cases, *mt* $^-$ chloroplast DNA does escape the degradation by the *mt* $^+$ restriction system. Such biparental zygotes containing the chloroplast DNAs from both cells are called **cytohets** (cytoplasmic heterozygotes). The cytohets, in fact, are very useful in studying the recombination of cytoplasmic genes. Ruth Sager by using the analysis of cytohets constructed a genetic map of *Chlamydomonas* chloroplast DNA.

SAQ 2

Fill in the blank spaces in the following statements by appropriate words.

i) The genetic determinants of maternal inheritance are present in ..... and ..... .

ii) The biparental zygote containing chloroplast DNAs of both parents in *Chlamydomonas* is called ..... .

iii) In *Chlamydomonas* crosses, the chloroplast DNA of the *mt* $^+$ strain survives because it has a gene for restriction modification enzyme to ..... its own *cp*DNA from ..... .

iv) The DNA fibrils in chloroplast have been observed by ..... studies and have also been ..... and characterised by biochemical studies.

7.4.2 Mitochondria

In support of extra-nuclear inheritance, we have already discussed the involvement of mitochondria in the poky mutants in *Neurospora*. However, the best studied examples of such a pattern of inheritance come from genetic as well as molecular studies of mitochondrial genetic system of baker's yeast, *Saccharomyces cerevisiae*. Yeast is a haploid organism and has two kinds of mating type alleles, α and *a*. The gametes carrying different mating type alleles, α and *a* fuse leading to the formation of a zygote.

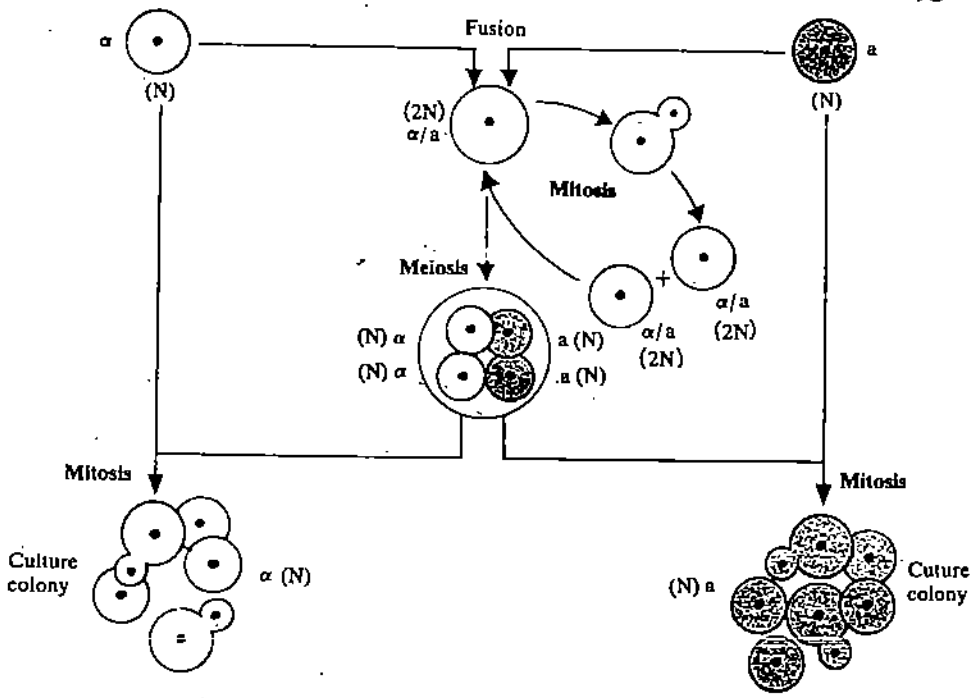


Fig. 7.9: The life cycle of baker's yeast (*Saccharomyces cerevisiae*). See text for details.

The diploid zygotes undergo meiosis and results in the formation of four ascospores. The ascospores germinate and form haploid yeast cells. The life cycle of yeast is given in Fig. 7.9.

Yeast cells normally form large colonies on agar plates which are called '**grande**'. Occasionally minute colonies, called '**petite**' (meaning small) also appear amongst them. When petites were crossed to grande strains, following three different kinds of results, as outlined in Fig. 7.10 were obtained:

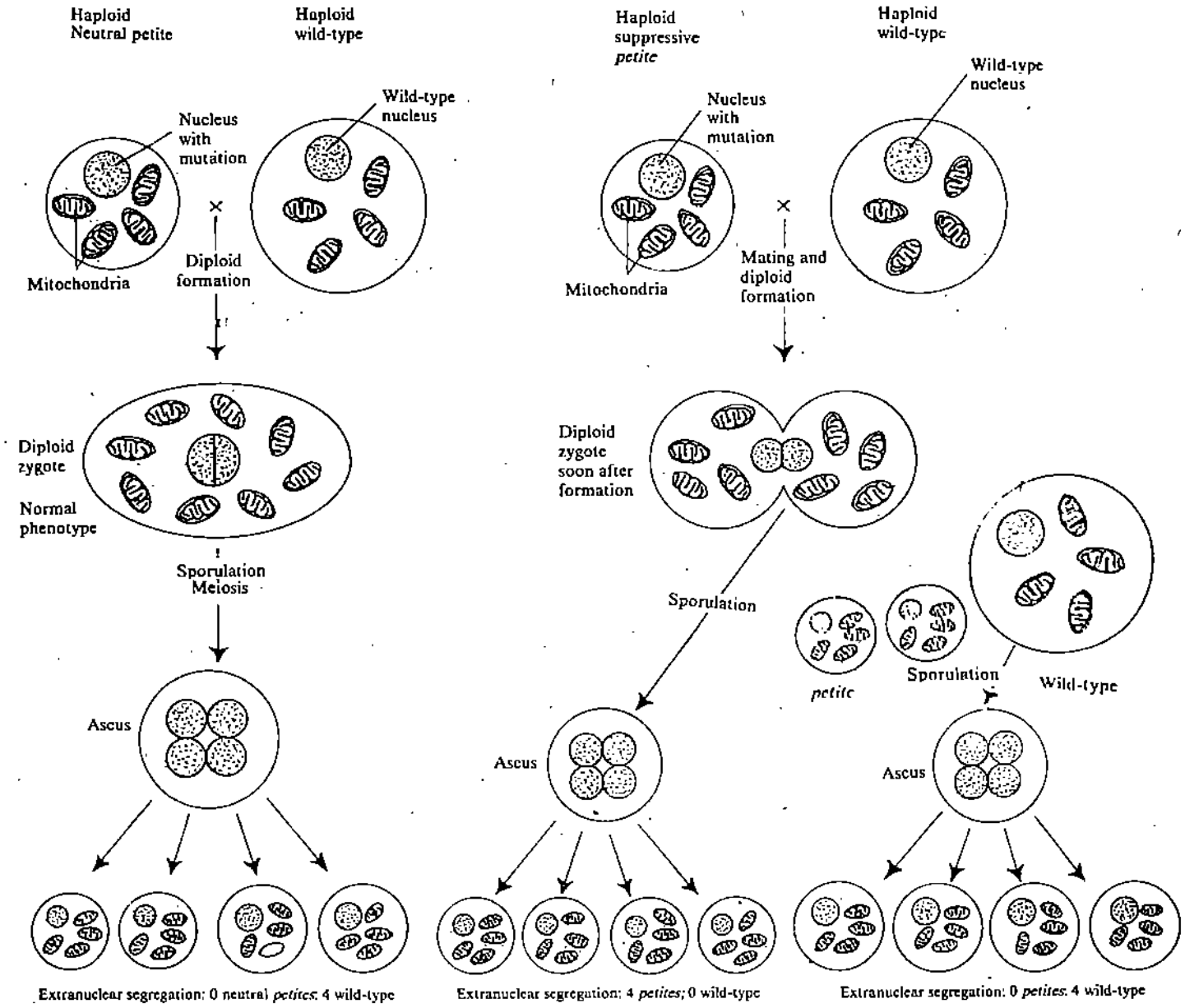


Fig. 7.10: ***Petite*** **mutations in genetic crosses. The nuclear** ***petites*** **(a) show mendelian segregation; the neutral** ***petites*** **(b) show 4+ : 0** ***petite*** **non-mendelian segregation; the highly suppressive** ***petites*** **(c) show 0 + : 4** ***petite*** **non-mendelian segregation.**

If the inheritance of a nuclear gene, such as the mating type locus is followed, the progeny ascospores segregate in the expected 2:2 ratio. The similar results (2:2 ratio with respect to grande and petite phenotype) were obtained in the first category of crosses, indicating that this petite phenotype is controlled by nuclear gene. Such petites governed by the mendelian inheritance are called **nuclear** or **segregational petites**. Therefore, we are not going to discuss them in this Unit.

The other two crosses produce 0:4 or 4:0 with respect to grande and petite phenotypes indicating the extra-nuclear pattern of inheritance. In one case, all the progeny is normal

(grande) and in the other, all are petites. In the former case, it appears that the more functional mitochondria take over resulting in all the normal progeny. The petites falling into this category are called **neutral petites**. The last category of cross, petite x grande producing all petite progeny, are called **suppressive petites**. Since the inheritance pattern of neutral and suppressive petites is extra-nuclear, they are also referred to as cytoplasmic petites. In a yeast cross, the two parental haploid cells fuse and apparently contribute equally to the cytoplasm of the resulting diploid cell. In extra-nuclear inheritance in yeast, the phenotype of the F_1 is not dependent on a particular mating type. In this respect yeast is clearly different from *Chlamydomonas*.

Several properties of the cytoplasmic petites which are listed below point to the involvement of mitochondria in their phenotype.

(i) In cytoplasmic petites, the mitochondrial electron transport chain is defective. Therefore, the petites have to rely on the less efficient process of fermentation for providing ATP. If they are placed in a medium containing non-fermentable energy-source such as glycerol, they do not grow.

(ii) Mitochondria normally possess their unique protein-synthesizing apparatus, which is different from the non-organellar or cytosolic system. The cytoplasmic petites are, therefore, defective in protein synthesis.

(iii) The suppressive petites have altered mitochondrial DNA (*mt*DNA) and the neutral petites contain no *mt*DNA at all.

During the derivation of petite mutants from grande strains, certain other mitochondrial genes are also lost. The coincidental or simultaneous loss of several determinants is indicative of deletion. This led to a proposition that petite phenotype may possibly be a consequence of deletion in mitochondrial DNA (*mt*DNA). It is now established that the petite phenotype in fact results from deletion of *mt*DNA. However, the size of *mt*DNA isolated from petite mitochondria was found to be similar to that of the *mt*DNA isolated from normal mitochondria (~78kb). Moreover, different petites have been shown to contain different genetic determinants indicating the deletions of different regions in mtDNA. To compensate the deleted segments of mtDNA, the DNA retained by the petite is amplified through tandem duplication to provide a chromosome of approximately the normal length. This phenomenon of deletion-amplification can be explained diagrammatically (Fig. 7.11).

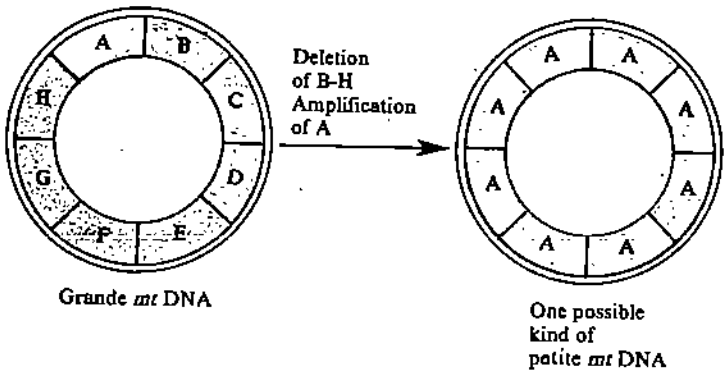


Fig. 7.11: When a petite is produced from a grande cell, a large region of *mt*DNA may be deleted. Apparently, the DNA region retained by the petite (in this example A) is amplified through tandem duplication to provide a chromosome of approximately the normal length.

Besides petites, several other mutations showing the pattern of extra-nuclear inheritance have been analysed in yeast. Prominent among them are (i) Mitochondrial mutations conferring resistance to various antibiotics such as chloramphenicol (cap^R), erythromycin (ery^R), spiromycin (spi^R), paramomycin (par^R) and oligomysin (oli^R). These mutants as a class are called antibiotic resistant (ant^R) mutants. (ii) The other category of mitochondrial mutants called *mit* $^-$ are defective in mitochondrial electron transport chain and they also produce small colonies. However, unlike petites they have normal protein synthesis and are able to revert back to normal. The reversion indicates that they resemble point mutations. The inheritance pattern of the *mit* $^-$ mutants is comparable to antibiotic mutants of *chlamydomonas*, mentioned earlier, showing uniparental inheritance at meiosis.

SAQ 3

In the following statements choose the alternate correct word given in the parenthesis.

i) Resistance to antibiotic in some strains of yeast is due to (chloroplast / mitochondrial / nuclear) genes.

ii) The *mit* $^-$ mutant strain of yeast shows (uniparental/biparental) inheritance.

7.5 CHARACTERISTICS OF EXTRA-NUCLEAR GENOME

So far we have discussed the concept of extra-nuclear inheritance derived from the study of pattern of inheritance in *Mirabilis*, *Chlamydomonas* and yeast. In this section, we will describe some of the studies which revealed the presence of DNA in chloroplasts and mitochondria. Overall gene organisation and the genetic maps of yeast and human and liverwort mitochondrial DNA (*mt*DNA) and of rice and liverwort chloroplast DNA (*cp*DNA) have already been constructed.

7.5.1 Chloroplasts

During late 60's and early 70's electron microscopic and biochemical techniques were used to study the DNA in chloroplasts. Electron microscopic studies using DNA specific strains revealed the existence of several areas containing DNA-like fibrils within chloroplasts. These areas are called **nucleoids**. When isolated chloroplasts are suspended in a hypotonic solution or water, they lyse releasing their DNA. Such a preparation of DNA can be visualized under an electron microscope (Fig. 7.12). Thus the presence of chloroplast genetic system in the form of their own DNA was unequivocally demonstrated.

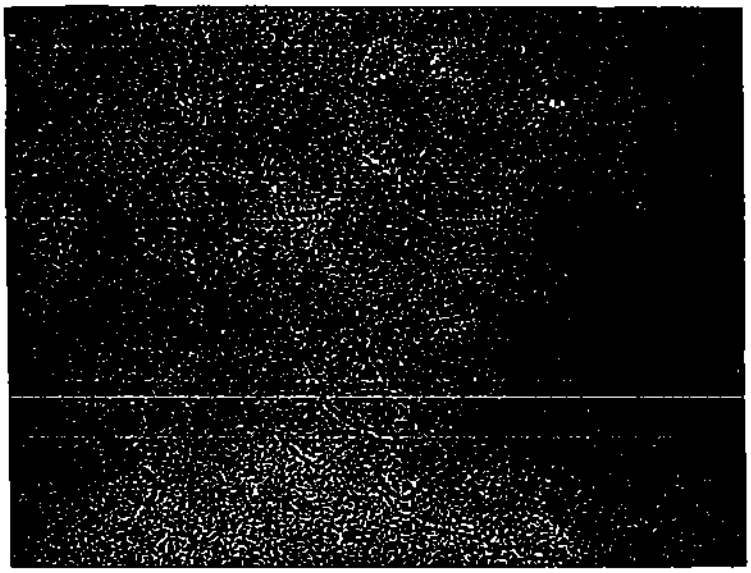


Fig. 7.12: EM picture of chloroplast DNA.

The number of copies of chloroplast DNA varies from species to species. Surprisingly, it can also vary within a single species. Each plant cell contains many chloroplasts, each chloroplast contains several nucleoids and each nucleoid contains 4-8 DNA molecules. For example, a single cell of the garden beet leaf contains about 40 chloroplasts and each chloroplast contains 4 to 18 nucleoids. Thus, each cell may contain as many as 5000 copies of chloroplast DNA molecules. Chlamydomonas has only one chloroplast per cell and the each chloroplast contains ~ 500–1500 chloroplast DNA molecules.

Chloroplast DNA of majority of plant species contain ~ 135 kb (one kilo base pairs = 1000 nucleotide pairs) of circular DNA, however, the size among the plant species ranges from 120 to 200 kb. In fact, we now know the complete nucleotide sequence of chloroplast DNA of two plant species, rice and liverwort (*Marchantia polymorpha*). By analysing the DNA sequence with the help of a computer, potential genetic information

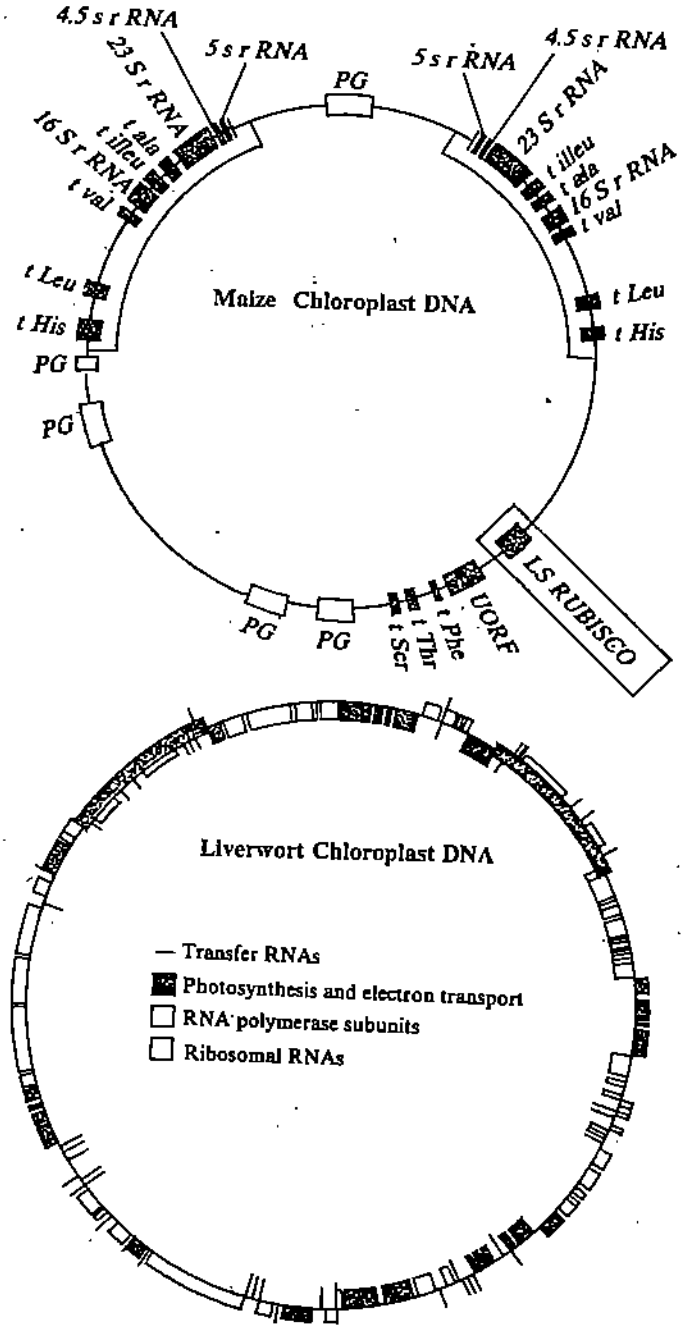


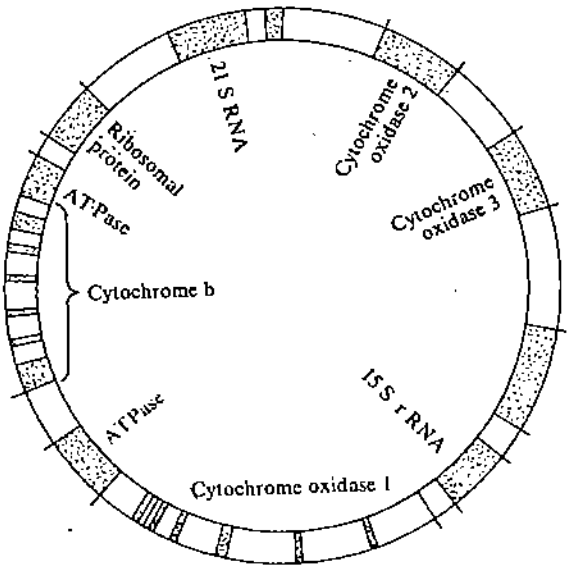
Fig. 7.13: Genetic map of liverwort and maize chloroplast DNA. In maize *cp*DNA note the position of two large inverted repeats at the top.

ncoded in the genome could be assessed. The liverwort chloroplast DNA which is 121 b long encodes for about 136 genes (Fig. 7.13). There are four kinds of rRNA, 31 inds of tRNA, and about 90 protein coding genes. Of the 90 protein coding genes, 20 ode for proteins involved in the process of photosynthesis. One half of the chloroplast enome is taken up by the genes coding for the translation functions required for rotein synthesis to occur in the organelle.

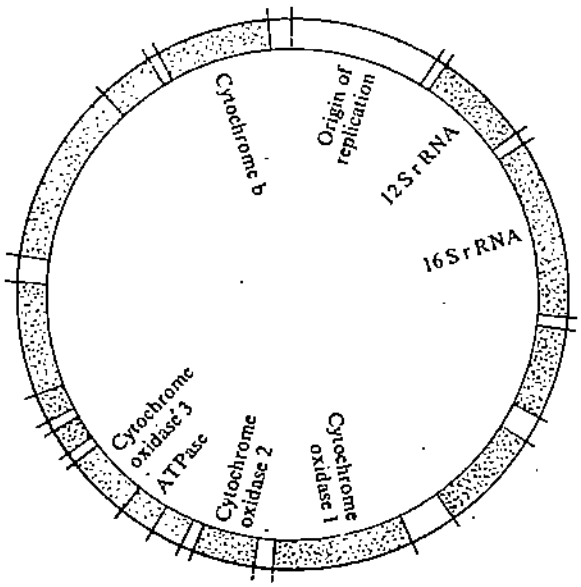
.5.2 Mitochondria

ıformation on molecular biology and genetics of mitochondrial DNA (*mt*DNA) has :en derived from yeast *mt*DNA (~ 78 kb) human *mt*DNA (~ 17 kb) or plant *t*DNA(> 180 – 1400 kb, depending on the plant species). A typical haploid yeast cell ıntains 1 to 45 mitochondria each having 10–30 nucleoids with four or five DNA

molecules in each nucleoid. The human *mt*DNA contains genes for cytochrome *b*, cytochrome oxidase I, cytochrome oxidase III, ATPase subunit 6, certain subunits of NADH dehydrogenase, large and small rRNAs, and various tRNA genes. Thus, the genetic map of human *mt*DNA is very compact. While the yeast *mt*DNA contains, besides the above mentioned genes, several other genes (Fig. 7.14). One surprising discovery in several yeast mitochondrial genes was the presence of introns in several odd mitochondrial genes. The genes for subunit I of cytochrome oxidase contains nine introns. It was much more surprising since introns are relatively rare in yeast nuclear genes. Another interesting aspect was the discovery that several yeast *mt*DNA introns contain genes encoding enzymes called maturases. Maturases have been shown to be involved in processing of mitochondrial RNA especially in the removal of intron-specific sequences.



Yeast Mitochondrial DNA



Human Mitochondrial DNA

Fig. 7.14: Physical maps of yeast and human mitochondrial chromosomes. The yeast chromosome is much larger than human one, but contains substantial amounts of repeated sequences.

7.6 ORGANELLES INHERITANCE DEPENDENT ON NUCLEAR GENOME

It is apparent from the foregoing discussion that animal and non-photosynthetic eukaryotic cells contain two genetic systems i.e. nuclear and mitochondrial; and plant and other eukaryotic photosynthetic cells contain three different genetic systems namely nuclear, chloroplast and mitochondrial. The presence of more than one genetic systems, each having its own distinct machinery for transcription and translation, requires an active cooperation and coordination between the genetic systems for various cellular functions. The systems of cooperation, as we will discuss in this section, is apparent in the synthesis and assembly of the chloroplast or mitochondrial proteins. We will describe the examples of organelle-proteins in which one or more subunits are encoded by nuclear genome and the others by the organelle genome.

7.6.1 Chloroplast Proteins

We have mentioned earlier that chloroplast DNA codes for several proteins which are involved in photosynthesis and also in translational machinery of chloroplasts. Some examples of the genes encoding the chloroplast proteins are large subunit of the enzyme, ribulose bisphosphate carboxylase (*rbc*), photosystem 1 (*psa*), photosystem 2 (*psb*), ATP synthethase (*atp*), RNA polymerase (*rpo*), permease (*mbpx*), 30 S and 50 S ribosomal subunit proteins (*rps* and *rpl*) and ribosomal and various transfer RNA genes. Nuclear genes also code for many of the chloroplast proteins or subunits of some proteins involved both in photosynthesis. The mRNAs of these proteins is translated on cytosolic ribosomes (80 S) and these proteins are then transported to chloroplasts. The cooperation of nuclear and chloroplast genes, each encoding a subunit of a single protein is well documented by the study of most abundant enzyme, rubisco (ribulose bisphosphate carboxylase–RuBP carboxylase) an important enzyme of the Calvin cycle. The enzyme is composed of two subunits: large and small. The large subunit is encoded by chloroplast DNA and synthesised on chloroplast ribosomes (70 S) while the small subunit encoded by nuclear genes is synthesised in cytosol and is then transported into chloroplasts, where it assembles with the large subunit to produce a functional RuBP carboxylase.

7.6.2 Mitochondrial Proteins

Overall function of the mitochondrial genome shows that it codes for some proteins that are actually involved or associated with the electron transfer chain, and all tRNAs and rRNA's necessary for mitochondrial protein synthesis. As you know that the three main components of the electron transfer chain of mitochondria—cytochrome *c* oxidase, cytochrome *b* and ATPase complex are located in its inner membrane. For these three functional units, *mt*DNA supplies six proteins, three subunits of cytochrome *c* oxidase, two subunits of A ! cytochrome *b*. The remaining necessary components are encoded by nuclear genes and the products are transported to mitochondria where they assemble into functional complexes.

It is apparent that all the genes required to fulfil either chloroplast or mitochondrial functions are not present in these organelles. The cytosolically synthesised subunits contain specific signal sequences which direct their transport to either of the organelles, where the signal peptides are cleaved off and the subunits are assembled to form a functional protein. The photosynthetic eukaryotes have three genetic systems i.e. mitochondrial, chloroplasts and nuclear and other eukaryotes have two i.e. mitochondrial and nuclear. It is clear from the foregoing discussion that in spite of the existence of a compartmentalisation in a eukaryotic cell, a degree of interdependence and communication exists between the genetic systems present in the cell. Therefore, the chloroplasts and mitochondrial systems cannot be considered autonomous. With the same logic and considering limited genetic potential of the organelles they cannot possibly survive outside the cellular system, even though they have their own distinct genetic system and protein synthesising machinery.

7.7 ORGANELLE-ASSOCIATED LINEAR PLASMID DNA

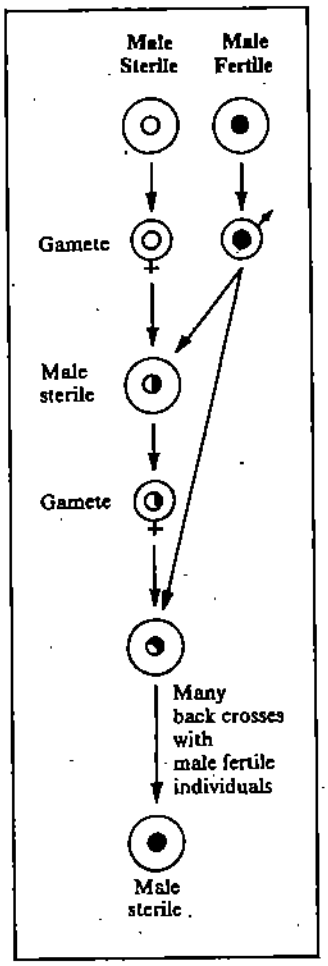


Fig. 7.15: Maternal inheritance of male sterility in maize.

Plasmids are autonomously replicating extra-chromosomal DNA molecules commonly occurring in bacteria. Occasionally they have also been found in eukaryotes. For example certain strains of yeast, *Saccharomyces cerevisiae* contain 2 micron (μ) plasmid circles which are present in nucleoplasm. It has been found that in certain cases, plasmid DNA is also associated with *mt*DNA. Two interesting cases of plasmid DNA associated with *mt*DNA are (i) Cytoplasmic male sterility in maize and (ii) Senescence (aging) in the fungus, *Neurospora*.

(i) Male sterility in plants has a great economic importance in agriculture, since male sterile plants eliminate the necessity of emasculation (removal of anthers) for producing hybird plants. A male sterile plant does not produce functional pollen. If a male sterile plant is used as a female parent in a cross with a normal fertile plant as a pollen parent, all the progeny is male sterile (Fig. 7.15). Thus, the inheritance pattern is exclusively maternal, indicating that the determinants of male sterility are located in organelles. It has been found that male sterile lines of maize have two linear plasmids, S1 and S2 within mitochondria. The presence of these plasmids makes the plants male sterile. They are also found to recombine with *mt*DNA. The exact role of S1 and S2 in male sterility is not yet fully understood.

(ii) Most Neurospora strains do not senesce and if nutrients in the medium are not limiting they will keep on growing for ever. However, a Hawaiian strain called "Kalilo" does senesce. Kalilo strain will grow only certain distance through the culture medium before it dies. This senescent property is maternally inherited or shows extra-nuclear inheritance. It has been established now that senescence is determined by a linear plasmid called *Kal* DNA. This plasmid exists autonomously and the onset of death is preceded by insertion of this plasmid DNA into *mt*DNA. At death, most *mt*DNA molecules contain full length plasmid DNA inserts. Presumably, the insertion of plasmid DNA into mitochondrial genome inactivates the genes affecting certain essential mitochondrial functions and thus results in cell death.

SAQ 4

a) In the following statements choose the alternate correct word given in parenthesis.

i) The DNA regions in chloroplast could be observed under (light microscope/electron microscope).

ii) The regions containing *cp*DNA are called (nucleoids/celluloids).

iii) Each nucleoid contains (a single/a few) copies of DNA.

iv) The *cp*DNA and *mt*DNA are usually (circular/linear) in nature.

v) Both chloroplast and mitochondria contain (a few/many) copies of DNA.

vi) The *mt*DNA of yeast is (bigger/smaller) than *mt*DNA of humans.

vii) The occurrence of introns is discovered in (yeast/human).

b) Which among the following statements are correct.

i) Mitochondria contain 80 S ribosomes. ☐

ii) The mRNA encoded by nuclear genes for the smaller subunit of rubisco is translated in the chloroplast. ☐

iii) The inheritance of chloroplast genome is independent of nuclear genome. ☐

iv) The chloroplast proteins or their subunits synthesised in the cytoplasm are transported to chloroplast. ☐

7.8 POSSIBLE ORIGIN OF ORGANELLES

We know that both chloroplasts and mitochondria perform certain essential functions for the cell. They contain their own distinct genetic system and protein synthesising apparatus. However, the genetic system, and transcription and translational machinery is similar to prokaryotes. You may recall that the eukaryotic ribosomes are of 80S whereas the prokaryotic ribosomes are of 70S type. The protein synthesis in both the organelles is similar to the prokaryotic system and can be inhibited by **chloramphenicol** whereas the cytoplasmic protein synthesis is inhibited by **cycloheximide**. The inhibition of protein synthesis by different sets of antibiotics, in fact, reflects the differences in the nature of ribosomal subunits present. The existence of genetic and protein synthetic machinery similar to prokaryotes within chloroplasts and mitochondria led to the question of their possible origin from a prokaryotic source i.e. the theory of endosymbiosis.

It has been widely conjectured that during pre-evolutionary time certain bacteria-like organism (cyanobacteria-like organism in the case of chloroplasts) invaded eukaryotic cell and established a symbiotic (endosymbiosis) like relationship within the cell. The chloroplasts conferred on the cell the photosynthetic ability and mitochondria endowed the cell with energy-transducing properties. The relationship was reinforced when the system became permanently interdependent, i.e. the cell cannot survive without mitochondria and/or chloroplasts and these organelles cannot survive ouside the cell as discussed earlier.

The above endosymbiotic theory has received a certain degree of experimental support from recent molecular studies. We have discussed in the earlier section that subunits of several enzymes of chloroplast and mitochondria are encoded by nuclear genes. These enzymes and other organelle protiens such as RNA polymerase and ribosomal protiens are very similar to bacteria or cyanobacteria. A certain degree of similarities in sequences, both at the protein and DNA levels, between organelles and bacteria exists. If chloroplast and mitochondria were indeed prokaryotes during a pre-evolutionary period, then they certainly have lost a considerable amount of genetic information. It appears that a massive transfer of genetic information from organelles to nuclei might have occurred during their long co-existence. The transfer would have resulted in certain modifications of prokaryotic genes in order for them to function in a eukaryotic manner.

Recent studies have shown that certain common DNA sequences are present between chloroplasts and nuclei; between chloroplasts and mitochondria; and between mitochondria and nuclei in plants. The results of these studies provide evidence for the inter-organellar movements of DNA sequences. Many of the such DNA sequences are not functional now but they certainly appear to have left their evolutionary footprints.

Other indications in support of the distinct origin of mitochondria are their unique genetic code and phenomenon of RNA editing. The genetic code is considered universal, however, certain variations exist in mitochondria. In fact, it is one of the exclusive property of mitochondria and it is the only exception to the universality of genetic code. A very recent discovery is related to RNA editing in plant mitochondria which has also been observed in certain parasites. After the synthesis of mRNA, certain nucleotides are edited and replaced by others in mitochondria.

The evidences discussed above provide support to the endosymbiotic theory. Basically many processes which are characteristics of prokaryotic systems and unique to the organellar system and not present in the nuclei support the distinct origin of the organelles.

7.9 SUMMARY

In this unit you have learnt that:

- The study of inheritance pattern of certain traits in plants and animals show non-conformity with Mendel's laws of heredity.

iv) cycloheximide

Prokaryotes, Chloroplast and Mitochondria:

i) circular
ii) 70 S
iii) 40 S and 30 S
iv) Chloroamphenicol

UNIT 8 HUMAN CHROMOSOMES

Structure

8.1 INTRODUCTION

While Mendel was investigating the mechanism of inheritance, a major area of interest in biology was the study of cell structure. The early microscopes were being rapidly refined, and in 1940, the 'cell theory' was proposed. Due to the transparent nature of cells, little advancement was made until 1870, when cell stains were developed. In 1879, Walther Flemming found that the nuclei of cells contained some material that at certain times took up stain and became very intensely – coloured. He named this material **chromatin** — a greek word meaning colour. In the early stages of nuclear division, chromatin appears as long, thread-like structures. In later stages, they appear as short, rod-like bodies. W. Waldeyer termed these bodies **chromosomes** (coloured bodies).

The number of chromosomes in each cell is characteristic of a species but the number of chromosomes in itself is not necessarily sufficient for species identification since several species may have the same number of chromosomes. For example, the red squirrel, the evening bat and humans all have forty – six chromosomes. Similarly, white oaks, rice, tomatoes, pines and grasshoppers all have twenty four chromosomes, yet they are clearly separate species. All the normal members of any species have the same number of chromosomes. Human beings, for example, irrespective of their age, sex, race, colour, size or geographical habitat, have a haploid number of twenty-three chromosomes. The haploid number of chromosomes for different species may vary from one in the nematode *Parascaris univalens* and in the ant *Myrmecia pilosula*, to 510 in the fern *Ophioglossum petiolatum*.

In 1902, two years after the rediscovery of Gregor Mendel's paper on inheritance, W.S. Sutton suggested that the transmission of Mendel's factors or genes might be associated with the behaviour of chromosomes during meiosis. Sutton's hypothesis was proved correct and it became known as the *Chromosome Theory of Heredity*. The theory proposed that genes are located in and remain with the chromosome during the formation of daughter cells (mitosis) or gametes (meiosis). In 1910, Thomas Morgan's work indicated that genes in chromosomes could be separated from each other by an

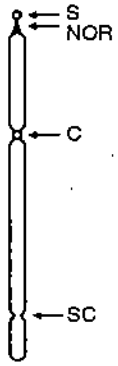


Fig. 8.4 : Diagrammatic representation of various components of a chromosome. For simplicity chromatids are not represented. C, centromere; S, satellite; SC, secondary constriction; NOR, nucleolar organising region.

then it is eventually deleted or lost from the cell during cell division. When two chromosomes have fractured ends, they can fuse with each other, resulting in chromosome morphology different from the normal.

In addition to the centromere, many chromosomes have one or more **secondary constrictions** (see Fig. 8.4). Such secondary constrictions usually appear as a single small body or a pair of such bodies attached to the remainder of the chromosome by a slender thread. The chromosome bodies are known as **satellites**. The connecting thread is known as **nucleolar organising region**. It is the site where **nucleolus** is organised or synthesised. The nucleolus as you know is involved in the production of ribosomes - the site of protein synthesis. During the first stage of nuclear division, this thread can be seen to pass through the nucleolus. At a later stage in division the nucleolus is no longer visible and only the nucleolar organising region can be seen.

When chromosomes are stained by various procedures such as the Feulgen reaction, which is specific for DNA, and are examined by light microscopy some regions of the chromosomes are observed to stain very darkly, whereas other regions stain only lightly. When examined by electron microscopy, the intensely staining chromatin called **heterochromatin**, is seen to consist of densely packed chromatin fibers. The lightly staining chromatin, called **euchromatin**, is composed of less tightly packed fibers.

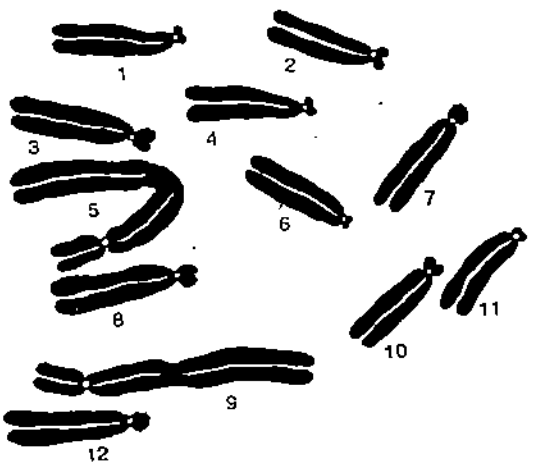
Constitutive heterochromatin : It refers to the occurrence of heterochromatin anywhere along a chromosome at the secondary constrictions, ends of the arms (**telomeric**) or as small intercalations between the euchromatic portions. The latter is referred to as **interstitial heterochromatin.** It has also been found to occur in the centromeric region (**centromeric heterochromatin**).

Facultative heterochromatin : It indicates the condition when one chromosome of a homologous pair becomes partially or wholly heterochromatic and inactive genetically. This is the case with one of the X chromosomes of the human female.

Heterochromatin was thought to be genetically inactive as most of the genes of eukaryotes that have been extensively characterised are located in euchromatic regions of the chromosomes. But the recent findings have shown that the heterochromatin includes some important genetically active regions. These include nucleolar organisers and genes for some of the RNA molecules as well as more genetically inert regions containing highly repeated nucleotide sequences referred to as *repetitive* or *satellite* DNA.

SAQ 1

Observe the figure given below carefully and answer the following questions:



Chromosomes from the root tip of *Vicia faba*.

i) How many homologous pairs does *Vicia faba* contain?

.....

.....

ii) Which chromosomes do you think are homologous pairs?

.....

.....

iii) What is the function of nucleolus?

.....

.....

iv) Which part of some chromosomes is connected with nucleolus function?

.....

8.3 CHROMOSOME PREPARATIONS FOR CYTOGENETIC STUDIES

The study of chromosomes — **cytogenetics** involves microscopic examination of chromosome shape, size, structure and movement. It is essential to have a good preparation with well spread chromosomes.

Chromosomes are best viewed in the metaphase stage of mitosis or meiosis. Cells undergoing mitosis are usually the ones that are examined. Chromosomes at this stage are bivalent and lie in one plane — the equatorial plane, and are most prominent. To get the finer details, chromosomes at prometaphase stage are also studied.

Dividing cells can be obtained from many tissues of the body, e.g., bone marrow, lymph nodes, spleen, regenerating liver, gonads and chorionic villi. Most of these tissues are difficult to obtain and may require some surgical procedures. The easiest source of obtaining cells is the peripheral blood. But the blood cells normally do not divide rapidly. However, with the use of certain chemicals these cells can be induced to divide.

We now discuss a commonly used technique for making chromosomal preparation (see Fig. 8.5) from the blood cells. This whole procedure is carried out in sterile conditions including the use of sterile glassware and nutrient medium.

1. **Obtaining Sample**

A small amount of blood is obtained through finger prick or heel prick. Five to six drops are collected directly on nutrient medium (culture medium) like TC 199. To prevent coagulation of red blood cells in the sample, an anticoagulant - heparin is added to it.

2. **Culture**

To the medium is also added **phytohaemagglutinin (PHA)**. It is a chemical extracted from the red kidney bean. Phytohaemagglutinin stimulates the white blood cells or the lymphocytes to undergo mitotic division.

The cells with the culture medium are incubated at 37°C for 72 hours. During this period most of the lymphocytes have divided twice.

An incubator is a chamber with controlled conditions such as temperature etc.

3. **Harvest**

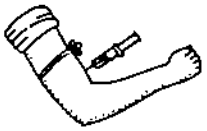
a. Colchicine treatment — To the culture is added a non-toxic chemical derivative of colchicine such as colcemid. This arrests the cell division at metaphase through interference with tubulin polymerisation which forms microtubules of the spindle. Thus if colchicine is present in a dividing cell culture, then cell division will proceed normally up to the metaphase but would not go beyond that stage. A large number of cells available at metaphase makes it easier to study the chromosomes.

Colchicine is a plant alkaloid derived from *Colchicum autumnale* (family Liliaceae). It is also known as meadow saffron and autumn crocus. Colchicine is extracted from the swollen underground stem of this plant. For over 2000 years this substance has been used in the treatment of pain associated with rheumatism and gout. Presently, colchicine is used routinely for chromosome study in virtually all human genetics laboratories.

b. Centrifuge — About 1 hour after the above treatment, centrifugation is done to isolate the cells.

c. Hypotonic treatment — The cells are then suspended for 20 minutes in a weak salt solution of KCl. As a result, the cells imbibe water and swell. This treatment helps to disperse the chromosomes over a large area within the cell, and disentangles the arms of the sister chromatids.

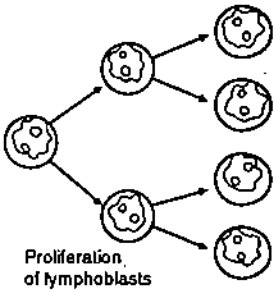
1. **Obtaining Sample**



Syringe containing inhibitor of clotting

2. **Culture**

White blood cells along with plasma

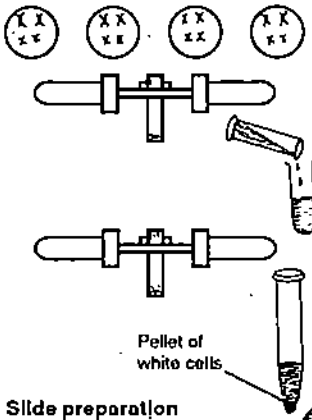


Growth medium containing salts, nutrients, phytohemagglutinin, fetal calf serum, pH indicator, antibiotics

Incubate at 37°C for 72 hours

Proliferation of lymphoblasts

3. **Harvest**



a. Colchicine treatment for 1 hour at 37°C.
Accumulation of cells in metaphase stage of division.

b. Centrifuge at 600-800 rpm for 10 minutes to sediment cells; discard liquid.

c. Hypotonic treatment : Add 0.075 ml KCl, incubate for 20 minutes to allow cells to swell

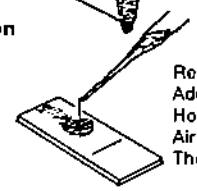
d. Centrifuge at 600-800 rpm for 10 minutes to sediment cells; discard liquid.

e. Fixation
Add cold solution of 3:1 carnoy's fluid (3:1 methanol/acetic acid)
incubate at room temperature for 10 minutes
Centrifuge and collect pellet.

Pellet of white cells

4. **Slide preparation**

Slide preparation is an art



Resuspend pellet in faxative until solution is turbid.
Add 2 drops of cell suspension on acid cleaned slide.
Hold slide at to 75° angle. Fixed cells flatten onto the slide.
Air dry and observe under microscope—check for well spread out nuclei.
The goal of a slide preparation is to have the chromosomes well spread out

5. **Staining and preparing for observation**

Step depends on whether a *banded* chromosome is to be prepared. A variety of stains are available for banded and nonbanded preparations—Giemsa stain is widely used.

Stained cells are flattened out under cover slip.

6. **Microscopic examination**

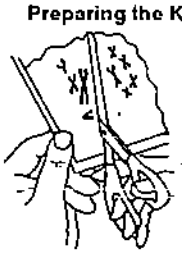
Observe chromosomes of cells in microscope and count the number of cells to obtain a *model* count; any abnormal event is recorded.

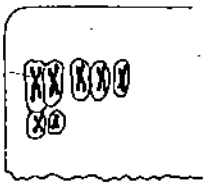
7. **Photography**

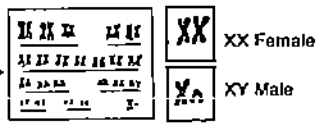
Favourable cell spread is photographed.

Photographs are printed and enlarged for a permanent record.

8. **Preparing the Karyotype**







Individual chromosomes are cut from the photographs.

Homologs are matched, arranged in order of diminishing size, and pasted for analysis and writing report.

Fig. 8.5 : Preparation of human karyotype from peripheral blood lymphocytes.

d. Centrifuge — The cell suspension is centrifuged to sediment the cells.

e. Fixation — The cells are then fixed in cold Carnoy's fluid, that is made up of methanol and acetic acid in the ratio of 3:1. This treatment preserves the cells retaining their internal structure. The pellet of cells is collected by centrifugation.

4. Slide Preparation

Following this, the cells are spread out on a glass slide and allowed to dry. Drying helps separate the chromosomes in a cell from one another, so they can be observed individually.

5. Staining and Preparation for Observation

The cells are then stained with a dye such as Geimsa and dried once again.

6. Microscopic Examination

The slides are examined under the light microscope, model count of the chromosome number is made. Model count refers to the number of chromosomes present in the majority of cells. In normal humans, the model number is forty six. It may be different in some abnormal cases. Areas of the slide with well spread chromosomes is selected for photography.

7. Photography

A few microscope fields having good chromosome spreads are photographed. From the negatives obtained, prints are made. Good pictures are magnified and used for chromosome analysis.

8. Preparing the Karyotype

Each chromosome is cut out from the enlarged print. The cut out images of chromosomes are sorted by size and shape. The homologous pairs are matched and the chromosome pictures are laid out on the karyotype board in a particular format (that you would study in Subsection 8.5.1) and pasted. This is the karyotype and can be used for analysing the chromosomes of that individual.

SAQ 2

i) What is PHA? Why is it used in tissue culture?

.....

.....

ii) What is the source of colchicine? Would there be any difference if the step involving treatment of blood cell cultures with colchicine is omitted?

.....

.....

.....

.....

iii) Why is it important to treat the cells with a hypotonic solution prior to fixation?

.....

.....

.....

8.4 CHROMOSOME BANDING

Several technical developments in the past two decades have revolutionised the study of human chromosomes. In the late 1960's and early 1970's new staining techniques were discovered that have made it possible for human cytogeneticists not only to specify every chromosome but even in many cases to identify the exact parts of chromosomes. This came about because the new techniques disclosed that each of the chromosomes possesses a unique banding pattern. A band is defined as part of a chromosome that is distinguishable from its adjacent segments by appearing lighter or darker as a result of the new staining methods. Bands that appear dark using one staining technique may appear light using another technique. The banded chromosomes are seen as a continuous series of light and dark bands, *there are no interbands*. It is important to understand that

the bands do not correlate to an individual gene. There may be some thirty four genes per small band.

Chromosome banding produces distinctive and reproducible pattern of dark and light bands that change noticeably if the chromosomes undergo some alteration, such as a rearrangement or the loss of a segment. Not only does banding permit the detection of changes, it also makes it possible to pinpoint to the location of change and the specific defects that occur due to a particular disease. The more stretched a chromosome, the more bands become visible and the more accurately its structure can be mapped. In short, banding marks the chromosomes in a way that makes it possible to detect many chemical and physical changes using a light microscope.

8.4.1 Q Banding

Distinct banding patterns are obtained when the chromosomes are stained with fluorochromes such as quinacrine mustard and quinacrine dihydrochloride. Fluorochromes are dyes that bind to the nucleoprotein complexes and show precise patterns of different brightness when viewed under a fluorescent microscope. These are called **Q bands** (Fig. 8.6). These banding patterns are specific and are reproducible for each chromosome pair. The number, size, distribution and intensity of each band is a fixed property of a particular chromosome. By using this technique it is possible to distinguish all the 22 pairs of autosomes and the two sex chromosomes from each other. Metaphase chromosomes collectively demonstrate some 320 Q bands. There are, however, some disadvantages to the Q-banding technique, particularly the impermanence of the fluorescence. This difficulty can be overcome by Giemsa staining.

8.4.2 G Banding

One of the most popular banding method used for routine karyotype analysis is Giemsa or G banding technique. In this method, the chromosomes are incubated in trypsin followed by incubation in Giemsa or in a trypsin-Giemsa mixture. Trypsin is a proteolytic enzyme used for cell dispersal. The preparation thus stained is examined under the light microscope. It is important to understand that although the G bands (Fig. 8.6) are obtained because of the use of the stain Giemsa, these same bands also occur when other stains are used. For example, with minor exceptions, Q bands are identical to G bands except that Q bands fluoresce.

The G bands represent the substructure of the chromosome and indicate the type and location of chromatin along the length of the chromosome. Of the two major types of chromatin, euchromatin and heterochromatin, euchromatin has been shown to be genetically active, and in G banding it is represented as light bands. Heterochromatin which is rich in A–T base pairs, is highly contracted throughout the life cycle, and is genetically inactive. These are represented in G banding as the dark bands.

8.4.3 R Banding

Reverse Giemsa or R banding (Fig. 8.6) is prepared by treating chromosomes with heated phosphate buffer followed by staining with Giemsa. R banding shows light heterochromatic bands and dark euchromatic bands, which are rich in G–C base pairs - the reverse of the pattern that results from G banding.

The problem with R banding is that there are no consistent banding patterns or a 1 : 1 ratio when compared to banding patterns produced by the Q–, and G– banding techniques. For this reason, R banding is not routinely applied but is used to detect minor deletions, inversions, chromosome polymorphism and to evaluate terminal bands that appear light after G banding.

The R banding technique stains the telomeric region especially well, and variations of the R-band procedure have been developed for this purpose. The resulting stained regions are known as **T-bands**.

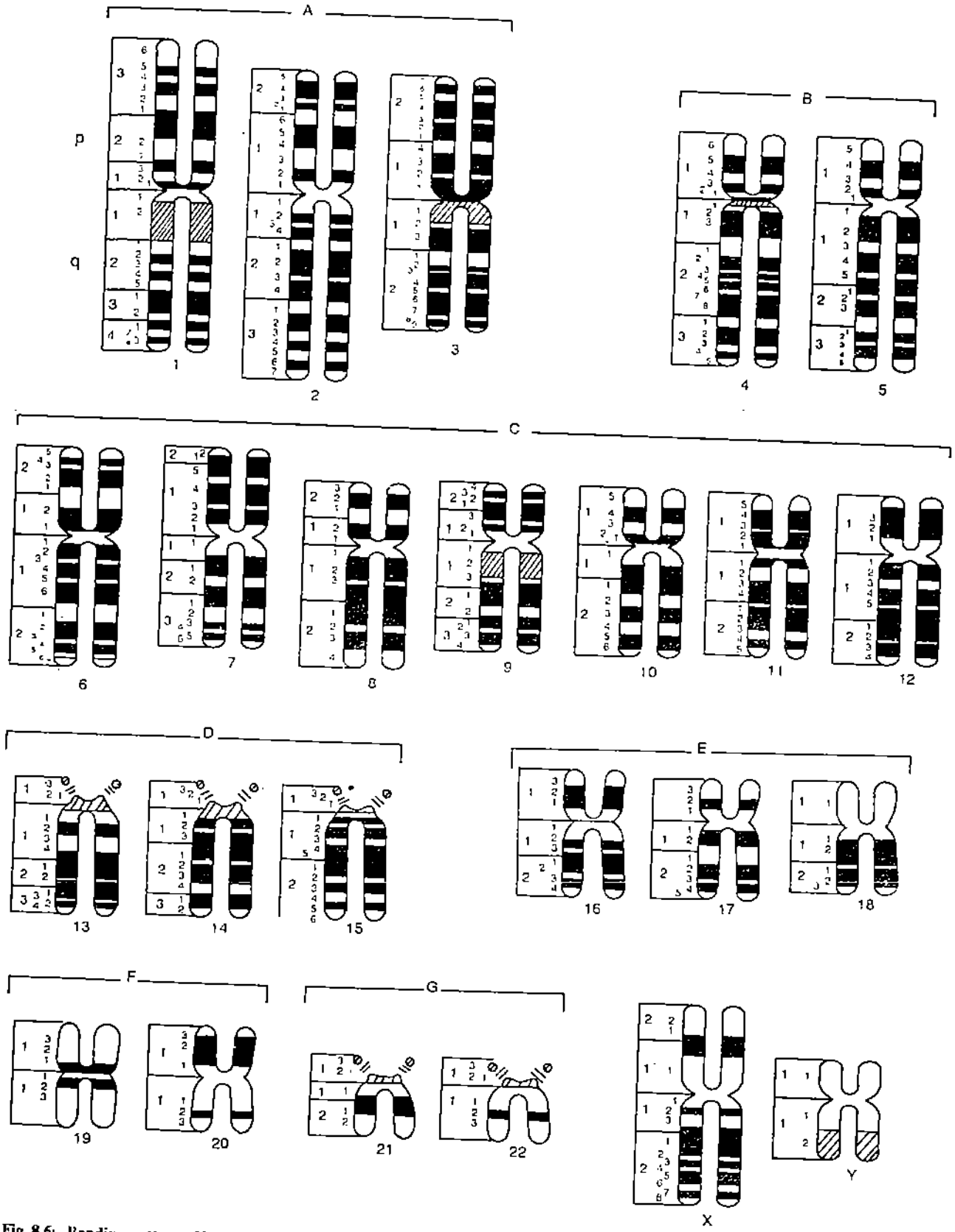


Fig. 8.6: Banding pattern of human male karyotype obtained as a composite of Q, G, and R staining procedures. The 1-22 are assigned to the autosomes and X and Y represent the sex chromosomes. The seven letters A to G represent the chromosome grouping based on the position of the centromere. There is a polymorphic or variable-sized region located beneath the centromere in chromosomes 1,9 and 16, and in the q arm of Y. In the chromosome 3, this polymorphic region is demonstrated in about 70 per cent of the karyotypes. These polymorphisms may appear in one or both homologues and are not associated with a known human disease. Note the satellite chromosome segments extending from chromosomes 13, 14, 15, 21 and 22 and the size difference between the X and the Y chromosomes.

chromosomes in preparing a karyotype. A common nomenclature has been developed for describing a karyotype. By international agreement, the forty six chromosomes are arranged by order of size, arm ratio and centromere position. Figure 8.6 shows a standardised human karyotype. The 46 chromosomes are categorised into two categories—autosomes comprising of 22 pairs and a sex chromosome pair. Each of the twenty two pairs of autosomes is assigned a number based on the length of the chromosome. Chromosome 1 is the longest and chromosome 22 the shortest.

You may recall that the letter p designates the short arm of those chromosomes demonstrating arms of unequal length. The letter q designates the long arm in such chromosomes (see Fig 8.10). The chromosomes 1, 3, 19 and 20 have arms of about equal length, but the remaining autosomes X and Y chromosomes demonstrate short and long arms. You should remember that when placing chromosomes on karyotype sheet, the short arms are placed up, the long arms down. For those chromosomes with arms of equal length, the placement of the p arm is based on the banding pattern.

Figure 8.6 shows chromosomes at metaphase and the p and q arms of each chromosome are divided into regions. The regions are presented in large Arabic numbers consecutively arranged outward from the centromere. In chromosome 1, the p arm contains regions 1, 2 and 3. The q arm contains region 1, 2, 3 and 4. Within each region, the individual adjacent light and dark bands are designated from the centromere out in small Arabic numbers.

In subsection 8.4.6, you have learnt that using high resolution banding techniques, that is, the application of banding methods to chromosomes at the prophase or prometaphase stage, when the chromosomes are more extended, three or four times the number of bands are found than at metaphase (Fig. 8.9). Electron microscopy of banded chromosomes of prophase and prometaphase stage has revealed as many as 1500 to

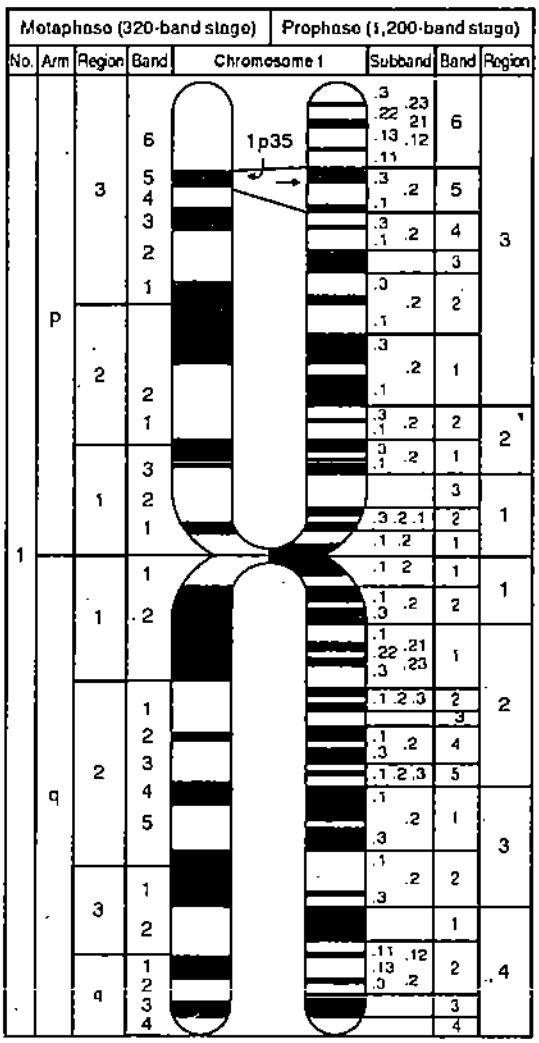


Fig.8.10: Diagrammatic representation of chromosome 1 showing Giemsa-stained bands at metaphase (left hand part of the figure) and prometaphase (right hand). At prometaphase stage, the chromosomes

3000 sub-bands, which when fused by chromosome compaction give rise to the characteristic G bands at metaphase. Banded chromosome preparations may yield sub-bands within bands. The sub-bands are annotated with decimels. For example, Figure 8.10 shows the comparison between a 320-band metaphase chromatid of chromosome 1 to its homologous 1,200 band prophase chromatid. Note that in the figure, the term 1p35 in the metaphase chromatid refers to a specific band location. Read backwards, the term identifies band 5 of region 3 in the p arm of chromosome 1; band 5 actually contains three sub-bands when chromosome 1 is seen in the prophase stage.

High resolution banding techniques have enabled us to detect even small changes that were once thought to be "normal" chromosomes. Many of these changes have now been associated with a number of congenital syndromes whose causes were previously unknown.

8.5.1 Organisation of Karyotype

Individual chromosomes cut from the photograph of a metaphase smear are arranged, pasted on the karyotype board or sheet into seven distinguishable groups (A-G), followed by sex chromosomes (XX or XY). The arrangement of the twenty two pairs of autosomes is based on length and centromere position. The criteria by which chromosomes are assigned to the groups A-G are listed in Table 8.1.

Table 8.1 : Characteristics of groups A-G and the XX and XY chromosomes.

Group	Chromosome Number	Descriptions
A	1 to 3	The longest metacentrics, distinguished from each other by centromere location.
B	4 and 5	The longest submetacentrics separated on the banding pattern.
C	6 to 12	Medium-sized metacentrics, identified only according to the banding.
D	13 to 15	Medium-sized acrocentrics with satellites.
E	16 to 18	Short metacentric 16, and submetacentrics 17 and 18.
F	19 and 20	Short metacentrics.
G	21 and 22	Short acrocentric with satellites.
XX	About the size and shape of C group chromosomes.	
Y	Similar in size and shape to the G-group chromosome but do not have satellites.	

8.5.2 Nomenclature of Karyotype

The report of the 1966 Chicago Conference that standardised the terminology for describing nonbanded normal and abnormal human chromosomes is being used till date. Because of a major breakthrough in preparing banded chromosomes, the Paris Conference (1971) was held to provide cytogeneticists with a basic standardised abbreviations of designating chromosome regions and bands. This system provides a means of describing chromosome structural rearrangements and variants with respect to their banding patterns. In 1976, the Fifth International Conference of Human Genetics was held in Mexico City to elect a standing committee on human chromosome nomenclature. This committee produced the 1978 document entitled "An International System for Human Cytogenetic Nomenclature" written in short as ISCN. It is an extensive document. A brief presentation of some of ISCN symbolism used in describing a karyotype provides an insight into the complexity of communicating the laboratory results among cytologists, from cytologists to genetic counsellor or physician, and from these individuals to the patients. Given below (in Table 8.2) are some of the commonly used symbols in describing human karyotype.

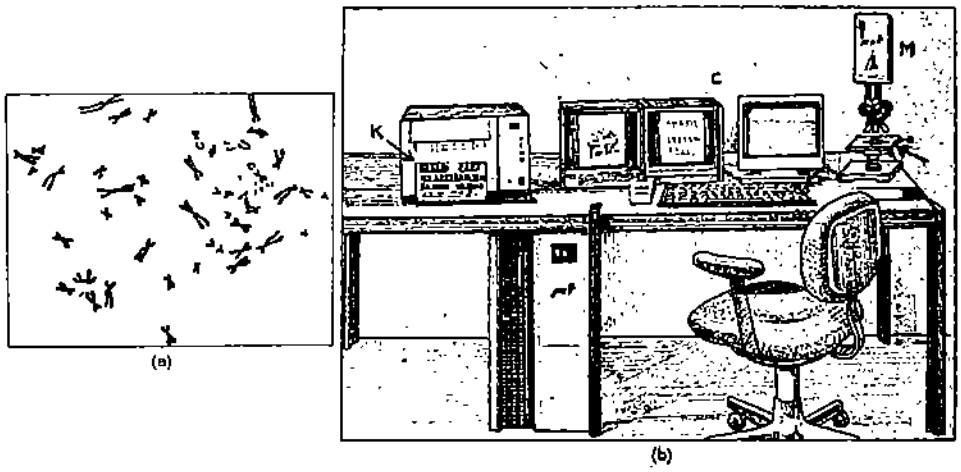


Fig.8.11: An automatic karyotype analysis set up. a) A metaphase spread of Giemsa-banded chromsomes. b) A slide placed on the motorised microscope stage (→) is scanned by the automatic light microscope (M) and each chromosome is sorted without human help. The computer (C) orients and arranges the chromosomes in the standard karyotype format. Chromosomes can be identified, oriented in the right plane, flipped, copied, and magnified in no time. And a printout of karyotype (K) alongwith the group and number classification appear below each chromosome.

SAQ 4

i) What is a human karyotype and how is it arranged?

.....

.....

.....

ii) What does the following karyotype symbolism p, q, 2q11, and ABCDEFG mean?

.....

.....

.....

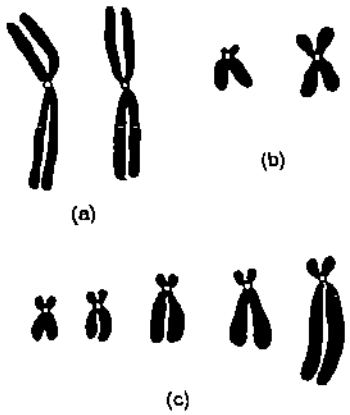


Fig. 8. 12 Some variants found in the normal chromosome complement. a) shows a longer chromosome 1 on the left and its normal homologue as right. b) one of the chromosomes of pair 21 (arrow) has a giant satellite. c) Variation in morphology of normal Y chromosome.

8.6 CHROMOSOME HETEROMORPHISM

Some chromosomes in human set do not look the same in every person, that is, minor variants of certain chromosomes do occur in human populations. Such variations have generally been called **chromosome polymorphism**. The term 'polymorphism' has for years been applied by population geneticists to the occurrence of two or more alleles in frequencies too great to be maintained by 'fresh mutation'. The proper term for cytogenetic chromosome variations as defined by the Paris Conference is **chromosome heteromorphism**. Chromosome heteromorphisms has been examined by many investigators, and the overall consensus is that a normal amount of variation exists between pairs of homologous chromosomes within an individual, a family and general population. Examples of such variants are shown in Fig. 8.12. Chromosome 1 at the left of the Figure 8.12a has a long arm that is noticeably longer than its homologous counterpart. Figure 8.12 b shows chromosome 21 (right) with giant satellites. Variation in the length of long arm of the Y chromosome is illustrated in Fig. 8.12 c.

It is believed that these variations do not have any detectable effects on the individuals. The differences have been aossociated to the degree of coiling of the chromosomes. The more tightly coiled a chromosome is, the shorter it will appear. It is also believed that variation in length occurs in parts of chromosomes that are inert genetically, that is, in the heteromorphic regions that are rich in repetitive and inactive DNA. Let us consider the Y chromosome. The genes on Y chromosome that determine maleness are

located on the short arm of Y, whereas the variation is seen commonly in the longer arm of this chromosome. Perhaps the long arm of the Y (and certain other chromosomal regions as well) carries so little or such a special kind of genetic information that the variation in size makes no difference.

Normal individuals can have chromosomal differences, that is, there is chromosomal (i.e., genetic) variation among normal people. Chromosomal variation is just the tip of the iceberg of genetic variation. You would study later in the course, that there is so much subtle variation in the DNA among individuals – variation not visible in the microscope that no two individuals (possibly with the exception of identical twins) are genetically alike.

8.7 MOLECULAR ORGANISATION

The human cell has partitioned its genetic material – DNA into forty six chromosomes. Each chromosome contains a linear duplex molecule with about 100 million base pairs. The DNA in the forty six chromosomes if laid in a line end to end would be about 1 meter in length.

Chemical analysis of isolated chromatin shows that it consists primarily of DNA and proteins with lesser amounts of RNA. The proteins are of two major classes: 1) **Histone** – the basic proteins that are positively charged at neutral pH; 2) **Nonhistone proteins** – are acidic at neutral pH. Histones are small proteins that contain between 100 and 200 amino acids and from 20 to 30 per cent of its amino acids are lysine and arginine. The positive charges on the histones enable them to bind to DNA, primarily by electrostatic attraction to the negatively-charged phosphate groups in the sugar-phosphate backbone of DNA. Histones also bind tightly to each other; both DNA–histone and histone–histone binding are important for chromatin structure. Histones play a major structural role in chromatin. They are present in the chromatin of all higher eukaryotes in amounts equivalent to the amount of DNA weight/weight. The histones of eukaryotes including man are composed of five different proteins. These five major histones called H1, H2a, H2b, H3 and H4 are present in all cell types. A few exceptions exist in the sperms of some eukaryotes where the histones are replaced by another class of small basic proteins called *protamines.*

There is a remarkable constancy of histones H2a, H2b, H3 and H4 in all cell types of an organism and even between widely divergent species. These histones are believed to be important in chromatin structure – "DNA packaging" and are involved in the regulation of gene expression.

The five histones are present in molar ratios of approximately 1H1 : 2H2a : 2H2b : 2H3 : 2H4. They are specifically complexed with DNA to produce the basic structural subunits of chromatin – the nucleosome. You would learn their details in subsection 8.7.1.

The histones have been highly conserved during evolution, four of the five types of histones are very similar in all higher eukaryotes.

On the other hand, the nonhistone protein fraction of chromatin consists of a large number of heterogenous proteins. The composition of the nonhistone chromosomal protein fraction varies widely among different cell types of the same organism. The nonhistone chromosomal proteins are thus, thought to be involved in regulation of expression of specific genes or sets of genes.

8.7.1 Packaging of DNA into Chromosome

The largest chromosome in the human genome contains about 85 mm (85,000 μm or 8.5×10^8 Å) of DNA that is believed to exist as one giant molecule. This DNA molecule somehow gets packaged into a metaphase structure that is about 0.5 μm in diameter and about 10 μm in length. This represents a condensation of almost 10^4 fold in length from the naked DNA molecule to the metaphase chromosome. Now, the questions arise how does this condensation occur? What are the components that are involved in the packaging processes? And what is the packaging scheme? You would find answers to these questions in the following text.

Nucleosome Structure

When isolated chromatin is examined under the electron microscope, it is found to be made up of a series of ellipsoidal beads joined by thread-like structures (Figs 8.13 and 8.14). These beads are nucleosomes that are joined by DNA known as the linker DNA. When chromatin is subjected to various nucleases, the DNase attacks the linker DNA in the region between the beads and then releases the beads, that is, nucleosome from the continuous strand (Fig. 8.14). But the DNA in the nucleosome itself is protected from digestion by the nuclease. When digestion by nuclease is limited, pieces of DNA about 200 base pairs long are produced indicating that there are about 200 base pairs of DNA from linker to linker. When digestion by nuclease is extended, "core" nucelosomes are produced in which the linker DNA is digested, leaving a piece of DNA about 146 nucleotides long. Core nucleotides are flattened spheres or discs with a diameter of 100

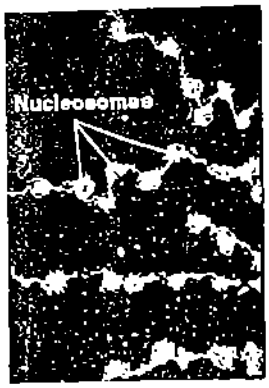


Fig. 8.13 : Electron micrograph of chromatin showing bead-like structures on strings. Each 'bead' is a nucelosome.

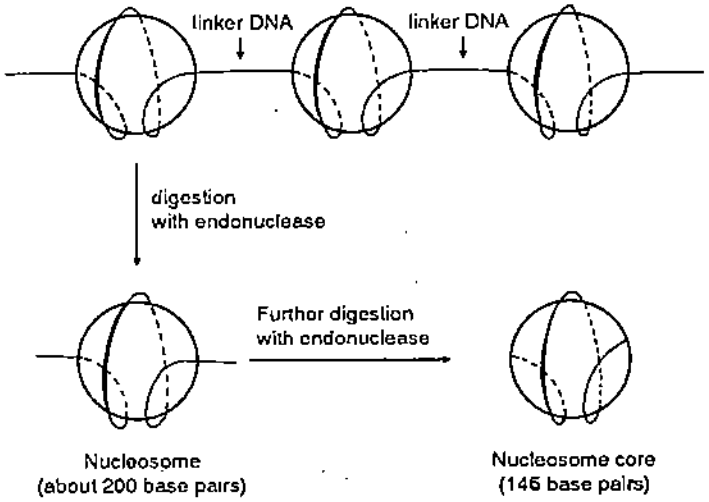


Fig. 8.14 : Production of nucleosome cores by cleavage and endonuclease.

Å and a thickness of about 50 Å. The DNA is on the outside of the nucleosome and winds around it about 1 and 3/4 turns. The inside of the core is composed of the basic structural chromosomal proteins – the histones.

The nucleosome core consists of 2 molecules each of histones H2A, H2B, H3, and H4 forming a histone octamer. The core-constitution histones have specific position within the nucleosomes and their contacts with DNA are indicated in Fig. 8.15. The composite structure of nucleosome is depicted in Fig. 8.16 wherein the aggregate of four pairs of histones is shown as a contoured structure. Did you note position of the histone

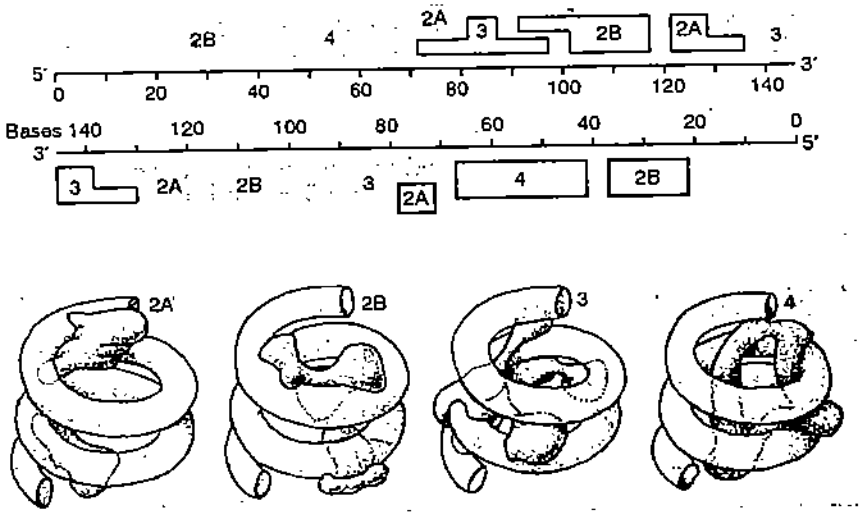


Fig.8.15: Position of histones H2a, H2b, H3 and H4 in nucleosomes. The coiled tube-like structures represent DNA of a nucleosome. The various stippled shapes represent histones. The top diagram shows where individual histones make contact with each strand of a double helix. (After Haraux and Ottensmeyer, 1984, Science Vol, 226, p. 936)

H1 in the Figure (8.16)? The histone H1 is not only different in location but also has a different function. This histone sits on the outside of the nucleosome core and binds to about 20 base pairs of the DNA as the DNA enters and exits the core particle (see Fig 8.16). By holding both the entering and leaving strands of DNA, histone H1 brings the

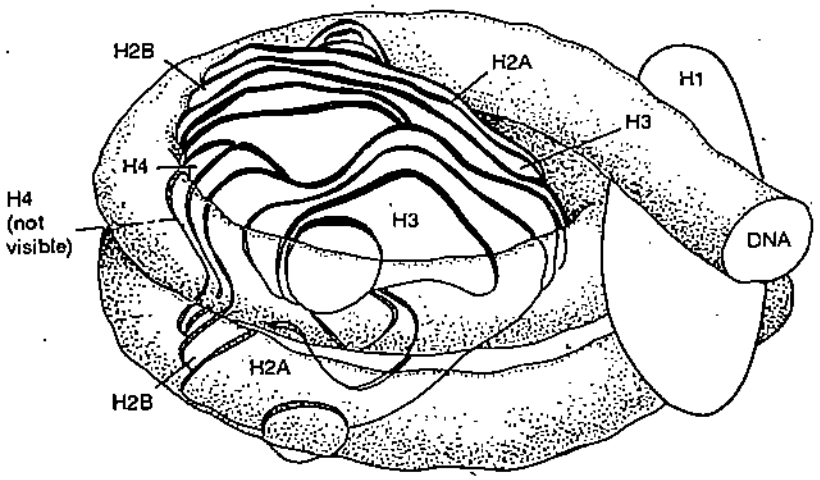


Fig. 8.16: A composite picture of a nucleosome sharing the position of individual histones within the core. Note, the histone H1 occupies a position outside the core and holds the two strands of DNA together as they enter and leave the core. (After Klug et al ;1980, Nature, Volume 287, p. 509).

nucleosomes together – a step considered to be necessary for formation of a 300 Å solenoid (Fig. 8.17). Each turn of the solenoid contains 6 to 7 nucleosomes, the H1 histone molecules being located on the inside surface of the solenoid.

A large number of acidic nonhistone chromosomal proteins have been thought to be involved in further hierarchical coiling or packaging of the 300 Å solenoid. The precise roles of these proteins in chromosomal structure and function are largely unknown. Some of these proteins are believed to have structural roles in the formation of the 2000 to 4000 Å supersolenoid (see Fig. 8.17).

8.8 CHROMOSOME BANDING AND TAXONOMIC RELATIONSHIP

Similarities in chromosome structure have been helpful in elucidating the taxonomic and evolutionary relationship among species. With the development of banding methods, such studies have revealed many interesting facts. One of the most interesting sets of findings has come from a comparison of banded karyotypes of man (*Homo sapiens*) where 2n = 46, with those of chimpanzee (*Pan troglodytes*) and gorilla (*Gorilla gorilla*) and the orangutan (*Pongo pygmaeus*). In these latter three species the diploid chromosome number is 48. Some chromosomes of the four karyotypes have been evolutionarily stable. For example, human chromosome 1 is closely similar to autosome 1 of the other three other species (see Fig. 8.18). Others like human autosome 3, is closely similar in all but the orangutan. Human autosome 2 has been found to have resulted from the fusion of two acrocentric ape chromosomes, numbers 12 and 13. This accounts for the difference in somatic chromosome number between humans and the three ape species. See Fig. 8.18 carefully, and match the chromosomes of the four species.

SAQ 5

i) Histones are basic proteins, which means that they interact with an acid. What is the acid with which they intereact?

.....

.....

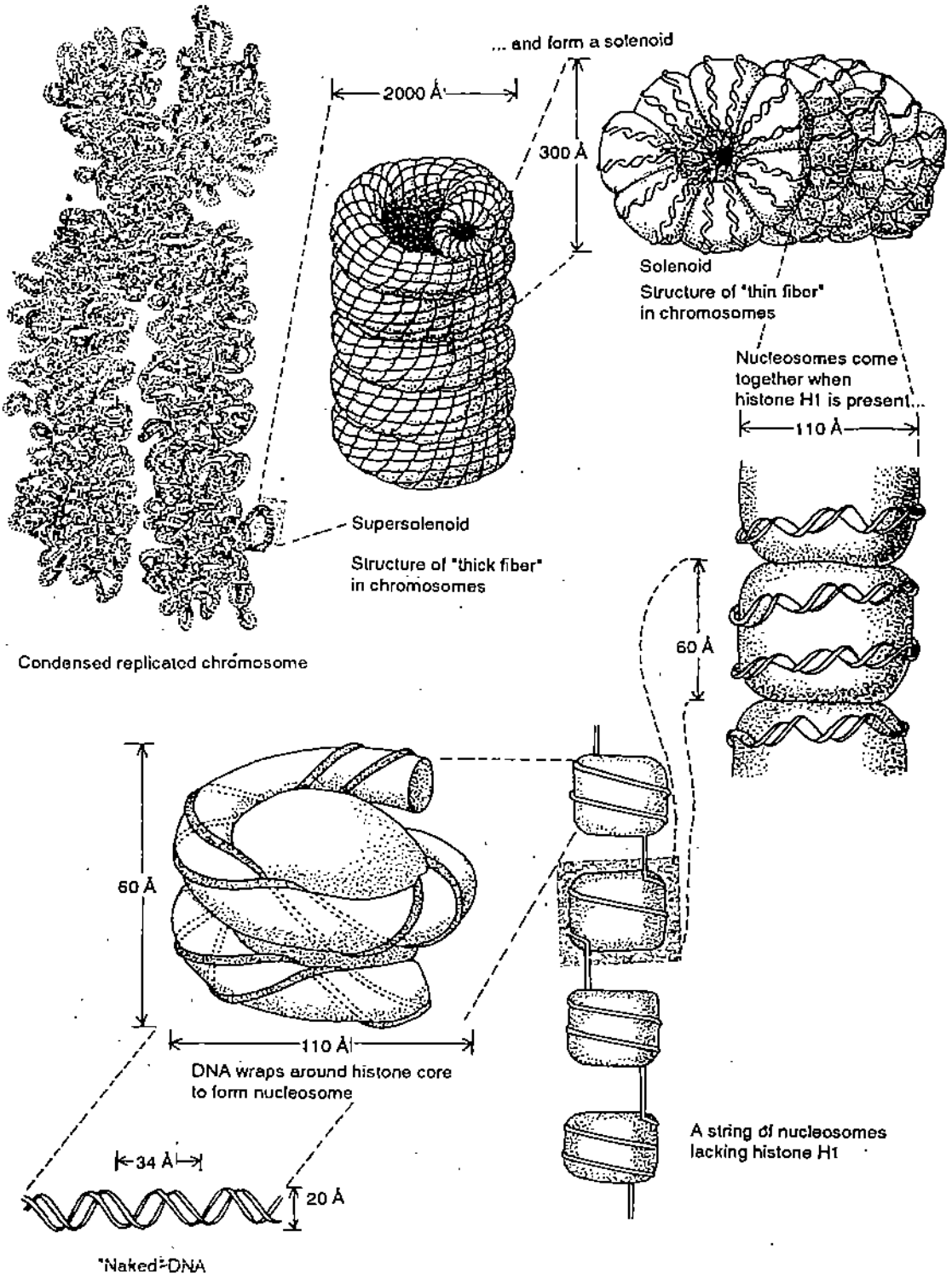
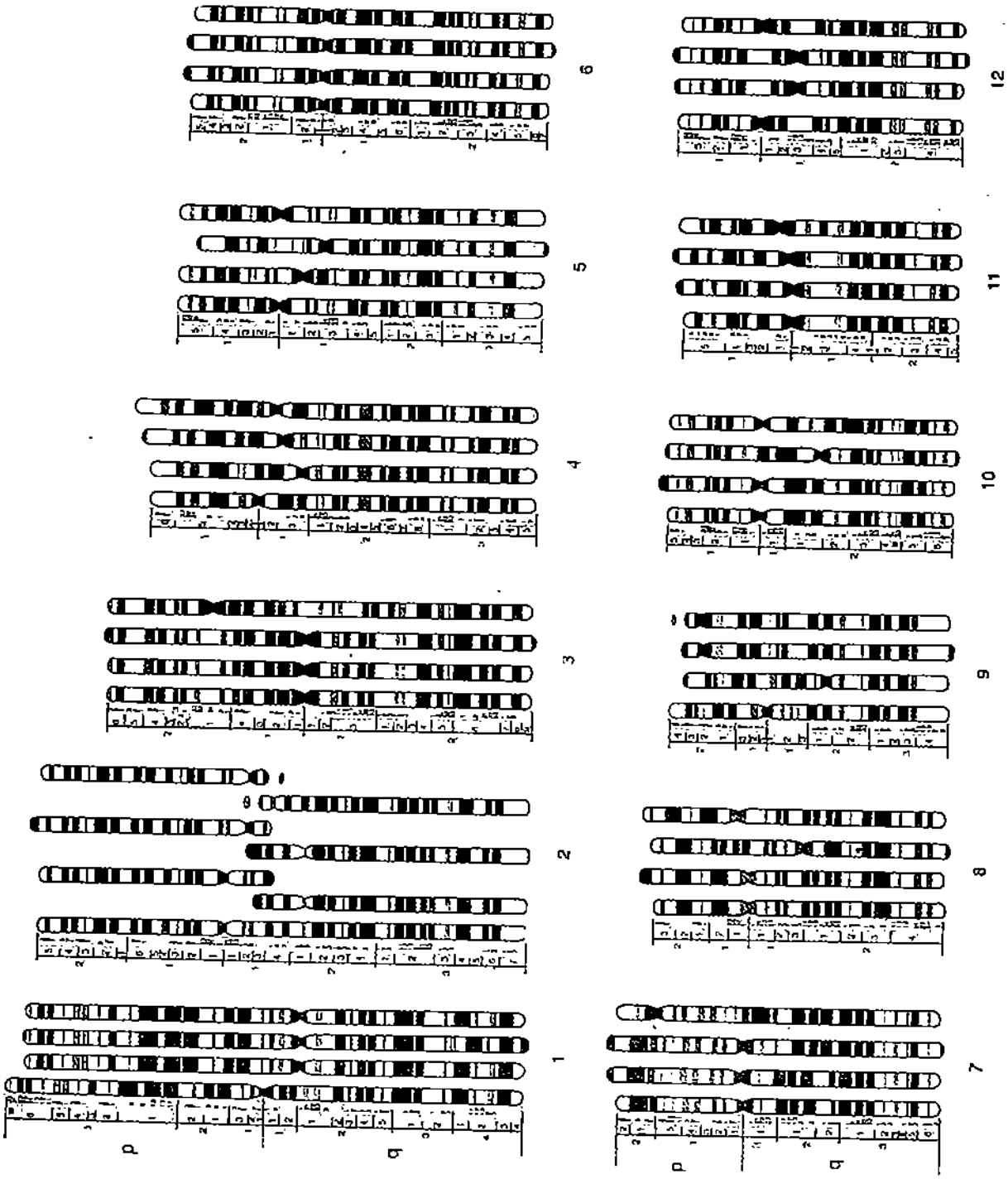


Fig. 8.17: Different levels of packaging of the chromosomal material ranging from a molecule of DNA to a supersolenoid that is part of a condensed chromosome.


6
12
5
11
4
10
3
9
2
8
1
7
p
q
p
q

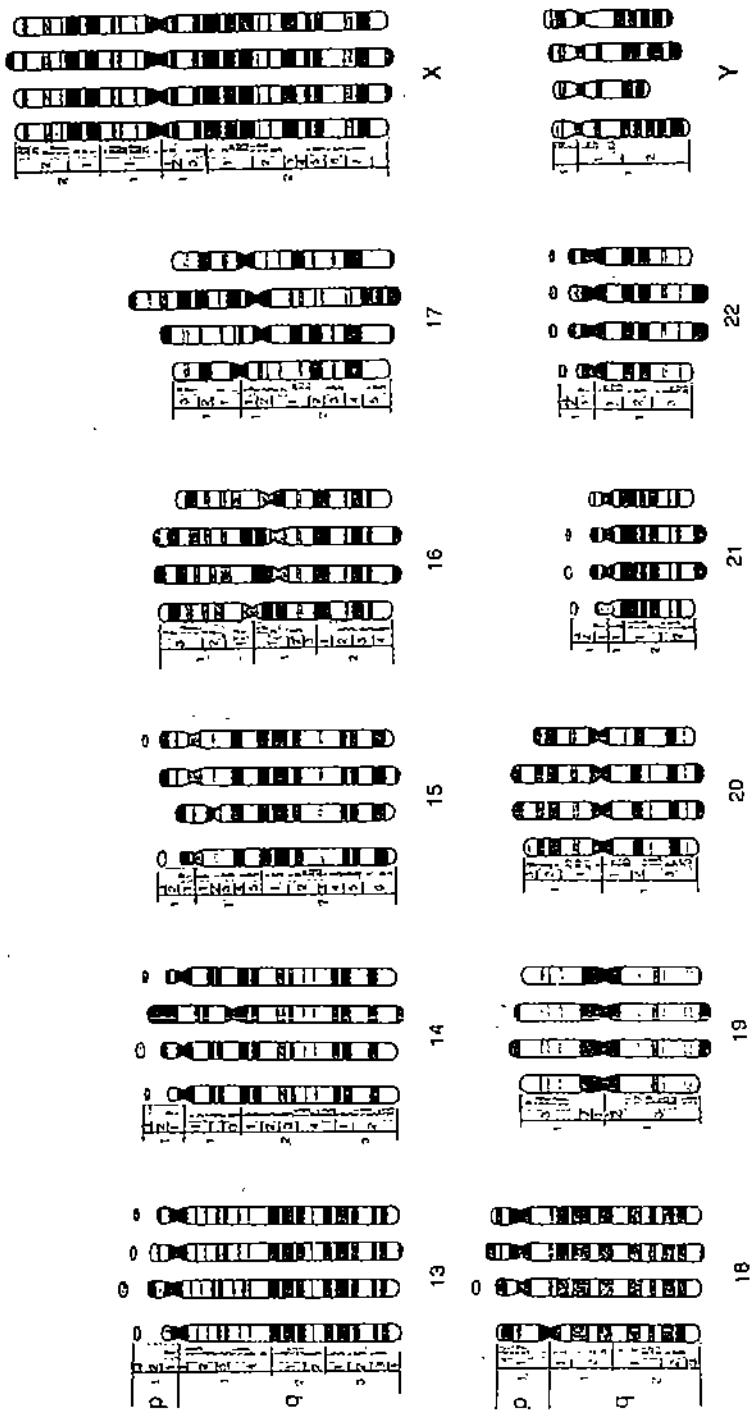


Fig. 8.18 a,b : Comparison of late-prophase chromosomes (left to right) of human, chimpanzee, gorilla and orangutan. These drawings represent the 1000 band stage. Note: human chromosome 2 is a fusion product of two chromsomes found in great apes. (From: J. J. Yunis and O. Prakash 1982. "The Origin of Man ; A Chromosomal Pictorial Legacy", Science, Vol. 215, pp. 1525-1530).

ii) How many individual protein molecules are in a nucleosome?

.....

iii) Which type of histone molecule is not present in pairs in nucleosomes?

.....

iv) Are the ratios of the different histone types the same in all cells of a eukaryotic organism?

.....
.....
.....
.....
.....

v) How do you imagine the intact chromosome pictured in Figure 8.17 would appear if it lacked histone H1?

.....
.....
.....
.....
.....

8.9 SUMMARY

In this unit you have learnt that:

- The normal diploid complement of chromosomes in humans consists of 46 chromosomes – 44 autosomes and 2 sex chromosomes. In females, the sex chromosomes are both X chromosomes; in males one sex chromosome is an X and other is a Y.
- The effective examination of human chromosome depends on several technical developments. These include the ability to culture cells, the use of colchicine to arrest cells in metaphase, the use of hypotonic treatment to spread chromosomes and in the case of lymphocytes, the stimulation of cell division by chemicals like PHA.
- Chromosomes contain both active and inactive chromatin.
- Each chromosome contains many millions of base pairs in a single DNA molecule.
- Human chromosomes can be identified by morphological differences and by the banding pattern they demonstrate.
- Specific chemical treatment of chromosomes produces Q, G, R, C and NOR bands.
- Morphology and banding pattern permit accurate karyotyping of human chromosomes.
- The human karyotype contains seven chromosome groups with each chromosome in a group maintaining its individual set of bands.
- Analysis of human karyotypes can reveal the cause or potential cause of a variety of genetic diseases.
- Normal chromosome variations that exist in a population are called chromosome heteromorphism.
- In eukaryotes including humans, DNA is packaged into the chromosomes in a hierarchy of coils. The primary structure is the nucleosome in which DNA is wound around the outside of a core of histones. The nucleosomes may themselves be coiled into a helical tube, a solenoid and the solenoid itself may be coiled into a supersolenoid.

- Chromosome banding patterns suggest that humans, chimpanzees and gorillas have a common ancestor.

8.10 TERMINAL QUESTIONS

1) List **five** different structures which could be identified on a chromosome.
2) List **four** ways in which chromosomes may differ structurally from each other.
3) Distinguish between the following terms:
   a) Chromatin fibre and chromosome
   b) Nucleolus and nucleolar organising region
4) Why are metaphase chromosomes more convenient to study than the ones at prophase? Why are prophase chromosomes often more informative than metaphase chromosomes?
5) What are the special advantages and disadvantages, if any, of a) Q-banding; b) G-banding; c) C- banding; and d) R-banding?
6) Without refering to the text, state the number of chromosomes normally found in each group – A to G, in the diploid set of : a) the human male; b) the human female.
7) Perform an exercise similar to 6 for the gametes of (a) the female; and (b) the male.
8) What are the accepted designations for 'long arm' and 'short arm' of a chromosome? What is the basis of these designations?
9) List an example of facultative heterochromatin.
10) Discuss the structure and role of nucleosomes.

8.11 ANSWERS

Self–assessment Questions

1) i) 6 pairs
   ii) 5/9, 6/11, 2/10, 1/4, 7/12, 3/8
   iii) The nucleolus produces ribosomes for protein synthesis.
   iv) The nucleolar organising region and satellite.
2) i) PHA is a phytohaemagglutinin. It is used to stimulate cell division.
   ii) It is extracted from the underground stem of *Colchicum autumnale* of family liliaceae. If the above said step is omitted, we may not get in the preparation, many cells at the desired stage (i.e., metaphase) of division.
   iii) This treatment disentangles the arms of sister chromatids and spreads the chromosomes over a large area in the cell.
3) i) A chromosome band is a segment of chromosome that is distinguishable from adjacent segments by a lighter and darker appearance as a result of a staining method.
   ii) Q – Quinacrine; G – Giemsa; C – centromere or constitutive heterochromatin; R – reverse Giemsa, NOR – Nucleolus organising region banding.
   iii) **Hint** : Can be observed with light microscope. If the chromosomes are pretreated with proteolytic enzymes and stained with Giemsa, generally Q-band like patterns are seen. And R banding as you know is produced on pretreating chromosomes with heated phosphate buffer while staining with Giemsa. The banding pattern is reverse of G and Q.
4) i) A human karyotype is a display of all the forty six chromosomes sequentially arranged by size. The largest pair is chromosome 1 and the smallest pair is

chromosome number 22.

ii) p = short arm; q = long arm; 2q11 means band 1 in region 1 in the long arm of chromosome 2; A, B, C, D, E, F, G are the seven groups into which the human karyotype has been divided.

5) i) Histones interact with deoxyribonucleic acid, which is a weak acid.

ii) Each nucleosome contains two molecules of each of four distinct types of histones forming the histone octamer, plus a single H1 histone in the linker.

iii) H1 is unpaired.

iv) There is always a 1:1:1:1 ratio of histones H2a, H2b, H3, and H4 in all eukaryotic cells because of the universality of the histone octamer. The ratio of H1 to the others varies slightly.

v) **Hint :** Absence of H1 would mean a lack of any structural conformation.

Terminal Questions

1) Centromere, nucleolar organiser, satellite, secondary constriction, telomere.

2) Length, position of centromere, presence/absence of nucleolar organiser, presence/absence of secondary constrictions and satellites.

3) a) Chromatin refers to the fibres of DNA and protein found in the cell nucleus. They become packed due to coiling to form the chromosomes prior to nuclear division.

Chromosomes are the thread-shaped bodies composed of packed chromatin fibres found in the nucleus. They are readily visible with the light microscope just prior to and during nuclear division.

b) The **nucleolus** is a darkly staining body found in the nucleus which is involved in ribosome manufacture. These are usually one or two per nucleus.

The *nucleolar organising region* is a portion of a chromosome narrower than the main part, representing the site at which the nucleolus is synthesised. The number of these regions corresponds to the number of nucleoli. Usually one or two chromosomes possess them.

4) Metaphase chromsomes are highly condensed and appear distinctly. The prophase chromosomes reveal more details of the chromosome through many fold increase in bands than the metaphase stage. A study of prophase chromosomes enable us to detect even the smallest structural change in the chromosome.

5) See Section 8.4.

6) To compare your answer refer Sub-section 8.5.1.

7) Check with Sub-section 8.5.1,

8) Long arm: q; short arm: p.

Short arm is designated 'p' from the french word 'petite' meaning short or small.

(You may also refer Section 8.2 and 8.5.)

9) Facultative heterochromatin – e.g. one of the two X chromosomes in mammalian (including human) females, exhibits this property.

10) Refer to Section 8.7.

UNIT 9 STRUCTURAL ABNORMALITIES IN CHROMOSOMES AND THEIR EFFECTS

9.1 INTRODUCTION

The process of cell division, both meiosis or mitosis, transmits chromosomes that are morphologically intact and unchanged from one generation to another. At times, however, errors occur that result in the modification of one or more chromosomes within a cell resulting in cells and individuals with abnormal structures or numbers of chromosomes. Such visible chromosomal abnormalities have been found to occur in about six or seven of every 1,000 live births. A study of chromosomal abnormalities is important as these are associated with problems like infertility, miscarriages, several kinds of birth defects, delayed puberty, hermaphroditism and cancer.

You will study chromosomal abnormalities due to change in their structure or number in both this unit and the next unit. Having studied the normal chromosomal structure in Unit 8, you can now appreciate the significance of such studies. You should keep in mind that these *two classes of abnormalities are not mutually exclusive.* There are *some individuals who possess a numerical chromosomal change in conjunction with a structural change.* Such changes are also referred to as chromosomal aberrations.

Aberrations of structure refer to the changes in the parts of individual chromosomes rather than whole chromosomes or sets of chromosomes in a genome. There are four types of structural alterations : **deletions** and **duplications**, which involve a change in the *amount* of genetic material on a chromosome; **inversions** that involve a change in the *arrangement* of a chromosomal segment, and **translocations** which involve a change in the *location* of a chromosomal segment. You would study these four basic types of aberrations and their effects in this unit.

Objectives

After studying this unit you would be able to:

- describe the origin, diagnostic features and the phenotypic effects resulting from the structural chromosomal aberrations, viz., deletions, duplications, inversions and translocations (Sections 9.2 to 9.5);

- explain the formation of ring chromosomes and isochromosomes and their effect on organisms' phenotypes (Sections 9.6 and 9.7).

9.2 DELETIONS

A deletion or deficiency is a chromosomal change resulting in the loss of a segment of the genetic material and the genetic information contained therein (Fig. 9.1). A deletion may be as small as a nucleotide (though such a short segment cannot be seen cytologically), or it may be a large chromosome fragment. Deletions may be caused by agents such as heat, radiation, viruses, chemicals, drugs or due to errors in recombination.

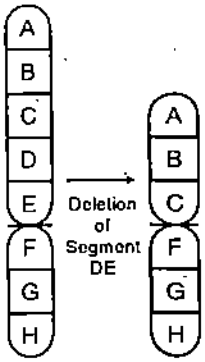


Fig.9.1 : A chromosome showing deletion of a segment, that is, D,E.

The deletion where the chromosome breaks off near the end it is called a **terminal deletion** (Fig.9.2 a). If there are two breaks within a chromosome arm and the segment is lost, it is called an **interstitial deletion** (see Fig. 9.2 b). The chromosome segment that lacks a centromere (known as the acentric fragment), will be unable to attach to the

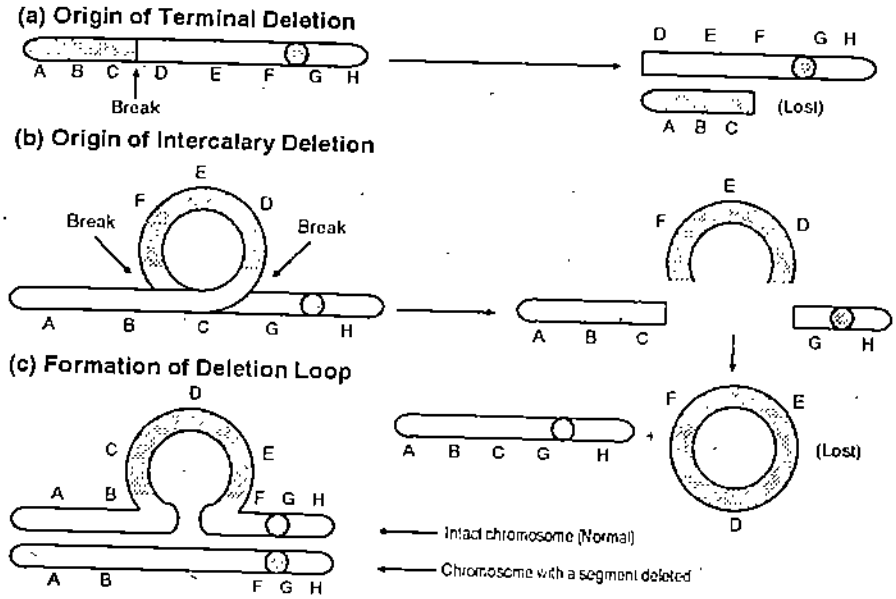


Fig.9.2 : Terminal (a), and interstitial (b) deletions. (c) Formation of deletion loop.

spindle fibre, attach at the next cell division, and consequently it will not be included in the nucleus of either of the daughter cell. Such fragments are ultimately degraded and are lost from the genome. Genetically, deletions are distinguished from gene mutations. In the case of mutations, the genetic material may back mutate but in deletions the loss is permanent, as the genetic material is no longer available for further change. Cytologically, deleted chromosome sections can be detected in prophase of meiosis I. If the cell is heterozygous for deletion, that is, it has a deficient and a normal chromosome, then during synapsis, the chromosomes pair precisely band for band, along all the homologous regions. In the deficient region, however, the normal chromosome will not have any region with which it can pair. Therefore, small unpaired chromosome segment buckles out to the side (Fig. 9.2 c). It is, therefore, known as **buckling effect**. Such a configuration is called a **deficiency loop** or **compensation loop**, and can be observed under the light microscope.

The consequences of deletion depend on the genes or parts of genes that have been removed. The consequences are severe especially when a euchromatic chromosome band is removed (see Fig. 9.3 arrows). At a 1,000 prometaphase band stage, each half band contains about 50,000 base pairs, and these can code for many genes. You can well imagine what would be the result when several base pairs are deleted. In diploid organisms, the effects may be lessened by the presence of another set of the deleted gene in the homologous chromosome. However, if the homolog contains recessive genes with deletrious effects, then the consequences can be severe. We, therefore, know mostly about those deletions that cause a detectable phenotypic change.

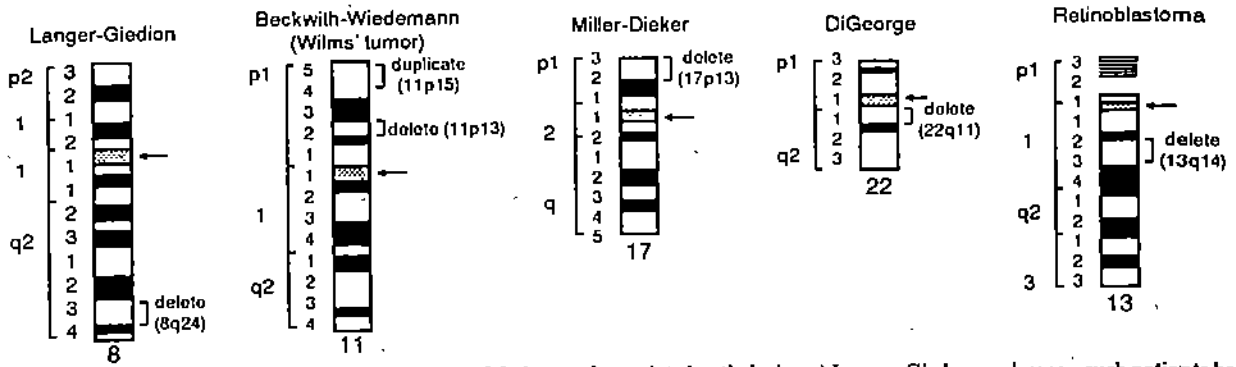


Fig. 9.3 : Some examples of interstitial chromosome deletions and associated pathologies. a) Langer-Gieden syndrome – such patients have large, pear-shaped nose, long philtrum, small jaw, thin upper jaw, thin upper lip, large protruding ears, sparse scalp hair, mental retardation and short stature. b) Beckwith —Wiedemann syndrome (Wilm's tumour) patients exhibit an enlarged tongue, gigantism, hypoglycemia and omphalocele (protrusion of part of the intestine at birth). c) Miller-Dieker syndrome patients exhibit a small jaw, multiple dysmorphic features, siezures, lack of muscle tone and mental retardation. d) DiGeorge syndrome patients exhibit congenital absence of the thymus and parathyroid glands and immunodeficiency. e) Retinoblastoma, cancer of the eye. (After Stine, G.J. 1989. *The New Human Genetics.* Web, USA.)

With the use of high resolution banding techniques a growing number of interstitial deletions have been found to be associated with diseases of previously unknown causes (Table 9.1). Deletions have also been identified in the X and Y chromosomes and most of the autosomes. Virtually all deletions of euchromatic bands result in mentally retarded babies having multiple congenital anomalies. Note that several of the deletions listed in Table 9.1 cause cancer.

Table 9.1: Some diseases associated with the interstitial chromosome deletions. After Stine, G.J., 1989. The New Human Genetics. Web, USA.

Disease:	Interstitial deletion
Wilms' tumor	del (11) (p 13)
DiGeorge's syndrome	del (22) (pter-q 11)
Duchenne's muscular dystrophy	del (X) (p 21)
Goldenhar's syndrome	del (4) (p 16)
Hallermann-Strieff syndrome	del (4) (p 16)
Langer-Giedion syndrome	del (8) (q 23.3)
Miller Dieker syndrome	del (17) (p 13.3)
Prader-Willi syndrome	del (15) (q 11-13)
Retinoblastoma	del (13) (q 14)
Seckel's syndrome	del (13) (q 22-qter)
Small-cell lung cancer	del (3) (p 14-23)

Since the Chicago Conference (1966) deletions have been indicated by a minus sign (–) after the arm involved. Xq – for example means that there is a deletion in the long arm of the X chromosome. The Paris Conference (1971) suggested adding "del" before the chromosome number. Thus the karyotype of a human being with deletion would be written as 46, X, del (X q or p), without need to use the minus sign. The letter X stands for the chromosome number, and it can be 1–23 or X or Y chromosome. If the break-point bands are known, they are added in another parenthesis, e.g., 46, X, del (X) (q 13 – q13 – 25) indicates that part of q13, part of q25 and the bands in between are deleted. In order to denote deletions, any of the two symbols, i.e., – sign or del may be used.

Deficiencies occur in humans as well as in other organisms. We shall discuss two examples that result due to such chromosomal aberrations in humans.

9.2.1 Cri-du- chat Syndrome

A well-known disorder associated with a deletion in humans is cri-du-chat syndrome, described by Lejeune and Colleagues in 1963. The children suffering from this disorder have a characteristic, high-pitched, plaintive cry, very similar to a cat in distress.

Accordingly, this disorder was called, in French, the cri-du-chat (cry of the cat) syndrome. This syndrome is associated with a malformation of the larynx. The cry, facial appearance and other abnormal physical features and an IQ (Intelligence Quotient that tells us about the level of mental capacity) usually below 20, are so characteristic of this deletion that the diagnosis can often be made prior to karyotype analysis. Other nonspecific malformations and dysfunctions involving the brain, heart, eyes, kidneys and skeleton may also occur. Nevertheless, many affected individuals attain adulthood. The frequency of this defect is about 1 in 50,000 live births. Twice as many live born females are affected than males.

This disorder results due to a deletion of a segment of the short arm of chromosome 5 (see Fig. 9.4), precisely the parts of bands 14 and 15 and the region in between. The karyotype is thus represented as 5,del (p14p15). In most cases, the deletion arises during gametogenesis in one parent or the other. Both parents of affected children, however, present normal karyotypes. No parental age effect is discernible.

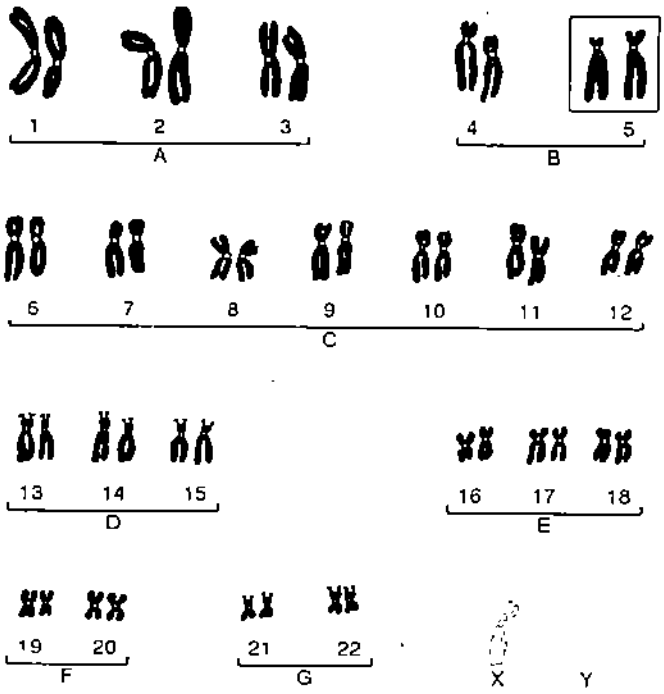


Fig. 9.4: Karyotype of a child with cri-du-chat syndrome. Note deletion of the short arm of one of the number 5 chromosome (see box).

9.2.3 Myelocytic Leukemia

Patients with chronic myelocytic leukemia carry a short chromosome (named the philadelphia chromosome – Ph^1, after the city where the first case was discovered) in all the bone marrow cells and cells originating in the bone marrow (i.e., neutrophils, eosinophils, normoblasts, and megakaryocytes). During remission Ph^1 persists in bone marrow and often in cells derived therefrom. Detailed cytological study discloses Ph^1 to be a number 22 chromosome that has lost most of the distal part of its long arm (22q–). This deleted part of autosome 22 is often seen to be attached (translocated) to one of the other autosomes, preferentially to the distal end of 9q, and reciprocally a small fragment from 9q is translocated to 22q.

9.3 DUPLICATIONS

When any part of the genetic material – a single locus or a large piece of a chromosome is present more than once in a genome, it is called duplication. Such chromosomes, that is, one with duplicated portion, and the other one normal, produce compensation loops at the time of synapsing similar to those described for chromosome deletions.

Duplications arise as a result of unequal crossing-over between synapsed chromosomes during meiosis (see Fig. 9.5) or through a replication error prior to meiosis. In the former case, both the duplication and a deficiency are produced.

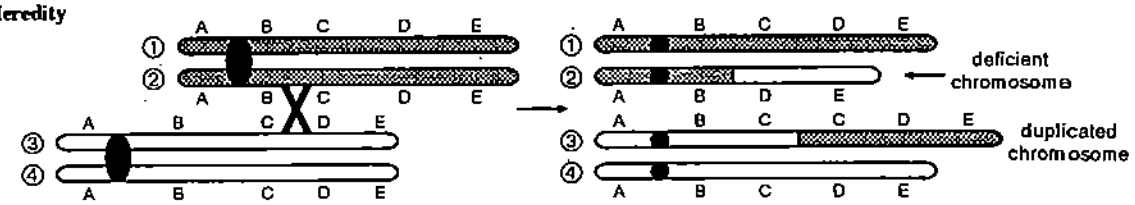


Fig. 9.5 : The formation of duplicated and deficient chromosomes as a result of unequal crossing-over.

The duplicated segment may be attached:

i) either to the chromosome whose loci are repeated,
ii) or to a different linkage group,
iii) or it may even be present as a separate fragment.

When present on the same chromosome (see Fig. 9.6 b), it may be present in a *tandem configuration*, e.g. *ABCBCDE*.FGH, that is, adjacent to one another and in the same order. Sometimes, however, the order of the genes in the duplicated segment are reverse of the original, e.g., *ABCCB*DE.FGH and it is known as *reverse tandem configuration* (see Fig. 9.6 c). And when the duplicated segments are tandemly arranged at the end of the chromosome, it is called a *terminal tandem duplication* (Fig. 9.6 d), e.g., *ABAB*CDE.FGH.

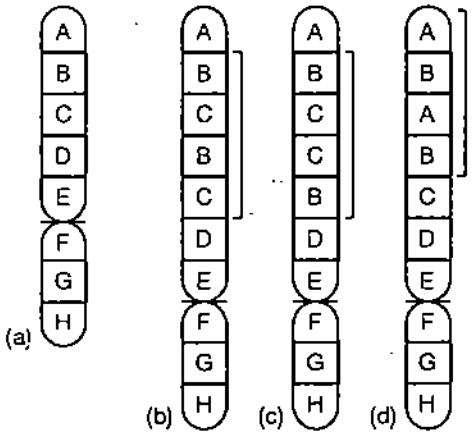


Fig. 9.6 : a) Normal chromosome. (b-d) Different forms of chromosome duplications : (b) tandem, (c) reverse-tandem, and (d) terminal tandem- configurations.

Small duplications do not usually have as drastic effects as deficiencies of comparable size. Thus a small duplication is rarely lethal even when homozygous, whereas small deficiencies may be lethal even in heterozygous condition. However, sometimes duplications can be pathological.

The first duplication to receive critical study was the bar eye variant in *Drosophila melanogaster*. The wild-type eye is oval shaped, and the bar eye phenotype is characterised by a narrower, oblong, bar-shaped eye with fewer facets. The classical studies of Bridges (1936) showed this trait to be associated with the duplication of a segment of the X chromosome, called section 16A as observed in salivary gland chromosomes. Each added section 16A intensifies the bar phenotype. However, the narrowing effect is greater if the duplicated segments are on the same chromosome. Supposing A represents one section 16A in a given X chromosome, we can recognise the genotypes and phenotypes listed in Table 9.2. These show clearly that the bar effect of a given number of duplicated 16A sections is intensified if the duplications occur in one X chromosome rather than being divided between the two X chromosomes of the female. Compare heterozygous ultrabar and homozygous bar eyes for example.

Such a change in the effect of the gene or genes in a chromosomal segment is known as the **position effect**. In bar eye, each added segment narrows the eye still further and this effect is enhanced as more duplications occur in one chromosome. Other duplications are known that produce the opposite effect, counteracting the effect of mutant genes. Moreover, duplications need not always be immediately adjacent to exert this position effect.

Table 9.2 : Comparison of genotypes and phenotypes for bar eye in *Drosophila* females.

X chromosomes	Phenotype	Mean number of facets
A/A	Wild-type	779
AA/A	Heterozygous, bar eye	358
AA/AA	Homozygous, bar eye	68
AAA/A	Heterozygous, Ultrabar	45
AAA/AAA	Homozygous, Ultrabar	25

A = One section 16A of the X Chromosome.

There are very few reports of duplication in human beings, as small duplications that have little or no phenotypic effect largely escape notice. One example of duplication is dup (9p). The affected persons are mentally retarded, having a good life expectancy.

SAQ 1

Assume a chromosome with the sequence ABCD.EFGH (. represents the centromere) shows the following aberrations. Identify the specific kind of aberration.

a) ABCD.EFH .....

b) ABCDCD.EFGH .....

SAQ 2

Name a well known deletion syndrome and present the deletion in chromosome symbolism.

.....

.....

9.4 INVERSIONS

Chromosomal inversions, another category of chromosomal structural variations, involve two breaks either in one arm of a chromosome or one break on either side of the centromere. Following the breaks, the chromosome fragments between the break points and rotates through an angle of 180°, and is reinserted into the chromosome in such a way that the genes are in *reversed order* (Fig. 9.7). For example, if a chromosome contained the following gene order ABCDEF and a break occurred between genes A and B and between genes D and E, after rotation and reinsertion, the chromosome would read ADCBEF. Figure 9.8 demonstrates how an inversion might arise. In this, the newly created "sticky ends" are brought close together and rejoined.

The inverted segment may be short or quite long and may or may not include the centromere. If the centromere is not part of the rearranged chromosome segment, the inversion is said to be **paracentric**. If the centromere is a part of the inverted segment, the term **pericentric** describes the inversion (see Fig. 9.8).

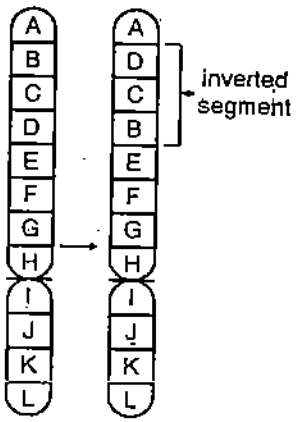


Fig. 9.7 : Inversion of the chromosome segment *BCD*.

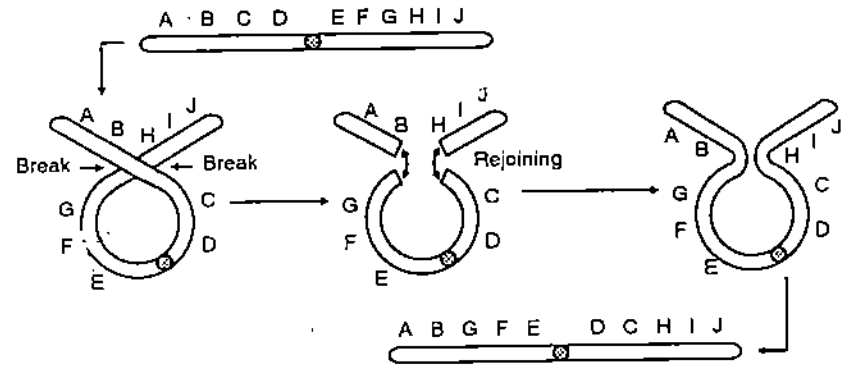


Fig. 9.8 : Diagram illustrating the mechanism by which some of the inversions might occur.

Although the gene sequence has been reversed in the paracentric inversion, the ratio of arm lengths extending from the centromere is unchanged. In contrast, some pericentric inversions create chromosomes with arms of different lengths than those of the noninverted chromosome. Thus, the **arm ratio** is often changed when a pericentric inversion is produced (Fig. 9.9). The change in arm lengths may sometimes be detected during the metaphase stage of mitotic or meiotic divisions.

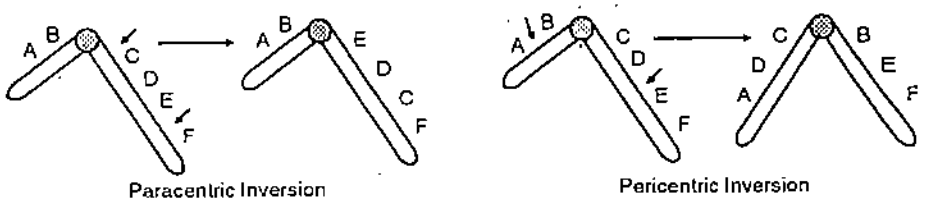


Fig. 9.9 : A comparison of the arm ratio before and after the paracentric and pericentric inversion.

Pericentric inversions are responsible for some of the changes in chromosome configuration that take place during evolution. For example, the human chromosome 17 is acrocentric, while the corresponding chimpanzee chromosome is metacentric (see Fig. 8.18 in Unit 8).

Inversions, in general, do not result in a loss or gain of genetic material, although these can be phenotypic consequences when inversions occur within genes or within regions that control gene expression. Genes when placed into a new setting (as happens in inversions) may lose their regulatory controls or alter the activity of genes in the area.

In individuals that are homozygous for a chromosomal inversion, the linkage order of genes becomes changed. A homozygote for the ADCBEF chromosomal arrangement will show A closely linked to D (rather than to B, as in the original sequence, ABCDEF) and E closely linked to B (rather than to D).

In inversion heterozyotes, synapsis of homologous chromosomes requires the formation of loops containing the inverted segments (Figs 9.10 and 9.11). Heterozygous inversions can be recognised by the presence of such loops, in the preparation of cells at the pachytene stage of meiosis.

Heterozygous inversions can also be identified in genetic experiments, since recombination is reduced significantly or suppressed. Actually, the frequency of crossing-over is not diminished in inversion heterozygotes in comparison with normal cells, but gametes derived from recombined chromatids are inviable, as illustrated in Fig. 9.10. In the figure, the effects of a single crossover in the inversion loop of an individual heterozygous for a paracentric inversion are shown. During the first meiotic anaphase, the two centromeres migrate to the opposite poles of the cell. As a result of the crossover between genes C and D in the inversion loop, one recombinant chromatid becomes stretched across the cell as the two centromeres begin to migrate, forming a dicentric bridge. As the two centromeres continue to migrate to the opposite poles in the cell, the dicentric bridge under physical tension, breaks randomly. The other recombinant product of the crossover event, a chromosome without centromere (an acentric fragment) is unable to continue through meiosis and is lost, that is, it is not found in the resultant gametes.

In the second meiotic division, the centromeres get divided, and the chromosomes are segregated to the four gametes. Two of the gametes have complete sets of genes and are viable : the gamete with the normal order of genes (ABCDEF), and the gamete with the

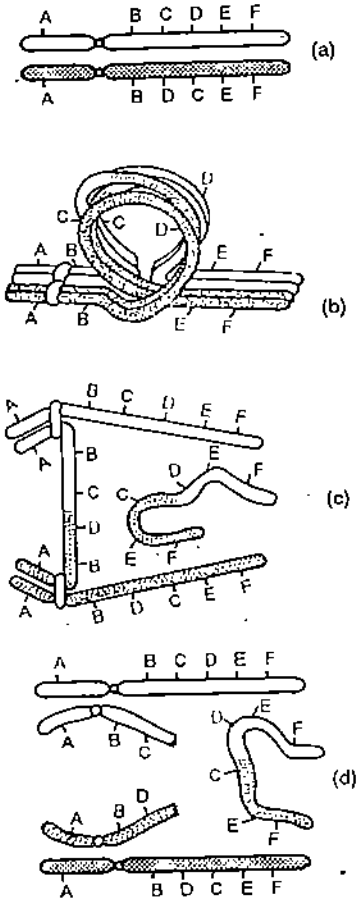


Fig. 9.10 : (a–d) : Crossing over in a heterozygote for paracentric inversion. a) Two homologous chromosomes, b) Pairing occurs during meiosis and crossing over takes place at the four-strand stage between two non-sister chromatids. c) Next there is separation of chromosomes during early anaphase of the first meiotic division. d) The resulting chromosomes. The chromosomal segment without a centromere fails to move toward the poles during meiosis I and is usually lost. The chromosome segment with two centromeres breaks and after meiosis II yields gametes with deletions.

inverted segment (ABDCEF). The other two gametes are inviable since they both are deficient for many of the genes of the chromosome. In general, the only gametes produced that can give rise to viable progeny are those containing *the chromosomes that were not involved in the crossover event.*

The effects of crossing over in an individual heterozygous for a pericentric inversions are shown in Figure 9.11. Of the four chromosomes resulting from the meiotic divisions, two form viable gametes, with the nonrecombinant chromosomes ABCDEF (normal) and ADCBEF (inversion), and two recombinant gametes that are inviable, each as a result of the deletion of some genes and the duplication of other genes. Usually only the gametes containing the two complete chromosomes can produce viable progeny, and usually, therefore, no genetic recombination is found in the progeny of heterozygotes for either pericentric or paracentric inversions.

It should be noted that not all crossover events within an inversion loop lead to inviable recombinants. One exception occurs, if a double crossover takes place in the inversion loop and the same two chromatids are involved in each crossover. A second exception occurs, when the duplicated and deleted segments of the recombinant chromatids do not affect gene expression and, hence, viability to a significant degree, as when the chromosome segments involved are very small.

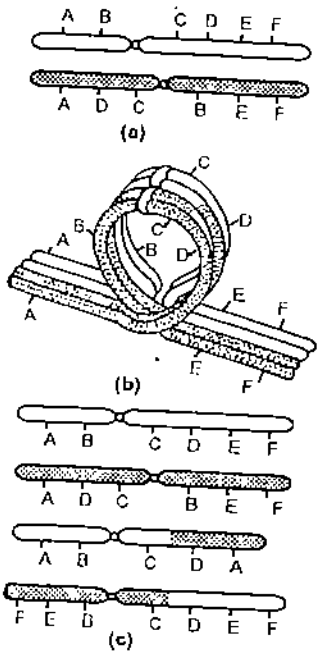


Fig. 9.11 : Crossing-over in a heterozygote for a pericentric inversion. a) Two homologous chromosomes. b) Crossing-over between two non-sister chromatids. c) The four resultant chromosomes. Only the top two chromosomes have complete sets of genes; they are non-crossover chromosomes with the same gene sequences as the two original chromosomes.

9.5 TRANSLOCATIONS

This is the fourth kind of structural alterations in chromosomes. It is also called a transposition, and involves the movement of a segment of a chromosome to a new place in the genome. Translocation may occur within a single chromosome – **intrachromosomal** (Fig. 9.12 a) or between non-homologous chromosomes – **interchromosomal** (Fig. 9.12 b,c). In the interchromosomal translocation there are two possibilities. One, there is transfer of one chromosome segment to another chromosome, and it is known as *non-reciprocal* translocation (Fig. 9.12 b). In the second case, there is reciprocal exchange of segments between the non-homologous chromosomes, and it is called a *reciprocal* translocation (Fig. 9.12 c). By definition, a reciprocal translocation is *balanced* if there is no loss of genetic material or the DNA, and *unbalanced* if there is loss of genetic material.

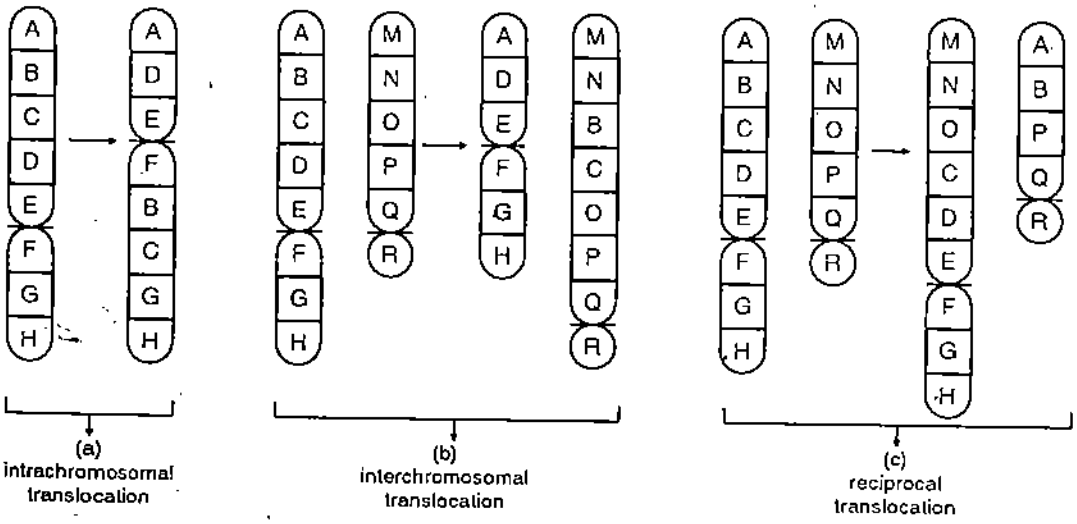


Fig. 9.12 : Translocations – a) intrachromosomal, b) interchromosomal nonreciprocal; c) reciprocal.

The origin of a relatively simple reciprocal exchange is illustrated in Fig. 9.13. The two nonhomologous chromosome arms come close to each other so that an exchange is facilitated. For this type of translocation in which the ends of chromosomes are involved, only two breaks are required. If the exchange includes internal chromosome segments, four breaks are required, two on each chromosome.

The genetic consequence of translocations is that, in homozygotes for translocations, the linkage relationships of genes are altered as compared to their normal counterparts. For example, in the nonreciprocal translocation shown in Fig. 9.12 a, the BC segment has moved to the other chromosome arm and has become inserted between the F and G

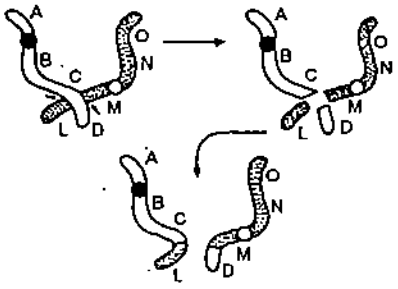


Fig.9.13: Origin of reciprocal translocation (zygotene stage).

segments. As a result, genes in the F and G segments are now farther apart than they are in normal situations, and the genes in the A and D segments are now more closely linked. Similarly, in reciprocal translocations, new linkage relationships are produced.

The consequences of translocations to the production of meiotic products depend on the type of translocation involved. In many cases, some of the gametes produced have duplications and/or deletions, and hence, are inviable. We concentrate here on reciprocal translocations, since they are the most frequent and are found in 1 in 500 live births.

In homozygotes for a reciporcal translocation, meiosis proceeds normally since all chromosome pairs can synapse to produce bivalents, whereas in heterozygotes for reciprocal translocations, unorthodox synapsis occurs during meiotic prophase-I. This results in the formation of cross–like configurations, owing to the requirements of pairing between homologous chromosome parts (Fig. 9.14). Instead of bivalents, that is, pairs of synapsed homologous chromosomes, there appear quadrivalents consisting of four associated chromosomes, each chromosome being partially homologous to two other chromosomes in the group. Segregation at meiotic anaphase may occur in three different ways, producing six types of gametes. Of these gametes formed only the two (as shown on the left in Fig. 9.14) – which result from 'alternate' segregation, that is, the two chromosomes derived from the same parent go to the same pole – contain all the chromosomal parts that too only once and no duplications or deletions occur. All the other gametes (the remaining four types, see Fig. 9.14), which result from 'adjacent' segregation in which the two chromosomes going to the same pole are derived one from each parent – have some chromosome segments duplicated and some deleted. These, therefore, cannot form in normal progeny.

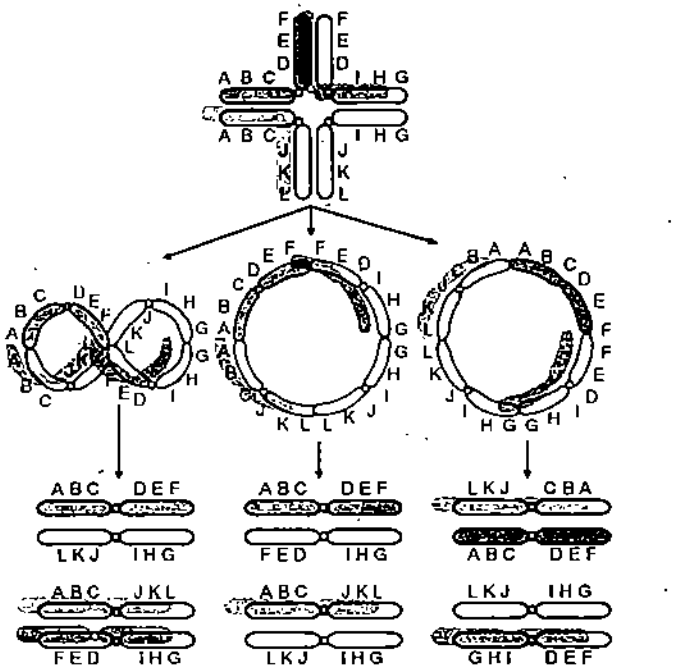


Fig.9.14: Meiosis in a heterozygote for reciprocal translocation. At the top is the cross-shaped configuration formed when the chromosomes of translocation heterozygotes pair during prophase of meiosis I. The second row shows the three configurations that may occur at metaphase I. The two lower rows show the six types of gametes that may be formed. Only the two, as shown on the left, resulting from "alternate" segregation contain complete sets of genes; the other four types, which result from adjacent segregation contain some duplicated and some deleted chromosome segments.

Translocation heterozygotes are semisterile, owing to the production of abnormal gametes. In plants, pollen grains, with duplicated or deleted chromosome segments, are usually aborted. Animal gametes with duplicated or deleted chromosome parts may function, but the zygotes formed by such gametes usually die. However, if the duplicated and deleted chromosome segments are small, the gametes may function and produce viable offspring.

9.5.1 Multiple Translocation Systems

In nature, there are many living organisms in which translocations involve more than two nonhomologous pairs of chromosomes. Such instances are more abundant in plants, and a remarkable example is *Oenothera lamarckiana*. It is heterozygous for translocations involving 12 of its 14 chromosomes (see Fig. 9.15).

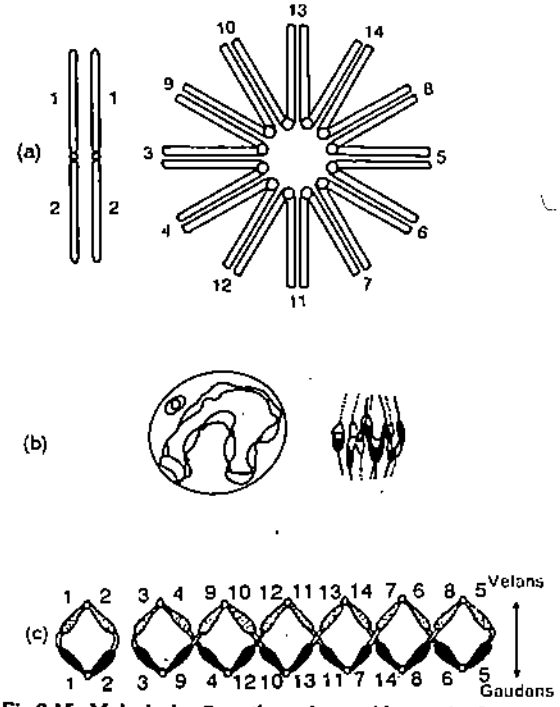


Fig.9.15: Meiosis in *Oenothera lamarckiana*. a) shows pairing between the nontranslocated pair of chromosomes and the six pairs of translocated chromosomes. b) The chromosomes during the prophase of meiosis I (left side); alternate chromosome segregation during anaphase (right side). c) Chromosome segregation in terms of the chromosome arms of the two complexes : velans and gaudens (After M.W. Strickberger, *Genetics*, Third Edition. 1985).

If we represent the 14 arms of the seven nontranslocated chromosomes as 1-2, 3-4, 5-6, 7-8, 9-10, 11-12 and 13-14, all the 14 chromosomes can be represented as follows : 1-2, 1-2, 3-4, 4-12, 12-11, 11-7, 7-6, 6-5, 5-8, 8-14, 14-13, 13-10, 10-9 and 9-3. With the exception of the first chromosome pair (1-2, 1-2) the homologues of the two arms of each chromosome appear one in each of two different chromosomes. For example, the homologues of 3-4 are found one in 4-12 and the other in 9-3. Pairing at meiosis produces one bivalent and one dodecavalent, consisting of 12 chromosomes in a multiarmed, star-like configuration.

Segregation during meiosis produces only two kinds of functional gametes in *O. lamarckiana* : 1-2, 3-4, 9-10, 12-11, 13-14, 7-6 and 8-5 (called velans) and 1-2, 3-9, 4-12, 10-13, 11-7, 14-8 and 6-5 (called gaudens). An interesting thing, however, is that except for the appearance of occasional 'mutants', *O. lamarckiana* breeds true. Only velans/gaudens heterozygotes and no velans/velans or gaudens/gaudens homozygotes are produced. This is due to the existence or recessive lethals in each of the two gametic complexes that make the homozygous combinations lethal.

9.5.2 Robertsonian Translocations

These are named after their discoverer, W.R.B. Robertson (1911), and are also referred to as **centric fusion translocations**. These result from chromosome breakage near the centromeres of the acrocentric chromosomes (Fig. 9.16). The larger segments fuse at their centromeric region, and produce a new large, submetacentric or metacentric chromosome. And the small segments may also fuse to form smaller chromosome, or the small fragments may not fuse and are eventually lost from the cell. The smaller chromosome is generally lost from the cell without consequence, because the short arms of acrocentric chromosomes contain heterochromatic or nonessential DNA.

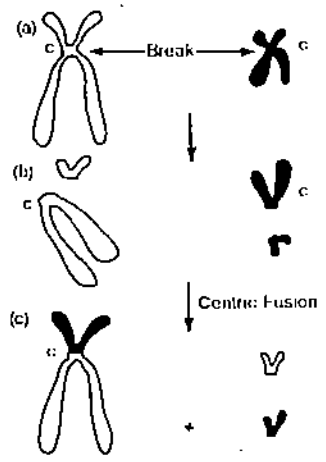


Fig. 9.16 : Origin of Robertsonian translocation. a) Two independent breaks occur at the centromeric region on two nonhomologous chromosomes. b) Free chromosome segments produced. c) Centric fusion of long arms of the two acrocentric chromosomes, the short arms may or may not fuse.

The chromosome pairs 13, 14 and 15 of the D group, and pairs 21 and 22 of the G group undergo such translocations. Robertsonian translocations may occur among the D group or G group chromosomes and between the D group or G group chromosomes. Persons who carry a balanced Robertsonian translocation will have only forty five chromosomes.

This aberration may produce only minor deficiencies, and in many cases, the affected individual may survive. First such case in humans was discovered in 1959 in an abnormal child, who was mentally retarded and suffered from severe spinal malformations. The condition was named **polydysspondylie**. Karyotypic analysis showed that the child's cell nuclei had only 45 chromosomes because of centric fusion of a small acrocentric chromosome of the G group and a large acrocentric chromosome of the D group. Deletions occurring during this translocation are apparently responsible for this phenotype.

Two other types of abnormal chromosome structures : ring chromosomes and isochromosomes deserve a brief discussion even though they are less common than deletions, duplications, inversions and translocations. These are being discussed in the following sections.

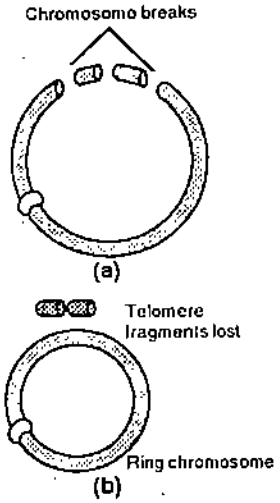


Fig.9.17 (a,b): Origin of ring chromosomes.

9.6 RING CHROMOSOMES

Ring chromosomes, i.e., a chromosome that forms a complete circle (Fig. 9.17) are formed following a break near the ends of the short and the long arm of a single chromosome. The two 'raw' or 'sticky' ends form an end-to-end reunion that result in a ring formation (Fig. 9.17). The two distal fragments, being acentric, are lost. Ring configurations are abundantly seen in the chromosomes 5, 10, 13, 15, 18 and 22. The presence of a ring is a qualitative proof that material has been deleted from both arms, whereas comparable deletion not involved in ring formation might be too small to be detected by cytological techniques. The nature of phenotypic defects associated with ring chromosomes depends on the specific genes lost in the deleted fragments.

Although ring formations are rarely observed in live born infants. Where they have been found, the newborn was birth defective and demonstrated a variety of mental and physical abnormalities.

How to represent ring chromosome in a symbolic form? Let us understand this with this example: 46, XXr(2)(p21q31) means loss of the material beyond bands 2p21 and 2q31, with the joining of the proximal segments to form a ring. This symbol can also be written in detail as 46XX, r(2)(p21 q31).

9.7 ISOCHROMOSOMES

Isochromosome is a chromosome in which the two arms are identical. It arises by an error in the plane of centromere division, yielding two metacentric, daughter chromosomes, each of which carries the information of one arm only but present twice. This process is illustrated in Fig. 9.18. Formation of isochromosomes can presumably occur either during mitosis or meiosis.

There can be isochromosomes made up of short arms or those made up of long arms. Usually, only the long-arm isochromosomes survive. The formation of isochromosomes out of any autosome, with the possible exception of the acrocentrics, is generally lethal due to the resulting genetic imbalance.

Isochromosome formation by the long arm of the X chromosome is a common abnormality. Such 46, X, i(Xq) patients have defects that are typical of the Turner's syndrome, and usually cannot be distinguished clinically. They are sex chromatin positive, the i(Xq) chromosome regularly forming the sex chromatin body. This is expected, since the i(Xq) lacks the genes of the short arm. At the time of X inactivation, those cells in which the i(Xq) is active would not survive. The remaining cells, in which the normal X is active, would form the embryo. But these cells would have only one copy of the distal segment of Xp and, hence, they would be functionally equivalent to 45, XO cells.

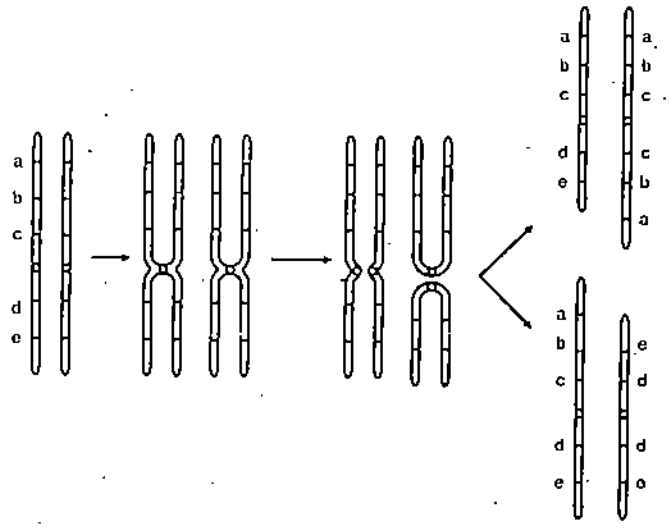


Fig. 9.18 :Formation of isochromosomes by division of the centromere in the wrong plane. Since the two arms of an isochromosome are identical, the chromosome always is metacentric.

SAQ 3

Name the two types of inversions and state which portions of the chromosome they involve?

.....

.....

.....

.....

.....

SAQ 4

What is an unbalanced translocation?

.....

.....

SAQ 5

To detect a translocation, the ..... of the progeny of various crosses must be compared.

a) phenotypes, b) genotypes, c) gametes, d) karyotype. (Fill in the correct choice in the blank space)

SAQ 6

The chromosomes in a translocation heterozygote consist of two original chromosomes : ABCDEF and GHIJKL, and two translocated chromosomes : ABCJKL and GHIDEF. Which of the following choices is a viable chromosome combination?

a) ABCDEF, FEDIHG

b) ABCJKL, LKJIHJ

c) ABCJKL, FEDIHG

d) LKJIGH, FEDIHG

Put a tick mark (√) on the correct answer.

SAQ 7

What is a Robertsonian translocation? Are they generally lethal?

.....

.....

.....

.....

SAQ 8

How many chromosomes will a carrier parent of a Robertsonian translocation have? A carrier of a reciprocal translocation?

.....

9.8 SUMMARY

In this unit you have learnt that:

- Most structural abnormalities in chromosomes are of one of four types — duplications, deletions, inversions and translocations.
- The basic mechanism of structural aberration consists of the breaking and rejoining of chromosomes.
- Linear linkage relation among genes in chromosomes are altered by structural aberrations.
- *Deletions* and *duplications* – both involve a change in the amount of DNA on chromosomes; *inversions* – a change in the arrangement of chromosomal segment; *translocations* – a change in the location of one or more DNA segments.
- An organism can often tolerate an imbalance of gene dosage resulting from small duplications or deletions, but large duplications or deletions are always harmful. In inversions though all the genetic material is present (in reverse order) during meiosis, crossing-over within an inversion loop produces unbalanced gametes. Organisms that contain a reciprocal translocation as well as normal homologous chromosomes of the translocation, exhibit semisterility because of abnormal segregation of chromosomes in meiosis.
- A Robertsonian translocation is a type of non-reciprocal translocation in which the long arm of two acrocentric chromosomes are attached to a common centromere.
- Chromosomes can often form rings if breaks occur very near the ends of the arms, with rejoining of the proximal ends to each other. Foetuses with ring chromosomes are often live born if the segments lost are small and have few loci.
- Isochromosomes form when the centromere divides in the wrong plane to produce daughter chromosomes each with two identical arms. Such a duplication/deficiency is not viable except in the case of sex chromosomes.

9.9 TERMINAL QUESTIONS

1) In the table given below, the left and the middle columns give details of the chromosomes. Identify and write in the column, on the extreme right, the type of chromosomal aberration.

S.No.	Description	Gene changes	Type
a)	ABCDEFGH	None	Normal
b)	broken segment reattached to original chromosome in reverse order	ABFEDCGH	—?—
c)	broken segment becomes attached to a nonhomolog resulting in a new linkage relations	LMNOPQRCDEFGH	—?—
d)	broken segment becomes attached to homolog that has experienced a break; homolog then bears one block of genes in duplicate	ABCDEFGEFGH	—?—
e)	no rejoining; chromosomal segment lost	ABFGH, CDEFGH etc.	—?—

2) How can (a) paracentric and (b) pericentric inversions act as "crossover suppressors"? (c) Is crossing-over really suppressed?

3) The chromosomes in a translocation heterozygote consist of two original chromosomes : ABCDEF and GHIJKL, and two translocated chromosomes : ABCJKL and GHIDEF. Make a diagram and explain the viable and lethal chromosome combinations.

4) What type of abnormal chromosome forms a bridge at anaphase? What kind of chromosomes cannot move during anaphase?

5) Which abnormality in chromosome structure can change a metacentric chromosome into a submetacentric chromosome? Which can fuse two acrocentric chromosomes into one metacentric chromosome?

6) What (a) genetic and (b) cytological evidence would indicate that a translocation was present in a plant material, such as, barley?

7) How is pollen sterility associated with translocations?

8) Correctly associate the items of column A with those of column B.

A		B	
a)	Chri-du-chat	1)	deletion
b)	*Oenothera lamarckiana*	2)	duplication
c)	bar eye in *Drosophila*	3)	inversion
d)	during synapsis, loops formed, containing inverted segments	4)	translocation
e)	terminal deletions and eventual rejoining of the sticky ends	5)	isochromosome
f)	error in centromere division	6)	ring chromosome

9.10 ANSWERS

Self - assessment Questions

1) a) deletion

   b) duplication

2) Cri-du-chat, 5p–

3) Paracentric and pericentric. Pericentric inversions involve the centromere and a section of each of the chromosome arms rotating 180 degrees. Paracentric inversions involve a 180 degree rotation of a chromosome fragment (without centromere) in either the long arm or short arm of a chromosome.

4) An unbalanced translocation is one in which essential DNA is lost from the cell.

5) d

6) c

7) A Robertsonian translocation is one involving the long arm of the two acrocentric chromosomes. It is generally not lethal because there is no loss of essential DNA.

8) 45 chromosomes, 46 chromosomes.

Terminal Questions

1) a) inversion

   b) translocation

   c) duplication

   d) deletion

2) a) When crossing over occurs in the area of a paracentric inversion, acentric and dicentric chromosomes occur that do not separate properly to the poles, during cell division. Gametes carrying crossover chromatids are abnormal and inviable.

   b) Crossovers within pericentric inversions result in unbalanced chromosome arrangements that make the crossover gametes or zygotes inviable.

   c) Crossing-over is reduced to some extent, but 'suppression' mainly results from inviable gametes or zygotes.

3) See Fig. 9.14. The homologous parts of a chromosome will associate with each other. As shown in the figure, such an association gives the gene sequence a definite order in a ring-shaped structure. At anaphase I of meiosis the ring is broken and the two chromosomes move to each pole of the spindle. If the chromosomes were distributed at random, six different types of gametes would form. The only sets of gametes that would be viable would be the ones that had the complete genetic sequences : ABCDEF, GHIJKL and ABCJKL, FEDIGH. The other types of gametes would be unbalanced and, hence, nonviable since they do not have a complete set of the genetic information. For instance, in the gamete set ABCDEF, ABCJKL, the sequence ABC occurs twice while the sequence GHI is missing entirely. The other gametes are lethal for similar reason.

4) A chromosome with two centromeres – dicentric; a chromosome with no centromere – acentric.

5) An inversion that includes the centromere, but with breakpoints that are not symmetrical around the centromere, moves the centromere from a central location in a metacentric chromosome to a noncentral location, forming a submetacentric chromosome. A metacentric chromosome can be formed from two chromosomes that have long arms about equal in length, by a Robertsonian translocation.

6) a) Altered linkage groups, pollen or ovule sterility and position effects suggest the presence of translocations.

   b) Cross configurations, rings and chains of chromosomes in the meiotic prophase that can be seen microscopically represent cytological evidence.

7) Unbalanced chromosome arrangements occur in gamete formation making pollen inviable.

8) a) 1

   b) 4

   c) 2

   d) 3

   e) 6

   f) 5

UNIT 10 NUMERICAL ABNORMALITIES IN CHROMOSOMES AND THEIR EFFECTS

Structure

10.1 INTRODUCTION

As you know, each species is characterised by a particular number of chromosomes. This chromosome set or **genome** in higher plants or animals is usually represented once in their haploid cells (e.g., gametes, spores etc.) and twice in their diploid (somatic) cells. It was pointed out in the previous unit that irregularities sometimes occur in nuclear divisions or 'accidents' (from radiations, chemicals and other agents) may befall interphase chromosomes so that cells or entire organisms with aberrant genomes may be formed. In the earlier unit, you have studied the genetic consequences of aberrant genomes caused mainly due to changes in chromosome structure. This unit will deal with the genetic effects of numerical chromosomal abnormalities.

Of the two classes of chromosome abnormalities, change in chromosome number is more frequently observed among the live-born children. A numerical change is also cytologically the easiest to observe. Basically, there are two types of numerical changes in cell chromosome number. These are **euploidy**, in which there is an addition or deletion of entire set(s) of chromosomes, and **aneuploidy**, which involves the addition or deletion of individual chromosomes within a set. Both euploidy and aneuploidy thus alter **gene dosage** in the affected individuals. In this unit we will examine the numerical chromosomal aberrations in detail.

Objectives

After studying this unit you should be able to:

- distinguish between euploids and aneuploids (Sections 10.2 and 10.3);
- list the possible ways of origin of euploidy, both monoploidy, and polyploidy (Section 10.2);
- explain the reasons of the prevalence of polyploidy in plants and its non-occurrence in animals (Section 10.2, Subsection 10.2.2);

- list and describe the main features for differentiating an autopolyploid from an allopolyploid (Section 10.2, Subsection 10.2.2);
- justify with suitable supporting examples the statement that "allopolyploidy has contributed immensely in the formation of new species, whereas autopolyploidy represents only the dead-ends with reference to species evolution" (Section 10.2, Subsection 10.2.2);
- explain the causes of aneuploidy (Section 10.3);
- describe the phenomenon of nondisjunction (Section 10.3);
- explain the variations in chromosome number (both autosomes and sex chromosomes) in human beings and the consequent phenotypic effects (Sections 10.4 and 10.5).

10.2 EUPLOIDY

The term 'euploidy' refers to the condition when a complete set (or sets) of chromsomes are present in an eukaryotic organism or in its gametes. When one complete set of chromosomes is present, the condition is called **monoploidy** (n). If two complete sets are present it is termed **diploidy** (2n), and when more than two sets of chromosomes are present it is called **polyploidy**. Let's discuss these one by one in the following subsections.

10.2.1 Monoploidy

Monoploidy is also referred to as haploidy. You already know that the n (one set) chromosome number is usual for gametes of diploid animals, but is unusual for their somatic cells. Monoploidy is seldom seen in animals, but is found in the male honeybee, ants and other insects in which male haploids develop from unfertilised eggs, a phenomenon known as parthenogenesis.

In contrast, plants have a gametophyte stage in their life cycle that is characterised by the reduced (n) chromosome number. In higher plants this stage is, however, brief and inconspicuous. In some lower plants, e.g., algae, fungi, bryophytes (liverworts and mosses) the monoploid phase constitutes the dominant part of the life cycle. The monoploid plants can be recognised in natural populations by observation and verification through cytological preparations. These plants are usually frail in appearance, and have small leaves, low viability and high degree of sterility. Sterility is largely because of *irregularities at meiosis*. Since only one set of chromosomes is present therefore, no pairing is possible. And if the meiotic process succeeds at all, the dispersal of chromosomes to the poles is irregular and the resulting gametes are highly inviable. The monoploids are valuable experimental tools because they undergo no segregation and carry a single set of genes.

10.2.2 Polyploidy

The euploids with three or more complete sets of chromosomes or genomes are called polyploids. More specifically, the polyploid organisms can be classified as *triploids* (3n) with three sets of chromosomes, *tetraploids* (4n) with 4 sets of chromosomes, *pentaploids* (5n) with five sets of chromosomes and so on.

Polyploidy results due to errors during meiosis or fertilisation. Given below are four ways of the origin of polyploid condition. **First**, double fertilisation of one egg by two sperms (*dispermy*) or fusion of an egg, a polar body and a sperm results in triploidy. **Second**, an unreduced (diploid or 2n) gamete uniting with a normal (haploid or n) gamete also produces triploidy. **Third**, many a time, tetraploidy occurs after fertilisation due to a mitotic error. Two cycles of chromosomal replication (*endoreduplication*) without cell division result in a tetraploid cell. **Fourth**, fusion of two diploid cells is another process whereby tetraploidy may result. In human beings, nearly all triploid and tetraploid foetuses are spontaneously aborted or are still-born.

Polyploidy is rather common in the plant kingdom, but is rare in bisexual animals. Most animals have well-defined separate sexes characterised by the presence of specific sex chromosomes. Multiple chromosome sets lead to variations in the numbers of the sex chromosomes (for example XXXY or XYYY), a situation incompatible with the normal sex-determining mechanism. As a result, the animal polyploids are restricted to groups that reproduce asexually or are hermaphroditic or parthenogenetic. Examples are the earthworms, some types of shrimp, and the parthenogenetic species among the insects and lizards.

Polyploidy has played a significant role in the evolution of plants. Polyploidy is possible in plants because of the general absence of sex chromosomes, and a large number of plant species being monoecious. Plants have an additional feature, that is, reproduction by self-fertilisation or by asexual means, which enables the partially sterile polyploid species to propagate. The examples include many mosses and ferns and a few gymnosperms, and almost half the angiosperms. New polyploid varieties of cultivated and ornamental plants are constantly developed and introduced by the plant breeders. Some examples of important polyploid crop plants are: apple, banana, barley, coffee, cotton, peanut, potato, strawberry, sugarcane, tobacco and wheat.

The phenotypic effects of polyploidy are generally those of increased vigourous growth and size, hardiness, ability to occupy less favourable habitats and overall superiority as compared to their normal counterparts. One reason for such superiority is the fact that the expression of any deleterious or lethal allele can be effectively masked by the activity of the two or more normal alleles that are present in the genotype. However, because of meiotic irregularities, polyploidy is often accompanied by partial or complete sterility, caused by nonfunctional gametes, particularly the male gametes.

Autopolyploidy and Allopolyploidy

We have classified the polyploids on the basis of the number of sets of chromosomes present. We can also divide them into two general groups according to the source of their chromosome sets. Polyploidy derived from the multiplication of a single ancestral set of chromosomes is known as **autopolyploidy**. However, in many cases, the polyploid species have complete sets of chromosomes from two or more different ancestral species. Such polyploids are known as **allopolyploids** and they originate from occasional hybrids that occur between distinct diploid species. Let us examine these two states in detail.

How to identify autopolyploids? The presence of varying numbers of quadrivalents, that is, four homologous chromosomes instead of the usual two pairing with one another during meiosis, suggests autopolyploidy. The unequal segregation of chromosomes in the quadrivalents is the main reason why the autopolyploids are sterile to varying degrees. In normal meiotic prophase, chromosomes pair with one another throughout their length. But in situations, when four similar chromosomes are present, they usually pair with different chromosomes, at different places along their length. This complicates the meiotic process and results in the nonviable gametes or zygotes. Further, sterility in autopolyploids also results due to unequal disjunction. *Chrysanthemum* is a well known example of autopolyploidy. In this genus, the basic number is 9. Its polyploid species are known to have 18, 36, 54, 72 and 90 chromosomes — all multiples of 9. Potato or *Solanum tuberosum* is another example. Its basic chromosome number is 12. Its autopolyploid forms, with chromosome numbers 24, 36, 48, 60, 72, 96, 108, 120 and even 144, are available.

Allopolyploidy is the most common natural form of polyploidy that has occurred in various plant groups. Many of the present-day plants have resulted from this kind of hybridisation. The polyploid species, that are established in nature, have genomes that correspond more or less completely to the combined chromosome complements of two different but related diploid plants (Fig. 10.1). In the allopolyploids, since two homologues of each specific chromosome are present, meiosis may occur normally and fertilisation may successfully propagate the tetraploid sexually. These allopolyploids, also known as the *allotetraploids* or *amphidiploids*, have undergone hybridisation somewhere in their ancestral history. Allopolyploidy thus represents a method by which new species may be formed almost immediately, whereas the autopolyploids result in meiotic anomalies and "dead ends" with reference to evolution.

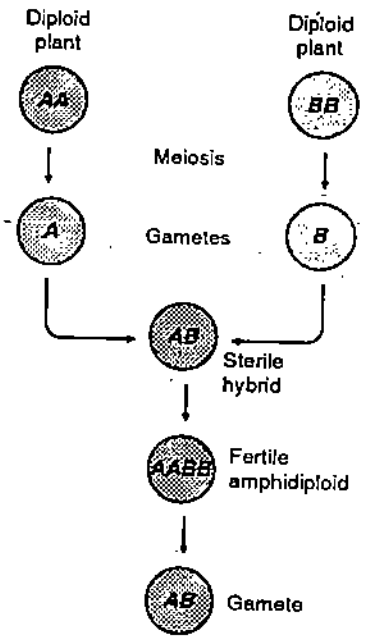


Fig.10.1 : The production of a fertile allotetraploid (AABB) by first crossing the diploids AA and BB, and then doubling the chromosomes of the sterile hybrid (AB)

Cultivated wheat is an excellent example of allopolyploidy. Cultivated bread wheat is a hexaploid with 42 chromosomes constituting a complete diploid genome of 14 chromosomes from each of the three ancestral species. The 42-chromosome allopolyploid is thought to have originated through hybridisation as shown in Fig. 10.2.

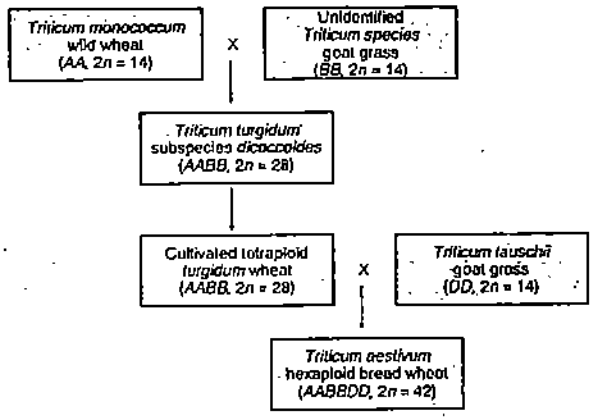


Fig. 10.2: Hybridisations that occurred in the ancestry of cultivated bread wheat (*Triticum aestivum*), which is an allohexaploid containing complete diploid genome (AA, BB, DD) from the ancestral species.

Another allopolyploid – *Raphanobrassica* was experimentally synthesised by G.D. Karpechenko crossing two common vegetables that belong to different genera. These are radish – *Raphanus sativus* and cabbage – *Brassica oleracea*. Both the species have 9 pairs of chromosomes. The diploid hybrid has 18 chromosomes, 9 from each parent, but is sterile and cannot perpetuate itself, largely because of the failure in pairing between the unlike chromosomes in meiosis. When the chromosomes of the F_1 hybrid were doubled, a fertile polyploid, named *Raphanobrassica*, was produced with 18 radish and 18 cabbage chromosomes. Because two sets of chromosomes were now present from each parent, pairing was quite regular and hence normal gametes were produced. This experiment suggested the possibility of incorporating desirable genotypes from two different species into a new polyploid species. Seed capsules of the parent, the sterile hybrid, and the tetraploid plants are shown in Fig. 10.3. Unfortunately, from the practical standpoint, *Raphanobrassica* has a foliage of a radish and the roots of a cabbage. But nevertheless, this experiment has shown a novel way of perpetuating the sterile hybrids.

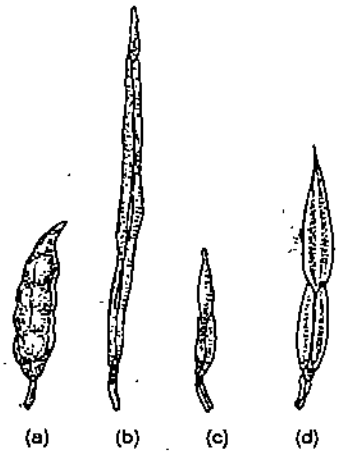


Fig.10.3: Seed pods of (a) radish, (b) cabbage, (c) sterile hybrid, d) fertile tetraploid resulting from chromosome doubling.

Induced Polyploidy

Polyploids may arise naturally or may be artificially induced. Anything that interferes with spindle formation during mitosis results in a doubling of chromosomes. Induced polyploidy was first demonstrated by subjecting growing plants to a higher than usual temperature. Maize and some other plants responded to such treatment with an increase in chromosome number in their cells. Application of the alkaloid colchicine, derived from the autumn crocus (*Colchicum autmnale*) disturbs spindle formation. You have studied the use of this chemical in karyotype preparation in Unit 8, Section 8.3. The cell division proceeds as usual, but since the spindle formation is inhibited, the cytoplasmic phase of the cell division does not occur. Instead, restitution nuclei with the double number of chromosomes are produced in the treated tissue. Subsequent nuclear divisions are normal, so that the polyploid cell line, once initiated, is maintained. Polyploidy may also be induced by other chemicals (acenaphthene and veratrine) and also by exposure to cold temperatures.

Applications of Polyploidy

Bread Wheat – Without polyploidy we would have a very different diet. Among the food we would miss might well be bread which is made from one of several species of wheat (*Triticum*). Wheat is the most important grain crop as over 20% of calories consumed by human beings come from wheat. Presently, several varieties of wheat are grown throughout the world.

The genus *Triticum* consists of a polyploid series of wheat. The diploid species, *T. monococcum* or *einkorn* wheat 2n = 14, was the first to be used for food about 10,000 - 12,000 years ago, about the time civilisation began. At that time, seeds of wild plants were collected for food and subsequently brought in cultivation. The *einkorn* wheat is hulled and is small as compared to its polyploid relatives.

The *emmer* and *durum* wheat belong to the tetraploid species *T. turgidum* (2n = 18). These wheats originated from a naturally occurring wild tetraploid variety, whose grains are larger than einkorn wheat. The *emmer* wheat is hulled, but the durum wheat is particularly desirable because of its free threshing quality, that is, the hull is easily removed by threshing; and unlike hulled wheat it does not need to be ground for human consumption.

Varieties of bread wheats *T. aestivum* are agriculturally the most important. Bread wheat is hexaploid, 2n = 42, derived from hybridisation of tetraploid *emmer* wheat with a diploid relative, the wild grass *Aegilops squarrosa* (2n = 14). Bread wheat is especially desirable, because it is free-threshing and has large grains, and also has a high content of proteins called the glutens. The glutens are sticky and elastic, allowing wheat breads to rise without collapsing when leavened, i.e., when mixed with yeasts which produce carbon dioxide gas.

Tobacco Resistance – Induced polyploidy has been exploited to a great extent. Practical applications may become more common as additional data has accumulated. By artificially inducing polyploidy, disease resistance and other desirable qualities have been incorporated into some commercial crop plants. Tobacco, *Nicotiana tabacum* is susceptible to the tobacco mosaic virus (TMV), whereas another species, *N. glutinosa*, appeared at first observation to be resistant. Further investigation, however, showed that in *N. glutinosa* the virus killed the cells that were invaded, and the virus particles became isolated in the dead cell. The apparent resistance was due to hypersensitivity. When the two tobacco species were crossed, the hybrid was found to be 'resistant' to the virus, but was totally sterile. So, when the chromosomes were doubled, it was possible to secure a fertile polyploid resistant to the virus.

Polyploid Fruits and Flowers

Many varieties of plants that fulfil human needs have been found to be polyploids. The polyploids were selected for cultivation because of their large size, vigour and ornamental value. In *McIntosh apple* trees, some twigs produced extra-large fruits. These twigs were known as 'giant sports' and were found to be tetraploid (4n). These 'giant sports' were propagated into whole trees, which produced extra large fruits. Similar sports with giant fruits are also seen in *Bartlett pears*, several varieties of *grapes* and *cranberries*. These giant sports developed naturally. But with colchicine treatment, a number of polyploids have been developed artificially. This technique has provided a way to explore the mechanism involved in the polyploid formation. The tetraploid *maize* is more vigourous than the diploid form and contains about 20 per cent more vitamin A. Polyploid *watermelons* have been developed from colchicine treatment. The tetraploid with 44 chromosomes is large and commercially useful. The triploid watermelons with 33 chromosomes have great market potential because they are sterile and lack seeds. Several tetraploid flower varieties, such as in *marigolds* and *snapdragons* are widely cultivated.

SAQ 1

Polyploidy is rare in animals, yet some tissues in the bodies of some diploid animal species show evidence of polyploidy. Why do numbers above 2n persist in somatic tissues when they do not occur in the whole animal?

.....

.....

.....

SAQ 2

How does colchicine treatment result in chromosome doubling in plants?

.....

.....

.....

.....

.....

.....

10.3 ANEUPLOIDY

The aneuploids are individuals with unbalanced chromosome sets, that is, with extra or missing chromosomes. Their chromosome complement contains at least one more or one less chromosome than the normal diploid number. If one homolog of the normally diploid complement is present three times, the condition is called **trisomy**, and an individual of this type is called a **trisomic** and the condition can be designated by the symbol 2n+1. **Tetrasomy** (2n + 2) indicates that one chromosome of the complement is present four times. A 2n+1+1 individual would be **doubly trisomic** for two different homologs and the 2n+2+2 individual would be **doubly tetrasomic** and so forth. The reciprocal situation involving loss of a chromosome from a pair is called **monosomy** (2n – 1), and loss of a whole pair of chromosomes is known as **nullisomy** (2n – 2).

The primary cause of aneuploidy is **nondisjunction** of homologs during meiosis (see Box 10.1). Nondisjunction is the failure of chromosomes to separate during cell division (Fig. 10.4). If one pair of homologs fails to disjoin, then n + 1 and n – 1 gametes will result. And if these gametes are fertilised with a normal gamete, then 2n + 1 and 2n – 1 individuals would be formed instead of the normal 2n. The chromosome constitution of the zygote is critical to normal development, because any chromosome abnormality present in the zygote will be passed on through each succeeding generation of cells (the single cell line produced by meiotic nondisjunction).

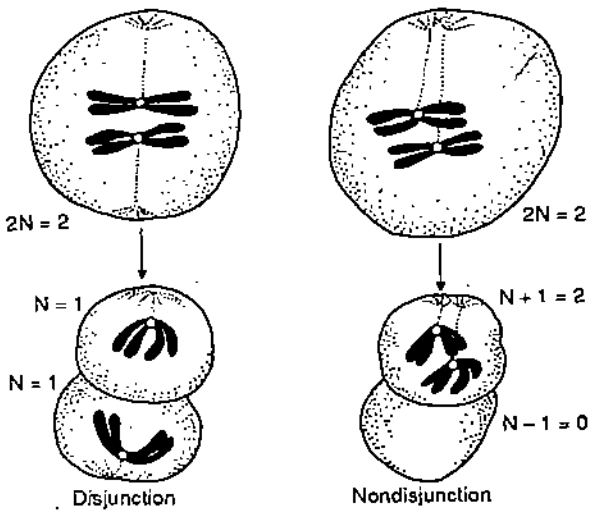


Fig.10.4 : The cell on the left shows a pair of homologous chromosomes during meiosis, with the centromere of each chromosome attached to its closest pole. Normal separation or disjunction converts the diploid (2n) to two haploid cells (n). On the right, each of the two homologs is attached to the same pole of the cell. This converts the diploid (2n) into nondisjunctional haploid cells: (n+1) and (n–1), one containing both homologs and one lacking both homologs respectively.

Box 10.1 : Meiotic nondisjunction

If chromosome nondisjunction occurs during meiosis I, it is referred to as first division nondisjunction (Fig. 10.5a). At meiosis II, it is called second-division nondisjunction (Fig. 10.5b). In the first-division nondisjunction, a pair of homologous chromosomes, (*either autosomes or sex chromosomes*) fail to separate from each other between first metaphase and first anaphase. In second-division nondisjunction, sister chromatids fail to separate from each other between metaphase II and anaphase II. In either case, nondisjunction results in some gametes (sperms and eggs) having two of a given chromosome, while the companion gametes lack that given chromosome. When such aneuploid gametes are involved with a normal gamete, in fertilisation, the resulting zygote would be monosomic or trisomic for the given chromosome.

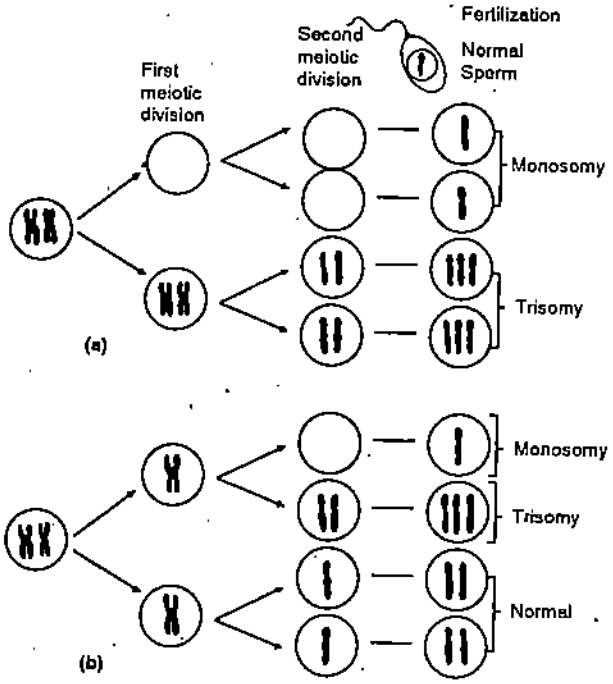


Fig. 10.5: First (a), and (b) second division meiotic nondisjunction. The diagram shows occurrence of nondisjunction in the egg, which is then fertilised by a chromosomally normal sperm. The reverse is also known to occur giving similar results.

All forms of aneuploidy have serious consequences in meiosis. In the following paragraphs, you will study some well-known examples of aneuploidy.

Trisomy in Jimson Weed

The jimson weed (*Datura stramonium*) shows a considerable amount of morphological (Fig. 10.6) variation in many traits, particularly the fruits. The normal chromosome number of this

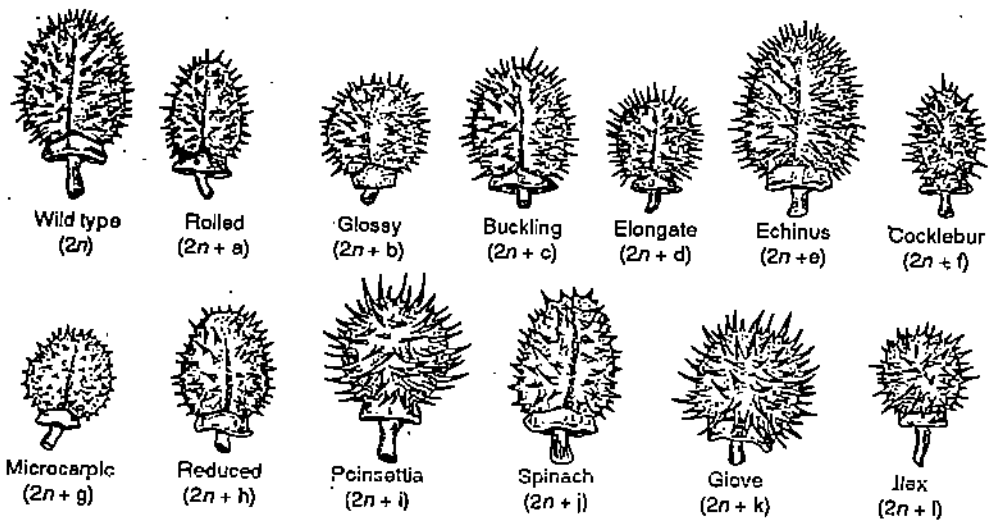


Fig. 10.6: Capsule phenotypes of the fruits of *Datura stramonium*. As compared with the wild type, each phenotype is the result of trisomy of one of the twelve chromosomes (denoted here by the alphabets a-l) characteristic of the haploid genome (After Blakeslee, 1934. J. Hered. Vol. 25, pp. 80-108).

plant is 2n = 24. Blakeslee and Belling (1924), who studied this system extensively showed that each of the particular morphological variants had 25 chromosomes. In each variant, one of the 12 kinds of chromosomes was found to be in triplicate. That is, the somatic cells were 2n+1, and it is a case of trisomy. These people thought, since the jimson weed has 12 pairs of chromosomes, 12 recognisable trisomics should be possible. And as expected, they could collect all the variants for the 12 chromosomes (see Fig. 10.6).

10.4 AUTOSOMAL ANEUPLOIDIES

After studying what is aneuploidy, its origin and the classic example of jimson weed, let us now take up some commonly found syndromes due to autosomal aneuploidy in man. The most common aneuploid conditions seen with appreciable frequency in live-born children are trisomics 13, 18, and 21. In order of increasing frequencies, these trisomies are Patau, Edward and Down syndromes respectively.

10.4.1 Trisomy 21 — Down Syndrome

It is one of the well-known and best studied examples of chromosome–caused anomalies in human beings. This condition, which is an autosomal trisomy, was first described by Langdon Down in 1866, and was originally called **mongoloid idiocy** or **mongolism**. Down chose this name because of the prominence of the epicanthic fold in the eyelid (see Fig. 10.7) a phenotype resembling of members of the mongoloid race.

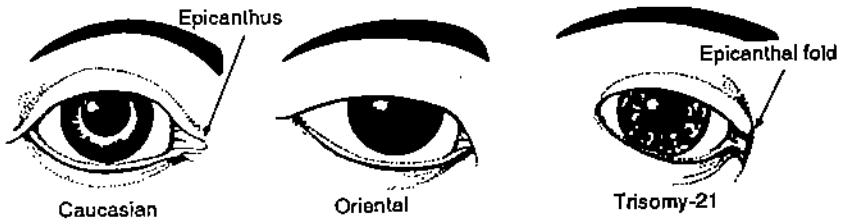


Fig.10.7: Trisomy – 21 and the epicanthal fold. The junction where the eye folds come together is called the canthus. In trisomy-21, a vertical fold of skin near the basal ridge (the epicanthus) pulls and tilts the eye slightly toward the nostrils. Note that the upper eyelid in the normal caucasian eye runs parallel to the skin fold below the eyebrow. In orientals, this skin fold covers a major portion of the upper eyelid. In caucasians with trisomy-21 the epicanthal fold does not cover a major portion of the upper eyelid, making their eyes distinct from the orientals. The eyes of persons with Trisomy-21 frequently show light patches on the iris. These spots are not found in the normal individuals.

This syndrome is the result of chromosome 21, one of the G group (Fig. 10.8) and is designated as 47, 21+. This trisomy is found in about 1 infant in every 600-700 live births.

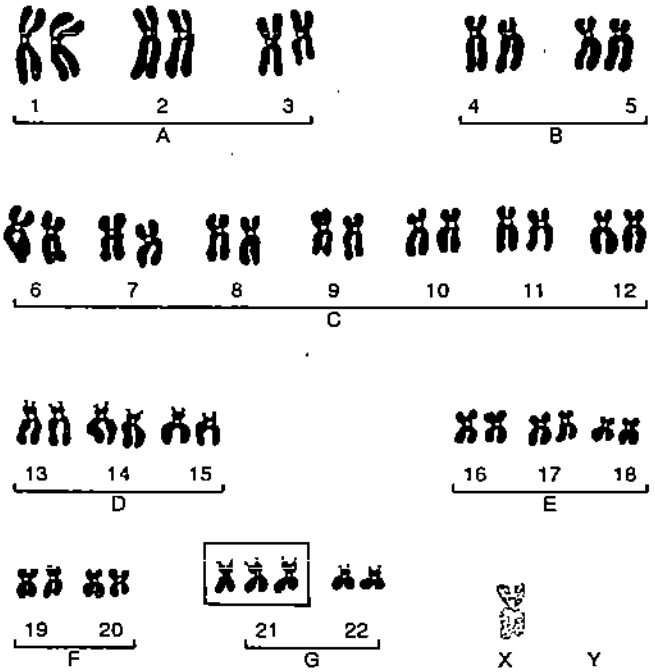


Fig.10.8 : Karyotype of a trisomic Down Syndrome male.

The Down Syndrome (DS) Phenotype — Some three hundred phenotypic characteristics have been associated with Down syndrome. A few dozen of these traits are present in most children suffering from this disorder, and many of them appear as siblings at first glance. A child with DS ages rapidly, has an IQ range of 25-70, and has a flat facial profile and a broad 'saddle' nose (see Fig. 10.9).

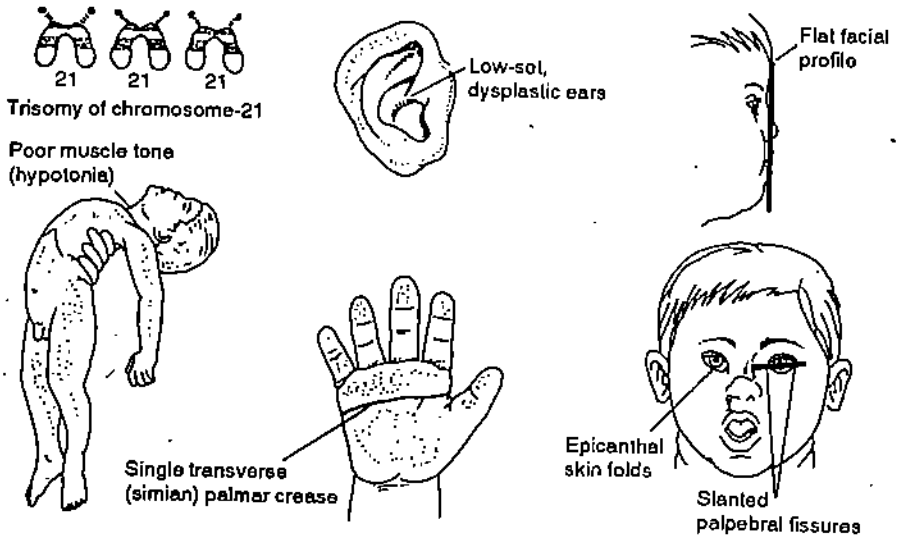


Fig.10.9: Down syndrome phenotype. Infants suffering from this syndrome have lax and hyperextensible joints. The flat facial profile, low-set, rounded ears, presence of simian crease in the palm and eyes slanted with epicanthal skin fold, are some of the characteristic features of this syndrome.

The ears are small and round; the tongue is large and flat. The child has a protruding lower lip; a broad short skull; thick, short hands with stubby fingers; and thick, short feet and body. As you have seen in Fig. 10.7, DS children usually have light-coloured specks in the iris of their eyes, referred to as *Brushfield spots*. About 30 per cent have a simian crease, and 70 per cent have heart trouble. Hernias and a marked susceptibility to infections are associated with the syndrome. They are usually retarded in physical development and are short-statured.

For example, a six-month old baby with DS often acts like a baby of about two months: slow to sit up, crawl and walk. Sometimes the child does not talk until the age of 6 years or more. In extreme cases some never develop speech. DS patients have a tendency to be diabetic and to develop leukemia than are persons in the general population. Thirty per cent of DS children die within the first year; 50 per cent within five years; very few reached the age of twenty before antibiotics were introduced. Presently, with the use of antibiotics, like span of DS children has been stretched into the forties.

Origin of Down Syndrome – One of the main ways in which this trisomic condition originates is through **nondisjunction of chromosome 21** during meiosis. Failure of paired homologs to disjoin during anaphase I or II can result in male or female gametes with the n+1 chromosome composition. Following fertilisation with a normal gamete, the trisomic condition results. Chromosome banding studies have shown that, while the additional chromosome may be derived from either mother or father, defective ovum is most often the source.

Maternal/Paternal Age and the Down's Syndrome – Clearly, DS syndrome more than any other syndrome, shows a marked relationship to maternal age (see Fig. 10.10). A female is born with all the oocytes she will ever produce – nearly 7 million. These remain in an arrested prophase I of meiosis from birth until ovulation, generally at the rate of one per menstrual cycle (about 28 days) after puberty. A given oocyte may remain in this state of suspended development for some 12 to 45 years or so. It has been observed that the accuracy of meiotic division decreases with age, and after thirty the frequency of nondisjuction associated with chromosome 21 increases substantially. For women over the age of 45, the probability of producing a DS child is about one in forty.

Higher than normal frequencies of children suffering from DS have also been discovered for mothers under the age of seventeen (Fig. 10.10). The exact reason, however, is not clear as yet. A convincing theory that may explain the relatively high

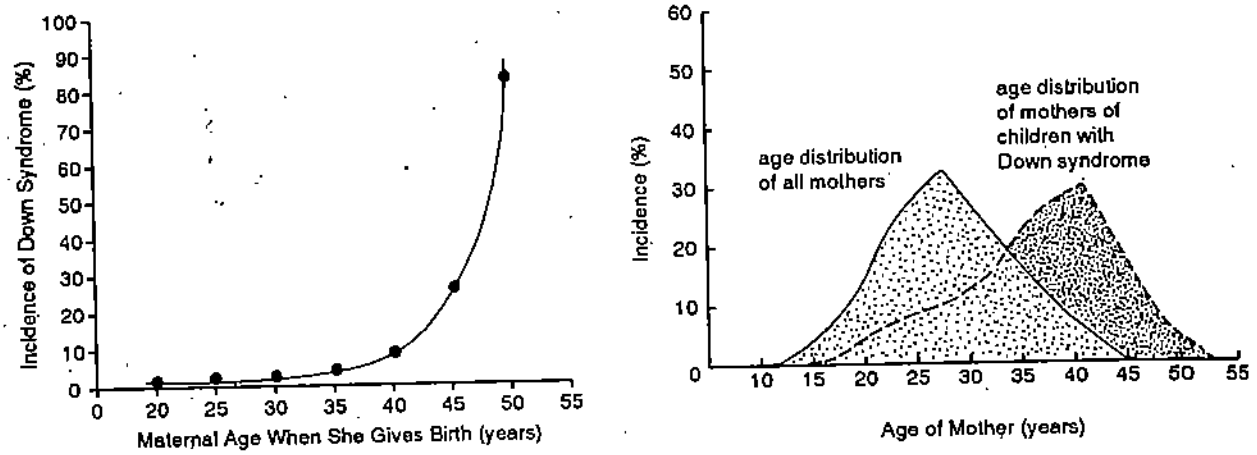
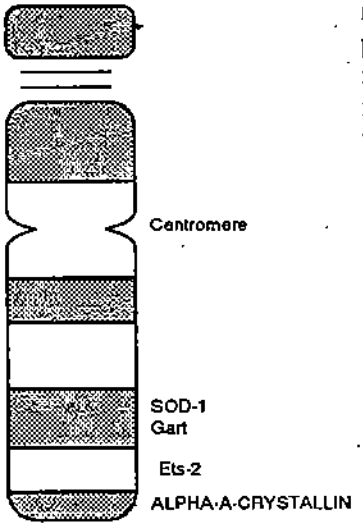


Fig.10.10: Down syndrome and age of the mother. The incidence of DS is about 1 in every 600-700 live births. (a) Among women below age 17, and over the age of 35 years, however, the incidence of delivering a child with DS increases. (b) The correlation between maternal and DS risk is striking when the age distribution for all mothers is compared to that of mothers of DS children.

incidence of DS for both older and younger mothers relates the frequency of nondisjunction to estrogen levels. Estrogen helps to control the rate of meiosis in primary oocytes. As estrogen levels decrease with age, a slowing of the meiotic process may make nondisjunction more likely. Estrogen levels also do not stabilise following puberty until about the age of twenty. The DS children produced by very young mothers thus may reflect fluctuations in hormone levels. If age or age-related factors increase the likelihood of nondisjunction, then trisomies involving chromosomes other than chromosome 21 should also be more common in older mothers, and even in younger mothers. In fact, the incidence of trisomies 13 and 18 is higher in older women. Other trisomies may also be common in these women, but because they are lethal, they are difficult to detect.

In the 1970s improved techniques for chromosomal banding analysis showed that in a considerable number of cases of DS, the extra chromosome is derived from the father also. The probability of a father having an affected child increases with age, as it does for mother. The rate of increase is relatively low until the age of fifty and steep thereafter.

This finding, thereby, implies that the chromosomes of both parents should always be investigated in trying to pinpoint the source of the genetic problem. Older mothers who have produced DS children have often borne an unnecessary burden of guilt, assuming they were the "cause" of their child's defect and misery.



Genes of Chromosome 21	Associated conditions
SOD-1	Premature aging: mental retardation
Gart	Mental retardation
Est-2	Acute myelogenous leukemia
Alpha-A-Crystallin	Cataracts

Fig. 10.11: A map of chromosome 21 and some of the genes that may cause phenotypic abnormalities associated with Down syndrome.

Chromosome 21— This chromosome has been referred to several times in this above discussion. Now, let us discuss some details of this chromosome. It is one of the smallest in the human genome and contains about 45 million nucleotide base pairs or 1.5% of the total nuclear DNA. Because of its small size and its association with several important genetic conditions, it is now one of the best known chromosomes. About 1,500 genes are thought to be located on chromosome 21. Only about 20 of these genes have been mapped so far. These genes give sufficient clues about the phenotypic aspects of DS and related conditions (see Fig. 10.11).

10.4.2 Trisomy 18 — Edward Syndrome

In 1960, John Edward and his colleagues reported on an infant trisomic for a chromosome in the E group, now known to be chromosome 18 (Fig. 10.12). It is designated as 47,18+. Incidence is about 1 in 4,000 live births. Eighty per cent of trisomy 18 cases are due to meiotic nondisjunction; 10 per cent are due to translocations; and 10 per cent are the result of trisomy mosaics. This syndrome occurs three times more frequently in females than in males. Of the live born, survival beyond one year is rare.

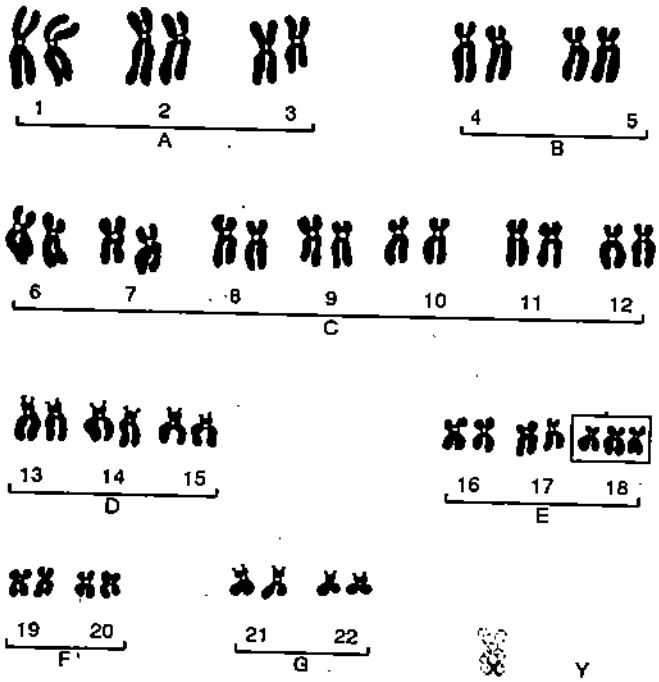


Fig. 10.12: Karyotype of a child with Edward syndrome.

Characteristics of trisomy 18 children are highly variable and pronounced. They demonstrate severe mental retardation, odd-shaped head, webbed-neck, short thumb, low-set ears, small mouth, receding chin, the index finger overlapping the middle and ring fingers, rocker bottom feet and clefts in the lip and palate. Heart defects are found in nearly all trisomy 18 births as well. Survival time is less than four months on the average. Death is usually caused by pneumonia or heart failure.

As in the Down syndrome, there is a pronounced maternal age effect. On plotting frequency against maternal age, a bimodal curve is obtained (Fig. 10.13). A secondary peak in the early 20s reflects the normal maternal age group of maximum births, whereas the pronounced peak from 35 to 45 years is clearly related to the greater age of the mother.

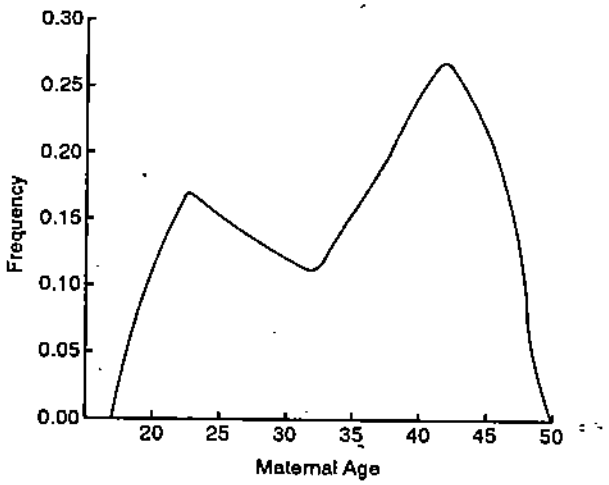


Fig. 10.13: The effect of maternal age in Edward's syndrome. Note the peak between maternal ages 35 and 45 indicating the age-group associated with maximum production of affected children.

10.4.3 Trisomy 13 — Patau Syndrome

In 1960s Klaus Patau and his associates observed an infant with gross developmental malformations with a karyotype of 47 chromosomes (Fig. 10.14). The additional chromosome was medium-sized, one of the acrocentric D group. The G and Q banding techniques have identified the extra D chromosome as number 13. The trisomy 13

condition has since been described in many newborns, and is called Patau syndrome, being designated as 47,13+. Seventy per cent of the trisomy 13 cases are caused by meiotic nondisjunction; 10 per cent of the cases are trisomy mosaics; and 20 per cent are partially trisomic due to translocation. As the chromosome 13 contains considerable active genetic material, most chromosome 13 trisomies abort. Of those that are live born, 90 per cent die during their first year of life. Those few that survive beyond the age of two years suffer from severe mental retardation.

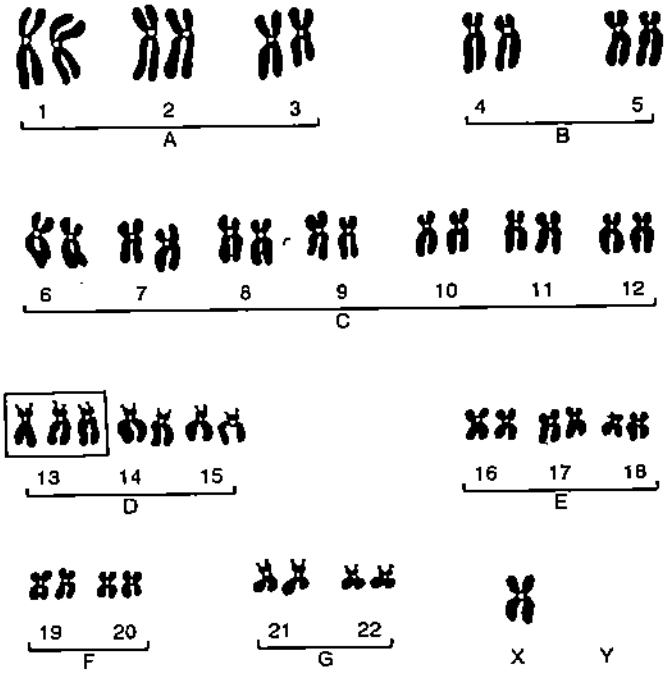


Fig. 10.14 : Karyotype of a child with Patau syndrome.

The characteristics of trisomy 13 include seizures, moderate microcephaly (small head), wide-spaced eyes, beaked nose, clefts of the lip and palate, low-set ears, scalp defects, and characteristic "trigger thumb" (the thumb and index fingers overlap the third finger). Trisomy 13 babies display a simian crease (one continuous crease across the palm of the hand), as in Down syndrome, as well as polydactyly (extra fingers and toes), congenital heart disease, kidney disorders and malformed genitalia.

The average maternal and paternal ages of parents of affected infants are higher than the ages of parents of normal children, but paternal ages are not as high as the average maternal age as in cases of Down syndrome. Both male and female parents average about 32 years of age when the affected child is born. Because the condition is so rare, it is not known if the origin of the extra chromosome is more often maternal, paternal or arises equally from either parent.

SAQ 3

Match the type of aneuploidy to its correct genetic designation. Use 2n to represent a diploid nucleus.

i) Trisomic	a) $2n - 2$
ii) Tetrasomic	b) $2n + 1$
iii) Nullisomic	c) $2n + 2$

SAQ 4

For a species with a diploid number of 10, indicate how many chromosomes will be present in the somatic nuclei of individuals that are haploid, triploid, tetraploid, trisomic and monosomic.

.....

.....

SAQ 5

a) What is nondisjunction?

.....

.....

.....

b) What are the other names for trisomics 13, 18 and 21?

.....

.....

c) If nondisjunction of chromosome 21 is known to have occurred in the division of secondary oocyte in a particular woman, what is the chance that a mature egg arising from the cell division will receive two, number 21 chromosomes?

.....

SAQ 6

Why are tetrasomics and nullisomics found in nature less frequently than trisomics?

.....

.....

10.5 SEX CHROMOSOME ANEUPLOIDIES

Aberrations in the number of the X and Y chromosomes are more readily found among humans than are the aberrations in the number of autosomes. The reason why they are more readily found is that a mechanism – called dosage compensation, exists in mammals to compensate the X chromosomes in females which are in excess of the number in males. You have already studied about it in Unit 4, and you may recall that dosage compensation does not occur for autosomes. As we know, normal human males are XY and normal females are XX. The nuclei of normal females thus have a Barr body. The abnormal number of Barr bodies is a key to the identification of variations in sex chromosome number. In this section you would study four kinds of aneuploidy in sex chromosomes and their effects discussed further in the following subsections. These are the XXX, XYY, XXY and XO conditions.

10.5.1 XXX – Trisomy X Syndrome

The presence of three X chromosomes (47,XXX) results in female differentiation (Fig. 10.15). This syndrome, also known as triple X syndrome, is estimated to occur in

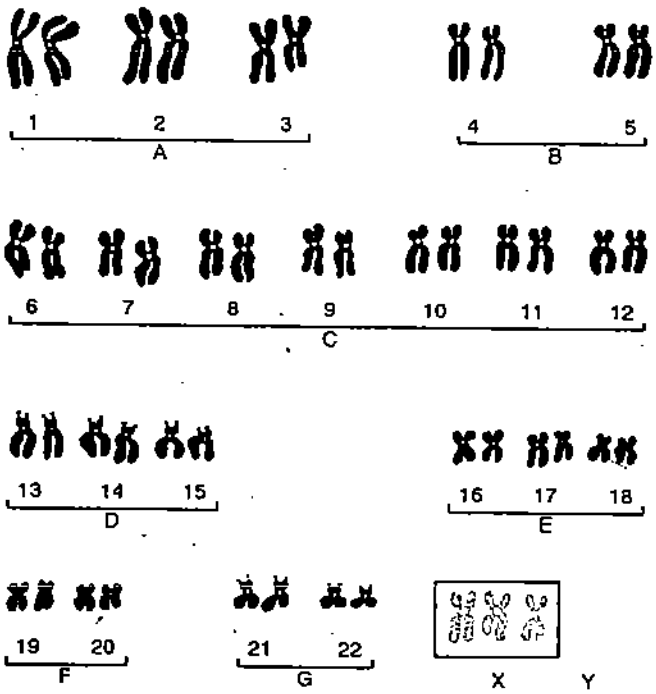


Fig. 10.15 : Karyotype of an individual with trisomy X syndrome bearing three X chromosomes.

1 in 1,000 live births. The buccal smears of such persons show two X-chromatin bodies in the epithelial cells as expected when three X chromosomes are present. This syndrome is highly variable in expression. Frequently 47,XXX women are perfectly normal. In other cases, underdeveloped secondary sex characteristics, sterility, and mental retardation may occur. In rare instances 48,XXXX and 49,XXXXX karyotypes have also been reported. The syndromes associated with these karyotypes are similar to but more pronounced than the 47,XXX. The presence of additional X chromosomes appears to disrupt the delicate balance of genetic information essential to normal female development.

10.5.2 XYY – Double Y Syndrome

Another sex-chromosomal abnormality associated with few physical manifestations is the XYY condition – designated as 47,XYY (Fig. 10.16). Affected males have 47 chromosomes; they have the full chromosome complement of normal males plus an extra Y chromosome. The XYY males tend to be tall, averaging over 6 ft. Amongst newborn males, this syndrome occurs with high frequency, that is, one in every 200 – 1000 live births. At one time it was thought that 47,XYY males developed severe personality disorders and were at a high risk of committing crimes of violence, a belief based on a higher incidence of 47,XYY among violent criminals. Recent studies indicate that most 47,XYY males have moderately impaired mental function and, although their rate of criminality is higher than that of normal males, the crimes being mainly due to their impulsive nature. The majority of 47,XYY males are phenotypically and psychologically normal and have no criminal convictions.

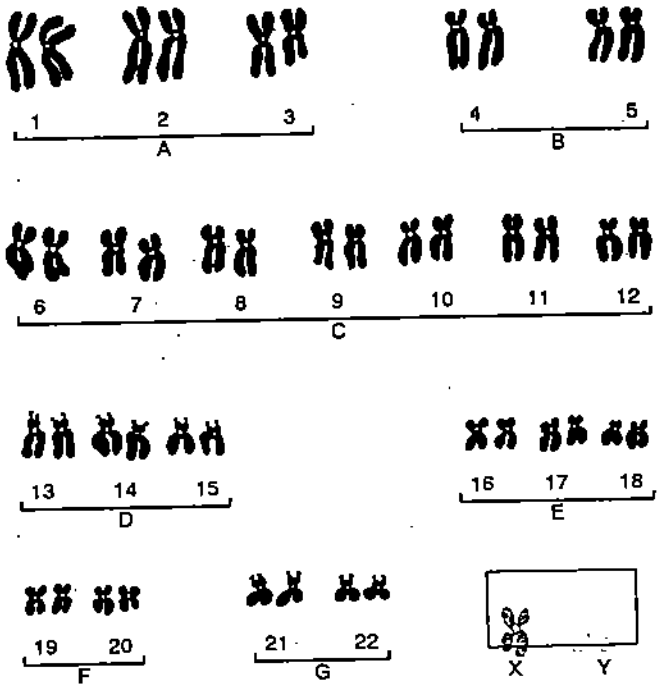


Fig.10.16 : Chromosome constitution of an individual with double Y syndrome, showing two Y chromosomes, and one X chromosome.

10.5.3 XXY — Klinefelter Syndrome

An extra X chromosome in addition to the usual male (XY) chromosome complement (47,XXY) has been associated with a male phenotype – the Klinefelter syndrome (Fig. 10.17). It was described by H.F. Klinefelter and his colleagues in 1942. It is estimated to occur in 1 per 500 live male births. The risk of a Klinefelter birth is associated with increasing maternal age. It is believed that the XXY constitution originates either by fertilisation of an exceptional XX egg by a Y sperm or of an X egg by an exceptional XY sperm.

The Klinefelter syndrome patients are always male by virtue of having Y chromosome, but they have variable male genital development. The characteristics of Klinefelter syndrome male are underdeveloped testes (**hypogonadism**) causing sterility, have small prostate glands and a tendency towards femaleness, particularly the secondary sex

characteristics: They show some female features such as breast development after puberty (gynecomastia), pubic hair and distribution of fat . Unlike adult males, they lack facial hair. Most of them demonstrate mild to severe retardation. Many adjust poorly to breast development and exhibit emotional problems.

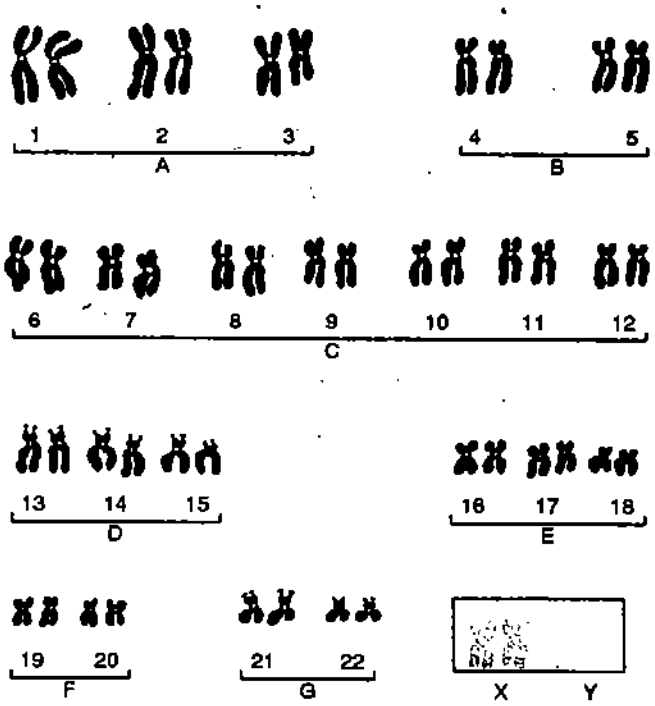


Fig.10.17: Karyotype of an individual with Klinefelter syndrome bearing two X chromosomes and a Y chromosome.

The Klinefelter syndrome patient is one of the most amenable to treatment, using hormones and surgery. This treatment involves surgical removal of Klinefelter males' breasts (mastectomy) and subsequent treatment with a male hormone – testosterone. In addition to the XXY Klinefelter syndrome, a number of cases have been reported with modified karyotypes: 48,XXXY; 48,XXYY; 49,XXXXY; and 49,XXXYY. They are all similar phenotypically to 47,XXY. Generally, more severe manifestations are seen in individuals with a greater number of X chromosomes. The number of X chromosomes can be ascertained by counting the number of Barr bodies.

10.5.4 XO – Turner Syndrome

A condition – XO, exemplifies monosomy of the X chromosome in humans. This condition is termed as Turner syndrome and was first described by Turner in 1938. These people are female, because of the absence of the Y chromosome. They have a genomic complement of 45,X indicating that they have a total of 45 chromosomes, in contrast to the normal 46 (Fig. 10.18). The sex chromosome complement consists of one X chromosome.

Turner syndrome is the least frequent of the live-born sex-chromosome aneuploidies. As 98 per cent of the XO conceptions are spontaneously aborted, yet it has a frequency of about 1 in 5,000 live-born infants. About 20 per cent of all spontaneous abortions are 45,XO making this one of the most common chromosome aneuploidies known. Because of what appears to be a lethal effect of this monosomy, it is believed that most if not all Turner syndrome cases females are mosaics, i.e., they carry a normal cell line within their important body tissues. In some cases of Turner syndrome, the somatic cells are chromosomally normal, but their gonads have 45,XO constitution. Vice versa, there are individuals who have 45,XO in the somatic cells, but they have normal gonadal cells. These females with Turner syndrome may have the potential to become pregnant. The possible ways of origin of X monosomics are from exceptional eggs or sperm with no sex chromosome, or from the loss of a sex chromosome in mitosis during early cleavage stages, after an XX or XY zygote has been formed.

The characteristic features of an adult with Turner syndrome are relatively short stature, sterility due to failure of the ovaries to develop, webbed neck, low set ears, wide-spaced

and underdeveloped breasts, sparse pubic hair, abnormal hormone levels, narrowing of the aorta, and failure to menstruate (**amenorrhea**). The majority of these females have a normal IQ, but demonstrate space-form perception difficulities. Epithelial cells of 45,X patients are X chromatin negative, i.e., lack Barr body as expected since only one X chromosome is present.

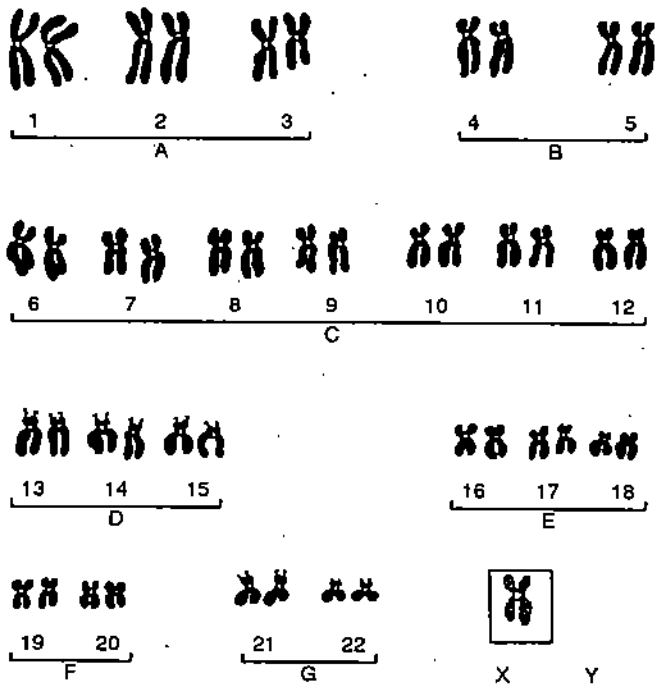


Fig.10.18: Chromosome constitution of an individual with Turner syndrome showing only one X chromosome without a Y chromosome.

SAQ 7

Based on Lyon's hypothesis (that you studied in Unit 4), can you tell the number of Barr bodies to be expected in the cells of persons with (a) Turner syndrome; (b) trisomic X; and (c) Down syndrome.

.....

.....

SAQ 8

A spontaneously aborted human foetus was found to have the karyotype 92,XXYY. What might have happened to the chromosomes in the zygote to result in such a karyotype?

.....

.....

.....

.....

SAQ 9

A spontaneously aborted human foetus is found to have 45 chromosomes. What is the most-probable karyotype? Had the foetus survived, what genetic disorder would it have?

.....

.....

.....

.....

.....

Colour-blindness in human beings is an X-linked trait. A man who is colour-blind has a 45,X (Turner syndrome) daughter who also is colour-blind. Did the nondisjunction that led to the 45,X child occur in father or mother? How can you tell?

.....

.....

.....

.....

.....

10.6 SUMMARY

In this unit you have learnt that:

- Abnormalities in chromosome number can either involve entire sets of chromosomes or individual chromosomes. When complete sets of chromosomes are added to the diploid genome, a polyploid is created. The sets may have identical or diverse genetic origin creating either autopolyploidy or allopolyploidy respectively. Relatively rare in animals, polyploidy has been an important force in the evolution of some plants groups. By using colchicine and some other agents that interfere with cell division, polyploids can be induced. In wheat, induced polyploidy has paralleled what has occurred in nature; induced polyploidy also has practical value in such plants as apples, grapes and tobacco.

- Aneuploidy is the gain or loss of one or more chromosomes from the diploid content. When one chromosome is lost or gained, the conditions of monosomy and trisomy, respectively are created.

- The major mechanism of the development of cells demonstrating either autosomal or sex-chromosome aneuploidy is nondisjunction. Nondisjunction is the failure of chromosomes to properly separate into their respective nuclei during nuclear division. Both first and second meiotic division nondisjunction will produce gametes having two of a given chromosome while the companion gametes lack that chromosome. These aneuploid gametes, when fertilised by a normal gamete, result in either a monosomic or trisomic zygote. All cells from that zygote will contain a monosomic or trisomic condition that is generally lethal, if it is autosomal. There is an increasing incidence of nondisjunction found in foetuses of older women, especially for trisomy 21 – the Down syndrome.

- Of the surviving autosomal aneuploid foetuses, trisomy 13, 18 and 21 occur in increasing number respectively. The only long-term survival occurs in trisomy-21 patients.

- The XXX, XYY, XXY (Klinefelter syndrome) and XO (Turner syndrome) conditions of sex-chromosome are examples of trisomy known to occur in man.

10.7 TERMINAL QUESTIONS

A human female with Turner syndrome also expresses the sex-linked trait haemophilia, as her father did. Which parent underwent nondisjunction during ...eiosis, giving rise to gamete responsible for the syndrome?

...w can the origin of trisomy in humans be explained?

19 Down syndrome occurs in about 1 in 700 and the Turner syndrome occurs in ... in 5000 in the general population, and each is separately and randomly

distributed in the population, what is the chance that a baby will be born with both these abnormalities?

4) a) What evidence concerning the influence of the Y chromosome on sex determination in humans can be obtained by comparing the characteristics of X, XXY, and XXX individuals?

   b) Compare the influence of the human Y chromosome on sex determination with that of *Drosophila* and *Melandrium* Y chromosomes (you may refer to Unit 3).

5) Explain the use of aneuploidy as a tool to determine the chromosome on which particular genes are located.

6) What potential monoploids have for genetic studies?

7) Give two methods by which polyploidy might occur in nature.

8) a) How does autopolyploidy and allopolyploidy originate?

   b) Discuss the significance that each might have in evolution.

9) Give one term each for the chromosome arrangement as symbolised below:

   a) n

   b) 2n

   c) 2n+1

   d) 2n–1

   e) 2n+2

   f) 3n

   g) 4n

10) Fill in the blanks with appropriate words:

   a) The abnormality in which there is one more or less than the normal number of chromosomes is an ..... .

   b) A person with an extra chromosome with three of one kind is termed ..... .

   c) A person with one less than the normal number of chromosomes, having only one member of a pair is termed ..... .

   d) The failure of two homologous chromosomes to separate normally during cell division is called ..... .

   e) A person with ..... is retarded in physical and mental development, has abnormalities of face, tongue, and eyelids, and is ..... for chromosome ..... .

   f) The XXY individual has the disorder known as ...., the XO individual has ....

10.8 ANSWERS

Self-assessment Questions

1) **The polyploid tissues grow through cell divisions that can occur regularly among polyploid cells. Sexual reproduction of animals requires gamete formation, fertilisation, and sex determination. Many irregularities associated with polyploidy result in inviability and sterility in animals.**

2) **Colchicine treatment interferes with spindle formation during cell division. Double number of chromosomes are thus included in the daughter cells.**

3) **i) b**

   ii) c

   iii) a

4) Haploid, 5; triploid, 15; tetraploid, 20; trisomic, 11; monosomic, 9.

5) a) Nondisjunction is the failure of members of a homologous pair of chromosomes to separate from each other during nuclear division.

   b) Patau syndrome, Edward syndrome and Down syndrome.

   c) 1/2

6) Tetrasomics and nullisomics are less viable and usually die before they are detected.

7) a) 0

   b) 2

   c) 1 in ♀ and nil in ♂

8) The chromosomes underwent replication with no cell division resulting in a tetraploid.

9) The probable karyotype is 45,X. It is the only 45-chromosome karyotype found at appreciable frequencies in spontaneous abortions. Had the foetus survived, it would have been a 45, X female with Turner syndrome.

10) Because the X chromosome in the 45,X daughter contains the colour-blindness allele, the 45,X daughter must have received the X chromosome from her father through a normal X-bearing sperm. The nondisjunction must, therefore, have occurred in the mother resulting in an egg cell lacking any X chromosome.

Terminal Questions

1) Since she received her single X chromosome from her father, her mother contributed the gamete devoid of sex chromosomes.

2) Nondisjunction of chromosomes during the production of gametes (eggs and sperms) is considered to be the major cause for most trisomies in humans.

3) 1/3,500,000

4) a) XO is basically female; XXY is male; and XXX is female. The human Y carries the male-determining capacity.

   b) The *Drosophila* Y chromosome has no influence on sex determination. The *Melandrium* Y chromosome carries male-determining genes.

5) Many of the plants and animals with extra chromosome can be recognised phenotypically. Genes that give trisomic ratios (e.g. genes for capsule shape in *Datura*) are located on the particular chromosome that makes the trisome. Monosomics are useful because recessive genes on a chromosome that has no homolog (e.g. a monosome) express themselves and can thus be associated with a particular chromosome.

6) Since monoploids have only one set of genes, the question of gene segregation does not arise. They could be used in experimental work where it is desirable to relate genes with phenotypes or traits.

7) Polyploidy may occur through a doubling of the chromosomes in the somatic cells or the failure of the reduction division of meiosis.

8) See Section 10.2.

9) a) haploid

   b) diploid

   c) trisomic

   d) monosomic

   e) tetrasomic

   f) triploid

   g) tetraploid

10) a) aneuploidy
b) trisomic
c) monosomic
d) nondisjunction
e) Down syndrome, trisomic, 21
f) Klinefelter syndrome, Turner syndrome

UNIT 11 THE NATURE AND STRUCTURE OF GENETIC MATERIAL

11.1 INTRODUCTION

In the preceding units, you have learnt that the systematic study of the science of genetics began through Mendel's investigations. You also sudied the patterns of transmission of traits from one generation to another, the origin of variations through recombination, and how the genes governing certain traits have been located on given chromosomes by establishing linkage and gene mapping.

You may recall that the involvement of discrete genes in determining phenotypic characters was known from Mendel's work. It also demonstrated that genes were transmitted unchanged between succeeding generations. But there was paucity of information regarding the nature and structure of the genetic material composing the gene. For decades after Mendel's work, many geneticists concentrated on establishing the chemical identity of the genetic material as well as describing its molecular structure.

We begin this unit by discussing the qualities that the genetic material possesses and also the evidences that DNA and RNA are the genetic materials. In addition, we shall also examine the molecular structure of the nucleic acids.

Study Guide

In this unit, some additional information, such as, details of various techniques used, data obtained through various research works and small writeups about scientists are provided. All these are given to enrich your knowledge, though test questions may not be based on these additional information.

In this course of genetics, of which the present unit forms a part, we have deliberately omitted the topics DNA replication and protein synthesis as you have already studied them earlier in Cell Biology (LSE-01) course. Now that you are going to study the detailed structure of the nucleic acids, it would be useful that you revise these topics in continuation with this unit. For revision on replication see Unit 13, p.65 (LSE-01) and for protein synthesis Unit 14, p.82 of the same course.

Objectives

After studying this unit you should be able to:

- describe the various investigations leading to the identification of the genetic material (Section 11.2);
- draw diagrams of a nucleotide and a polynucleotide (Section 11.3);
- distinguish between purines and pyrimidines (Section 11.3);
- state the complementary bases found in nucleic acids (Section 11.3);
- describe the Watson-Crick model of DNA (Section 11.4);
- state the chemical and structural differences between DNA and RNA (Section 11.4);
- describe the conformational flexibility of DNA molecules (Section 11.4).

11.2 THE SEARCH FOR GENETIC MATERIAL

You already know that the genetic or hereditary material for all species is nucleic acid. In most organisms, it occurs in the form of deoxyribonucleic acid – DNA; and in a few viruses it occurs as ribonucleic acid – RNA. Long before DNA and RNA were proved to carry genetic information, geneticists recognised that special molecules must fulfil that function. They postulated that the material responsible for inheritable traits should have the following three principal characteristics.

i) It must contain **all the information** for an organism's cell structure, function, development and reproduction in a stable form.

ii) It must **replicate accurately** so that the progeny cells have the same genetic information as the parental cells.

iii) It must be capable of undergoing **variation**, and exist in an almost infinite variety of forms, as without variation (through mutation and recombination), organisms would be incapable of change, and adaptation and evolution would not occur.

In mid to late 1800's and early 1900's, scientists considered **proteins** rather than the nucleic acids as the prime candidates for hereditary material. But critical experiments that were performed subsequently, the details of which are discussed in the unit later on, went on to show that DNA is the genetic material. But let us first look into some salient features of proteins that made these likely candidates for hereditary material.

11.2.1 Proteins

Since proteins were known to be highly complex polymeric molecules, they were considered strong candidates for genetic substance. About 66% of the weight of the human body is water; 47% of the remainder (that is, 16% of the total weight) is protein. Protein fulfils a variety of crucial functions for the living organism. One category of proteins called **enzymes** are of particular importance, as they catalyse thousands of crucial biochemical reactions in the body.

Some proteins function to provide **physical support**, both for individual cells and for the entire organisms. Keratin is a structural protein that is the primary component of skin and hair. Tubulin forms part of the "**skeletal**" or supportive system of an individual cell. Muscle tissue consists of protein specialised for movement. **Antibodies** are proteins that protect the body from bacteria and other foreign substances. **Haemoglobin** is a protein specialised for the transport of gases in the blood. Some **hormones** are proteins that provide communication among some tissues of the body. We have seen the various roles that proteins play in the body. A brief introduction to the chemical structure of proteins can help to illustrate how proteins fulfil so many different functions, and why they seemed to be the most likely chemical to bear hereditary information. We shall now briefly consider some aspects of protein chemistry.

Many of the salient facts about protein chemistry that we still accept today were known in the 1930's. Proteins are composed of linear polymers of **amino acids** called polypeptides. Typically, amino acids have a basic amino ($-NH_2$) group and an acidic carboxyl ($-COOH$) group attached to a common carbon atom from which a variety of

side chains extend, (see Fig. 11.1). Twenty amino acids are commonly found in polypetides. Amino acids are joined together in a linear polymer by bonds called

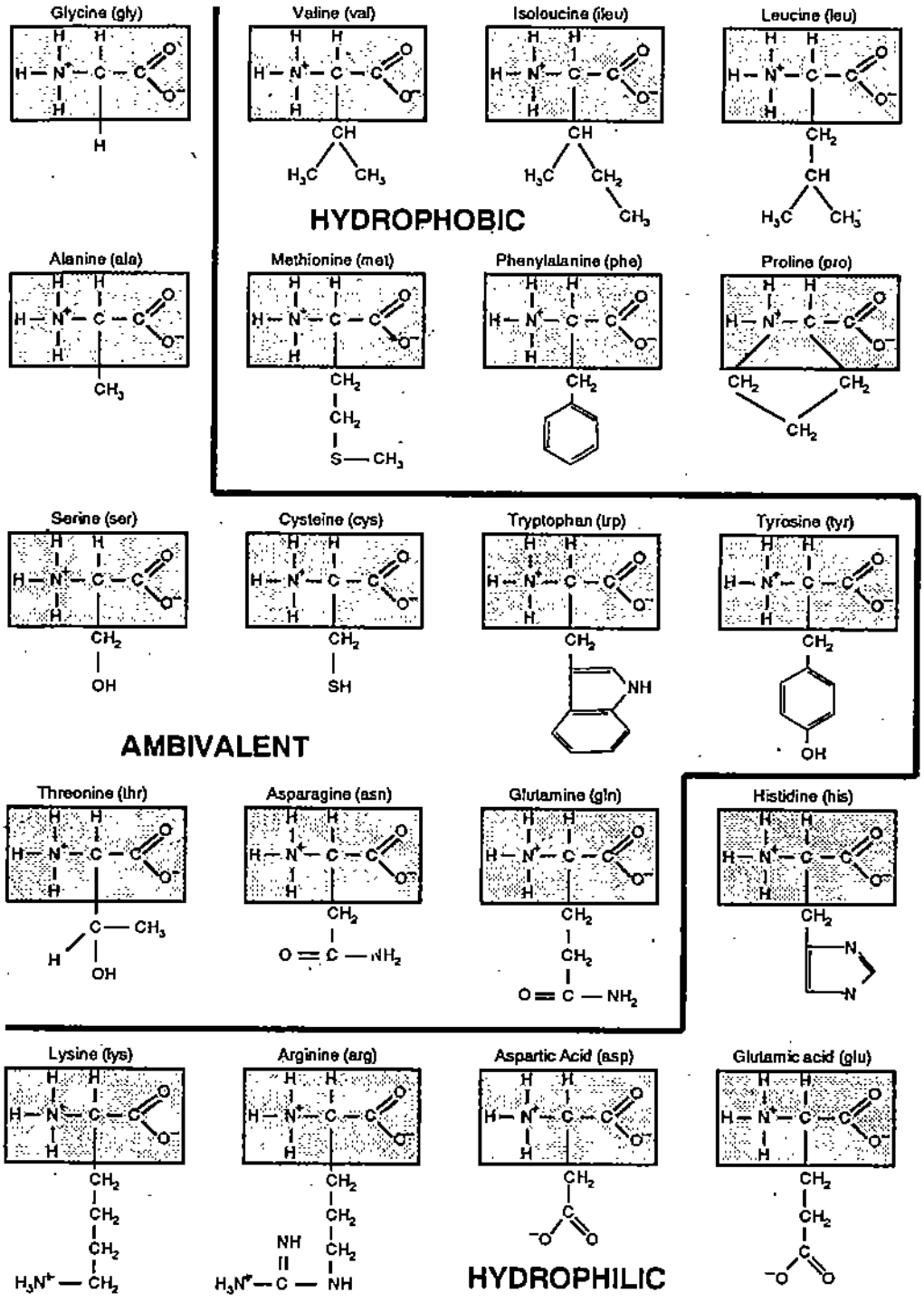


Fig.11.1 : The structures of twenty amino acids consisiuting proteins.

peptide bonds, between the amino group of one amino acid and the carboxyl group of an adjacent amino acid (Fig. 11.2). As a result of these linkages, most polypeptides have a free, unbound amino group at one end (the amino terminal or N-terminal amino acid) and a free carboxyl group at the other end (the carboxyl terminal or the C-terminal amino acid) of the linear polymer.

If two amino acids are bound by one polypeptide bond, the molecule thus formed is known as a **dipeptide**. And in the case of more than two amino acids, or if a chain of many amino acids is involved, it is known as a **polypeptide**. Proteins may contain only one polypeptide chain, several polypeptide chains, or polypeptide chains combined with

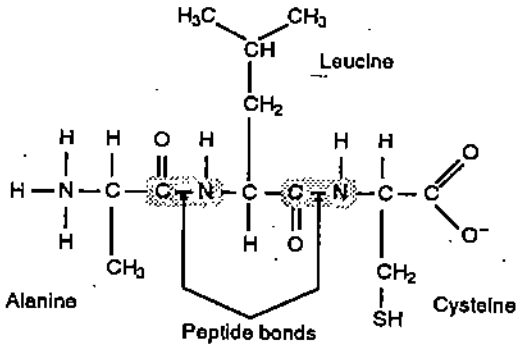


Fig. 11.2 : A polypeptide formed from three amino acids. Note the two peptide bonds.

other molecules. Haemoglobin, for example, is a protein consisting of four polypeptide chains, each bound to an iron-containing molecule (see Fig.11.3).

Different proteins have different compositions and sequences of amino acids. The amino acid *composition* of a protein is the relative proportion of the different kinds of amino acids present in a protein. The amino acid *sequence* of a protein is the order in which the amino acids follow one after the other in the linear polypeptide chain. By convention, the amino acid sequence of a polypeptide is listed starting with the N-terminal amino acid or residue, and continues through the C-terminal amino acid. The smallest proteins or polypeptides have fewer than one hundred amino acids, for example, the hormone insulin has 51 amino acids, and some large proteins like myosin (involved in muscle contraction) have over 1,800 amino acids long polypeptide chain.

Fig. 11.3. The protien molecule haemoglobin showing four polypetide chains.

The number of different possible sequences of amino acids in a protein is 20^n, where n is the number of amino acids in the protein. For a polypeptide only four amino acids long, some 160,000 different combinations are possible. And ten amino acids can produce 10,240,000,000,000 unique combinations. For a protein one hundred amino acids long, the number of combinations are unimaginably large. The biochemists of the nineteenth century knew that a large molecule, as that of protein, manifests a great diversity of structure to fulfil the many functions as well as encode the immense amount of hereditary information needed to direct various biochemical reactions.

In addition to being of many different sizes, proteins also have widely varying physical properties because of varying amino acid content and the differing chemical and physical properties of the side chains of the amino acids. Most of the side chains are neutral, but three (viz. lysine, histidine and arginine) are basic, and two (viz. aspartic acid and glutamic acid) are acidic. Furthermore, the various side chains have different affections for water. The basic and acidic amino acids are water loving, i.e., **hydrophilic** or water soluble (see Fig. 11.1). Other amino acids, such as leucine, isoleucine, phenylalanine, proline, valine and methionine are water-hating, i.e., **hydrophobic** or water- insoluble. But most amino acids, e.g., glycine, tyrosine and serine are **ambivalent**, they can take water or leave it.

The **primary structure** of protein refers to the sequence of amino acids in the polypeptide chain. Commonly, protein in a polypeptide chain coils into a spiral or helix which is the **secondary structure** of proteins. The helix in turn is twisted or folded precisely into a complex three-dimensional conformation, crucial for the protein function, and is known as **tertiary structure**. Finally many proteins contain two or more polypeptides, forming the **quaternary structure**.

In the 1930's the biochemists knew that enzyme activities, the ability to catalyse chemical reactions in cells, were associated with proteins, implying that enzymes were proteins. What emerged in 1930's was the realisation that proteins are large, highly complex molecules, with a variety of physical properties and functional activities. This is still valid though greatly augmented by present knowledge.

11.2.2 Nuclein

In 1869, only a few years after the publication of Mendel's work, a Swiss biochemist, Johann Friedrich Miescher studied the nuclei of pus cells from waste surgical bandages. He found that these nuclei contained a novel phosphorus-bearing compound, which was

a non-protein material that he named nuclein. He named it so based on the site of its origin. Nuclein is mostly **chromatin** – a complex of deoxyribonucleic acid (DNA) and chromosomal proteins.

The work by Miescher and other scientists, subsequently, demonstrated that the nuclei of the cells of all species contain nuclein and that nuclein was an important constituent of chromosomes. Quantitative analysis of chromosomes showed that they consisted of about 60% protein and 40% nuclein (DNA). This led the scientists of that time to infer that protein was the genetic material. DNA was not considered to be identical with the genetic material as it was found to lack chemical diversity needed for the genetic substance. In contrast, proteins are exceedingly diverse collection of molecules, so it was widely believed that protein was the genetic material and that DNA was merely the structural framework of chromosomes.

11.2.3 Amounts of DNA and Protein

Description of the meiotic behaviour of chromosomes supported their role in heredity. The diploid number was constant in the somatic cells of all members of a species, but was halved in the germinal cells and the diploid condition was resumed only after fertilisation.

Careful measurements of the amounts of DNA and protein in somatic and sperm cells indicated the likely hereditary role of DNA, but not of proteins. Table 11.1 gives data on the DNA content of somatic and gametic cells in some organisms. Within a species, the amount of protein in a sperm cell varied considerably. Furthermore, the amounts of protein in sperm and somatic cells bore no simple relationship to each other. On the other hand, the amount of DNA in sperm was relatively constant. Additionally, the amount of DNA in sperm cells was exactly one-half that in the somatic cells.

Table 11.1 : DNA content in the somatic and gametic cells of some organisms. The unit for measuring DNA content is picograms 10^{-12} gram, or one trillionth of a gram). After Maxson & Daugherty, *Human Genetics*, 1992, p.70

Organism	Somatic Cell (2n)	Gametic Cell (n)
Chicken	2.6	1.3
Cow	6.6	3.3
Diatom	0.08	0.04
Fern	326.4	163.2
Frog	15.0	7.5
Fruit fly	0.4	0.2
Fungus	3.0	1.5
Man	6.4	3.2
Onion	36.7	18.4
Pine	138.7	69.4
Rat	6.0	3.0
Salamander	336.0	168.0

11.2.4 DNA is the Genetic Material

The identity between DNA and the genetic material was shown directly in the experiments that are described below:

Transformation in Pneumococcus

Frederick Griffith laid the foundation for the identification of DNA as the genetic material, in 1928, with his experiments on transformation in bacterium *Diplococcus pneumoniae* or *pneumococcus*, now named *Streptococcus pneumoniae*. In his experiment, he used two naturally occurring strains of *pneumococcus*. These two strains

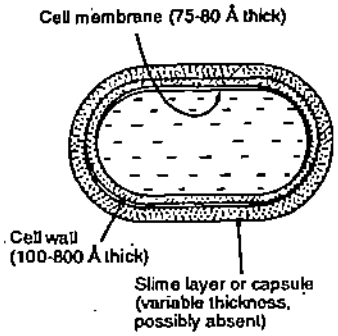


Fig. 11.4: A typical bacterial cell showing the surface layers including the capsule.

have markedly different properties. One strain possesses a smooth polysaccharide capsule (Fig.11.4), and thus produces smooth colonies. This strain was called strain S. It causes pneumonia in mice. Another strain, which lacks polysaccharide capsule produces rough colonies. This strain was called strain R and did not have a pathological effect on mice. Bacteria of the rough strain are engulfed by the white blood cells of the mice; but the virulent smooth-strain bacteria survive because they are protected by their polysaccharide coating. Griffith found that neither the heat-killed S-type nor the live R-type cells, by themselves, caused pneumonia in mice. However, if a mixture of the live R-type and the heat-killed S-type cells was injected into the mice, they developed an infection identical to that caused by the living S-type cells (Fig. 11.5). Thus, something in the heat-killed S-cells transformed the R-type bacteria into the S-type cells. Griffith felt that some principle transferred from the killed S-type cells converted the R strain to virulence by enabling it to synthesise a new polysaccharide coat. In Griffith's terms, the transforming principle from the dead S cells served as a "pabulum" for the R cells. This experiment of Griffith's has also been dealt with in an elaborate manner in Unit 12, Subsection 12.9.1, while explaining the concept of 'transformation' in bacteria. You are advised to read this particular subsection in conjunction with the description given above.

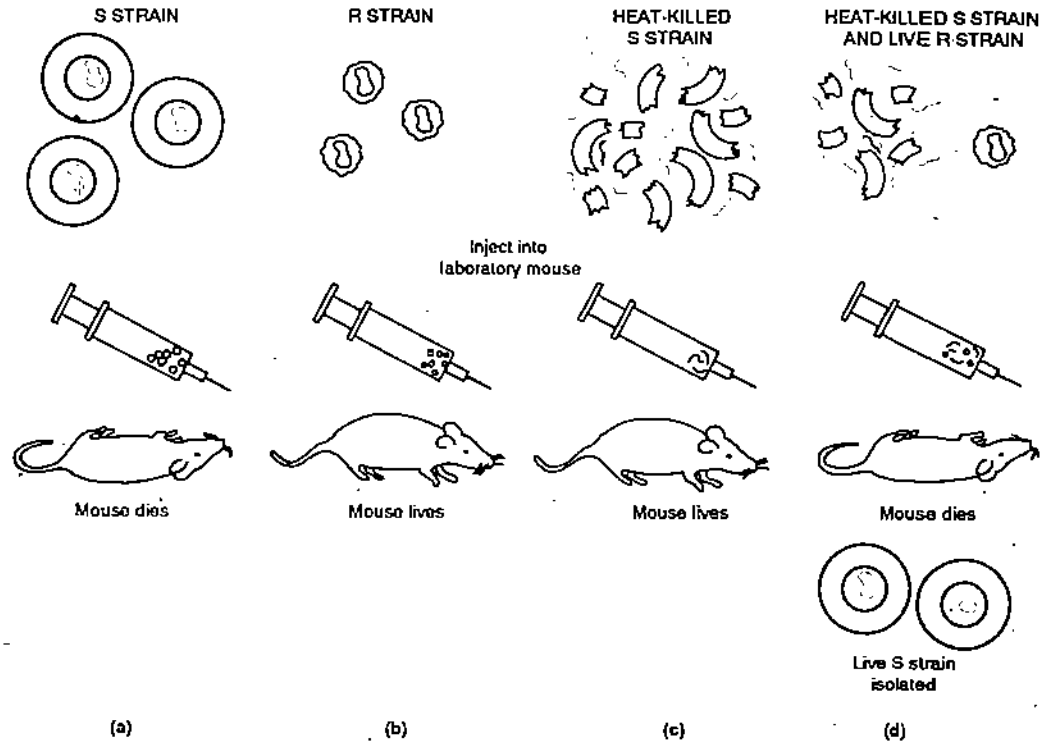


Fig. 11.5: Griffith's transformation experiment with smooth (S) and rough (R) strain of *Streptococcus pneumoniae*.

Fig. 11.6: Oswald T. Avery b.21 Oct. 1877, Nova Scotia; d, 2 Feb. 1955 Nashville.

Avery, born in Canada, studied medicine at Columbia University in New York City, obtaining his MD in 1904. He joined the Rockfeller Institute Hospital in 1933, remaining there until 1948, when he retired and moved to Nashville, Tennessee. Avery was a bachelor and a dedicated bacteriologist. He was actively engaged in research even towards the end of his career. He made several contributions to the study of pneumonia causing bacteria. His meticulous plan of work on the transforming principle though lasting several years, eventually culminated in 1944 with the clear identification that the active agent is DNA. René Dubos has written a biography of Avery: 'The Professor, the Institute and DNA', Rockfeller, University Press., New York, 1976.

Proof that the Transforming Principle is DNA

Over the next decade and a half, a research team headed by Oswald T. Avery (Fig. 11.6) at the Rockfeller Institute earnestly pursued to identify the transforming principle. Between 1930 and 1933, they achieved transformation outside the body of a living mouse. Avery and his coworkers did most of their work under *in vitro* conditions, rather than the occurrence of pneumonia in the mice as evidence of transformation.

In 1944, sixteen years after Griffith's publication, Avery, MacLeod and McCarty published a remarkable paper in which they reported that they had purified the transforming factor. Analysis of molecular composition and weight indicated this factor to be DNA. Their strategy was to prepare a cell-free filtrate from the heat-killed S-cells (see Fig. 11.7, steps 1 to 4), containing the transforming principle and then to digest the individual components of this filtrate with specific degradative enzymes (see steps 5, 6, and 7, in Fig. 11.7). For instance, any of the three enzymes – protease, trypsin or chymotrypsin–could be used to degrade all the proteins in the filterate (step 5). Similarly, a ribonuclease (step 6) can be employed to destroy the ribonucleic acid (RNA) – the second type of nucleic acid (step 7) present in living cells (though, unlike

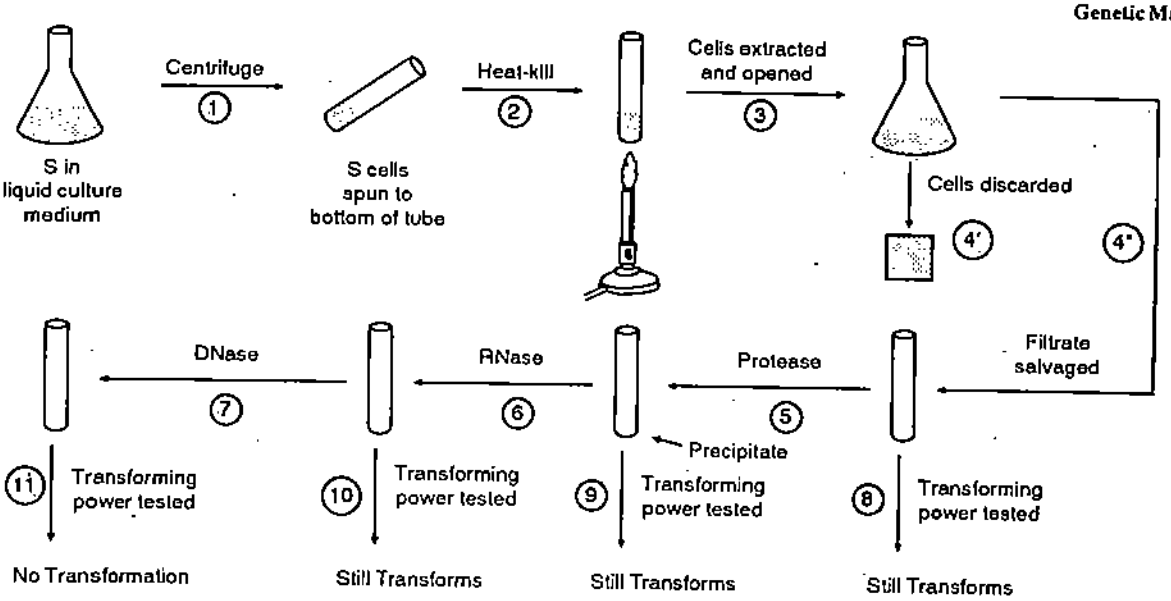


Fig. 11.7: An outline of Avery, McLeod and McCarty's experiment demonstrating that DNA is the transforming principle.

DNA, not a major constituent of chromosomes). The filtrate salvaged (step 4) was tested for its transforming ability (see step 8); that was found to be positive. Likewise after the respective enzymatic digestion the filtrate was also tested for retention of its transforming ability. Surprisingly (as genes were still thought to be made of protein), treatment with protease had no effect on the ability of the filtrate to transform the R cells into the S form (step 9). Similarly, ribonuclease treatment had no effect on the transforming activity (step 10). However, digestion of DNA with deoxyribonuclease (see step in Fig. 11.7) totally destroyed the transforming ability, so that the filtrate was no longer able to convert one cell type into the other.

Gradually, it became clear to Avery and his colleagues that the transforming principle must be DNA. This idea was supported not only by the results of the enzymatic digestion experiments, but also by the analysis of purified transforming principle. The analytical tools that Avery et. al. used were: *ultracentrifugation, electrophoresis, ultraviolet absorption spectrophotometry, and elementary chemical analysis*. For their details see Box 11.1.

On reading Avery et. al's paper even today, one has little doubt over the reliability of their results or the validity of their conclusions. Scientists, in 1944, were more sceptical. Doubts were raised about the efficiency of the enzymes used to digest the protein component of the S cell filtrate. It was suggested that as much as 2% residual protein in the 'pure' preparation of transforming principle would be undetected and might be responsible for the transformation of the R cells to S, if genes were made of proteins. These nagging doubts although unfounded prevented the results of Avery's group from being unanimously accepted by the scientific community. Therefore, a second experimental identification of genetic material was required.

Box 11.1: Techniques used in Avery et. al's experiment.

1. Electrophoresis – By this method, the movement of the charged molecules in an electric field is studied. The DNA molecules like proteins and many other biological compounds carry an electric charge. DNA carries negative charge. And when it is placed in an electric field (Fig. 11.8) the DNA molecules migrate to the positive pole. In an aqueous solution, the rate of migration of a molecule depends on two factors: its shape and its electrical charge, meaning that most DNA molecules will migrate at the same speed in an aqueous electrophoresis system. Avery's group subjected the purified transforming principle to electrophoresis and demonstrated that its migration rate is the same as that of a sample of pure DNA.

Presently, electrophoresis of DNA is carried in agarose gel, polyacrylamide or a mixture of the two. In a gel, the migration rate of macromolecules is influenced by a third factor, its size. This is because the gel consists of a complex network of pores through which the molecules have to move to reach the electrode. The smaller

the molecule, the faster it can migrate through the gel. Gel electrophoresis therefore separates the DNA molecule.

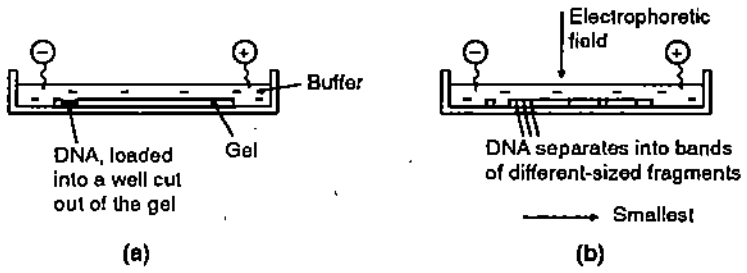


Fig. 11.8 (a, b): Outline diagram showing the setup of electrophoretic apparatus.

2. Ultracentrifugation – This technique was developed by Svedberg in 1925. The sample to be tested is subjected to intense centrifugal force, up to several hundred thousand revolutions per minute (rpm). In this process, the various cell components sediment at a rate depending on their size, shape, density and molecular weight. Avery showed that the transforming principle sedimented in the ultracentrifuge at a rate similar to pure DNA.

Two modifications of ultracentrifugation are now widely used in studying DNA. The first is called *velocity sedimentation analysis.* It involves measuring the rate at which a macromolecule or particle sediments through a dense solution of sucrose when it is subjected to a high centrifugal force. The rate of sedimentation is a measure of the size of the molecule or particle, although the shape and density also influence the rate. It is expressed as a sedimentation coefficient.

The second method is called *density gradient centrifugation.* A density gradient is produced by centrifuging, usually using caesium chloride, as a high centrifugal force will pull the caesium and chloride ions towards the bottom of the tube. Their downward migration is counter-balanced by diffusion, so a concentration gradient is setup. The CsCl density is greater towards the bottom of the tube. Depending on the buoyant density the macromolecules present in CsCl solution form different bands (see Fig.11.9).

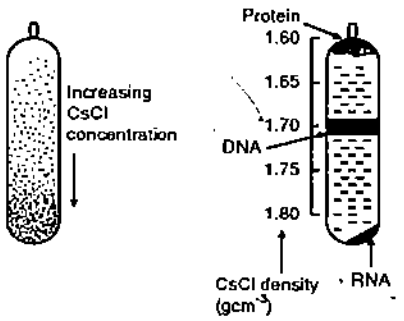


Fig. 11.9: Density gradient centrifugation using CsCl

Since DNA has a buoyant density of about 1.7 gm cm^{-3}, it migrates to the point in the gradient where the CsCl density is also 1.7 gm cm^{-3}. Proteins have lower, whereas RNA has higher buoyant densities.

3. Ultraviolet Spectroscopy – This involves analysis of spectra produced by a particular substance. Every substance has a characteristic spectrum of the light it emits or absorbs. For example, DNA strongly absorbs ultraviolet radiation with a wavelength of 260 nm; proteins have a strong absorbance at 280 nm. Avery showed that the ultraviolet absorbance spectrum of the transforming principle is the same as that of pure DNA.

4. Elementary Chemical Analysis – This involves the chemical analysis of compounds. This yielded an average nitrogen/phosphorous ratio of 1.67, about what one would expect for DNA, which is rich in both elements, but vastly lower than the value expected for a protein which is rich in nitrogen but poor in phosphorous. Even a slight protein contamination would have revealed itself by raising the nitrogen / phosphorous ratio.

You may take a short break here, after attempting the SAQ given below.

SAQ 1

Insert the correct words in the blank spaces.

i) Experiments on mice and ..... showed that structural and physiological character can be ..... from one bacterium to another (....., 1928). This transfer of characters appears to be associated with a transfer of ..... (....., ... and ..... , 19 .....).

The Hershey Chase Experiment

The third experiment, that was directed at chemical characterisation of the genetic material, was carried out in 1951-52, seven years after Avery's work was published. By then experimental techniques had progressed, and new approaches to problem solving in genetics were available. In particular, the use of bacteriophages as experimental tools for studying molecular genetics had become fully established.

Bacteriophages or phages as they are commonly known, are viruses that specifically infect bacteria. Phage T2 (Fig. 11.10), for example, is one of several types of phages that are specific for the bacterium *Escherichia coli*. The phage consists of a protein coat which surrounds a core of DNA. Electron micrographs have revealed that the protein coat portion that contains the genetic material is hexagonal and this entire portion is called head. In addition, there is a tail also. The life cycle of bacteriophage T2 is shown in Fig. 11.11.

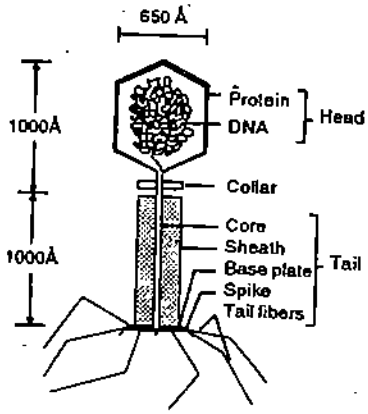


Fig. 11.10 : Diagrammatic sketch of the *E. coli* phage T2 showing its various parts.

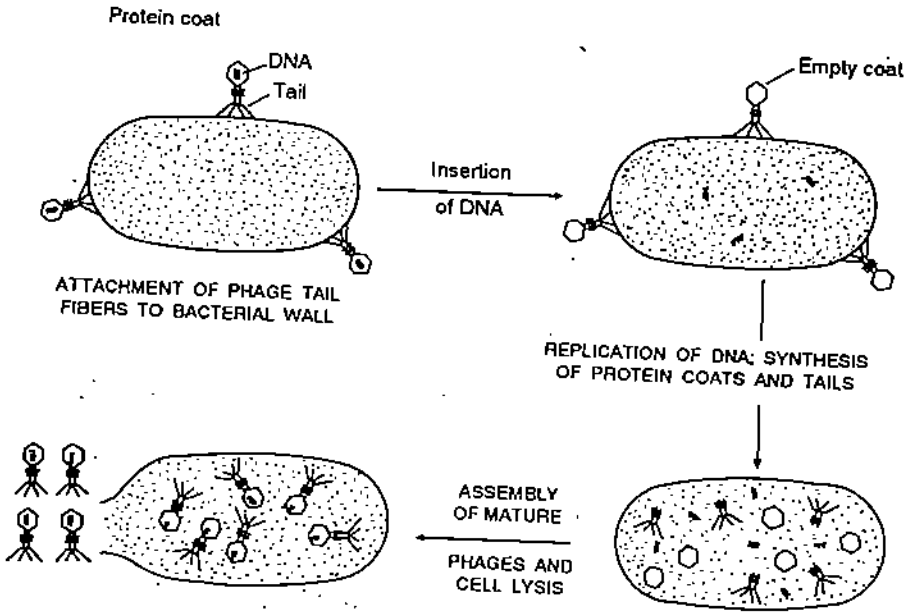


Fig. 11.11 : The life cycle of bacteriophage T2.

Alfred Hershey (Fig. 11.12) and Martha Chase in 1952 published the results of experiments designed to clarify the events leading to phage reproduction. Several of the experiments clearly established the independent function of phage's coat protein and nucleic acid in the reproductive process associated with the bacterial cell. From the existing data Hershey and Chase knew that: (i) T2 phages consist of approximately 50 per cent protein and 50 per cent DNA; (ii) infection is initiated by adsorption of the phage through its tail fibers to the bacterial cell; (iii) the reproduction of new viruses occurs within the bacterial cell.

While working with this phage, it appeared to Hershey and Chase that some molecular component of the phage DNA and/or protein enters the bacterial cell and directs viral reproduction. What could possibly this component be, was the main question arising in their minds. Both DNA and protein contain elements most characteristic of living organisms —carbon, hydrogen, oxygen and nitrogen. They differ in one major aspect, that is, sulfur occurs only in proteins not in DNA while phosphorous occurs only in

Fig.11.12: Alfred Hershey (b. 4 Dec, 1908, Owosso, Michigan)

Alfred Hershey began his career as a microbiologist, obtaining his Ph.D. from the Michigan State College in 1934 for his work on the *Brucella* bacteria. He then moved to Washington University, St. Louis where he stayed until 1950. Then he joined the Genetics Department of the Carnegie Institute of Washington at Cold Spring Harbor, New York, becoming Director in 1962. He retired in 1974 and now lives near Cold Spring Harbor on Long island. He was a founder member of Delbruck's Phage Group, and made several discoveries that led to an understanding of the phage infection cycle. The finding that phage DNA is injected into the host cell was really very crucial that enabled scientists to use phages as an invaluable tool for the genetic research. In 1969 he shared the Nobel Prize with Salvador Luria and Max Delbruck.

DNA and not proteins. Based on the information that the two macromolecules are different, they designed their experiment. They differentially radiolabelled the macromolecules (DNA or protein) with radioisotopes ('marker') to follow the molecular components of phages during infection. Both P^{32} and S^{35} – radioactive isotopes of phosphorus and sulfur were used respectively.

Hershey and Chase began their experiment (Fig. 11.13) by growing cultures of phages T2 (Fig. 11.10) in suspensions of *E. coli* rich in one of two radioactive isotopes, sulfur –35 (S^{35}) or phosphorus – 32 (P^{32}). They then proceeded to identify the material injected into the cell by phages attached to the bacterial wall.

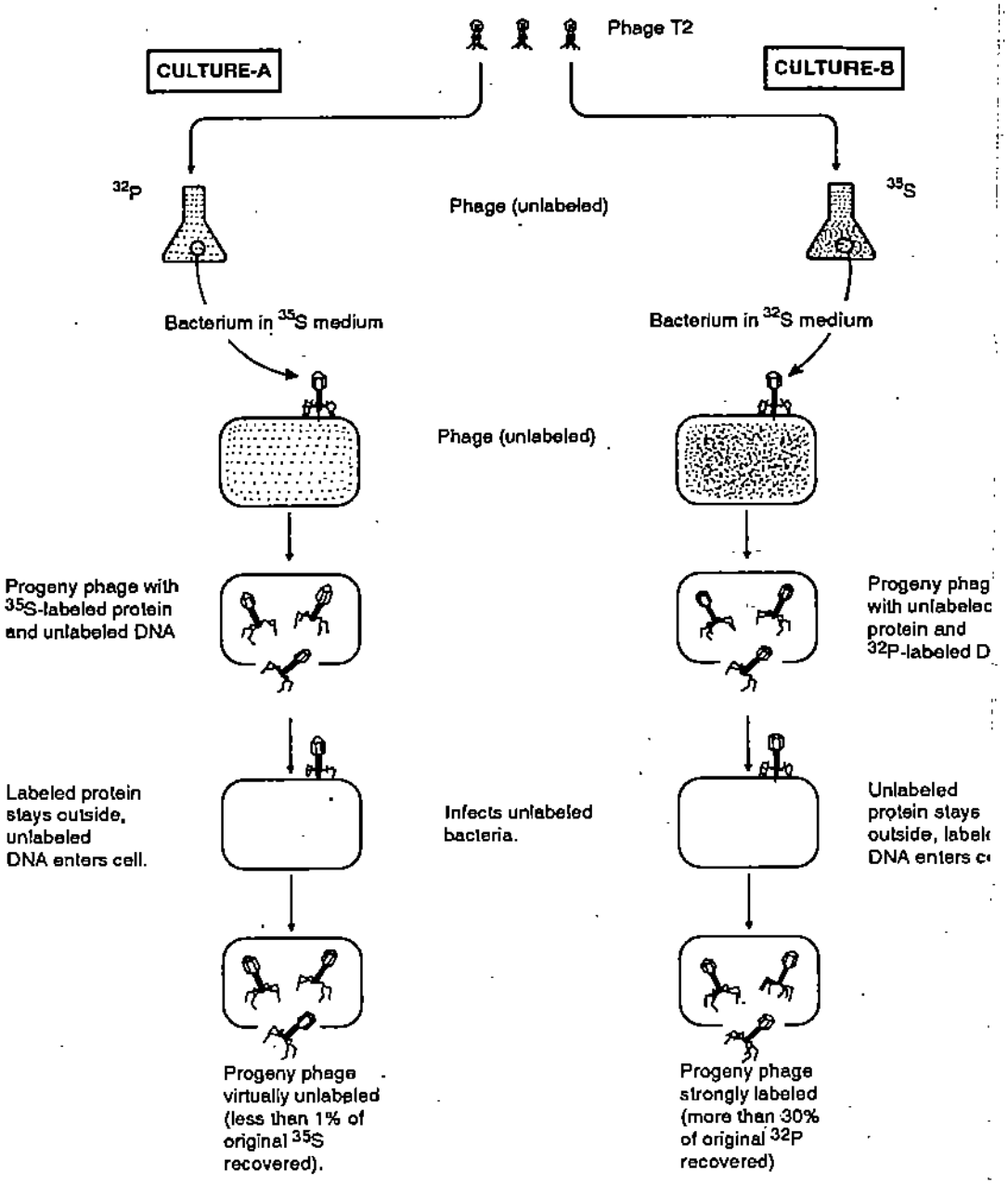


Fig. 11. 13 :Diagrammatic representation of Hershey and Chase's experiment, that demonstrated that DNA and not protein is responsible for directing the reproduction of phage T2 during the infection of *E. coli*. This is based on the fact that nucleic acid lebel (P^{32}) enters the bacterium during infection. The protein label (S^{35}) does not.

All elements exist in a number of forms called isotopes. Each isotope differsonly in the numbersof neutrons in the atomic nucleus. Different isotopes of the same atomic number have different atomic masses.

Isotopes may be stable (e.g., C^{12}) or unstable (e.g., C^{14}). The unstable ones decay into a stable form (possibly not of the same element) with emission of radiation of characteristic energy. With the help of Scintillation Counter, we can measure the amount of radioactive disintegration and also identify the radioactive atom involved. The biological systems or cells cannot distinguish between isotopes, e.g., a bacterium cannot differentiate between C^{14} – glucose and C^{12}– glucose, and will use each equally efficiently as carbon source. Based on this Hershey and Chase prepared bacteriophages P^{32} and S^{35}.

Have you studied Figure 11.13 carefully? Then how about answering the following questions?

SAQ 2

a) Which molecules of the phage in culture A will be radioactively labelled?

.....

b) And which molecules of phage in B are radioactively labelled?

.....

All the T2 derived from these cultures thus had their DNA labelled with P^{32} or the protein coats labelled with S^{35}. Next, the unlabelled cultures of *E. coli* were infected separately with S^{35}- labelled and P^{32}– labelled T2. In the case of the P^{32}– labelled phages that were mixed with *E. coli* they found that the P^{32} label entered the bacterial cells and the next generation of phages that burst from the infected cells carried a significant amount of P^{32} label. And in the S^{35} labelled phages mixed with unlabelled *E. coli*, the researchers found that the S^{35} label stayed on the outside of the bacteria in the phage's coats. From these observations they concluded that the outer protein coat of a phage does not enter the bacterium it infects, whereas the phage's inner material consisting of DNA does enter the bacterial cell. Since the DNA is responsible for the production of new phages during the infection process, the DNA, not the protein must be the genetic material.

Transfection Experiments – During the eight years following the publication of the Hershey-Chase experiment, further research provided additional proof that DNA is the genetic material. These studies involved the same organisms as were used by Hershey and Chase.

In 1957, there were several reports showing that naked cells of *E.coli* could be obtained on treating the cells with the enzyme *lysozyme*. This enzyme dissolves the outer wall without destroying the bacterium. Such naked structures are called *protoplasts* or *spheroplasts*. John Spizizen and Dean Fraser independently reported that by using protoplasts, they were able to initiate phage multiplication with disrupted T2 particles. That is, provided protoplasts are used, it is not necessary for virus to be intact in order for infection to occur.

Later, in 1960, similar but refined experiments were reported by George Gutherie and Robert Sinsheimer. They purified DNA from the bacteriophage ϕ X-174. It is a small phage consisting of a single-stranded, circular DNA molecule of some 5500 nucleotides. When added to the *E.coli* protoplasts, the purified DNA resulted in the production of complete ϕ X-174 bacteriophages. This process of infection by only the viral nucleic acid, is called **transfection**, proves beyond doubt that the phage DNA alone contains all the necessary information for production of mature viruses. Thus, the evidence that DNA serves as the genetic material was further strengthened, even though all direct evidence had been obtained from bacterial and viral studies.

11.2.5 RNA as Genetic Material

All viruses are not limited to bacterial hosts. Other kinds of viruses infect and parasitise animal and plant cells. One plant virus, *tobacco mosaic virus* (TMV), that contains RNA, not DNA was an important tool for genetic experiments in the late 1940's. TMV infects tobacco, causing the infected regions on leaves to become discoloured and bristled. Different strains of TMV produce clearly different inherited lesions on the infected leaves. The common virus (TMV-*common*) produces a green mosaic disease, but a variant Holmes Ribgrass (TMV-*HR*), produces ringspot lesions. Moreover, the amino acid compositions of the proteins of the two strains differ.

Heinz Fraenkel-Conrat and Bea Singer found that the virus could be broken into component parts and these could again be reassembled or reconstituted to form functional virus. From the two strains of TMV, they were able to reconstitute viruses with the RNA from TMV-*common* enclosed in TMV-*HR*–protein and TMV-*HR* RNA with TMV-*common* protein (Fig. 11.14). When these reassembled viruses were used to infect tobacco leaves, the progeny viruses produced were always found to be phenotypically and genotypically identical to the parent strain from which the RNA had been obtained. The reassembled viruses with TMV-*common* RNA and TMV-*HR* protein produced a green mosaic disease characteristic of TMV- *common*. Recovered progeny virus had protein characteristic of TMV-*common*. Similarly, the virus containing TMV-*HR* RNA with TMV-*common* proteins produced ringspot lesions and the virus progeny had protein characteristic of TMV-*HR*. The source of the RNA determined the nature of the mosaic infection and of the proteins present in the progeny viruses indicated that nucleic acid in the progeny viruses indicated that nucleic acid in

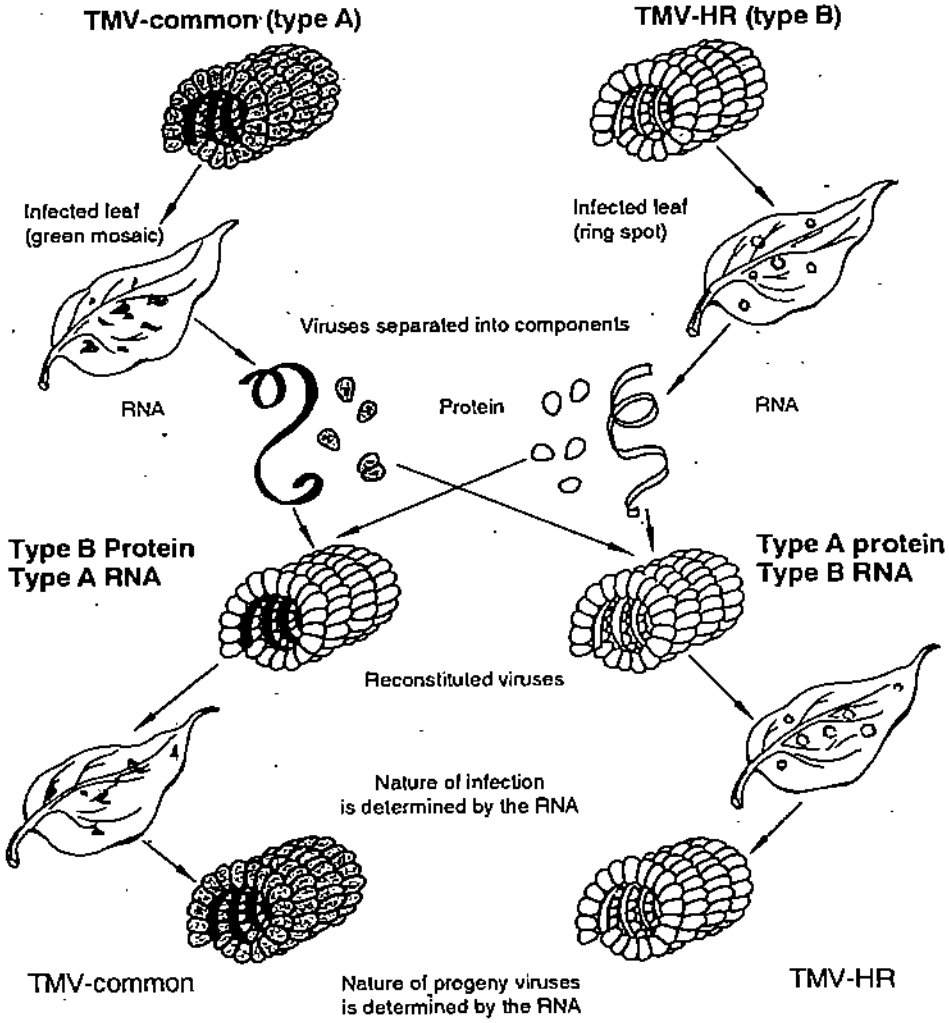


Fig. 1L14: Genetic effects of reconstituting tobacco mosaic viruses by using RNA and proteins from two different strains.

this case was RNA. Hence, RNA carries genetic information not proteins.

SAQ 3

Choose the correct answer:

i) DNA is believed to be the genetic material because:

   a) all the body cells of an individual seem to have identical amounts and composition of DNA, while reproductive cells have half the amount of DNA found in body cells.

   b) the proteins are the same from cell to cell in an individual, but DNA differs, thus DNA must be the material that makes tissues different.

   c) DNA is the largest type of macromolecule found in living organisms.

   d) DNA is found in the cell nucleus.

ii) By differentially labelling the coat protein and the DNA of phage T2, Hershey and Chase demonstrated that:

   a) only the proteins enter the infected cell

   b) the entire virus enters the infected cell

   c) the phage genetic material is most probably DNA

   d) the phage coat protein directs synthesis of new progeny phage.

SAQ 4

What is the macromolecular composition of a bacterial virus or bacteriophage such as a phage T2?

.....

.....

.....

.....

.....

.....

.....

.....

.....

SAQ 5

With respect to which particular component do the proteins and DNA differ from each other and that allowed researchers to specifically label one or the other of these macromolecules with radioactive isotope?

.....

.....

.....

.....

.....

.....

.....

.....

.....

11.3 THE CHEMICAL COMPOSITION OF DNA AND RNA

The results of the experiments that we have described above, all point to nucleic acid, DNA or RNA as the carrier of genetic information. Now we need to examine the chemical structure of these molecules. Their structure will tell us a good deal about how they function. Although you have already learnt the basics of nucleotide structure in LSE-01, Block 1, Unit 5. Nevertheless, we shall briefly recapitulate the structure of nucleic acids before going into further details.

Nucleic acids, both DNA and RNA are **macromolecules**, which means they have a molecular weight of atleast a few thousand daltons. One dalton is equivalent to a twelfth of the mass of the carbon^{12} atom or 1.67×10^{-24}g. In this respect they differ from many other molecules important to cell function such as sugars and amino acids, which weigh from a hundred to a few hundred daltons.

Both the nucleic acids are polymeric molecules and are composed of repeating subunits called **nucleotides**. Each nucleotide is made up of three components: (1) a phosphate group, (2) a five-carbon sugar or pentose, and (3) a cyclic nitrogen containing compound called a nitrogenous base or simply base (Table 11.2; Fig.11.15).

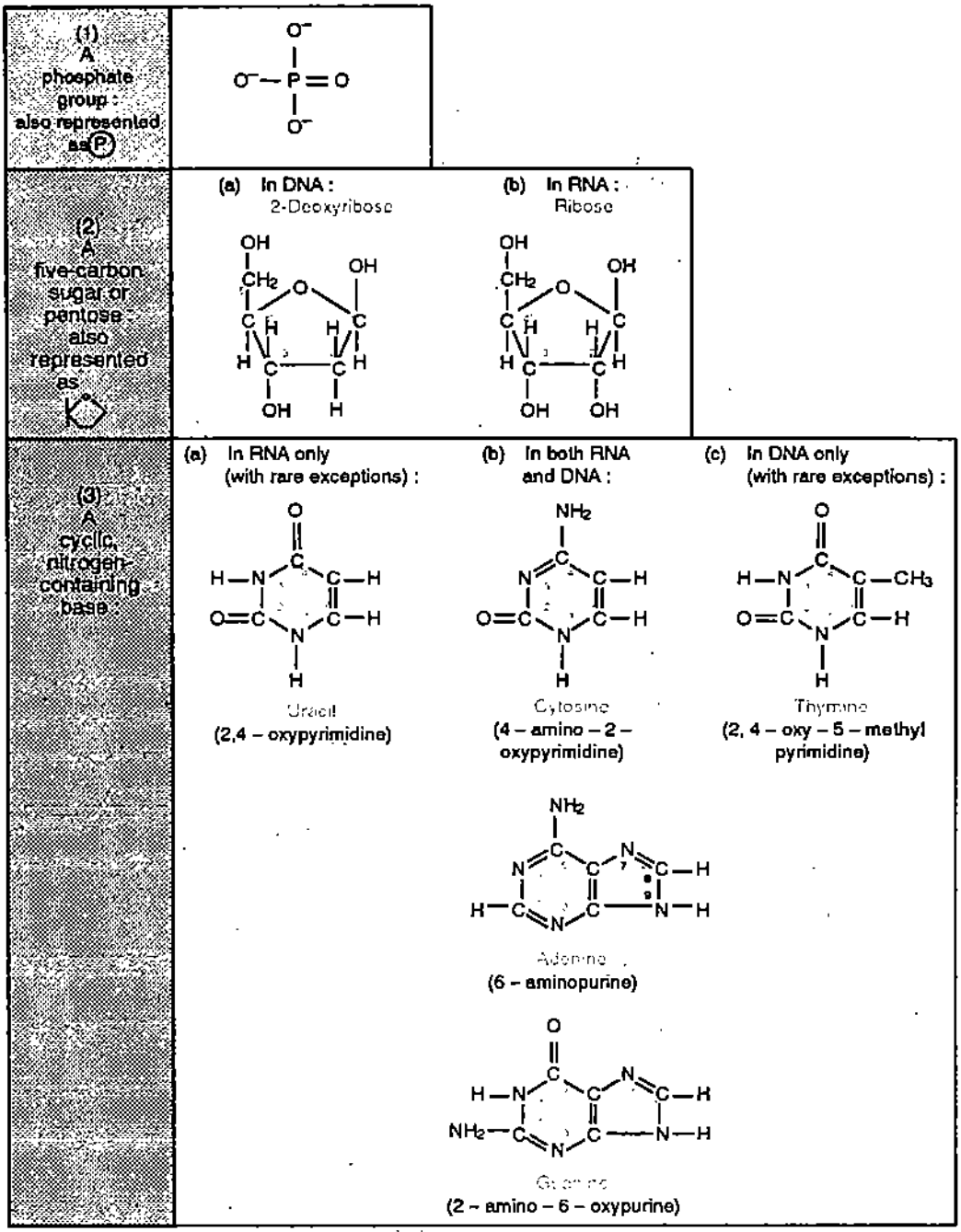


Fig. 11.15: Structural formulae of the components of nucleic acids: phosphate, sugars and bases. When the pentoses are present in nucleosides, nucleotides or nucleic acids, the five carbons are numbered

When incorporated into a nucleic acid, a nucleotide contains one each of the three components. But, when free in the cell pool, nucleotides usually occur as triphosphates. The energy held in extra phosphates is used, among other purposes to synthesise the polymer. A **nucleoside** is a sugar-base compound. Nucleotides are therefore nucleoside phosphates (Fig. 11.16). Note that ATP, adenosine triphosphate, the energy currency of the cell, is a nucleoside triphosphate.

The sugars differ only in the presence (ribose in RNA) or absence (deoxysibose in DNA) of an oxygen in the 2'position. The carbons of the sugar are numbered 1' to 5'. The primes are used to avoid confusion with the numbering system of the bases (see Fig. 11.15). There are four different bases found in DNA: **adenine, guanine, thymine, and cytosine** (see Table 11.2 again). RNA also usually contains adenine, guanine, and cytosine, but has a different base, **uracil**, in place of thymine. Thus three of the nitrogenous bases are found in both DNA and RNA, whereas thymine is unique to DNA and uracil is unique to RNA. Adenine and guanine are double-ring bases called **purines**; cytosine, thymine and uracil are single-ring bases called **pyrimidines** (Fig. 11.2).

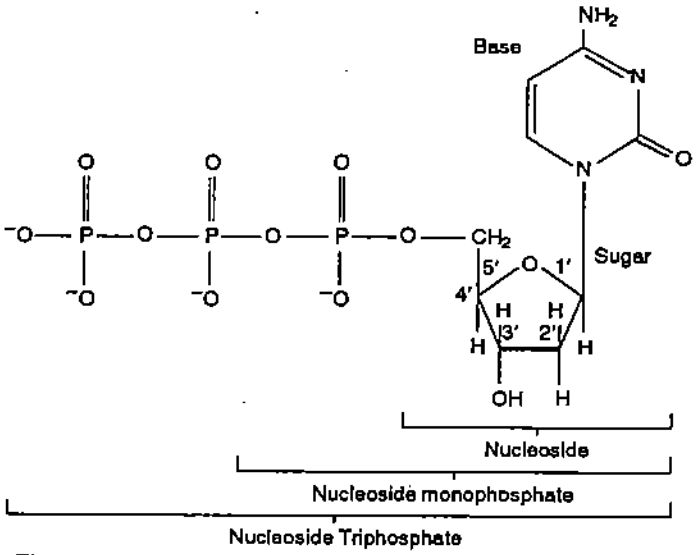


Fig.11.16 : The structure of a nucleoside and nucleotides: a nucleoside monophosphate and a nucleoside triphosphate.

Table 11.2 : Components of Nucleic Acid.

			Base	
	Phosphate	**Sugar**	**Purine**	**Pyrimidine**
DNA	Present	Deoxyribose	Guanine Adenine	cytosine *Thymine*
RNA	Present	Ribose	Guanine Adenine	Cytosine *Uracil*

A nucleotide is formed, by the attachment of a base to the 1′ carbon of the sugar and attachment of a phosphate to the 5′ carbon of the same sugar (Fig. 11.17) – the first

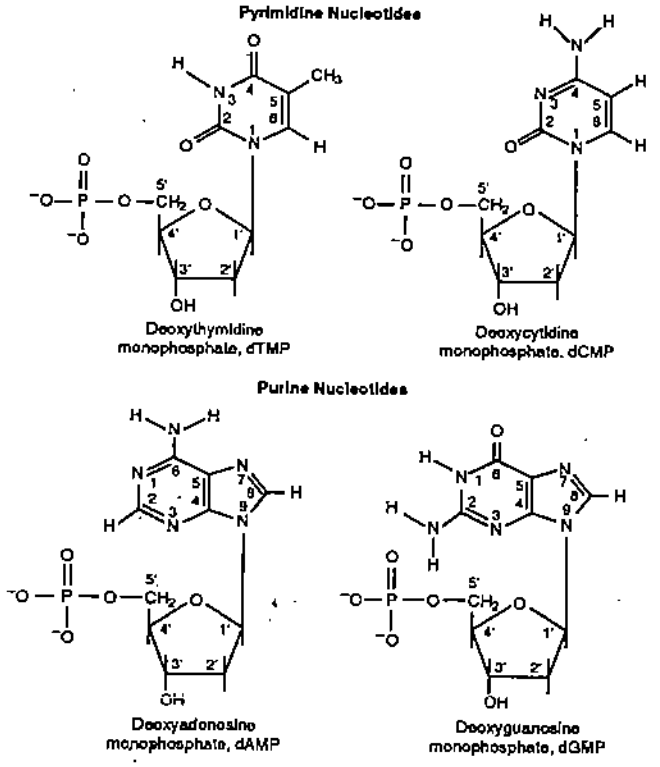


Fig. 11.17: Structure of the four common deoxyribonucleotides of DNA. RNA contains similar nucleotides, which contain the pyrimidines uracil and cytosine and the purines adenine and guanine.

phosphate ester bond (see Fig.11.18) ; it takes its name from the base (Table 11.3). Nucleotides are linked together or **polymerised** by the formation of a bond between the phosphate of one nucleotide and hydroxyl (OH) group at the 3′ carbon of an adjacent molecule – the second phosphate ester bond (see Fig.11.18). Very long strings of nucleotides can be thus polymerised by the **phosphodiester bonding** (Fig. 11.18 and Box 11.4).

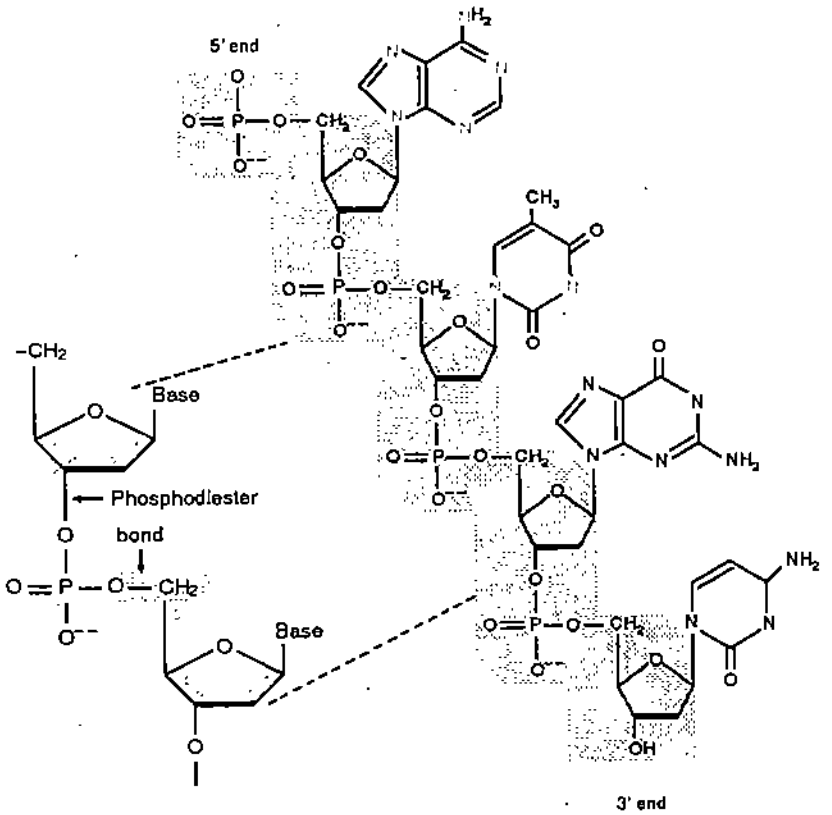


Fig. 11.18: Formation of a short polydeoxy ribonucleotide strand by polymerisation of adjacent nucleotides. Note the role of phosphodiester bonds. There is no limit to the length that the strand can be or the type of base attached to each nucleotide residue.

Table 11.3: Nucleotide Nomenclature.

Nucleotide (Nucleoside monophosphate)		Symbol					
		Monophosphate		Diphosphate		Triphosphate	
		Ribose	Deoxyribose	Ribose	Deoxyribose	Ribose	Deoxyribose
Guanine	Guanosine monophosphate	GMP		GDP		GTP	
	Deoxyguanosine monophosphate		dGMP		dGDP		dGTP
Adenine	Adenosine monophosphate	AMP		ADP		ATP	
	Deoxyadenosine monophosphate		dAMP		dADP		dATP
Cytosine	Cytidine monophosphate	CMP		CDP		CTP	
	Deoxycytidine monophosphate		dCMP		dCDP		dCTP
Thymine	Deoxythymidine monophosphate		dTMP		dTDP		dTTP
Uracil	Uridine monophosphate	UMP		UDP		UTP	

Both DNA and RNA, therefore contain four different subunits or nucleotides, two purine nucleotides and two pyrimidine nucleotides (Fig. 11.17). RNA usually exists as a single-stranded polymer that is composed of a long sequence of nucleotides. DNA, however, has one very important additional level of organisation, it is usually a double stranded molecule.

After studying the chemical composition of DNA and RNA, how about attempting a couple of SAQs now?

SAQ 6

a) What are the four bases commonly found in DNA?

.....

b) What five-carbon sugar is found in DNA?

.....

c) What is the difference between a nucleoside and a nucleotide?

.....

.....

d) Which base is unique to DNA and which base is unique to RNA?

.....

SAQ 7

a) How many phosphate groups are there per base in DNA, and how many phosphates are there in each precursor for DNA synthesis?

.....

.....

.....

.....

b) Which chemical groups are at the ends of a single polynucleotide strand?

.....

.....

.....

.....

c) Does the incorporation of nucleotides into a polynucleotide chain impose any limitations on the type of bases in the chain or the order in which these bases are arranged along the chain?

.....

.....

.....

.....

11.4 THE PHYSICAL STRUCTURE OF DNA: THE DOUBLE HELIX

The discovery of the double helix by James Watson and Francis Crick at Cambridge, in 1953, was one of the great deductive triumphs in the history of science, and it has influenced every aspect of molecular genetics, indeed of biology as a whole.

11.4.1 The Double Helix

Watson and Crick made use of the results of several lines of investigations concerning the structure of DNA. Based on these findings, they began to build models of various possible structures of DNA. Only the double helix was compatible with all the data. Their double helix model of DNA structure was based on two major kinds of evidence.

1. **Chargaff's Base Ratios** – E. Chargaff (Fig. 11.19) together with his colleagues E. Vischer and S. Zamenhof analysed the DNA of many different organisms. They used sensitive paper chromatographic techniques (see Box 11.2) to determine the exact amounts of each of the four nitrogenous bases in samples of DNA from different tissues of different organisms. Their results were quite startling. They observed that *the concentration of thymine was always equal to the concentration of adenine and the concentration of cytosine was always equal to concentration of guanine.* That is A = T and G = C. This strongly suggested that thymine and adenine as well as cytosine and guanine were present in DNA with some fixed interrelationship. It implies that the total concentration of pyrimidines (thymine plus cytosine) always equals the total

Fig. 11.19: Erwin Chargaff b. 11 August, 1905, Austria

Chargaff received his Ph.D. from the University of Vienna in 1928. He then joined Columbia University, in 1935. He worked in many fields of biochemistry such as lipid metabolism and blood coagulation. His most known contribution was the demonstration during late 1940's that the base ratios in DNA are constant. Although this was one of the crucial discoveries that eventually led to the understanding of structure of DNA, Chargaff has never claimed any allegiance to molecular biology and he has become one of its sternest critics (see Nature 326, 199-200, 1987). He explains his objection to molecular biology by its claim to be able to explain everything. He felt, it actually impedes the flow of scientific explanation. His autobiography is entitled '*Heraclitean Fire: Sketches From a Life Before Nature*', Rockefeller University Press, New York, 1978.

concentration of purines (adenine plus guanine). This rule is illustrated by Table 11.4. However, the (thymine + adenine)/(cytosine + guanine) ratio is found to vary in different species (see Table 11.5).

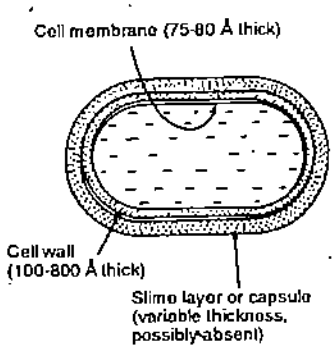


Table 11.4: Chargaff's results in selected organisms indicating a consistent pattern of ratios of bases in DNA from different organisms. (Data from E. Chargaff, 1963, Essays on nucleic acids. Elsevier, p.30).

Organism	Ratio of bases in DNA sample	
	A:T	G:C
Cow	1.04	1.00
man	1.00	1.00
Salmon	1.02	1.02
E. coli	1.09	0.99

Table 11.5: Base composition of DNA from some organisms (data from H. Sober ed). Handbook of Biochemistry, Selected Data of Molecular Biology. The Chemical Rubber Co. Cleveland, 1970.

	Species	Nitrogenous bases (in %)				Molar Ratios	
		ADENINE	GUANINE	CYTOSINE	THYMINE	$\frac{A+G}{T+C}$	$\frac{A+T}{G+C}$
I.	Viruses						
	Bacteriophage	26.0	23.8	24.3	25.8	0.99	1.08
	Bacteriophage T2	32.6	18.1	16.6	32.6	1.03	1.88
	Herpes simplex	13.8	37.7	35.6	12.8	1.06	0.36
II.	Bacteria						
	Escherichia coli	26.0	24.9	25.2	23.9	1.04	1.00
	Diplococcus pneumoniae	29.8	20.5	18.0	31.6	1.02	1.59
	Micrococcus lysodeikticus	14.4	37.3	34.6	13.7	1.07	0.39
	Ramibacterium ramosum	35.1	14.9	15.2	34.8	1.00	2.32
III.	Fungi						
	Neurospora crassa	23.0	27.1	26.6	23.3	1.00	0.86
	Aspergillus niger	25.0	25.1	25.0	24.9	1.00	1.00
	Saccharomyces cerevisiae	31.7	18.3	17.4	32.6	1.00	1.80
IV.	Higher Eukaryotes						
	Zea mays (corn)	25.6	24.5	24.6	25.3	1.00	1.04
	Nicotiana tabacum (tobacco)	29.3	23.5	16.5	30.7	1.12	1.50
	Arachis hypogaea (peanut)	32.1	17.6	18.0	32.2	0.99	1.80
	Drosophila melanogaster (fruit fly)	30.7	19.6	20.2	29.4	1.01	1.51
	Bombyx moris (silkworm)	30.7	18.9	19.4	31.1	0.98	1.61
	Rana pipiens (frog)	26.3	23.5	23.8	26.4	0.99	1.11
	Homo sapiens (human)						
	liver	30.3	19.5	19.9	30.3	0.99	1.53
	thymus	29.8	20.2	18.2	31.8	1.02	1.60
	spermatozoa	30.5	19.9	20.6	28.9	1.02	1.47

Box 11.2: Chromatography

Chromatography techniques are used to separate compounds on the basis of their relative affinities or absorption on to a solid matrix. They are particularly useful for separating individual proteins, amino acids or nucleotides.

There are several different chromatography methods distinguished primarily by the nature of the solid matrix used (see Fig. 11.20). In paper chromatography this matrix is a strip of filter paper. A sample of the compounds to be analysed is placed at one end of the strip which is dipped in a solvent that may be aqueous or an organic solution. As the solvent moves along the paper it carries the compounds with it, but at different rates depending on the relative affinities of the compounds for absorption on the matrix.

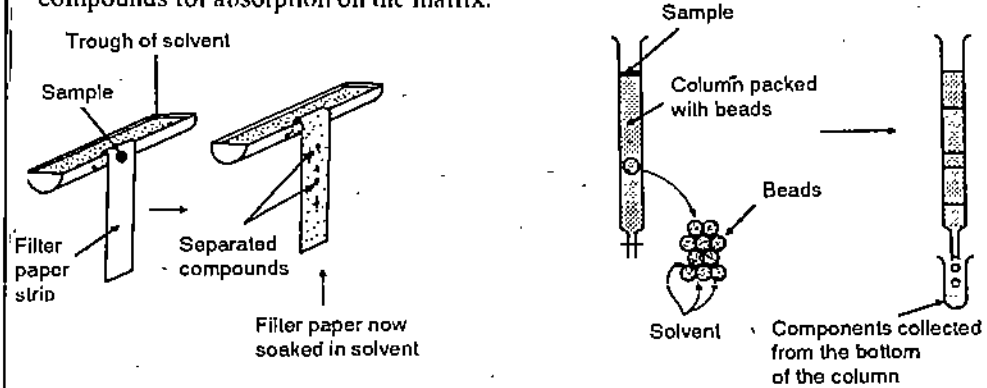


Fig. 11.20 : Set up for chromatography

Chargaff used this method for analysing the base composition of DNA. He first carried out the chemical breakdown of the molecule into its component nucleotides by treatment with acid or alkali. The chromatograph was run by the method described above, and the resulting spots cut out. The pure nucleotides (from the spots) were recovered from the paper by soaking in an aqueous solution, and their concentrations measured.

Now a days chromatography is also carried out in a glass column packed with tiny beads, composed of substances such as cellulose or agarose, immersed in the solvent. the sample is layered on to the top of the column and eluted by passing through more solvent. The separated compounds are then collected as they drip out of the bottom of the column.

Units of length

1 micrometer or micron (μm) = 10^{-6} meters = one-thousandth of a millimeter

1 nanometer (nm) = 10^{-9} meters = one-millionth of a millimetre

1 Angstrom unit (Å) = 10^{-10} meters

2. **X-ray Diffraction Analysis** — The second piece of evidence utilised by Watson and Crick was the **X-ray diffraction pattern.** Maurice Wilkins, Rosalind Franklin (see Fig. 11.21) and their colleagues in 1952 employed **X-ray chrystallography** (also see Box 11.3) to analyse the structure of DNA. On analysing these pictures they found that DNA was a highly ordered, multiple-stranded structure with repeating subunits placed every 3.4 Angstroms [1 Angstrom (Å) = 10^{-8} cm] along the axis of the molecule.

Box 11.3: X-ray Diffraction Analysis

When X-rays are focussed through isolated macromolecules or crystals of purified molecules, theX-rays are deflected by the atoms of the molecules in a specific pattern called the **diffraction pattern.** These diffraction patterns can be recorded on an X-ray sensitive film just like one photographs different objects with a camera (see Fig. 11.22). The result is a pattern of spots, and from the positions and intensities of which the structural organisation of the macromolecule or the crystal can be deduced. The X-ray diffracton pattern of DNA shows a cross-like structure (see the Figs 11.22 and 11.23) which indicates that the molecule is a helix; the dark areas at the top and bottom are due to the presence of bases, stacked perpendicularly to the main axis of the molecule.

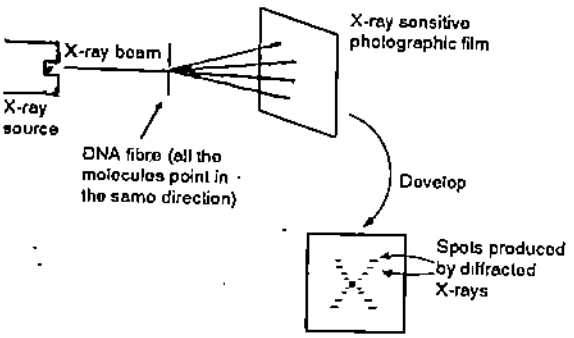


Fig. 11.22 : X-ray diffraction analysis set up.

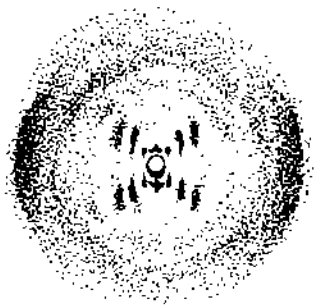


Fig. 11.23 Diagrammatic representation of X-ray diffraction pattern of DNA. Note the central cross-shaped structure.

Fig. 11.21: (a) Maurice Wilkins born–15 Dec, 1916 New Zealand.

(b) Rosalind Franklin born — 25th July 1920, London died — 16th April, 1958, London.

Maurice Wilkins began his career as a physicist, but his experience on the atomic bomb project during the war turned him away from physical sciences, towards biology. He joined the Kings College (London) in 1947. There he began studying the structure of chromosomes and genes using physical methods. In May 1950, he obtained a sample of DNA (that was believed to be the purest at that time) from Rudolf Singer of Bern. He made detailed X-ray diffraction analysis of the sample and the diffraction patterns thusobtained proved extremely useful in determining the structure of DNA.

Rosalind Franklin joined King's College in 1951 from Paris. She too started working on a DNA sample obtained from Singer. She analysed the pictures from DNA analysis and found that it had a highly ordered structure. Unfortunately there was a personality clash between Franklin and Wilkins and the intellectual discussion needed to solve the structure of DNA (which Watson and Crick enjoyed) never took place. Nevertheless, Franklin was close to the idea that the structure was a double helix. She moved to Birkbeck College in 1953 and remained there until she died of cancer.

Watson and Crick (see Fig. 11.24) considered all the evidence just described and began to build three- dimensional models for the structure of DNA. The model they devised, which fits all the known data on the composition of DNA molecule was the now famous double helix model for DNA. It is also known as the Watson-Crick model. According to this model two polynucleotide chains constitute the double helix. These two chains are coiled about one another in a spiral with the sugar phosphate backbones on the outside and bases on the inside (Fig. 11.25). Each polynucleotide chain is made up of nucleotides arranged in a sequence and linked together by **phosphodiester bonds,** joining adjacent deoxyribose moieties (see Fig. 11.25 and Box 11.4). The two polynucleotide strands are held together in their helical configuration by **hydrogen bonding** (Box 11.4) between the bases of the opposite strands, the resulting base pairs being stacked between the two chains perpendicular to the axis of the molecule like the steps of a spiral staircase (Fig. 11.25). The **base pairing is specific.** That is, **adenine** is always **paired with thymine** and **guanine** is always **paired with**

Fig 11.24 : a) James Watson born – 6th April, 1928, Chicago.

b) Francis Crick born - 8th June, 1916, Northampton

James Watson entered the University of Chicago at the age of 15 and graduated in 1947. He obtained his Ph.D. from the University of Indana in 1950. He was then awarded a fellowship to study in Copenhagen, but after a year there, he moved to Cavandish Laboratory, Cambridge with the specific objective of studying the gene. There he met Francis Crick, who was working on the structure of proteins for his Ph.D. Crick started as a physicist but became diverted to biology after the Second World War. Both of them worked together and the famous fusion of minds led Watson and Crick to deduce that DNA is a double helix. The two shared a Nobel Prize in 1962 with Maurice Wilkins.

After the double helix, Watson and Crick went separate ways. Watson returned to United States in 1953 and then to Harvard, and in 1976 to Cold Spring Harbor Laboratory. He worked on RNA synthesis, protein synthesis and the role of virus in cancer. Crick stayed in Cambridge until moving to Salk Institute in California in 1977. During the 1950s and 1960s he was the major influence in molecular biology and made great advances in the understanding of genes and gene expression.

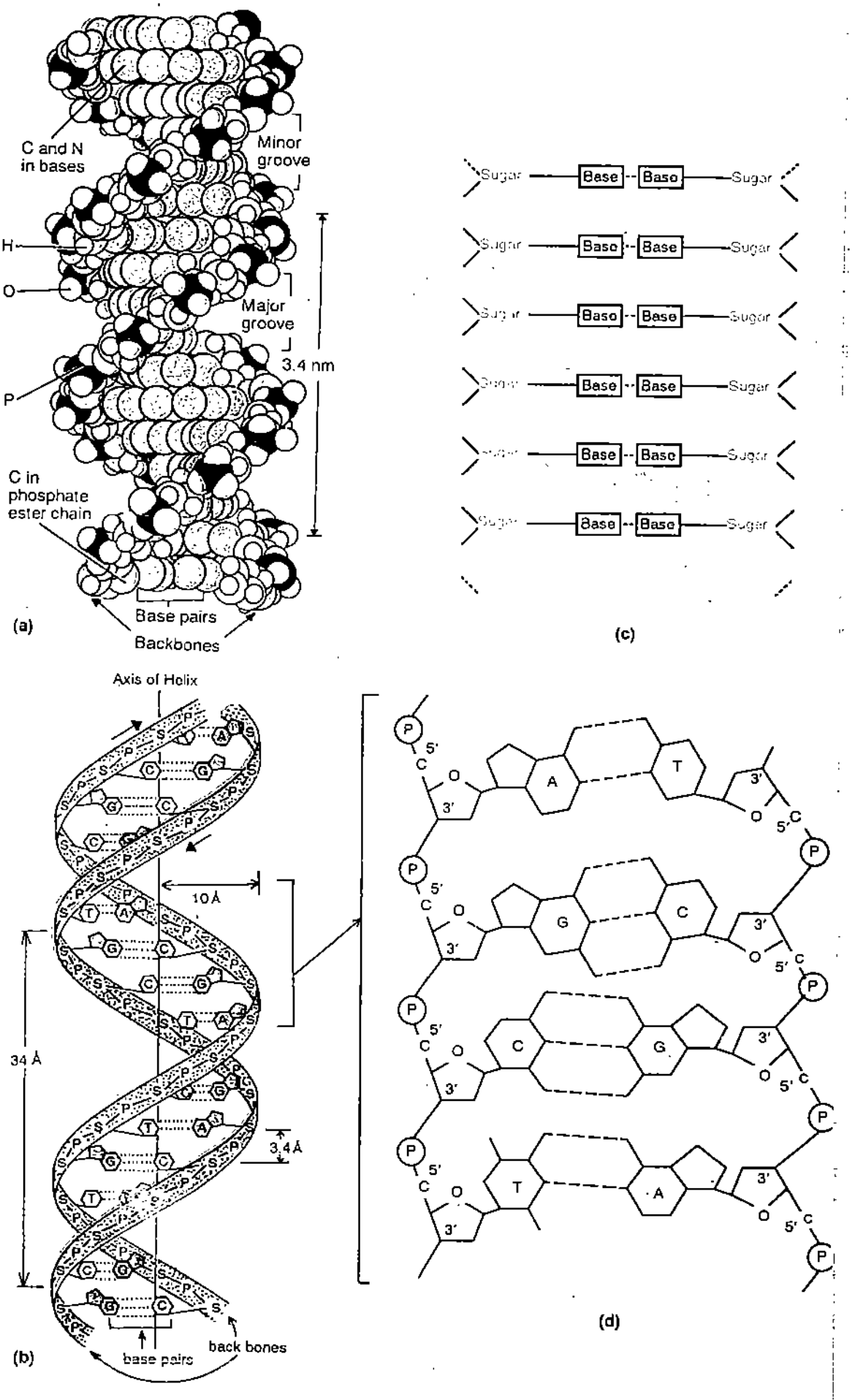


Fig. 11.25: Molecular structure of DNA, (a) space-filling model of the double helix, (b) diagrammatic representation of DNA double helix, (c) the arrangement of nucleotide components in the double helix, and (d) molecular structure of a fragment of DNA double helix.

cytosine (see Fig. 11.25). Thus all base pairs consist of one purine and one pyrimidine. The specificity of base pairing results from the hydrogen-bonding capacities of bases in their normal configurations. Adenine and thymine form two hydrogen bonds and guanine and cytosine form three hydrogen bonds (Fig. 11.26).

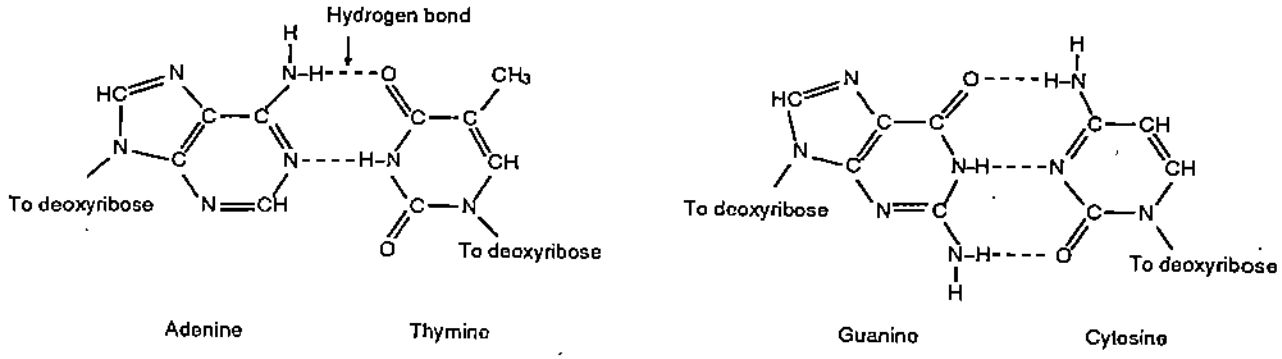


Fig. 11.26: Hydrogen bonding between nitrogenous bases in DNA.

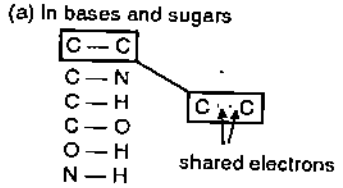
Box 11.4: Chemical Bonds Important in DNA Sequence.

1. Covalent Bond

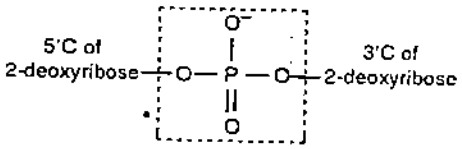
Strong bonds formed by sharing of electrons between atoms. Such bonds are about 1 Å long and generally have energies greater than 50 KCal mol^{-1}

2. Hydrogen Bond

(a) In bases and sugars

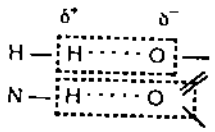


(b) Phosphodiestor linkages

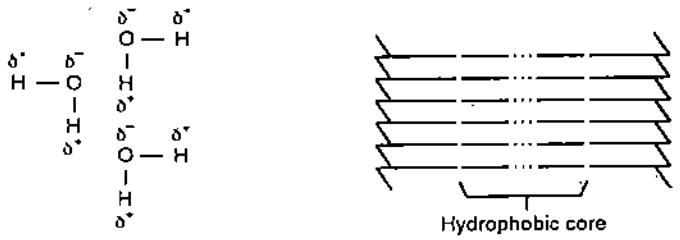


A hydrogen bond is a weak electrostatic attraction between an electronegative atom, such as oxygen and nitrogen and a hydrogen atom attached to a second electronegative atom. It arises because the electrons in a covalent bond such as O – H or N – H are displaced towards the electronegative atom so that the hydrogen atom possesses a partial positive charge and can therefore form a weak bond with a second electronegative atom. Actually, the hydrogen atom is shared between two electronegative atoms. Hydrogen bonds are longer than covalent bonds, about 2Å or more. A hydrogen bond usually has a bond energy between 3 and 10 KCal mol^{-1} at 25°C. A large number of hydrogen bonds can cooperate to form a relatively strong attractive force. Hydrogen bonds play an important role in the three-dimensional structures of macromolecules such as nucleic acids and proteins.

3. Hydrophobic bonds



The association of nonpolar groups with each other when present in aqueous solutions is because of their insolubility in water. Water molecules are very polar (δ^-O and δ^+ Hs). Therefore, it acts as a good solvent for polar compounds but is a poor solvent for nonpolar compounds. Compounds which are polar are soluble in water ("hydrophilic"). Compounds which are non-polar (no longer charged



groups) are insoluble in water ("hydrophobic"). The stacked base pairs provide hydrophobic core.

Once the sequence of bases in one strand of a DNA double helix is known, the sequence of bases in the other strand is also known because of the specific base pairing. The two strands of a DNA double helix are thus said to be **complementary (not identical)**. This property of complementarity of the two strands makes DNA uniquely suited to store and transmit genetic information.

Describing the lengths of DNA molecules.

The basic monomeric unit of the double helix is the base pair,with ten base pairs occurring per turn of helix. The length of a DNA molecule is usually described as so many base pairs, abbreviated to bp (for example 100 bp, 250bp). A kilobase pair, usually called a kilobase or kb is 1000 bp.

The dimensions of the DNA double helix are indicated in Fig. 11.27. The distance between the backbones is 20Å and each base pair is 3.4Å apart. The orientation of the bases rotates with the rotation of the double helix, each base pair being offset by 36° from the adjacent ones. Thus a 360-degree rotation of the helix occurs every ten base pairs, or every 34Å. We can also identify two grooves, a **major groove** and a **minor groove**, bound by the sugar- phosphate backbones. The two grooves differ in size, because the two attachment points of the base pairs to the backbones are not opposite to each other (Fig. 11.27).

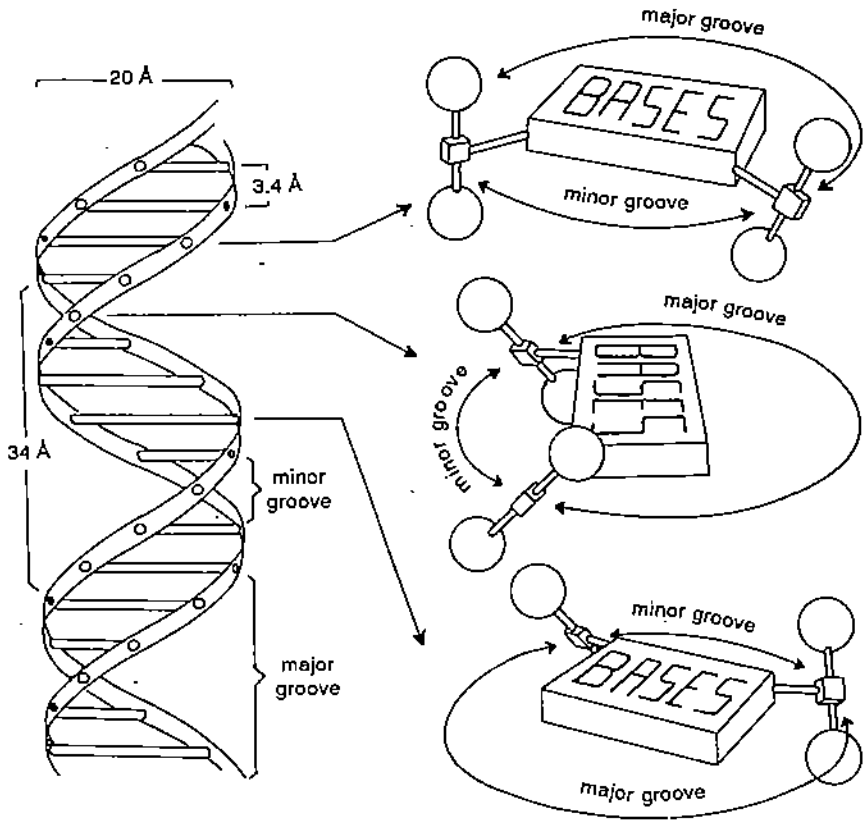


Fig. 11.27: Diagrammatic representation of the double helical structure of DNA (left). The horizontal rods represent base pairs connecting the two ribbon-like sugar-phosphate backbones. On the right are shown the rotation of base pairs in double helix DNA. Cubes represent sugar molecules; spheres represent phosphate molecules.

The sugar-phosphate backbones of the two complementary strands are antiparallel; that is, they have opposite chemical polarity (Fig. 11.28). As one moves unidirectionally along a DNA double helix, the phosphodiester bonds in one strand go from a 3′ carbon

of one nucleotide to a 5′ carbon of the adjacent nucleotide, whereas those of complementary strand go from a 5′ carbon to a 3′ carbon. This opposite polarity of the complementary strands is very important in considering the mechanism of replication of DNA.

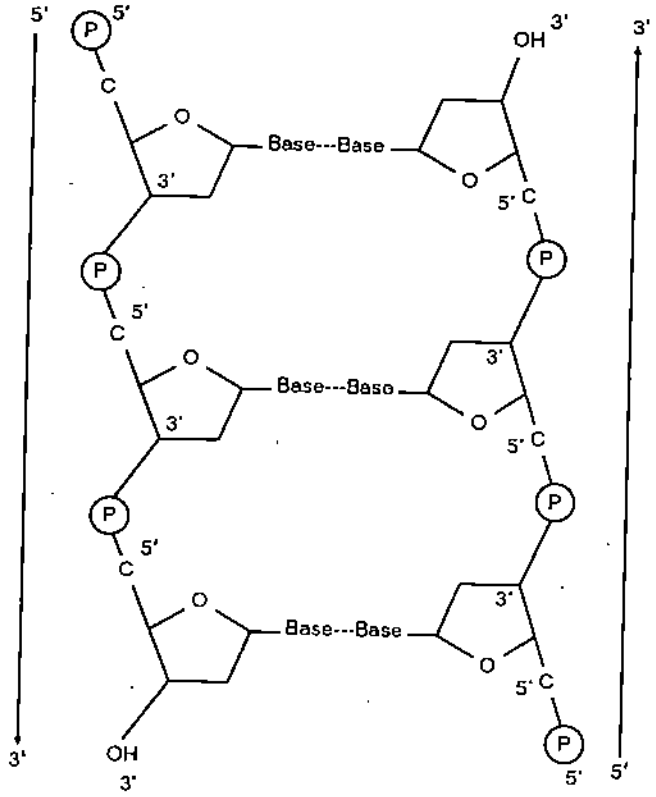


Fig. 11.28 : Polarity of the DNA strands.

The high degree of structural stability of DNA double helices results in part from the large number of hydrogen bonds between the base-pairs (even though each hydrogen bond by itself is quite weak, much weaker than a covalent bond) and in part from the hydrophobic bonding (or 'stacking forces') between the stacked base pairs (see box 11.4 and Fig. 11.25). The planar sides of the base pairs are relatively nonpolar and tend to be water insoluble ("hydrophobic"). This hydrophobic core of base-pairs provide considerable stability to DNA molecules present in the aqueous protoplasm of living cells.

11.4.2 DNA Denaturation and Renaturation

When **denaturation** of double-stranded DNA occurs, the hydrogen bonds of the double helix break, the helix unwinds and the strands separate. However, the covalent bonds do not break. The strand separation can be induced by heat or chemical treatment. While heating a specimen of DNA, a point is reached when the thermal agitation overcomes hydrogen bonding and the molecule becomes denatured. It is also referred to as melting. It is logical that the more hydrogen bonds DNA contains, the higher the temperature needed to denature it. It follows that since a G – C (guanine-cytosine) base pair has three hydrogen bonds to every two in an A – T (adenine-thymine) base pair, the higher the G – C content in a given molecule of DNA, the higher the temperature required to denature that DNA. This relationship has been found to exist (Fig. 11.29).

Thermal denaturation of DNA is reversible by slow cooling. This renaturation process is due to the affinity between the long sequence of complementary A – T and G – C base pairs.

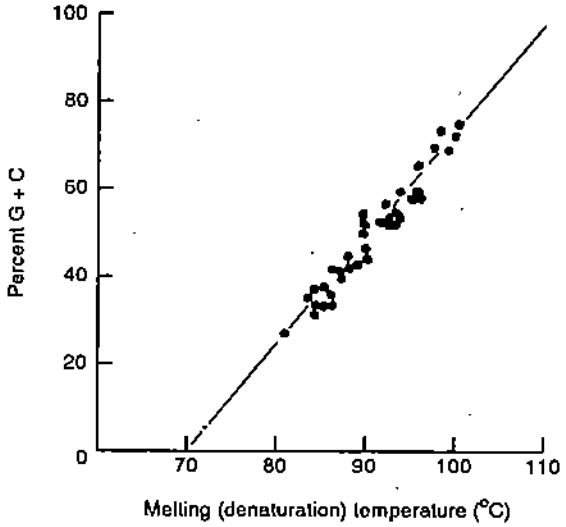


Fig. 11.29: Relationship of the number of hydrogen bonds (G–C content) and the thermal stability of DNA from different sources.

11.5 CONFORMATIONAL FLEXIBILITY OF DNA MOLECULES

The form of DNA we have described so far is called B–DNA. It is a right-handed helix. It turns in a clockwise manner when viewed down its axis. The bases are stacked almost exactly perpendicular to the main axis with about ten base pairs per turn (34Å; see Fig. 11.25). However, DNA can exist in other forms. DNA has also been described as molecule exhibiting a considerable amount of conformational flexibility.

The structures of DNA molecules changes as a function of their environment. The exact conformation of a given DNA molecule or segment of a DNA molecule will depend on the nature of the molecules with which it is interacting. When humidity is high DNA is in what is called the B-form (Watson-Crick's model). When humidity is relatively low or in dehydrated state, DNA exists in **A-form**, which has 11 nucleotides per turn. It is not yet clear to what extent A-DNA occurs under biological conditions. However, under certain conditions, short synthetic double-stranded fragments (e.g., CGG and GGTATACC along each strand) show a preference for assuming the A-configuration. A-T rich polymers show a preference for the B-form. Since A-DNA occurs under conditions of decreased hydration it has been suggested that it may be formed physiologically as a result of interaction with hydrophobic molecules or changing cellular conditions.

Recently, certain DNA sequences have been discovered by Andrew Wang, Alexander Rich and their colleagues in 1979. They exist in a unique left-handed, double helical form called Z-DNA (Z for the zigzagged path of sugar-phosphate backbones of the structure, see Fig. 11.30). The Z-DNA helix contains 12 base pairs per turn. The helices of A- and B-form DNA are wound in a right-handed manner. Moreover, specific segments of DNA molecules can undergo conformational shifts from the B-form to the Z-form and vice-versa. It has been found that certain regulatory proteins may bind only to the Z-form (or B-form) of a DNA sequence and cause it to shift to the B-form (or Z-form). There is some speculation that Z-DNA may be important in the regulatory aspects of gene expression.

SAQ 8

a) If the nucleotide sequence of one polynucleotide of a DNA double helix is 5'-ATAGCAATGCAA-3', what will be the sequence of the complementary polynucleotide?

.....

.....

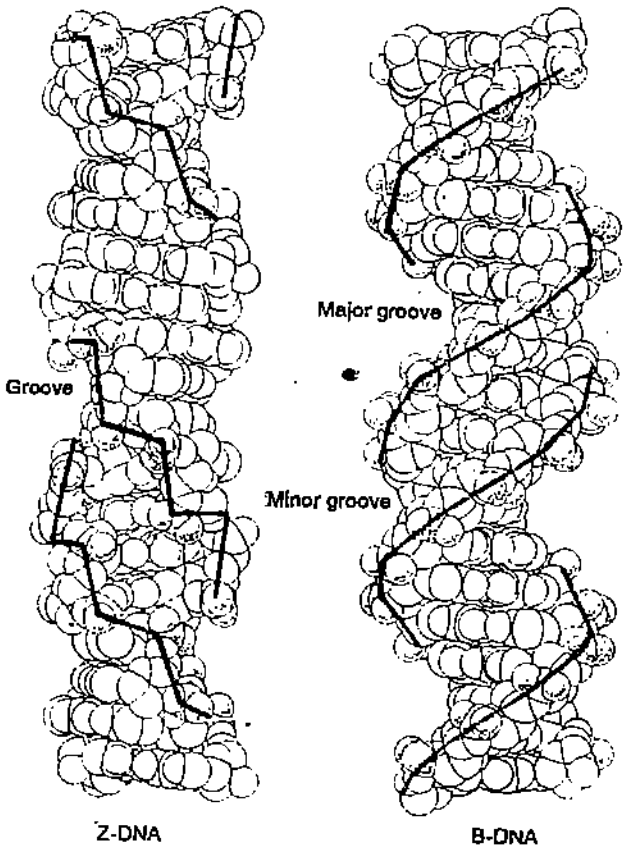


Fig. 11.30 Molecular models of Z-DNA and B-DNA. The irregularity of the Z- DNA backbone is illustrated by the dark lines that go from phosphate to phosphate residue along the chain. The groove in Z-DNA is quite deep and it extends to the axis of the double helix. The B-DNA on the other hand has a smooth line connecting the phosphate groups and two grooves.

b) DNA from the fungus *Neurospora crassa* has an AT content of 46%. What is the GC content?

.....

.....

.....

.....

.....

c) Is an adenine-thymine or guanine-cytosine base pair harder to break apart? Explain.

.....

.....

.....

.....

.....

d) For double-stranded DNA, which of the following base ratios always equals 1?

i) (A + T) / (G + C)

ii) (A + G) / (C + T)

iii) C / G

iv) (G + T) / (A + C)

v) A / G

.....

.....

SAQ 9

A double-stranded DNA molecule is 100,000 base pairs (100 kilobases) long.

a) How many nucleotides does it contain?

.....

.....

b) How many complete turns are there in the molecule?

.....

.....

c) How long is the DNA molecule?

.....

.....

SAQ 10

a) In what sense are the two strands of DNA antiparallel?

.....

.....

.....

.....

.....

b) Compare and contrast between B-DNA and Z-DNA.

.....

.....

.....

.....

.....

11.6 SUMMARY

In this unit you have learnt the following:

- Proteins fulfil many important functions for the living organism. In particular, proteins that are enzymes catalyse the biochemical reactions essential for life. Proteins consist of amino acids joined together by peptide bonds into polypeptides. Either one polypeptide chain or several chains bound together may form a protein. Proteins were once thought to be the hereditary molecules because they are long molecules of variable amino acids sequence, and are, therefore, capable of encoding immense amounts of information.

- The main feature that the geneticists were looking for in the genetic material was that it must be able to control the phenotype of a cell or organism (direct protein synthesis), replicate, and must be located in the chromosomes. Avery and his colleagues demonstrated that DNA was the genetic material when they showed that the transforming agent was DNA. Griffith had originally demonstrated the phenomenon of transformation of *Pneumococcus* in mice. Hershey and Chase demonstrated that it was the DNA of bacteriophage T2 that entered the bacterial cell. Fraenkel-Conrat and Singer showed that in viruses without DNA (RNA viruses) such as tobacco mosaic virus, the RNA acted as the genetic material. Thus by 1953, the evidence was strongly supportive of nucleic acids (DNA, or in its absence, RNA) as the genetic material.

- Chargaff showed a 1:1 relationship of adenine (A) to thymine (T) and cytosine (C) to guanine (G) in DNA. Wilkins, Franklin and their colleagues showed by X-ray crystallography, that DNA was a helix of specific dimensions. Following these lines of evidence, Watson and Crick in 1953 suggested the double-helical model of the structure of DNA. In their model, DNA is made up of two strands, running in opposite directions, with sugar-phosphate backbones and bases facing inward. Bases from the two strands form hydrogen bonds with each other with the restriction that only A and T and only G and C can pair. This explains the quantitative relationships that Chargaff found among the bases. Melting temperatures of DNA also support their structural hypothesis because DNAs with higher G-C contents have higher melting, or denaturation temperatures; G–C base pairs have three hydrogen bonds versus only two in an A–T base pair. The model of DNA presented by Watson and Crick is that of the B–form. DNA can exist in other forms, including the Z-form, a left handed double helix that may be important in gene expression.

11.7 TERMINAL QUESTIONS

1) Explain why biologists in the early 1900s thought that genes must be made of proteins?

2) One possible explanation for Griffith's transformation results was that avirulent (R) bacteria somehow restored the dead virulent (S) cells to life. How did the experiments of Avery *et al.* eliminate this unlikely possibility?

3) Imagine that you are one of Avery's co-workers and helped conduct the transformation experiments demonstrating the hereditary function of DNA. How might you respond to the following criticisms of your conclusions?

   a) DNA is not a hereditary molecule, but an agent, like ultraviolet light, that causes mutation.

   b) Genes are not made up of DNA or proteins, but of DNA and proteins.

   c) Like painting a house, transformation changes only the physical appearance of a bacterium.

4) Would C^{14} and H^{3} have been suitable radioactive tracers for Hershey-Chase experiment?

5) a) What was the objective of the experiment carried out by Hershey and Chase?

   b) How was the objective accomplished?

   c) What is the significance of this experiment?

6) How are the carbon and nitrogen atoms of the sugars, purines and pyrimidines numbered?

7) A double-stranded DNA molecule of a newly discovered virus was found by electron microscopy to have a length of 34 μm.

   a) How many nucleotide pairs are present in one of these molecules?

   b) How many complete turns of the two polynucleotide chains are present in such a double helix?

8) a) Why was a double-helix chosen for the basic pattern of the DNA molecule?

   b) Why were hydrogen bonds placed in the model to connect the bases?

9) a) If a virus particle contains double-stranded DNA with 200,000 base pairs, how many nucleotides would be present?

   b) How many complete spirals would occur on each strand?

   c) How many atoms of phosphorus would be present?

   d) What would be the length of the DNA configuration in the virus?

10) Indicate whether each of the following statements about the structure of DNA is true or false. Each letter is used to refer to the concentration of that base in DNA.

a) A + T = G + C

b) A = G ; C = T

c) A/T = C/G

d) T/A = C/G

e) A + G = C + T

f) G/C = 1

g) A = T within each single strand.

h) Hydrogen bonding provides stability to the double helix in aqueous solutions.

i) Hydrophobic bonding provides stability to the double helix in aqueous cytoplasms.

j) When separated, the two strands of a double helix are identical.

k) Once the base sequence of one strand of a DNA double helix is known, the base sequence of the second strand can be deduced.

l) Each nucleotide pair contains two phosphate groups, two deoxyribose molecules, and two bases.

11.8 ANSWERS

Self-assessment Questions

1) Pneumococcus; transformed; Griffith; transforming substance; Avery, McCleod, McCarty; 43.

2) a) Proteins

b) DNA

3) i) a

ii) c

4) About 1/2 protein, 1/2 DNA. A single long molecule of DNA is enclosed within a complete 'coat' that is composed of many proteins.

5) DNA contains phosphorus but no sulfur; it can be labelled with P^{32}. Proteins contain sulfur but usually no phosphorus; therefore, they can be labelled with S^{35}.

6) a) Adenine, guanine, thymine, cytosine.

b) The sugar is deoxyribose.

c) A nucleoside is a base attached to a sugar.

A nucleotide with one or more phosphate groups attached.

d) Thymine is unique to DNA and uracil is unique to RNA.

7) a) One phosphate group per base in DNA. Each precursor nucleotide contains three phosphates.

b) One end terminates with a 3'-OH group, that is, free hydroxyl on the 3' carbon of the sugar and the other terminates with a 5'-P group, that is, a free phosphate group on the 5' carbon of the sugar.

c) No. Any of the four bases may be incorporated into a polynucleotide chain in any order.

8) a) 5-TTGCATTGCTAT-3'

b 54%

c) The adenine-thymine base pair is held together by two hydrogen bonds but the

guanine-cytosine base pair is held together by three hydrogen bonds. Thus, the guanine-cytosine base pair requires more energy to break it and is the harder pair to break apart.

d) ii), iii), iv)

9) a) 200,000

b) 10,000

c) 3.4×10^4 nm

10) a) They are antiparallel in the sense that at the end of a double helix one strand terminates with a 3′-OH group and the other strand terminates with a 5′-P group.

b) See Section 11.5.

Terminal Questions

1) See Subsections 11.2.1 to 11.2.3.

2) Avery *et al.* used cell-free extracts of virulent (S) cells as the transforming agent. Since no virulent cells were mixed with the avirulent ones, there was no possibility of resurrecting any virulent cells.

3) Write your view point. You may refer to Subsection 11.2.4.

4) No. Both protein and DNA contain C and H.

5) a) The objective was to demonstrate whether the genetic material was DNA or protein.

b) By labelling phosphorus – a constituent of DNA, and sulfur – a constituent of protein, in a virus, it was possible to demonstrate that only the labelled phosphorus was introduced into the host cell during the viral reproductive cycle. This was enough to produce new phages.

c) That DNA not protein is the genetic material.

6) See Fig. 11.16.

7) a) The distance between nucleotide pairs is 3.4Å or 0.34 nm. The length of the molecule is 34 m or 34×10^3 nm. Therefore the number of nucleotide pairs equals $(34 \times 10^3)/0.34 = 10^5$.

b) There are ten nucleotide pairs per turn of the helix, and so the total number of turns equals $10^5/10 = 10^4$.

8) a) A multi-stranded, spiral structure was envisaged by the X-ray diffraction patterns. A double-stranded helix with specific base-pairing nicely fits the 1:1 ratio observed for A:T and G:C in DNA.

b) Use of the known hydrogen-bonding potential of the bases provides a means of holding the two complementary strands in a stable configuration in such a double helix.

9) a) 400,000 nucleotides

b) 20,000 complete spirals

c) 400,000 phosphorous atoms

d) 680,000 Å

10) a) False

b) False

c) True

d) True

e) True

f) True

The basic genetic principles and analysis are very similar in prokaryotes and eukaryotes. However the methods used to analyze the inheritance pattern in prokaryotes are different from those used for eukaryotes.

The workhorses of prokaryote genetics are bacteria particularly, *Esclerichia coli* and their viruses. The viruses which parasitize bacteria are called bacteriophages or simply phages. Bacteria and phages have been favorite organisms to study and analyze the various genetic processes. In fact the demonstration of DNA as the genetic material, one of the important events in modern biology, paving the foundation for the microbial and molecular genetics, was made using the bacterium *Pneumococcus* in transformation experiments. Similarly the bacteriophages were the tools in pioneering works that led to the understanding of the concept of genes, complementaion and of high resolution gene mapping.

This unit will be concerned with the genetic mechanisms of bacteriophages and bacteria. More importantly this unit will lay emphasis on the processes of exchange of genetic information in bacteria and viruses.

Objectives

After reading this unit you will be able to :

- discuss various methods used in studying bacterial and bacteriophage genetics,
- appreciate that the simplest organisms like bacteriophages can also be subjected to genetic analysis,
- distinguish between complementation and recombination,
- learn that a gene which is a functional unit is not unitary but divisible into number of mutational and recombination sites,
- describe the gene transfer processes in bacteria such as transformation, conjugation and transduction,
- learn about extrachromosomal genetic entities i.e., plasmids,
- discuss that sexuality in bacteria is determined by F-plasmid.

12.2 STRUCTURE OF A BACTERIOPHAGE

A bacteriophage (Fig. 12.1) is a bacterial parasite and it can replicate only within a bacterial cell. Normally referred to as phage, a bacteriophage or the virus consists of a nucleic acid (either DNA or RNA) surrounded by a coat of protein molecules. A typical phage contains only a few different types of molecules, usually several hundred protein molecules of one to ten types depending on the complexity of the phage and one nucleic acid molecule, either DNA or RNA. The protein molecules are organized into one of the three ways.

1. **Tailed form** (Fig 12.1 a):

This is the most common form. The protein molecules form the protein shell called the coat or phage head or capsid surrounding the nucleic acid. The tail, a continuation of head, is a complex multicomponent structure often ending in tail fibers.

2. **Tailless form** (Fig. 12.1 b) :

These forms resemble the tailed form in their head structure. But they lack the tail.

3. **Filamentous form** (Fig 12.1 c):

The protein molecules of this type of phage form a tubular structure in which the nucleic acid is embedded in an extended helical form.

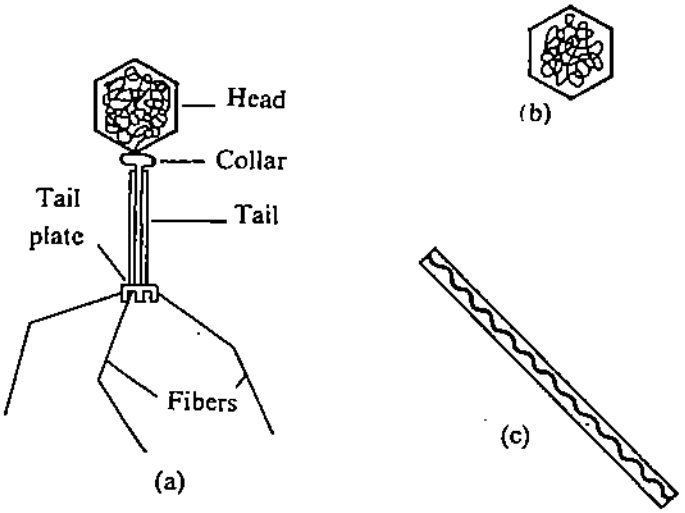


Fig. 12.1 : Three diffrent types of phage structures (a) Tailed form (b) Tailless form (c) Filamentous form.

Phages are known to contain double stranded DNA, single stranded DNA, single stranded RNA and double stranded RNA. Among these, double stranded DNA is the most common and the double stranded RNA is the least common. DNA which forms 50% of the weight of certain phages is a useful source of double stranded DNA for physical studies.

12.3 LIFE CYCLE OF PHAGES

Phage life cycles are of two types -(i) **lytic** and (ii) **lysogenic** cycles. Phages which undergo only lytic cycle are known as virulent phages. Lysogenic cycle is observed in phages with double stranded DNA. Those phages which are capable of lysogeny are known as **temperate** phages. Studies have shown that temperate phages can also undergo lytic cycle.

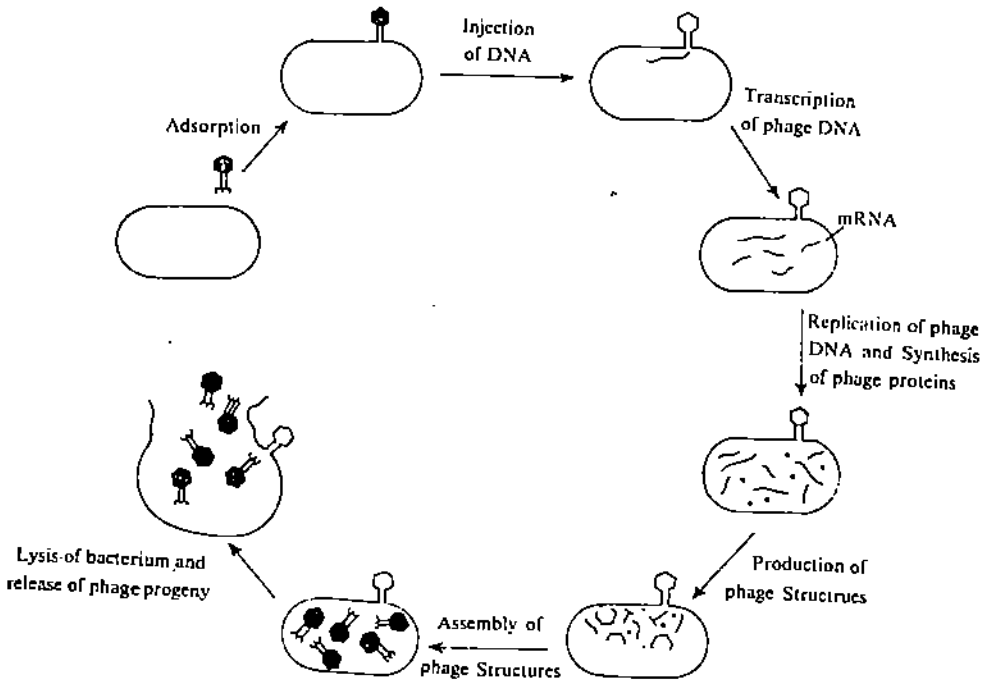


Fig. 12.2 : Lytic cycle of a bacteriophage.

12.3.1 Lytic Cycle of a Bacteriophage

A typical lytic cycle of a bacteriophage is depicted in Fig. 12.2. The lytic cycle of a phage begins when a phage particle adsorbs itself to the surface of a susceptible bacterium by means of tail fibers and injects its genetic material into the bacterial cell.

The phage genetic information then switches off the replication and transcription processes of the DNA of the host bacterium but directs the replication process of its own nucleic acid. Also the bacterial protein synthesizing machinery is used to make protein components of the phage.

The phage genetic material is then packaged into phage heads. The bacterial cell wall is disrupted releasing the phage particles numbering between 50 to 1000. This breaking open process is called **lysis**. During the propagation of phage progeny, the bacterial genetic material is degraded and the substrates consisting of sugar, phosphate and nitrogenous bases are utilized for the synthesis of phage genetic material. The infection of bacterial cell by phage particles leading to the lysis of bacterial cell releasing the phage progeny is called the lytic cycle. Thus a phage in the lytic cycle converts an infected cell to a phage producing factory, and many phage progenies are produced.

The liberated phage particles from a single bacterium infect the neighbouring bacterial cells and also induce their lysis. During the process of multiple cycles of infection originating from a single phage particle, the area of lysed bacterial cells becomes clear and visible to the naked eye. Contrarily the uninfected colony forms an opaque lawn of bacteria on the surface of a plate on solid medium. The clear areas are called **plaques**. (Fig. 12.3). A single plaque contains thousands of phage particles which are descendants of a single phage. The suspension of newly synthesized phage is called **phage lysate.**

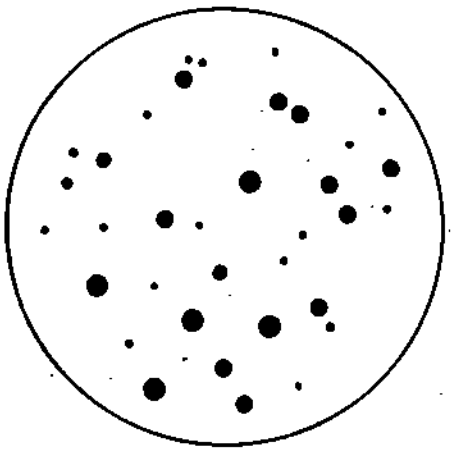


Fig. 12.3 :The appearance of phage plaques. The smaller plaques are of wild type. The larger plaques are mutants .

Phage multiply much faster than bacteria. A typical bacterium doubles in about 30 minutes, while a single phage gives rise to more than a 100 progenies in the same period. Each of these phage particles can then infect more bacteria and those released in the second cycle of infection can infect even more. Thus in two hours there are four cycles of infection for both a bacterium and a phage, yet a single bacterium becomes 2^4 that is 16 bacteria and a single phage becomes 100^4, that is 10^8, phage particles.

12.3.2 Lysogenic Cycle

The lysogenic cycle observed in temperate phages is one in which no progeny particles are produced and phage DNA gets integrated into the bacterial chromosome. As in lytic cycle, here also the phage first injects the DNA molecule into a bacterium. The phage DNA using bacterial machinery transcribes an mRNA which codes for a repressor protein and a DNA insertion enzyme. The repressor protein inhibits the synthesis of mRNA which codes for lytic functions. The DNA insertion enzyme catalyses the integration of phage DNA into bacterial chromosome. The bacterium multiplies in the usual way, and as a part of bacterial chromosome the phage genome also multiplies (Fig 12.4). The phage DNA in a lysogenic state remains more or less silent without disturbing the various processes of a bacterial cell. This process is called **lysogenization** and the phage genome present in an integrated state is called **prophage.** The prophage can be induced to undergo excision from the bacterial chromosome leading to lytic cycle.

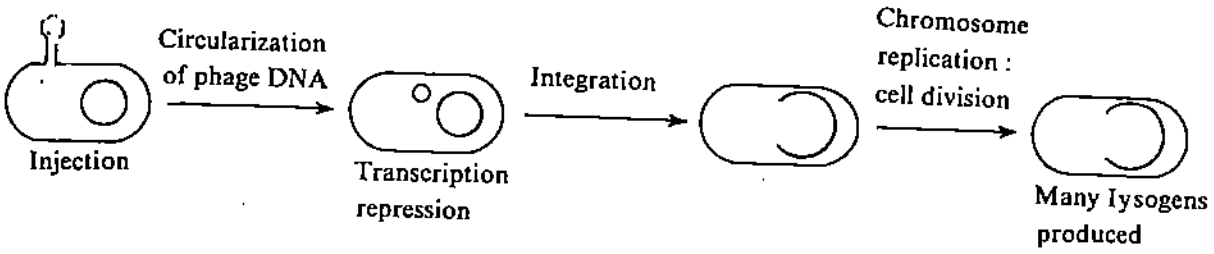


Fig. 12.4 : The lysogenization process by the insertion of phage DNA into a bacterial chromosome.

SAQ 1

Say which of the following statements is true or false. Write T or F in the box provided.

a. A bacteriophage is formed of a protein molecule surrounded by a coat of DNA or RNA. ☐

b. Filamentous phage contains a tailless head. ☐

c. Phage most commonly contain double stranded RNA. ☐

d. Phage which undergo only lytic cycle are known as virulent phages. ☐

e. Once inside the bacterium, the phage genetic information switches off its own replication process but directs the replication of host's nucleic acid. ☐

f. Lytic cycle in a phage refers to multiplication of the phage inside a bacterium and their subsequent release by the disruption of bacterial cell wall. ☐

g. Plaques refer to clear areas of a bacterial lawn caused by the lysis of bacteria by the phage particles. ☐

h. In two hours, after four cycles of infection, a single bacterium becomes 10^8 bacteria and a single phage becomes 2^4 phage particles. ☐

i. Lysogeny refers to integration of phage DNA with bacterial chromosome. ☐

j. Prophage is the state in which phage genome is integrated with bacterial chromosome. ☐

12.4 PLAQUE ASSAY TECHNIQUE

Before we study phenomena such as recombination and complementation in bacteriophage, we briefly look into a technique quite often used in phage studies known as plaque assay technique. The technique is useful for quantification of number of phage particles produced following infection under specific culture conditions. Such quantitative studies have been extensively made use of in mapping of the genes of the bacteriophage. The technique is as follows.

First a serial dilution of a phage infected bacterial culture is obtained (Fig. 12.5). To a 3 ml of melted agar containing a few drops of healthy culture, a 0.1 ml portion of one of the above mentioned serial dilutions is added. The mixture is then layered evenly over a base of solid nutrient agar, allowed to solidify and incubated. Phage plaques would appear at each place where a single phage has infected one bacterium in the culture medium plate and has grown during the incubation period. In such dilution studies three different kinds of results can be obtained.

(1) If the dilution factor is low, then, many plaques appear on the plate, all of which fuse and the bacterial lawn is completely lysed (Fig. 12.5a).

(2) If the dilution factor is increased, then plaques can be counted on the petriplate and the density of phage particles in the initial culture can be determined. (Fig. 12.5 b).

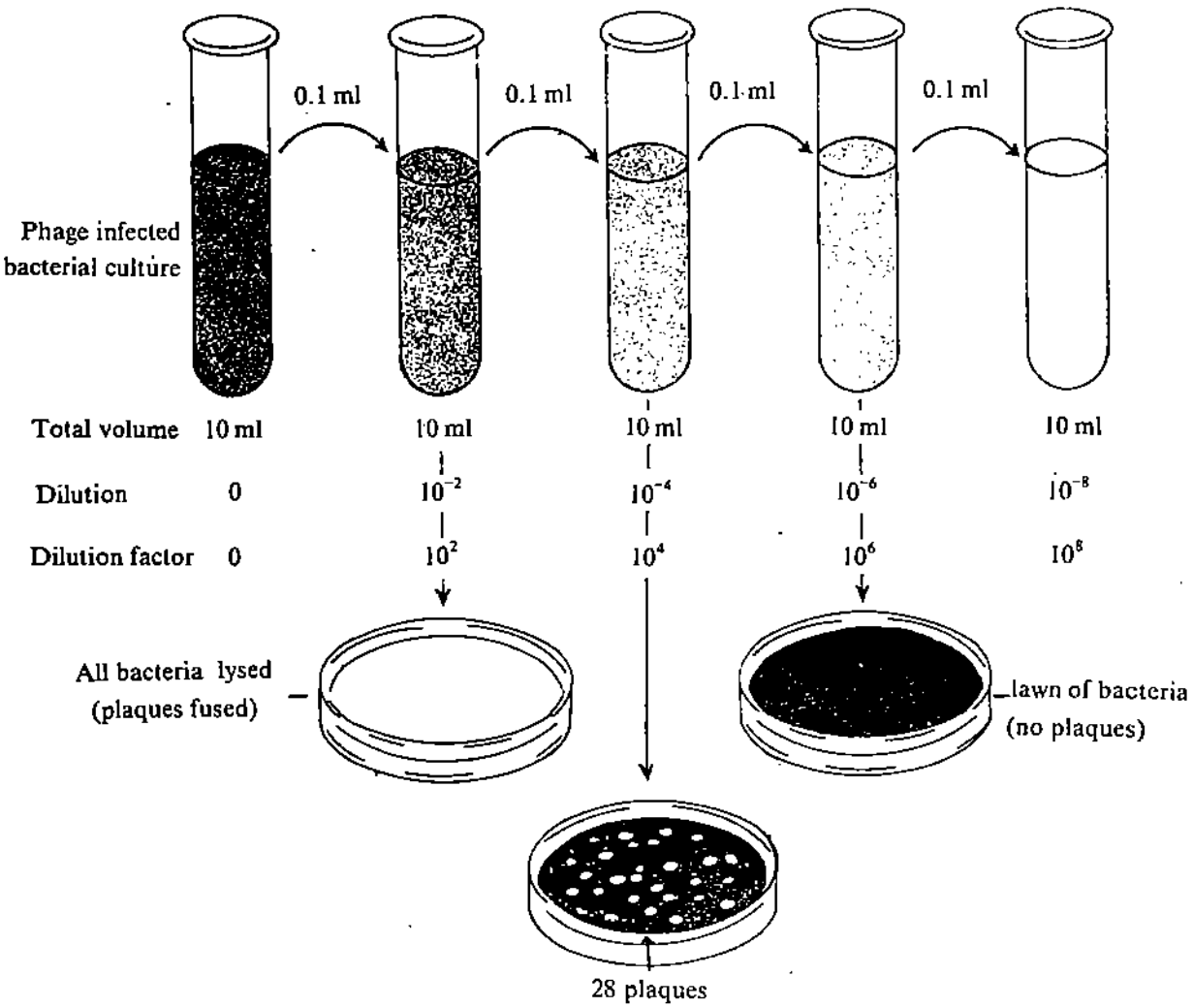


Fig. 12.5: Plaque assay technique for bacteriophage.

(3) If the dilution factor is very high, the sufficient number of phages are not available for infection and the culture plate has only a lawn of bacteria without any phage plaques (Fig. 12.5c).

In cases where plaques can be counted (2nd case) the estimation of phage numbers is done as follows. Supposing there are 28 plaques on a culture plate, then it essentially means that in 0.1 ml of the solution used for infecting, 28 phage particles were present (since each phage infects a single bacterium). In other words, there will be 280 phage particles per ml of this dilution. Assuming the dilution factor is 10^{-4}, then the undiluted sample would contain.

$$280 \text{ phages/ml} \times 10^4 = 2.8 \times 10^6/\text{ml}.$$

At 10^{-4} dilution, we said that 0.1 ml contains 28 phage particles. If the dilution were to be 10^{-6} instead of 10^{-4} then there will be only 0.28 particle per 0.1ml. There is not even a single phage particle per 0.1 ml. Therefore when the assay is made from 0.1 ml of 10^{-6} dilution no plaques are to be expected. This is the reason why there is only a bacterial lawn without any phage in it (again refer to Fig. 12.5 c).

12.5 GENETIC RECOMBINATION IN PHAGE

We discussed earlier that when a phage infects a bacterium, it injects its genetic material into the bacterial cell. We also said that on a solid surface lawn of bacteria, plaques originating from a single phage particle can be visualized. Genetic

recombination in phage was demonstrated in experiments in which two phage mutants were allowed to infect simultaneously a bacterial culture. This kind of infection is called double infection or mixed infection. Phages may differ in the spectrum of bacterial strains that can be infected and lysed. However in the inheritance pattern and recombination between phage genes, a partial diploid is to be obtained. Partial diploid refers to the diploid condition present in a part of chromosome. The creation of a partial diploid condition is the purpose of the double infection process. The infection of a bacterium by two phage particles can be considered equivalent to a phage cross. For instance in a study of the bacterium *E.coli* and phage T_2 system, two mutants of T_2 are known. One mutant causes a *rapid lysis* of the host cell and the other exhibits *extended host range* characteristic. The wild types are known to exhibit properties of *restricted host range*, that is, the host range is limited and *normal lysis* of host cells. The mutants, extended host range and rapid lysis are designated respectively as h and r and the wild types as h^+ and r^+. Let us say that two genotypes $h^+ r$ (wild type host range restriction and mutant rapid lysis) and $h r^+$ (mutant extended host range and wild type normal lysis) were allowed to infect an *E.coli* bacterium (Fig.12.6).

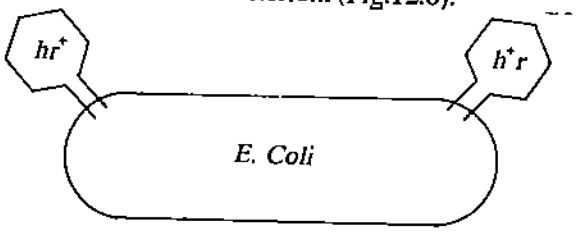


Fig. 12.6: A double infection of *E. coli* by phages.

If there is no recombination between the two phage genotypes, only the parental types could be expected as progeny. But in addition to two parental phage genotypes, that is $h^+ r$ and hr^+, two new combinations $h^+ r^+$ and hr were detected as progeny (Fig. 12.7).

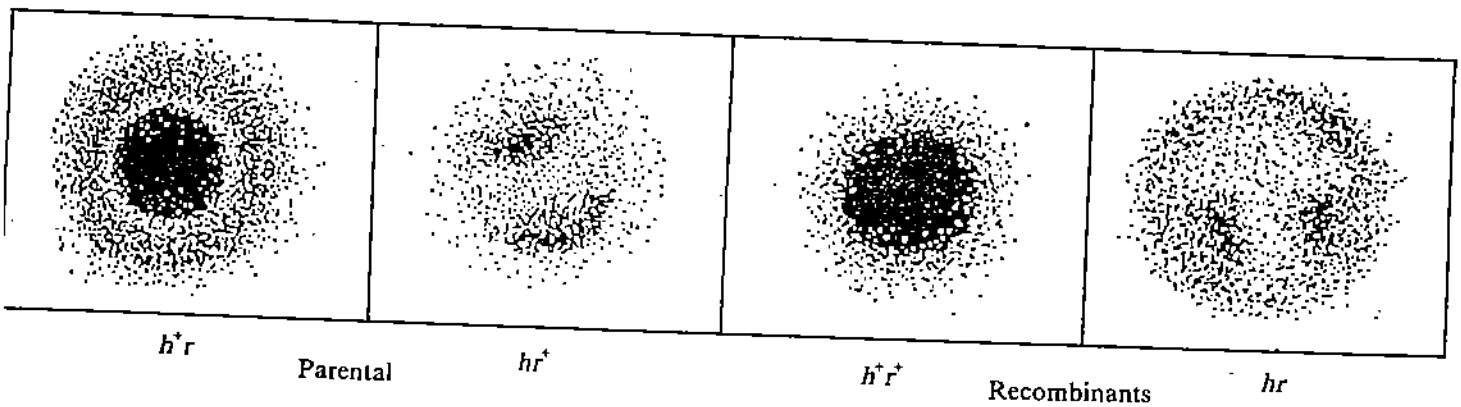


Fig. 12.7 :Plaque phenotypes obtained by the simultaneous infection of h^+r and hr^+ T_2 phages into a bacterium.

$$h^+r \times hr^+$$

$$\downarrow$$

h^+r hr^+ — parental types; h^+r^+ hr — recombinants

Evidence is available to suggest that as in eukaryotic crossing over, the recombination between phage chromosomes involves a breakage and reunion process. In the example we cited, each of the recombinant chromosomes ($h^+ r^+$ and hr) may undergo replication and are free to recombine with other recombinants as well as parental types. As the phage development occurs, the chromosomes are randomly removed from the pool and packed into phage head, producing parental types and recombinants.

Since we can not handle a single bacterium or single phage particle, a mixture of phage with a high enough ratio of phages to bacteria is used to ensure virtually that every bacterium will be infected with at least one phage of each type. Once inside the cell, DNA from each phage has an opportunity to recombine with one another generating recombinant phages, which are recovered among the progeny. The number of recombinants can be analyzed by infecting the appropriate host by the progeny

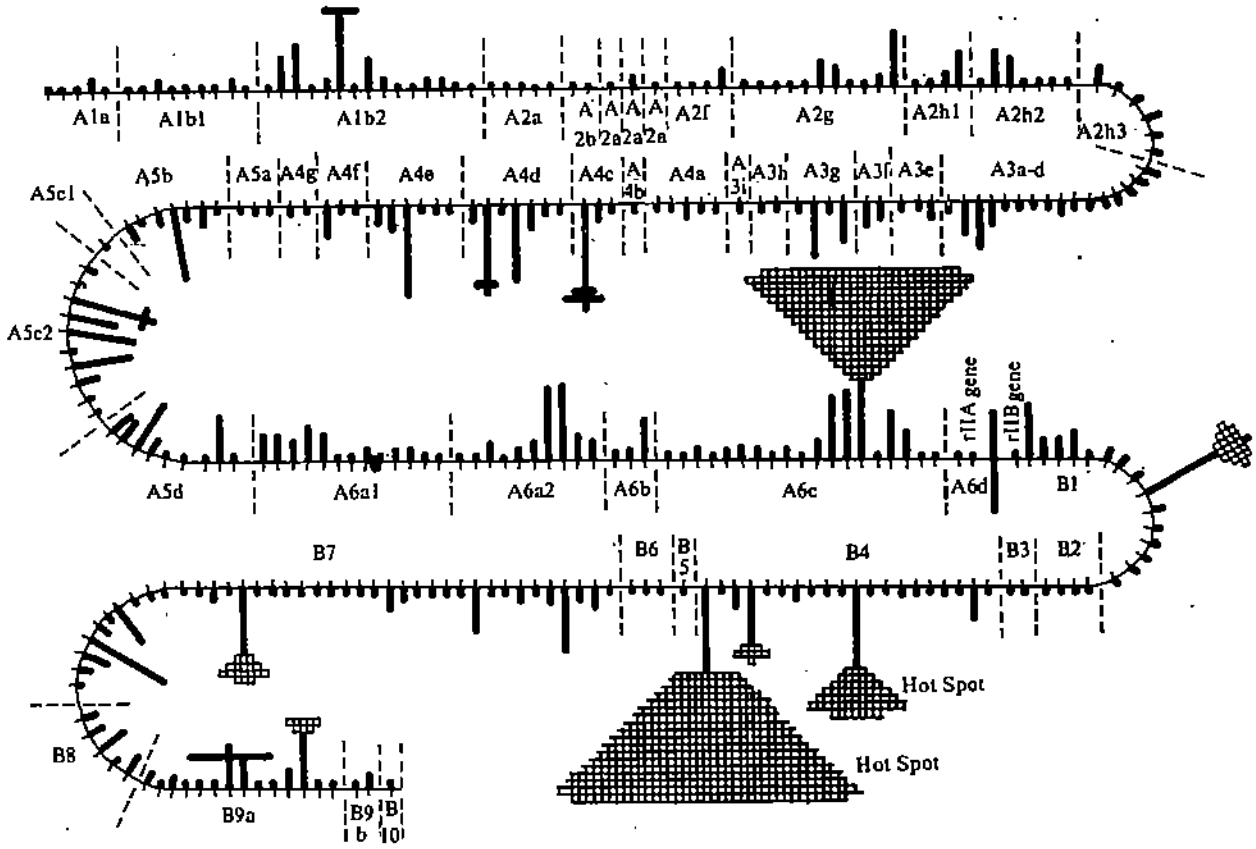


Fig. 12.9: The fine structure of rII locus of phage T_4

SAQ 3 :

(1) Match the items in column I with those in column II

Column I	Column II
a) rapid lysis phenotype	(i) egulation of length of T_4 life cycle
b) wild type rIIA & rIIB	(ii) large plaque
c) cistron	(iii) cis configuration
d) mutants rIIA & rIIB	(iv) mutant
e) $\frac{a_1\ a_2}{+\ +}$	(v) a portion of DNA molecule that codes for a polypeptide chain
f) $\frac{a_1\ +}{+\ a_2}$	(vi) trans configuration

(2) Fill in the blanks.

a) The approximate number of mutations in rIIA and rIIB cistrons is around .....

b) The distinct sites within a gene which harbour many mutations are known as .....

c) The smallest segment of DNA that can be altered to produce a mutation is known as a .....

d) The smallest unit of DNA that can undergo a recombinational event is known as .....

12.7 BACTERIAL GENETIC SYSTEM

Bacteria can be grown as pure cultures in a liquid medium or on a solid medium containing agar as long as basic nutrients are supplied. The bacteria divide by binary fission. They multiply geometrically i. e., 1 – 2 – 4 – 8 – 16 and so on, until the nutrients are exhausted or until waste products (toxic substances) accumulate to the levels that halt the population growth. A small amount of liquid culture can be transferred on to a petriplate containing solid medium and spread evenly on the surface with a sterile spreader. Each cell then reproduces by binary fission. Because the cells are immobilized on the agar surface, all the daughter cells remain together in a clump. When this mass reaches 10^7 cells, it becomes visible to naked eyes as a colony. This process is called **plating** (Fig. 12. 10). If the initially plated samples contain very few cells, then each distinct colony on the plate will be derived from a single original cell.

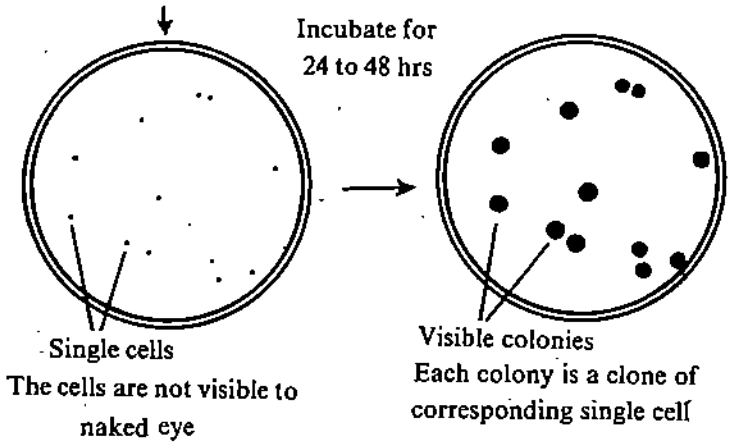


Fig 12.10: Visible colonies of bacteria . Each colony is a clone of corresponding single cell.

Members of a colony that share a single genetic ancestor are known as clones. For many characters, the phenotype of a clone can be determined readily through visual inspection or by simple chemical tests. This phenotype then can be assigned to the original cell of the clone and the frequency of the various phenotypes in the plated sample can be determined. Useful phenotypic characters are auxotrophic or nutritional mutations. **Auxotrophs** are deficient in the synthesis of a certain amino acid or a growth regulator and require for their growth the particular nutritional supplement to the minimal medium. The cells of a pure culture grown on the complete medium can be replicated exactly by the technique of replica plating onto the plates with nutritional supplement for determination of a particular mutation [see box for replica plating technique].

Replica Plating Technique :

Joshua and Esther Lederberg developed a technique for screening spontaneous mutants in cultures of bacteria. Assuming if one wants to screen for *met*$^-$ mutants (mutant bacteria requiring methionine in the medium for normal growth) the following procedure is followed.

First the cells are plated onto an agar plate containing all nutrients (complete medium) and each cell forms a colony (a). Then an adsorbant material such as filter paper or velvet is brought into contact with the agar surface (b).This results in the adherence of some cells to the adsorbant (c). The paper or velvet is now pressed into fresh media, one the complete medium (d) and the other medium lacking methionine (e). The bacteria are allowed to multiply in these plates and the comparison of plates allows the identification of *met*$^-$ mutants (Fig. 12.11).

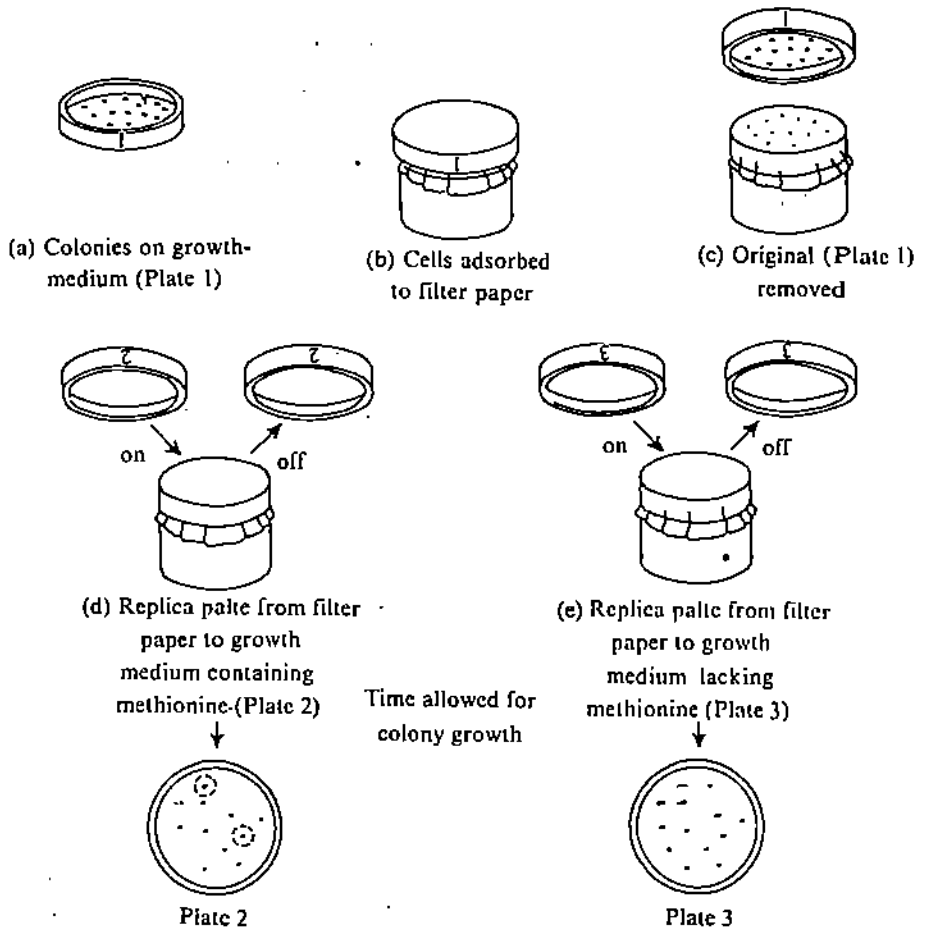


Fig. 12.11: Replica plating technique

12.8 RECOMBINATION IN BACTERIA

Let us recall that the genetic analysis in the case of eukaryotes involves crosses and examination of the progeny with respect to their phenotypes and genotypes. The offspring result as a consequence of gametic fusion, and the formation of gametes involves chromosomal pairing, crossing over and exchange of genetic material (recombination). In the process of recombination, the two sets of chromosomal complements, one paternal and the other maternal come together. In prokaryotes which are haploids , two different kinds of genetic material have to be brought together to produce a partial diploid in order to analyze the recombinants. Prokaryotes contain a single DNA duplex as their genetic material which unlike that of eukaryotes is not associated with histones.

Three processes in bacteria can transfer the genetic material from one cell to another : (i) **Transformation**, (ii) **Conjugation** and (iii) **Transduction**. Let us look into the details of these genetic exchange processes.

12.9 TRANSFORMATION

The **transformation** is a process in which conversion of one genotype into another occurs when a recipient cell acquires genes from free DNA molecules in the surrounding medium. The discovery of transformation process in bacteria laid the foundation for studies on microbial and molecular genetics and led to the demonstration of DNA as the genetic material. Briefly, transformation can be explained as follows.

If an auxotrophic mutant of *E.coli* for example *met* $^-$ (methionine-requiring, i.e., minimal medium has to be supplemented with methionine for the bacterial growth) is incubated with the DNA isolated from a *met* $^+$ strain (the strain can synthesize methionine) *met* $^+$ recipient cells can be obtained. The conversion of *met* $^-$ recipient cells into *met* $^+$ in such a case would have resulted from transformation.

12.9.1 Griffith's Experiment

Fred Griffith in 1928 first demonstrated the transformation phenomenon in the bacterium *Pneumococcus* now known as *Staphylococcus pneumoniae* which causes pneumonia in mammals. There are two strains of this bacterium. Cells of one strain known as **smooth** strain (S strain) have a polysaccharide coat around them and hence protected from the immune system of the infected mammal. Such strains cause severe pneumonia and the organism usually dies. A mutant strain of the bacterium lacks the enzyme to synthesize the polysaccharide and therefore the cells have a rough surface. Such a strain is known as **rough** strain (R strain). S and R phenotypes are controlled by two alleles of a single gene. Griffith showed that mice injected with S strain died of pneumonia, but those injected with R strain survived. Heat killed S strain does not cause the death of the organism. Mice when injected with a mixture of a small number of live R strain and a large number of heat killed S strain were dead.

Obviously the dead S cells in some way could change the R cells to survive the defense reaction of the host, so that the latter multiplied and caused pneumonia in mice. The blood of the dead mice when analyzed contained only S-strains of the bacterium and the culture of these changed or transformed bacteria remained virulent (Fig. 12.12).

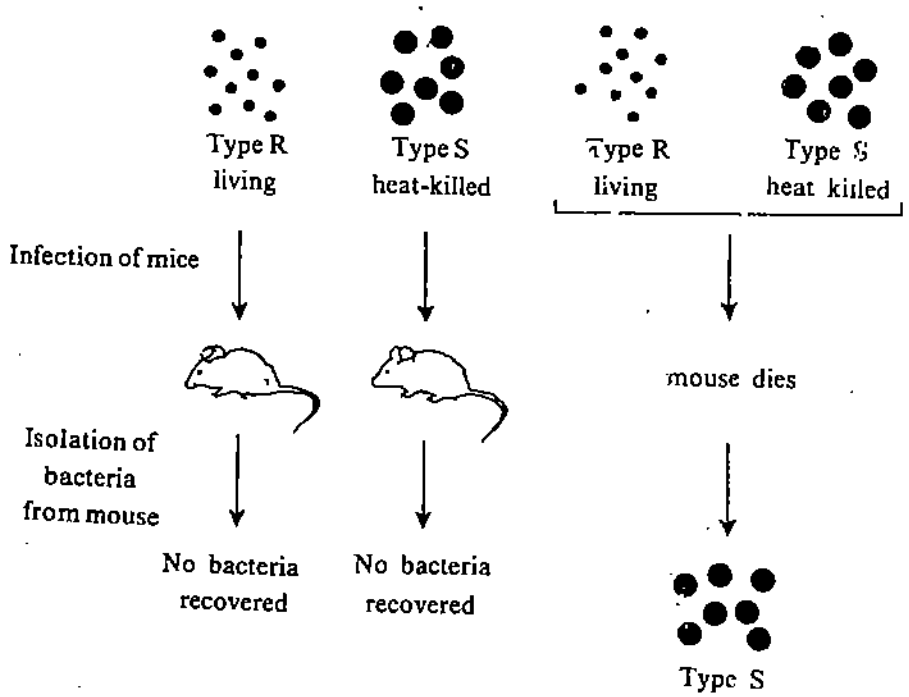


Fig. 12.12 : Griffith's transformation experiment.

The question is what confers the virulence to the non- virulent strain. The answer was provided by Oswald Avery, Colin Macleod and Maclyn McCarthy in 1949 through a critical experiment. They found that addition of minute amount of purified DNA of S cells to the cultures of R cells resulted in the transformation of R strain into S strain with polysaccharide coat around them. They further showed that the transforming activity was

completely disrupted if DNA were treated previously with DNAase, whereas treatment of DNA with RNAase or proteinase did not alter the transforming activity (Fig. 12.13).

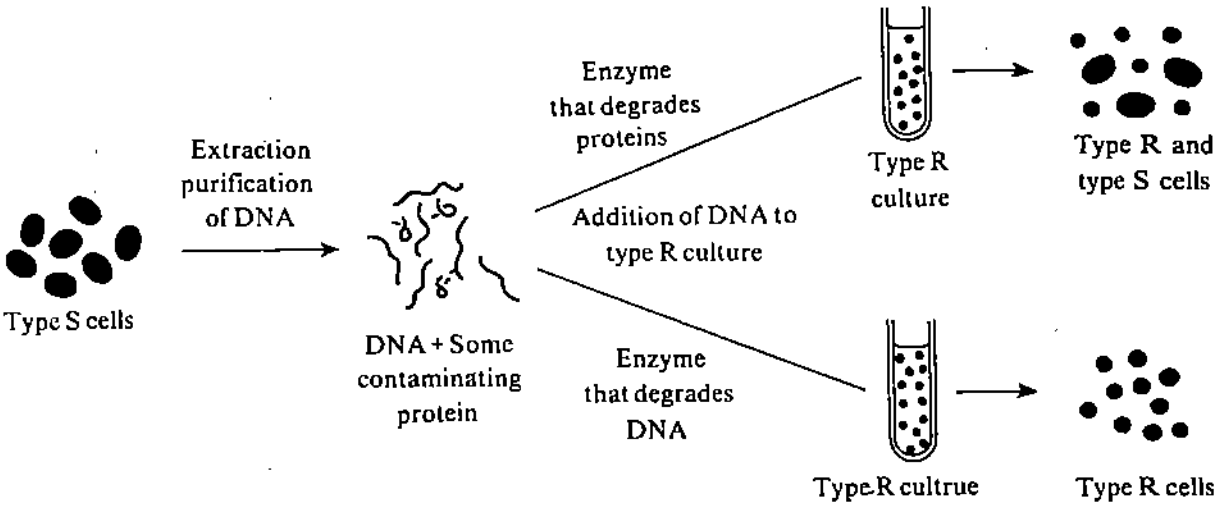


Fig. 12.13: Illustration to show that DNA is the active material involved in bacterial transformation.

Such and similar other experiments not only confirmed Griffith's findings or the occurrence of the transformation but showed that the substance responsible for genetic information was DNA, in this case the DNA from donor cells. Let us now discuss the mechanism of transformation.

12.9.2 Molecular Mechanism for Transformation

In the previous subsection we discussed transformation phenomenon in *Pneumococcus*. *Pneumococcus* is one bacterium where transformation process is fairly well understood.

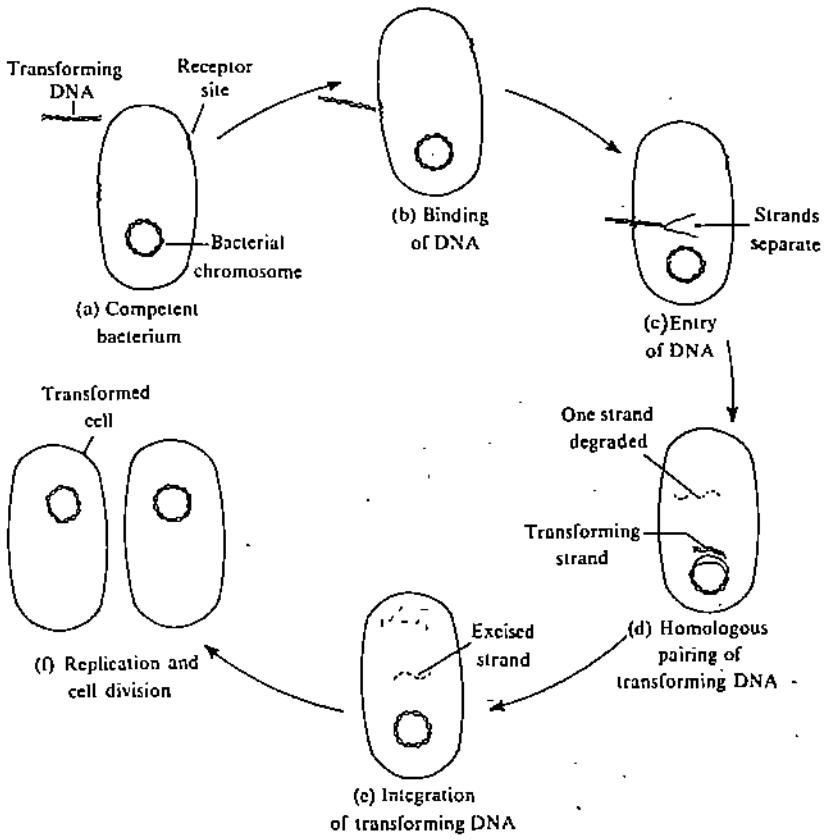


Fig. 12.14 : Proposed mechanism of bacterial transformation.

It is known that during transformation process, of the two strands of a double stranded DNA that enters into the bacterium , one strand is digested by the nuclease that is contained in the bacterial cell membrane. The other single strand in some way is protected from the nuclease attack. The single DNA strand, once inside the cell, causes the local unwinding of the recipient DNA and invades it (Fig .12.14).

Such an action is probably due to a protein, very similar to Rec A protein of E. coli (you may refer to Unit 6 of the Block 1 of Genetics course-Linkage and Crossing over to know more about Rec A protein) which brings about pairing of DNA during recombination. The invading single strand is cut by certain unknown mechanism and is base paired to the recipient strand. The action of nucleases removes the free ends of both donor and recepient strands and finally the DNA ligase seals the break. There might be some mismatching of the base pairs at the end of the recombination process, but the mismatch repair mechanism corrects the error in base pairing (Fig. 12.15).

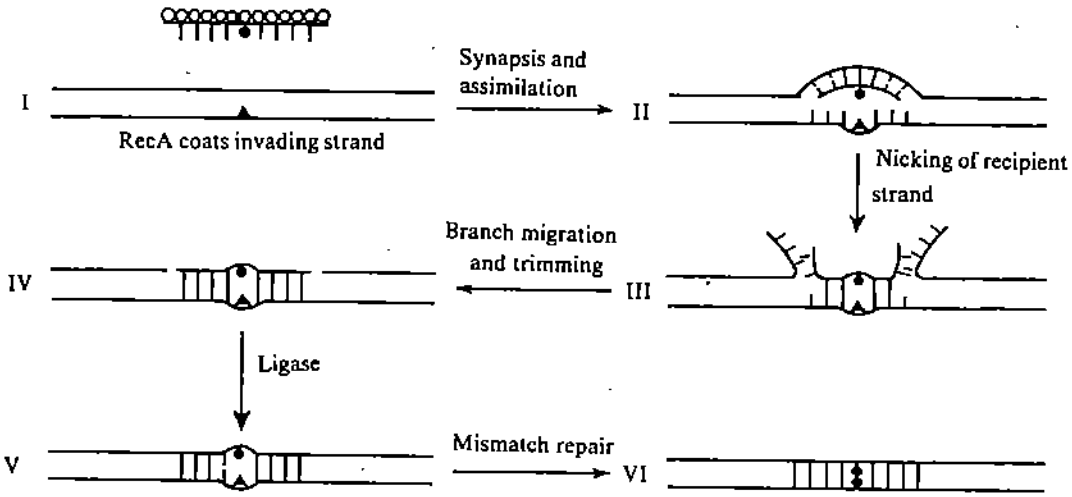


Fig. 12.15: Molecular mechanism for transformation.

12.9.3 Competence

Many bacterial genera such as *Bacillus*, *Salmonella*, *Diplococcus*, *Hemophilus*, *Pneumococcus* and *Neisseria* are naturally amenable to transformation i.e., uptake of DNA. However , this DNA uptake process occurs during a particular stage, typically during the late log phase of the cell cycle. The cells during this state are said to be **competent.** For example in *Pneumococcus*, competence is generated just as the cells are entering the stationary phase and the competence stage lasts only for a few minutes (Fig. 12. 16).

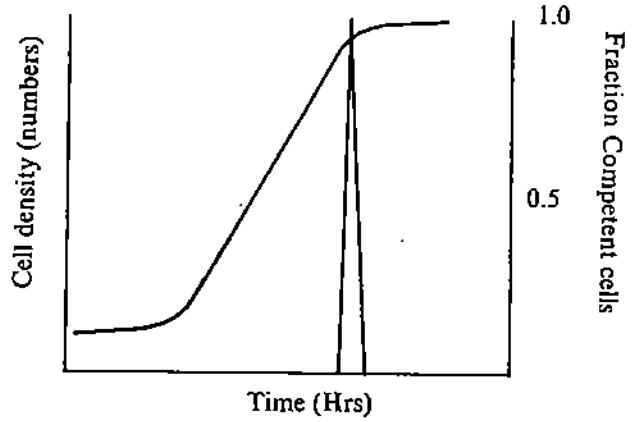


Fig. 12.16: The period in which *Pneumococcus* cells exhibit competence for transformation.

The development of competence is a consequence of synthesis of a specific protein (competence factor) at the cell surface. The protein facilitates the entry of DNA into the cell. The bacterial cells such as *E.coli* are not naturally amenable to transformation. However , their cell wall/membrane can be made permeable artificially by certain treatments to take up DNA. In fact this process for *E.coli* transformation is extensively used in recombinant DNA experiments.

12.9.4 Transformation and Mapping of Genes

Transformation can only be used to derive some information on the relative distance of genes or linkage. When the bacterial DNA is extracted for transformation experiments, some breakage into smaller pieces always occurs. The frequency of transformation is monitored by the conversion of auxotroph to a prototroph using DNA isolated from a prototroph. Assuming we are monitoring two markers let us say, a^+, b^+ and the recipient cell is a^-, b^-, there are two ways by which recipient cell can become a^+, b^+.

(i) The cell takes up a single DNA molecule carrying both the markers or (ii) it has taken up two DNA fragments each carrying a single marker. The second phenomenon is known as **co-transformation.** In the first case the frequency of single i.e., a^+ or b^+ or double i.e., a^+ b^+ transformants will be same since a single molecule carries both the markers. In the latter case, the frequency of double transformant will be much lower than the single transformants, since the probability of uptake of single DNA molecule is higher than the uptake of two molecules simultaneously. If the two markers are close to each other, they have a higher probability to be included within a single DNA molecule. If they are located far apart, they more often would be located on separate molecules in the transforming DNA. Thus the process of transformation can only be used for assessing the linkage relationship between two genes. In fact this analysis, using transformation, depends on many factors and therefore it is not normally used for analysis of bacterial genes now a days. However, transformation is a routine technique in recombinant DNA research.

SAQ 4

Pick out the correct answer from the alternatives provided.

a) Members of a colony that share a single genetic ancestor are known as plaque/clone.

b) In prokaryotes two different kinds of genetic material have to be brought together to produce complete/partial diploid to analyse recombinants.

c) Bacteria requiring nutritional supplements in their culture medium are prototrophs/auxotrophs.

d) The transformation phenomenon was first demonstrated by Griffith/Avery and others.

e) The pairing of DNA during recombination is brought about by RecA protein/DNA ligase.

f) Competence/Transformation refers to a pheneomenon in which during the particular stage of the cell cycle the bacteria are amenable to the uptake of DNA.

g) Transformation can be used to derive information on linkage/to precisely map the genes.

h) Cotransformation refers to uptake of single DNA molecule carrying both the markers/two DNA fragments each carrying a single marker.

12.10 CONJUGATION

The process by which the genetic material from one bacterium is transformed to and recombined with that of another when they are in contact with each other is called **conjugation.**

In the process one cell acts as a donor and the other as a recipient, and recombinants could be scored in the recipient cells. We have already seen earlier that the transfer of genetic markers also occurs during the process of transformation. Therefore, to understand the mechanism of the process of conjugation, a number of question were asked.

(i) Whether the resulting recombinants could be accounted for by back mutations?

(ii) Whether it is necessary to have the cell to cell contact in this process of gene transfer?

(iii) Is there any specificity in the direction of gene transfer?

(iv) What is the nature of so-called sexuality in bacteria?

(v) How could this process of gene transfer be used in mapping bacterial genes?

The answer to the first question was obtained by Joshua Lederberg and Edward Tatum in 1946, who studied two auxotroph strains of *E. coli* with different nutritional requirements. Strain A would grow on a minimal medium only if the medium was supplemented with methionine and biotin; strain B would grow on a minimal medium supplemented with threonine, leucine, and thiamine. Thus we can designate strain A as *met⁻ bio⁻ thr⁺ leu⁺ thi⁺* and strain B as *met⁺ bio⁺ thr⁻ leu⁻ thi⁻*. In their experiments, Lederberg and Tatum plated bacteria on plates containing only strain A bacteria, some with only strain B bacteria, and some with a mixture of strain A and B bacteria which had been incubated together for several hours in a liquid medium containing all of the supplements (Fig. 12.17).

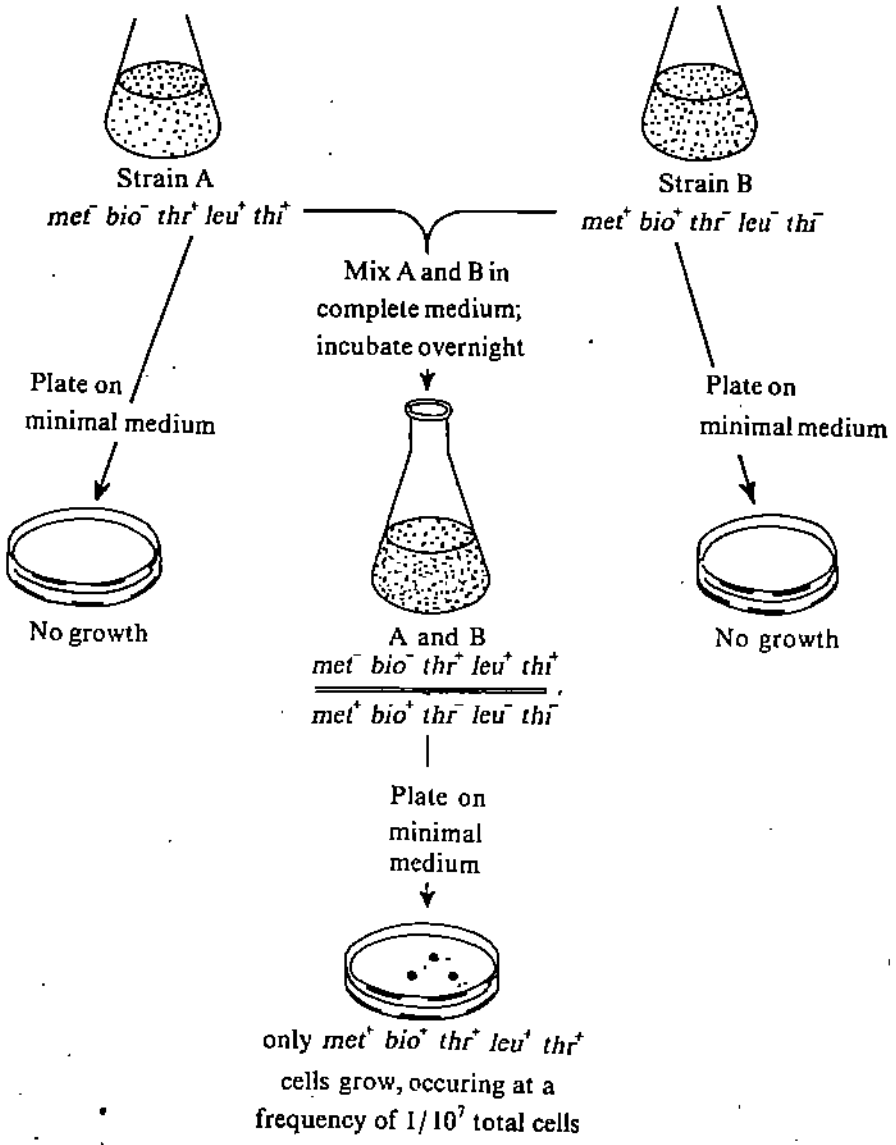


Fig 12.17: An experiment to demonstrate recombination in bacteria during conjugation.

The plates which received the mixture of strain A and strain B produced growing colonies at a frequency of one in every 10 million cells plated (say 1×10^{-7}). And no colonies arose on the plates containing either strain A or strain B alone. Remember only prototrophic (wild type) *met⁺ bio⁺ thr⁺ leu⁺ thi⁺* bacteria can grow on the minimal medium. Two kinds of inferences can be drawn from the above experiments. You may observe from fig 12.16 that no colonies arose on minimal medium when strain A and strain B bacteria were plated separately. Therefore, the origin of colonies on the plate receiving the mixture of two strains

can not be accounted for by back mutation. These observations suggest that some form of recombination of genes has occurred between the two strains to produce wild type colonies.

It could reasonably be argued that the cells of the two strains do not really exchange genes but leak substances, which the other cells can take up and use for growing. This possibility was ruled out by the classic U-tube experiment. A U- tube was designed in which the two arms were separated by a fine filter. The pores of the filter were too small to allow bacteria to pass through but large enough to allow the free passage of the liquid medium including all the dissolved substances. The strain A was kept in one arm and the strain B in the other (Fig. 12.18). The cells were incubated for some time, the liquid of the two arms was intermixed by applying pressure or suction. When the contents of the two arms were tested for growth on minimal medium, no colonies appeared.

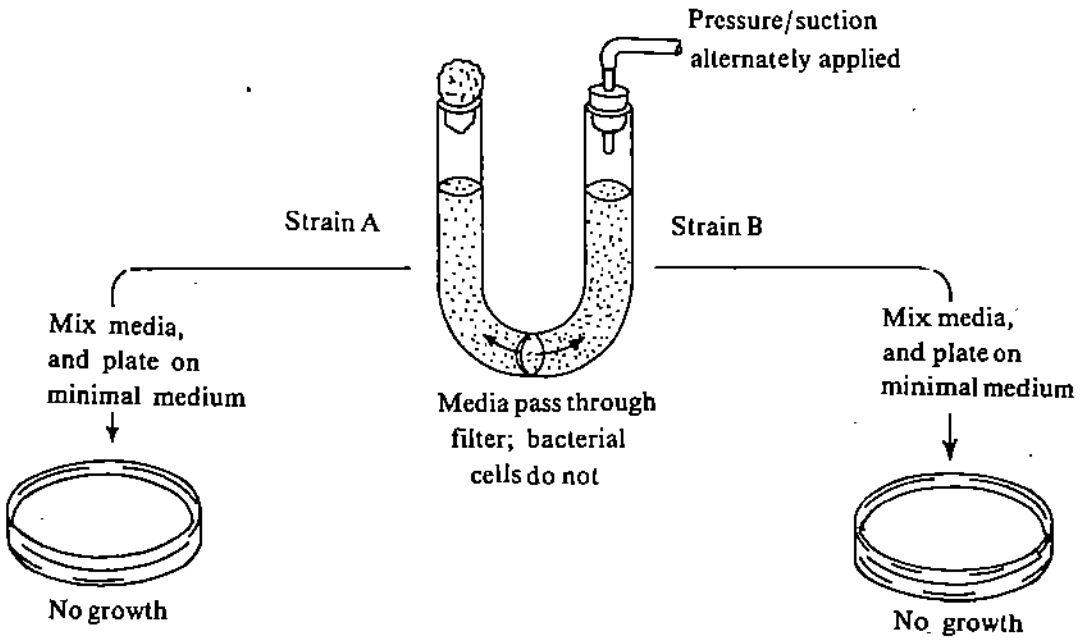


Fig .12. 18: Experiment to demonstrate that physical contact is required for transfer of genetic material during conjugation.

It is apparent from this experiment that the physical contact between the two strains is very essential for the wild type cells to form. Such a physical contact between two bacteria is mediated through structures called **pili** (Fig. 12. 19). The pili are microscopic extension of bacterial cell. After the establishment of physical contact between two bacteria through pili, a conjugation tube is formed between mating pairs. It is believed by some workers that a single pilus may give rise to conjugation canal.

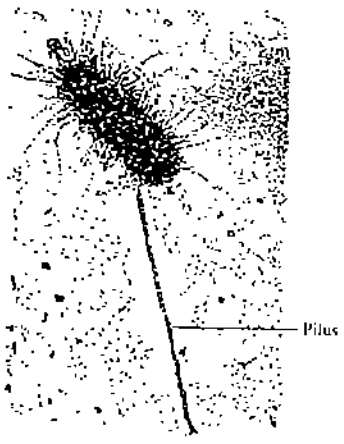


Fig. 12.19: Bacterium showing the appendage pilus. The pilus could act as a conjugation canal.

The next question one would like to ask is, when there is a physical contact between two bacteria during conjugation, is there an exchange of genetic material ? Or does the movement of genetic information occur only in one direction?

12.10. 1 Direction of Transfer of Genetic Information

By way of answering the question asked in the last section, William Hays in 1952 performed experiments with auxotrophs and antibiotic streptomycin and demonstrated that transfer of genetic material is unidirectional during conjugation. Streptomycin prevents cell division and subsequently kills the sensitive bacteria. The two E. coli strains mentioned earlier were constructed to have either *str*^{s} or *str*^{r} gene. The *str*^{r} strains can be grown in the presence of streptomycin but the *str*^{s} strains do not survive in medium containing streptomycin. The following two kinds of combinations were plated on the minimal medium plates to check for the colonies of recombinants.

i)

Strain A: $str^S\ met^-\ bio^-\ thr^+\ leu^+\ thi^+$ x Strain B: $str^r\ met^+\ bio^+\ thr^-\ leu^-\ thi^-$

↓

Plated on minimal medium containing streptomycin

↓

Colonies

$str^r\ met^+\ bio^+\ thr^+\ leu^+\ thi^+$

ii)

Strain A: $str^r\ met^-\ bio^-\ thr^+\ leu^+\ thi^+$ x Strain B: $str^S\ met^+\ bio^+\ thr^-\ leu^-\ thi^-$

↓

Plated on minimal medium containing streptomycin

↓

No colonies

In experiment (i) strain A may not survive since it is streptomycin sensitive strain. But there is a transfer of genetic material from strain A to strain B and a prototroph progeny is formed. In experiment (ii), strain B cells die and no recombinants arise. This means there is no transfer of genetic material. The results indicate that the recombination occurs in strain B and the genetic information is transferred from strain A to B. Therefore, in conjugation the genetic transfer is not reciprocal but is only unidirectional. Once the strain A bacteria have donated the genetic material their survival is not necessary. Thus, strain A acts as the donor and the other strain acts as the recipient.

This kind of unidirectional transfer of genes was originally considered equivalent to the existence of sexuality in bacteria, with the donor being termed "male" and the recipient "female". Although the terms "male and female" in terms of bacterial genetics still persist, it should be emphasized that this type of gene transfer is not a sexual reproduction. In bacterial conjugation one organism receives genetic information from a donor and the recipient is changed by this information. In the case of above matings, it was observed that the recombinants acquire the capability to act as donors. It indicates that the fertility factor which confers on the cell the so called "maleness" is transferable to the recipient cells. It was thus suggested that this donor ability which is hereditary, is imposed by the fertility factor (F). "Females" lack F and therefore are recipients. Thus, "female" or recipient cells can be designated F^- and "males" or donors F^+

12.10.2 F^+ and F^- Factors

In the previous sub-section, it was pointed out that the ability to transfer DNA by conjugation is dependent on the presence of an entity termed the F or fertility factor. Under certain conditions, the bacterium could lose its donor ability, resulting in the conversion of a donor into a recipient cell. But the fertility can be re-established if these cells were grown along with fertile donor cells. Subsequently it was shown that F factor is a granular plasmid DNA containing approximately 6×10^4 nucleotide bases and varying number of genes. F plasmids are capable of replicating independently of the main chromosome. The genes in plasmid DNA code for fibrous pili proteins.

Whenever a single cell of F^+ type is present in a population of F^- cells, within a matter of few hours all the F^- are converted to F^+ cells. During the process of bacterial

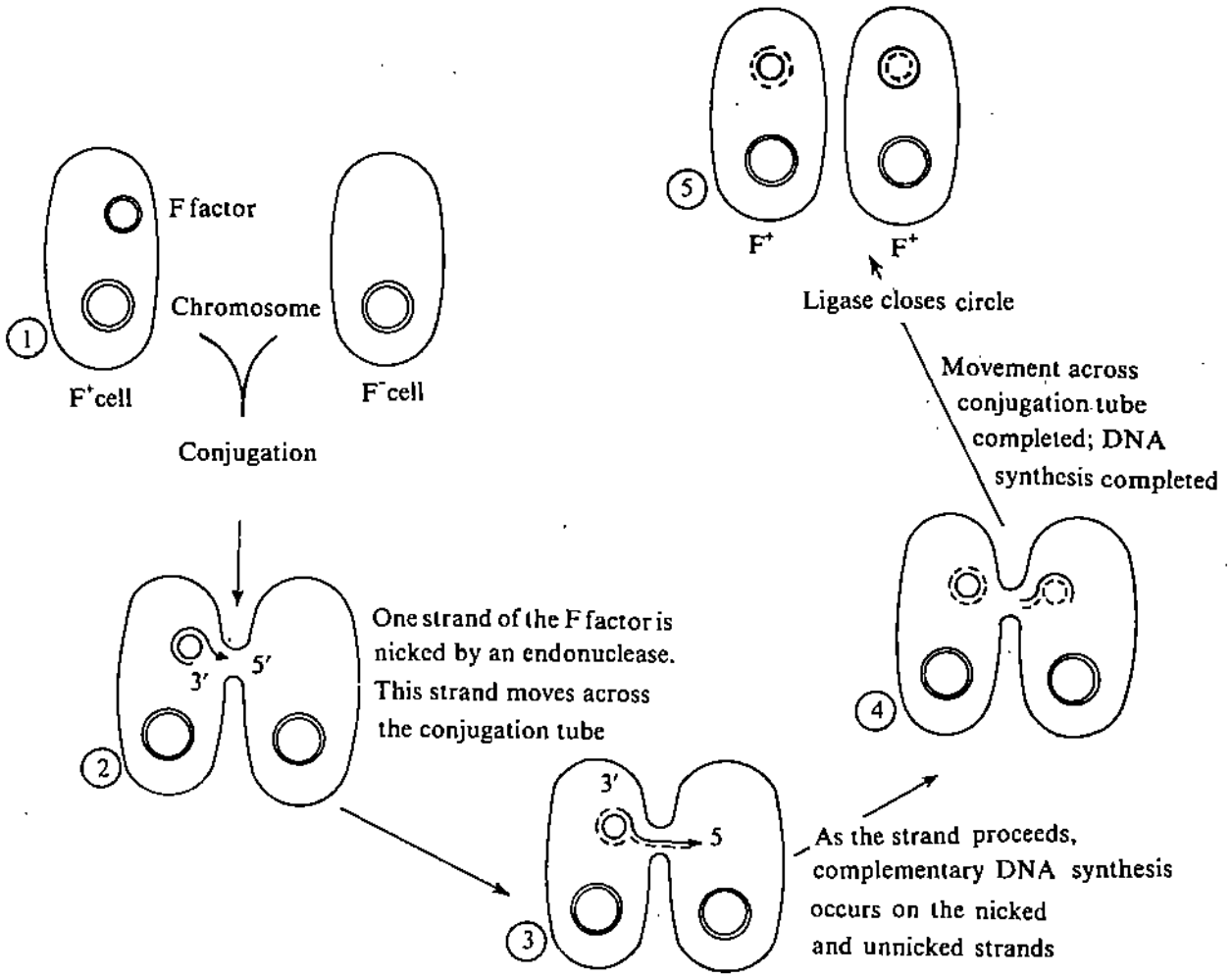


Fig. 12.20: Diagram showing the transfer of F DNA to F^- cell. One strand of F factor is cleaved and replication commences. The newly replicated F factor passes into F^- cell via the pilus. You may note that only F DNA replicates.

conjugation, after the formation of a conjugation tube between F^+ and F^- cells, a copy of the plasmid DNA moves to F^- cell converting it to F^+ (Fig. 12.20). At the end of conjugation every recipient cell becomes an F^+ cell thereby showing the transfer of F factor through the conjugation tube during each mating.

However, in only 1×10^{-7} cells there is a transfer of any part of the donor's chromosome along with the transfer of fertility factor. The transfer of a segment of a donor chromosome to the recipient is a random event and as such genes of the donor have only a random and low probability of entering into the recipient cell. The recipient bacterium appears diploid for those genes that it receives from the donor. This temporary diploid condition leads to a genetic recombination and a genetic exchange takes place between the homologous regions of recipient chromosome and donor chromosomal segment. Such an exchange resembles a double cross over event that you have studied in unit 6 of Block 1 of LSE-03. At the end of this process (Fig. 12.21) the donor genes are integrated with the recipient chromosome and the corresponding recipient fragment is degraded in the cell.

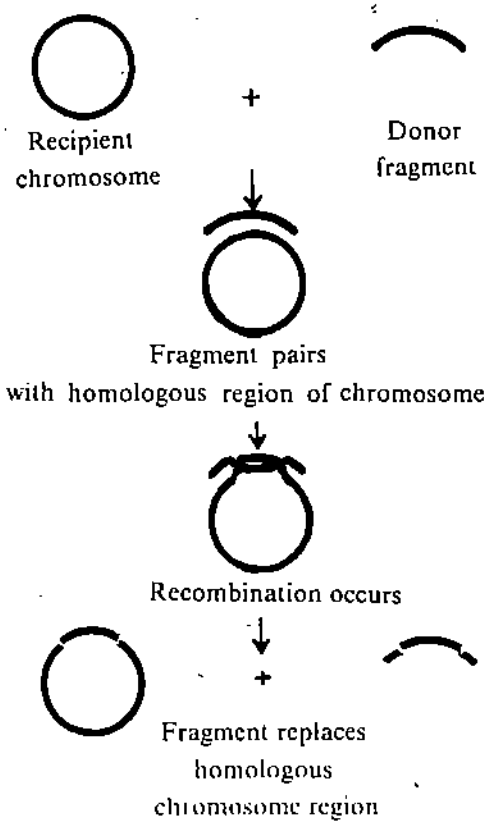


Fig. 12. 21: Recombination between donor chromosomal segment and recipient chromosome.

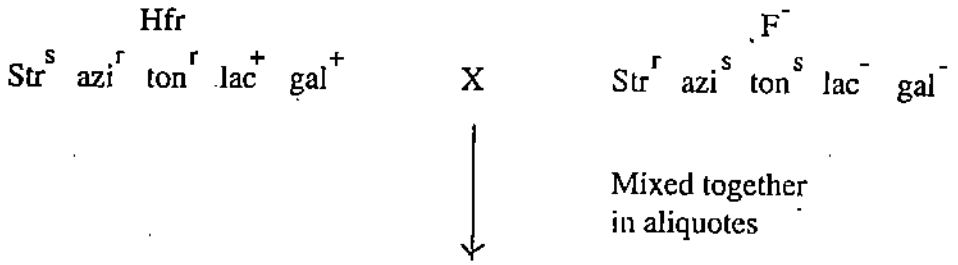
12.10.3 High Frequency Recombinant Bacteria (Hfr bacteria

We earlier said that recombination in bacteria occurs at a low frequency of 1 10^{-7} cells i.e., one at every 10 million cells. However, a F^+ strain of E.coli K12 w. s identified which could undergo recombination at a frequency of 1×10^{-4} cells i.e., 1000 times faster than the ones contained in $F^+ \times F^-$ crosses. Such a strain of bacterium was named as **Hfr strain** or **High Frequency Recombinant strain** Hfr strains behaved as donor F^+ cells; however in crosses between $Hfr \times F^-$ strains no transfer of fertility factor took place. In other words Hfr $\times F^-$ produced only F^- cells. Unlike F^+ donor strains, conjugation involving Hfr cells results in recombination of strain specific genes. And certain genes are recombined more frequently than others. One question that might interest you is why Hfr cells do not transfer the fertility factor. It was found that in various Hfr strains, the F plasmid DNA integrates into bacterial chromosome at different points. And at the time of conjugation, the circular bacterial chromosome breaks open at the point where the F factor is integrated with it. In an $Hfr \times F^-$ mating, it is the bacterial chromosome and the genes present in it are the ones to be transferred first to the F^- cell. The F factor would be the last segment to be transferred. In practice, conjugation process does not last long enough to allow the entire chromosome to enter into an F^- cell. And this explains that the recipient cells namely F^- cells when mated with Hfr cells do not acquire the fertility factor.

12.10.4 Conjugation and Genetic Mapping

In 1950, Wollman and Jacob showed that Hfr strains could be successfully used to prepare the genetic map of the E.coli chromosome. Wollman and Jacob mixed Hfr and F^- strains and analyzed the recombination of specific genes in relation to time. At various time intervals, the conjugating Hfr and F^- strains were removed from the culture and transferred to a kitchen blender for a few seconds. The shearing forces in the blender separated the conjugating bacteria and the mating is interrupted. The interruption prevented the transfer of entire chromosome and the cells were analyzed for recombination. The technique known as *interrupted mating technique* is useful in demonstrating that in a given Hfr strain certain genes were transferred and recombined slower than others.

Normally the Hfr and F^- strains used are streptomycin sensitive (str^s) and streptomycin resistant (str^r) respectively. After interrupting the matings, the cells are plated on to a medium containing streptomycin to kill the Hfr donor cells. The str^r cells are then tested for the presence of marker alleles from the donor. Let us assume the following were the genetic constitution of the Hfr and F cells.



Aliquots are withdrawn at specific time intervals and matings are interrupted. Cells are transfered onto the plates containing streptomycin. The antibiotic streptomycin kills the Hfr cells but allows the str^r, F^- cells to grow and to be tested for their ability to carry out the four metabolic steps.

Those str^s cells bearing the donor marker alleles must have been involved in conjugation; such cells are called **exconjugants**. The donor markers that appeared in the experiment are plotted in Fig .12.22.

The *azi*^{r} (azide resistance) is transferred to F^- cells in less then 10 minutes closely followed by *ton*^{r} (T-one resistance). The *lac*$^+$ is the third marker which takes about 17 minutes and *gal*$^+$ takes about 25 minutes for the transfer to the recipient cell. Thus a linkage map can be constructed for *E.coli* chromosome from the above interrupted mating studies using the time at which the donor alleles first appear after mating. The units of distance are given in minutes.

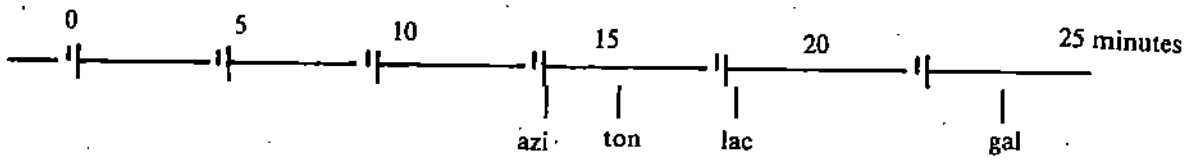


Fig. 12.22 A time map of genes obtained from the interrupted mating technique. The arrow indicates the direction of the movement of Hfr chromosome to the recipient cell.

The most striking feature of these results is that each donor allele first appears in F^- recipient at a specified time after mating begins. Further more, the donor alleles appear in a specific sequence. The maximal yield of the cells for those donor markers that enter late into the recipient is small. For interpretation of the results,all these observations are taken together. The genetic map of the bacterial chromosome can be constructed as a measure of "distance" in minutes at which the donor alleles first appear after mating. If the *lac* marker begins to enter F^- cell 7 minutes after *ton* begins to enter, then *ton* and *lac* are 7 minutes apart.

SAQ 5

Provide brief answer to the following questions.

a) Define the term conjugation.

.....

.....

.....

b) Assuming a strain of *E.coli* is $met^- bio^- thr^+ leu^+ thi^+$, what are the nutrients that it would require in the culture medium for its growth ?

c) When strain A which is pur^+, gal^- is mixed and cultured with strain B which is pur^- gal^+ the progeny were found to the pur^+ and gal^+. What inference do you draw from this experiment?

d) Does the presence of an F plasmid make a cell a donor or recipient?

e) What is the difference between an F^+ and F^- cell?

f) What is an F factor?

g) What do the genes in an F plasmid DNA code for?

h) What is the frequency at which a donor chromosome is transferred to a recipient in an $F^+ \times F^-$ mating?

i) How does an Hfr cell differ from an F^+ cell?

j) What is the frequency at which Hfr cells transfer their genetic material to the recipients?

k) How are the F^+ plasmids produced?

12.11 TRANSDUCTION

The process by which gene transfer from one bacterium (donor) to another bacterium (recipient) is mediated by bacteriophage is called **transduction**. The transfer of genes facilitates the production of partial diploid and recombination of genetic markers of the two strains. The involvement of phage particles in this transfer of genetic information was demonstrated with help of the classical U-tube experiment discussed in conjugation. Two different strains of *Salmonella typhimurium* were used. One was $phe^- trp^- tyr^- met^+ his^+$ and the other was $phe^+ trp^+ tyr^+ met^- his^-$. When either of the strains was plated on a minimal medium, no wild type cells were observed. However, by mixing the two strains, wild type cells ($phe^+ trp^+ tyr^+ met^+ his^+$) appeared at a frequency of 1 in 10^5. This seems similar to what was observed in conjugation. However when the two strains were placed in two separate arms of a U-tube in which cell to cell contact was prevented by a filter as described in Section 12.10, the recombinants (wild type) were recovered. It was hypothesized that the genes must have been transferred by a filterable agent. Since the filterable agent and F22 bacteriophage were found to be identical in size, sensitivity to antiserum and resistance to hydrolytic enzymes, it was concluded that the filterable agent was in fact a phage. Thus, in this process the transfer of genes is mediated by a virus (bacteriophage) and since it is distinct both from conjugation and transformation, it has been called *transduction.*

The process of transduction has been shown to be mediated by virulent and temperate phages. Two very different types of transduction are known; (i) **generalized and (ii) specialized or restricted** transduction.

12.11.1 Generalized Transduction

The process of mode of transfer of DNA during transduction begins with the entry of phage into lytic cycle. The resulting progeny virus infect and lyse the other bacteria. During such infections, the bacterial DNA is degraded into small segments and the viral DNA is replicated. The viral DNA is packaged into phage head. Occassionally mistakes are made in packaging DNA. Instead of phage DNA bacterial DNA of an equivalent length is packaged. Such abnormal phage particles without their own genetic material are released into the culture medium. The protein coat of such aberrant phages infect the unlysed bacterium; however, the bacterial segment does not undergo replication. Rather it remains in the cytoplasm or recombines with the homologous regions of bacterial chromosome. Assuming the recombination does not take place, the segment remains in one of the bacterial progeny cells when division occurs. This means that one of the cells is partially diploid for the transduced genes. This phenomenon is known as *abortive transduction.*

In the case of a recombination between the chromosomal segment and the bacterial chromosome, the transduced genes undergo replication as a part of the host chromosome and are passed on to the daughter cells. This phenomenon is *complete transduction.* Since during phage maturation, any part of bacterial chromosome can be packaged randomly, every portion of the bacterial chromosome has an equal probability of being transferred. Since all the genes of a donor are represented in a population of transducing particles, this type of transduction has been named *generalized transduction.* Both abortive and complete transduction fall under generalized transduction category. But at the same time you should know, that any one transducing particle contains only one segment of host DNA, representing 1/100 or 1/50 of the total donor chromosome.

The process of generalized transduction is of limited use in gene mapping, because as we pointed out earlier, the frequency of transduction particles in a phage progeny is very low. Nevertheless, information on linkage or on the relative distance between genes can be derived by analyzing the recipients for the donor markers. The linkage relationship between the genes is estimated from the frequency at which two markers are included on a single DNA fragment and being transduced together. The markers transduced together are called **cotransductants.** Linkages are usually expressed as contransduction. The frequency of

contransduction depends on the closeness of the location of genetic markers. If contransductants can not be obtained, then it provides a negative information on the linkage phenomenon, namely that the two genes are not linked. Generalized transduction is mediated by some virulent bacteriophages and by certain temperate phages whose chromosomes are not integrated at specified site in the host chromosome. As we mentioned earlier, the transducing particles are produced during the lytic cycles of these phages. Here again, certain virulent phages such as T even bacteriophages (T_2, T_4, T_6) do not mediate transduction as they degrade the host DNA and utilize the mononucleotides for the synthesis of phage DNA. The host DNA is not available for packaging into progency particles. Other phages may not produce transducing particles, since the host chromosome may be too large to be packaged into transducing particles. In still other cases, packing process will be highly specific that there may not be any error during packaging process. It could therefore be said that only a limited number of phages are known to mediate transduction. Some of the transducing phages and their hosts are phage Pi and *E.coli*, phage PBSI and SPIO of *Bacillus subtilis*.

12.11.2 Specialized or Restricted Transduction

In the last sub-section we said that in generalized transduction there is an equal probability of all bacterial genes to be transduced. But in **specialized transduction** the process is restricted to certain genes. Transduction in *E.coli* involving λ phage has often been cited as one of the good examples of specialised transduction process.

λ DNA contains a 15 nucleotide sequence known as *att* gene responsible for the integration of the phage DNA with *E.coli* chromosome during lysogeny (refer to section 12.3.2 to recall your studies on lysogeny) A homologous segment of *att* also occurs in *E. coli* between *gal* and *bio* loci. Obviously the integration of λ DNA with *E. coli* chromosome occurs at a location between these two genes (Fig. 12.23 a).

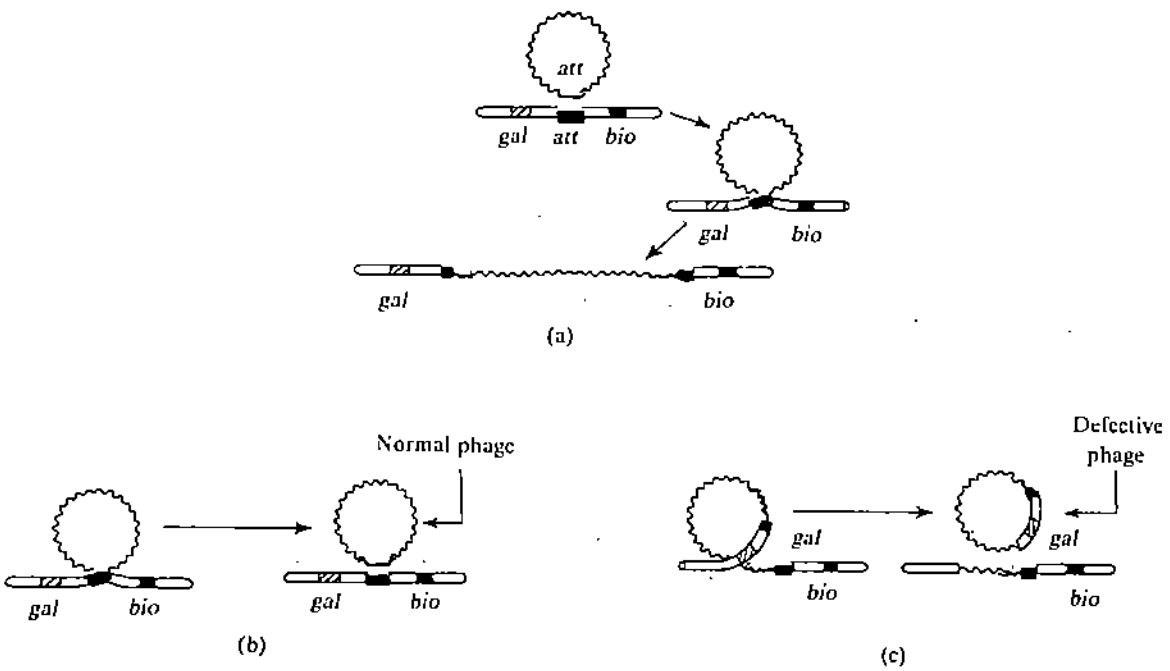


Fig. 12.23: (a) Integration of phage DNA with bacterial chromosome. (b) Normal detachment process. (c) Defective detachment process.

Once the integration of phage DNA with bacterial chromosome has occurred, the phage has entered into lysogenic cycle. The integration involves recombinational events in the presence of certain specific enzymes. In the integrated stage the phage chromosome is called a **prophage**. A lysogenic cell you may recall is immune to infection by the same strain of phage. The transition from lysogenic stage to lytic stage is rare. One out of hundred thousand cell divisions (1×10^{-5}) may show a spontaneous transition from lysogenic to lytic cycle. Such transition can be induced by UV irradiation. During such transition the excision of phage DNA from the bacterial chromosome occurs. The excision process like

the integration process is site specific. In other words, the excision occurs at the same site at which the integration occurred and the phage chromosome is obtained exactly in the same form in which it occurred prior to integration (Fig. 12.23 b). But occasionally mistakes may occur in the excision process. In other words, the excision, instead of at the original attachment site occurs at a different site (Fig. 12.23 c). Such mistakes during the prophage excision are responsible for the formation of specialized transducting particles. Only the host genes located close to the site of integration can be excised with phage DNA and packaged into phage particles. As you can see from the Fig. 12.23 c , phages can transduce only a segment restricted to bacterial genes located within a short distance on each side of the prophage attachment site. Hence this mode of transduction is known as **restricted** or **specialized transduction.**

As you could observe from the discussion on generalized and restricted transduction that the two processes are quite different. We earlier said that the generalized transduction is of limited use in gene mapping as the transduction frequency is quite low. And the specialized transduction is also not useful in studies on linkage or mapping, as the transduction process is limited to certain genes only.

SAQ 6

Fill in the blanks with suitable words.

a) Phage mediated gene transfer from one bacterium to another is known as .....

b) Instead of phage DNA if bacterial DNA of equivalent length were to be transferred to another bacterium, if the DNA is unable to recombine and remains in one of the cells as a partial diploid the condition is known as .....

c) ..... refers to recombination between a bacterial chromosome and bacterial chromosomal segment and passing of the recombination products to the daughter cells

d) In E. Coli the *att* locus is located between ..... and ..... loci

e) UV rays can induce transition in a phage from ..... to ..... cycle

f) Transduction of bacterial genes located between a short distance on each side of the prophage attachment site is known as .....

12.12 SEXDUCTION

In Section 12.10.3 you studied about Hfr strains. You may recall that Hfr strains are formed when there is an integration of F factor or plasmid to the bacterial chromosome. Experiments with Hfr strains of *E. coli* showed that the F factor could get itself excised from the bacterial chromosome and the bacterial cell reverted from its Hfr state to F^+ state (Fig. 12.24).

At the time of excision, once again because of errors in excision, the F factor could carry with it several adjacent bacterial genes along with it. Such F factors were designated as F' factors to distinguish them from F^+ factors . You may say that F' is a special type of F^+. Since F' is not packaged like a phage DNA, there are no size restrictions as to how many genes the plasmid can carry during the anomalous excision process. The extent of bacterial genes carried by F' factor varies from a single marker to a sizable portion of bacterial chromosome.

Once inside the bacterium, an F' confers on the host F^+ status. That is, it initiates conjugation with an F^- cell and as a result the F factor along with the chromosomal genes is transferred to the F^- cell. The process of transfer of genes mediated by a F' plasmid as termed **sexduction** or **F duction**. The F' factor may either be lost or be involved in the recombinational process resulting in wild type. The bacterial chromosomal genes as a part of F factor form a diploid condition in the recipient cell's chromosome. This results in a

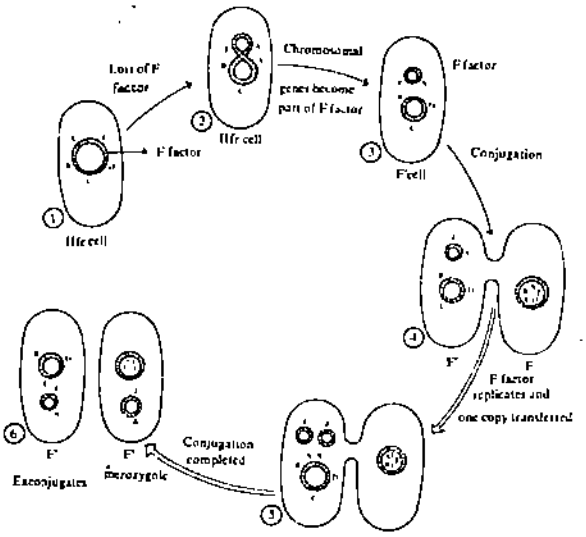


Fig. 12.24: Conversion of an Hfr bacterium to an F state.

partial diploid cell called **merozygote**. Merozygotes have been useful in studies on genetic analysis and regulation of bacterial genes.

12.13 PLASMIDS AND EPISOMES

Bacterial cells besides the main chromosome may possess from one to several copies of extra-chromosomal DNA molecules, called plasmids. By definition, a plasmid is an autonomously replicating genetic entity which may carry certain genes. They have an origin of replication through which the plasmid DNA replicates confering on it the property of autonomous repliction. They by nature are dispensable i.e., they are not required for survival of the cell in which they reside. However, whenever they are present in a cell they confer on it certain selective advantages; for example if the plasmid carries the gene for certain antibiotic resistance e.g. ampicillin resistance (*amp*^{r}), the cell carrying the plasmid would be able to grow in the presence of ampicillin.

The cells due to the presence of the plasmid acquire the capability of degrading the particular antibiotic. Under such conditions, they are essential for the cell survival. Besides the origin of replication and certain other genes, plasmid has its own genes for regulating the time of synthesis and number of copies per cell. Plasmid may be present in one or two to 50-100 copies per cell depending on the type of plasmid. The replication of low copy (1-2 copies/cell) plasmid is tightly linked to bacterial chromosomal replication; such plasmids are said to be under **stringent** control. Their replication can not be delinked from the bacterial DNA replication. The replication of higher copy number (10-100 copies per cell) plasmid can be delinked from the replicatioin of bacterial chromosome under certain conditions and copy number of such plasmids can be further enhanced. Such plasmids are said to be in **relaxed** replication.

The process of enchancement of copy number of these relaxed plasmids is termed **amplification**. The plasmids under stringent control are usually large and one of the prominent examples is F plasmid. The relaxed plasmids are usually small. By removing undesirable sequences and addition of resistance markers and other desirable features they are designed to be useful in recombinant DNA technology.

Plasmids or phage DNA which can exist both in the free state, autonomous of bacterial chromosome as well as in the integrated form as a part of bacterial chormosome are called **episomes**. The plasmid and episome are not synonyms. Many plasmids do not exist in integrated state and thus are not episomes; similarly many temperate phage chromosomes

such as phage lambda genomes, are episomes but not plasmid.

The importance of plasmids has become increasingly recognized during the last two decades. Plasmids have been identified in almost all strains of bacteria tested. They have major practical significance in two areas : (i) Spread of multiple antibiotic and drug resistance in pathogenic bacteria and (i) industrially important microorganisms that are known to carry genes for certain steps in fermentation process such as cheese making.

SAQ 7

Match the items in column I with those of column II.

COLUMN I	COLUMN II
(a) plasmid	(i) tight linkage of plasmid and bacterial chromosomal replication.
(b) relaxed replication	(ii) Plasmid as integral part of bacterial chromosome.
(c) amplification	(iii) self replicating extra chromosomal DNA molecule
(d) stringent replication	(iv) Delinking of plasmid replication from that of bacterial chromosome.
(e) episomes	(v) the process of transfer of genes mediated by F factor
(f) sexduction	(vi) enhancement of copy number of relaxed plasmid.

12.14 SUMMARY

In this unit we outlined the expansion in the knowledge of bacterial and viral genetics over the past fifty years. The structure and life cycle of bacteriophages, viruses which are parasitizing bacteria, have been discussed in detail. Phages consist of DNA or RNA surrounded by a protein coat. Two type of phage life cycles are known to occur. One, the lytic cycle in which a virulent phage enters the bacterium, replicates and transcribes at the expense of bacterial DNA and uses the bacterial protein synthesizing machinery for making its own protein. The assembly of phage particles occurs inside the bacterial cell, and at the end of this process the lysis of bacterial cell takes place, releasing the progeny particles to reinfect other bacterial cells. The other cycle, the lysogenic cycle witnessed in temperate phages consists of integration of viral DNA with the bacterial genome. Such a phage in a lysogenised state is called a prophage. A prophage can be induced to get excised from bacteria and enter into a lytic cycle.

Plaque assay technique is an useful technique for assessing the phage population in a culture. The technique requires making serial dilutions of phage infected bacterial cultures and using such dilution for infecting fresh bacterial cultures on agar petriplates. Usually optimal dilutions give rise to countable number of phage plaques on petriplates. The number of plaques produced for a specific dilution together with complementation test is useful in detecting mutations occurring within the same gene. Infection of a bacteria with two phage mutants would result in a wild type phenotype if there is a complementation and the mutants are nonallelic. And they tend to produce a mutant phenotype if they do not complement each other and the mutants are allelic.

Bacterial multiplication is on a geometrical scale and a bacterial clone produced from a single cell may have a distinct phenotype by visual inspection or by simple chemical tests. Nutritional mutants or auxotrophs, can be replicated exactly on a complete medium, for the detection of specific mutations.

Transfer of genetic material from one bacterium to another may take place by three different ways. (i) Transformation (i) conjugation (iii) Transduction.

In transformation the genetic material in the form of free DNA nucleotides are acquired from the surrounding medium.

In conjugation there is a physical contact between two bacteria and the unidirectional flow of genetic material occurs through the conjugation canal. Bacteria with a fertility factor known as *F* act as donors during conjugation and transfer the genetic material to F^- bacteria. F factor is shown to be a circular self-replicating plasmid containing 6×10^4 nucleotides. The F bacteria could integrate with bacteria of chromosome and convert it into a high frequency recombination (Hfr) bacteria. In Hfr bacteria recombination frequency is 1000 times faster than obtained in $F^+ \times F^-$ matings. Conjugation using Hfr strains is very useful in mapping the bacterial chromosome.

Transduction is another process in which the transfer of genetic material from one bacterium to another is mediated by bacteriophage. Transduction is regarded as a generalized one when every portion of the donor DNA has an equal probability of being trasnferred, and any transduction particle has 1/100 to 1/50 portion of donor chromosome. Specialised or restricted transduction limits the transfer of only certain genes through bacteriophage. Both the processes are only of limited value in genetic mapping of bacterial chromosome. Sexduction is a process of transmission of chromosomal genes from a donor bacterium to a recipient by an F factor. In this process, the F factor is a plasmid excised from a Hfr strain and carries with it a sizable portion of bacterial chromosome. Such an F factor is referred to as F′ factor to distinguish it from an F^+ factor.

Plasmids and episomes are autonomously replicating extra chromosomal DNA molecules. Plasmids may carry certain genes such as ones for antibiotic resistance and may ensure the survival of susceptible strains of bacteria in which they are found. Plasmids which are capable of integrating with bacterial chromosome are known as episomes.

12.15 TERMINAL QUESTIONS

1. What is lysogeny? Briefly describe the events occurring during lysogeny.

.....

.....

.....

.....

.....

.....

.....

.....

2. What is meant by positive complementation test ? Illustrate your answer with the help of an example.

.....

.....

.....

.....

.....

.....

.....

.....

.....

3. Briefly describe the experiment that provided evidence that the transforming principle and the genetic material are one and the same.

4. Draw a diagram to explain the molecular mechanism of the transformation process as proposed for *Pneumococcus*.

5. Describe the experiment of Lederberg and Tatum which demonstrated that recombination did occur during conjugation.

6. Describe the experiment done to show that a physical contact between bacteria is necessary for transfer of genetic material during conjugation.

7. Briefly describe the F^+, F^-, Hfr and F′ cells.

8. Distinguish the terms generalised and restricted transduction with the help of examples.

12.16 ANSWERS

Self-Assessment Questions

1. (a) F (b) F (c) F (d) T (e) F . (f) T (g) T (h) F
   (i) T (j) T

2. (a) plaque assay technique (b) 1.9×10^7
   (c) double infection (d) $h^+ r$ and $h\ r^+$
   (e) billion

3. (1) (a) iv (b) i (c) v (d) ii
   (e) iii (f) vi
   (2) (a) 1000 (b) hot spots (c) muton, (d) recon.

4. (a) clones (b) partial (c) auxotrophs (d) Griffith
   (e) Rec A protein (f) Competence (g) linkage
   (h) two DNA fragments each carrying a single marker.

5. (a) Conjugation could be defined as the temporary fusion of two single celled organisms for the transfer of genetic information from one to another. The transfer could be unidirectional as in the case of bacteria.
   (b) It would require methionine and biotin in the culture medium.
   (c) There should have been transfer of gal^+ or pur^+ gene from one bacterium to the other. The chromosome segment from the donor carrying pur^+ or gal^+ should have recombined with recipient chromosome.
   (d) A donor
   (e) An F^+ cell contains the fertility factor or F plasmid and F^- cell does not.
   (f) A F factor is a circular plasmid of DNA found inside a bacterial cell capable of replicating independently of the main bacterial chromosome and contains roughly 6×10^4 nucleotides.
   (g) The genes in F plasmid DNA code for the fibrous pili proteins of the bacterium.
   (h) 1×10^{-7} cells.

(i) Whereas in an F^+ cell the F plasmid is having an independent existence, in an Hfr cell the F plasmid is integrated with bacterial chromosome.

(j) 1×10^{-4} cells. The Hfr cells transfer genetic material 100 times faster than the F^+ cells.

6. (a) transduction (b) abortive transduction
   (c) complete transduction (d) bio., gal.
   (e) lysogenic to lytic (f) restricted or specialised transduction

7. (a) iii, (b) iv, (c) vi (d) i (e) ii (f) v

TERMINAL QUESTIONS

1. The process of integration of phage DNA with bacterial chromosome is known as lysogeny. During this process, the phage attaches itself to the bacterium and injects its DNA into the bacterium. The phage DNA using bacterial machinery transcribes an mRNA which codes for a repressor protein and a DNA insertion enzyme. The repressor protein inhibits the synthesis of mRNA which codes for lytic functions. The DNA insertion enzyme catalyses the integration of phage DNA into bacterial chromosome.

2. Complementation test is the genetic test which shows whether two mutations are present within the same gene. For the example refer to Section 12.6.

3. Griffith chose the bacterium *Pneumococcus* which causes pneumonia in mammals. Of the two strains of the bacterium, R strain is non-virulent and S strain causes pneumonia. Mice infected with S strain die but ones with R strain survived. Heat killed S strain does not cause the death of mice. When infected with a mixture of a small number of live R strain and a large number of heat killed S strain, the mice were dead. Analysis of the blood of dead mice showed bacteria of S strains. Avery, Macleod and McCarty showed that addition of minute amounts of purified DNA from S-cells to the cultures of R cells resulted in the transformation of R strain to S strain. They further showed that transforming activity to be disrupted if DNA were treated previously with DNAase, suggesting that transforming principle and the genetic material are one and the same.

4. Refer to Fig. 12.14.

5. Refer to the Section 12.10 for the experiment of Lederberg and Tatum.

6. Refer to the Section 12.10 and fig. 12.17 which describe the classic U tube experiment that shows a physical contact between bacteria is necessary for conjugation to occur.

7. F^+ cell: a bacterium having a fertility (F) factor. The cell is a donor of genetic material during conjugation. The fertility factor itself is an independently existing, self-replicating plasmid. F^- cell: a bacterial cell that does not contain a fertility (F) factor. The cell is a recipient during the conjugation process and receives the genetic material from the donor.

   Hfr cell : a strain of bacteria exhibiting a high frequency of recombination. The cells have the F factor integrated into the chromosome. At the time of conjugation, all or part of the chromosome could be transferred to the recipient (F) cell. F′ cell : cell containing fertility factor which also includes a portion of the bacterial chromosome, when the fertility factor is induced to excise itself from the Hfr cell. Mistakes occurring during the excision process result in the transfer a portion of bacterial chromosome along with F factor.

8. Refer to Sections 12.11.1 to 12.11. 2.

GLOSSARY

Acentric chromosome : chromosome fragment lacking a centromere.

Acrocentric : a chromosome or chromatid that has its centromere near the end.

Adenine : a purine base found in RNA and DNA.

Amphidiploid : a species or type of plant derived from doubling the chromosomes in the F_1 hybrid of two species; an allopolyploid. In an amphidiploid the two species are known whereas in other allopolyploids they may not be known.

Aneuploid : an organism or cell having a chromosome number that is not an exact multiple of the monoploid (n) with one genome, that is, hyperploid has a number (e.g. 2n–1). Also applied to cases where part of a chromosome is duplicated or deficient.

Autopolyploid : a polyploid that has multiple and identical or nearly identical sets of chromosomes (genomes). A polyploid species with genomes derived from the same original species.

Autosome : all chromosomes other than sex chromosomes.

Auxotroph : mutant microorganisms that cannot synthesise a specific nutrient or substance needed for multiplication and require it in the medium. Wild type strains do synthesise the substance.

Centromere : the region of a chromosome associated with spindle fibers and participating in normal chromosome movement in mitosis and meiosis.

Chromatid : in mitosis or meiosis, one of the two identical strands resulting from self-duplication of a chromosome.

Chromatin : the deoxyribonucleohistone in chromosomes; originally named because of the readiness with which it stains with certain dyes (chromaticity).

Chromosomic aberration : abnormal structure or number of chromosomes; includes deficiency, duplication, inversion, translocation, aneuploidy, polyploidy or any other change from the normal pattern.

Chromosome banding : staining of chromosomes in such a way that light and dark areas occur along the length of the chromosomes. Lateral comparisons identify pairs. Each human chromosome can be identified on its banding pattern.

Colchicine : a chemical that prevents formation of the spindle in nuclear division.

Complementation test : a genetic test to assess whether two mutations occur within the same gene. If the two mutations introduced in a cell produce a wild type phenotype then they are non-allelic (they complement each other). If they produce a mutant phenotype, they are allelic (they do not complement).

Cistron : the region in a DNA molecule that specifies a polypeptide chain. Genetically defined as a region within which two mutations cannot complement each other.

Cis-configuration : the arrangement of two mutant sites within the same gene on the same homologue.

Clone : genetically identical cells or organisms derived from a single ancestor by asexual or parasexual methods..

Conjugation : the sexual exchange of genetic material in two single celled organisms by temporary fusion. In bacteria the exchange is unidirectional.

Deficiency (deletion) : absence of a segment of a chromosome reducing the number of loci.

Denaturation : loss of native configuration of a macromolecule, usually accompanied by loss of biological activity.

Duplication : the occurrence of a segment more than once in the same chromosome or genome.

Euchromatin : genetic material that is not stained so intensely by certain dyes during interphase and that comprises many different kinds of genes.

Euploid : an organism or cell having a chromosome number that is an exact multiple of the monoploid (n) or haploid number. Terms used to identify different levels in an euploid series are diploid, triploid, tetraploid and so on.

Episome : extrachromosomal circular DNA molecule capable of self replication either independent of, or as an integral part of host chromosome.

Genome : a complete set (n) of chromosomes (hence, of genes) inherited as a unit from one parent.

Gene expression : the multistep process by which a gene is regulated and its product synthesised.

Haploid (monoploid) : an organism or cell having only one complete set (n) of chromosomes or one genome.

Heterochromatin : chromatin staining darkly even during interphase, often containing repetitive DNA with few genes.

Hot spots : those sites in a locus that are more susceptible to mutation than others.

Lytic cycle : a condition in which the phage DNA replicates inside the bacterium which the phage infects, forms several phage progeny and cause the lysis of the bacterium.

Lysogeny : a condition in which the phage DNA becomes integrated with the chromosome of the bacterium which it infects.

Monosomic : A diploid organism lacking one chromosome of its proper complement (2n–1); an aneuploid. Monosome refers to the single chromosome, disome to two chromosomes of a kind, and trisome to three chromosomes of a kind.

Minimal medium : a medium containing nutrients that can support growth and reproduction of only wild type strains of an organism.

Muton : the smallest unit of DNA (a single base) that can be altered to produce a charge.

Nucleic acid : a macromolecule composed of phosphoric acid, pentose sugar, and organic bases; DNA and RNA.

Nucleoprotein : conjugated protein composed of nucleic acid and protein; the material of which the chromosomes are made.

Nucleotide : a unit of DNA and RNA molecules containing a phosphate, a sugar and an organic base.

Paracentric inversion : an inversion that is entirely in one arm of a chromosome and does not include the centromere.

Parthenogenesis : the development of a new individual from an egg without fertilisation.

Phosphodiester bond : in nucleic acids, the covalent bond formed between the phosphate group of one nucleotide and the 3′-OH group of the next nucleotide in line; these bonds form the backbone of nucleic acids.

Peptide : a compound containing amino acids. A break-down or build-up unit in protein metabolism.

Peptide bond : a chemical bond holding amino acid subunits together in proteins.

Pericentric inversion : an inversion including the centromere, hence involving both arms of a chromosome.

Polynucleotide : a linear sequence of joined nucleotides in DNA or RNA.

Polyploid : an organism with more than two sets of chromosomes (2n diploid) or genomes (e.g. triploid - 3n, tetraploid - 4n, pentaploid - 5n) .....so on.

Prototroph : a strain of microorganism that can grow and reproduce in a well defined minimal medium, usually the wild type strains are protoprophs.

Pilus : a filamentous projection from the bacterial surface, associated with cells possessing F factors.

Plasmid : extrachromosomal circular DNA molecules capable of self-replication independent of host chromosomes.

Plaques : clear areas formed on an opaque bacterial lawn due to the growth and reproduction of phages.

Prophage : the phage genome that is in an integrated state in a bacterium.

Renaturation : the restoration of a molecule to its native form, in nucleic acids, it refers to the formation of a double-stranded helix from complementary single-stranded molecules.

Recombination : the process by which new combination of genes is produced in a chromosome.

Recon : the smallest unit of DNA between which recombination could occur.

Satellite DNA : a component of the genome that can be isolated from the rest of the DNA by centrifugation, usually it consists of short, highly repetitive sequences.

Sex chromosomes : chromosomes that are connected with the determination of sex.

Sexduction : a process of transfer of chromosomal genes from a donor bacterium to the recipient through F factor.

Syndrome : a group of symptoms that occur together and represent a particular disease.

Thymine : a pyrimidine base found in DNA. The other three organic bases — adenine, cytosine, and guanine are found in both RNA and DNA, but in RNA, thymine is replaced by uracil.

Trans-configuration : the arrangement of two mutant sites within the same gene on the opposite homologue.

Transformation : conversion of one genotype into another when the recipient cell acquires exogenous DNA from the surrounding medium.

Transduction : the recombination in a bacterial chromosome mediated by a phage.

X chromosome : a chromosome associated with sex determination. In most animals, the female has two and the male has one X chromosome.

Y chromosome : the partner of the X chromosome in the male of many animal species.

FURTHER READINGS

1. Ayala, F. J. & Kiger, J. A. Jr. 1984. Second Edition. *Modern Genetics*. The Benjamin Cummings Publishing Company Inc., California.
2. Carlson, E.A. 1985. *Human genetics*. Tata McGraw Hill Publishing Co. Limited, New Delhi.
3. Gardner, E. J.: Simmons, M. J. & Snustad, D.P. 1991. Eighth Edition. *Principles of Genetics*. John Wiley & Sons, Inc. New York.
4. Gupta, P. K. 1987. Second Edition. *Genetics*. Rastogi Publications, Meerut.
5. Klug, W. S. & Cummings, M. R. 1986. *Concepts of Genetics*. Second Edition. Scott Foresman & Co., illinois.
6. Lewin, B. 1983. *Genes*. Wiley Eastern Ltd., New Delhi.
7. Rao, K. R. 1986. Third Edition. *Text Book of Biochemistry*. Prentice Hall of India (Pvt) Ltd., New Delhi.
8. Strickberger, M.W. 1985. Third Edition. *Genetics*. Macmillan Publishing Company, New York.
9. Swanson, C. P; Merz, T. and Young, W.J. 1990. Second Edition. *Cytogenetics*. Prentice Hall of India (Pvt) Ltd., New Delhi.
10. Lloyd, J. R. 1986. *Genes and Chromosomes*. English Language Book Society, Macmillan, Hongkong.

Dear Student,

While studying these units you may have found certain portions of the text difficult to comprehend. We wish to know your difficulties and suggestions in order to improve the course. Therefore, we request you to fill and send us the following questionnaire which pertains to this block.

QUESTIONNAIRE

LSE-03
Block-2

Enrolment No. | | | | | | | | | |

1) How many hours did you need for studying the units?

Unit Number	7	8	9	10	11	12
No. of hours						

2) How Many hours (approximately) did you take to do the assignments pertaining to this block?

Assignment Number	
No. of hours	

3) In the following table we have listed 4 kinds of difficulties that we thought you might have come across. Kindly tick (√) the type of difficulty and give the relevant page number in the appropriate columns.

Page Number	Types of difficulties			
	Presentation is not clear	Language is difficult	Diagram is not clear	Terms are not explained

4) It is possible that you could not attempt some SAQs and TQs.

In the following table are listed the possible difficulties. Kindly tick (√) the type of difficulty and the relevant unit and question numbers in the appropriate columns.

Unit No.	SAQ No.	TQ No.	Types of difficulties			
			Not clearly posed	Cannot answer on basis of information given	Answer given (at end of Unit) not clear	Answer given is not sufficient

5) Were all the difficult terms included in the glossary. If not, please list in the space given below.

6) Any Other Suggestion(s)

To

The Course Coordinator (LSE-03; Genetics)
School of Sciences
Indira Gandhi National Open University
Maidan Garhi
New Dehli-110 068.

Uttar Pradesh
Rajarshi Tandon Open University

UGZY/BY-07
Genetics

Block

3

GENE STRUCTURE AND FUNCTION

BLOCK 3 GENE STRUCTURE AND FUNCTION

In this block we have attempted to convey some of the excitement that has resulted from the recent developments in the field of genetics. This block consists of six units, Units 13 to 18, focussing on different aspects of gene structure and function.

The first unit of this block (Unit 13) examines the fine structure of gene in light of the recent findings. The bacteriophages have proved to be valuable genetic tools for investigating the fine structural details of the gene. The crucial experiments with phage *T4* involving recombination and deletion mapping have brought to light several new facts about fine structure of gene. Based on these studies now a much clearer picture of gene – its fine structure and function, has emerged. The ultimate fine structure mapping of the gene has made it possible to locate precisely the defective sequences within genes and to correct them at molecular level.

Next you would study an important gene function, that is, regulation of gene expression. This aspect has been covered in two units. Unit 14 focusses on gene expression in prokaryotes and Unit 15 in eukaryotes. In Unit 14, we discuss the regulation of gene function in prokaryotes and the operon as a unit of gene regulation. Current molecular approaches to the regulation of gene expression in bacterial operons and regulation of gene expression in bacteriophages are also discussed in this unit. In Unit 15, you will study how gene expression is regulated in eukaryotes.The genetic control of development in higher eukaryotes is very complex. In this unit we have dealt with some of the exciting developments in the field.

In Unit 16 we discuss the mechanism that generates variability in genetic material, namely the mutations. Mutations along with recombinations form the source of variability from which the selection process promotes adaptations to suit specific environments. In this unit we discuss the different types of mutations. There is a discussion on the rates at which the mutations occur. Many mutations are spontaneous and random events. Mutations can also be induced by using mutagenic agents. The unit on 'Mutations and Mutagenesis' further describes the various agents that would induce mutations in the genetic material.

One of the dreaded diseases for which a definite cure is not available is cancer. Is cancer a heritable disease? What are the agents that cause cancer in human beings? Does cancer occur because of the alterations of the genetic material? These questions are discussed in Unit 17 - 'Carcinogenesis and Teratogenesis'. Teratogenesis relates to abnormalities caused during foetal development by chemicals and radiations. You may be familiar with the thalidomide syndrome – a foetal abnormality caused by the drug thalidomide. The unit lays emphasis on foetal abnormalities that may be caused by extraneous agents.

The last unit in this block discusses the genetics of immune system. Mammals in general and humans in particular can defend themselves from countless microorganisms that might possibly invade their body. The immune system consists of numerous cells generally known as lymphocytes that defend the body either by cell mediated or humoral mediated mechanism. Cell mediated immunity functions mostly by phagocytosis and humoral mediated immunity functions by the production of high molecular globular proteins called antibodies. In this unit we describe extensively the specificity of antigen antibody reactions and the antibody diversity responsible for such specificity. More importantly the discussion would centre around the genetics of antibody diversity, the mechanism that generates millions of antibodies from a relatively a fewer number of genes.

As in the previous two blocks, there is a feedback form at the end of this block too. After completing your study of this block, do send your feedback on this block.

Objectives :

After studying this block you should be able to :

- discuss modern viewpoint of the gene fine structure;
- explain the basic mechanisms of regulation of gene expression in bacteria and bacteriophages;
- describe regulation of gene expression in eukaryotes and comment on gene regulation during development of eukaryotes;
- describe the different types of mutations that arise in the genetic material and the rates at which they occur;

- distinguish between the spontaneous and induced mutations and list the various mutagenic agents that would induce mutations;
- discuss that cancer is caused by the alterations in the genetic material and the carcinogenic agents are responsible for such alterations, and the teratogenesis relates to abnormalities during the foetal development and certain chemicals and radiations could act as teratogenic agents;
- List various types of cells that are involved in the defense of human body against the microorganisms and describe that the antigen-antibody reactions form the basis of immune response;
- appreciate that the genetics of antibody diversity is an unique mechanism that generates more than a billion antibodies during the life-time of an individual.

ACKNOWLEDGEMENTS

Shri Rajbir Singh and Ms. Usha Marwah for word-processing.

UNIT 13 GENE FINE STRUCTURE

Structure

13.1 INTRODUCTION

The term 'gene' till date is the most important word in Genetics. The present understanding of its fine structure has come a long way from its early vague concept. At the time of acceptance of chromosome theory of inheritance, the gene was considered as a bead on the chromosomal string. One gene was thought to be separated from its adjacent genes by non-genetic material. Mutant alleles of a gene were considered as beads of different colours, with only one bead of a particular colour (representing a particular gene) in each string. It was also thought that recombination occurred at positions between the beads whereas mutation caused a change in the bead thereby resulting in an alteration of gene expression. The gene was regarded as an *indivisible* entity.

The bead theory of gene was popular until about 1940 owing to the lack of resolving power of the genetic systems being studied. The genetic studies of 1950s and early 1960s suggested that *gene* was the *unit of function, mutation and recombination*. And subsequently with the introduction of microbial genetic systems, alongwith their great resolving power, questions concerning the possible substructure of a gene began to be settled. These latter studies sharply defined the basic units of mutation, recombination and genetic function. The genetic investigations revealing that the gene can be split into its subunits, began at about the same time when the Watson-Crick model for the structure of DNA was proposed. The elucidation of the fine structure of gene bridged the gap between the genetic map and the physical structure of the genetic material itself. As a result of these studies, the concept of the gene as a sequence of nucleotide pairs in a DNA molecule evolved. In this unit, you would study the various experimental approaches revealing the fine structure of gene.

Objectives

After studying this unit you should be able to:

- define and use in proper context the terms related to the use of bacteriophages in studying the fine structure of the gene (Section 13.2);
- analyse the recombination and deletion mapping experimental approaches for the fine structure of gene (Section 13.3);
- explain how the cis-trans test has added new dimension to our understanding of the fine structure of the gene (Section 13.4);
- interpet data obtained from the recombination and complementation experiments (Sections 13.3 and 13.4);

- describe the ultimate fine structure mapping and list its advantages and applications (Section 13.5); and
- explain molecular organisation, functioning, and genetic implications of overlapping genes (Section 13.6).

13.2 A GENETIC SYSTEM - BACTERIOPHAGE

The use of bacteriophages in genetic experiments led to considerable understanding of the detailed structure of the gene. Before beginning to discuss the experiments exploring the gene structure, let us know some important features of these systems. Virtually all genetics experiments with phages are performed without direct observation of either parents or progeny, meaning thereby that indirect methods are employed to determine the phenotype of genetic variants. In Unit 12, you have studied that the life cycles of phages fit into two distinct categories — the *lytic* and the *lysogenic* cycles (see Section 12.3). You may recall that in the lytic cycle, phage nucleic acid enters a cell and replicates repeatedly, the bacterium is killed and hundreds of phage progeny result. All phage species can undergo a lytic cycle. A phage capable only of lytic growth is called *virulent*. In the alternative lysogenic cycle, no progeny particles are produced and the bacterium survives, and a phage DNA molecule is transmitted to each bacterial daughter cell. In most cases, transmission of this kind is accomplished by integration of the phage chromosome into the bacterial chromosome. A phage capable of such a life cycle is called *temperate*.

Plaque Formation and Phage Mutants : The presence of phages can be easily detected because of some distinctly observable manifestations. In a lytic cycle, an infected cell gets lysed releasing phage particles onto the growth medium. Let us elaborate it further. You have already studied the plaque assay technique in Unit 12, Section 12.4. As you know, a large number of bacteria, that is about 10^8 are placed on a solid nutrient medium. After a period of growth, a continuous turbid layer of bacteria results. If a phage is present at the time the bacteria are placed on the medium, it adsorbs to a bacterial cell and shortly afterwards, the infected cell lyses and releases many phages: each of these progeny phages get adsorbed onto a nearby bacterium, and after another lytic cycle, this bacterium in turn releases phages that can infect many other bacteria in the vicinity. These multiple cycles of infection continue, and after several hours the phages destroy all of the bacteria in a localised area, giving rise to a clear transparent region in the otherwise turbid layer of confluent bacterial growth known as a *plaque* (see Fig. 13.1)

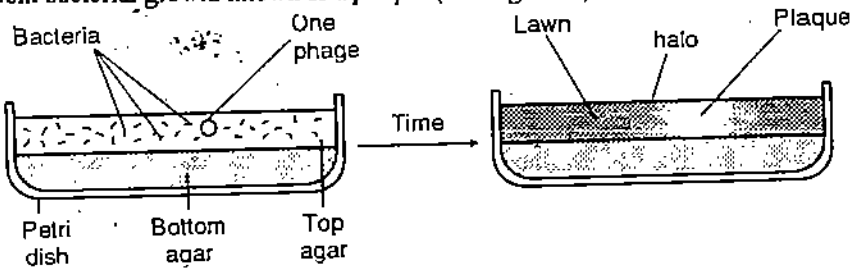


Fig. 13.1 : Diagrammatic representation of plaque formation in culture medium. Bacteria grow and form a translucent lawn. Due to the lysis of bacterial cells, in the vicinity of the site of initial phage, clear transparent area appears. Each such clear area is known as a plaque. Note the halo around the large plaques – it is the result of large amounts of the lysis enzyme diffusing outward and lysing uninfected cells.

Phages can multiply only in growing bacterial cells, and so exhaustion of nutrients in the growth medium limits phage multiplication and eventually the size of the plaque. As a plaque is a result of an initial infection by one phage particle, the number of individual plaques originally present on the medium can be counted.

The plaque of a wild-type of virus is quite distinct. It may be small with smooth edges, whereas a mutant form may be large. Some plaques are cloudy, others are clear. Some forms of plaques are shown in Figure 13.2.

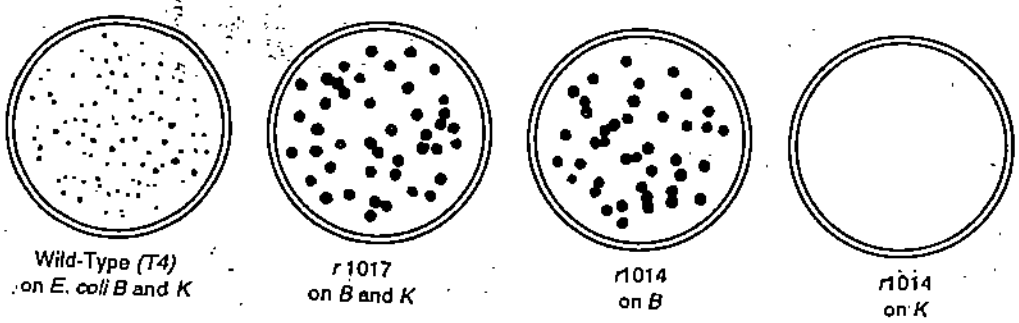


Fig.13.2: Plaques formed by wild-type phage and its different mutants. Phages may differ in the type of plaque they form and also in the range of hosts in which they can reproduce. Phage *T4* produces small, wild-type plaque on *E.coli B* and *K*. Certain mutants (*r* mutants) produce large plaques on *B* and *K*. Still others, the *rII* mutants, produce the large plaques on *B* but no plaques at all on *K*.

The genotype of the phage mutants can be determined by studying the plaques. In some ases the appearance of the plaque is sufficient. Some mutations affect the *size* of the laque. Phage mutations that decrease the number of phage progeny from infected cells ften yield smaller plaques. Large plaques can be produced by mutants that cause remature lysis of infected cells so that each round of infection proceeds more quickly Fig. 13.3). Another type of phage mutation can be identified by the ability or inability of ie phage to form plaques on a particular bacterial strain.

esides the phage morphology mutants, there are phage mutants with extended host range. hat is, the mutant phages are able to attach to the bacterial strain to whom their wild-pes would not attach. Such mutants are ideal materials for identifying and studying the nes responsible for phage adsorption. Phage morphology mutations and extended host-nge mutations were among the first phage mutations to be studied. However, they occur only a small number of all the genes present in the phage genome.

ie majority of phage genes are associated with functions that are essential for the olication and production of progeny phages. Any mutations in these genes that prevent production of progeny are thus lethal. In experimental conditions this can be detected no plaques are formed. Such lethal mutations, except under special circumstances inot be propagated, as phages are haploids. *Conditional lethal mutations* are those that lethal under one set of conditions – *restrictive conditions*, but can be propagated under ther set of conditions – *permissive conditions*. These mutations permit most phage ies to be identified and studied. The first conditional lethal mutations to be studied in ge genetics were the *rII* mutations of *T4*. You are already familiar with this mutant n your study of Unit 12, Section 12.6.

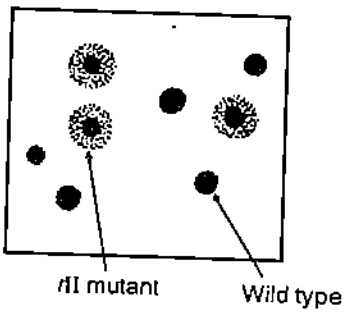


Fig. 13.3: Plaques of *E.coli* phage *T4*. Two type of plaques are present. The smaller plaques are made by wild-type phage; the larger plaques are made by a *rII* mutant phage. Beside large size of plaques, the plaques of mutant phages, have halo around them.

Q 1

in the blanks with appropriate words:

A ..... refers to a clear area in a bacterial lawn caused by phage reproduction.

Conditional lethal mutation is the one that kills the affected organism under one set of environmental conditions known as .....conditions but are not lethal in another set of conditions called the ..... conditions.

In the ..... life cycle of the phage there is integration of the phage chromosome into the bacterial chromosome as a prophage.

The results of ..... cycle of the phage are : the production of progeny phages and the lysis of the bacterial cell.

A ..... phage can only enter into the lytic cycle of growth.

A bacteriophage capable of conferring lysogeny upon a host cell is known as the ..... phage.

FINE STRUCTURE ANALYSIS OF A GENE

Unit 6 of this course, you have studied about the construction of genetic maps (also as chromosome maps), based on the frequency of recombination between genes.

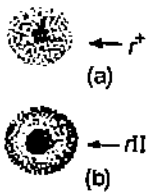


Fig. 13.4 : The r^+ and mutant *rII* plaques on a lawn of *E.coli B*. Note the *r* plaques (a) is turbid with a fuzzy edge, and (b) the *rII* plaque is larger, clear, with a distinct boundary.

Such recombination-based mapping of the distance between genes is called *intergenic mapping* (inter – between). The same general principle of recombinational mapping can be applied to mapping the distance between mutational sites within the same gene. Such a type of mapping is called *intragenic mapping* (intra = within). We shall more specifically deal with the intragenic or fine-structure mapping in this unit. On the basis of such studies, the fine structure of a particular gene could be understood.

The classical fine structure analysis of the gene were carried out by Seymour Benzer with *rII* mutants or phage *T4*, during the 1950s. In his experiments, Benzer used phage *T4* principally because bacteriophages produce large number of progeny that facilitate the potential determination of very low recombination frequencies. Another advantage was that *rII* mutants can easily be selected in large number as each has distinctive plaque morphology when plated on *Eschericlua coli* strain *B*. Benzer found that when the cells of *E.coli B* (same as strain *B*) or *E.coli K12* (λ) growing on solid medium in a petridish are infected with the wild-type *T4* phages, i.e., r^+, then small, turbid plaques with fuzzy edges are formed (see Fig.13.4a). An interesting feature about the *E.coli* strain *K12* (λ) is the presence of incorporated lambda phage genome. On the other hand, when cells of *E.coli* strain *B* are infected with a mutant *r* (rapid lysis) of phage *T4, rII,* then large clear plaques with distinct edges are produced (Fig. 13.4b). From the intergenic mapping experiments the *r mutations were found to map at several locations in the phage's genome* suggesting there are several *r* genes (see Fig. 13.5).

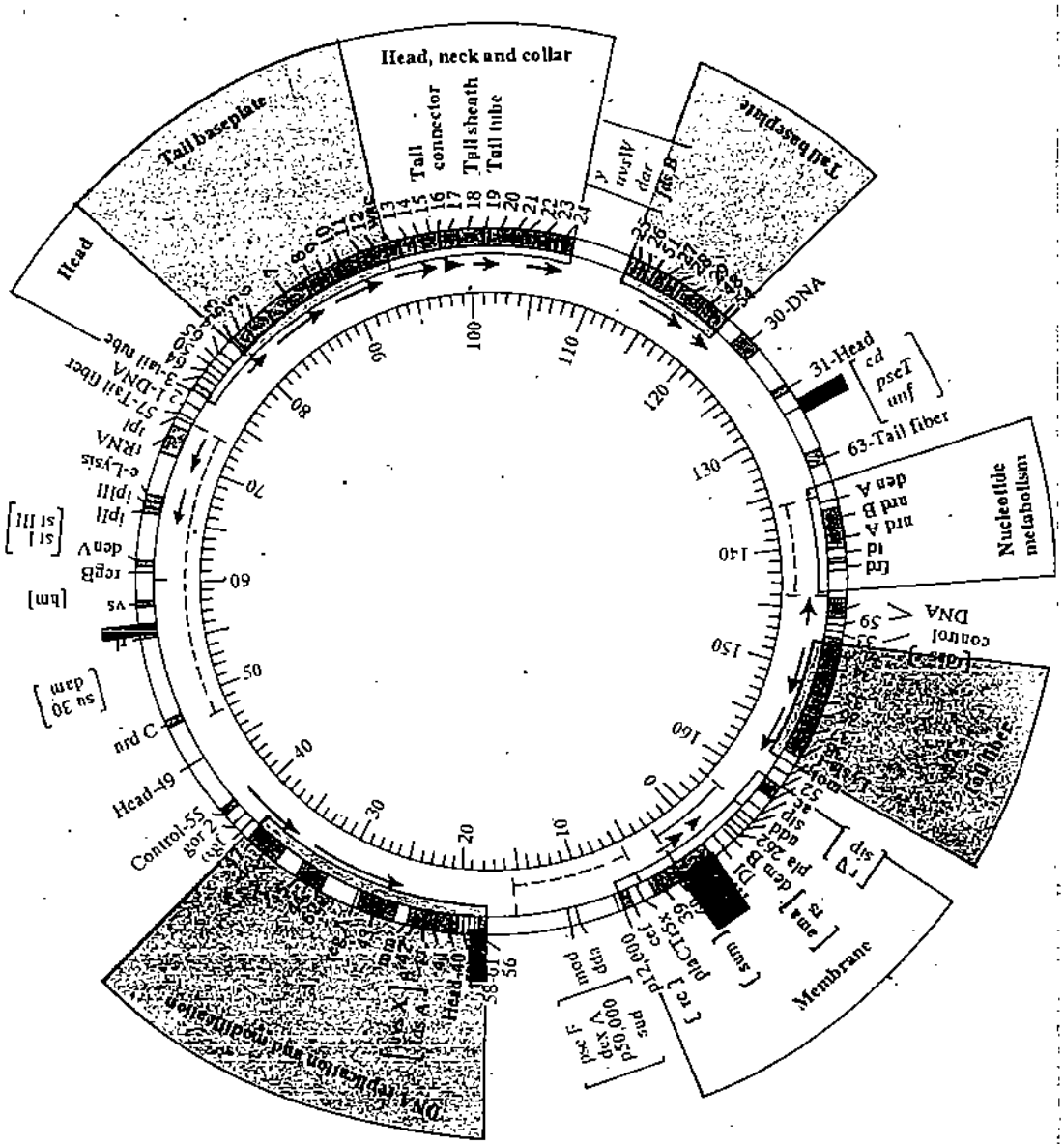


Fig. 13.5 : Genetic map of the bacteriophage *T4*. The genes with related functions are clustered around the perimeter of the map. Note the locations of *r* loci.

In his experiments Benzer used particular *r* mutants from two adjacent genes *rIIA* and *rIIB* in the *rII* region on *T4* map. These two genes affected the same phenotypic trait. These *rII* mutants in addition to their *plaque morphology* also have another feature not found in

their wild counterparts, and that is, *host range properties*. Such mutants can grow only on certain strains of *E.coli*. The wild-type *T4* phages can grow both on *E.coli B* as well as *E.coli K12* (λ) strains. Whereas, on the other hand, the *rII* mutants can grow in and lyse the cells of strain *B* but are unable to grow in the cells of *K12* (λ) strain. Thus for the *rII* mutants *E.coli* strain *B* is a *permissive host* whereas *E.coli* strain *K12* (λ) is *nonpermissive host*. In other words, the *rII* mutants are *conditional mutants* as they can grow under one set of conditions and not under the other.

13.3.1 Recombination in *rII* Mutants

Seymour Benzer realised that the growth defect of *rII* mutants on *E.coli K12* (λ) could serve as a powerful selective tool for detecting the presence of a nearly small proportion of r^+ (wild-type) phages within a large population of *rII* mutants. Based on this finding he first collected thousands of *rII* mutants, some of which had arisen spontaneously and some had been induced by mutagen treatment.

From the mutants collected, he set out to construct a fine structure genetic map of the *rII* region. For this purpose he first crossed about 60 independently isolated *rII* mutants, in all possible combinations using the permissive host, that is, strain *B*. The progeny phages were collected after the host cells had lysed. These progeny phages obtained from each cross were examined from the following angles: i) their total number (per millilitre), and ii) quantification of any rare r^+ recombinants produced, (an outline of his experiments is given in Fig.13.6). In order to determine (i) above, he plated a sample of phage progeny on *E.coli B* – the permissive host. And for (ii) another sample of phage progeny was plated on *E.coli K12* (λ) – the nonpermissive host. In this manner, Benzer was able to calculate the percentage of the rare r^+ recombinants between closely- linked genetic sites.

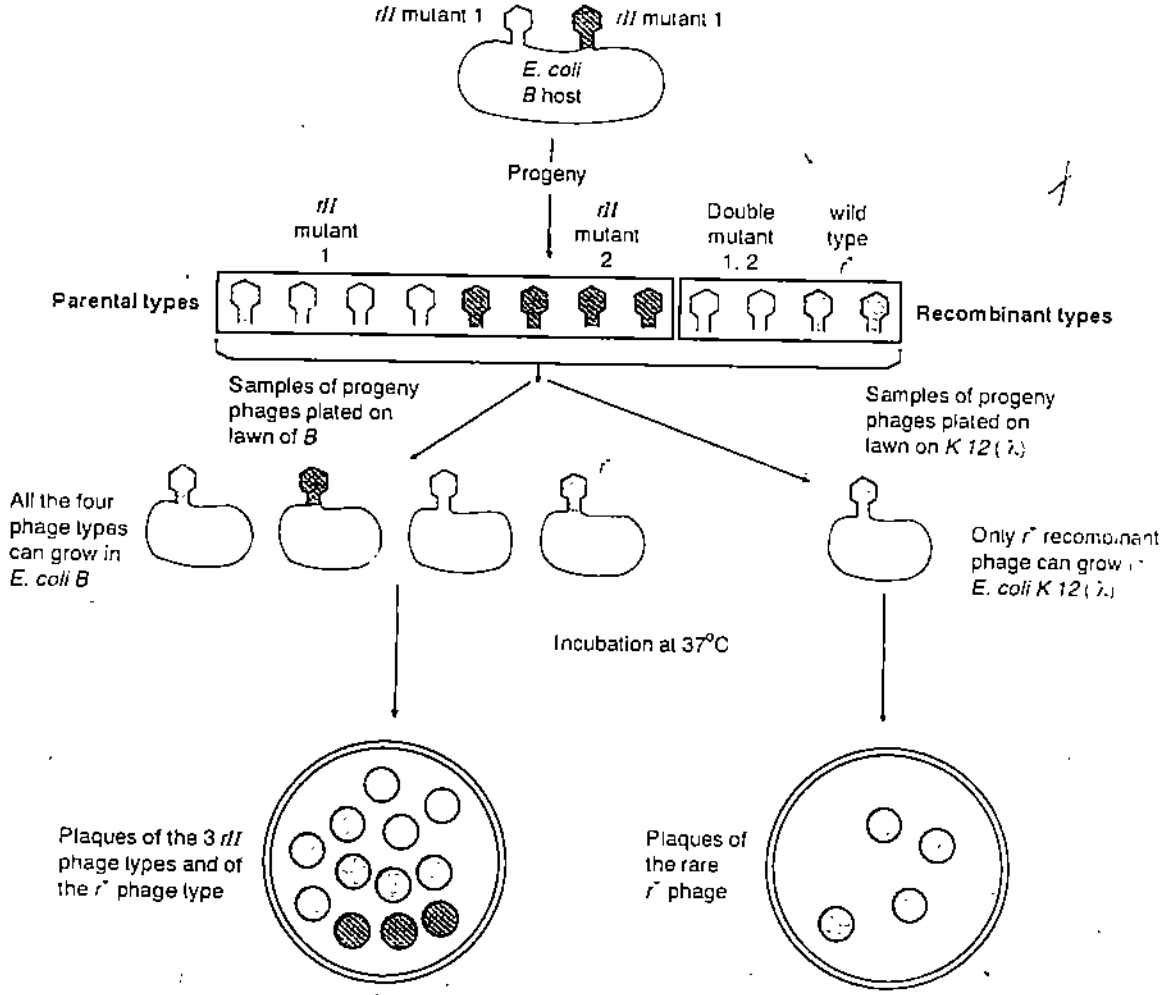


Fig. 13.6: Outline of the procedure used by Benzer for determining the number of r^+ recombinants from a cross involving two *rII* mutants of *T4*.

In each cross between two *rII* mutants, as a result of single cross over event, four genetic classes of progeny were usually found: two parentals and two recombinants. Such a cross and its results are diagrammatically depicted in Fig. 13.7. A cross between mutants *rIII*

and *rII2* results in two parentals *rII1*, *rII2*, and two recombinants *rII1,2* (the double mutant) and r^+ (the wild-type). As you know the map distance between two mutants is determined by the percentage of recombinant progeny among all the progeny of the two mutants. For the crosses involving *rII* mutants, a single crossover event between the two mutations will result in the r^+ and double *rII* mutant recombinants. Therefore, the frequencies of the two recombinants classes of phages are expected to be the same. Consequently, the total number of recombinants is approximated by twice the number of r^+ plaques counted on plates of strain *K12* (λ). The general formula for the map distance between the *rII* mutations is as given below.

$$\text{Map distance} = \frac{2 \times \text{number of } r^+ \text{ recombinations}}{\text{Total number of progeny}} \times 100$$

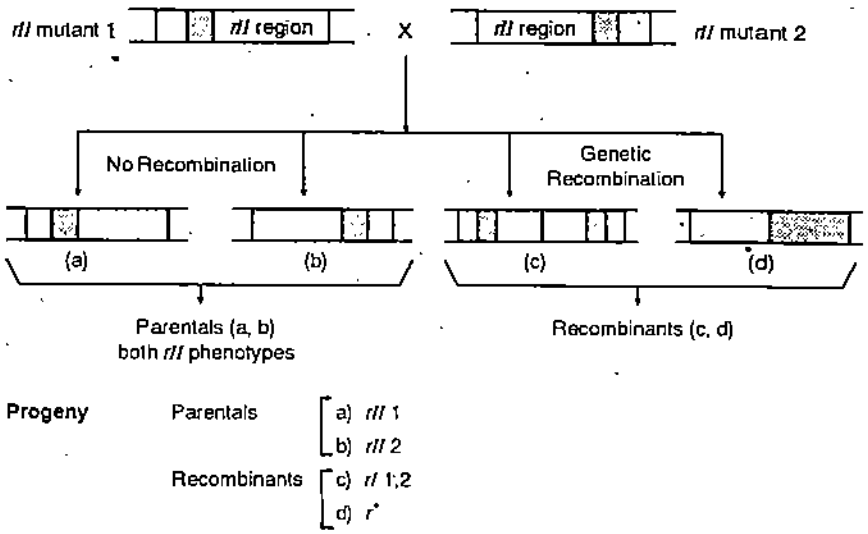
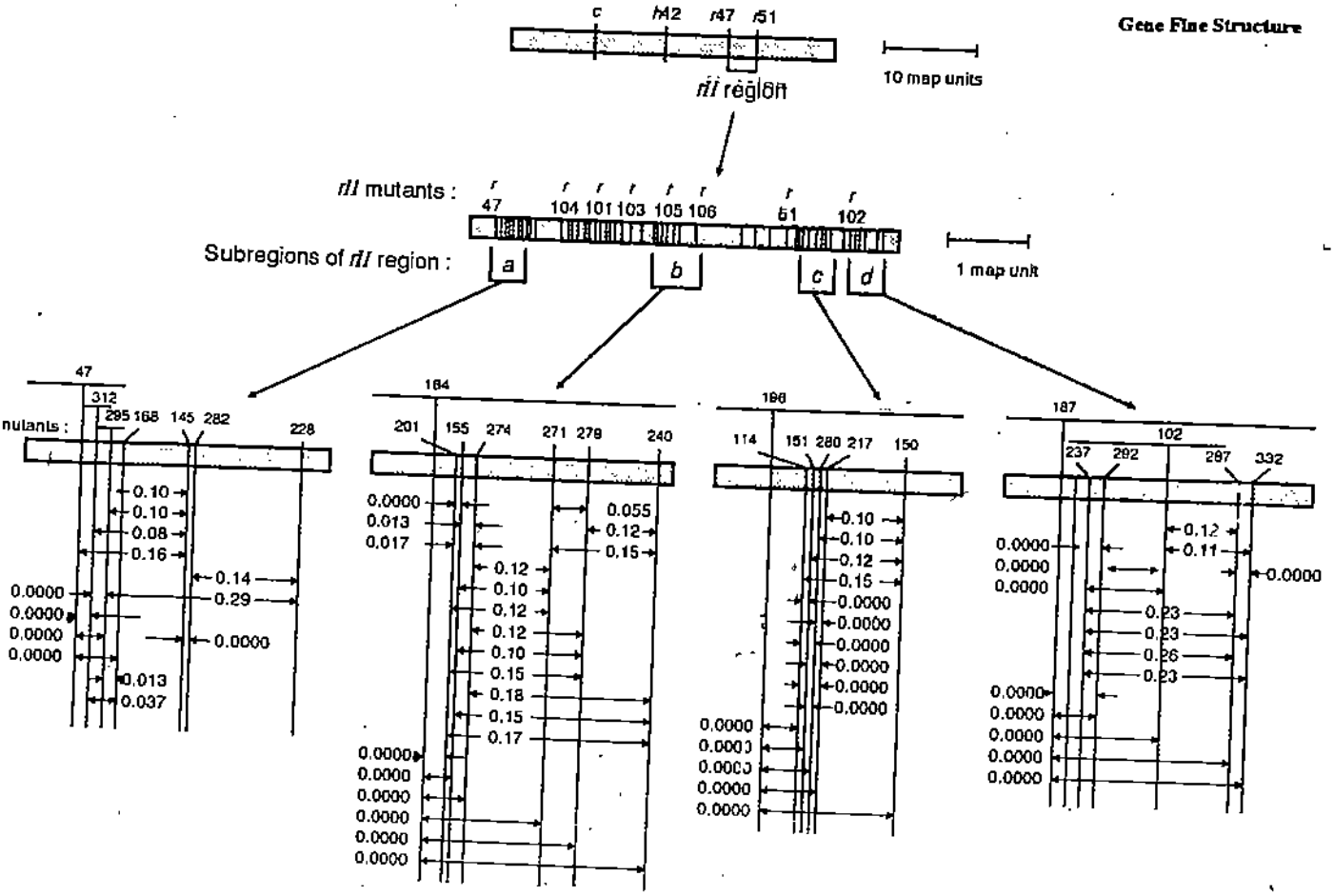


Fig. 13.7: A cross between two *rII* mutants with mutation in different sites within the *rII* region showing the production of parental and recombinant progeny. Note, if no crossing over occurs, then progeny with parental genotypes are produced. If a single cross over occurs between the sites of the two mutants, progeny phages with recombinant genotypes are produced.

Based on the recombination data obtained from all possible pairwise crosses of the 60 *rII* mutants, Benzer was able to construct a linear genetic map (Fig. 13.8). From the data obtained from his crosses he also observed that some mutant pairs, on crossing, did not produce any r^+ recombinants. From these results he concluded that those pairs carried mutations at exactly the same site; that is, the same nucleotide pair in the DNA had been changed. Thus, there was no possibility of recombination between the mutations. Such mutations that change the same nucleotide pair within the gene are termed *homoallelic.* Besides the above situation, a majority of mutants on crossing did produce r^+ recombinants: This indicates that the mutations they carried had altered different nucleotide pairs in the DNA. Such mutants that change different nucleotide pairs within a gene are known as *heteroallelic* (heteroalleles also known as pseudoalleles).

From the map (Fig. 13.8) the minimum map distance (that denotes the lowest frequency with which r^+ recombinants were formed) was found to be 0.01 per cent. This minimum map distance figure of 0.01 per cent can be utilised to calculate the molecular distance or the distance between the base pairs that are involved in recombination event. The circular genetic map of *T4* is known to consist of about 1,500 map units. One map unit is the distance separating two loci that recombine with a frequency of 1% (see Unit 6). If two *rII* mutants produce 0.01 per cent r^+ recombinants, the mutations are separated by 0.02 map units or by $0.02/1500 = (1.3 \times 10^{-5}) \times (2 \times 10^5)$, or about 2 base pairs. This suggests that genetic recombination can occur in distances of 2 base pairs or less. We also know that by



13.8: Benzer's preliminary fine structure genetic map of the *rII* region of phage *T4*. This map was constructed based on the information obtained from crosses of an initial set of 60 *rII* mutants. Lower levels in the figure show finer details of the map. In the lowest level the numbered vertical lines indicate individual *rII* mutants. The decimals indicate the percentage of *r*⁺ recombinants found in crosses between the two *rII* mutants connected by an arrow.

ınd large genes consist of a few hundred base pairs. So, Benzer's mapping data clearly ndicated that *recombination could occur within a gene.* We now know that *nucleotide ›air is the unit of recombination also know as recon.* Benzer's work also established that he nucleotide pair is also the *unit of mutation or muton.* These definitions clearly eplaced the classical ones that the whole gene is the unit of recombination and the unit of nutation.

AQ 2

Vrite short answer to the following questions.

) Distinguish between intergenic and intragenic mapping.

.....

.....

.....

) Why did Benzer prefer to use phaqe *T4* as his genetic tool for investigating the fine structure of gene?

.....

.....

.....

.....

iii) What are the phenotypic differences in the wild-type and *T4* mutants in permissive and nonpermissive hosts. Present your answer in a tabular form.

iv) Which two phenotypic traits did Benzer select for his studies on recombination in the *rII* mutants?

.....

v) How could Benzer detect and calculate the percentage of the rare wild-type (r^+) progeny in his recombination experiment?

.....

.....

.....

.....

.....

vi) In Benzer's experiment, on crossing the 60 *rII* mutants in all possible combinations, some mutant pairs did not yield any r^+ type progeny. What is the interpretation of this finding?

.....

.....

.....

vii) What aspect of gene fine structure emerged from Benzer's experiments on recombination in *rII* mutants?

.....

.....

.....

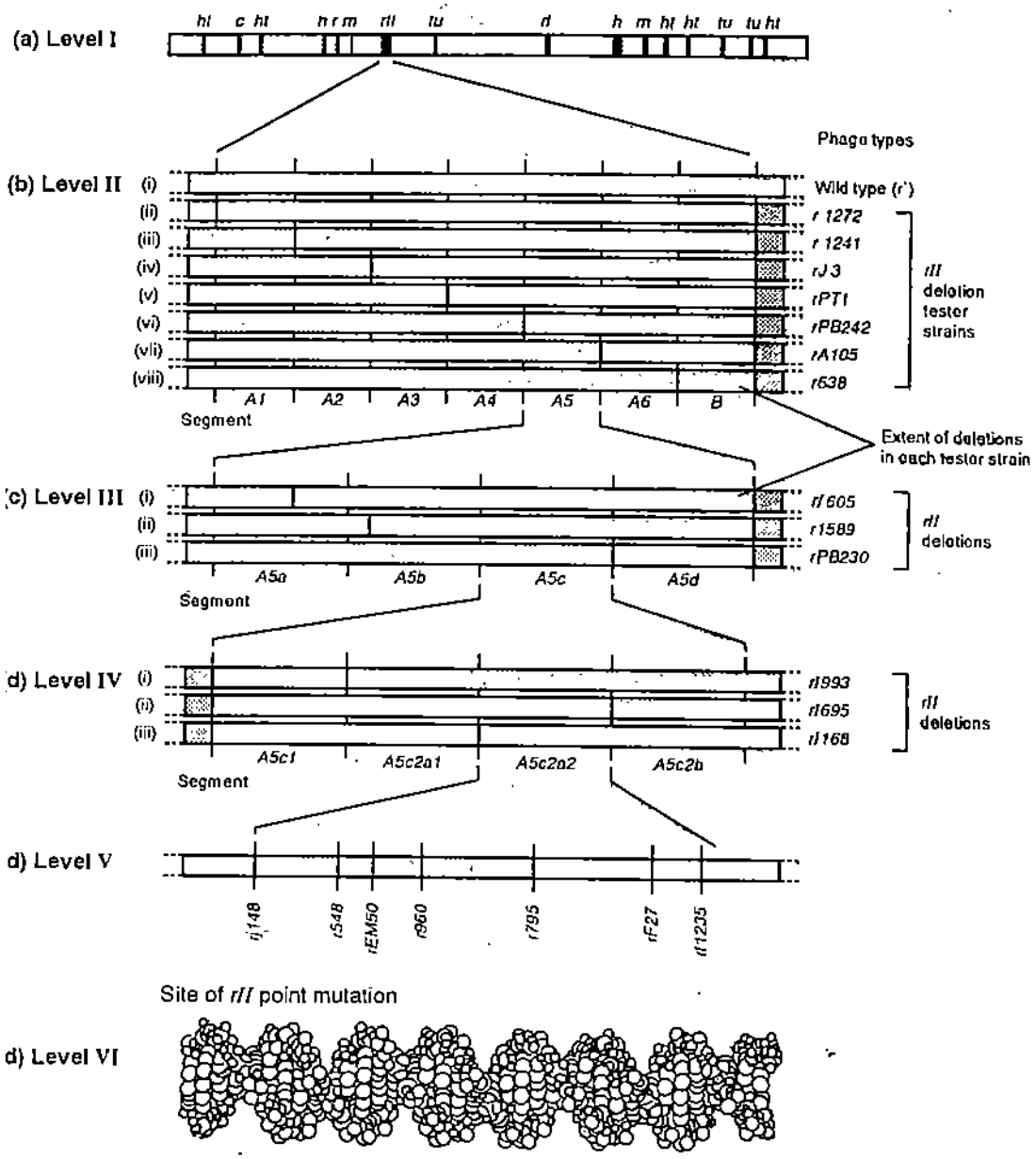
.....

13.3.2 Deletion Mapping

Benzer continued his genetic experiments with the odd 3000 mutants *rII*, in order to complete the fine structure map. He realised this overwhelming task would require carrying out some 5 million crosses, with upto 50 crosses per day. He found a simpler alternative to the process, that is, the technique of *deletion mapping*. Deletion mapping enables one to determine the order of a group of deletions at a locus. Deletions are the same as deficiencies that you have studied in Unit 9. As a result of deletion, several bases are lost for good from the organism's genome. Benzer, subsequently made use of deletion mapping technique to localise unknown mutations.

Most of the *rII* mutants isolated by him were point mutants, that is, their changed phenotype was the result of an alternation of a single nucleotide pair. A point mutant can revert to the wild-type form (also known as revertant) spontaneously or following treatment with an appropriate mutagen. But some of the Benzer's *rII* mutants did not revert (non-revertants) and also they did not produce the wild types (r^+ recombinants) in crosses with a number of known deletion mutants. A known deletion mutant is one, with a loss of a segment of DNA whose position on the *rII* map is known. The *rII* deletion mutants that Benzer studied, differed widely in the extent and deletion of genetic material

We shall now discuss how Benzer identified the position of deleted base pair of an unknown point mutant. An unknown *rII* point mutant was first crossed with each of the seven standard deletion mutants that define seven main segments of the *rII* region. These are segments *A1 – A6*, and *B* (see Fig. 13.9). For example, if a *rII* point mutant is crossed.



g.13.9: The *rII* region phage *T4* showing segmental subdivisions. (a) the whole *T4* genetic map. (b) Seven deletions define seven segments of the *rII* region. (c) Three deletions define four subsegments of the *A5* segments. (d) Three deletions define four subsub-segments of the *A5c* subsegment. (e) The order and spacing of the sites of the *rII* mutations in the *A5c2a2* subsegment, as established by pairwise crosses of seven point mutants. (f) The model of DNA double helix to indicate the approximate scale of the map shown in (e).

/ith the seven standard deletion mutants, the wild-type recombinants (r^+) are only formed ı the crosses with mutants, say *rA105* and *r638* defining deletion *A6* and *B* respectively, ut not with the other five deletions (see Fig. 13.9b). The point mutation must be in egment *A5*. That means the r^+ recombinants can only be produced in crosses with eletions if the deleted segment does not include the region of DNA containing the point ıutation. Since the r^+ recombinants were not formed in crosses involving mutants with ıe deletion *r1272, r1241 , rJ3 , rPT1* and *rPB242*, then the point mutation must be in the egment of DNA that these mutants lack. In the example that we have taken above, the ıct that r^+ recombinants are produced with *rA105* and *r638* indicates that the point ıutation must be in the *A5* region, which is not missing in either deletion mutant.

s the main segment in which the mutation occurred was known (see Fig. 13.9b), the point ıutant was crossed with each of the relevent secondary set of reference deletions *r1605, 1589* and *rPB230* (see Fig. 13.9c). For instance, in the segment *A5* there are three

deletions that divide *A5* in four segments namely *A5a, A5b, A5c* and *A5d*. Again crosses were made using the *A5rII* mutants with the secondary set of deletions. The presence or absence of r^+ recombinants in the progeny enabled Benzer to more precisely localise the mutation to a smaller region of the DNA. For example, if the mutation was in segment *A5c*, then r^+ recombinants will be produced with standard deletion mutant *rPB230* but not with the other two standard deletion mutants, that is *r1993* and *r1695*. Similarly other deletion mutants defined smaller regions of each of the four sub-segments, i.e., *A5a, A5b, A5c, A5d*. Let us understand this with the example of *A5c*. It was divided into *A5c1, A5c2a1, A5c2a2 and A5c2b* by deletions *r1993, r1695* and *r1168* (see Fig.13.9d).

Thus, in the three sequential sets of crosses of point mutants with deletion mutants Benzer was able to localise any given *rII* point mutant to one of 47 regions defined by the deletions as indicated in Fig.13.10. Then all those point mutants within a given region could be crossed in all possible pairwise crosses to construct a detailed genetic map. In this way Benzer used the odd 3000 *rII* mutants to prove that the *rII* region is subdivisible into more than 300 mutable sites (nucleotide pairs) that were separated by recombination (see Fig. 12.9 in Unit 12.) The distribution of mutants is not random, certain sites (called hot spots) are represented by a large number of independently isolated mutants.

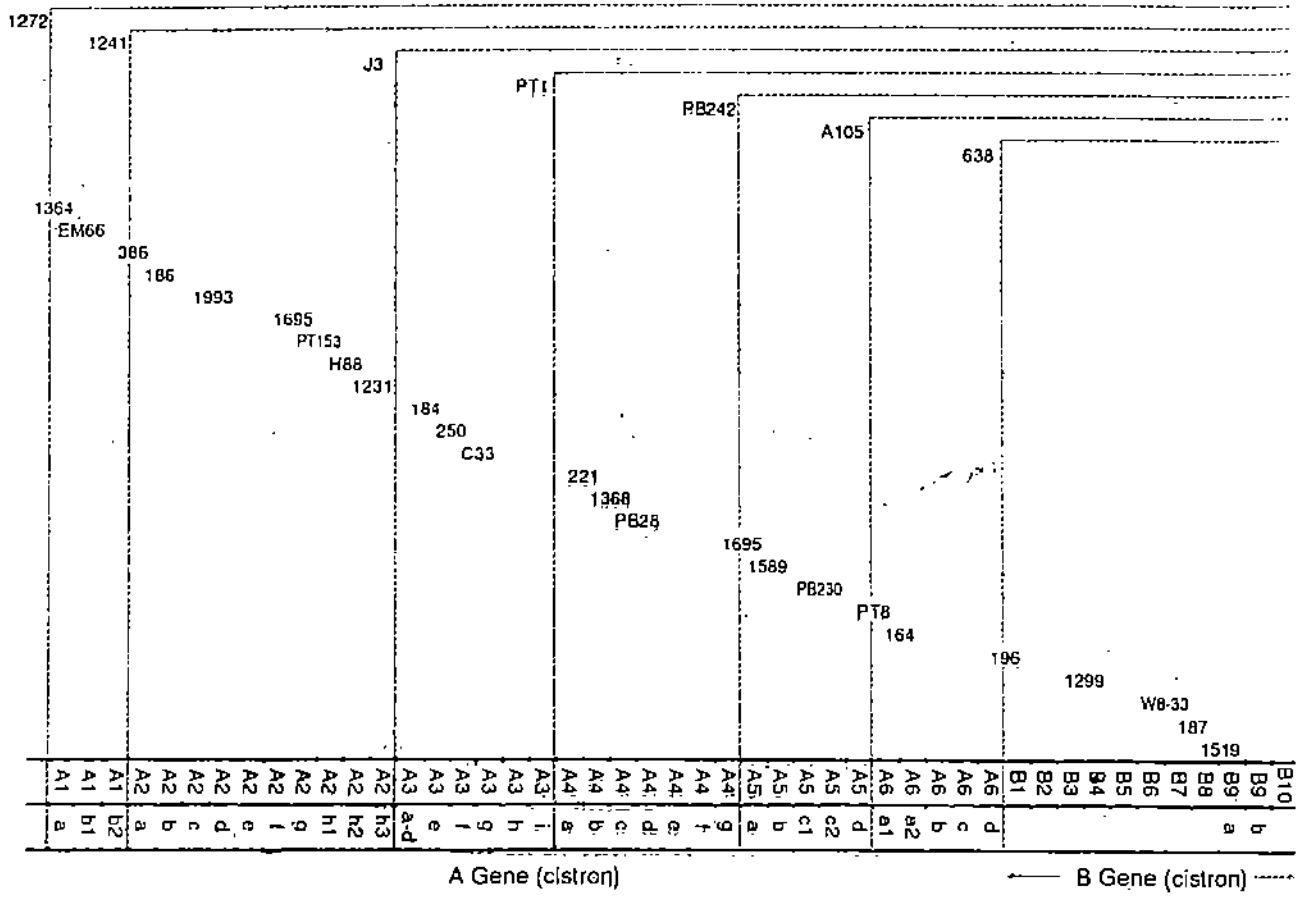


Fig.13.10: Map of *rII* deletions that subdivide the locus into forty-seven segments. The two genes (*rIIA* and *rIIB*) are divided into forty-seven segments (*A1a* through *B10*) according to the ends of the deletions denoted by heavy horizontal bars. The names of the deletion mutants are given on the left of the bars. Some ends do not define segments; these are not connected to a vertical line. Major divisions are denoted by thick lines, subdivisions within these are represented by thinner lines.

To sum up, Benzer's detailed analysis of the *rII* region of the bacteriophage *T4* indicated the spatial relationship between the sites in the gene. It emerged clearly that the unit of mutation and recombination is the base pair in DNA. This showed beyond doubt that the *gene was divisible*. This definition (that the gene was indivisible by mutation and recombination) replaced the classical view.

SAQ 3

Match the items of cloumn I with that of column II.

I	II
i) By employing the deletion mapping technique	a) are seven in number and are designated as *A*1 – *A*6 and *B*.
ii) A deletion mutant is one	b) the order of a group of deletion of a locus could be determined.
iii) The position of an unknown point mutant could be identified	c) the deleted segment does not include the region of DNA with point mutation.
iv) Standard deletion segments	d) with a loss of DNA segment whose position in the map is known.
v) Wild-type (r^+) progeny are produced only in crosses with deletions where	e) based on the presence or absence of r^+ recombinants.
vi) Benzer could precisely locate even a small mutation	f) by crossing with standard deletion mutants.

13.4 APPLYING COMPLEMENTATION TEST TO THE *rII* LOCUS

From the classical point of view, gene is a unit of function, that is, the unit of genetic material that controls the inheritance of one "unit character" or one attribute of phenotype. In molecular terms, it refers to a sequence of nucleotide pairs in DNA that specifies the sequence of amino acids in functional proteins coded for by the gene. Benzer, designed genetic experiments to determine whether this classical view was true for the *rII* region also. He applied the cis-trans or the complementation test to find out whether two different *rII* mutants belonged to the same gene (unit of function).

Before we proceed further, it would be useful to recapitulate briefly the complementation test and see how it can be used to provide insights into gene structure and function.

To test for complementation, two mutations of independent origin that produce a similar phenotype are introduced by mating, into a single cell or organism so that they are in a diploid condition. The phenotype of the resulting cell or organism carrying the mutations is evaluated. If the phenotype is normal, it means the mutations complement each other (also see the adjacent margin remark). From this we conclude that mutations are on two different genes and are nonallelic. On the other hand, if the phenotype is mutant, the two mutations are non-complementing. From this we can conclude that the same gene is involved, and the mutations are allelic. Let us examine the molecular bases of these conclusions.

As an aid to understand the complementation concept, consider the following analogy. There are two factories in a town that make the same kind of automobile. Each has two assembly lines, one for the engines and one for the bodies. One day the engine line of one factory and the body line of the other factory break down. Nevertheless, the two factories are able to continue producing cars by cooperation. The workers put the engines from the second factory together with the bodies from the first. This is complementation, and you can see that it depends on the portable nature of the product. On the other hand, if the engine assembly lines of both the factories cannot complement each other, there is no way to continue making cars.

Complementing Mutations are the ones in which two mutations result in similar phenotype, but are present on different genes. These can be determined if they are introduced in a cell, so that they are in a diploid condition (Fig.13.11a). If the mutations are in two different genes, the 'diploid cells' should also contain non-mutant forms of both genes. A double heterozygote results and provided the mutant conditions are recessive, the phenotype of the cell or the organism is normal. The two mutations complement each other. Although neither mutant gene can direct the formation of a normal polypeptide, there is complementation because the transcripts of the normal forms of both genes direct the synthesis of normal ploypeptides. Hence a normal phenotype results. Compare the mechanism described above with the example discussed in the margin remark.

Non-complementing Mutations –This refers to two indepently isolated mutations that produce like phenotypes (Fig.13.11b) . Each mutant is the result of a change in the base sequence at separate position or **mutation site** within one gene. Since both the mutations involve the same gene, only mutant forms of that gene are present when a diploid is formed. And this implies that transcription using either allele of the gene as a template must lead to the formation of an abnormal or defective polypeptide. Any such condition

indicating the absence of complementation means that the mutations involve the same gene, that is, they are allelic.

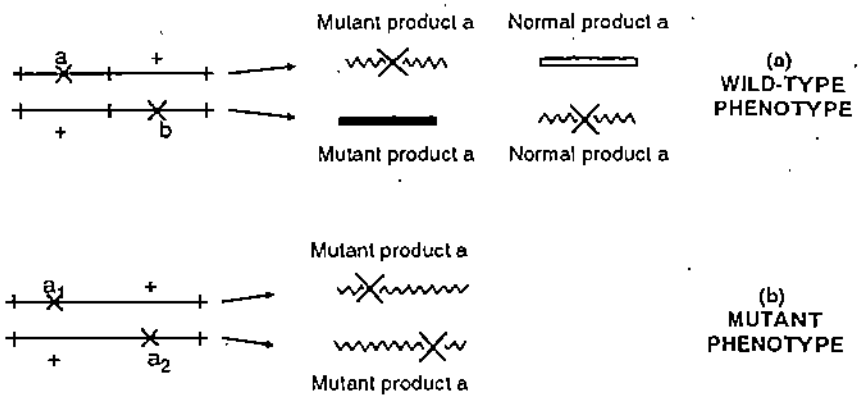


Fig.13.11: Expected results of complementation tests. a) Formation of wild-type phenotype and b) the mutant phenotype.

The complementation tests not only suggest whether two mutants are allelic, but also help us to find out the number of genes that produce a phenotype. And in conjunction with mapping studies, it enables one to define gene boundaries too. Let us understand it more clearly with the help of an example. The prime requirement for carrying out the complementation test is to unite pairs of mutant genomes to produce a diploid condition. Usually, pairs of different mutations sharing a similar phenotype are tested. By determining which of the mutants complement each other, one can know, how many separate genes are involved in the production of that particular phenotype.

Benzer's work in the late 1950s and early 1960s on *rII* mutants is an excellent example of joint mapping and complementation studies. We shall now discuss it briefly. To perform complementation tests, Benzer infected the non-permissive host *E.coli K* cells simultaneouly with two *rII* mutant phages of independent origin, that is, *rIIA* and *rIIB* (see Fig 13.12). Simultaneous infection of a bacterium with two phage genomes, creates a *diploid condition* like that of an F_1 cell, because the two phage genomes exist together in the same cell. In situations where the *rII* mutants cannot produce plaques on *E.coli* strain *K* indicates that both mutants have been derived from changes in a single gene (Fig. 13.12b). But if the *rII* locus has more than one gene, some of the test pairs would involve non-allelic mutants. Such pairs will produce progeny phages, lysis of infected bacteria, and ultimately plaques on bacterial lawns (see Fig. 13.12a).

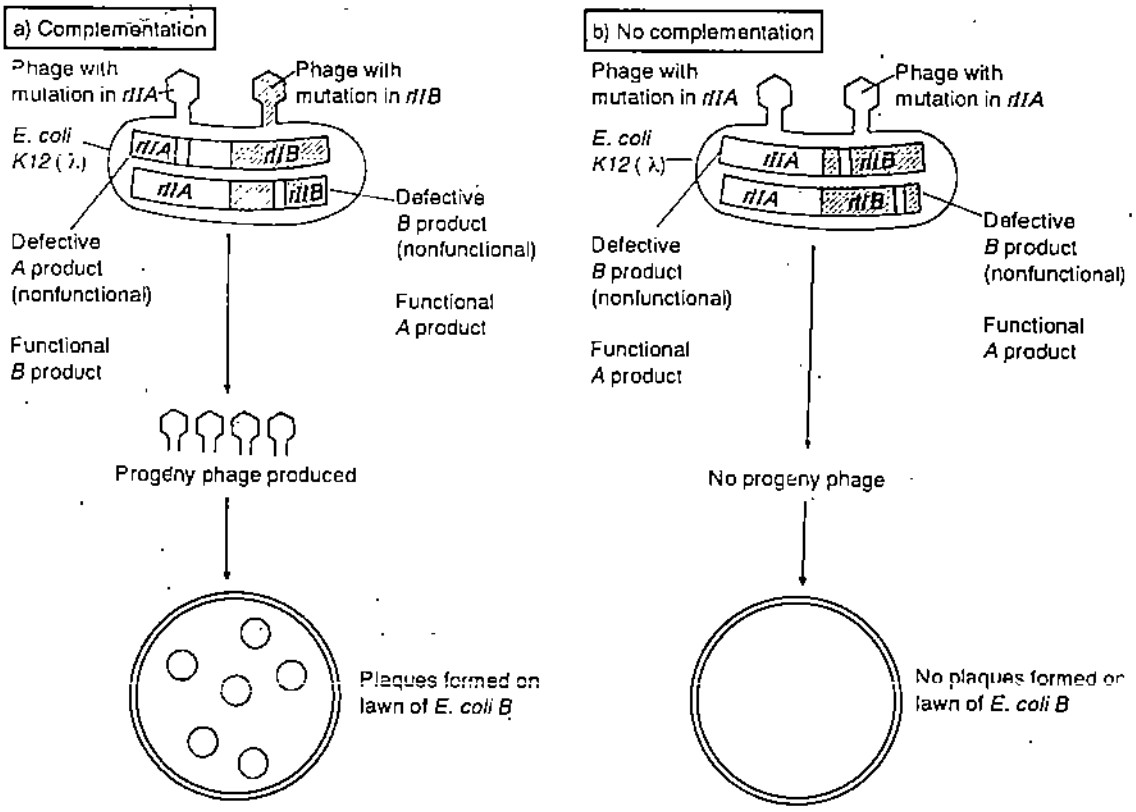


Fig.13.12: Tests for complementation of *rII* mutants of *T4* phage using *E.coli K*. The non-permissive host *K12* (λ) is infected with two different *rII* mutants. a) complementation occurs; b) complementation does not occur.

Benzer found that most of the mutants that he studied could be categorised into two distinct classes or complementation groups. i) One class consists of all *rIIA* mutants complementing all *rIIB* mutants; and ii) The second class encompasses all those that fail to complement each other, e.g., the *rIIA* mutants fail to complement other *rIIA* mutants and *rIIB* mutants fail to complement other *rIIB* mutants. These findings indicated that the two groups of mutants originated from separate modifications in two different genes at the *rII* locus. We shall refer to the gene for the first group as 'gene A' and that of the second group as 'gene B'.

From the discussion so far it emerges that the mutation sites of gene *A* alleles occupy one region at the *rII* locus and the mutation sites of the gene *B* alleles occupy a neighbouring, non-overlapping region at the locus. If it is so, then in a genetic map, gene *A* mutation sites should map together and gene *B* mutation sites should map together; and there should be no intermixing of each gene's sites. The genetic map of different alleles of *rII* region (see Fig.13.13) shows that all mutation sites for gene *A* alleles map on one side of the locus and all mutation sites of gene *B* alleles map on the other side. The map positions of the mutation sites of the two groups do not intermix. The distribution of the mutation sites into two separate but contiguous complementation groups is consistent with Benzer's hypothesis that two genes exist side by side at the *rII* locus. These studies also pinpoint the boundary between the *A* and *B* genes somewhere between the rightmost mutation site of the *A* gene and the leftmost site of the *B* gene (see Fig. 13.13). Also note in the same figure that the most widely separated mutation sites are close to the outside boundaries of the *A* and *B* genes.

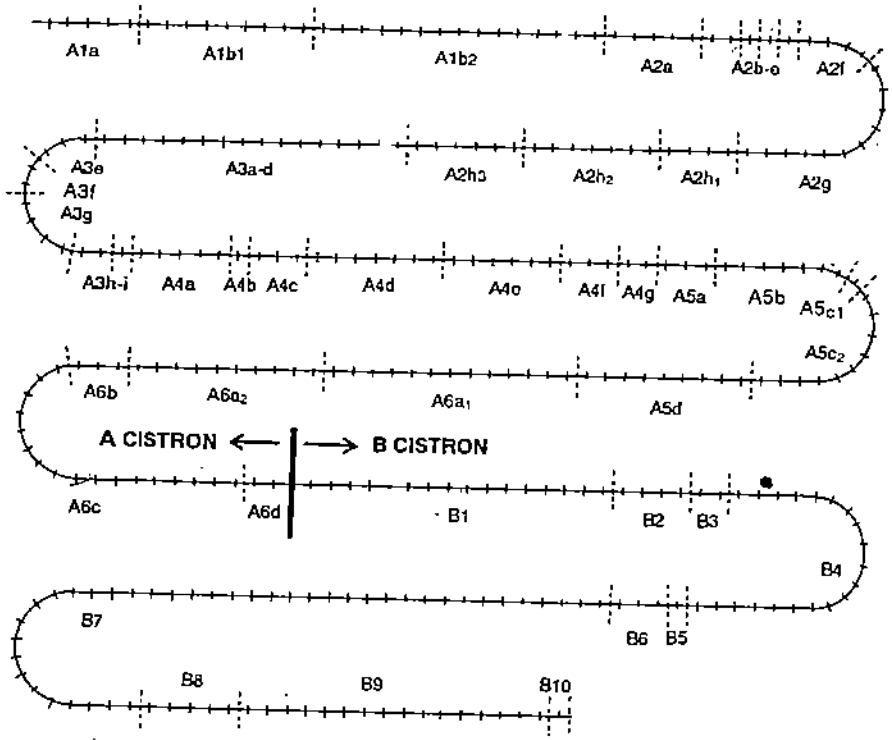


Fig.13.13: Map of the *rII* region of the phage *T4*. Each small line dividing the entire chromosome length indicates the region where a group of mutational sites has been mapped. Each segment such as *A6b*, *B8* etc., refers to a region defined by deletion mapping. (After Benzer, Proceedings of the National Academy of Sciences, U.S. Vol. *47*, p.403, 1961).

lenzer also observed that in a diploid condition two alleles with different mutational sites ı a single gene produce a mutant phenotype when the sites are on opposite chromosome ı the *trans configuration*, (see Fig.13.14a). When the two mutation sites are carried in the iploid condition on a single chromosome, as shown in Fig.13.14b, that is, in the *cis onfiguration*, the phenotype of the cell is normal. In the trans configuration neither copy f the gene can direct the synthesis of a normal polypeptide (see Fig.13.14a). When both ıe mutation sites are in the cis configuration, the other copy of the gene has a normal base :quence and can direct the formation of a normal polypeptide.

enzer referred to the genetic unit of function revealed by the cis-trans test as the *cistron*. t that time (i.e., around 1955) this term was considered as a genetic synonym of gene. esently, the term gene is more commonly used and cistron is being used less.

Nonetheless, gene and cistron are equivalent in referring to the genetic unit of function. In the light of the above the functional units *A* and *B* can also be referred to as *rIIA* and *rIIB* cistrons or genes. It has been observed that the *rIIA* cistron is about 6 map units and has 800 base pairs, and the *rIIB* cistron is about 4 map units and 500 base pairs.

SAQ 4

Provide brief answers to the following.

i) What was Benzer's objective of applying cis-trans test to the rII locus?

.....

.....

ii) What information does one derive from the results of complementaion tests?

.....

.....

iii) What is the minimal requirement for cis-trans test?

.....

.....

iv) In the light of complementation test state whether the given statement is complete? The statement is — "The terms 'Gene' and 'Cistron' can be used interchangeably". Or would you suggest some modification.

.....

13.5 THE ULTIMATE IN FINE STRUCTURE MAPPING

We have seen how recombination can serve as a useful tool for mapping genes. On the other hand, it is also possible to do genetic mapping without recombination. Though it requires more labour than the traditional recombination mapping, but it provides considerably more information. In fact, the ultimate in fine structure genetic mapping is to obtain the nucleotide-pair sequence complete with the information of all nucleotide-pair changes that after the function of that gene or chromosome. Prior to 1975, the thought of trying to sequence entire genomes or chromosomes was considered an uphill task, requiring years of work. By the end of 1976, Frederick Sanger and his colleagues sequenced the entire 5,387 nucleotide long chromosome of a tiny phage $\phi X174$. Today the entire chromosomes of several viruses, including the complete 48,502 nucleotide-pair sequence of the phage λ chromosome, and segments of several eukaryotic chromosomes tens of thousands of nucleotide-pairs in length have been sequenced. Within the next few years, the nucleotide sequences of hundreds of genes and even the entire eukaryotic chromosomes would be determined and the information would be stored in computer data banks for future reference. The goal of the 'Human Genome Project' (inaugurated in 1988) is to sequence the entire human genome (having about 1,00,000 genes) in the coming decade or two.

What kind of information can we gather from the nucleotide sequence data? First, we can locate exactly the coding regions of all the genes. This tells us the spatial rlationships among genes and the distances between them, to the exact nucleotide, without any guesswork about recombination frequency. Then how we recognise a coding region, may be your next question. It contains an *open reading frame*, i.e., a sequence of bases which if translated into one frame, contains no stop codons for a relatively long distance—long enough to code for a polypeptide. In addition, an open reading frame must start with an ATG (or occassionally a GTG triplet, corresponding to an AUG or (GUG) translation initiation codon. An open reading frame, therefore, is the same as a gene's coding region.

Once the base sequence of DNA is known, from it one can also find out the amino acid sequence of a particular protein. All we have to do is to use the genetic code to translate the DNA base sequence of each open reading frame into the corresponding amino acid sequence. This sounds like a laborious process, but there are computer programs available that perform this task in a spilt second.

Let us consider a specific example, that is, phage $\phi X174$. Sanger et al analysed the open reading frame of this phage's DNA. Many new fascinating facts, not known earlier, came into light. You would study them in the Section 13.6 below.

The ultimate fine structure mapping studies, in the times to come would transform biological research considerably. It may be that on one hand many puzzling dilemmas would be resolved and on the other hand new, unexpected genetic areas would be thrown open for research. In human context, by knowing the chromosomal positions of genes for hundreds of genetic illnesses, the isolation of such defective DNA segments may be possible, the study of this aspect may help us know their corrective measures at the molecular level of the gene. It is hoped that several persons suffering from diseases such as muscular dystrophy and cystic fibrosis may be cured.

13.6 OVERLAPPING GENES

In this section you would study an exception to the general picture of the gene, that is, overlapping gene also reffered to as gene-within-gene. These were first discovered in small bacteriophages like $\phi X174$ and *G4*. The bacteriophage $\phi X174$ is a small virus that stores the genetic information in a single-stranded circular molecule of DNA. The phage $\phi X174$'s genome has been found to contain nine genes coding for nine different polypeptides. All these nine gene products were characterised and also their molecular weights were determined. From the known molecular weights of the nine polypeptides, the length of DNA required to code for all the above polypeptides was calculated, based on the fact that genetic code is a tripet code (i.e., three nucleotides per amino acid). The minimum length of DNA required to code for the nine polypeptides was found to be 6,078 nucleotides. But the $\phi X174$ genome has been found to have only 5,387 nucleotides.

And this number of nucleotides, as estimated, is not enough to carry the coding sequence of all nine polypeptides if there is nonoverlapping condition in genes. For several years (until 1976), this remained a puzzling dilemma. And this was resolved when the complete nucleotide sequencing was determined. This is quite a feat in itself, the sequencing of the entire hereditary material of a genetic system. The question that did arise then was how to account for the disparity in the number of nucleotide sequences obtained for the $\phi X174$ genome. In the first numerical figure (of 6078 nucleotides) the amount of DNA required to code for the nine proteins is far in excess of the DNA actually present (5387 nucleotides).

The research findings of Frederick Sanger and his colleagues provided answer for this question. They studied in detail several genetic regions of $\phi X174$, one of which contains two genes *D* and *E*, which code for proteins *D* and *E* respectively. The protein *D*, whose exact amino acid sequence was known, is produced in great amounts in the host cell and is required for the production of single-stranded viral DNA. The *E* protein is required for lysis of the host cell when hundreds of new viruses have been produced and are to be released. The two genes *D* and *E*, in all counts seem like any other separate and distinct genes.

An exciting finding was made that a certain stretch of the phage DNA contains the coded information for both the *D* and the *E* proteins and that the *D* and *E* genes overlap (see Fig. 13.14 and 13.15). The start of the *E* gene is in the middle of the *D* gene. The ends of the both the genes close to the same spot. However, the *E* gene is not just a segment of the *D* gene (see Fig.13.14). The two proteins, *D* and *E* are very different in amino acid sequence. The DNA, however, in the *D* and *E* genetic region is read in different frames (see Figure 13.14 again).

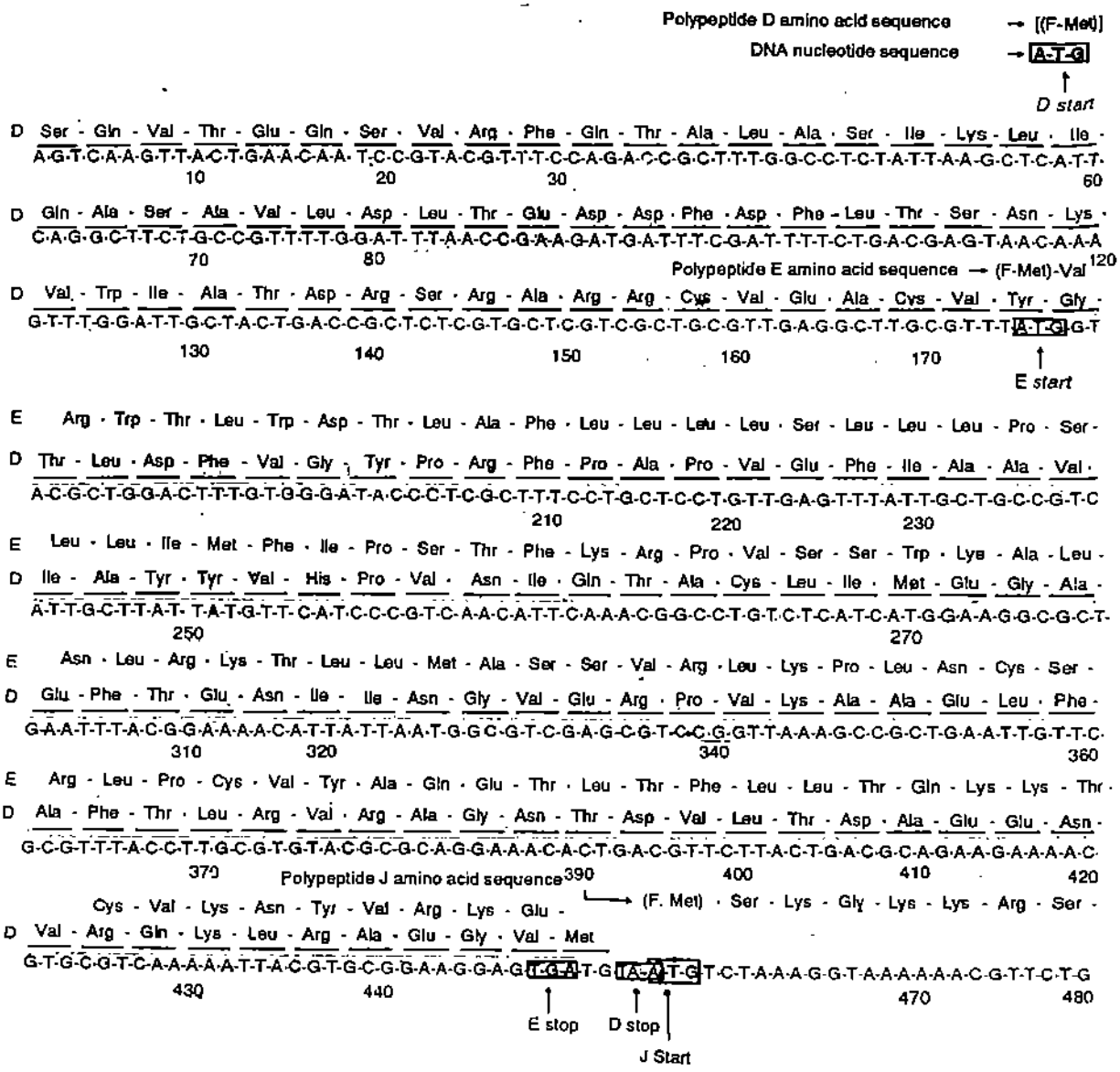


Fig.13.14: A segment of $\phi X174$ genome, showing overlapping genes. The figure above shows the coding sequence of gene *E* which is located within the coding sequence of gene *D*. In the overlapping genes, codons are read in different frames. Note, the reading frame of gene *E* is offset to the right by one nucleotide from the reading frame of gene *D*. And also there is overlapping between the terminating codon of gene *D* (TAA) and start codon of gene *J* (ATG). The DNA triplets shown here have the same sense as the RNA triplets of the messenger RNA, because, the strand shown here is not the one that undergoes transcription; but is the antisense strand. Since it has the same sense as the mRNA, the mRNA sequence can be

Another point to be noted is that the termination-codon of the *D* gene overlaps by one nucleotide to the start codon of the next gene, i.e., gene *J*. Moreover, overlapping of reading frames does not seem to be confined to just this one region of the phage. The gene *A*, which contains coded information for another protein needed for viral construction, overlaps with gene *B*, coded for a protein that produces a nick in the DNA when new copies of DNA are needed.

These surprising results have important genetic implications. Clearly, the amino acid sequences and genes with overlapping coding sequences cannot evolve totally independent of one another. They can have some independence due to the degeneracy of the genetic code. Moreover, a single mutation can result in the loss of two gene product activities.

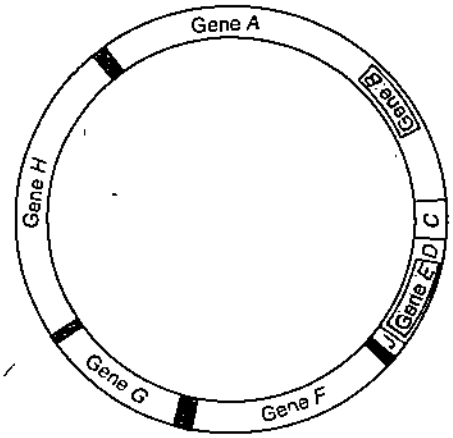


Fig.13.15: The genetic map of phage $\phi X174$ showing two cases of overlapping genes. The coding sequences of genes *B* and *E* are located entirely within the coding sequences of genes A and D respectively. The four, short, noncoding segments of the chromosomes are shown in grey colour. (Based on data of Sanger et al 1977, Nature *265*: 687, and Tessman et al 1980, Journal of Virology *33*: 557).

verlapping genes have also been studied in detail in a *simian virus 40 (SV40)*. As stated rlier, an overlapping gene is an exception to the general picture of the gene. Earlier, the enetic code was considered to be non-overlapping, with each single nucleotide being part f only one code word or codon. Exceptions to these have been found in $\phi X174$ and *V40*, where a single nucleotide may be a part of two code words depending on the ading frame. The finding of the overlapping genes helped solve the problem of fitting the nes of $\phi X174$ into a supposedly less than adequate length of DNA. It is believed that the verlapping stet may have evolved in viruses as a device to accomodate a larger number of nes than would otherwise be possible in so small an amount of DNA.

AQ 5

atch the items of column I with that of column II.

I	II
Frederick Sanger	a) 1988, sequencing of the entire genome
Nucleotide sequence data	b) sequencing and analysis of $\phi X174$ genome
) Coding region	c) the ultimate in fine structure mapping
) Human Genome Project	d) open reading frame
Genetic mapping without recombination	e) spatial relationship of genes

AQ 6

Find the longest open reading frame in this sequence of bases:

CCAGATGCCTAAATGAGTTGGCCAGCAGAGCGAGCATGGATGTAATCAG

What can be some of the genetic implications of overlapping genes or "genes-within-genes"?

.....

.....

.....

c) What is the maximum number of different amino acid sequences that can be produced

i) from the same segment of one strand of DNA?

.....

ii) From both strands of a DNA double helix?

.....

13.7 SUMMARY

In this unit you have been that :

- Bacteriophage such as *T4*, has been found to be an invaluable genetic system that has provided deeper insights into the fine structure.
- Fine structure mapping, including deletion mappings makes it possible to determine the order of mutation sites within a gene, and in conjuction with the complementation tests to identify the genetic boundaries of a gene.
- The nucleotide sequences of genes, alongwith the data on any nucleotide changes that alter gene function, constitute the ultimate fine structure maps.
- The nucleotide sequencing of the phage $\phi X174$ has revealed the occurrence of overlapping genes. Of the 9 genes of this phage, two are located entirely within the coding sequence of two different genes.

13.8 TERMINAL QUESTIONS

1) In what ways does the present concept of the gene differ from the pre-1940 or classical concept of gene?
2) What is the currently accepted operational defination of the gene?
3) Of what value are the conditional lethal mutations for genetic fine structure analysis?
4) a) How many plaques can be formed by a single bacteriophage particle?

   b) A bacteriophage adsorbs to a bacterium in a liquid growth medium. Before lysis, the infected cell is added to a suspension of cells that are plated on solid medium to form a lawn. How many plaques will result?
5) Phage *T2* (a relative of *T4*) normally forms small, clear plaques on a lawn of *E.coli* strain *B*. Another strain of *E.coli*, called *B/2* is unable to adsorb *T2* phage particles, and no plaques are formed. *T2h* is a host-range mutant capable of absorbing to *E.coli B* and to *B/2* and it forms normal looking plaques. If *E.*coli *B* and mutant *B/2* are mixed in equal proportions and used to produce a lawn, how can plaques made by *T2* and *T2h* be distinguished by their appearance?
6) What is the difference between a pair of homoalleles and a pair of heteroalleles?
7) Are the following statements concerning the genetic element referred to as the gene, true or false?

   a) The classical (pre-1940) concept of gene was:

      i) a unit of physiological function or expression; ii) the smallest unit that could undergo mutation; and iii) a unit not subdivisible by recombinaion.

   b) The cis-trans test provides an operational definition by which we can often identify a gene as the unit which specifies one mRNA molecule.

   c) Our present knowledge of the gene indicates that the units defined by criteria (ii) and (iii) in statement (a) above are both equivalent to a single nucleotide-pair.

   d) Studies of the 1940s demonstrated the existence of heteroalleles (pseudoalleles) clearly indicating that many mutations which were allelic by the functional criterion could be separated by recombination, thereby indicating that the units of function, mutation and recombination are *not* equivalent.

e) Homoalleles are functionally and structurally allelic, heteroalleles are functionally allelic but structurally nonallelic.

8) What is the correnr answer for a and b?

a) If we want to know whether two different *rII* point mutants lie at exactly the same site (nucleotide pair), one should :

i) coinfect *E.coli K12* (λ) with both mutants. If phage are produced they lie at the same site.

ii) coinfect *E.coli K12* (λ) with both mutants. If phage are not produced, they lie at the same site.

iii) coinfect *E. coli B* with both mutants and plate the progeny phage on both *E.coli B* and *E.coli K12* (λ). If plaques appear on *B* but not on *K12* (λ) they lie at the same site.

iv) coinfect *E.coli K12* (λ) with both mutants and plate the progeny phage on both *E.coli B* and *E.coli K12* (λ). If plaques appear on *K12* (λ) but not *B*, they lie at the same site.

b) If one wants to know whether two different *rII* point mutants lie in the same cistron, one should :

i) coinfect *E.coli K12* (λ) with both mutants. If phage are produced they lie at the same cistron.

ii) coinfect *E.coli K12* (λ) with both mutants. If phage are not produced they lie in the same cistron.

iii) coinfect *E. coli B* with both mutants and plate the progeny phage on both *E. coli B* and *E. coli K12* (λ). If plaques appear on *B* but not *K12* (λ). they lie in the same cistron.

iv) coinfect *E.coli K12* (λ) with both mutants and plate the progeny phage on both *E.coli B* and *E.coli K12* (λ). If plaques appear on *K12* (λ) but not *B*, they lie in the same cistron.

Construct a map from the following two-factor phage cross data indicating map distances.

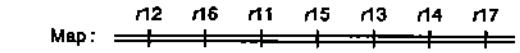
	% recombination
$r_1 \times r_2$	0.10
$r_1 \times r_3$	0.05
$r_1 \times r_4$	0.17
$r_2 \times r_3$	0.15
$r_2 \times r_4$	0.10
$r_3 \times r_4$	0.32

Five different *rII* deletion strains of phage *T4* were tested for recombination by pair-wise crossing in *E.coli B*. The following results were obtained, where $r^+ = r^+$ recombinants produced, and 0 = no r^+ recombinants produced :

	A	B	C	D	E
E	0	+	0	+	0
D	0	0	0	0	
C	0	0	0		
B	+	0			
A	0				

Draw a deletion map compatible with the above data.

Given the following map with point mutants, and given the data in the table below, make a topological representation of deletion mutants *r21*, *r22*, *r23*, *r24*, and *r25*. Indicate clearly the endpoints of the deletions (+means r^+ recombinants are obtained. 0 means r^+ recombinants are not obtained).



Deletion Mutants	Point Mutants r11	r12	r13	r14	r15	r16	r17
r21	0	+	0	+	0	+	+
r22	+	+	0	0	+	+	0
r23	0	0	0	+	0	0	+
r24	+	+	0	0	+	+	+
r25	+	+	0	0	0	+	+

12) The *T4rII* deletions chosen by Benzer to delineate the *rII* region into 47 subregions (Fig. 13.10) were chosen from a much larger set of *rII* mutations. The deletions can be ordered with respect to one another by pair-wise crosses in which only the presence or absence of wild-type recombinants is scored. The table below records the results of such crosses employing 45 of Benzer's nonreverting *rII* mutations. Arrange these mutations in order on a genetic map. Assume that each mutation deletes only one contiguous region; that is, there are no double mutations in the group (***Hint*** : Arrange the larger deletions first).

13) Three mutations were discovered with a similar phenotype. In complementation testing, mutants 1 and 2 were shown to be allelic but neither was allelic to mutant 3. Complete the table given below indicating + for positive complementation an – for negative complementation.

Mutants	condition	
	cis	trans
1,2		
1,3		
2,3		

14) In complementation studies of the *rII* locus of phage *T4*, three groups of three different mutations were tested. For each group, only two combinations were tested. On the basis of each set of data, predict the results of the third experiment.

Group A	Group B	Group C
d × e — lysis	g × h — no lysis	j × k — lysis
d × f — no lysis	g × i — no lysis	j × l — lysis
e × f — . . . . ? . . .	h × i — . . ?	k × l — . . . ? . .

15) In an analysis of other *rII* mutants, complementation testing yielded the following results :

Mutants	Results
1,2	+
1,3	+
1,4	¾
1,5	¾

a) Predict the results of testing 2 and 3; 2 and 4; and 3 and 4 together.

b) If further testing yielded the following results what would you conclude about mutant 5?

Mutants	Results
2,5	–
3,5	–
4,5	–

13.9 ANSWERS

Self Assessment Questions

1) i) plaque

ii) restrictive, permissive

iii) lysogenic

iv) lytic

v) virulent

vi) temperate

2) i) The intergenic mapping refers to mapping of the distance between different genes. Whereas intragenic mapping involves the various sites within a particular gene.

ii) Firstly, because it produces large number of progeny phages in manageable space and short time. Secondly, it has distinct plaque morphology. This makes it possible to detect even very low recombination frequencies.

iii)

Phage	Phenotype when plated on	
	E.coli B (permissive host)	*E.coli K12* (λ) (non-permissive host)
T4 wild type (*T4IIr*$^+$)	turbid plaques with fuzzy ends	turbid plaques with fuzzy ends
T4 rII	large clear plaques with distinct edges	No plaques

iv) plaque morphology, host range properties

v) The wild-type phages could grow on both *B* strain and *K12* (λ) strain, but mutants could not grow on the latter. On plating the progeny (obtained from the cross between the mutants) on *K12* (λ) strain, he could detect and quantify the rare wild-type phages produced.

vi) It is quite likely that those nucleotide pairs carried mutations at the same site of recombination. So, in the absence of recombination between mutants no r$^+$ were produced.

vii) A nucleotide pair is the unit of recombination or mutation, and that the gene was subdivisible.

3) i) b

ii) d

iii) f

iv) a

v) c

vi) e

4) i) The main objective was to determine whether the classical view of gene, i.e., a gene is a unit of function, true for the *rII* region also.

ii) We can know : one, if the two mutants are allelic; and two, the number of genes involved in producting a phenotype.

iii) It is to bring together the mutant genomes to produce a diploid condition.

iv) The statement would be complete by specifying that these two terms are equivalent while referring to the genetic unit of function only.

5) i) b

ii) e

iii) d

iv) a

v) c

6) a) ATGAGTTGGCCAGCAGAGCGAGCATGGATGTAA

b) The amino acid sequences of polypeptides that are specified by overlapping genes or genes-within-genes would not be able to evolve independently. However, the degeneracy of the code will allow for some independence. In addition a single mutation within a shared nucleotide sequence could result in nonfunctional products for both genes.

c) i) 3

ii) 6

Terminal Questions

1) Prior to 1940, the gene was considered a "bead-on-a-string" not subdivisible by recombination or mutation. Today, the gene is considered to be a unit of genetic material coding for one polypeptide. And in terms of unit of structure, not subdivisible by recombination or mutation is known to be the single nucleotide-pair.

2) The cis-trans test, which defines the unit of genetic material specifying the amino acid sequence of one polypeptide.

3) They provide powerful selective sieves for identifying rare recombinants. This is accomplished by using the restrictive environmental conditions to select wild-type recombinant progeny. From crosses between pairs of conditional lethal mutants.

4) a) One plaque per phage.

b) One plaque, because the plating was done before lysis, and so the progeny phage are confined to one tiny area.

5) The *T2* plaques will be turbid because *T2* fails to lyse the resistant bacteria. Plaques made by the *T2h* mutant will be clear because the mutant can lyse both the normal and resistant cells..

6) Homoalleles are structuraly and functionally allelic, they are not separable by recombination. Heteroalleles are functionally allelic (based on cis-trans test) but are structurally non-allelic (based on recombination test). Heteroalleles thus result from mutations occurring at different sites within a gene.

7) a) True

b) False

c) True

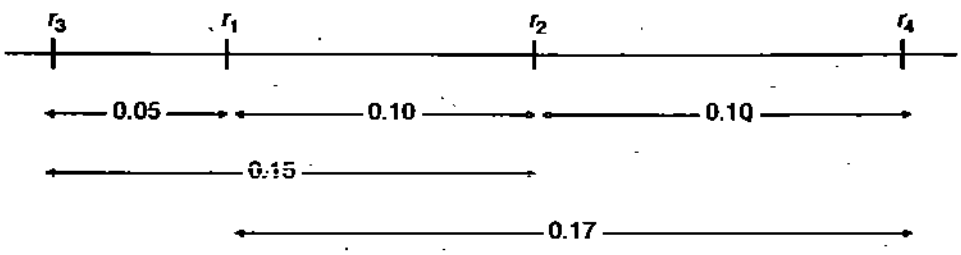
d) True

e) True

8) a) ii

b) i

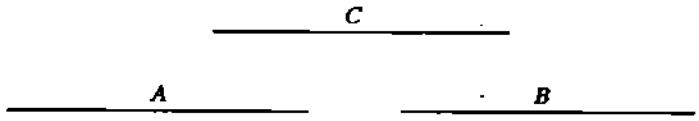
9)



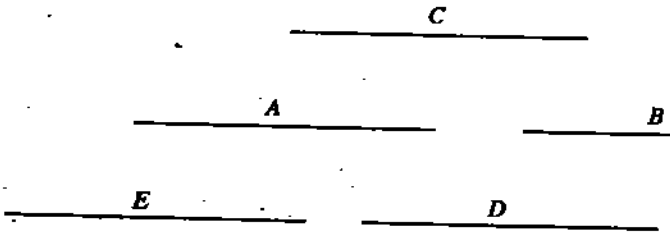
10) This problem is based on the principle that if two deletion mutations overlap, then no r^+ recombinants can be produced. Conversely, if two deletion mutations do not overlap, then r^+ recombinants can be produced. To approach a question of this kind, you should draw overlapping and non-overlapping lines from the given data.

Starting with *A* and *B* these two deletions do not overlap since r^+ recombinants are produced. Therefore, these two mutations can be represented as follows :

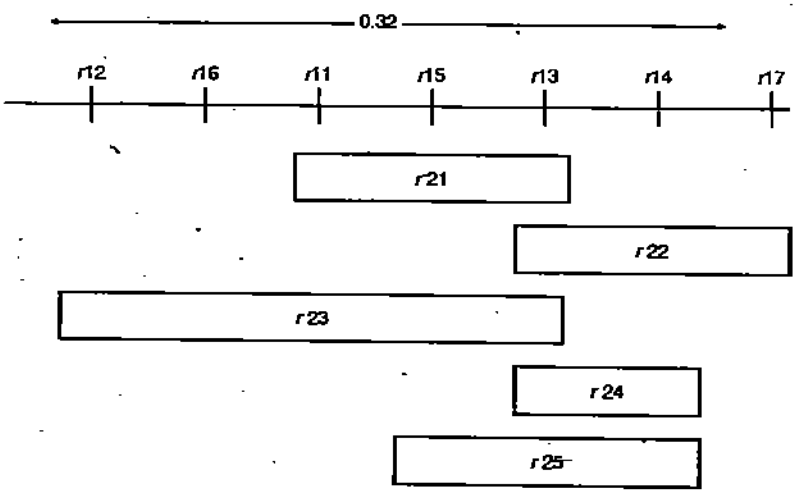
The next deletion *C*, does not give r^+ recombinants with any of the other four deletions. We must conclude, therefore, that *C* is an extensive deletion that overlaps all of the other four, with endpoints that cannot be determined from the data given. One possibility is as follows :



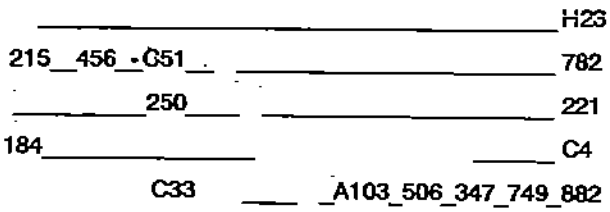
Deletion *D* does not give r^+ recombinants with *A*, *B* or *C* but it does with *E*. In turn, *E* gives r^+ recombinants with *B* and *D* but not with *A* or *C*.Thus *D* must overlap both *A* and *B* but not *E*, and *E* must overlap *A* and *C* but not *B*. *A* map for this situation can be as follows. Other maps can be drawn in terms of the endpoints of the deletions.



11)



12)



13)

Mutants	condition	
	cis	trans
1,2	+	–
1,3	+	+
2,3	+	+

14) Group A : e x f ⟶ lysis

Group B : h x i ⟶ no lysis

Group C : k x i ⟶ no lysis

15) a) 2 and 3 : –

2 and 4 : +

3 and 4 : +

b) Since 5 does not complement any of the other mutations, it is either a double mutation in both cistrons or a deletion which overlaps both cistron.

UNIT 14 REGULATION OF GENE EXPRESSION IN PROKARYOTES

Structure

14.1 INTRODUCTION

Every cell contains enormous amount of genetic information stored in its genes. But all this information is not used at all times, and therefore, not all the genes are expressed continuously. The level of gene expression may vary from one cell type to another or according to the stage in the cell cycle. For example, the genes for haemoglobin are expressed at high levels only in the precursors of red blood cells. The activity of the genes thus varies according to the function of the cell. A complex vertebrate animal contains about 200 different kinds of cells. By and large, all cell types contain the same genetic complement. The cell types differ only in respect of which particular gene is active. In general, the synthesis of particular gene products is controlled by mechanisms collectively called ***gene regulation***. This unit and the next pertain to gene regulation. We have discussed how the activities of genes are coordinated, first in prokaryotes and then in eukaryotes.

In this unit you will study the mechanism by which gene expression is regulated in prokaryotic organisms particularly in bacteria and viruses. We have dealt with in detail, some of the basic regulation mechanisms in bacteria. We shall also discuss how this regulation is related to the organisation of genes in the bacterial genome.

Since some of the viruses – bacteriophages are parasites. They rely on the activities of host bacteria to reproduce. Therefore, they must be capable of directing their own reproductive cycle in a precisely programmed way. This is directed through the regulation of gene expression. The last section gives you an overview of temporal sequences of gene expression during phage infection.

Before commencing your study of this unit, we advise you to revise the sequence of events during protein synthesis. You may refer to Unit 14 of LSE-01. It would help you to understand this and the following unit better. You can appreciate the regulatory processes well when you have a clear picture in your mind about the various components and steps of protein biosynthesis.

Objectives

After studying this unit, you should be able to:

- describe the various control points of gene expression (Section 14.2);
- explain with examples, the phenomenon of enzyme induction and repression in prokaryotes (Section 14.3);
- describe the components of an inducible operon such as lactose operon and explain the functions of the operator and promoter regions (Section 14.4, Subsection 14.4.2);
- describe the structure of a mRNA molecule produced by the lactose operon (Section 14.4, Subsection 14.4.3);
- interpret genetic and other experimental evidence for the components of *lac* operon (Section 14.4, Subsection 14.4.3);
- explain the differences between the positive and negative control, and discuss how both types of control are used in the lactose operon (Section 14.4, Subsections 14.4.5 and 14.4.6);
- describe the components of a repressible operon such as tryptophan operon (Section 14.5, Subsection 14.5.1);
- explain the two levels of regulation of the *trp* operon (Section 14.5, Subsection 14.5.2); and
- outline the temporal sequences of gene expression during phage infection (Section 14.6).

14.2 CONTROL POINTS OF GENE EXPRESSION

Control of gene expression is a highly complex business and may involve variable strategies in different organisms. In many cases, gene activity is regulated at the level of transcription, either through the signals originating within the cell itself or in response to external conditions. Many gene products are needed only in certain situations, and transcription can be regulated in an on-off manner that enables such products to be present only when demanded by external conditions. Besides transcription, flow of genetic information is regulated in other ways also. The possible control points for gene expression are given discussed below.

1) ***DNA Rearrangements*** in which gene expression depends on programmed changes in the position of DNA sequences in the genome.

2) ***Transcriptional Regulation*** of the occurrence or rate of synthesis of the RNA transcripts. If the number of transcripts synthesised per unit of time decreases, the amount of gene product in the cell would also decrease.

3) ***mRNA Processing*** – Events such as capping, polyadenylation and splicing of eukaryotic mRNAs are in most cases prerequisites for translation. If these processing events are slowed, the product synthesis will fall accordingly.

4) ***mRNA Turnover*** – If mRNA molecules are degraded before translation can occur, synthesis of gene product will be limited.

5) ***Translation Control*** – A control could be exerted over the number of ribosomes that can attach to single mRNA, or over the rate at which individual ribosomes translate a message.

6) ***Post-translation Control*** – It includes a great diversity of mechanisms affecting enzyme inhibition, activation, stability and so forth.

The control points of gene expression given above are not merely the theoretical possibilities, but these are, in actual fact, operative in different organisms. However, extensive data indicate that *regulation of transcription is the most prevalent mode of control of gene expression in prokaryotes*. It does not mean that regulation cannot occur at other levels. Regulation at translational level is also important in the overall control of metabolic processes in living organisms. The regulatory mechanism with the largest effects on phenotypes, however, have been shown to act at the level of transcription.

Based on our present understanding of regulation of transcription in both prokaryotes and eukaryotes, the various regulatory mechanisms seem to fit into two general categories.

(i) *The first category includes mechanisms that involve rapid turn-on and turn-off of gene expression in response to environmental changes.* Regulating mechanisms of this type are very important in microorganisms because of their frequent exposure to sudden changes in environment. These provide the micro-organisms with a great deal of plasticity and enable them to rapidly adjust their metabolic processes in order to achieve maximal growth and reproduction under the changed environmental conditions. These quick responding on-off switches seem to be less important to the higher eukaryotes. This may be because the circulating systems of higher eukaryotes buffer their cells against many sudden environmental changes.

(ii) The *second* major category of regulatory mechanisms is referred to as *pre-programmed circuits of gene expression* . In these cases some event, say infection by a virus, triggers the expression of one set of genes. The product(s) of one or more of these genes functions by turning off the transcription of the first set of genes and/or turning on the transcription of a second set of genes. In turn, one or more of the products of the second set acts by turning on a third set, and so on. In these cases, the sequential expression of genes is genetically preprogrammed, and the genes usually *cannot* be turned on out of sequence. Such preprogrammed sequences of gene expression have been well documented in viral infections. In most of these preprogrammed sequences, the circuitry is cyclical. For example, in viral infections some event associated with the packaging of the viral DNA or RNA inside the protein coat resets the program so that the first set of genes are again expressed when a progeny virus subsequently infects another host cell.

14.3 INDUCTION AND REPRESSION IN PROKARYOTES

Certain gene-products, such as tRNA and rRNA molecules, ribosomal proteins, RNA polymerase components (polypeptides), and other enzymes catalysing metabolic processes are expressed continually no matter what the life supporting environmental conditions are. That is because these products are essential for normal functioning of living cells. These are also referred to as cellular 'housekeeping' functions. Genes that specify products of this type, therefore are *continually being expressed* in most cells. Such genes are said to be expressed constitutively and are known as **constitutive genes or housekeeping genes.**

On the other hand, the **regulated genes** have a more specialised role, as their products are needed in certain circumstances only. The activity of the gene thus depends on the needs of the cells or the organism. For example, certain gene products are needed for cell growth only under certain environmental conditions. Constitutive synthesis of such gene-products would waste energy, that could otherwise be utilised for more rapid growth and reproduction under the existing environmental conditions. During the process of evolution, the organisms having developed regulatory mechanisms have a selective advantage over the organisms lacking these mechanisms. This undoubtely explains why presently existing organisms including the primitive bacteria and viruses, exhibit highly developed and very efficient mechanisms for the control of gene expression .

Escherichia coli and most other bacteria are capable of growing using any one of several carbohydrates, such as glucose, sucrose, galactose, lactose etc., as an energy source. *E.coli* can preferentially metabolise glucose. In the absence of glucose, however, they can grow very well on other carbohydrates. Let us consider the example of the sugar lactose. The bacterial cells growing on medium containing lactose as the sole carbon source start synthesising two enzymes β-*galcatosidase* and β-*galactoside permease* that are required for the catabolism of lactose (see Fig.14.1) . A third enzyme β-*galactoside transacetylase* is also synthesised. It's metabolic function, however, is not known. β-galactosidase cleaves lactose into glucose and galacotse. And β-galactoside permease transports lactose from medium into *E.coli* (see Fig. 14.1). Neither of these enzyme is of any use to *E.coli* cells when lactose is not present in its surrounding environments. The synthesis of these two enzymes, of course, requires the utilisation of considerable energy in the form of ATP and GTP. Thus, *E.coli* cells have evolved a regulatory mechanism by which the synthesis of these lactose catalysing enzymes is turned on in the presence of lactose and turned off in its absence.

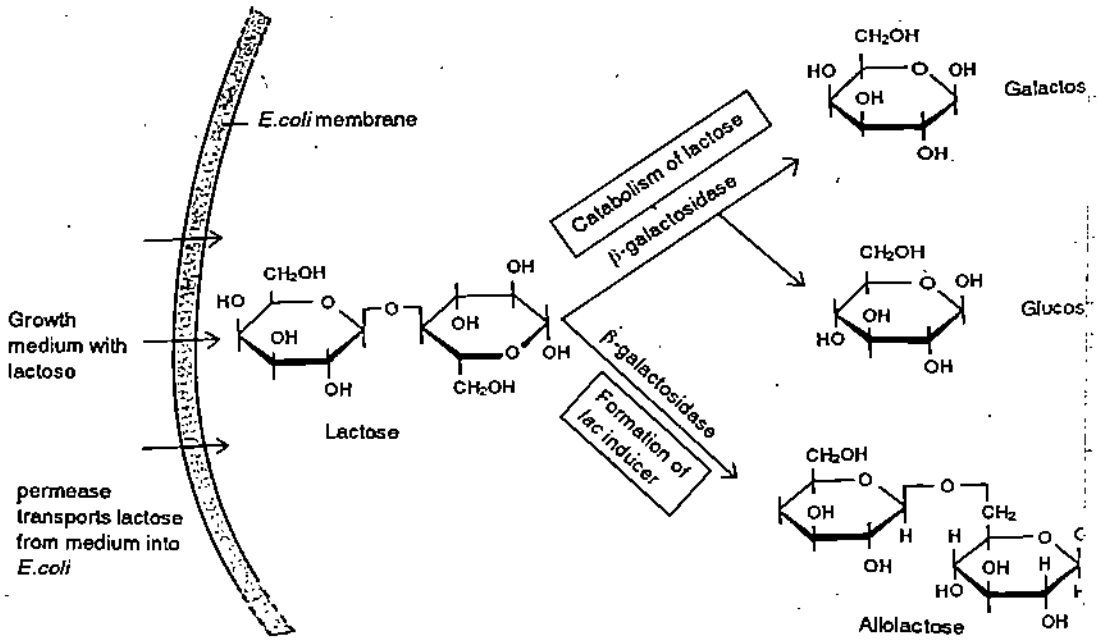


Fig.14.1: Two physiologically important reactions catalysed by the enzyme β-galactosidase. Lactose is either cleaved by β-galactosidase to form glucose and galactose (top), or is converted to allolactose (bottom) which is the inducer of *lac* Operon.

In the natural environments of *E.coli*, that is, in the intestinal tracts and sewers, glucose is generally absent and the presence of lactose is relatively infrequent. Therefore, most of the time, the genes coding for enzymes involved in lactose utilisation are not expressed. On the other hand, if cells growing on a carbohydrate other than lactose are, transferred to medium containing lactose, they rapidly begin synthesising the enzymes required for lactose utilisation (see Fig.14.2a). This process, by which the expression of genes is turned on in response to a substance in the environment is called **induction**. Genes whose expression are thus regulated are called **inducible genes**; their products if enzymes, are called **inducible enzymes**. The substance, or molecules responsible for induction are known as **inducers**.

Enzymes that are involved in **catabolic** or **degradative pathways** such as in lactose and other carbohydrate utilisation are characteristically inducible. It will become clear to you as you go through the following sections of this unit, that *induction occurs at the level of transcription.* It should be kept in mind that it alters the rate of synthesis of enzymes, not the activity of existing enzyme molecules. Therefore, *induction should not be confused with enzyme activation,* in which the binding of a small molecule to an enzyme increases the activity of the enzyme but does not affect its rate of synthesis.

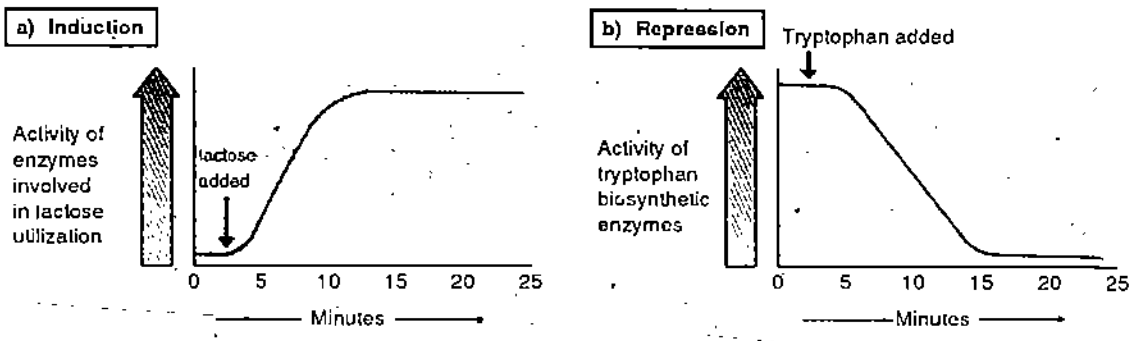


Fig: 14.2: The figures a and b show diagrammatically the phenomenon of induction and repression respectively of enzyme synthesis in bacteria. Induction is a characteristic feature of catabolic (degradative) pathways and repression is of anabolic (biosynthetic pathway).

Figure (a) illustrates the effect of the presence or absence of lactose on the synthesis of enzymes required for lactose utilisation in *E. coli*. Note, in the absence of lactose, the enzyme synthesis is very low. And as lactose is added (see arrow), the synthesis of enzymes required for lactose utilisation are rapidly induced or turned on.

The figure (b) represents the repression of synthesis of enzymes required for biosynthesis of tryptophan in *E.coli*. In conditions when tryptophan is not present, there is an increased synthesis of enzymes. As soon as tryptophan is added (see arrow), there is rapid repression in the synthesis of enzymes.

Bacteria have the inherent capacity of synthesising most of the organic molecules, (such as amino acids, purine, vitamins etc.) required for their growth. For instance, *E. coli* can synthesise an amino acid – tryptophan when not available from its environment. On the other hand, if tryptophan is available in the external milieu then the bacterium readily utilises it and stops its synthesis. This switch-off regulatory mechanism enables the bacterium to save on the energy. *E.coli* possesses five genes that code for the enzymes required in the biosynthesis of tryptophan. These five genes are, therefore, expressed only in conditions devoid of tryptophan. This process of "turning off" the expression of sets of genes is called **repression**. The gene whose expression has been turned off in this way are said to be **repressed**, and when its expression is turned on, the gene is then said to be **derepressed**.

Enzymes that are components of anabolic or biosynthetic pathways are frequently subject to repression, and are known as *repressible. Repression like induction, too occurs at the level of transcription.* Again a word of caution! *Repression should not be confused with feedback inhibition*, in which the binding of an end product to the first enzyme in a biosynthetic pathway inhibits the activity of the enzyme, but does not affect its synthesis.

The next two sections are concerned with two important, and well investigated systems of regulation in prokaryotes. They also serve as an introduction to the next unit which deals with regulation in eukaryotes.

SAQ 1

Write short answers to the following questions in the space provided.

i) What term describes a gene that is expressed continually?

.....

ii) The metabolic pathway for glycolysis is responsible for the degradation of glucose and is one of the fundamental energy-producing systems in living cells. Would you expect the enzymes in this pathway to be regulated? Why or why not?

.....

.....

.....

.....

.....

iii) Explain why is it desirable that organisms are able to regulate expression of their genes?

.....

.....

.....

.....

iv) How can inducible and repressible enzymes of micro-organisms be distinguished?

.....

.....

.....

.....

.....

14.4 LACTOSE METABOLISM AND THE OPERON

F. Jacob and J. Monod (1961). Genetic regulatory mechanisms in the synthesis of proteins. In Journal of Molecular Biology, Volume 3, pp. 318-56. It is the classic paper in which the operon theory was first proposed.

Metabolic regulation or the first operon was first studied in detail by F. Jacob and J.Monod. They studied the regulation of lactose metabolism in *E. coli* and presented their findings in a paper published in 1961 (see its reference in the adjacent margin remark). Theirs was a brilliant and exemplary piece of scientific investigation which proceeded by a combination of model building and genetic analyses. They received Nobel Prize for this work in the year 1965. The operon concept has had a powerful influence on all subsequent works on gene expression in both prokaryotes and eukaryotes. In fact, most of the terminology used to describe regulation has come from the genetic analysis of this system. We shall first discuss the salient features of *lac* operon, and then the model of *lac* operon.

14.4.1 Lactose as a Carbon Source

Escherichia coli grows very well in a simple medium containing salts and a carbon source such as glucose. These chemicals provide molecules that are utilised by these bacteria to produce everything they need to grow and reproduce, such as nucleic acids, proteins, lipids etc. The energy for these biochemical reactions comes from the metabolism of glucose, a process that is of central importance to the function of a bacterial cell and also to the function of cells of all organisms. The enzymes required for the metabolism of glucose are, therefore, coded for by the constitutive genes.

Let us see as to what happens when instead of glucose as a carbon source, some other sugar say lactose is provided. Lactose is a disaccharide composed of a glucose unit attached to a single galactose unit (see Fig. 14.1). When lactose is provided to *E. coli*, it starts rapidly synthesising a number of enzymes that are needed for the metabolism of lactose. The synthesis of these enzymes is triggered because the genes that code for them become actively transcribed in the presence of lactose. The very same genes are rendered inactive if lactose is absent. In addition, the mRNAs for these enzymes have a relatively short lifetime (see Fig. 14.3), that is, about an average half life of 1-2 minutes at 37°C. Therefore, the transcripts must continually be made in order for the enzymes to be produced. These genes are examples of regulated genes whose products are needed only at certain times.

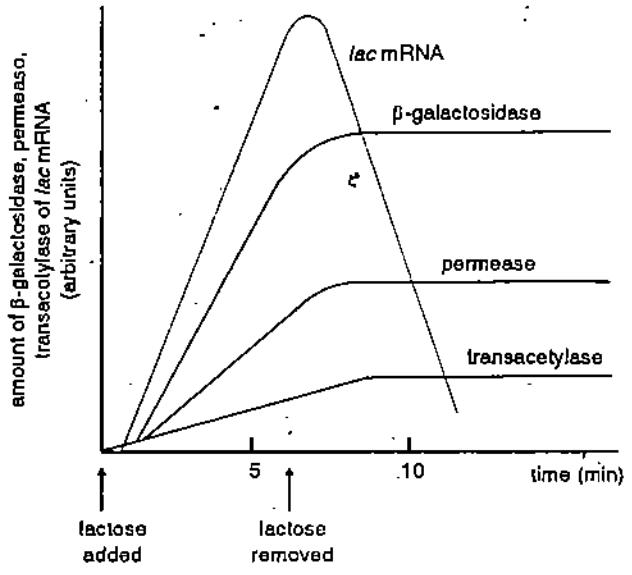


Fig : 14.3 : The " on-off" nature of the synthesis of three lactose metabolising enzymes. The *lac* mRNA appears soon after lactose is added. β-galactosidase and permease appear at nearly the same time but are delayed with respect to mRNA synthesis because of the time required for translation. When lactose is removed, no more *lac* mRNA is made, and the amount of *lac* mRNA decreases owing to the degradation of mRNA already present.

When lactose is the sole carbon source present in the growth medium, three different enzymes are synthesised in *E.coli*. But only two of these are abundantly utilised in metabolising lactose. These three enzymes are : 1) β-*galactosidase*; 2) *lactose permease*, and 3) *transacetylase*.

1) β-*galactosidase* : It has two functions (see Fig. 14.1 again). *First*, it catalyses the *isomerisation* (conversion to a different form) of lactose to allolactose. It is a compound that is important in the regulation of expression of lactose metabolising genes. *Second*, it catalyses the break-down of lactose into two component monosaccharides, glucose and galactose. In the growing cell the galactose is subsequently converted to glucose by certain enzymes that are coded by gene system specific for galactose catabolism, and the glucose in metabolised by the constitutively produced enzymes.

2) *Lactose permease*: It is also called **M protein.** It is located in the *E.coli* membrane and is needed in the active transport of lactose from the environment, which in this case is the growth medium, into the cell.

3) *Transacetylase*: The function of this enzyme is still obscure and perhaps not directly related to lactose utilisation.

In a wild-type *E.coli* that is growing in a medium containing glucose, but no lactose, only a few molecules of each of the three enzymes are produced. It indicates there is a low level of expression of the three genes that code for the proteins. If only lactose is present in the growth medium, however, the number of molecules of each of the three enzymes increases by about a thousand fold, indicating that the three genes which were inactive earlier have become highly active as they are being actively transcribed and translated. The inducer molecule directly responsible for the increased production of the enzymes is actually the allolactose (see Fig.14.1). Allolactose is produced from lactose as a result of one of the activities of the β-galactosidase enzyme.

14.4.2 Organisation of the Lactose Metabolic Genes

The mapping experiments and complementation tests done with mutants affecting the function of the three enzymes inducted by lactose, have revealed that three **structural genes** are involved in lactose metabolism. And these three genes lie adjacent to one another or are clustered in the genome.

When lactose is present, the three structural genes are **coordinately induced.** It means that all the three genes are simultaneously transcribed and their products, the enzymes are rapidly produced. The regulation of these three genes as a unit is under the control of a regulator protein which interacts with a controlling site, called an **operator.** It lies adjacent to the gene cluster (see Fig. 14.4). The operator, does not code for any product. It, however, has a unique sequence. The **promoter** is located next to the operator. A cluster or group of genes whose expressions are regulated together by the operator-regulator protein interaction including the promoter and operator regions, is called an **operon.** This is how the cluster of genes governing the metabolism of lactose are collectively named as **lactose operon** or simply **lac operon** (Fig.14.4). Thus lac operon consists of three structural genes (for β-galactosidase, permease and transacetylase) and the adjacent operator and promoter.

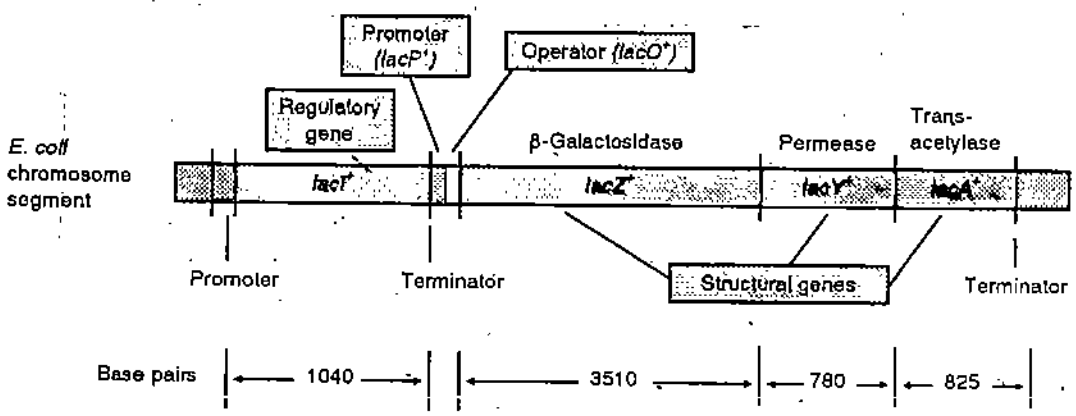


Fig.14.4: The constituents of *lac* operon in *E.coli*. The number of base pairs for each gene is given at the lowermost end of the figure.

Observe Figure 14.4 carefully and answer the following questions.

SAQ 2

i) List the components of *lac* operon?

.....

.....

ii) How many structural genes constitute this operon?

.....

iii) What is the order of the structural genes?

.....

iv) What is meant by coordinate regulation?

.....

The arrangement of the five components of the *lac* operon and the names of products of the structural genes must be clear to you from figure given above. You must have noticed in the figure that all the genes are denoted with a + sign. This sign indicates the genes in a wild-type cell. The expression of the *lac* operon is controlled by a repressor molecule, which is the product of the *lac I* gene (*lac I*$^+$ deontes the wild-type). Have you noticed in the figure (14.4) that this gene maps a short distance away from the promoter for the *lac* operon. The promoter and terminator for the *lac I* gene are distinct from those elements for the *lac* operon itself. Expression of the *lac I* gene is under the control of promoter, and therefore, this gene is another example of a constitutive gene.

14.4.3 Experimental Evidence for the Regulation of the *lac* Operon

After studying the organisation of the genes constituting *lac* operon, let us look at the genetic evidences for each of the components and their manner of action.

A. Analysis of Mutants of the Structural Genes: In one set of experiments, Jacob and Monad induced mutations in the structural genes by using appropriate mutagens. The mutants produced were identified on the basis of enzyme assays since in every case the activity of a particular enzyme was reduced greatly as a consequence of a mutation in the structural gene coding for that enzyme. The study of mutants *lac Z*$^-$, lac *Y*$^-$, and lac *A*$^-$ revealed the location of the three genes. And it was established that the order of the structural genes in the genome was *lac Z*$^-$ *lac Y* *lac A* and that the three genes were tightly linked in a cluster or group.

In the *lac* operon some missense mutants were also obtained, the study of which brought to light another aspect of the structural genes. The missense mutations result in the substitution of one amino acid for another in a polypeptide. Thus such a mutation only affects the expression of the gene concerned. Let us understand it with the help of an example. A lac *Z*$^-$ missense mutation results in a nonfunctional or partially functional β-galactosidase, but permease and transacetylase are totally normal.

In contrast, the effect of a chain terminating or nonsense mutation will depend on the location of the mutation in the gene. For instance, if a nonsense mutation occurring near the chain termination codon of the *lac Z* gene, active or partially active β-galactosidase may still be produced, with little or no effect on permease and transacetylase production. However, the closer the nonsense mutation to the 5' end of the *lac Z* gene the more serious are the effects. Firstly, a short β-galactosidase polypeptide is produced, which is mostly nonfunctional. Secondly, no functional products from the *lac Y* and *lac A* genes are produced even though there are no mutations in those genes.

If there are nonsense mutations of the *lac Y* gene, the result is loss of permease function, as well as loss of the transacetylase function, but no effect on β-galactosidase activity.

And the nonsense mutations in the *lac A* gene results in the loss of only transacetylase function. The activities of β-galactosidase and permease remain unaffected.

From this discussion, it can be deduced that the nonsense mutations depending on their occurrence in a particular structural gene, have different effects. Extending this further, the nonsense mutations exhibit *polar effects*, and the phenomenon is called **polarity.**

Based on the polar effects exhibited by the nonsense mutations of the *lac* operon structural genes, one important interpretation was made. It was that all the three genes are transcribed onto a single mRNA molecule (known as a **polygenic mRNA**) rather than onto three separate mRNAs. Findings from subsequent experiments also supported the above interpretation. Recent studies have shown that RNA polymerase initiates transcription at the promoter, and a polygenic mRNA is synthesised in the order 5' – *lac Z*$^+$ – *lac Y*$^+$ – *lac A*$^+$ – 3'. Translation of the polygenic mRNA begins near the 5' end and proceeds towards the 3' end. So, the activity of ribosome would be as follows : *first*, synthesise β-galactosidase; *second* slide along the mRNA until it recognises the initiation sequence of permease; *third*, synthesise permease; *fourth*, slide along mRNA until it recognises the initiation sequence for transacetylase; *fifth* synthesise transacetylase. And finally dissociate from mRNA.

B. Analysis of Promoter Mutants : The mutants of promoter that is located at the *lac Z* end of the group of *lac* genes (see Fig. 14.4) were also available. Most of the promoter mutants, represented as *lac* P$^-$ affected all three structural genes. In such promoter mutants, even if lactose is present, the enzymes required for lactose metabolism, are either not at all synthesised, or synthesised at very low rates. Since the promoter is the recognition sequence for RNA polymerase, the effect of a *lac P*$^-$ mutation is confined to the genes that it controls on the same chromosomal strand.

C. Analysis of Operator Mutants : These mutants are very important as they affect the regulation of the lactose operon. You may recall that in wild-type *E. coli* the three structural gene products are not made in the absence of lactose. And in the presence of lactose all the three genes are induced. Jacob and Monod isolated a number of mutants in which all gene products of the lactose operon were synthesised constitutively. That is, all the three enzymes were synthesised continually whether lactose was present or absent. These mutants also known as *regulatory mutants*, did not affect the functions of the enzymes but affect the cellular mechanisms responsible for regulating the expression of the genes coding for the enzymes. As a result of their mapping experiments, Jacob and Monod recognised two classes of constitutive mutants : one class mapped to a relatively small DNA region, i.e., operator (*lac O*) which lies adjacent to the *lac Z* gene. The second mapped to a gene-sized DNA region a short distance away from the *lac* Z gene, and it is known as the *lac I* gene (see Fig. 14.4).

The operator mutants were called operator-constitutive or *lac O*^{c} mutants. All the *lac O*^{c} mutants synthesise enzymes for lactose metabolism whether lactose is present or not. By the use of partial diploid strains (see Unit 12, Fig. 12.20, Subsection 12.10.2), Jacob and Monod were able to precisely define the role of the operator in regulating the expression of the *lac* operon. One such partial diploid they used was *lac O*$^+$ *lac Z*$^-$ *lac Y*$^+$ / *lac O*^{c} *lac Z*$^+$ lac *Y*$^-$. In this example the gene *lac A* gene is deliberately omitted, since it does not pertain to our present discussion. In the partial diploid mentioned above, the genotype of the plasmid is given on the left of the line and that of the chromosome at the right. Thus the plasmid genotype has normal operator (*lac O*$^+$), a mutant β-galactosidase (*lac Z*$^+$) and a normal permease (*lac Y*$^+$) gene. The bacterial chromosome has a constitutive operator mutant, (*lac O*^{c}) a normal β-galactosidase gene (*lac Z*$^+$), and a mutant permease gene (*lac Y*$^-$). This partial diploid was tested for the production of β-galactosidase from the *lac Z*$^+$ gene, and of permease from *lac Y*$^+$ gene both in the presence and absence of the inducer lactose. Their analysis revealed that in the absence of inducer, the β-galactosidase is still synthesised, but permease is not. Only when lactose is added to the culture, permease synthesis occurs. In other words, in this partial diploid *lac Z*$^+$ gene is constitutively expressed, whereas the gene *lac Y*$^+$ is expressed under normal inducible conditions. The interpretation of these results was that *lac O*^{c} mutations affect only the genes that are adjacent to it on the same chromosome strand. Furthermore, the *lac O*$^+$ region may control those genes adjacent to it having no effect on the genes on the other chromosome. Thus the *lac O*$^+$ mutation is cis-dominant as it affects the adjacent genes only and cannot be dominated by a normal *lac O*$^+$ region elsewhere in the chromosome. From these experiments Jacob and Monod also concluded that the operator region does not produce a diffusible product that functions in the cell. If it was produced then in the *lac O*$^+$/*lac O*^{c} diploid state one or the other regions would have controlled all the lactose utilisation genes wherever they were.

D. Analysis of lac *I Gene* Regulatory Mutants : The mutants of the *lac I* gene constitute another class of mutants of the *lac* operon. You have seen above that not only three *lac genes* (*Z, Y* and *A*) are induced together, but they are also adjacent to one another. They are in fact transcribed on a single polycistronic mRNA. The above given induction process involves the protein product of another gene, called the *regulator gene* or *I* gene (*lac I*). Although the regulator gene is located adjacent to the three other *lac* genes, it is a totally independent transcriptional entity. The *repressor* synthesised by the regulator interferes with the transcription of the genes involved in lactose metabolism.

The study of the *lac I*$^-$ mutants in partial diploid strains was instrumental in obtaining an understanding of the normal regulation of the lac operon. A partial diploid with genotype *lac I*$^+$ *lac O*$^+$ *lac Z*$^-$ *lac Y*$^+$ / *lac I*$^-$ *lac O*$^+$ *lac Z*$^+$ *lac Y*$^-$ was used. In this case both the *lac* operons have normal operators and normal promoters; one operon has a normal *lac I* gene (*lac I*$^+$), a mutant β-galactosidase gene (*lac Z*$^+$), and a wild-type permease gene (*lac Y*$^+$). The other operon has a constitutive mutant *lac I* gene (*lac I*$^-$) a normal β-galactosidase gene (*lac Z*$^+$) and a mutant permease gene (*lac Y*$^-$). The partial diploid was tested for structural gene expression in the absence and presence of the inducer lactose. In the absence of lactose, no β-galactosidase or permease was produced, but both were synthesised in the presence of lactose. Therefore, the expression of both the genes (*Z* and *Y*) was inducible. This implies that the *lac I*$^+$ gene in the cell can overcome the defect of *lac I*$^+$ gene in the cell can overcome the defect of *lac I*$^-$. Hence *lac I*$^+$ is dominant over *lac I*$^-$ and the *lac I*$^-$ mutants are recessive. Since the *lac I*$^+$ gene did control the genes on the other chromosome strand it was believed that the *lac I* gene must be producing a diffusible product. Jacob and Monod proposed that the *lac I*$^+$ gene produces a functional repressor molecule. Based on which the *lac I* gene also became known as a *repressor gene*. The mutants *lac I*$^-$ do not produce any functional repressor molecules. Thus in a haploid bacterial strain that has a *lac I*$^-$ mutation, the *lac* operon is constitutive. In a partial diploid with both a *lac I*$^+$ and a *lac Z*$^-$, the functional repressor molecules produced by the *lac I*$^+$ gene control the expression of both *lac* operons present in the cell, making both of them inducible.

14.4.4 The Model for Regulation of the *lac* Operon

Jacob and Monod, on the basis of their investigations on the genetic mutants of lactose metabolising enzyme, proposed their famous operon model, which is a milestone in the genetic investigations. We present below an updated version of the Jacob-Monod model for the regulation of the *lac* operon.

First let us consider the state of *lac* operon in a wild-type *E. coli* cell growing in a medium lacking lactose. What happens at this condition is represented diagrammatically in Fig. 14.5. In this situation, the RNA polymerase binds to the promoter of the repressor

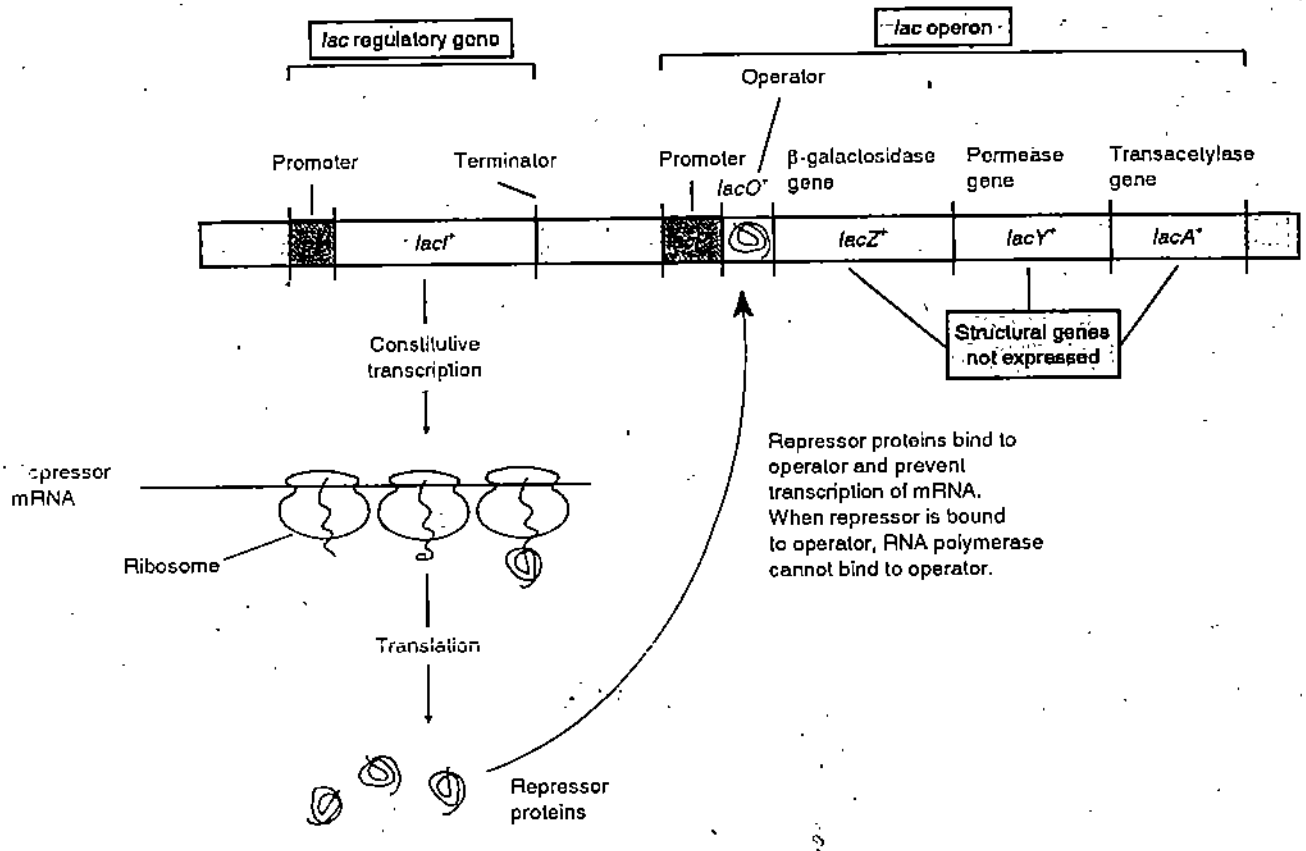


Fig. 14.5 : Diagrammatic representation of the functional state of the *lac* operon in wild type *E. coli* growing in a medium lacking lactose

gene (*lac I*$^+$) and transcribes it. The mRNA produced, on translation produces a polypeptide having 360 amino acids. Four of these polypeptides associate together to form the repressor protein. This repressor has affinity for the base-pair sequence of the operator to which it binds subsequently. And if the repressor is bound to the operator, RNA polymerase cannot bind to the operon's promoter which is situated adjacent to the three structural genes. Hence, the transcription of the three structural genes in the operon cannot occur. The *lac* operon in this state is said to be under **negative control**. The repressor, however, does not just bind and stay at the operator site, but binds and unbinds continuously. In the split second time interval, when one repressor molecule unbinds, and before another repressor binds, RNA polymerase can initiate transcription of the operon even when lactose is absent. As a result this low level of transcription of the structural genes (*Z*, *Y* and *A*), a few molecules of their respective products are present even in the absence of lactose.

Let us now see the functional state of the wild-type of *E. coli* growing in the presence of lactose (see Fig. 14.6). Some of the lactose which is transported into the cell, is converted to allolactose by the existing molecules of β-galactosidase. The allolactose thus formed induces the production of the *lac* operon enzymes. The *lac* repressor protein, in addition to having a recognition site for the *lac* operator, also has a recognition site for allolactose. Once the allolactose binds to the repressor, the shape of the repressor changes and hence its recognition site for the operator. As a result, the repressor loses its affinity for the *lac* operator and dissociates from the site. The free repressor proteins are also altered so that they cannot bind to the operator.

In the absence of repressor, RNA polymerase binds to the operon's promoter and initiates the synthesis of single polygenic mRNA molecule that contains the transcripts for *lac Z*$^+$, *lac Y*$^+$ and *lac A*$^+$ gene. The mRNA thus produced is subsequently translated by ribosomes forming the proteins specified by the operon.

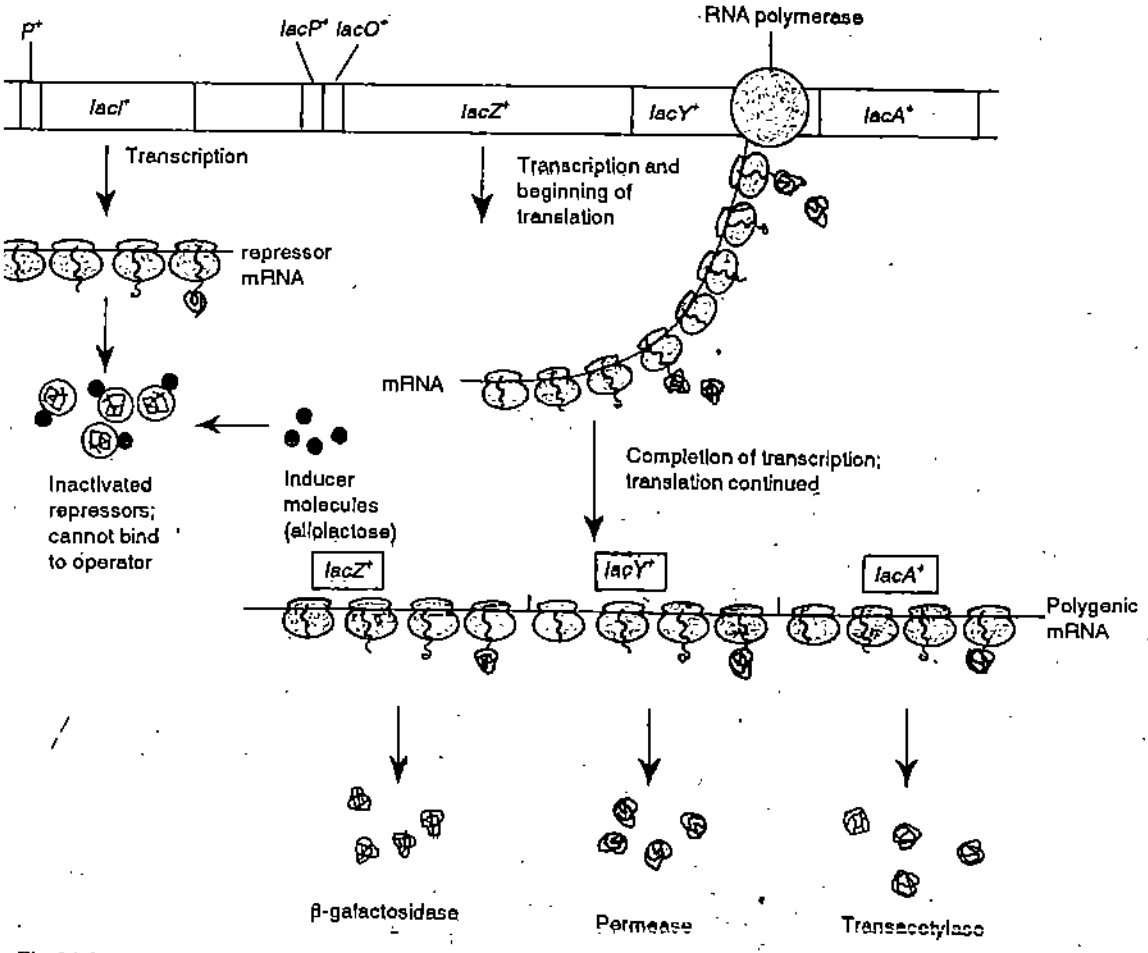


Fig. 14.6 : Diagrammatic representation of the functional state of the *lac* operon in wild-type *E. coli* growing in a medium containing lactose as a sole carbon source.

SAQ 3

Match the elements of column I with those of column II. Write your answer in the space provided.

I			II	
1)	Regulator gene	[ ]	a)	binding site of RNA polymerase and CAP-cAMP complex
2)	Operator	[ ]	b)	encodes β-galactosidase permease
3)	Promoter	[ ]	c)	codes for repressor
4)	Structural Gene *Z*	[ ]	d)	binding site for repressor
5)	Structural Gene *Y*	[ ]	e)	encodes β-galactosidase

SAQ 4

Write short answers for the following questions in the space provided.

1) What would be the result of inactivation by mutation of the following genes or sites in the *E. coli lac* operon : a) regulator, b) operator, c) promoter, d) structural gene *Z*, and e) structural gene *Y*?

.....

.....

.....

.....

ii) Other than the ability to turn on and off a set of genes in an operon by a single regulatory element, what else is accomplished by having a set of genes contained in one polycistronic mRNA molecule ?

.....

.....

iii) Would synthesis of an enzyme that is needed continually be regulated ?

.....

.....

14.4.5 Positive Control of the *lac* Operon by CAP and Cyclic AMP

In the previous subsection, you studied the *lac* operon model. It explains how the biosynthesis of the enzymes involved in lactose metabolism is induced when lactose is added to the medium in which *E. coli* cells are growing. An interesting property of *lac* operon is that its induction is prevented by the presence of glucose in its environment. That is, glucose is catabolised in preference to other sugars (such as lactose, arabinose, galactose etc.) present in the environment. This phenomenon is known as **catabolite repression or the glucose effect.**

Catabolite repression of the *lac* operon is mediated via positive control of transcription by a regulatory protein known as *Catabolite Activator Protein* (a dimer) abbreviatd as CAP in conjunction with a small effector molecule called *cyclic-AMP* (adenosine-3,5-phosphate, see Fig. 14.7). The *CAP* is also called as *cyclic AMP receptor protein*. The *lac*

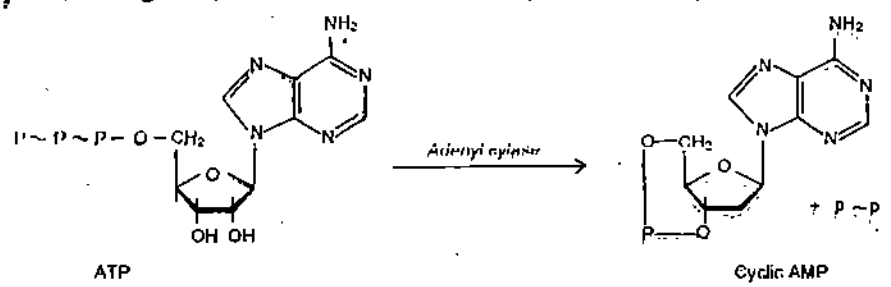


Fig. 14.7 : Formation of the regulatory molecule cyclic AMP (adenosine-3',5'-phosphate) from ATP by the action of the enzyme adenylcyclase.

promoter actually contains two separate binding sites, one for RNA polymerase and the second for the CAP-cAMP complex (see Fig. 14.8). The complete nucleotide pair sequence of the *lac* operon regulatory region — the promoter (CAP-cAMP site and the RNA polymerase site) and operator is now known (see Fig. 14.8). In the absence of glucose, cAMP binds to the CAP protein; the CAP-cAMP complex then binds to the CAP site of the *lac* promoter. This binding stimulates the binding of RNA polymerase to its binding site at the *lac* promoter leading to the transcription of the structural gene. In the absence of sufficient quantities of cAMP, CAP cannot bind to the *lac* promoter and the *lac* operon cannot be induced. The intercellular cAMP concentration is sensitive to the presence of glucose. The increase in level of glucose in the *E. coli* cells causes a loss of cAMP from the cell by some unknown mechanism. Perhaps glucose, or some metabolite that forms in the presence of sufficient concentration of glucose, inhibits the activity of *adenylcylase*, – the enzyme that catalyses the formation of cAMP from ATP. Whatever is the mechanism, this is for sure that the presence of glucose results in a decrease in the intercellular concentration of cAMP, thus lowering the CAP – cAMP level too. The result

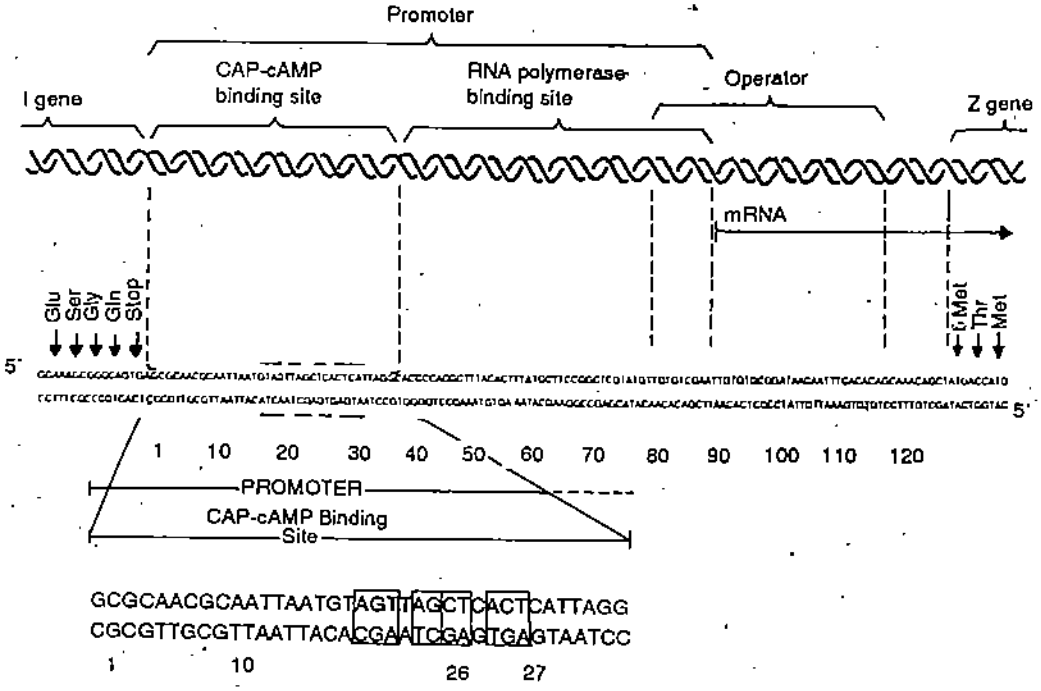


Fig. 14.8 : Organisation and nucleotide sequence of the promoter-operator region of the *lac* operon. Situated to its left is the *I* gene (repressor) and to its right is the *Z* structural gene. The promoter consists of two sites : one, the site that binds the CAP – cAMP complex, and second, the RNA polymerase binding site. Note, that the promoter and the operator (repressor binding site) overlap slightly. The line labelled mRNA (just below the operator) indicates the position at which transcription of the operon begins. There are some numbers near the lower end (i.e., 1, 10, 20 . . 120), these denote the distances in nucleotide pairs, from the end of the *I* gene, , or, the beginning of the promoter. The dot between the nucleotide chain sequences in the CAP – cAMP binding site is the centre of symmetry of an imperfect palindrome indicated by parallel bars. This region is the potential *site for interaction with the CAP–cAMP complex*.

is reduced rate of transcription of the *lac* operon (Fig. 14.9). The same reduction of transcription rates is noticed in mutant strains of *E. coli* when the CAP–cAMP complex binding site of the promoter is deleted.

Catabolite repression is an example of positive regulation. That is, binding of the CAP – cAMP complex at the CAP site enhances the transcription rate of that transcription unit. Thus, the lac operon is both positively and negatively regulated. The repressor exerts a negative control and the CAP – cAMP complex exerts positive control of transcription.

Palindromes are nucleotide-pair sequences that read almost the same in both directions, i.e., forward and backward from a central axis of symmetry, like the non sense phrase

AND MADAM DNA

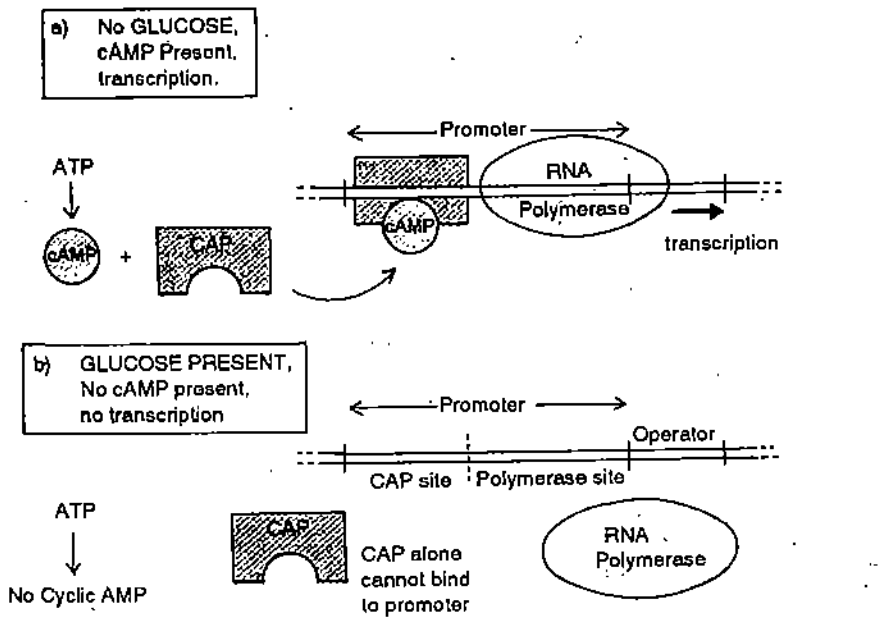


Fig. 14.9 : Catabolite repression. a) The CAP–cAMP complex formed, binds to the CAP site and enhances the transcription of the operon. b) In the presence of glucose, no CAP – cAMP complex is formed, and therefore, the transcription of the same operon is reduced.

SAQ 5

i) Which type of regulation, positive or negative involves removal of an inhibitor ?

.....

ii) When glucose is present, is the concentration of cyclic AMP high or low ?

.....

iii) Can a mutant with either an inactive adenyl cyclase gene or an inactive CAP gene synthesise β-galactosidase?

.....

.....

iv) Does the binding of cAMP–CAP to the operon affect the binding of a repressor in any way?

.....

v) Of what biological significance is the phenomenon of catabolite repression ?

.....

.....

.....

14.5 THE TRYPTOPHAN OPERON

The tryptophan (*trp*) operon is responsible for the synthesis of the amino acid tryptophan. Regulation of this operon occurs in such a way that, when tryptophan is present in the growth medium, the *trp* operon is not active. That is, when adequate amount of tryptophan is present, transcription of the operon is repressed. However, when the supply is insufficient, transcription occurs. The *trp* operon is quite different from the *lac* operon in that tryptophan acts directly in repression rather than as an inducer. Furthermore, because the *trp* operon codes for a set of biosynthetic rather than degradative enzymes, neither glucose nor cAMP–CAP functions in operon activity.

A simple on-off system, as in the *lac* operon is not optimal for a biosynthetic pathway. For example, a situation may arise in which some tryptophan is present in the growth medium, but not enough to sustain the optimal growth. Under these conditions, it is necessary to synthesise tryptophan, but at less than the maximum possible rate. Cells adjust to this

situation by means of regulatory mechanism in which the amount of transcription in the derepressed (not repressed) state is determined by the concentration of tryptophan in the cell. Such regulatory mechanism is found in many operons responsible for amino acid biosynthesis.

14.5.1 Tryptophan Biosynthesis

Tryptophan is synthesised in five steps as shown in Fig. 14.10, each requiring a particular enzyme. The genes coding for these enzymes lie adjacent to one another in the *E.coli*

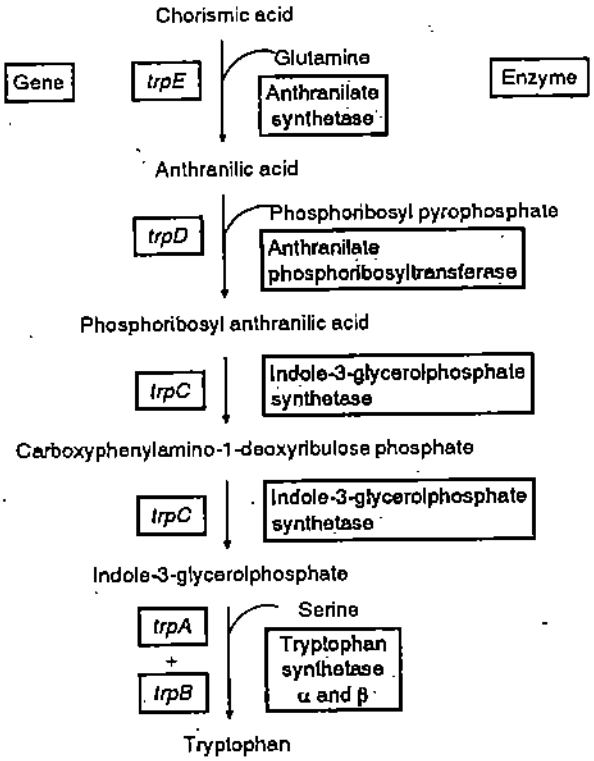


Fig. 14.10 : Biosynthesis of tryptophan in *E. coli*. The enzymes produced by the trp operon catalyse the conversion of chorismic acid to tryptophan.

chromosome (see Fig. 14.11). These genes are in the same order as that in which the enzymes are used in the biosynthetic pathway. The genes are called *trp E*, *trp D*, *trp C*, *trp B* and *trp A*, and the enzymes are translated from a single polycistronic mRNA molecule. The *trp E* gene is the first one to be translated. Upstream, i.e., near the *trp E* end of the operon (or the 5'-end) are the **promoter**, the **operator** and two regions called the **leader** and the **attenuator**, which are designated as *trp L* and *trp a* respectively (see Fig. 14.11). The repressor gene, *trp R*, is located quite far away from this operon and is not shown in the figure.

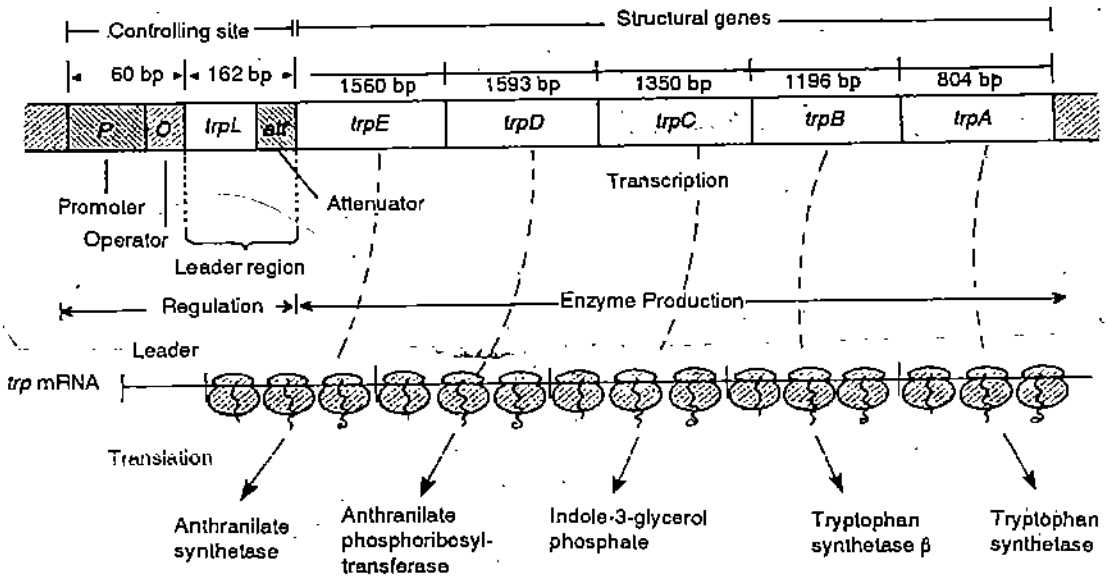


Fig. 14.11 : The organisation of regulation and enzyme production sites of the *E. coli trp* operon. The size of each region is indicated by the number of base pairs.

14.5.2 Regulation of *trp* Operon

The regulatory protein of the *trp* operon is the product of the *trp R* gene. Mutations in either this gene or the operator cause constitutive initiation of transcription of the *trp* mRNA, as in the *lac* operon. The *trp R* gene product is called the ***trp aporepressor.*** It does not bind to the operator unless it is first bound to tryptophan, that is, the aporepressor and the tryptophan molecule join together to form the ***active trp repressor***, which binds to the operator (see Fig. 14.12). When there is not enough tryptophan, the aporepressor has such a three-dimensional conformation that it is unable to bind with the *trp* operator, and transcription of the operon occurs. On the other hand when tryptophan is present in high enough concentration, some molecules bind with the aporepressor and cause it to change conformation into the active repressor. The active repressor binds with the *trp* operator and prevents transcription (see Fig. 14.12). Thus, only when tryptophan is present in sufficient amounts is the active repressor molecule formed. This is the basic on-off regulatory mechanism.

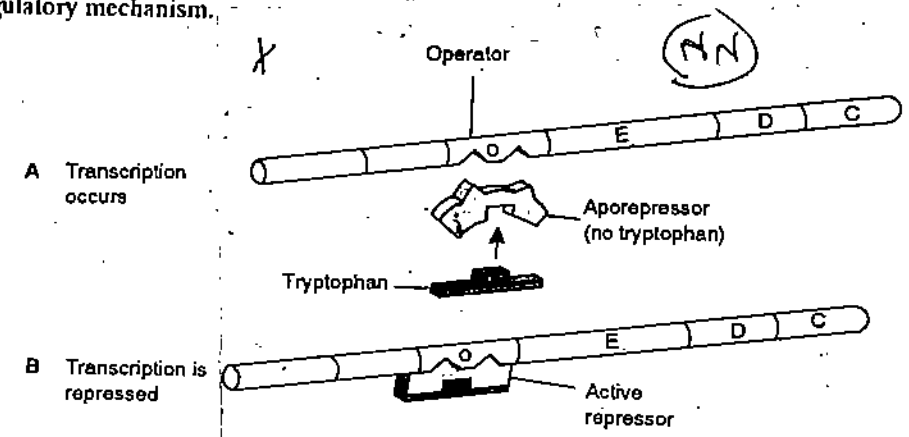


Fig. 14.12 : Regulation of *trp* operon. Transcription is inhibited when the active repressor binds to the operator.

14.5.3 Attenuation

In addition to what is stated above a second level of regulatory mechanism of *trp* operon exists and you would study about it in this subsection.

When the *trp* operon is in the on state, a finer control in which enzyme concentration is adjusted by amino acid concentration is effected by : 1) premature termination of transcription before the first structural gene is reached, and 2) regulation of the frequency of this termination by the internal concentration of tryptophan. This type of regulation uses translation to control transcription.

Attenuation occurs through interactions between DNA sequences present in the leader region of the *trp* transcript. In the wild-type cells, transcription of the *trp* opern is often initiated. However, in the presence of even small amounts of tryptophan, most of the mRNA molecules terminate in a specific 28-base region within the leader sequence. The result of such termination is an RNA molecule that contains only 140 nucleotides and stops short of the genes coding for the *trp* enzymes. This 28-base pair region, where termination occurs is called the **attenuator**. The base sequence (see Fig. 14.13) of the region in which termination occurs has the usual features of a termination site, i.e., a stem-and-loop configuration in the mRNA followed by a sequence of eight AT base pairs.

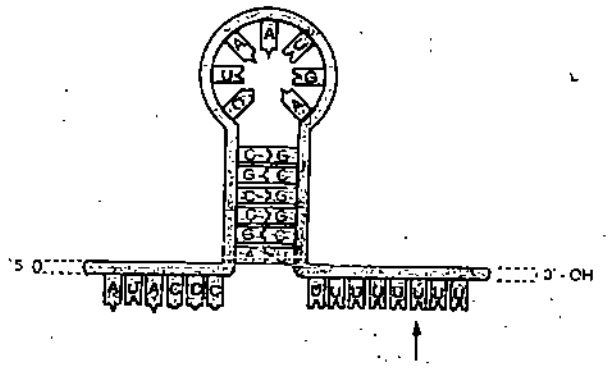


Fig. 14.13: The terminal portion of the *trp* attenuator sequence. The arrow indicates the final uridine in attenuated RNA. Nonattenuated RNA continues past that base.

Let us view some salient features of the leader sequence.

First, it codes for a polypeptide known as leader polypeptide, which is composed of 14 amino acids (Fig. 14.14).

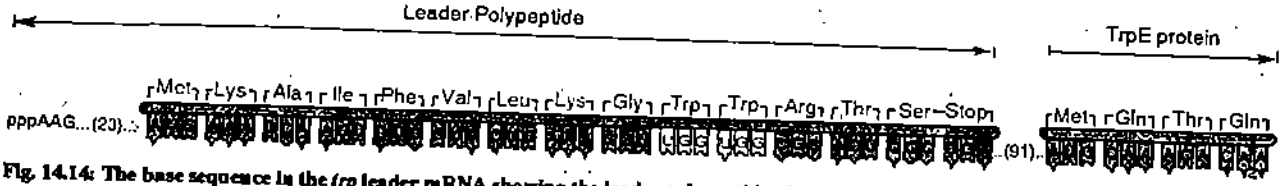


Fig. 14.14: The base sequence in the *trp* leader mRNA showing the leader polypeptide, the two tryptophan codons and the beginning of *trp E* protein. The numbers 23 and 91 are the numbers of bases in sequence that, for clarity are not shown.

Second, two adjacent tryptophan codons are located in the leader polypeptide at positions 10 and 11. We shall be discussing the significance of these repeated codons shortly.

Third, four segments of the leader RNA — denoted 1,2,3 and 4 are capable of base-pairing in two different ways (see Fig. 14.15) : either the base-paired regions, i.e., 1-2 and 3-4 or just the region 3-4 pair or 2-3 pair (see Fig. 14.15 b and c respectively). The 1-2 and 3-4 pairs form in purified *trp* leader region. Region 1 overlaps the coding region of the leader polypeptide.

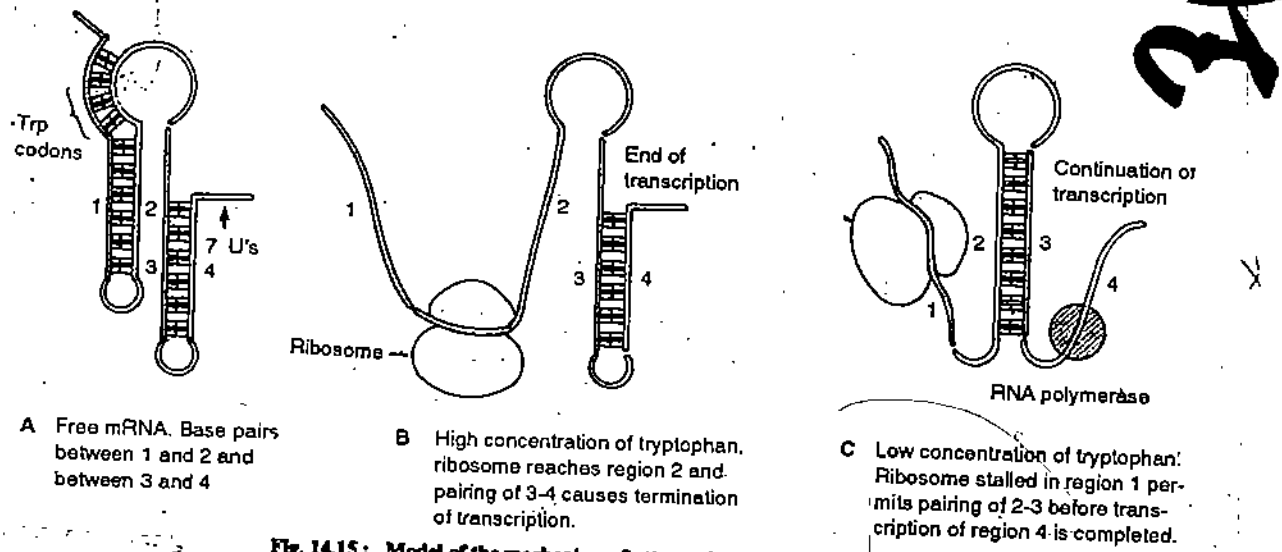


Fig. 14.15: Model of the mechanism of attenuation in the *E.coli trp* operon.

Such a base sequence organisation enables premature termination in the *trp* leader region by the following mechanism.

The termination of transcription is brought about by translation of the leader peptide region. Since there are two tryptophan codons in this sequence, the translation of the sequence is sensitive to the concentration of charged $tRNA^{trp}$. That is, if the tryptophan supply is not adequate, then amount of charged $tRNA^{trp}$ will be insufficient and hence translation will be slowed at the tryptophan codon.

Figure 14.15 shows the end of the *trp* leader peptide in segment 1. A translating ribosome is in contact with about ten bases in the mRNA past the codons being translated, and all base-pairing is eliminated in the segment of mRNA that is in contact with the ribosome.

Thus, when the codons of the leader polypeptide are being translated, segments 1 and 2 are not paired (Fig. 14.15b). Because transcription and translation are coupled in prokaryotes the leading ribosome is not far behind the RNA polymerase (Fig. 14.15c). Thus, if the ribosome is in contact with segment 2 when transcription of segment 4 is in progress, then segments 3 and 4 will form the duplex region 3-4, (Fig 14.15b) without segment 2 competing for segment 3. The presence of the 3-4 stem-and-loop configuration allows termination to occur when the terminating sequence of seven uridines is reached. If the concentration of charged $tRNA^{trp}$ is too low, occasionally a translating ribosome is stalled at the tryptophan codon (Fig. 14.15c). These codons are located 16 bases before the beginning of segment 2. Thus, segment 2 is free before segment 4 has been transcribed, and the 2-3 duplex can form. This prevents the 3-4 stem and loop from forming, and so termination does not occur and the complete mRNA molecule is made including the coding sequences for the *trp* genes. Hence, if tryptophan is present in excess, termination

occurs and little enzyme is synthesised; and if tryptophan is absent, termination does not occur and the enzymes are made. At intermediate concentrations, the fraction of initiation events that result in completion of *trp* mRNA will depend on how often translation is stalled, which in turn depends on the concentration of tryptophan.

You should remember that attenuation cannot occur in eukaryotes because transcription and translation are not coupled.

Regulation of the *lac* and *trp* operons that you have studied above typifies some of the important mechanisms controlling transcription in prokaryotes. In the following section you will study briefly transcriptional regulation in bacteriophages.

14.6 TEMPORAL SEQUENCES OF GENE EXPRESSION DURING PHAGE INFECTION

The regulation of gene expression in bacteriophages is quite different from the reversible on-off switches characteristic of bacterial operons. The viral genes are expressed in *genetically preprogrammed sequences.* Although different bacterial viruses exhibit variations of the specific mechanisms involved, there are some commonalities in them. One set of phage genes, usually called 'early' genes, is expressed immediately after infection. The product(s) of one or more of the early genes is involved in turning off the expression of the 'early' genes and turning on the expression of the next set of genes and so on. Two to four sets of genes depending on the kind of virus are characteristically involved. In all cases studied so far, the regulation of *sequential gene expression* during phage infection *occurs primarily at the level of transcription.*

The phages *T4* and *T7* in *E.coli,* and phage *SPO1* in *Bacillus subtilis* are the examples of three of the most extensively studied bacterial viruses. In these viruses the sequential gene expression is controlled by modifying the promoter specificity of RNA polymerase as in *T7* or by phage-induced alterations of the host cell's RNA polymerase as seen in *T4* and *SPO1.*

In the cells infected with the phage *T7* the 'early' genes are transcribed by the *E.coli* RNA polymerase, which then transcribes all the 'late' genes, that code for *T7* structural proteins, lysozyme etc. *Bacillus subtilis* phage *SPO1* exhibits a slightly more complex pathway of sequential gene expression, involving three sets of genes. These three sets of genes are called 'early', 'middle' and 'late' in reference to their time of gene expression during the reproductive cycle of phage. The 'early'genes of *SPO1* are transcribed by *B. subtilis* RNA polymerase. One of the products of the 'early' gene is a polypeptide that binds to the host cell's RNA polymerase, changing its specificity such that it transcribes the 'middle' genes of *SPO1.* Two of the products of 'middle' genes are, in turn, polypeptides that associate with the *B. subtilis* RNA polymerase, further changing its specificity so that it then transcribes the 'late' genes of *SPO1.*

The phage *T4* exhibits even more complex pattern of sequential gene expression involving several different modifications of the RNA polymerase of the host cell. Thus, in these bacterial viruses the control of the observed sequential gene expression occurs primarily at the level of transcription and is mediated by specific RNA polymerase – promoter sequence interactions.

SAQ 6

i) Both repressors and aporepressors bind molecules that are substrates or products of the metabolic pathway encoded by the genes in an operon. Generally speaking, which binds the substrate of a metabolic pathway and which the product ?

.....

.....

.....

.....

.....

ii) Is an attenuator a protein binding site like an operator? Is RNA synthesis ever initiated at an attenuator ?

.....

.....

.....

14.7 SUMMARY

In this unit you have studied the following :

- Most cells do not synthesise molecules that are not needed. There are important opportunities for controlling gene expression in transcription, RNA processing, and translation. Gene expression can also be regulated through the stability of mRNA, and the activity of proteins can be regulated in a variety of ways after translation. The process of gene regulation generally differ between prokaryotes and eukaryotes.
- The *lac* operon in *E.coli* responds to the presence of lactose by switching on the enzymes responsible for its metabolism. The *lac* operon consists of three contiguous structural genes *Z, Y,* and *A* for β-galactosidase, permease and transacetylase located immediately down stream of a promoter/operator (p/o) region. A repressor protein encoded by the regulator gene *I*, binds to the p/o region to prevent transcription unless an inducer molecule lactose is present. The inducer binds to the repressor altering its shape so that it falls off the p/o, thus allowing transcription to occur. The three structural genes are transcribed as a unit ; The polycistronic mRNA, which is translated into three independent polypeptides.
- The *lac* operon is subject to higher level control by a cAMP receptor protein (CAP). CAP forms a complex with cAMP which by binding next to the promoter region, facilitates RNA polymerase to start transcription. Metabolic conditions that lower cAMP levels in the cell render the *lac* operon uninducible.
- Biosynthetic enzyme systems exemplify another type of transcriptional regulation. In the synthesis of tryptophan, transcription of the genes coding *trp* enzymes is controlled by the concentration of tryptophan in the medium. When excess tryptophan is present, it binds with *trp* aporepressor to form the active repressor that inhibits transcription. The *trp* operon is also regulated by attenuation, in which transcription is initiated continually but the transcript forms a hair pin structure resulting in premature termination. The frequency of termination of transcription is determined by the availability of tryptophan; with decreasing concentration of tryptophan, termination occurs less often and the enzymes for tryptophan synthesis are made, thereby increasing the concentration of tryptophan. Attenuators also regulate operons for the synthesis of other amino acids.
- Preprogrammed sequences of viral gene expression occur in bacteriophage infected cells where the products of genes expressed early after infection interact with RNA polymerase to change its promoter specificity and switch transcription to a second set of genes. In the case of some of the more complex bacteriophages, this switching process may be repeated several times during the life cycle of the virus.

14.8 TERMINAL QUESTIONS

1) Define the following terms :
   a) housekeeping genes
   b) structural gene
   c) regulatory gene
   d) constitutive expression
   e) operator
   f) catabolite repression
   g) attenuation

b) inducer

2) Compare and contrast the following :

a) Promoter and operator

b) Repressor and aporepressor

c) Repression and derepression

d) Attenuator and terminator

e) Constitutive gene and regulated gene

f) Positive regulator and negative regulator

3) If a mutant strain of *E. coli* synthesises β-galactosidase irrespective of the presence or absence of the inducer, what genetic defect(s) might be responsible for this phenotype?

4) In the *lac operon* of *E. coli*, what would be the partial-diploid genotype for a strain that will produce β-galactosidase constitutively and permease by induction ?

5) How might the concentration of glucose in the medium in which an *E. coli* cell is growing, regulate the intracellular level of cyclic AMP?

6) Is the CAP–cAMP effect on the transcription of the *lac* operon an example of positive or negative regulation ? Why ?

7) Would it be possible to isolate *E. coli* mutants in which the transcription of the *lac* operon is not sensitive to catabolite repression ? If so, in what genes might the mutations be located ?

8) Using examples, disinguish between regulation based on negative mechanisms and that based on positive regulatory mechanisms.

9) Given below are ten strains of *E. coli* with the following lactose operon genotypes :

(1) *lacI*$^+$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^+$;

(2) *lacI*$^-$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^+$;

(3) *lacI*$^+$ *lacP*$^+$ *lacO*^{c} *lacZ*$^+$;

(4) *lacI*$^-$ *lacP*$^+$ *lacO*^{c} *lacZ*$^+$

(5) *lacI*$^+$ *lacP*$^+$ *lacO*^{c} *lacZ*$^-$

(6) *F lacI*$^+$ *lacP*$^+$ *lacO*^{c} *lacZ*$^-$/*lacI*$^+$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^+$;

(7) *F lacI*$^+$ *lacO*$^+$ *lacZ*$^-$/*lacI*$^+$ *lacP*$^+$ *lacO*^{c} *lacZ*$^+$;

(8) *F lacI*$^+$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^+$/*lacI*$^-$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^-$;

(9) *F lacI*$^+$ *lacP*$^+$ *lacO*^{c} *lacZ*$^-$/*lacI*$^-$ *lacP*$^+$ *lacO*$^+$ *lacZ*$^+$;

(10) *F lacI*$^-$ *lacP*$^+$ *lacO*^{c} *lacZ*$^-$/*lacI*$^-$ *lacP*$^+$ *lacO*^{c} *lacZ*$^+$;

Note : In the partial diploid strains (6-10) one copy of the *lac* operon is in the host chromosome and the other copy is in the *F* episome.

Predict for each strain, whether β-galactosidase will be produced : a) if lactose is absent from the growth medium, and b) if lactose is present in the growth medium.

10) Describe the role of cAMP in transcriptional control in *E. coli*.

11) What consequences would a mutation in the catabolite activator protein (CAP) gene of *E. coli* have for the expression of a wild type *lac* operon ?

12) Explain why attenuators are common in repressible operons, but are absent in inducible operon such as *lac*.

13) The rate of transcription of the *trp* operon in *E. coli* is controlled by both (a) repression/derepression, and (b) attenuation. Explain, by which mechanisms do these two regulatory processes modulate *trp* operon transcript levels ?

14) If there is deletion of the *trp L* region of the *trp* operon, then what effect would it have on the rates of synthesis of enzymes encoded by five genes of the *trp* operon in *E. coli* cells growing in the presence of tryptophan ?

15) Describe the a) similarities and b) differences between the ways in which gene expression is regulated at the *E. coli lac* and *trp* operons. Confine your answer to a consideration of just the repressor-operator systems.

14.9 ANSWERS

Self-assessment Questions

1) i) Constitutive

ii) The enzymes are expressed constitutively because glucose is metabolised in virtually all cells. However, the levels of enzyme are regulated to prevent runaway synthesis or inadequate synthesis.

iii) The organisms that are able to regulate expression of their genes can adjust quickly to the changed environment, by either switching on or off the related metabolic processes.

iv) By studying the synthesis or lack of synthesis of enzyme in cells grown on chemically defined media. If the enzyme is synthesised only in the presence of a certain metabolite or a particular set of metabolites, it is inducible. If it is synthesised in the absence but not in the presence of a particular metabolite or group of metabolites it is repressible.

2) i) to iii) See Subsection 14.4.1

iv) Turning on and off a set of genes as a unit, simultaneously or by a signal.

3) 1) c

2) d

3) a

4) e

5) b

4) i) a) Constitutive synthesis of the *lac* enzymes

b) Constitutive synthesis of the *lac* enzymes

c) Uninducibility of the *lac* enzymes

d) No β-galactosidase activity

e) No β-galactoside permease activity

ii) The ratio of the amount of each components is maintained constant without having each one separately regulated.

iii) Not always, especially if needed in large quantities.

5) i) Negative regulation

ii) Low

iii) No. In the first case it would not make cAMP. In the second case, it would not make CAP.

iv) No

v) Catabolite repression has apparently evolved to assure the use of glucose as a carbon source when this carbohydrate is available, rather than less efficient energy sources.

6) i) With a repressor, the inducer is usually an early (often the first) substrate in the pathway and the repressor is inactivated by combining with the inducer. With an aporepressor, the effector molecule is usually the product of the pathway, and the aporepressor is activated by the binding.

ii) An attenuator is not a binding site for any protein, and RNA synthesis does not begin at an attenuator. An attenuator is strictly a potential termination site for transcription.

Terminal Questions

1) See text for clarification.

2) See text for clarification.

3) A constitutive phenotype can be the result of either a *lac I*$^-$ or *lac O*^{c} mutant.

4) For the *E. coli lac* operon, write the partial diploid genotype for a strain that will produce β-galactosidase constitutively and permease by induction *lac I*$^+$ *lac O*^{c} *lac Z*$^+$ *lac Y*$^+$ / *lac I*$^+$ *lac O*$^+$ *lacZ* − *lacY*$^+$. It cannot be ruled out that one of the repressor gene is *lac I*−.

5) Possibly by inhibiting the enzyme adenyl cyclase which catalyses the synthesis of cAMP from ATP.

6) Positive regulation, the CAP–cAMP complex has a positive effect on the expression of *lac* operon. It functions by switching on the transcription of structural genes in the operon.

7) Yes, in the gene coding for CAP; some mutation in this gene may result in a CAP that binds to the promoter in the absence of cAMP. Also, the mutations in the gene(s) that code for protein(s) that regulate the cAMP level as a function of glucose concentration.

8) Negative regulatory mechanisms such as that involving the repressor in the lactose operon block the transcription of structural genes of the operon, whereas positive mechanisms such as the CAP–cAMP effect on the *lac* operon promote the transcription of the structural genes of the operon.

9) The sign '+' denote β-galactosidase is produced and '—' sign represents not produced.

Genotype	Lactose absent	Lactose present
1	—	+
2	+	+
3	+	+
4	+	+
5	—	—
6	—	+
7	+	+
8	—	+
9	—	+
10	+	+

10) See subsection 14.4.5.

11) The CAP in a complex with cAMP is required to facilitate RNA polymerase binding to the *lac* promoter. The RNA polymerase binding occurs only in the absence of glucose and only if the operator is not occupied by repressor, i.e., if lactose is absent. A mutation in the CAP gene would render the *lac* operon incapable of expression because RNA polymerase would not be able to recognise the promotor.

12) Transport of lactose and many other sugars in bacteria are mediated by a transport protein in the cell membrane, e.g., lactose permease. Therefore, it is essential that some synthesis of *lac* proteins occurs and complete shutdown of transcription must not occur. In contrast, as long as sufficient quantities of amino acid, e.g., tryptophan are present in a cell, no synthesis of the enzymes needed to make the amino acid is required. Furthermore, attenuation depends on pauses in polypeptide synthesis and can be affected by variations in amino acid concentration, e.g., tryptophan, but not by variation in other metabolites.

Repression/derepression of *trp* operon occurs at the level of transcript intiation modulating the frequency at which RNA polymerase initiates transcription from the *trp* operon promoters. Attenuation modulates *trp* transcript levels by altering the frequency of termination of transcription within the *trp* operon leader region (*trp L*).

14) Deletion of *trp L* region would result in the levels of the tryptophan biosynthetic enzymes (in cells growing in the presence of tryptophan) being increased because attenuation would no longer occur if this region were absent.

15) Write in your own words.

UNIT 15 REGULATION OF GENE EXPRESSION AND DEVELOPMENT IN EUKARYOTES

Structure

15.1 INTRODUCTION

In Unit 14 you have seen that in prokaryotes gene expression is commonly regulated by an operon consisting of structural genes and adjacent controlling sites. The discovery of operons in prokaryotes initiated search for similar regulatory systems in eukaryotes. But so far no operons have been found in eukaryotes. Nevertheless, there are some commonalities between the gene regulation systems of lower eukaryotes and prokaryotes. It implies that gene regulation in eukaryotes must occur by ways other than by operons. In this unit we shall discuss some salient aspects of regulatory systems in eukaryotes.

The eukaryotic gene regulation can be classified into two categories: short-term and long-term regulation. The short-term regulation is just like the regulation by operons as seen in bacteria and viruses. It involves regulatory events in which gene sets are quickly turned on or off in response to changes in environmental or physiological conditions. On the other hand, the long-term regulation involves those events that are required for an organism to develop and differentiate.

Objectives

After studying this unit you should be able to:

- compare the genetic organisation of the eukaryotes and prokaryotes, and relate these features to regulation of gene expression in eukaryotes (Sections 15.2 to 15.5);
- explain short-term gene regulation pointing out at which level it operates (Section 15.3);
- describe the key roles of gene regulation in development and differentiation (Section 15.4); and
- explain the most accepted model for regulation of gene expression in eukaryotes (Section 15.5).

15.2 GENETIC ORGANISATION OF PROKARYOTES VS EUKARYOTES

The prokaryotes and eukaryotes differ from each other with regard to transcription, translation and spatial organisation of DNA. Given below are some of the differential features of the two, that are relevant to regulation.

1) In an eukaryote, usually one type of the polypeptide chain is translated from a completed mRNA molecule. Thus, polycistronic mRNA of the type seen in prokaryotes is not present in eukaryotes.

2) The eukaryotic DNA is associated with histones forming chromatin, and to numerous nonhistone proteins. Only a small fraction of the DNA is bare. In bacteria, most of their DNA is free, and sometimes proteins are present in the folded chromosome.

3) A significant fraction of the DNA of eukaryotes consists of moderately or highly repetitive nucleotide sequences. Some of these sequences are repeated in tandem copies, but others are not. Bacteria contain camparatively lesser amount of repetitive DNA, most of which is confined to the rRNA and tRNA genes, and a few specific sequences such as the insertion sequences.

4) A large fraction of the eukaryotic DNA is untranslated, as most of the nucleotide sequences do not code for proteins.

5) Some eukaryotic genes are expressed and regulated by certain mechanisms for rearranging some DNA segments in a controlled way and for increasing the number of specific genes when needed.

6) The eukaryotic genes can be split into exons and introns. The introns must be removed during processing of the RNA transcript before translation begins.

7) The mRNA, in eukaryotes is synthesised in their nucleus from where it is transported across the nuclear envelope to the cytoplasm where it is utilised. Owing to the absence of nucleus in bacteria, such a compartmentalisation is not seen.

Now that you have studied the salient differences between the genetic organisation in prokaryotes and eukaryotes, you should note in the following sections how some of these features are incorporated into particular modes of eukaryotic gene regulation.

15.3 THE SHORT-TERM GENE REGULATION

In this section, we shall take up some examples of short term regulation of gene expression in eukaryotes. Short-term regulation as you know involves rapid responses to changes in environmental or physiological conditions. *It operates at the transcriptional level.* As you study the following examples, keep in mind the short-term regulation of gene expression in bacteria and compare the molecular events involved.

15.3.1 Quinic Acid Metabolic Genes in *Neurospora crassa*

In *Neurospora*, the genes involved in metabolism of quinic acid (qa) as a source of carbon, constitute a simple and genetically well-characterised system for studying gene regulation in eukaryotes. This system comprises a cluster of five structural genes and two regulatory genes arranged as shown in Figure 15.1.

These five structural genes are *inducible* and their products are synthesised *coordinately* thus showing parallels with some bacterial operons. Two regulatory genes are present immediately adjacent to the structural genes. The regulatory gene *qa-1F* codes for a protein — *activator* and the second one *qa-1S* codes for a *repressor* protein. The activator protein (gene qa-1F) is needed both for the transcription of its own mRNA and for the transcription of all the structural genes except qa-x. The gene qa-1S on the other hand, codes for a protein that has a repressor (negative) function. If quinic acid is absent, the repressor protein interacts with qa, the effector molecule, and blocks transcription of qa-1F. Transcription of the qa-x structural gene seems to be controlled mainly by qa-1S and to a much lesser extent by qa-1F. Each of the four structural genes under qa-1F control is transcribed from two to four qa-1F independent promoters. Unlike bacterial operons, this gene cluster gives no evidence for operator-like controlling regions, and so it is not an operon. This regulatory system, however, functions to maintain cellular levels of particular chemicals, enabling the organism to adjust to changing physiological environment.

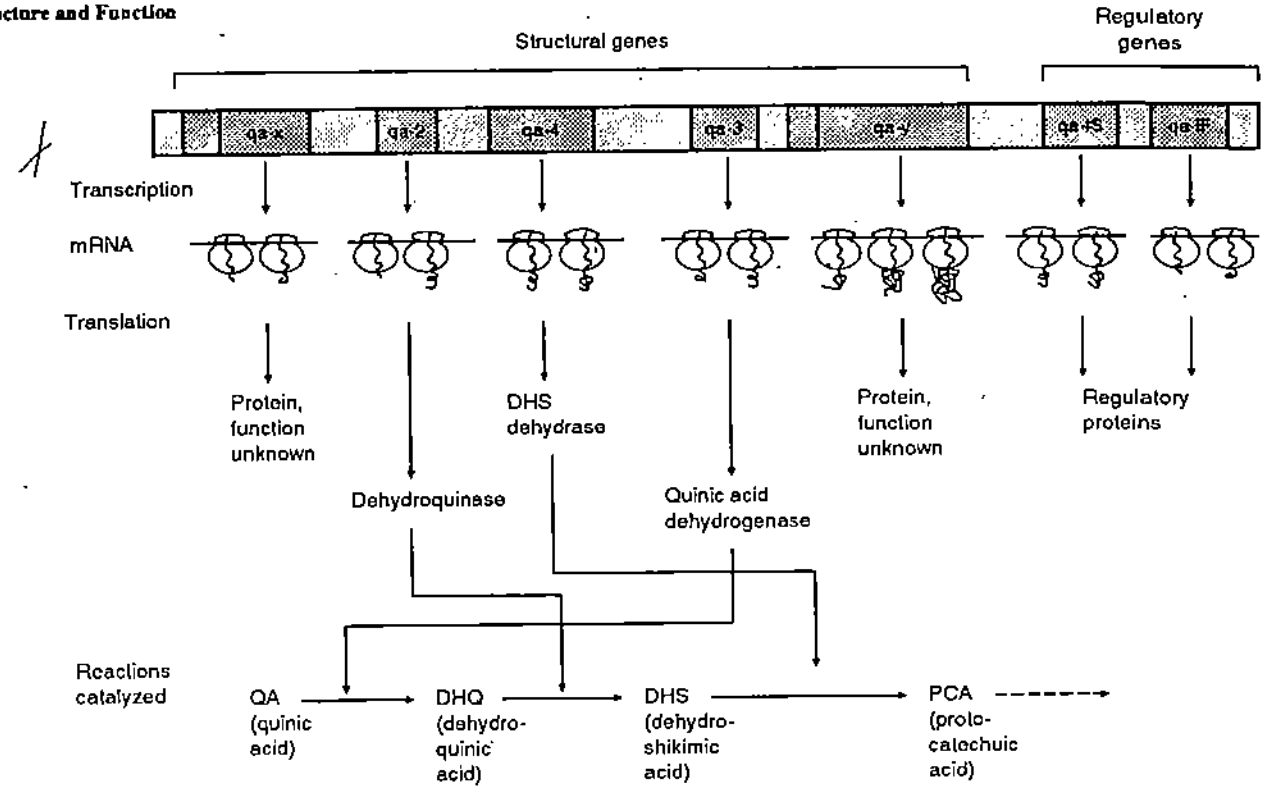
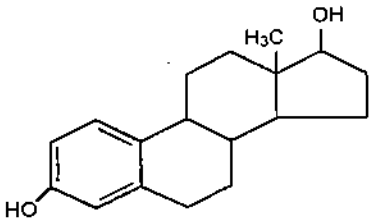
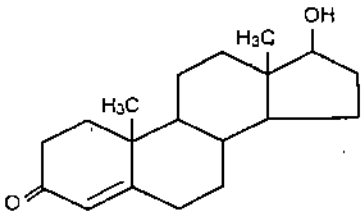


Fig.15.1: Organisation of quinic acid (qa) catabolising genes in *Neurospora crassa*.



Estrogen



Testosterone

Fig. 15.2 : Chemical structures of the steroid sex hormones estrogen (female) and testosterone (male). The steroid hormones are relatively small molecules with molecular weights around 300. They have a four-ring structure and are synthesised from cholestrol. The various steroid hormones have different side chains and different banding patterns within the rings. These differences permit them to be recognised by different receptor proteins that are present in the cytoplasm of various target cells.

15.3.2 Hormonal Regulation

Hormones are also known to regulate transcription in higher eukaryotes. They are small molecules or polypeptides that are carried from *hormone producing cells* to the *target cells*. The steroid hormones constitute a prominent class amongst all the hormones. They are made up of small molecules synthesised from cholestrol. The principal sex hormones are steroids (see Fig. 15.2). Many of the steroid hormones act by turning on the *transcription* of specific sets of genes. It, therefore, means that a hormone that regulates transcription must somehow signal the DNA. The mechanisms by which this occurs is outlined in Figure 15.3. A steroid hormone penetrates a target cell through diffusion because steroids are hydrophobic (nonpolar) molecules and they pass freely through the cell membrane and the nuclear envelope. The nuclei of target cells contain specific receptor proteins for the steroid hormones. These receptor proteins form a complex with the hormone. During this process the receptor protein undergoes modifications in three-dimensional shape. This enables the hormone-receptor complex to bind with particular sequences in the DNA and stimulate transcription. The non-target cells do not contain the specific receptor proteins and so are unaffected by a particular hormone.

A well studied example of induction of transcription by a hormone is the stimulation of synthesis of ovalbumin in the chicken oviduct by the steroid sex hormone estrogen (Fig.15.2). When chickens are injected with estrogen, oviduct tissue responds by synthesising ovalbumin mRNA. This synthesis continues as long as estrogen is administered.

Once the hormone is withdrawn, the rate of synthesis decreases. Both before the injection of the hormone and sixty hours after withdrawal, no ovalbumin mRNA is detectable. When estrogen is given to chickens, only the oviduct synthesises mRNA because other tissues lack the hormone receptors.

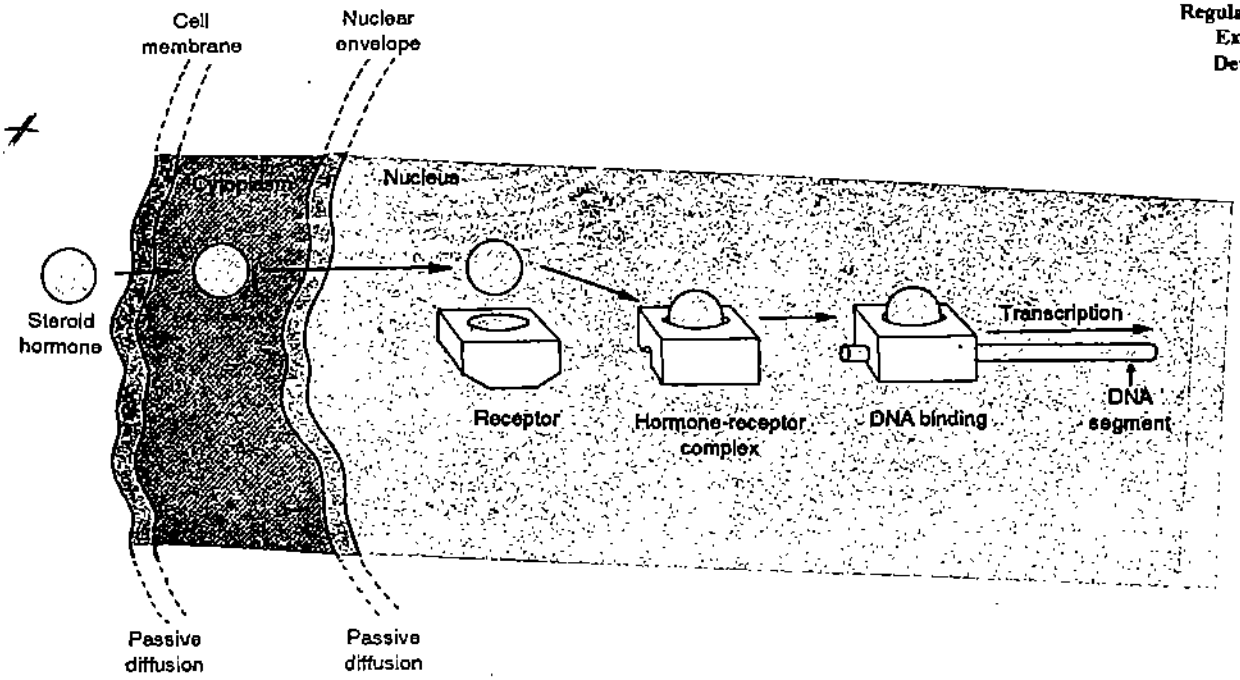


Fig. 15.3: Diagrammatic representation of the action of steroid hormone. The steroid hormone reaches a DNA molecule and triggers transcription by binding with a receptor in the nucleus to form a transcriptional activator. The steroid hormone enters the cytoplasm and nucleus by diffusion.

Besides animals, there are several examples of hormonal regulation of gene expression in plants too. Plant hormones are categorised into the following five main classes: ethylene gas, abscisic acid, auxins, cytokinins and gibberellins (Fig. 15.4). Amongst them, the role of gibberellins has been well-studied. These hormones have

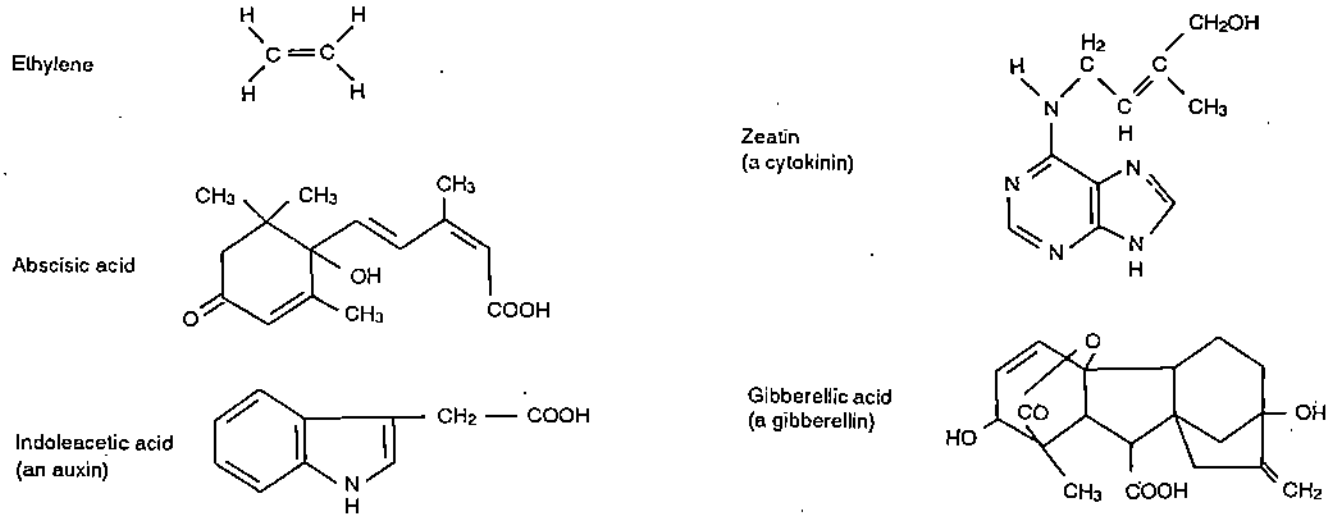


Fig. 15.4: Chemical structure of five plant hormones.

actions similar to those of steroid hormones. Gibberellins *stimulate transcription* and thereby increase the production of specific proteins. These proteins bring out profound changes in plant cell form during its differentiation. The use of gibberellins in making dwarf plants grow tall is well known. Apart from that, gibberellins are also known to stimulate germination of seeds. It's role in germination of the wheat seeds is diagrammatically shown in Figure 15.5. Have a thorough look at this figure.

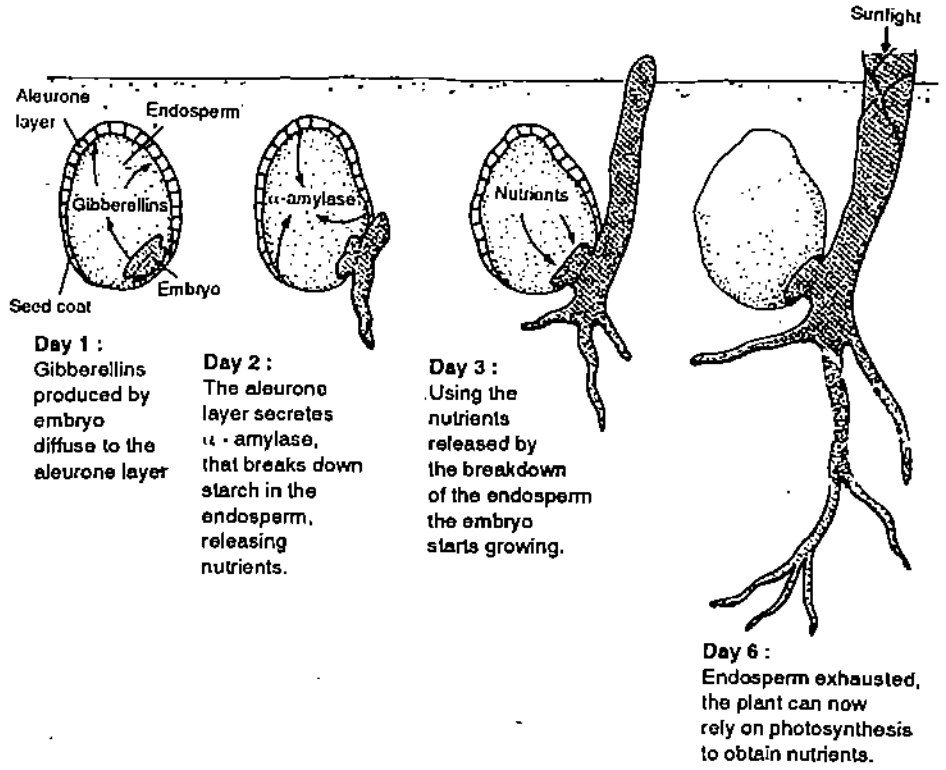


Fig.15.5: The effect of gibberellins on the germination of wheat.

During germination, the wheat seed produces gibberellins that diffuse to the aleurone layer, which is the outermost layer of the endosperm. As a result of the effect of the gibberellins on gene expression, and particularly on the genes for the enzyme ∝-amylase, the aleurone layer synthesises and secretes ∝-amylase, which breaks down the endosperm, releasing nutrients for the embryo growth. Eventhough the action of gibberellins is well known, the molecular details of their action are still not clear.

SAQ 1

Which of the following is not a true statement?

i) In eukaryotes too polycistronic mRNA as seen in prokaryotes is present.

ii) The DNA of eukaryotes is unique as it is associated with various histone and non-histone proteins.

iii) A large portion of eukaryotic DNA is untranslated.

iv) The genes of prokaryotes are composed of exons and introns.

v) The site of synthesis of mRNA in eukaryotes is the nucleus, from where it is transported to the cytoplasm where it is utilised.

vi) The quinic acid metabolic genes of *Neurospora* being inducible, and their products being synthesised coordinately, show parallels with the bacterial operons.

vii) The expression of quinic acid metabolising genes is regulated in operon like manner.

viii) The steroid hormones act mostly at the post-translational level.

15.4 GENE-REGULATION IN DEVELOPMENT AND DIFFERENTIATION

In this section you will study the long-term gene regulation. There are two key terms used in describing the long-term gene regulation. First one is **development**. It refers to the process of regulated growth resulting from the interaction of the genome with cytoplasm and the environment. It involves a programmed sequence of

phenotypic events that are irreversible. The total phenotypic changes of an organism constitutes its life cycle. The second term is **differentiation**. This is the most spectacular aspect of development. It involves the formation of different types of cells, tissues and organs from a zygote through the process of specific regulation of gene expression. The process of differentiation results in cells having characteristic structural and functional properties. Therefore, we can say that the processes in differentiation and development are the result of highly programmed patterns of gene activation and gene repression.

A great deal is known to us about the *lac* operon in bacteria, but in eukaryotes information in this regard is meagre. Equally mysterious is how gene expression is regulated and coordinated during the development of a eukaryotic organism from a fertilised egg or zygote to an adult. In these organisms, the additional complication is that the activities of a multitude of different cells must be coordinated. So, we are dealing not with the control of gene expression in a single cell but with the coordinated regulation of gene activity in a number of different cells, perhaps in different parts of the organism. Despite all these factors, it is not an impossible task. Considerable progress has been made in understanding the regulation of gene expression in the fruit fly *Drosophila melanogaster*.

Drosophila is an ideal system for studying development. It is a relatively simple, segmented organism that develops through distinct larval stages. Since the embryo develops externally, as opposed to the *in utero*, changes can be readily observed. The entire life cycle, from fertilised egg to adult is accomplished in a matter of 2 weeks (see Fig. 15.6).

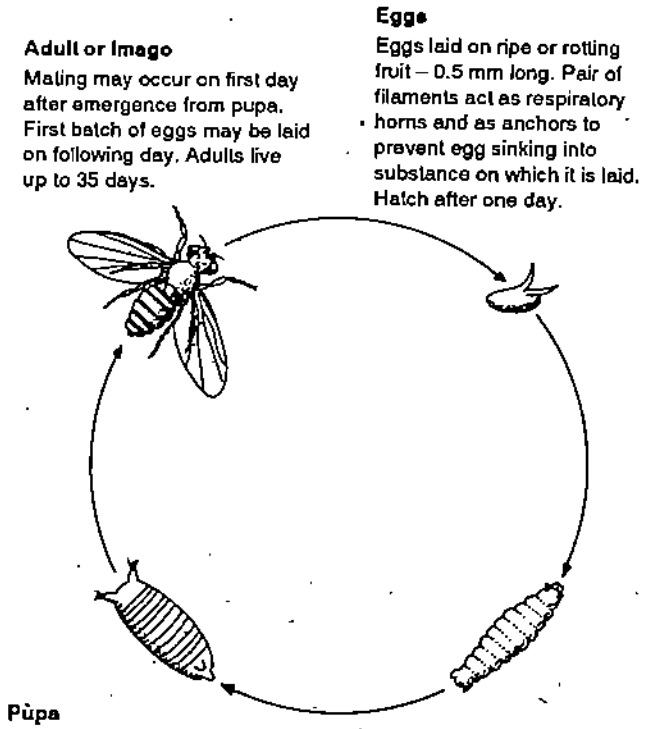


Fig. 15.6: Life cycle of *Drosophila*

The discovery over the years of a number of bizzare developmental mutations in *Drosophila* has been very crucial, as these mutants have turned out to be invaluable tools in understanding differentiation and development. These aberrations which alter the fates of cells, during embryogenesis and commit them to new developmental pathways, are known as "homeotic" mutations. For example, the mutation *Antennapedia* causes a pair of normal-looking legs to grow from the head segment, replacing the antenna (see Fig. 15.7a). The *Bithorax* mutations lead to the development of a second pair of partial or complete wings in an inappropriate thoracic segment.

Genes that get altered by these developmental mutations have been studied, and are found to be extremely complex both in terms of structure and of expression. However, one intriguing feature of several of these genes has emerged: *the homoeobox*, a short segment of 180 bp coding for a 60 amino acid portion of the gene product which is similar in atleast ten of the relevant *Drosophila* genes (Fig. 15.7b). The fact that these genes have a structural relationship, that is, the biological information that each carries is similar in some respect, suggests that possibly the gene products are related in terms of function. But the biggest surprise to molecular biologists has been the discovery that homoeoboxes also exist in genes of frog, mouse and man. These genes in vertebrates are totally different from the *Drosophila* genes, except for the presence of homoeobox and the fact that the genes in vertebrates are also believed to be involved in development. From this one question arises. Are the genetic instructions for developmental processes universal? Thomas Hunt Morgan and his colleagues showed that the study of a simple organism like *Drosophila* could provide information of general relevance to genetics. It is exciting to think that *Drosophila* could also provide the key to understanding the complex events that underlie development in man.

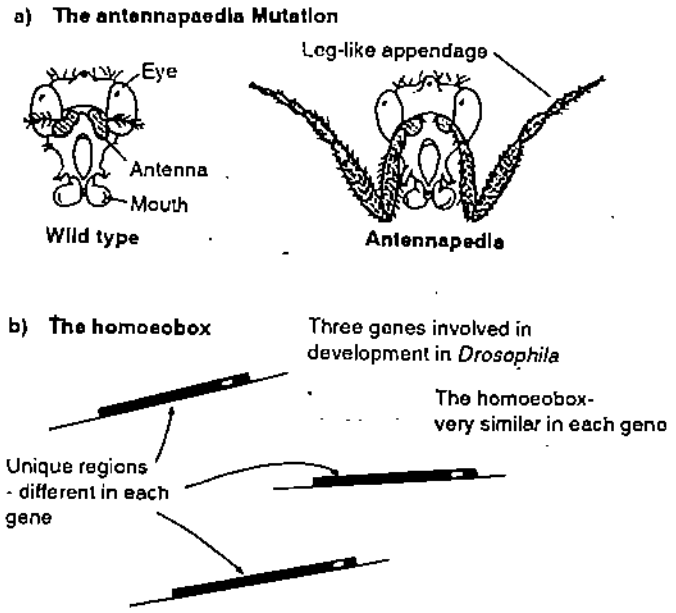


Fig. 15.7 : The development genetics of *Drosophila*. (a) The wild type (left) and *Antennapedia* mutant (right). (b) different genes involved in development contain homoeoboxes.

15.5 GENE REGULATION IN EUKARYOTES

The control of gene expression in eukaryotes is much more complex and less well defined than the prokaryotes and viruses. The eukaryotic cell also contains greater amounts of genetic information, and its DNA is complexed with a wide variety of proteins to form chromatin. They also have more than one chromosome contained in a nuclear envelope. Translation occurs in the cytoplasm. In multicellular organisms, tissue and cell-specific gene products are restricted, even though each cell contains a full genetic complement. In comparison to bacteria and viruses, the eukaryotic organisms are not amenable to mutagenesis and experimentation. Thus, our understanding of the regulatory mechanisms in eukaryotes is not as precise as that of the prokaryotes and bacteria. Many speculations have been made and many models have been proposed to explain regulation of gene expression in eukaryotes.

One of the popular early models was proposed by R.J. Britten and E.H. Davidson (Fig. 15.8). This model proposes an integrated regulation of sets of structural genes by means of moderately repetitive regulator genes. It takes into account the observed interspersion of single copy DNA sequences and repetitive DNA sequences. See Figure 15.8 carefully.

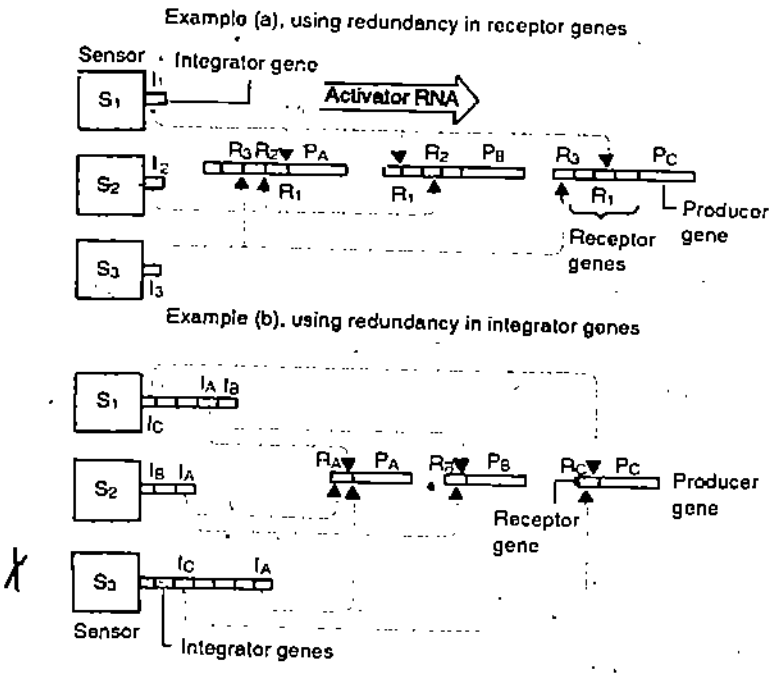


Fig.15.8: The Britten-Davidson model of regulation of gene expression in eukaryotes showing two variations of integrated regulation. (a) represents a system based on redundancy of 'receptor' genes, and in (b) system based on redundancy of integrator gene is shown. S_1, S_2, and S_3 are three sensor genes that respond to three different signals, that may be hormone-receptor complexes. The diagrams represent the events that occur after sensor genes have triggered the transcription of their respective integrator genes (I_1, I_2, I_3 or I_A, I_B, I_C). The products of integrator genes - 'activator RNAs' diffuse from their site of synthesis (integrator genes) to their sites of action (receptor genes). With the binding of activator RNAs to their respective receptor genes somehow triggers the transcription of the contiguous producer genes P_A, P_B, P_C. Depending on which integrator gene(s) is activated by sensor gene(s), one, two or all three of the producer (structural) genes may be turned on. From: R.J. Britten and E.H. Davidson, 1969. Gene Regulation for Higher Cells : A Theory. Science, Vol. 165, pp 349-357.

In the Britten-Davidson model, the specific *Sensor genes* represent sequence-specific binding sites, that respond to specific signals such as hormone-receptor protein complexes. When the sensor genes receive the right signal, they activate the transcription of the adjacent *integrator genes*. The products of integrator gene then interact in a sequence-specific manner with the *receptor genes*. According to Britten and Davidson, the integrator gene products were *activator RNAs* that interact directly with the receptor genes to trigger the transcription of the contiguous *producer genes*. The producer genes are considered analogous to the structural genes in prokaryotic operons. They also point that, it would make no difference whether the active integrator gene-products are RNA molecules or proteins.

When either the receptor genes (see Fig. 15.8a) or the integrator genes (see Fig. 15.8b) are redundant, various combinations of the producer genes can be turned on in response to different signals. And if both the integrator genes and the receptor genes are redundant, complex integrated circuits of gene expression can easily be devised, although testing the validity of such models is far more difficult.

The most attractive feature of this model is that it provides a plausible reason for observed pattern of interspersion of moderately repetitive sequences and single-copy DNA sequences. There are strong evidences showing that most structural genes (producer genes) are actually single copy DNA sequences. According to the above model, the adjacent moderately repetitive DNA sequences contain various kinds of regulator genes (sensor, integrator and receptor genes).

Further studies comparing the complexity of heterogeneous nuclear RNA (hnRNA) populations and mRNA populations in different types of cells have shown that hnRNA populations are usually more complex (as they contain more distinct sequences) than mRNA populations. These findings suggest that considerable regulation occurs post-transcriptionally during RNA processing, that is, in the hnRNA → mRNA stage.

Based on these observations, Davidson and Britten proposed a second model, the Davidson-Britten model (see Fig. 15.9). This model proposes that gene expression is regulated at the level of RNA processing. According to this model, most of the structural genes are located in *constitutive transcription units*, which are transcribed at basal levels in all cells. These constitutive transcripts are processed, only in cells that contain the appropriate *integrating regulatory transcripts (IRTs)*. The *IRTs* are transcribed in a cell-specific manner, and these must be present before the *constitutive transcripts (CTs)* of the structural genes can be processed into mRNAs.

These 'integrating regulatory transcripts' contain repetitive sequences that interact with different structural gene transcripts like the repetitive 'integrator' genes interacted with different 'receptor' genes in the original Britten-Davidson model. The main difference is that the regulation occurs post-transcriptionally during RNA in the new Davidson-Britten model, rather than transcriptionally as in the original model.

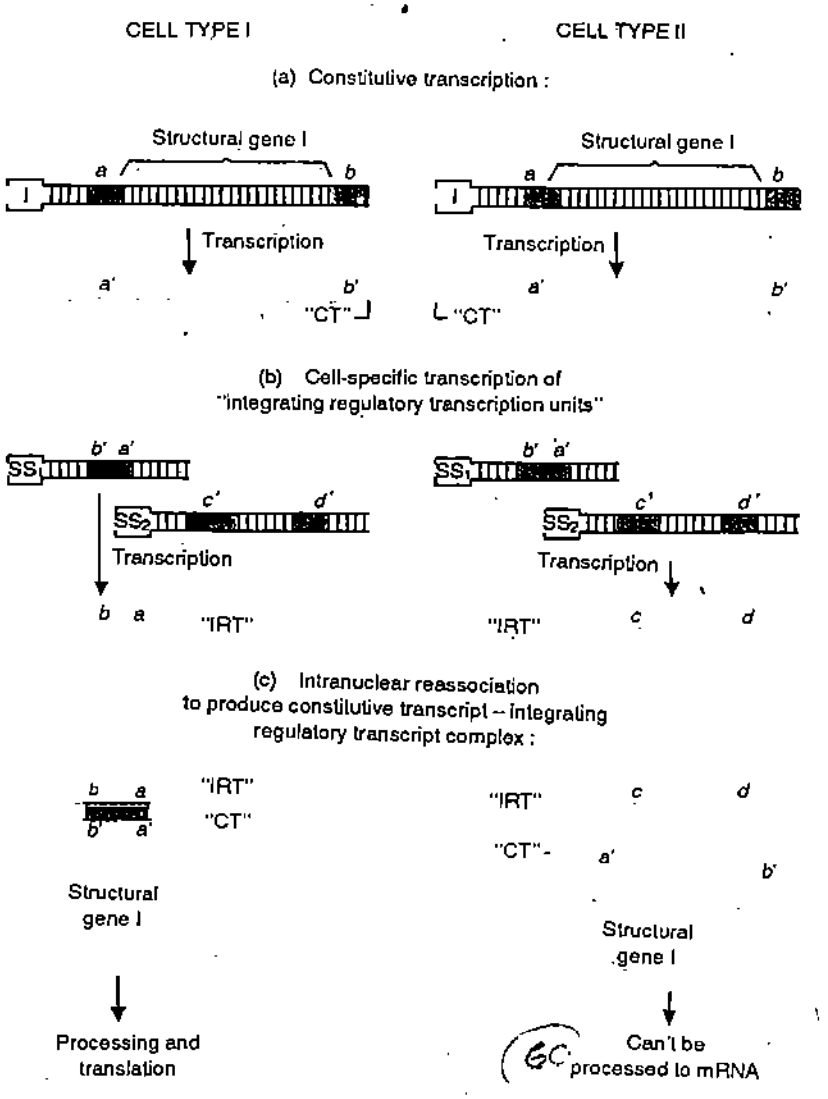


Fig.15.9: The new Davidson-Britten model for the regulation of gene expression. It is based on RNA processing level in eukaryotes. a) The majority of the structural genes are assumed to be located in 'constitutive transcription units'. These are transcribed continually in all cells. The letter *I* denotes the transcription initiation site. The letters *a*, *a'*, *b* and so on represent middle-repetitive sequences that are proposed to be involved in regulating gene expression. The sequences *a'*, *b'*, *c'* and *d'* are complementary to sequences *a*, *b*, *c* and *d* respectively. (b) Structural gene expression is regulated by repetitive RNA sequences transcribed in a tissue or cell-specific manner as components of "integrating regulatory transcription units (IRTUs). The transcription of IRTU is controlled by nucleoprotein 'sensors' (*SS*), which respond to specific external signals. Different tissues or cell types transcribe different IRTUs. (c) A give structural gene constitutive transcript (*CT*) can be processed into mRNA and thus expressed only if it forms a complex with 'integrating regulatory transcript' (*IRT*). Thus, for the expression of a particular structural gene in any given cells, the presence of appropriate nuclear

IRT is required. Different types of cells will have *IRTs* carrying overlapping populations of repetitive sequences, so that they process some of the same structural gene transcripts and some different gene transcripts. The complex *IRT* populations of different cells are, however, not illustrated in this diagram. (Based on E.H. Davidson and R.J. Britten, 1979. Science, Vol. 204 : 1052-1059).

SAQ 2

Match the items given in column I with those of column II. Write your answer in the space provided.

I		II
i) Long-term gene regulation	[ ]	a) several genes having similar base pair sequence of a short segment.
ii) Antennapedia	[ ]	b) respond to certain hormone-receptor proteins
iii) Homoeobox	[ ]	c) development and differentiation
iv) Sensor genes	[ ]	d) regulation of gene expression at the level of RNA processing
v) Davidson-Britten model	[ ]	e) homoeotic mutation

15.6 SUMMARY

In this unit you have learnt that:

- Eukaryotes being more complex also contain more genetic material. They are assumed to have a more complex system of regulation.
- The discovery of operon system in bacteria prompted a search for similar control systems in eukaryotes. No operons however, have been found in eukaryotes, although the genes for related functions in some biochemical pathways are closely linked or are contiguous. More often the genes that are coordinately regulated are dispersed throughout the genome.
- The regulation in higher eukaryotes can be either short-term or long-term. In higher eukaryotes one of the well studied systems of short-term gene regulation is the control of enzyme synthesis by steroid hormones. The specificity of hormone action is caused by the specific array of receptors in the genome of each hormone. The long-term regulation of gene expression refers to the regulation during the development and differentiation of organisms.
- One of the most popular models of regulation of gene expression in eukaryotes is the Davidson-Britten model. According to this model, regulation occurs at the level of RNA processing.

15.7 TERMINAL QUESTIONS

1) Why is gene regulation assumed to be more complex in a multicellular eukaryote than in a prokaryote? Why is the study of this phenomenon in eukaryotes more difficult?
2) Are operons more common in bacteria or in higher organisms?
3) How do steroid hormones induce transcription of eukaryotic genes?
4) Why is the model proposed by Britten and Davidson more acceptable than the operon model for explaining regulation in cells of higher animals?

15.8 ANSWERS

Self-assessment Questions

1) i)
   iv)
   vii)
   viii)

2) i) c
   ii) e
   iii) a
   iv) b
   v) d

Terminal Questions

1) **Hint**: One, they contain greater amount of genetic information. Two, their DNA is associated with various proteins, and it is present in chromosomes enclosed in nuclear envelope. In addition, various kinds of tissues/cells have cell-specific gene products that are associated with the particular genes.
   You may also refer to Section 15.2 and Section 15.5.

2) Operons are common in prokaryotes. Gene clusters that resemble operons (actually are not operons) exist in several microorganisms, e.g. *Neurospora*. No operon is known in higher eukaryotes.

3) Steroid hormones pass through the cell to the nucleus where they combine with transcriptional activator proteins, and stimulate transcription.

4) In prokaryotes, the structural genes specifying the enzymes in a metabolic pathway are usually arranged as groups of contiguous genes. This facilitates regulation by operon mechanism. In higher eukaryotes, such genes are usually not in clusters and are often unlinked. The complex patterns of gene expression during development in higher animals almost certainly require complex integrated controls that can govern expression. Genomes of higher eukaryotes contain single copy DNA sequences (structural genes) that are interspersed with middle-repetitive

UNIT 16 MUTATIONS AND MUTAGENESIS

Structure

16.1 INTRODUCTION

Natural selection, as Darwin recognised more than a century ago, favours individuals and populations that inherit characters conducive to their survival and reproduction. [For a detailed discussion on Natural Selection refer to Units 11 and 12 of LSE 07 course.] We must understand therefore, that the generation of biological variation is a central requirement for evolution of species in diverse and changing environment. You have already learnt in the previous blocks that the occurrence of phenotypic variability amongst the living organisms is an absolute requirement also for understanding the inheritance of the genetic characters. One of the basic mechanisms that creates genotypic as well as the phenotypic variation is the mutation or the alteration of DNA within the gene. In recent years, the terms mutation has generally been restricted to processes that result in a direct alteration of gene contents generating the new alleles of a gene that end up coding for a different sequence of amino acids.

The natural occurrence of mutations is very slow. In order to make significant progress on genetice studies in any living system it becomes imperative that the new mutations must be isolated and accumulated at a faster rate. This necessity has led to an active field of study on experimental mutations called mutagenesis. Ever since the time of M.J. Muller (1927) who reported that X-ray could induce mutations in *Drosophila* followed by L.J. Stadler's finding that the same is true in barley, a large number of agents called mutagenes have become available to induce mutations.

In this unit we will discuss the concept, types and molecular basis of mutations and detection of mutations. You will also read about various mutagens, their effects and repair mechanisms to rectify these effects. With the availabity of whole range of mutagens it has become possible to exploit these for advancement of genetic studies and improve the agricultural crops. The study of mutations has also enhanced our understanding of the cancer process as well as abnormal development in detail.

Objectives

After studying this unit you will be able to:

- explain what are mutations, how they arise and how they can be detected,
- differentiate the types of mutations and explain their origin on molecular basis,
- discuss the role of mutagens in inducing mutations and the cellular repair mechanism to correct the DNA damage, and
- discuss how these mutations can be exploited for advancement of genetic studies, and betterment of agricultural crops.

16.2 WHAT IS MUTATION ?

In the year 1791, a New England farmer, Seth Wright, noticed a peculiar male lamb in his small flock of sheep. This lamb had unusually short legs which were somewhat deformed. Mr. Wright recognised the unique advantage of sheeps like this one, as they would not be able to jump the low stone fences and therefore not damage the crops. He carefully bred his sheep and was able to raise a short legged flock. This breed was named **ancon** (Greek—elbow) as the crooked legs resembled the human elbow. The same mutation appeared in a flock of sheep in Norway in 1925 and a separate breed of ancon sheep was established from it.

A similiar event occurred during the latter part of 19th century when a worker in South America spotted a peculiar type of orange. At one end of the fruit there was a shrivelled indented portion which resembled a human navel, but pulp of this orange fruit contained no seeds. You should be able to recognise the potential value of having the seedless orange. By careful budding and grafting this new character could be propagated. An American tourist brought a twig back to California, and thus the great navel orange industry of America was established. What do you get from these examples? Probably that some changes have suddenly arisen in these organisms which became beneficial to mankind.

When such changes occur in natural populations, mutation is said to have taken place. Mutation, therefore, is a **sudden, heritable change in genotype** that involves qualitative change in the genetic material. The change may lead to a corresponding change in the phenotype. Mutations are an extremely important source of genetic variability in living populations. They are the deviations from normal genotypic and phenotypic conditions. The normal conditions are referred to as **wild-type** conditions.

Mutations include changes occurring at chromosomal level as well as at gene level. At chromosomal level, the change in the organisation and structure of chromosome (s) is called **chromosomal aberration** or **chromosomal mutations**. You have already read about chromosomal aberration in previous block. In this unit you will study about the mutations occurring at gene level. When we say mutation we refer to the changes occurring at gene level, i.e. **gene mutation.**

The possibility that new types of inherited characters may appear suddenly was first suggested by Hugo De Vries in 1901 as a result of his experiments on the plant evening primrose, *Oenothera lamarckiana.* He coined the term mutation to explain the variations he observed in crosses involving this plant. Most of the variations observed by De Vries, however, were later found to be chromosomal aberrations rather than mutations. Neverthless, he deserves credit for the formulation of the concept of mutation and its importance from the point of view of evolution.

In addition to their contributions to the genetic variation, mutations are working tools for the geneticist in order to understand the structure and functioning of the gene. Mutations provide insights into basic biochemical processes such as gene expression and development (see Fig. 16.1). Certain mutations that can be easily detected and studied are induced in organisms like bacteria, fungi, fruit flies, certain plants, mice etc. These organisms have short life cycles and therefore are normally used for studying mutations and mutagenesis.

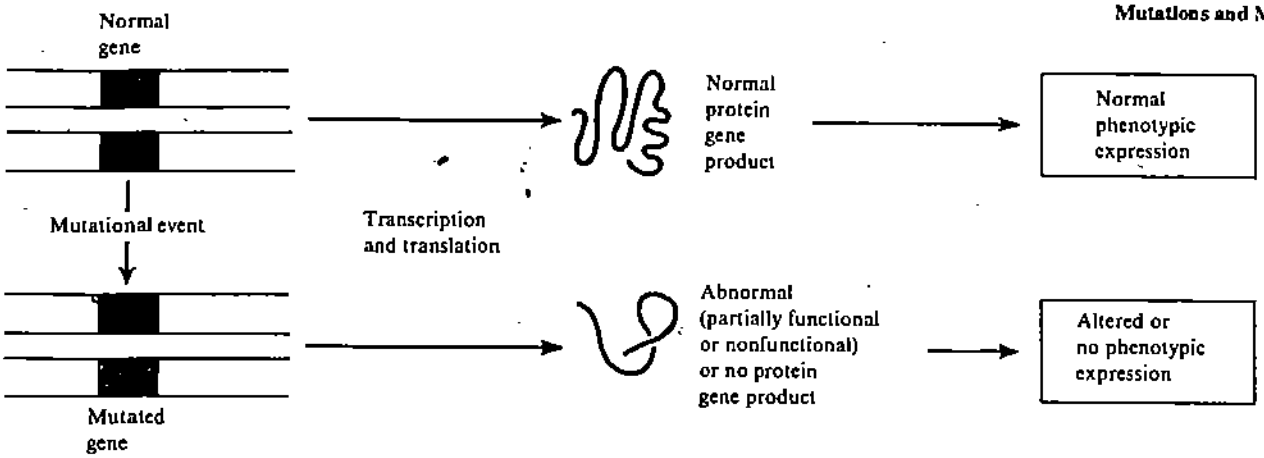


Fig. 16.1: Illustration to show the effect of mutation on gene expression.

After having understood the meaning of the word mutation and its importance let us now discuss various classes of mutations.

Mutations can be classified on the basis of several criteria. None of these are mutually exclusive. Instead they depend on simply which aspect of mutations are being discussed. In the following subsections we will discuss three ways of classifying mutations. But first do the following SAQ.

SAQ 1

Explain what do you understand by mutation ? Give any four examples of mutations.

.....

.....

.....

.....

.....

.....

.....

.....

16.2.1 Somatic and Gametic Mutations

Mutations may occur in somatic cells as well as in germ cells. Mutations arising in **somatic cells**, i.e. body cells are not transmitted to future generations and die with the death of the individual. These are called **somatic mutations.** Let us see how mutations in somatic cells affect the organisms.

In the tissues of adult organisms many cells perform similar function. So mutation in a single cell of a tissue may not impair the organism even if the mutation is detrimental. This has two aspects: first, mutation might occur in a gene that are not active: in other words it may occur in genes that are not essential to the functions of that cell. Second, even if the critical gene in a cell or a group of cells is affected there are still thousands of normal, unaffected cells to perform the same function in that tissue. However, mutations in somatic cells may sometimes cause damage, including cancer, to the parts of the body that arise from the mutated cells. As you have read earlier, somatic mutations are not significant from the stand point of heredity. However, if these mutations arise early enough during embryonic development, they may express themselves in the somatic cells. Some of such somatic mutations can be highly beneficial to mankind. You may remember the earlier examples of navel orange. Navel orange arose due to mutation in the

meristematic cells of a bud. Such mutations can be propagated vegetatively and retain a constant phenotype. You can see an example of somatic mutation in Fig. 16.2.

Fig. 16.2 : Mutant flower heads of Zinnia. The somatic mutation was from splotched colour (peppermint) to solid colour (a).

Gametic mutations are of greater concern because they have the potential of being expressed in all cells of an offspring. Such mutations perferably occur in **germinal cells** i.e., reproductive cells of an organism. The germinal cells give rise to gametes which are the carriers of genetic information from generation to generation. So you see that mutations in germinal cells are bound to be heritable. It is possible that mutations in germ cells may cause no noticeable abnormality in the individual in which they occur, but will be passed on to its offspring and may be expressed in the offspring's phenotype. Gametic mutations may occur on autosomes or sex chromosomes of the germ cells or gametes. These are called **autosomal** and **sex-linked mutations** respectively.

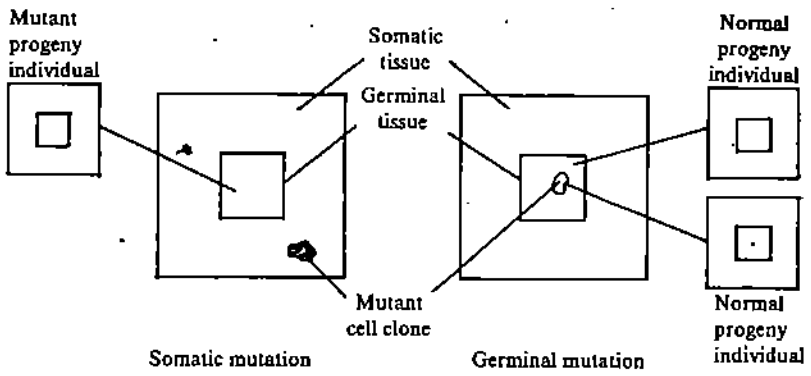


Fig. 16.3: Diagrammatic representations of the consequences of somatic and germinal mutations.

16.2.2 Spontaneous and Induced Mutations

All the mutations are categorised as either spontaneous or induced, on the basis of the way they arise in the organism.

Spontaneous mutations are those mutations that arise randomly in nature, strictly by chance. No specific agent may be required for their origin. They essentially arise due to inherent errors in DNA replication and transmission processes. The internal environment of the cell also plays a significant role in inducing mutational events. Each organism has a characteristic spontaneous mutation rate. The mutation rate is the probability that a gene undergoes mutation in a single generation. Some examples of which are shown in Table 16.1. You can infer by studying the table that spontaneous mutation rates are relatively low, most occurring one in a million genes or more. Also the spontaneous rate for different genes of the same organism may vary greatly. Measurement of mutation rates is important in population genetics, studies of evolution and in analysing the effects of environmental mutagen.

Table 16.1: Some examples of spontaneous mutations rates for various organisms.

Organism	Trait	Mutation Rate
E. coil	histidine auxotrophy	2×10^{-6}
(various sources)	streptomycin sensitivity	1×10^{-8}
	phage T_1 resistance	2.3×10^{-8}
Drosophila males	brown eyes	3×10^{-5}
	eyeless	6×10^{-5}
	yellow body	12×10^{-5}
	colourless kernel	2×10^{-6}
	shrunken kernel	1.2×10^{-6}
Human	achondroplasia	1×10^{-5}
	aniridia	2.9×10^{-6}
	retinoblastoma	6.7×10^{-6}
	muscular dystrophy (Duchenne type)	9.2×10^{-5}

Mutation rate for bacteria expressed as the number of mutations per cell per generation; and for eukaryotes as the number of mutations per gamete per generation.

Induced mutations are the mutations that occur in response to obvious externally applied agents. We referred to earlier H.J. Muller's findings that x-rays could induce mutations in *Drosophila*. Muller later received a Nobel Prize for his contribution. The agents that induce mutations are called mutagenic agents. Various forms of radiation sources and many chemicals are the known as mutagenic agents. We will discuss about mutagenic agents later in this unit. It should however, be clarified that induced mutations are not new types of mutations. Instead mutagenic agents accelerate the spontaneous mutation rates several folds, thus enhancing the chance of occurrence of a mutation.

16.2.3 Categories of Mutations

Mutations are also classified on the basis of their effects on the organism. In this subsection we will discuss these categories of mutations.

Morphological mutations, the most obvious mutations are those which affect the **morphological traits** of the organism. These are **visible mutations** which result in the phenotypic change that alters the morophology, i.e. visible characters of an organism. These are usually outward characters such as shape, colour or size. Curl wings in **Drosophila** and dwarfism in peas are some examples of morphological mutations (see Fig. 16.4).

Fig 16.4: A rare morphological mutation in *Rubus parviflorus* (thimble berry). The mutation arose spontaneously in nature. Flowers 1,2 and 3 are normal wild type and flower 4 is a mutant. The mutation causes increase in the number of petals. Such mutants are use in horticulture to increase the showiness of the flowers.

Nutritional mutations may not always be observed visually, but can be identified only by biochemical analysis. The inability to synthesise an amino acid or a vitamin in bacteria and fungi is a typical nutritional mutation. Nutritional mutants are called **auxotrophs** and they require a specific substance to grow as against the wild types which are self-sufficient and called **prototrophs**.

Behaviour mutations affect the behaviour pattern of an organism. For example, mating behaviour of a fruit fly may be imparied if it cannot beat its wings. Behaviour mutations have greatly helped the study of behavioural pattern in various organisms.

Regulatory mutations are the mutations which affect the regulation of gene activities. Such mutations may permanently activate or inactivate the gene by affecting the regulatory gene. You have already read about the regulation of genes earlier in this course.

Lethal mutation means that the organism carrying the mutation cannot survive. **Biochemical mutations** may also fall in this category. For example a mutant bacterium that cannot synthesise a specific amino acid needed for its growth will die if grown on a medium lacking that amino acid.

Detrimental mutations influence the viability of the organisms. **Conditional mutations** are those which are expressed under certain conditions. The conditions may be **restrictive** when the expression is inhibited and **permissive** when they allow the normal expression of a phenotype. The best examples are the **temperature-sensitive mutations** found in various organisms. In such organisms a mutant gene product functions normally at certain temperatures but loses its functional capability at other temperatures. For example, a certain class of mutations in *Drosophila* is known as dominant heat-sensitive lethal. Heterozygotes in this class (say H^s/H) are normal at 20°C (permissive condition) but die if the temperature is raised to 30°C (restrictive condition). The study of conditional mutations are extremely important for experimental genetics. You can also classify mutations as **dominant** or **recessive**. They are dominant if they are expressed in a heterozygous condition, i.e., in the presence of wild type allele. If they are not expressed, it means they are recessive in nature. Mutations could also be calssified by the direction in which the mutations occur. Accordingly they are known as **forward mutations** if wild type allele mutates to a mutant allele and **reverse mutations** if a mutant allele mutates back into wild type allele. The process of regaining the original phenotype is called **reversion** and the organism that has reverted is called **revertant.**

A supressor or **second site mutation** occurs at a second site that completely or partially restores a function lost at another site because of earlier or primary mutation. Supressor mutation does not reverse the original mutation, instead compensates for its effects and in fact the organism becomes a double mutant. It can occur within the same gene in which primary mutation occurs. After studying the various categories of mutations let us now discuss detection of mutation in various organisms. But before that try the following SAQ to check that you have understood the concept of mutation, its classes and basis for its classification.

SAQ 2

a) Fill in the blanks with appropriate words from the text.

i) Mutations are ..... changes in the genetic material that are ..... .

ii) Mutations are the primary source of..... , and therefore, are essential for .....

b) Match the items given in column II with those given in column I. Write your answers in the box provided.

Column I		Column II
i) Somatic mutations	☐	a) Sources of spontaneous mutations.
ii) Errors in DNA replication	☐	b) Mutants requiring a nutrient in their medium to grow.
iii) Deformed legs of ancon sheep.	☐	c) The mutations in the meristematic cells of navel orange were vegetatively propogated.
iv) Auxotrophs	☐	d) This is the example of visible mutation.

16.3 DETECTION OF MUTATIONS

Mutations are usually detected by a change in the phenotype. For different systems different methods of detection of mutations have been designed. The detection methods are also based on the fact that most mutations are recessive in nature and their existence can be confirmed depending upon whether they are studied in diploid or haploid organisms or whether they are located on autosomes or sex chromosomes. The detections become

more significant as more and more methods of induction of mutations are becoming available resulting into the accumulation of a large variety of mutants. We shall now discuss some of the methods which are used in humans, fruit-fly, *Drosophila melanogaster* and haploid organisms.

16.3.1 Detection in Humans

In humans beings, the detection of mutation depends upon pedigree analysis and birth statistics. It is because designed matings are not possible or desirable in human beings.

For the **pedigree analysis** one has to prepare a family tree. Once any trait has been shown to be inherited, it is possible to predict whether the mutant allele is behaving as a dominant or a recessive and whether it is sex-linked or autosomal.

The detection of **recessive mutations** by this method is sometimes difficult because one cannot determine with certainity whether a recessive trait has arisen because of the mating of the two heterozygous parents or through the mutation in one of them.

In case of **recessive autosomal alleles**, mutation is hidden in heterozygous condition. An affected individual and a homozygous normal individual will produce unaffected carrier children. Mating between two carriers will produce on an average, one-fourth affected offspring. Because of the presence of a single x-chromosome in human males, **recessive sex-linked mutations** can be easily detected by pedigree analysis. The woman, who is **carrier** of the mutant trait, is not affected by the mutation because of the presence of dominant allele. However, she passes it to half of her male offspring which will show up the affected trait. The most famous case of sex-linked mutation in humans is that of hemophilia. The classical case of hemophilia in the royal family descendants of Queen Victoria can be cited here. The pedigree analysis of seven generations has shown how the gene for this lethal genetic disease for which Queen Victoria was the carrier has been transmitted to the offspring.

The **dominant mutants** are however easy to detect as they exert their effect immediately. If they are present on x-chromosomes, affected fathers pass the phenotypic trait to all their daughters. If the mutation is located on one of the x-chromosomes of female it will be passed on to 50% of her male offspring. If the condition is homozygous in female all her male offsprings will be affected. If dominant mutations are autosomol, approximately 50 per cent of the offspring of an affected heterozygous individual are expected to show the trait.

Another method to detect mutations in humans and other organisms is to screen various proteins and enzymes for minor biochemical changes. The technique used for this purpose is known as electrophoresis and is based on the differential migration of variant proteins in an eletric field. It is very useful as many of the variants or mutants may not produce obvious visible morphological behavioural effects but indicate the altered amino acid sequence of a portein, caused by mutation. For example, the mutant sickle cell haemoglobin has an altered electrophoretic mobility as compared to normal haemoglobin (Fig.16.5).

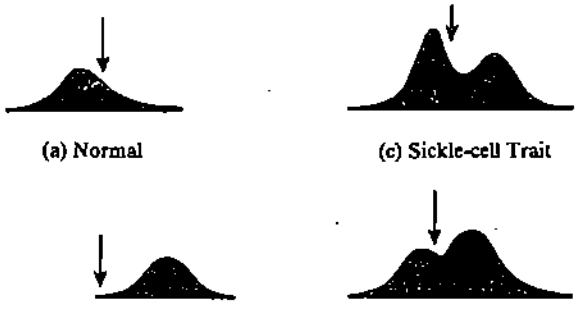


Fig.16.5: Differences in the electrophoretic mobility of normal and sickle cell haemoglobin.

16.3.2 Detection in Drosophila

Muller developed a number of systems for the detection in *Drosophila melanogaster*. We will in this subsection discuss two techniques to detect recessive sex-linked mutations devised by Muller; **the CIB** and the **attached-x-procedures.**

CIB test system detects the rate of induction of sex-linked recessive lethal mutations. The stock of CIB involves C an inversion that suppresses the crossing over, l a recessive lethal mutation and B a dominant gene duplicate causing bar shape of the eye. Genes for all these traits are present on x-chromosome.

In this technique, wild-type (P) males are treated with a mutagenic agent (X-ray in this case) and are crossed with nontreated heterozygous female having one x-chromosome with ClB markers, i.e. genes. As you can see in Fig. 16.6, the F_1 generation consists of four types of offspring. One type of offspring i.e. ClB male dies becasue it contains recessive lethal l gene which is expressed in hemizygous conditions showing its fatal effect. The remaining type of male offspring is wild type. This wild type male is crossed with bar eyed female of F_1 generation. These bar eyed females receive ClB chromosome from their mother and irradiated x-chromosome from their fater. F_2 generation, as a result of crossing over, will consist of bar eyed and wild-type females and two types of males. One of which will be ClB type, and will die because of the reason we have discussed above. In the other class of male offspring we might find a mutant, if induced at P_1 level due to x-ray irradiation. A lethal mutation means no males will survive in F_2. However, if morphological mutation is induced it will show up in this class of males. So we can detect the sex-linked morphological or lethal mutation by this method.

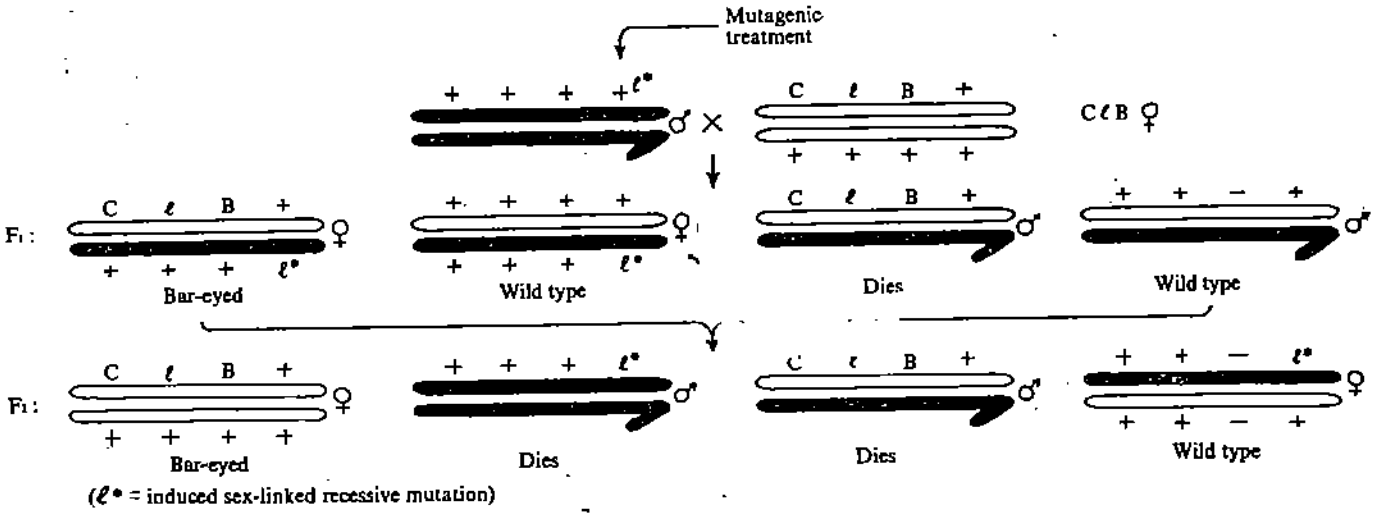


Fig. 16.6: Muller's ClB technique for detection of induced sex-linked mutation in *Drosophila*. In this case mutation is recessive and lethal.

Attached x-chromosome technique is comparatively a simpler technique as mutation can be detected in first generation itself. You must have learnt earlier that many times the two x-chromosomes fail to separate at anaphase and go to the same pole. This condition is known as attached -x and if this type of gamete is fertilised by Y sperm, a female fly of XXY composition is produced. Now look at the Fig. 16.7, you will see that when such XXY or attached -X female is crossed with a wild-type or a normal XY male P_1, which has been exposed to a mutagen, four types of progenies are produced in F_1 generation. These are as follows:

- Triplo-X (XXX) females which die
- Attached -.XY (XXY) females which are viable.
- YY males that also die
- Normal XY males that are viable.

The induced sex-linked morphological mutations of P male will be expressed in both viable progenies. Lethal mutation will once again wipe out the males.

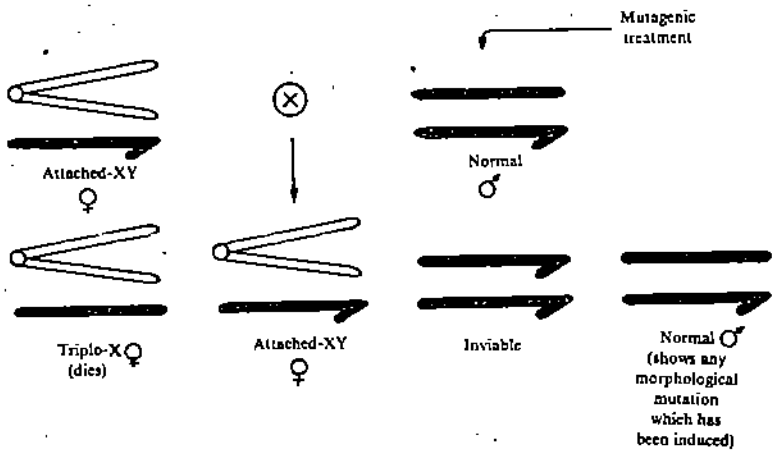


Fig. 16.7 : The attached -X method for detection of induced mutations in *Drosophila*.

16.3.3 Detection in Haploid Organisms

The detection methods in haploid organisms are much more simpler and straight forward. This is based on the fact that they contain only one set of chromosome (n) and thus also one set of alleles. The mutations, if generated, whether recessive or dominant will express themselves in the very next generation. Detection depends on selection system where mutant cells are isolated easily from non-mutant cells. For example, the microorganisms that are nutritional wild type i.e., prototrophs will grow on **minimal culture medium** which consists of glucose, a few inorganic acids and salts, a nitrogen source such as ammonium nitrate and vitamin biotin. But in comparison the induced nutritional mutants of the same microorganism, i.e. auxotrophs will not grow on minimal medium. These will grow only on **complete culture medium** which apart from contents of minimal medium also contains aminoacids, vitamins and nucleic acid derivatives. However, these mutants also grow on minimal medium supplemented with the compound, which as a result of mutations, they are not able to synthesise. So when all of their nutritional requirements are complete, in complete medium or suppplemented minimal medium, the mutants grow well. For instance a *bio*$^-$ mutant (an auxotroph which cannot synthesise biotin) will grow in a minimal medium supplemented with biotin. You can now understand that how the nutritional mutants can be detected and isolated by their failure to grow on minimal medium and their ability to grow on complete medium. You may read more about such techniques in detail in the books enlisted in further reading.

With this we end our discussion on detection of mutations in various organisms. Before we proceed further try the following SAQ to check your progress.

SAQ 3

Tick mark (✓) the correct option in the following statements.

i) Method for detection of mutation depends/does not depend upon whether it is located on autosome or sex chromosome.

ii) Pedigree analysis is helpful in analysing mutations in humans/microorganisms.

iii) It is necessary/not necessary to irradiate the P_1 male with mutagenic agent to detect mutations in Drosophila.

iv) Mutations are detected in F_1/F_2 generation in ClB technique.

v) Triplo-X females of F_1 generation in attached -X technique are viable/non viable.

vi) Mutation detection in microoganisms are simpler beacuse they contain haploid/ diploid set of chromosomes.

16.4 MOLECULAR BASIS OF MUTATIONS

As we all know, genes are made up of stretches of DNA having a specific base sequences. Mutations involve alterations of the DNA sequence including substitutions or addition or deletion of one or more bases. All such mutations can be collectively called **point mutations.** The technical advances in the last two decades have made possible the determination of base sequence of the large segments of DNA so that any change in the sequence can be detected. In this section we will deal with some of exciting aspects of molecular basis of mutations which are so central to any genetic analysis.

16.4.1 Base Pair Substitution

It is a change in a gene in which one base pair is replaced by another base pair e.g., an AT may be replaced by a GC pair. This change can be of two types:

* **Transition** is a change when a purine replaces a purine (A ⇔ G) or a pyrimidine replaces a pyrimidine (C ⇔ T)
* **Transversion** involves change where a purine base is replaced by a pyrimidine base or vice versa

$$\begin{bmatrix} A \Leftrightarrow T \\ G \Leftrightarrow C \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} A \Leftrightarrow C \\ G \Leftrightarrow T \end{bmatrix}$$

These two types of substitutions may change the composition of a triplet codon, so that it may code for a different amino acid, which in turn may change the property of the protein.

16.4.2 Frame Shift Mutation

This type of mutation arises by the addition or deletion of a base pair in the gene causing changes in the reading frame of DNA. You may recall that DNA's base sequence is read in

the form of a string composed of units of three bases, i.e triplet codon. Any addition or deletion of base pair will shift the entire reading frame of that sequence from that point, therefore, disturbing the amino acid sequence. In Fig. 16.8 we have used a sentence composed of three letter words as analogy for triple codon sequence. You can see how, the replacements, additions and deletions change the sense of the sequence.

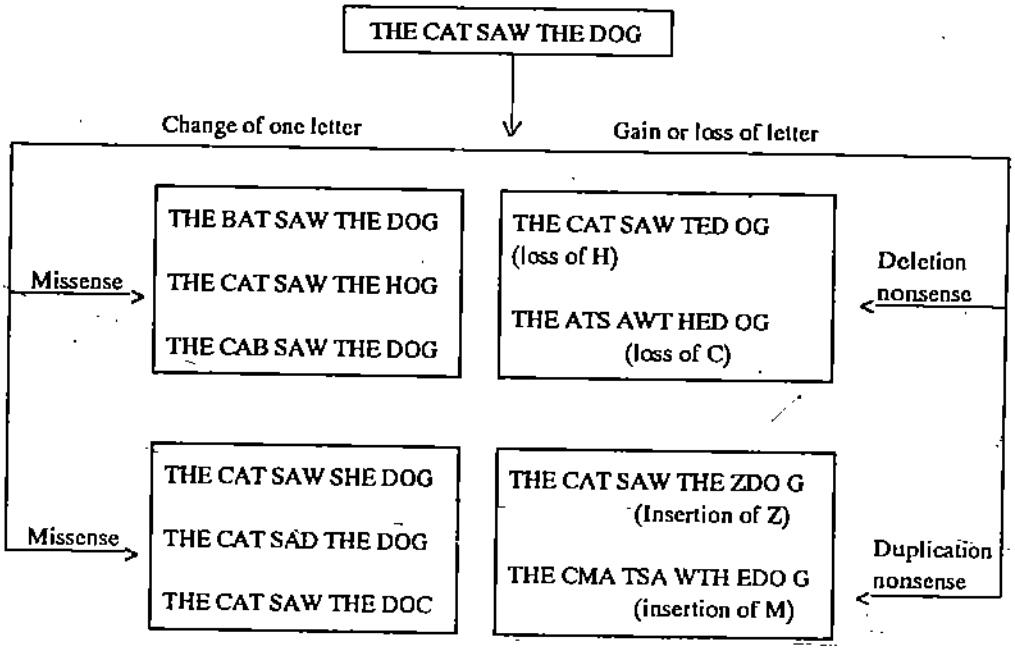


Fig. 16.8: The impact of the replacements, additions and deletions in a sentence composed of three letter words used as the analogy for triple codon sequence.

In the following subsections we are going to discuss various pathways by which single base pair changes take place.

16.4.3 Tautomerization

The purines and pyrimidines in DNA and RNA may exist in several alternate forms called tautomers. Tautomerism, i.e. change in chemical forms occurs through rearrangement of electrons and protons in the molecule. As a result some single bonds become double bonds and vice versa. Such change in chemical forms of the base is called **tautomeric shift.** An example of a tautomer of a purine and a pyrimidine is shown in Fig. 16.9. Although the normal bases posses potentially unstable bonds, they remain chemically stable in one tautomeric form most of the times. This stability is a significant genetic attribute.

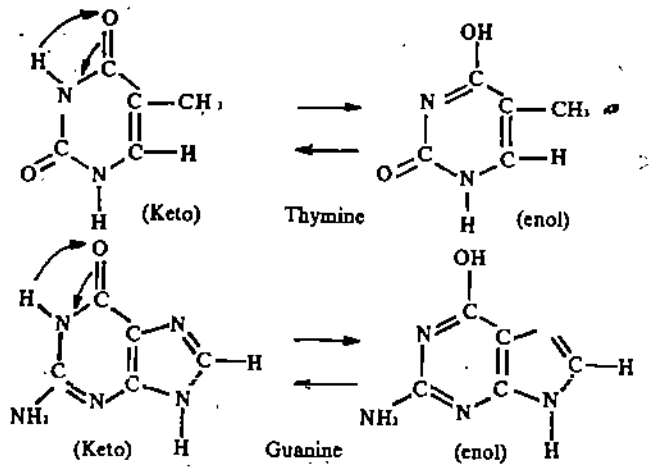


Fig. 16.9: Tautomers of thymine and guanine

16.4.4 Deamination and Depurination

Deamination and depurination of particular bases are the two common chemical events that produce spontaneous chemical mutations. Let us now discuss these processes. Removal of an amino group from a base is called **deamination**. For example, deamination of cytosine produces uracil (see Fig. 16.10). In case uracil is not repaired back, it will direct the incorporation of adenine in the new DNA strand during replication. This ultimately results in conversion of CG base pair to a TA base pair i.e., a transition mutation. Deamination is also caused by certain chemical mutagens about which you will read later in this unit.

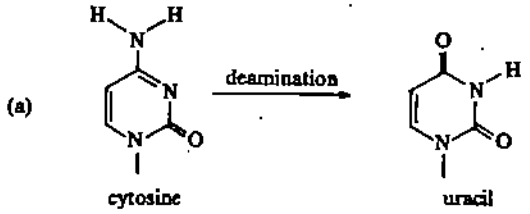


Fig. 16.10: Deamination of cytosine into uracil.

Depurination, as you can guess from the name, involves removal of a purine, either guanine or adenine, from the DNA. This removal occurs due to the breakage of bond between purine and the deoxyribose (see Fig. 16.11). If this fault is not repaired, there will be no base to specify a complementary base during DNA replication. In case a randomly chosen base is inserted a mismatched base pair will be produced resulting into genetic mutation. So in both cases mutation will occur.

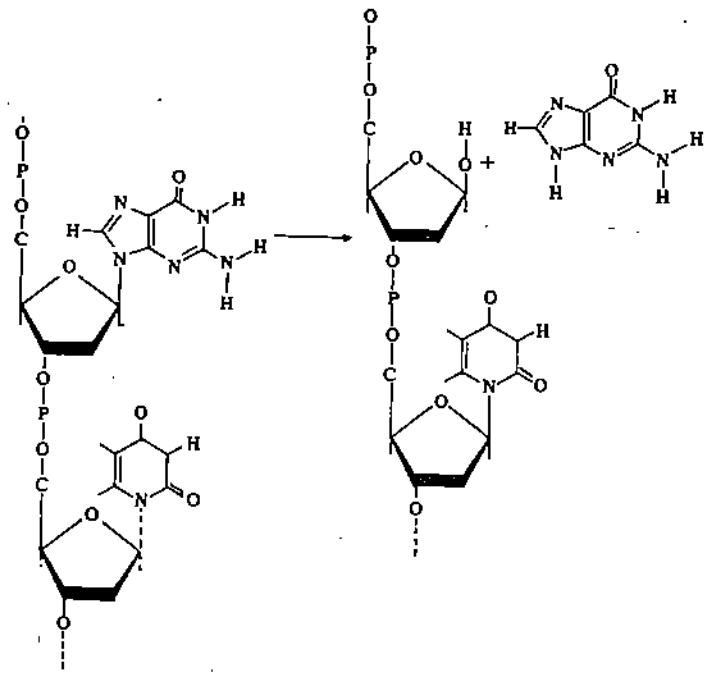


Fig. 16.11: Depurination from a single strand of DNA. You can see that sugar phosphate backbone is not broken. In this case purine adenine is lost.

16.4.5 Base Analogues

Base analogues are the chemicals that have molecular structure that are extremely similar to bases of DNA. These chemicals act as mutagens and during DNA replication get incorporated so as to form base pairs with usual bases. One such chemical is 5-bromouracil (5 BU). 5 BU is a base analogue of thymine and usually pairs with adenine. The bromine atom in 5 BUdR so alters the charge distribution of the molecule that it may tautomerise to a 5 BUdR* form quite frequently. After tautomerisation it possesses the base pairing properties of cytosine, that is, it behaves like cytosine. Fig.16.12 shows how the shift generates G → A transition.

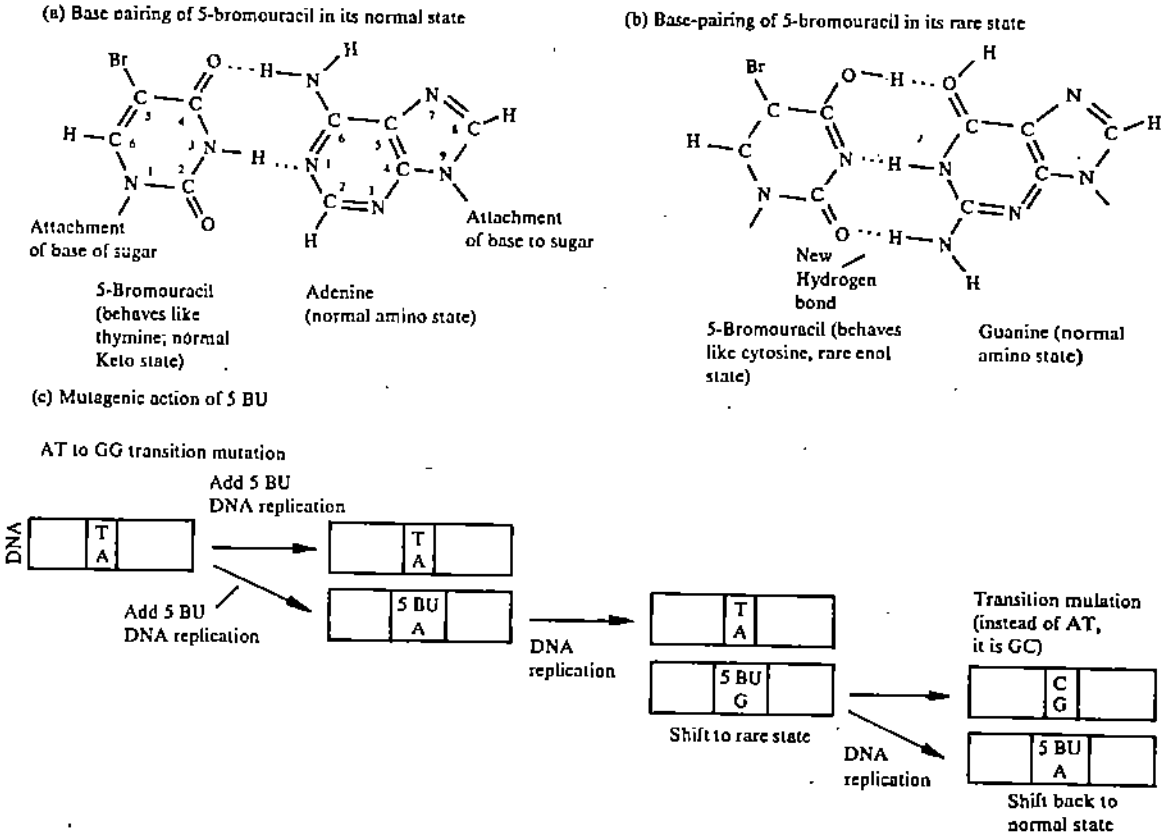


Fig. 16.12: Mutagenic effects of base analogue 5-bromouracil (5BU). (a) In normal state 5BU pairs with adenine; (b) In rare state it pairs with guanine; (c) mutations are induced when 5 BU gets incorporated into DNA in one form and then shifts to its alternate form during next replication stage.

The changes we have discussesd in the preceeding subsections can bring out mutational effect in various ways. When a base pair change causes the change in mRNA codon resulting into a modified protein in place of one specified by wild-type, a **missense mutation** occurs. **Nonsense mutation** occurs when base pair changes result in the change in mRNA that generate a chain terminating codon which will lead to premature termination of protein synthesis. **Neutral Mutation** may go unnoticed as it is a base pair change that changes a codon in mRNA such that the resulting amino acid substitution does not alter the function of the protein. For example, change from codon AGG to AAG which substitutes aminoacid lysine for arginine. Both amino acids are similar in properties, so function of protein is not altered significantly. When base pair change alters a codon in mRNA which may still code for same amino acid, **silent mutation** occurs. For example, change from mRNA codon AGG to AGA both of which specify arginine.

It is time now that you should do another SAQ to see whether you have understood the molecular basis of mutations.

SAQ 4

a) Tick mark (✓) the correct statements and (✗) the incorrect statements in the space provided:

i) When base A is replaced by base T, transition mutation occurs ( ).

ii) The rearrangements of electrons and protons in the bases is called frameshift mutation ( ).

iii) When an amino group is removed from a base, the process is deamination ( ).

iv) Base analogs are chemicals that cause mutations ( ).

16.5 TRANSPOSABLE GENETIC ELEMENTS

Studies have shown that mutations are also caused by addition of long stretches of DNA to the genome. These genetic elements have the capacity to move from one location to another and insert themselves at one or more sites in the genome. This leads to the breaking down of the reading frame and thus produces the mutational effect. The mobile

genetic elements are known by many names such as controlling elements, jumping genes, insertion sequences and transposons. However, these are generally called as **transposable genetic elements** (TGE), a widely accepted name. As you can see in Fig. 16.13 the copies of transposable elements can also be inserted at other points in the same chromosome causing disruption of the gene into which these are inserted. This often produces multiple physiological effects.

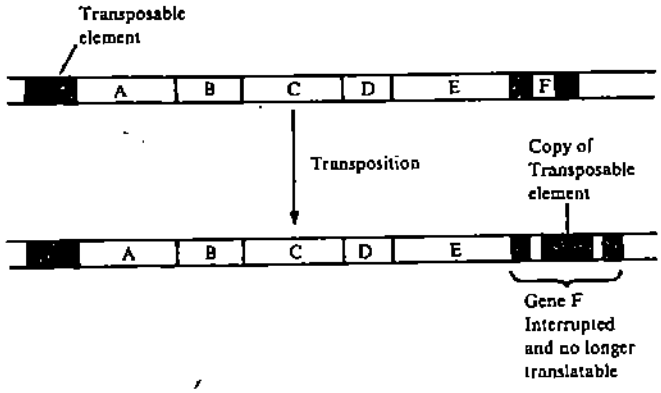


Fig. 16.13: When a transposable element is inserted in the middle of a gene, as in case of gene F in this figure the gene is disrupted. Such genes cannot be transcribed appropriately and thus do not function.

TGEs were first discovered by Barbara McClintock in 1940s. From her study of pigment patterns on kernels of the maize (*Zea mays*), McClintock inferred the existence of mobile genetic elements that influenced gene expression in the corn. She described these elements as "controlling elements". McClintock was awarded Nobel Prize in 1983 for this discovery.

Some examples of eucaryotic transposons are **Ty elements** in Yeast, **copia elements** in *Drosophila* and **retroviruses** in birds and mammals.

One of the more interesting discoveries to emerge regarding transposable elements is that in the process of moving from site to site, they influence the expression and organization of neighbouring genes. Since TGEs can migrate with considerable frequency, the process has the potential for causing rather rapid and dramatic genetic changes. Such changes are now belived to be a major contributing factors to an organism's ability to undergo evolutionary change.

16.6 MUTAGENESIS

In the above sections you have been familiarised with the concept of mutation, a concept which is so central to any genetic or evolutionary studies. The natural occurrence of mutation is very slow. In order to make significant progress in genetics studies, in any living system it becomes imperative that new mutations must be isolated and accumulated at a faster rate. This necessity has led to an active field of study on experimental mutations called **mutagenesis**. Eversince the time of H.J. Muller (1927) who reported that x-ray could induce mutations in *Drosophila* followed by L.J. Stadler's finding that same is true in the plant barley, a large number of agents have become available, to induce mutations. This infact has been responsible for the rapid strides that the field of genetics has taken in this century. The agents that can cause mutations are known as mutagens. All mutagenic agents induce mutations in any of the following pathways:

- They may replace a base in the DNA.
- They may alter the base in such a way that it specifically mispairs with another base.
- They may damage the base so much that it can no longer pair with any base.
- They may intercalate themselves in the DNA paving way for addition or deletion of bases.

With the initial discovery of X-rays as mutagenic agents a large number of other radiations and chemicals have been identified to have the capacity to induce mutations. In fact this list is growing as more and more agents are being studied. We will discuss various types of mutagens in the following subsections.

16.6.1 Physical Mutagens

These consist of high energy radiations which could penetrate living cells and affect the genetic material. The effect of radiations on living cells and tissues is directly

proportional to the degree of penetration of the radiation. Radiations are of two types: electromagnetic radiations and particulate radiations.

X-rays,γ-rays and UV rays are some short wave length **electromagnetic radiations** which have more penetrations in cells and tissues. As a rule the penetrating power of electromagnetic radiation is inversely proportional to their wave length.

Particulate radiations are in the form of subatomic particles emitted from the atoms with high energy. Alpha particles, beta particles and neutrons fall in their category. Alpha particles and beta particles are charged particles. However beta particles being smaller in size are more penetrating than alpha particles. Neutrons ejected from radioactive isotopes do not carry any charge and hence are not deflected when they travel through living matter. Thus they are extremely penetrant and can cause severe damage to the living tissues as well as genetic material. Cosmic radiations that pour down on us from the outer space have the properties of both particulate and electromagnetic radiations.

These physical mutagens are also divided as high energy ionizing radiations which include cosmic rays, X-rays, γ-rays and particulate radiations and low energy non-ionizing radiations which include ultraviolet light. Let us now see how these radiations tamper with the genetic material.

The **high energy radiations** create ionizations in the living cells. While passing through cells and tissues they collide with molecules such as water and cause the expulsion of electrons. This expulsion creates a positively(charged ion. The ejected electron can not remain in free state and therefore, is picked up by another ion creating a negative ion. The generation of free ions by radiations is the basis of extensive damage caused by them at the somatic and gametic level. These ions may combine with oxygen producing highly reactive chemical which may act on genes, chromosomes and other parts of the cells. Peroxides which are mutagenic may be formed in the presence of oxygen following the splitting of water. Experiments have shown that sensitivity to a radiation and the rate of mutation are much lower in the organisms maintained in oxygen-free environment.

Non-ionizing radiations such as UV have more precise mode of action. One of the major effects of UV is the formation of dimers whereby adjacent pyrimidine bases become linked to one another by carbon to carbon bonds. For example, thymine-thymine, cytosine-thymine and cytosine-cytosine dimers of which the first is most common type. Dimerization results in intrastrand or interstrand cross linking which distorts the DNA conformation, thereby affecting the normal replication. Several studies have shown that the affected cells employ a specific repair process to counteract the effects of UV. We will discuss about the repair mechanism at a later stage in this unit. UV can also cause hydroxylation of cytosine that results into weakening of bonds with guanine permitting localised strand separation thus affecting the replication.

The potential hazards that these radiations may cause to all living beings have led to several investigations working out a **dose-effect relationships**. Such studies were necessitated by the fact that many of these radiations have found use in medicine, agriculture and warfare. It is therefore important to know the effects on living matter when a range of doses of radiations are used.

The dosage of ionizing radiations is measured by the **roentgen unit (r)**. It is defined by physicists as the amount of radiation that yields 2.08×10^9 ion pairs per cubic centimeter of air under standard conditions of temperature and pressure. In biological terms this amount of radiation produces two ionizatioins per cubic micron of tissue or water. Most of the research studies with ionizing radiations have been conducted with X-rays and have led to following conclusions.

Irrespective of the wave length used ($0.1A^0$ to $10.0\ A^0$) the number of lethal mutations induced by X-rays is directly proportional to the dose in r units. Thus as you see in Fig. 16.14 when lethal mutation rate is plotted against the dose a linear relationship is observed. As this relationship holds true for very low doses, it is suggested that there is no dose which may be absolutely ineffective. Each doubling of dose results into the doubling of mutations induced. The dosage of an ionizing radiation is based on the amount of ion pairs produced or the amount of energy deposited in the tissue.

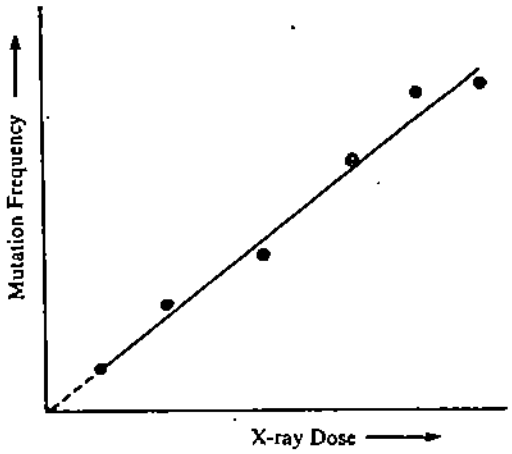


Fig. 16.14: Linear relationship between X-ray dose and sex-linked mutation in Drosophila.

Another interesting finding is that in general the effect of ionizing radiations is cumulative. This means that a given dosage administered in a single acute dose or cumulatively in the form of smaller doses given over an extended period of time will produce same mutagenic effect. However, an exception to this generalisation occurs in mice and probably all mammals. It has also been found that chromosomes are more prone to the radiation damage when they are highly condensed during mitosis. This observation formed the basis of radiations used in treating human malignancy where cells undergoing fast and uncontrolled division will provide more targets for radiations.

Before we proceed further on chemical mutagens try the following SAQ to check your progress.

SAQ 5

a) What are transposons? How are they responsible for mutations?

.....

.....

.....

.....

.....

.....

b) Write answers to the following statements in the space provided along with.

i) What are the radiations that have the properties of both electromagnetic and particulate radiations? .....

ii) Which subatomic particles emitted with high energy have no charge? .....

iii) Which radiations do cause mutations by creating positive and negative ions? ....

iv) Which radiations do cause the distortion of DNA helix due to dimerization? .....

16.6.2 Chemical Mutagens

As we have discussed earlier a large number of chemicals are known to cause mutation employing different pathways. These chemical mutagens are classified into four major groups on the basis of their specific reaction with DNA. We will now discuss these groups.

Base analogues are the chemical compounds structurally very similar to the normal nitrogenous bases of DNA. The analogues can get incorporated into replicating DNA in place of normal base leading to base pair substitution mutations. Like the normal bases they also exist in two alternative forms i.e., keto or enol form or amino and imino form, and change spontaneously from one to another form (tautomeric shift). As previously discussed, 5-bromo uracil is one such mutagen.

Alkylating agents such as ethyl methane sulphonate (EMS), ethyl ethane sulphonate (EES), nitrogen and sulphur mustards and diethyl sulphate (DES) act on DNA by adding alkyl group (ethyl or methyl) to all four bases. However these agents show a strong

The nitrogen atoms attached to the purines and pyrimidines are usually in the *amino* (NH_2) form and only rarely assume the *imino* (NH) configuration. Likewise the oxygen atoms attached to the C 6 atoms of guanine and thymine normally have the keto (C=O) form and only rarely take up the enol (COOH) configuration.

Intercalation is a process in which mutagens such as acridines which are flat, aromatic molecules interact with DNA in such a way that they become wedged between the stacked bases of double helix.

preference for base guanine. This results either in mispairing of affected base or its loss entirely, creating a gap and thus causing mutations.

Acridine dyes are the chemicals that intercalate between the bases of DNA. They include proflavin, acridine orange compounds which can mimic base pairs and are able to slip themselves in between the nitrogenous bases. This results in deletion or addition of base pairs during replication.

Direct acting chemicals like nitrous acid reacts with the nitrogenous base and deaminates them by removing amino group from adenine, cytosine and guanine by oxidative deamination. This results in mispairing and base pair substitution mutations.

16.6.3 Environmental Mutagens

In addition to the above mentioned chemical and physical mutagens, there are a number of chemicals present in the environment that are potentially mutagenic. A wide variety of mutagens occur naturally. Some major sources of natural mutagenic agents include parasitic fungi of field crops, mushrooms, certain vegetables and medicinal herbs. Other environmental mutagens include air and water pollutants, food additives and preservatives, agricultural chemicals, cosmetics, drugs, pesticides, cigarettes and industrial products such as benzidine, vinyl chloride, asbestos etc. In addition many potentially mutagenic compounds may be **carcinogens** or capable of inducing cancer in humans. We will discuss carcinogenesis later in this block.

16.6.4 DNA Repair Mechanism

Both prokaryotic and eukaryotic cells have repair systems to deal with DNA damage. All the systems consist of enzymes to repair DNA. Damages are corrected directly or by excision of base pairs. Let us now see how damages are repaired by these two processes.

Direct Correction of DNA damage

We shall discuss here two systems of direct corrections of mutational lesions. One system involves repairs of UV induced pyramidine dimers. The first relevant discovery on UV repair mechanism was made in 1949 when Kelner observed that UV damage to DNA of *E.Coli* could be reversed if UV irradiated cells are exposed to visible light in the blue range (blue light has wave length range of 320-370 nm). This process known as **photoreactivation or light repair** was found to be temperature sensitive. Photoreactivation means that light induces an enzymatically controlled chemical reaction. The enzym was subsequently isolated from *E.Coli* cells and was named as **photoreactivation enzyme or photolyase.** You can see in Fig. 16.15 that the enzymes cleave the bond between thymine dimer thereby, restoring the structure. The enzyme is also known to bind to the dimers in dark but is activated only when it absorbs a photon of light. Photolyases are apparently very effective since few thymine dimers and mutations are left after photoreactivation.

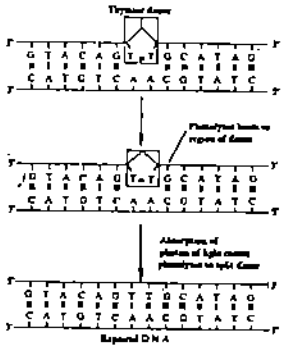


Fig. 16.15: Repair of thymine dimer induced by UV radiation by photoreactivation.

Another system of this kind corrects bases that have become alkylated by chemical mutagens such as EMS. An enzyme called **alkyltransferase** catalyzes the removal of the alkyl group from the base. So now you know that in direct rapair system the modified base is not removed from DNA, but corrected there itself. We will now discuss the repair process involving excision of base pairs.

Repair Involving Excision of Base Pairs

As you have read earlier excision of a base pair is second way of correction of DNA damage. The repair process of UV light induced pyrimidine dimers is called **excision repair or dark repair** as it does not depend on visible light as an energy source.

This repair mechanism was discovered in 1964, independently R.P. Boyce and P. Howard Flanders and by R.Setlow and W.Carrier. As you can see in Fig. 16.16 this multienzyme repair process involves several steps. First, the UV induced dimers are recognised by an endonuclease enzyme which produces a single-strand break in the DNA helix backbone on either side of the distorted region. The removal or **excision** of this distorted region occurs with the help of exonuclease enzyme. The gap created is filled by **repair synthesis** of DNA. DNA polymerase, specifically DNA polymerase I in *E. Coli*, catalyses the synthesis using the opposite strand as a template. Final closure of the break to repair the gap is carried out by polynucleotide ligase.

Defects in excision repair porcess leads to genetic diseases. In humans, an inherited disease called **xeroderma pigmentosum** is caused by a recessive mutation that blocks this excision repair process. People with this disease are very sensitive to sunlight and particularly the UV wavelength in the sunlight. They demonstrate a very high incidence of skin cancer, including malignant melanomas on the areas of the skin exposed to sunlight. Cell cultures from humans with xeroderma pigmentosum are defective in excision repair and are killed at very low doses of UV light.

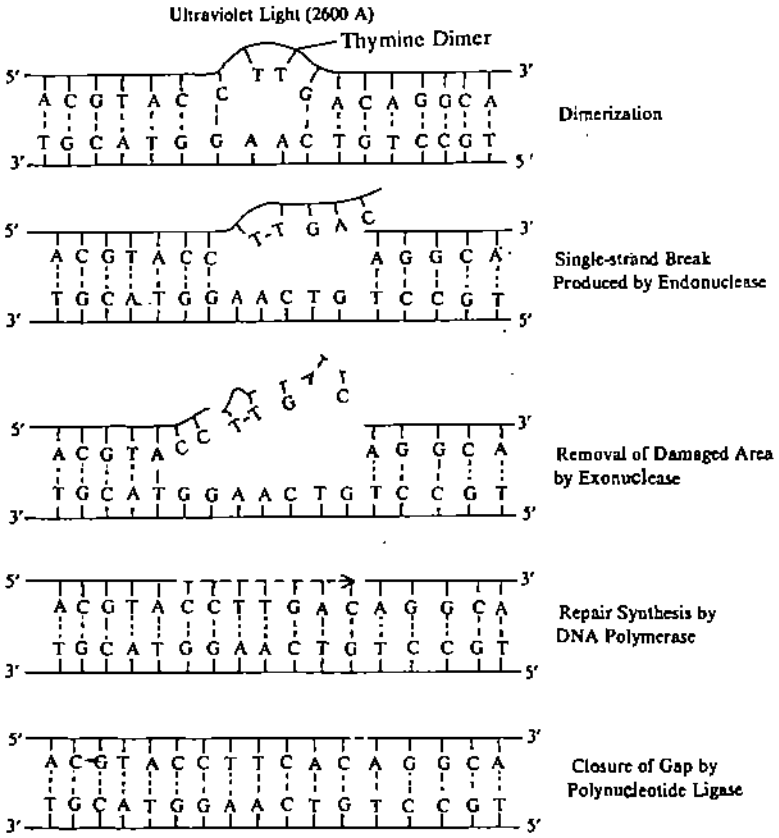


Fig. 16.16: Excision repair of a thymine dimer induced in DNA by ultraviolet radiation.

If UV induced pyrimidine dimer survives the photoreactivation and excision repair systems, another repair system acts during replication. This process is called **post replication repair**. This repair process was discovered in 1968 in an excision defective strain of *E.Coli*. During post-replication, the replicating enzyme slips over the dimers creating gaps on the newly synthesized strand opposite the site of a dimer present in the template strand. These gaps after the subsequent rounds of replication gradually disapper due to the activities of **rec loci** or genes : **rec A, rec B, rec C** and others but in particular rec A gene. You have already read about rec genes earlier in the course.

Another excision repair process to correct chemically induced DNA danage is **repair by glycosylases.** Enzyme glycosylases can detect and excise the altered base from deoxyribose sugar to which it is attached. Different glycosylases have been discovered that can remove deaminated bases, alkylated bases, ring open purines and also UV photodimers. They cleave base-sugar bond (N-glycosidic) and liberate the altered base consequently creating a hole called **AP site.** AP site can be apurinic where there is no A or G or apyrimidinic where there is no C or T.

Such AP sites, produced either by glycosylase action or by spontaneous loss of a purine or pyrimidine base are repaired by AP endonuclease. AP endonuclease breaks the chain by cleaving the phosphodiester bond at AP sites. This initiates excision repair process in

which exonuclease removes few nucleotides ahead of missing base, DNA polymerase I fills in the gap and DNA ligase seals the nucleotides. You can see various steps involved in this repair system in Fig. 16.17.

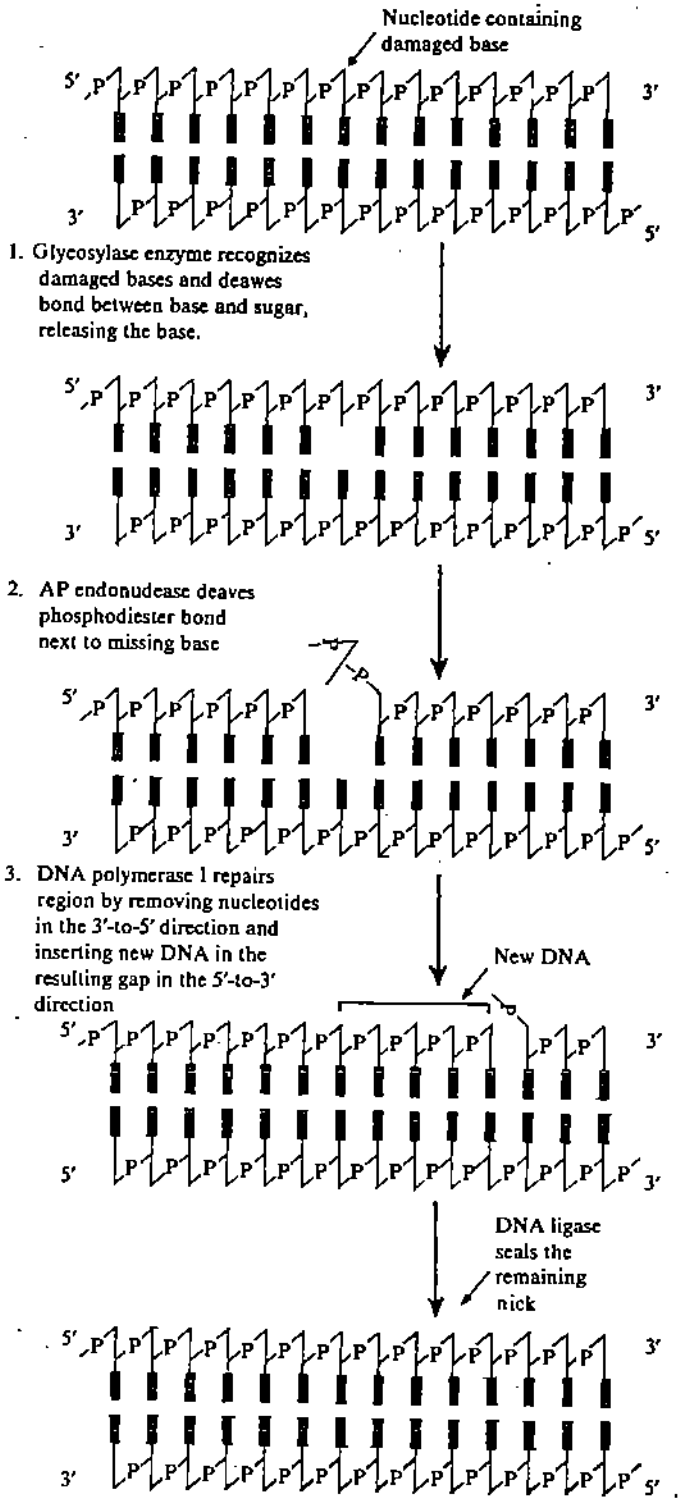


Fig. 16.17: Repair of damaged DNA bases involving the action of glycosylases.

So far you have studied how the mutagenic effects of physical and chemical mutagens can be successfully combated by repair mechanism. However, the errors introduced during the repair cause the mutation. Also not all mutational lesions are repaired and hence mutations are produced. We will now proceed to the last section of this unit which deals with the applications of mutations. But before that try the following SAQ to check your progress.

SAQ 6

The mutagen causes deaminations of adenine into hypoxanthine in the DNA. Name the repair mechanism to rectify this alteration and briefly write the steps involved in this process in the space given below.

.....

.....

.....

.....

.....

.....

.....

.....

.....

16.7 USE OF MUTATIONS

In the previous sections you have read that mutations bring about the changes in the genetic material of the organisms which may adversely affect their normal functioning. Such changes sometimes lead to serious consequences. But mutations are known to produce certain benefits as well. Let us now look into the beneficial aspects of mutations.

Genetic recombination helps in the construction of an array that indicates the position of genes on a chromosome with respect to one another. When this is done by genetic technique, the array is called a genetic map.

Mutations are raw material for evolution as they bring about variations that are inheritable. It is rather difficult to conceive any evolutionary process in the absence of mutations Secondly., they are the working materials for the geneticists. Some of our significant knowledge in microbial genetics has come from the study of mutant organisms. By noting the specific aberration in a particular mutant organism we can find about various aspects such as the functions, metabolic pathway and its regulations, site of action and interrelationship of various normal cells and systems, and genetic mapping. For example, if a mutant *E.coli* unable to synthesise lactose is found to lack a specific gene group, we can say that the particular gene group is responsible for lactose synthesis. Without mutants, we would still be in the dark as to how gene expression is regulated, and we would still be unaware of many essential features of DNA replication.

Thirdly, mutations being the principal source of variations, are very important for agriculatural purposes. Plant breeders throughout the world rely on the genetic variability available in crop plants for their improvement.

Induced mutagenesis has been practised by plant breeder since the time of Muller and Stadler. More and more mutagens have been tried as and when they are discovered. Many mutants have been able to attain the status of a variety. The application of induced mutations can be specially rewarding under the following conditions:

- When the naturally occurring variablility is low.
- Where sexual reproduction is absent and, therefore, variability does not arise by recombinations.
- Where change in a specific trait is required in an otherwise superior cultivar.
- When a phenotypic change brings a distinct economic value.
- When generation time is very long as in plantation crops and fruit trees.
- Where a plant product is to be improved by manipulating its biosynthetic pathway.
- Where a desirable gene is tightly linked to an undesirable gene.
- Where plants can be easily propagated by vegetative means as many mutations may lead to seed sterility.

Infact, one could find many other situations where breeding mutants becomes very beneficial. Highly useful mutations have been obtained. For characters like, improved variety, increased growth, reproduction and high yield, quality of the plant product, disease and pest resistance, and stress tolerance to various agronomic factors like pH, temperature, salinity and deficiency of water and nutrients. Some of the examples where mutations have helped are crops like wheat, maize barley, mustard, sorghum, cotton and pea for increased yield and improved variety. Synchrony of flowering has been obtained through mutation in castor beans. Seedless mutation has been obtained in grapes and banana. Here you may recall the examples of navel orange and ancon sheep which after spontaneous mutations have been of advantage to human beings.

The above discussion should make it very clear that induced mutations can be exploited for advancement of genetics as well as to yield useful mutations in any type of crop plants. The applicability of this technique is dependent on the type of the organism and the type of character in question.

With this we end our discussion on mutations and mutagenesis. In case you want to know more about any aspect of mutations you can read the books enlisted under further reading. In the next unit we will discuss about carcinogenesis and teratogenesis. Before we summarise what we have learnt in the unit, do the following SAQ to check upon your progress.

SAQ7

Fill in the blanks in the following sentences with appropriate words:

(i) Mutations enable one to learn about ..... regulation.

(ii) Mutations can indicate..... between apparently unrelated systems.

iii) Induced mutations can be applied when..... is long.

(iv) The technique to induce mutation depends upon the ..... of plants.

16.8 SUMMARY

In this unit you have studied that:

- A mutation is a sudden and heritable change in the genetic material. Mutations can be classified into several types depending upon the criteria used and therefore can be somatic, germinal or zygotic; spontaneous or induced by application of mutagen; visible, detrimental, or lethal; forward or reverse, and other categories like nutritional biochemical etc. Although all genes are mutable, they do not change with the same frequency or at the same rate.
- Mutations can be detected by simple observation or by employing a specific technique. Various detection methods have been devised keeping in mind the type of mutation and the organism under study.
- At DNA level mutations may be generated by replacing the original base by a new base, by shifting the reading frame of the codons, by base analogues or by tautomerization. These phenomena will lead to a protein which will differ in its function as compared to the original one producing the mutational change.
- Transposable genetic elelments are unique DNA segments that can insert themselves at one or more sites in the genome in both prokaryotes and eukaryotes.
- Mutations can be generated at an accelerated rate by employing various agents called mutagens. Mutagens can be physical e.g. , X-rays, γ-rays UV radiations, subatomic particles and cosmic radiations; chemical e.g. alkylating agents, base analogues, acridines and many other chemical compounds or environmental mutagens like pollutants, preservatives, pesticides etc.
- Cells can employ various repair mechanisms to correct the radiation damage caused by chemicals such as light repair, excision repair, postreplication repair and repair by glycosylases.
- Despite having their adverse effects on the normal functioning of the organisms, mutations offer certain benefits to human beings. Apart from bringing out variability in organisms they are helpful in advancement of studies in genetics and improvement of various characteristics of several crop plants that has been achieved to some extent through induction of mutations.

16.9 TERMINAL QUESTIONS

1) Mutations, the working tools of genetics are studied in a variety of organisms. Suggest any two advantages of using microorganisms for mutation studies.

.....

.....

.....

.....

2) Explain the following terms.

i) Induced mutations

ii)Auxotrophs

iii)Mutagenic agents

iv)Base analogues

v) CIB method detecting mutation.

(3) What are mobile genetic elements ? How can they cause mutations ?

(4) DNA damage caused by mutagens is repaired directly or by excision of damaged DNA. Give one example of each system of repair.

(5) How mutations can be applied to increase the agricultural yield ?

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

16.10 ANSWERS

Self Assessment Questions

1) Mutations are sudden hertitable changes in the genetic material and involve a qualitative or quantitative change in them. The change results in a corresponding change in the phenotype. Some of the examples are the sickle cell anemia, haemophilia, ability to taste PTC (phenyl thiocarbomide) bitter, navel orange.
2) Student can give many other examples of such types.
   a) i) sudden, hertitable
      ii) genetic variation, evolution
   b) i) c, (ii) a, (iii) d , (iv) b.
3) i) depends, (ii) humans, (iii) necessary, (iv) F_2, (v) non-viable, (vi) haploid
4) a) (i)x, (ii) x , (iii) ✓ , (iv) ✓
5) a) Transposons are genetic elements of long stretches of DNA that have the capacity to move from one location to another. They insert themselves at one or more sites of the genomes. The mutational effect is caused when they cause a break down in the reading frame.
   b) i) cosmic radiation, ii) neutrons, iii) ionizing radiations, iv) UV radiations
6) Deamination of adenine will be repaired by glycosylases. In this process enzyme glycosylase detects and excises the altered base creating an AP- site. The AP- site gap is filled by the removal activity of exonuclease and gap filling activity of DNA polymerase I.
7) i) metabolic, ii) relations iii) generation time, iv) type

Terminal Questions

1) i) They carry only one set of allele (haploid) and therefore mutations can be easily detected.
   ii) They have a short life span and, therefore, results can be easily obtained.
2) i) Induced mutations arise as a reponse to an extremally applied agent. Various forms of radiations and many chemicals can induce mutations.
   ii) Auxotrophs can otherwise be called nutritional mutants. These mutants lose their ability to synthesise specific nutrients and such nutrients have to besupplemented in their medium for their normal growth. For instance strains of *E-coli* may not synthesise galactose or tryptophan and are respectively known as *gal* and *try* mutants.
   iii) Agents which can induce mutations are known as mutagenic agents. They include various kinds of radiations such as eletromagnetic radiations and particulate radiations. Mutagenic chemicals such as base analogues, alkylating agents acridine dyes and compounds such as nitric acid bring bring about changes in the genetic material. Environmental mutagens such as water pollutants food additives and preservatives, cosmetics, drugs and industrial products can induce mutations.
   iv) Base analogues are chemical compounds structurally similar to the normal nitrogenous bases of DNA. They tend to substitute the purines and pyrimidines in

the DNA and cause mutations. 5 Bromouracil, for instance, is one such base analogue which substitutes thymine.

v) **CIB** method is a method devised by Muller for detecting mutations in *Drosophila*. This procedure can detect the rate of induction of sex linked recessive lethal mutations. For more detailed account of CIB method refer to section 16.3.2.

3) These elements are able to move from one position to another. They get inserted in a gene, break its reading frame and thus cause the mutation.

4) The examples are repair of damage caused by UV radiations and any of the excision repair mechanism like dark repair or postreplication reapir. Student can explain these from the text.

5) It has been long since known that many of the present day agricultural varieties arose as a result of mutations from the wild type. Crop scientists use mutations as a tool for obtaining increased agricultural yield. Mutations can result in improved variety, increased growth and reproduction, high yield, better quality of the plant product, disease and pest resistance and stress tolerance to various physical factors. Mutations have helped crops like wheat, maize, barely, sorghum, cotton and pea for increased yield and improved variety.

UNIT 17 CARCINOGENESIS AND TERATOGENESIS

Structure Page No.

17.1 INTRODUCTION

Cell growth is a process that is regulated by the specific needs of the body of the organism. At occasions when the controls that regulate cell multiplication break down, the cell begins to grow and divide even though the body has no need for further cells of its type. When the descendents of such a cell inherit this property to grow and divide without responding to regulation, a clone of cells is formed that can expand to a considerable size. So ultimately a mass of these unwanted cells called tumor is formed. Some tumors are benign, but some have devastating effects on the organism leading to cancer. It is because of their malignancy that much research has been done to understand various aspects of transformation of normal cell into cancerous cell.

In this unit we will discuss various aspects of carcinogenesis which includes cancer causing agents, processes involved in cellular transformation, and prevention and cure of cancer. Later in the unit you will read about teratogenesis, that is, deformities in developing embryos produced by various agents called teratogens. Development is a highly integrated and complex process of phenotypic changes which are guided largely by a precisely programmed sequence of genetic events. This precision of development, makes the normal developmental sequence vulnerable to disturbances caused by various agents, thus leading to physical defects in the new born.

Objectives

After studying this unit you will be able to:

- discuss the basic mechanism by which a normal cell is transformed into cancer cell,
- enlist various types of carcinogens and explain their role in cancer induction,
- correlate between mutations and cancer, and explain that carcinogens are also mutagens,
- realise the importance of cancer prevention and discuss the ways to cure cancer, and
- explain the role of teratogens in causing developmental abnormalities in the newborn infants.

17.2 CANCER

Cell division is one of the key features of living cells and is one of the most basic aspects of multicellular existence. Throughout the embryonic and adult life the differentiated cells

have to take a decision as to divide or not to divide. The correctness of the decision is important for the normal functioning of cells. The cells do undergo division with precision. It is only when they lose this control that they begin to divide mitotically in uncontrolled manner and give rise to tissue masses called **tumors**. Tumors may be either **benign** that is without threat to the life of the organism, or **malignant**, in which case the life of the organism slips into danger. In some cases, the malignant tumors invade surrounding tissues and spread to the other parts of the body. This spreading is called **metastasis** and metastasizing tumors are called cancers. Metastasis occurs in two stages: first, the extension of cancer cells into surrounding tissues and then their entry into either the blood stream or the lymphatic system. In Fig. 17.1 you can see a tumor cell of rat squeezing into lumen of a small vein through intercellular space in the wall of a blood vessel. Metastasis of cancer cells through lymph vessels is slowed by lymph nodes as the cancer cells pause there before proceeding to the next node. The removal of lymph nodes and the ducts between them, as in mastectomy for breast cancer, can often end the disease in a patient. But, metastasis through blood vessels is rapid, non-specific and very commonly fatal.

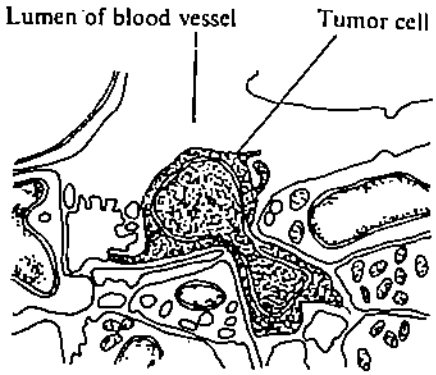


Fig. 17.1: Metastasis of a tumor cell in rat into the lumen of a vein through a cell in the blood vessel wall.

Tumor initiation in an organism, is called **oncogenesis** (**onkos means 'mass' or 'bulk'**; genesis means 'birth'). The **transformation** of a normal cell to a cancer cell may begin with the changes in the nucleus. Although in many types of cancer, the transformed cells may retain a large number of morphological and physiological characteristics of their precursor normal cells, the nuclear changes are followed by changes in other parts of the cell. Besides the potential to grow into tumor cells, the transformed cells exhibit the other characteristics such as rounder shape, disorganised arrangement of cytoplasmic microtubules and micro filaments, and insensitivity to contact inhibition; they grow to higher cell densities and appear nutritionally less fastidious. Apart from various morphological deviations, several biochemical differences have also been observed. For example, the cancer cells that form solid tumors excrete much larger quantities of lactic acid than do their normal counterparts. The normal cells show cell adhesion which is very specific. For example a normal liver cell will adhere to liver cell only and the same specificity is shown by other cell types. However, cancer cells do not express such cellular affinities. Fibronectin, a compound implicated in gluing activity in normal cells is found either in a very low concentration or is completely absent in tumor cells. Many cytoskeletal proteins may appear in altered forms. The membrane is modified in such a way that nature and kind of glycolipids and glycoproteins that are inserted normally into a bilayer are changed. Surface proteins of the cancer cells do not limit the cell growth so the cells continue to multiply forming tumors. You can see in Fig. 17.2 the difference between the division pattern of a normal and a cancer cell.

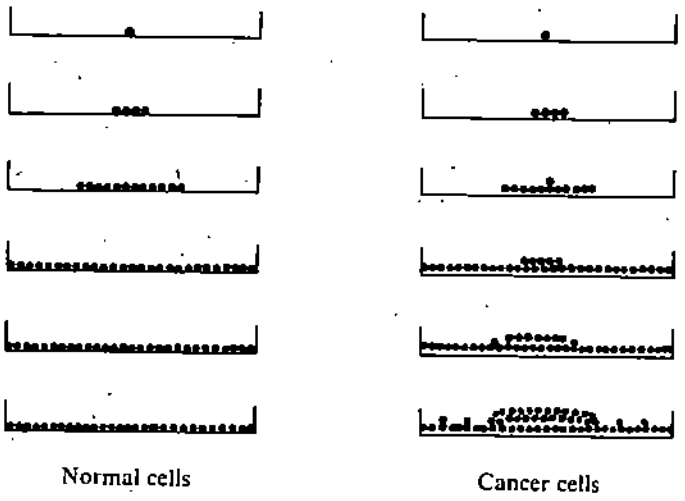


Fig. 17.2 : Diagramatic representation of comparison of the multiplication of a normal cell and of a cancer cell. The normal cells divide until they form a monolayer. Cancer cells, however, can form irregular masses, several layers deep.

Mouse 3T3-cell lines are the progeny of the cells of endothelial tissues of mouse that acquire the ability to multiply indefinitely in the culture. These 3T3-cell lines, although derived from normal mouse embryos, have about 65 ' chromosomes rather than normal diploid number of 40. It seems the possession of extra chromosomes (aneuploidy) frequently helps cells to grow indefinitely in culture from cell lines.

There can be many causes of cellular transformation to the cancerous state, such as spontaneous genetic change which could be gene mutation or chromosome aberrations, exposure to mutagens, activation of one or more normally 'silent' genes or induction by cancer inducing viruses (tumor viruses). The cancer causing mutagens are known as carcinogens. Cancer inducing viruses carry the cancer inducing genes called **oncogenes** as part of their genome. These are also known as **v-ones** (viral oncogenes). Similarly, normally silent cellular genes with the potential to become oncogenes are called **proto-oncogenes** or **c-ones** (cellular oncogenes). Cellular transformation is genetically stable so that a single cell produces a clone of cancer cells. The cancerous transformation of the cells of normal growth properties can occur spontaneously though rarely. Some cell lines like mouse 3T3-line show a significantly high frequency (10^{-5} -- 10^{-6}) of such transformation. High frequency of transformation is also expected when cells are infected with tumor viruses or exposed to carcinogenic chemicals or radiation source. Such an experimental induction of cancer is known as carcinogenesis. With a better understanding of the mechanism of action of carcinogens it is now clear that the majority of agents that can cause cancer act at the level of DNA. You will read about various aspects of cancer induction in the following subsections.

17.2.1 Chemical Carcinogens

Heavy smokers develop lung cancer more frequently than light smokers, and long-term smokers develop lung cancer frequently than shor-term smokers.

The first description of a link between an environmental chemical and humans dates back to 1761, when a London doctor John Hill reported that individuals who used snuff suffer an abnormally high incidence of nasal cancer. Now of course, it has become increasingly apparent that cancer is often caused by exposure to environmental chemicals. In many cases the ability of a particular chemical to cause cancer has only become obvious after industrial workers exposed to the agent on a regular basis develop an abnormally high incidence of a specific kind of cancer. Today, in the industrial society, a large number of organic and inorganic compounds have been known to be potential carcinogens. You can see a list of some such chemicals and the type of cancer induced by them in Table 17.1. Millions of workers are exposed to the hazards of occupation induced cancer. However, exposure to the chemical carcinogens is not always work related or involuntary. Tobacco smoke, which contains more than 30 known carcinogens is knowingly inhaled by a large number of people, despite extensive warnings. Tobacco being one of the number one culprits is smoked in many different ways and is also chewed. Apart from cigarettes and bidis, it is smoked in cigar, pipes, chillums and hookah. It is also chewed in paan and used for brushing teeth. According to the available data, out of an estimated 5 million deaths in adult Indian population, it is likely that approximately 1 million deaths can be traced to tobacco leaf. So you can understand the urgent need to curb the tobacco habit, which in turn will reduce the incidence of oral cancer.

Table 17.1: Some of the chemical carcinogens present in our environment and the type of cancer these can induce.

Agent	Type of Cancer Induced
Asbestos	Lung, esophagus, stomach
α - Naphthylamine	Bladder
4 - Aminodiphenyl	Bladder
Carbon tetrachloride	Liver
Acrylonitrile	Colon, lung
Mustard gas	Lung, larynx.
Wood and leather dust	Nasal sinuses
Arsenic compounds	Lung, skin
Chromium and chromates	Lung, nasal sinuses
Vinyl chloride	Liver, lung, brain
Benzene	Leukemia
Soot and tars	Skin, scrotum, lung, bladder
Organo-chloride posticides	Liver
Polychlorinated biphenyls	Liver
Aniline derivatives	Bladder
Lead	Kidney
Cadmium salts	Prostate, lung
Diethylstilbestrol (DES)	Uterus, vagina
Tobacco smoke components	Lung
Aminostilbenn, arsenic, benz (a) anthracene.	
benz (a) pyrene, benzene,	
benzo (b) fluoranthene.	
benzo (c) phenanthrene,	
benzo (j) fluoranthene, cadmium, chrysene.	
dibenz (c,c)anthracene.	
dibenzo (a,e) fluoranthene.	
dibenz (a.h) acridine, dibenz (a,j) acridine.	
dibenzo (c,e) carbazone.	
N-dibutylnitrosamine. 2.3-dimethylchrysene.	
indeno (1.2.3-c,d) pyrene. 5-methylchrysene.	
5-methylfluoranthene. 5-methylchrysene.	
nickel compounds, N-nitrosodimethylamine.	
N-nitrosomethylethylamine, polonium-210.	
N-nitrosodiethylamine. N-nitrosonornicotine.	
N-nitrosoanabasine. N-nitrosopiperidine	

It has been possible to learn about the mechanism of chemical carcinogenesis by studying the effects of known carcinogens on animals or cell cultures. Peyton Rous and his associates in early 1940s discovered that tumors induced by painting the skin of rabbits with coal tar regress when application of this carcinogen is stopped. However, when the skin was again treated with irritants like chloroform or turpentine, tumors reappeared. These agents by themselves do not cause tumors. Based on such observations, Rous proposed that carcinogenesis occurs in two stages, termed **initiation** and **promotion.** In the initiation phase, normal cells exposed to a carcinogenic agent are irreversibly altered to a state called **preneoplastic state.** These preneoplastic cells, when stimulated by a

promoting agent undergo cancerous or neoplastic type of growth. Subsequent experiments on various other animal tumors and cell culture systems have led to the conclusion that distinct initiation and promotion phases are a general phenomenon in carcinogenesis. Let us now discuss the mechanisms underlying initiation and promotion phases.

Evidence suggests that initiation effect of carcinogen arises from its ability to chemicallly modify DNA, i.e. by producing mutation in DNA. The carcinogenic potency of chemical agents correlate with their ability to bind to DNA. Most carcinogens are mutagenic when tested in bacteria. So we can say that much of carcinogenesis is due to changes in DNA resulting into stable and inheritable change in the properties of cells.

Promotion phase in contrast to initiation is a gradual process that requires prolonged exposure to promoting agents and to a little extent is partially reversible. Promoting agents cause the stimulation in cell divisions thereby, increasing the population of damaged cells, resulting into tumors. Sometimes it so happens that cancer may not develop until many years after the exposure to a carcinogenic agent. This could be because of the long time taken by the promotion phase.

If you have learnt about the above concepts on cancer and cancer causing chemicals, try the following SAQ.

SAQ 1

Write one word term for each of the following statements:

(i) Transformed cells invade the surrounding tissues.

.....

(ii) These genes have the potential to transform the normal cells into cancerous cell.

.....

(iii) Carcinogens act on DNA, in this phase of cancer induction.

.....

(iv) Uncontrolled cell division is stimulated during this phase of carcinogenesis.

.....

17.2.2 Radiation Induced Cancer

You have read in the previous unit that we on earth are normally exposed to various types of radiations. These include ultraviolet rays, subatomic particles, ionizing radiations, cosmic rays and emissions from radioactive elements. Radiations differ in energy and wavelengths and many of them are carcinogenic , i.e. capable of causing cancer. Ultraviolet and ionizing form of radiations have long known to be highly mutagenic agents. Excessive exposure of skin to ultraviolet light causes skin cancer. In the previous unit we have discussed a inherited disease *Xeroderma pigmentosum* caused by recessive mutation. The patients suffering from this disease have a defective DNA repair mechanism and are very sensitive to the UV wavelengths in the sunlight. They demonstrate a very high incidence of skin cancer, including malignant melanomas on areas of the skin exposed to sunlight. Unlike UV radiations which are too weak to pass through the skin into the interior of the body, ionizing radiations are more serious problems, as they are strong enough to penetrate the skin and affect internal organs. In countries where medical x-rays have been used for the treatment of superficial skin conditions of head and neck, like ringworm and acne, the rate of thyroid cancer is much higher. Thus we can say that practical benefits to be gained from every exposure to x-ray should be properly weighed against the increased cancer risk. Now-a-days, in medicine, soft x-rays have been introduced for diagnostic purposes. With their use the risk of cancer has been considerably reduced.

The carcinogenic effects of radiations emitted by radioactive elements resemble those of x-rays. Marie Curie, the co-discoverer of radioactive elements polonium and radium , later died of a form of leukemia apparently caused by her extensive exposure to radioactivity.

Were you ever told of the incidences of atomic bomb explosions over Hiroshima and Nagasaki in 1945 ? There was a massive radioactive fallout produced as a result of these explosions. The radiation thus produced led to dramatic increase in the incidence of leukemia, lymphomas, and cancer of thyroid, breast, uterus and gastrointestinal tract.

Radiation-induced carcinogenesis resembles chemical carcinogenesis in its basic mode of action. Like most chemical carcinogens, radiation initiates malignant transformation by inducing changes in DNA. Likewise a promotion phase is required after exposure to radiation before malignancy is fully expressed. After having known chemical and radiation carcinogens let us now discuss how viruses can cause cancer.

17.2.3 Oncogenic Viruses

Peyton Rous in 1910 showed that when pieces of sarcoma (a particular kind of tumor) of a chicken were transplanted into other chickens from the same stock, they developed sarcomas. He also showed that cell-free filtrates of sarcoma inoculated into chickens also resulted in tumor development. We now know that results of experiments done by Rous are explained by the existence of a tumor causing virus named as Rous Sarcoma Virus (RSV). You can see the structure of RSV in Fig. 17.3 which is RNA virus of retrovirus type. **Retroviruses** are the viruses surrounded by an envelope that infects animal cell. They have single-stranded RNA genome. Two copies of RNA genome are present in each virus particle, that is, retroviruses are diploid. Evidence shows that a number of viruses containing either DNA or RNA as genetic material are oncogenic, that is tumor-inducing in animals and plants. You can see Table 17.2 for some example of such viruses and types of tumor induced by them.

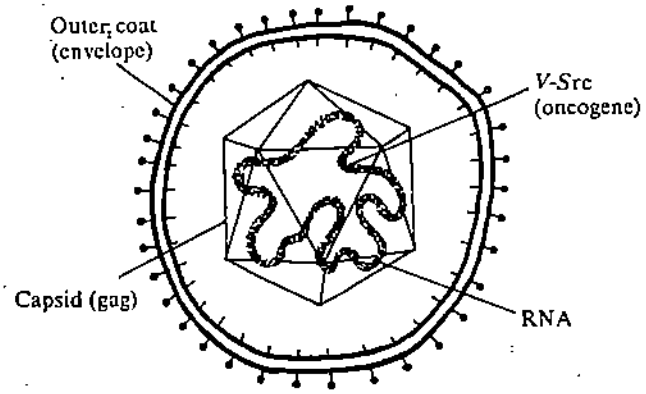


Fig 17.3 : Schematic diagram of an RNA virus, The Rous sarcoma virus.

Table 17.2 : Some examples of tumor inducing viruses and tumors caused by them.

Class	Examples	Tumors Induced
	DNA VIRUSES	
Herpesviruses	Lucke virus	Kidney adenocarcinoma (frog)
	Epstein-Barr virus	Burkitt's lymphoma and nasopharyngeal carcinoma (humans)
	Marek's disease virus	Lymphoma (chickens)
Papovaviruses	Shope papilloma virus	Papillomas (rabbits)
	SV-10	Subcutaneous, kidney and lung sarcomas (hamsters)
	Polyoma	Cancer of liver, kidney, lung, bone, blood vessels, nervous tissues and connective tissues (mice)
	Human papillomaviruses (HPVs)	Cervical cancer (humans)
Adenoviruses	Human adenoviruses (many types)	Subcutaneous, intraperitoneal and intracranial tumors (hamsters)
	RNA VIRUSES	
B-type viruses	Mouse mammary tumor virus (Bittner)	Mammary carcinoma (mice)
C-type viruses	Rous sarcoma virus	Sarcomas (birds, mammals)
	Murine leukemia viruses	Leukemia (mice)
	Feline leukemia virus	Leukemia (cats)
	Murine sarcoma virus	Sarcoma (mice)
	Feline sarcoma virus	Sarcoma (cats)
	Avian leukemia viruses (avian myeloblastosis and others)	Leukemia (chickens)
	Human T-cell lymphotropic viruses (HTLV-I & HTLV-II)	Leukemias/lymphomas (humans)
Plant viruses	Wound tumor virus	Roots and stems

T-Cells also called T-lymphocytes are a type of white blood cells that develop in thymus and are responsible for cell mediated immunity. The other type of lymphocytes are B-cells which differentiate in lymph nodes and vessels and secrete antibodies. You will read in detail about these in the unit on immunogenetics.

In case of humans, **Epstein Barr Virus (EBV),** responsible for Burkitt's lymphoma i.e. tumor of jaw was the first one to be discovered. The other viruses which have been found to be associated with human cancers are, **human T-Cell lymphotropic virus of HTLVs,** a family of related RNA-containing viruses. Two members of this group, **HTLV-1** and **HTLV-II,** cause certain kinds of human leukemias and lymphomas. The third member, **HTLV-III** which is also called **HIV** causes the disease called acquired immune deficiency syndrome **(AIDS).** Each of these viruses acts by infecting white blood cells. HTLV-I and HTLV-II cause the abnormal proliferation of the infected lymphocytes, i.e. cancer, whereas HTLV-III leads to the death of WBC which eventually incapacitates the immune system and thus causes AIDS.

We will now discuss in brief cellular transformation caused by DNA and RNA viruses separately.

● Transformation by DNA Viruses

Burkitt's Lymphoma, the cancerous tumor of the jaw is more prevalent in the areas where mosquito transmitted infections are common. The epidemic nature of this lymphoma led to the proposal that it is transmitted by a mosquito-borne infectious agent. Later studies have shown this agent to be virus.

Oncogenic DNA viruses consist of three major groups; herpes viruses, adenoviruses and papovaviruses. **Herpes viruses** cause a variety of malignant and nonmalignant diseases. Some examples of malignant tumors are Burkitt's lymphoma. (the tumor of jaw) and nasopharyngeal carcinoma in humans, kidney cancer in frogs and contagious type of lymphoma in chickens known as Mereb's disease. Among non-malignant diseases caused by herpes virus groups are chicken pox, wild sores and sexually transmitted diseases.

Adenoviruses are a group of DNA viruses commonly found in respiratory tract. Of more than 30 kinds of adenoviruses isolated from human sources, more than a dozen cause cancer when injected into animals. Although these viruses cause various respiratory ailments in humans, there is no evidence that they cause cancer in humans.

Papovaviruses are another group of DNA viruses that induce tumor. The most thoroughly studied papovaviruses are polyoma and SV-40 both of which cause cancer in laboratory animals but apparently not in humans. Another example is of papillomaviruses (HPVs), which include several dozen viruses that have been isolated from a variety of human tissues. Benign growths such as warts are caused by these viruses. Also HPV genes have been detected in the cancer cells of most women suffering from the cancer of the uterine cervix.

Let us now study how DNA tumor viruses act on the host cells. The mode of action of all the three groups of DNA tumor viruses appears to be similar. The viral DNA after entering the host cell is transcribed into mRNA molecules, which are translated into viral proteins. After this both cellular and viral DNA replicate and then cell division follows. Some weeks after the virus-stimulated cell division, one or more copies of the viral DNA become integrated into the host cell DNA. Thus viral genetic information becomes a permanent part of the cell's genetic material, and the cell is said to be transformed. During the entire process, no new virus particle is produced or released. So you can understand that if infective virus particle cannot be isolated from a human tumor, viral origin of the tumor cannot be ruled out. Oncogenic adenoviruses and papovaviruses have much smaller number of genes. Therefore, it has been possible to pinpoint the particular gene required for triggering the tumor development. On the other hand oncogenic herpes viruses code for at least 50 polypeptides, so it has been difficult to identify the genes responsible for malignant transformation.

● Transformation by RNA Viruses

RNA viruses have been linked to a wide variety of animal and plant malignancies, as well as to a few types of human cancer. Some examples of oncogenic RNA viruses are Rous Sarcoma Virus, Bittner mouse mammary virus and human T-cell lymphotropic viruses. Such viruses after infecting the cell can remain there for a longer time in a 'latent' form, i.e. inactive form. Such latent virus usually becomes active when the cell is exposed to triggering conditions such as radiations, chemicals, hormones or even other viruses. For example, feline leukemia virus can remain latent in otherwise healthy cats for many years. The exposure to stressful situations like mild respiratory infection activates the latent virus to induce the formation of malignancy . You have already known about the structure of RSV virus which is a retrovirus. All the RNA tumor viruses are retroviruses. We will now discuss the mode of action of an RNA virus. Look for the steps involved in cellular transformation by RNA tumor virus in Fig. 17.4.

When a retrovirus of RNA tumor virus infects the cell, the cell produces virally encoded enzyme called **reverse trancriptase**. This enzyme catalyses the reverse transcription or synthesis of DNA using viral RNA as tempelate. This synthesised DNA, called **provirus**, integrates into the host chromosome and undergoes replication along with the chromosomal DNA.

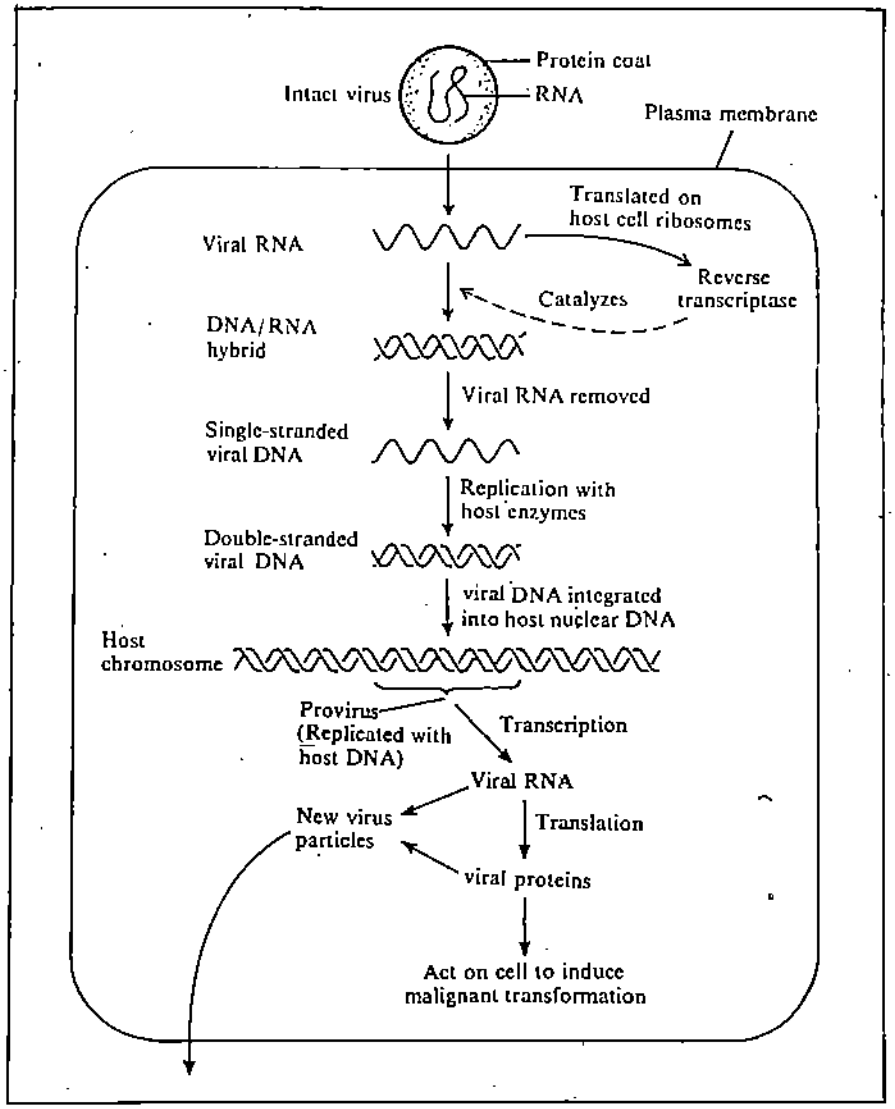


Fig. 17.4 : Mode of action of oncogenic RNA virus in the cellular transformation.

The portion of the integrated viral DNA responsible for transforming a normal cell into a malignant one is called an oncogene. When oncogene is expressed, the host cell undergoes malignant transformation. RSV's transforming gene is called *src* after the sarcoma tumor it induces. Therefore, you can say that V-onc gene (viral oncogene) of RSV is *v-src*. All the tumor inducing retroviruses are morphologically alike and have probably descended from one ancestral virus.

It has become possible to isolate and characterise specific oncogenes associated with the development of a cancer. This has helped us to come a step closer to understanding what causes cells to become malignant. Ultimately, however, we need to know how the oncogene encoded protein products cause malignancy. Let us briefly go through the classes of protein products encoded by various known oncogenes.

Classes of Oncogene Products

Oncogene products known till now can be divided into four major categories.

(1) Proteins resembling protein kinases are encoded by about a dozen oncogenes. Some of these are tyrosine-specific kinases, while others catalyse the phosphorylation of serine and threonine residues.

(2) Second type of oncogene codes for proteins related to growth factors. For example a protein that resembles platelet-derived growth factor (PDGF).

(3) Third type of proteins are GTP binding proteins, that are encoded by small group of oncogenes. These proteins are associated with the inner surface of plasma membrane and are thought to play a role in controlling the transmission of growth control signals from plasma membrane to the interior of the cell.

(4) The fourth category includes the protein products of oncogenes that function within the nucleus. Such proteins are thought to control the rate of DNA transcription and /or replication.

17.2.4 Genetic Aspects of Cancer

You know that changes in DNA induced by chemicals, radiations and viruses can all cause cancer. One aspect of cancerous transformation of the cell is that mutations have also a role in the formation of malignant tumor. It is the susceptibility to develop cancer rather than cancer itself that is inherited. For example, in USA, the probability of the average women to develop breast cancer is 1 in 12. But among women who have blood relatives with breast cancer the chance increases to 1 in 5. Another case is the Xeroderma pigmentosum about which you have already read. You may recall that in patients suffering from this disease a specific enzyme involved in DNA repair is deficient. Because the skin cells of such persons are less able to repair DNA damage, an elevated susceptibility to skin cancer results.

One of the ways in which the carcinogens may act to cause cancer is by converting a protooncogene into oncogene or, in the case of oncogenic viruses by directly introducing an oncogene. However, the fact is that oncogenes till now have been detected in less than a quarter of all human cancers. Therefore, it is premature to assume that protooncogenes and oncogenes will be able to explain all types of cancer.

The other aspect of cellular transformation is that they are caused by somatic mutations. The evidence for this hypothesis comes from the fact that many cancers have one or more distinctive abnormal chromosomes. In such cases of cancers, every cell in the tumor possesses the abnormal chromosome, but usually none of the normal cells have these. There is every possibility that some of them upset the normal control devices regulating cell division and thus cause cancer.

Thus we can say that transformation of normal cell into a cancerous one invariably involves changes in the cell's genome causing the cell to lose growth control. The next few years will bring about a revaluation in our understanding of the molecular basis of cancer as well as that of the regulation of the cellular activity. This understanding is expected to provide not only the clue but also cure for this dreadful disease. In the next section we will discuss the prevention and cure of cancer, but before that try the following SAQ.

SAQ 2

(a) Match the items given in column I with those of column II:

	Column I		Column II
(i)	Retroviruses.	(a)	DNA molecules are incorporated into host cell DNA after cell division.
(ii)	Viral DNA.	(b)	Two copies of RNA genome are present.
(iii)	Provirus.	(c)	Somatic mutations can be one of the aspect of cellular transforamtion.
(iv)	Cancer genetics.	(d)	DNA is synthesised with viral RNA as a template with the help of reverse transcriptase.

(b) Draw the diagram of a retrovirus and label it.

17.3 PREVENTION AND CURE FOR CANCER

Until recently, more attention was paid to cancer treatment than prevention of cancer. But now that we know that 75 per cent or more of all types of cancer is caused by identifiable factors, cancer prevention is attracting more attention. Let us first discuss how cancer can be prevented.

17.3.1 Prevention of Cancer

The first step to prevent environmentally induced cancer is to identify the agents that are carcinogenic. For example, with the knowledge of carcinogenic properties of sunlight and x-rays, we can reduce radiation induced cancers by avoiding excess and unnecessary exposure to sunlight and medical x-rays. Identification of tumor viruses might lead to the development of preventive vaccines. Such vaccines have already been developed for leukemia occurring in cats and chickens, but the applicability of vaccine to human cancer is not yet clear. The greatest help to prevent cancer could be made by eliminating the chemical carcinogens currently present in our food, air, water, clothing and drugs. Such carcinogens account for major proportion of all human cancer. To remove such carcinogens present in the environment, one has to first identify them. There are two ways of identification: epidemiological analysis and laboratory testing. Let us discuss these one by one.

In the **epidemiological approach** the occurrence of different types of cancer is compared among different populations. It has been observed that in different parts of the world different types of cancer occur with different frequencies. For example, stomach cancer is usually frequent in Japan, breast cancer is prominent in the United States, rectal cancer rates are high in Denmark and esophageal cancer is prevalent in Iran. Theoretically the reasons for such differences could be either hereditary or environmental. However, studies on the individuals who migrated from one country to another suggest that of the two factors, the environment is more important. For example in Japanese families immigrating to United States, the incidence of stomach cancer is gradually reduced to a great extent, indicating the role of environmental factors.

The identification of the nature of environmental factors that cause cancer is also possible with the help of epidemiological data. You can take the example of lung cancer, a disease that has increased over tenfolds in frequency since 1930 as shown in Fig.17.5. When the environmental factors potentially responsible for this epidemic of lung cancer were investigated, it was found that virtually all the victims of this disease share one environmental component in common i.e. cigarette smoking. You know from Table 17.1 that cigarette smoke contains several carcinogenic chemicals.

The major drawback of epidemiological approach is that carcinogens are identified only after they have caused a considerable damage to human health and have increased the mortality rate. It is also limited by the fact that there is a long time-lag between the exposure to a carcinogen and development of cancer. As you can see in Fig. 17.5 that more than 20 years elapsed before the increase in cigarette smoking could be correlated to increased rate of lung cancer.

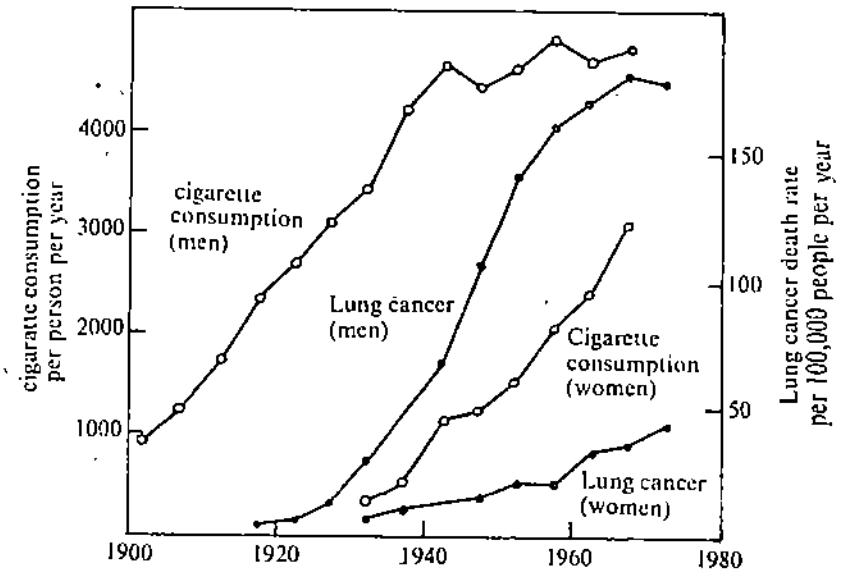


Fig. 17.5: The relationship between increased smoking and increased incidence of lung cancer in a developed country. Increased incidence of lung cancer occurred after 20-30 years of increased cigarette consumption. In females, lung cancer just started to increase because of recent increase in cigarette consumption.

The second approach for identifying carcinogenic agents is direct **laboratory testing.** Till recently the normal practice to conduct such tests was to administer potential carcinogens to laboratory animals. But limitations with such tests are that they take several years to complete, are quite expensive and data obtained may not be applicable to human populations. Also with recent additions of thousands of chemical pollutants to our environment, laboratory testing of such type has become highly impractical.

To overcome this problem, Bruce Ames and his associates devised a lab test based on the fact that most carcinogens are mutagens. Let us now discuss what the Ames test is all about. Ames test uses a mutant strain of bacterium *Salmonella typhimurium* , which cannot synthesise amino acid histidine. The bacterial cells are added to a culture plate containing growth medium which lacks histidine but contains the chemical to be tested for mutagenicity and a component of rat liver cells – the microsomal fraction containing liver enzymes. At the same time a control experiment is run in which no potential carcinogen is present.

In the living organism, as in humans the liver enzymes perform the task of detoxifying and toxifying various chemicals including many potential mutagens. Therefore, the presence of liver enzymes makes it possible to determine whether a chemical which by itself is not a mutant, can become mutagenic by being metabolised in the liver.

These culture plates are incubated for two days and then examined for bacterial growth. Only the bacteria that have undergone mutation which makes them produce their own histidine can grow into colonies. You can infer from presence of mutants on the plates spread with the chemical-treated mixtures that the chemical is a mutagen. A positive result for mutagenicity on the Ames test is generally accepted as an indication of carcinogenic nature of the substance. Over 90% of all the chemicals known to be carcinogenic in humans have given positive results on Ames test. The control plate shows no bacterial growth, meaning thereby, the histidine deficient bacteria do not get mutated in the absence of chemicals. You can see in Fig. 17.6 the steps involved in performing the Ames test for potential mutagens.

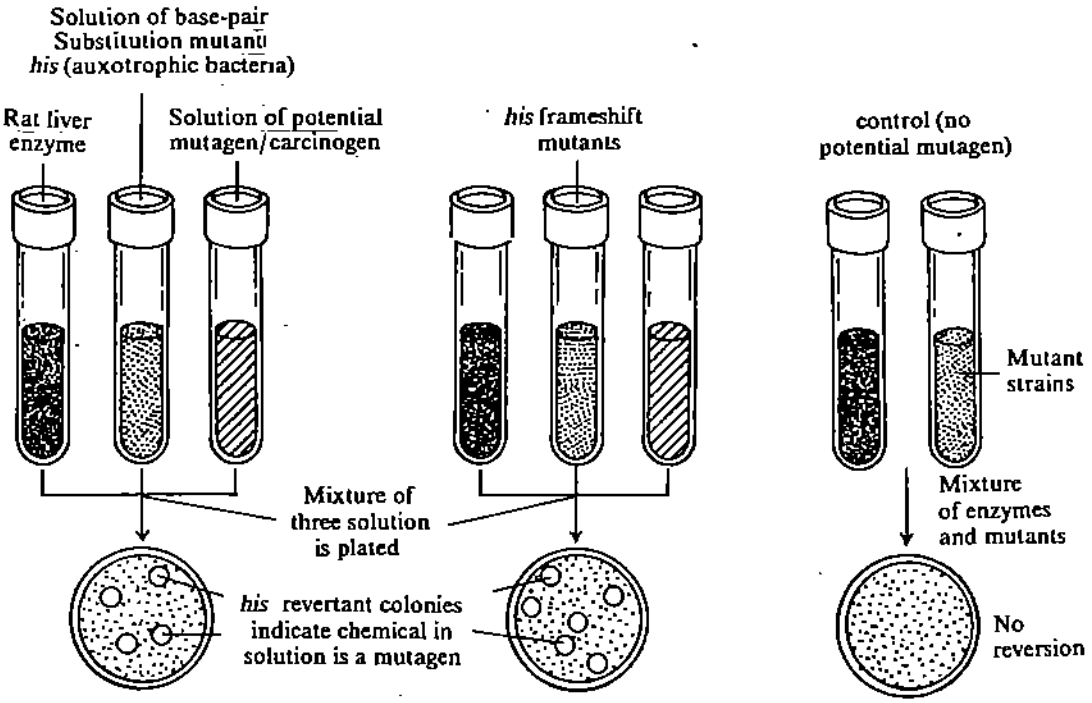


Fig. 17.6: Ames test showing the growth of histidine deficit bacteria due to the mutation by the chemical present in the growth medium. The control plate shows no growth. The number of bacterial colonies indicate the mutagenic potency of the chemical being tested.

The identification of potential carcinogens by any of the above methods is only the first step in assessing the danger posed by the substances present in the environment. Once a particular carcinogen is pinpointed it becomes important to know about its tolerance level in the environment before declaring that a clear hazard exists. Such an understanding helps in public policy decisions about its permissible environmental levels. On personal level careful attention to habits such as smoking and diet would allow the average person to reduce his or her risk of developing cancer. After prevention let us now discuss about treatment of the cancer.

17.3.2 Treatment of Cancer

No matter how seriously we all at personal and governmental level pursue ways for cancer prevention, it is impossible to eliminate every environmental carcinogen. Even if all human made carcinogens are barred, which is of course, highly unlikely, natural carcinogens still exist. You can take the examples of radiations such as cosmic and sunlight and also the naturally occurring chemicals. In addition inherited mutations in DNA cause a certain amount of cancer that cannot be easily prevented. Thus despite our best efforts to prevent cancer, methods for treating cancer will always be needed. Presently surgery, radiations and chemotherapy are the three procedures to treat cancer. We will now discuss briefly about these procedures.

Surgery is used to remove tumors that are localised to a particular part of the body. Breast cancer at an early stage in women is quite successfully treated by surgical approach. Surgery is effective in two aspects. Firstly, when cancer is of nonmetastasizing type so that it does not spread to the other parts of the body. Secondly, when the cancer is detected early enough so that the chances of metastasis are minimized. However, cancers in internal organs often do not produce clinical symptoms until they are in advanced stage and disseminated throughout the body. This reduces the effectiveness of the surgical approach. Now you can understand why early detection of cancer is important.

Radiation therapy is used to treat certain kinds of cancers, because the cells involved in DNA synthesis or mitosis are sensitive to the killing effects of radiation as they can damage the DNA. Irradiation with radioactive substances such as cobalt-60 or X-rays has been effective in treating cancers like skin cancer, Hodgkin's disease (a kind of lymphoma) and certain form of bone cancers. One of the major disadvantages of radiation therapy is

that along with cancerous cells it destroys the normal dividing cells. To overcome this problem limited doses of radiations can be used. Also recall that radiations themselves are carcinogenic, thus causing additional risk to develop cancer.

Chemotherapy is also widely used way for cancer treatment. In this approach the drugs are carried by the circulatory system throughout the body and so it is effective for treatment of tumors that have already metastasized. The problem with drug therapy is that drugs inhibit the division of normal as well as cancer cells producing toxic side effects such as diarrhoea, loss of hair and susceptibility to infections. These side effects are caused by destruction of intestinal lining cells, hair follicle cells and blood cells respectively. In certain cancer cases, drugs alone are responsible for the improved survival rate, while in others chemotherapy is used along with surgery or radiation therapy.

However, till now it has not become possible to completely eradicate cancer. Cancer biologists continue to search for a way by which only cancer cells are destroyed without damaging normal cells in process. Till the time we are able to find ways to completely cure cancer, prevention is the best cure: a check on smoking, proper dietary habits and cleaner environment. Now it is time you do another SAQ. After that we will proceed to discuss teratogenesis.

SAQ 3

Fill in the blanks with appropriate words from the text.

(i) Carcinogens to be removed from the environment are needed to be .....first.

(ii) ..... was introduced to overcome the limitations or previously practised way of carcinogen identification.

(iii) ..... is not very effective treatment for highly metastasizing tumors.

(iv) Drugs are effective in treating metastasizing cancers as they reach the effective parts of the body through .....

17.4 TERATOGENESIS

You are aware by now that various mutagenic agents such as chemicals and radiations can adversely affect human health by causing mutations. Many of such environmental agents are probably also teratogenic in nature, that is they produce physical defects or deformities in developing embryos. It should be clear to you that mutations and teratogenesis are not the same things. However, it is likely that mutations are an important step for teratogenesis to occur. It is because developmental abnormalities are often associated with genetic damage. So you can define **teratogenesis** as the phenomenon by which abnormalities in developing embryo are caused due to the damage of genetic material or its expression. It is well known that several agents (teratogens) produce changes that interfere with normal cellular differentiation and produce terata (terata means monster). The teratogens that damage the genetic material are **genotoxic** in nature, whereas **nonogenotoxic** agents affect the genetic expression without affecting the genetic material as such. In recent years, *in vitro* cell cultures, embryo cultures and invertebrate models have been developed which are based on detecting the changes in molecular event such as protein synthesis or cell organization produced by various agents. These tests have turned out to be predictive tests for teratogenicity.

The first three months of pregnancy is a particularly vulnerable time for a developing foetus because it is during this period that the basic tissues and organ systems are established. Chemicals that normally have little or no harmful effect on humans may cause severe damage to foetus during this period. Let us briefly go through the agents that can have teratogenic effects on a new born.

Chemical teratogens: The most notorious teratogen in the recent decades has been the drug **thalidomide**. It was widely used as a sedative in Europe in early 1960s. Pregnant women took thalidomide as a medication for nausea. Children of women who took this medicine during first trimester of pregnancy were born with abnormalities in bones of the arms and/or legs known as thalidomide syndrome. A rare genetic disease with same phenotypic effects is called **phocomelia**. So it was discovered that thalidomide produced

phenocopies of phocomelia. As soon as the teratogenic effects of thalidomide were recognised, the drug was removed from the market.

Another drug where teratogenic effects have been recognised recently is **alcohol.** The other effects of alcohol addiction and misuse are well known. Alcohol easily crosses placenta and reaches the foetus. Its concentration in foetus is same as in mother's blood stream. Alcohol causes abnormal development of head and nervous system. Mothers who drink alcohol have an increased incidence of still birth. For women taking the anticonvulsant phenylhydantoin, the risk of birth defect is very high i.e. about 40 per cent, with cardiac defects being most common. You can now understand why pregnant women should avoid coffee, tobacco, alcohol and drugs unless advised by a physician who is knowledgeable about teratogens. During pregnancy, even such drugs as aspirin should be taken only on doctor's advice.

Radiation teratogens: Radiations such as X-rays in large doses, can seriously disturb the developing embryo. Many deformed babies were born following the Hiroshima and Nagasaki atomic bombing. Developmental abnormalities in the new born of the women who were exposed to radiation at first trimester of their pregnancy were more severe than those exposed during the last trimester of pregnancy. It is true that a woman in her second trimester or later month of pregnancy is not at risk for teratogenic effects for her child from chest X-rays; however, there may be a very slight risk of gene mutation to a subsequent generation or even cancer for that child.

Viral teratogens or cellular parasites: Women, who become exposed to viruses such as **Rubella** (German measles) in the first trimester of their pregnancy can give birth to new borns with development defects. Other viral infections such as cytomegalovirus can also cause birth defects. Cellular parasite called **taxoplasma** can be inhaled from cat faeces and cause damage to developing embryos. It must be clear to you now that it is imperative on the part of women and physicians to have knowledge of possible teratogens and to avoid them during pregnancy at least. Also no medication should be given except for the most compelling reasons during the first three months of pregnancy.

An important concern has arisen regarding the male contribution of teratogenicity. While males can theoretically contribute to genetically derived congenital malformations, reproductive deficiencies, spontaneous abortions and others, the male contribution to nongenetic teratogenecity needs careful investigations.

With this we end our discussions on carcinogenesis and teratogenesis. If you want to know further about these aspects, read the books enlisted under further reading.

SAQ 4

Tick mark (✓) the correct options in the following statements.

i) Phenocopies of phocomelia are the teratogenic effects caused by alcohol/ thalidomide.

ii) Abnormalities in developing embryos are/are not caused by genetic/non-genetic factors.

iii) Developmental abnormalities in women who are exposed to teratogens in first trimester are more/less severe.

v) Medical experts need be/not be careful while treating pregnant women.

17.5 SUMMARY

In this unit you have studied that:

- Normal cells undergo cell division with such precision that it forms the basis of normal differentiation and function. Sometimes when this control is lost, the cell is transformed into cancer cell and the phenomenon is called carcinogenesis.
- Cancer cells are different in many aspects from normal cells. High correlation between mutation and cancer has been shown. Most of the cancer causing agents are mutagens.
- Since many of the cancers are caused by identifiable environmental agents, it is important to plant strategies for cancer prevention at governmental and personal level. Presently surgery, chemotherapy and radiation therapy are the three approaches to treat cancers.

- Various environmental agents can be teratogenic and they cause developmental abnormalities in the embryo. It becomes the responsibility of health professionals to avoid prescription of drugs to the pregnant women especially in first trimester, until of course it becomes very much necessary. On the part of pregnant women it is necessary that they should acquire the knowledge and concern about teratogens. Of course, it is also the responsibility of the governmental agencies to check and control the manufacture and distribution of teratogenic drugs in the market.

17.6 TERMINAL QUESTIONS

1) Write one difference between each of the following:

   a) Oncogenesis and Metastasis.

   .....

   .....

   .....

   .....

   b) v-oncs and c-oncs.

   .....

   .....

   .....

   .....

   c) Initiation and Promotion.

   .....

   .....

   .....

   .....

(2) Briefly explain the mode of action of RNA virus in cellular transformation.

.....

.....

.....

.....

(3) Name three modes of cancer treatment and write one limitation of each.

.....

.....

.....

.....

(4) Do you think that it is possible to completely cure teratogenic effects of a drug in the new born? Give reasons in support of your answer.

.....

.....

.....

.....

17.7 ANSWERS

Self Assessment Questions

1) i) Metastasis
ii) C-oncs.
iii) Initiation.
iv) Promotion.

2) a) i) b ii) a. iii) d. iv) c.

3) i) Identified, ii) Ames test, iii) Surgery iv) Blood

4) i) thalidomide, ii) genetic, iii) more, iv) need be.

Terminal Questions

1) a) Oncogenesis is initiation of cancer in an organism. Metastasis is spreading of tumor cells in other parts of the body.

b) v-oncs are cancer inducing genes in a virus, c-oncs are cellular genes that have potential to become oncogenes.

c) Initiation is the alteration in DNA of cells due to the effect of carcinogens.

d) Promotion is a process of stimulation of division of damged cells.

2) After RNA virus infects the cell the host cell produces enzyme reverse transcriptase. With the help of this enzyme DNA is synthesised on RNA template. This DNA integrates into host cell DNA and helps in replication and division of the cell.

3) i) Surgery—It is less effective in highly metastasizing cancers.

ii) Radiation therapy—Radiations destroy the normal dividing cells and are also carcinogenic.

iii) Chemotherapy—Drugs can produce side effects like diarrhoea and susceptibility to infections. In addition they can also be mutagenic/carcinogenic.

4) It is not possible to completely cure the teratogenic effects of a drug in the child as the abnormalities occur during its embryonic and developmental stages. The only way to avoid the problem is to prevent the pregnant mother from use of such drugs.

UNIT 18 IMMUNOGENETICS

Structure

18.1 INTRODUCTION

The environment is at constant biological threat by a large number and variety of microorganisms which cannot be seen with naked eye. Infact, a drop of tap water or a speck of dust contains enough organisms which could at times be harmful to humans, animals and plants. The fact that animals, humans, and plants are healthy in spite of being surrounded by pathogens suggests that by some mechanism the invading pathogens are inactivated in the host body. Such specific devices are called specific immune responses as against the non-specific immunity which is present from birth that forms the first line of defense. This unit will deal with certain aspects of immune response and the genetics of antibody diversity.

The specific immune defense is a highly specialised system that recognises the invader and discriminates between self constituents and invading agents (the nonself), and has the ability to retain immunological memory of a particular antigen. Specific immune responses are of two types-one that involves direct attack by certain white blood cells (lymphocytes) on foreign antigens that have become established within the cells and is known as **cell mediated immunity** (CMI), the other type is called **humoral immune response** in which highly specific proteins called antibodies are secreted in blood against invading microorganisms such as viruses and bacteria that have not yet invaded the cells.

Although there is no evidence that plants produce antibodies against infection by pathogens, there is some evidence that such reactions do occur is plants. An antibody like response showing resistance is exhibited by the introduction of foreign proteins (bacteria, viruses and fungi) as shown by serological studies with **linseed rust** and **angular leaf spot** of cotton. This type of response in plants is also genetically controlled. Some non-specific defenses such as surface barrier and chemical inhibitors produced by plants against disease causing organisms are well known. Certain fungi infected plants are stimulated to produce substances called **phytolexins** that are toxic to fungi and thus protect the host plants. This shows that plants have some inbuilt mechanisms of defense that protect them to survive several diseases that attack them and that no diseases has been able to wipe them out from the earth. You will read more about defense mechanisms in plants in higher level courses.

Immunogenetics has now become one of the most challenging disciplines in modern biology. In recent times with the advant of modern techniques like, genetic engineering and monoclonal antibodies, it has come to the forefront of modern sciences. (See LSE-01 for the reference)

In this unit we will study certain fundamental concepts of immune response and the genetics of antibody diversity in relation to human and animal systems,

Objectives

- After reading this unit you should be able to:
- define immune response, antigen, antibody, recognition,
- self versus nonself, specific and non specific defenses,
- describe classes, sturcture and properties of antibody,
- differentiate between types of cells involved and compare the humoral with cell mediated immunity,
- describe the genetic basis of antibody sturcture.

18.2 SOME BASIC CONCEPTS

18.2.1 Non-Specific Defense System

Animals and plants are protected from the invasion of harmful agents by a variety of physical (skin, mucous membrane) and chemical (pH, enzymes ect) barriers. They are summarised below in Table 18.1.

Table 18.1: Physical and chemical agents that defend the organisms from foregin invaders.

1.	Intact skin which only a few bacteria can penetrate. Plants living on land develop water-proof tough surfaces, which are impenetrable to pathogenic bacteria and fungi. These invaders can grow only inside the plant's internal environment,
2.	Ciliated mucous membranes, that line parts of the respiratory tract and that act like sticky brooms to sweep out bacteria and various inhaled particles.
3.	Secretions from exocrine glands on surface epithelium (one such secretion is lysozyme, an enzyme that helps degrade the cell walls of many bacteria). Some plant cells synthesize organic compounds that interfere with the metabolic process of the invaders.
4.	Gastric fluid in the stomach, the acidic nature of which destroys many potential invaders.
5.	Microbes that normally inhabit the gut (and, in females, also the vagina), and that compete effectively with many potential microbial invaders, thereby helping to keep them in check.

18.2.2 Specific Immune Response

Specific immune responses are processes in which animals synthesise specifically reactive protein molecules and cells in response to a large variety of organic macromolecules and pathogens which are foreign to the organism. These responses are known to occur mostly in vertebrates and save them from infections. There are two ways in which the immune system responds to infection (Fig. 18.1).

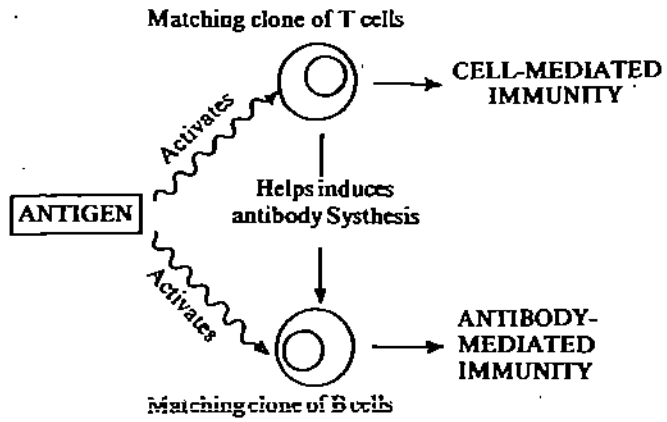


Fig. 18.1 : A schematic diagram showing two types of immunity , the cell mediated and humoral or antibody mediated. Also the interaction between T cells and B cells in the production of antibody is shown.

One is known as **humoral immune response** in which proteins known as antibody molecules are produced by specific B lymphocytes. The antibody is secreted in the blood

plasma, lymph and tissue fluid that is in circulation, hence the immune response is called humoral mediated antibody response. The second type, called **cellular immune response** involves direct interaction between lymphocytes (T-cells) and the invading agents. There are different classes of T-cells. Each T-cell carries out a specific function (Refer section 18.3). The immunity provided by these mechanisms is calle **active immunity.** Antibodies produced by the mother provide immunity to the infants through placenta, till their own immune system starts functioning. This type of immunity; is termed as **passive immunity.**

Antibodies react against specific microorganisms, their toxic products and other compounds. They can be used in the treatment of infection caused by their respective organism and also prevent infection and diseases caused by these agents.

You may have heard the story of Edward Jenner, the founder of immunology. In the late eighteenth century, Jenner observed that the farm workers who had been exposed to cows infected with cowpox, (a non-pathogen to man) did not contract small pox which at that time was very common and deadly disease. He reasoned that the cow pox must be doing something to the farm workers so that they were protected from further infection by small pox. This intelligent guess led to a pioneering experiment in which he took out live organisms from a milkmaid having cow pox infection and inoculated into a child. The child was subsequently infected with live organisms from a person suffering from small pox. This child did not develop small pox. This scientific study led to development of the concept of vaccine. Some diseases like small pox and polio have almost been eliminated due to vaccination process.

The immune system functions by **learning, recognising** and **retaining the memory** of the specific causative agent as a foreign agent. It was observed that if host is infected again by the same organism the latter will be killed. This was the first immunological experiment that initiated the concept of **vaccines** – agents that cause protection against pathogens. Later, Louis Pasteur and a long line of other workers developed vaccines for many other infectious diseases. Pasteur while working with anthrax and cholera organisms observed that exposure to a crippled or attenuated form of pathogen led to protection against the virulent form. This formed the basis of modern vaccines. In fact, the first experiment directed to understanding the nature of immunity was carried out by Behring, who took out serum from rabbits infected with tetanus, and injected it into animals having tetanus infection. He observed that under these conditions most of the animals survived. This was the first demonstration that something produced in the body in response to the pathogen provides immunity. Later, it was found that a specific class of molecules called antibody, synthesised by the body in response to an invading agent (like bacteria, viruses and parasites) offer protection. The antibodies fight the invading pathogens and kill them.

An important observation of the immune responses was made in *Mycobacterium tuberculosis* (the agent that causes tuberculosis). When this bacterium was injected under the skin of an individual, with an earlier exposure to this pathogen, it resulted in a red swelling. However, if one had not been exposed to the pathgen, no such swelling was seen. This particular response now can be transferred from one animal to another not through serum but through cells isolated from the animal exposed to pathogen. This kind of response which is transferred from one person to another with the help of cells is an important component of **acquired immune response.** We have now a number of vaccines being usesd for protection against diseases (Table 18.2).

Tabel 18.2: Principal vaccines used to prevent bacterial diseases in humans

Disease	Vaccine
Cholera	Crude fraction of *Virbrio Cholerae*
Diphtheria	Purified diphtheria toxoid
Meningococcal meningitis	Purified polysaccharide from *Neisseria meningitidis*
Pertussis (whooping cough)	Killed *Bordetella pertussis*
Plague	Crude fraction of *Yorsin pestis*
Pneumococci pneumonia	Purified polysaccharide from *Streptococcus pneumoniae*
Telanus	Purified tetanus toxoid
Tuberculosis	*Mycobacterium bovis* BCG
Typhoid fever and paratyphoid fever	killed *Salmonella typhi, S. schottmulleri,* and *S. paratyphi*

18.2.4 Antigen

Any foreign material when introduced into the body inducing the production of antibody is called an **antigen**. An antigen specifically reacts with its own antibody produced in response to its presence.

The immune system of the organism is endowed with the property of recognising any foreign material that enters into the body.

The antigen site which a given antibody molecule recognises and binds is called **antigenic determinant**. For example a bacterium may have from 10-1000 distinct macromolecules on its surface that are recognized by its antibodies. This specific recognition and binding of an antigenic determinant to the antibody forms the basis of the formation of **antigen antibody complex**. The antigen and antibody fit together just like **lock and key** or an **enzyme and its substrate** by hydrophobic, ionic and van der Waals forces, but no covalent bond is formed between an antigen and its antibody.

An antigen that stimulates a specific immune response, (the synthesis of antibodies) is called an **immunogen**. Macromolecules such as glycoproteins, liporoteins, nucleoproteins that are part of the invaders such as cilia flagella, pili, toxins, coats or viruses and many other types of cell surface molecules are potent stimulators of antibody production. The blood serum or cells of one species of animal may be antigenic when introduced into another species. Antigens present in human red blood cells in some persons react with the antibody called "isoantibody" present in the red blood cells or other persons. Since there are several antigens that are associated with human RBCs there are several blood groups based on these antigens such as A,B,O, and AB (refer to Unit of Block 4 of LSE-03 course). Other familiar antigens are pollen, egg white and transplanted tissues and organs. Substances with a molecular weight of less than 10,000 are not good **immunogens**. Thus high molecular weight, foreignness and chemical complexity of the molecules are basic requirements for an immunogen. However, small molecules when covalently conjugated to a good immunogenic protein, elicit a good immune response. Such small molecules are called "**haptens**" and the proteins to which they are conjugated are called "**carriers**". For example, a very small molecule like, 2, 4-dinitrophenol can be used to elicit good antibody response after conjugating it to a carrier. Although by itself it can not evoke antibody production, a hapten can react with its antibody even in the absensce of the carrier (Fig. 18.2). A well known example of hapten is drug penicillin which itself is not antigenic, but sometimes combines with serum protein of some persons and causes an immune response. Haptens are useful for the study of the structure of antibody.

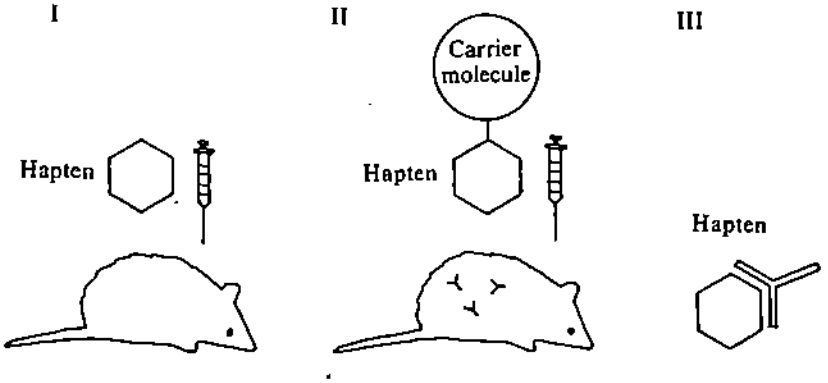


Fig. 18.2 : Antibody production for haptens. The non-immunogenic hapten (I) is conjugated with a high molecular weight carrier molecule (II) . When the antibody is produced, it reacts with the hapten even in the absence of carrier molecule (III).

18.2.5 Antibodies

All antibodies are globular proteins produced by B lymphocytes. They belong to a class of serum proteins called globulins. They are found in the blood serum and body's most important defense system. They remain in blood for a long time inactivating or destroying antigens.

Because of their globular structure and their migratory properties in an electrophoretic field (you have read about electrophoresis in cell Biology course - LSE-01, Block 1, Unit

4,) antibodies are termed γ-globulins. They are the slowest moving proteins in relation to albumin, alpha and beta-globulins. Since they carry a positive charge they move towards anode (Fig. 18.3). They are all collectively called as immunoglobulins, abbreviated as Ig. IgG (Immunoglobulin G) is the most common type of immunoglobulin molecule present in circulation and constitutes about 70-80 per cent of the total Ig population. It readily crosses the walls of blood vessels and tissue fluids. Maternal IgG can cross the placenta and reach the circulation of foetus where it confers passive immunity to the child till it is 5 to 6 months old.

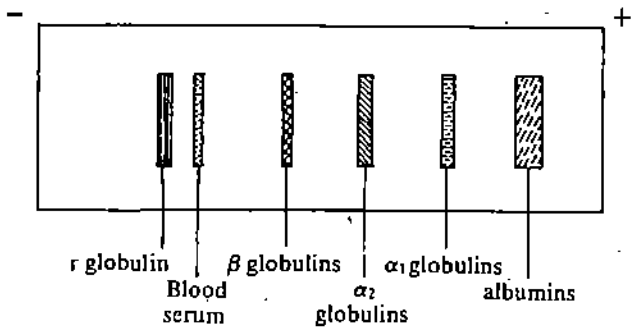


Fig. 18.3 : The movement of different classes of serum proteins under the influence of an electric field. Note the γ globulins move towards the anode as they are positively charged molecules.

If different components of the serum are separated by various immunological techniques, one could isolate a variety of fractions comprising different molecular species of antibodies such as IgG, IgE, IgD, IgM and IgA. They have a number of biological functions such as precipitation, neutralisation of toxins, agglutination of cells and their lysis. If you compare serum components containing antibodies from a healthy animal with the one infected with a pathogen, only one of the components termed gamma globulins (γ) increases considerably in the infected animal. Moreover, injection of purified gamma globulins from an infected animal into a healthy one confers on the latter immunity for the specific infection, suggesting that gamma globulins are involved in immune response.

SAQ 1

Say True or False.

i) The cell mediated immune response is directed against intracellular phase of infection by producing specific antibodies by B lymphocytes.

ii) Passive immunity refers to the type of immunity conferred on the baby by its mother through the antibodies that diffuse through the placenta, till such a time when the baby's, immune system starts functioning.

iii) Acquired immune response refers to the type of immune response transferred through serum from the individual exposed to the pathogen to another individual.

iv) The term antigenic determinant refers to the specific antigenic site which a given antibody molecule recognises and binds.

v) Immunogen is antibody molecule produced as a specific immune response to an antigen.

vi) Molecules having a weight less than 10,000 and known as haptens elicit an immune response only when conjugated covalently with a larger molecule, the carrier.

vii) Gamma globulins are high molecular weight serum proteins carrying a net negative charge.

18.3 CELLS INVOLVED IN IMMUNE RESPONSE

Immune response is dependent on the interaction among different tyes of cells inolved in the immune system (Fig. 18.4). Each one of the different cell types that are to be described below has a specific role to play in the defence of the host body against any foreign invasion.

18.3.1 The Cell Types of Immune System

1. **Macrophages :** Macrophages are a group of specialised leucocytes which are considered as the first line of defence in any immune reaction. Although macrophages do not exhibit any antigenic specificity, they are known for their extensive phagocytic activity. The entry of any foreign organism, be it a bacterium, virus or protozoan invites

the immediate attention of the macrophages which virtually rush to the site of infection and engulf them. The engulfing and subsequent digestion is achieved by the degradative enzymes found in their lysosomes. Another function of macrophages is to attach a portion of the antigen of the invading organism they engulf to their surface and display it for recognition by apporpriate and specific T lymphocytes. In other words, the macrophages breakdown the trapped invader, digest them partially and present them on their surfaces to specific T cells.

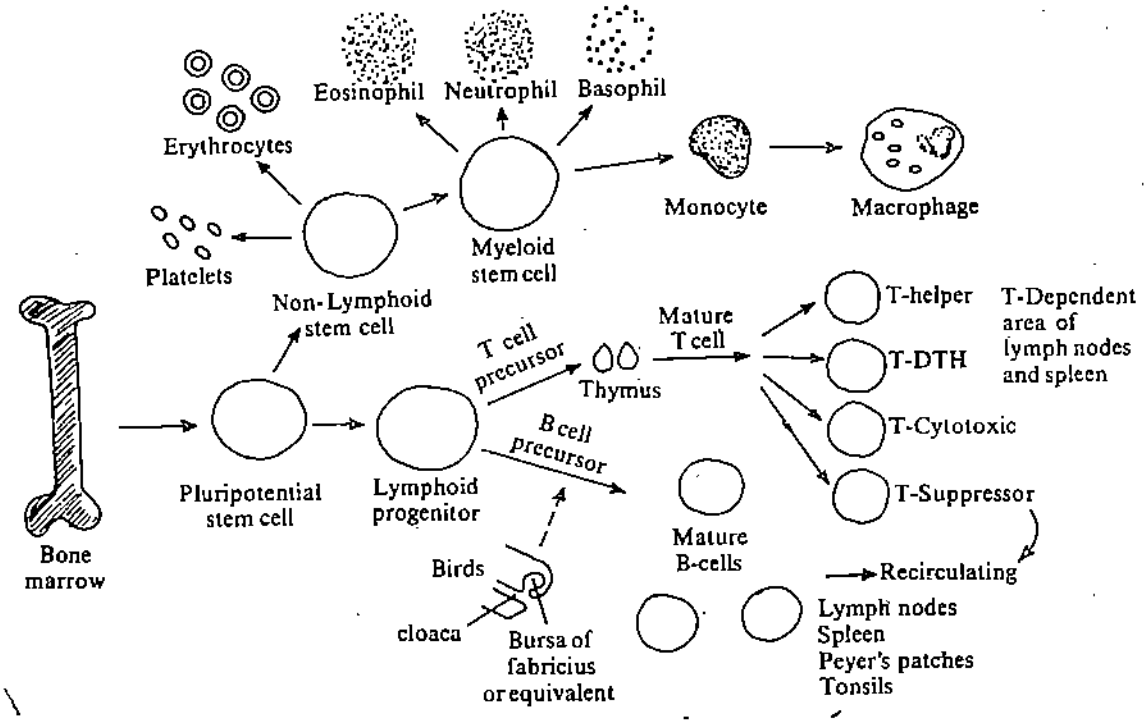


Fig. 18.4 : The different types of T lymphocytes and B lymphocyte that take part in eliciting immune response.

2. **T lymphocytes:** These cells originate in bone marrow as stem cells and differentiate into lymphocytes in thymus gland. Since the differentiation takes place in thymus these are termed T cells or thymus derived cells. T lymphocytes are prime structures involved in cell mediated immunity. They are not capable of synthesising antibodies but do help the other class of lymphocytes, the β lymphocytes to synthesise antibodies. Further, they also promote the proliferation of other types of T lymphocytes such as helper T cells, killer T cells, memory T cells and supressor T cells. In short, T lymphocytes are the important agents in the regulation of immune response.

3. **B lymphocytes:** B lymphocytes also originate as stem cells in bone marrow but the site of their differentiation in mammals is not exactly known. It is believed that cell differentiation may possibly occur in the lymph nodes or other lymphoid organs. It is also believed that they may differentiate in the bone marrow itself. The cells are so termed due to their discovery first in an organ called bursa Fabricius of birds. They are the bursa derived or B cells. An analogous organ is not found in mammals, but in mammals also the term B cells is continued to be used. B cells exhibit a very high degree of antigenic specificity and their differentiation and proliferation are in response to their coming in contact with specific antigens. These are the antibody synthesising cells.

The Role of T and B lymphocytes in Eliciting an Immune Response (Fig. 18.5)

When an organism is invaded by any foreign material, as we said earlier, the phagocytic cells are the first ones to arrive at the site of invasion. Such phagocytes include, besides macrophages, other cells as neutrophils, dendrite cells and kupfer cells of the liver. Earlier we mentioned that a macrophage while phagocytosing the foreign body, also attaches antigens of the pathogen to its surface in order to present it to the specific T cells. Such a presentation is the first signal for the activation of T cells. Additionally, the macrophages release a second signal in the form of a chemical messenger called lymphokines.

Interleukin I (IL-I) is the lymphokine released by the accessory cells to activate the T cells. The activated T cells may undergo proliferation and produce more of helper T cells which are themselves capable of phagocytic activity. The activated T cells release another lymphokine, the interleukin II (IL-II) which induces differentiation of some of the T lymphocytes to killer T cells on cytotoxic T cells which are also involved in the removal of the foreign body.

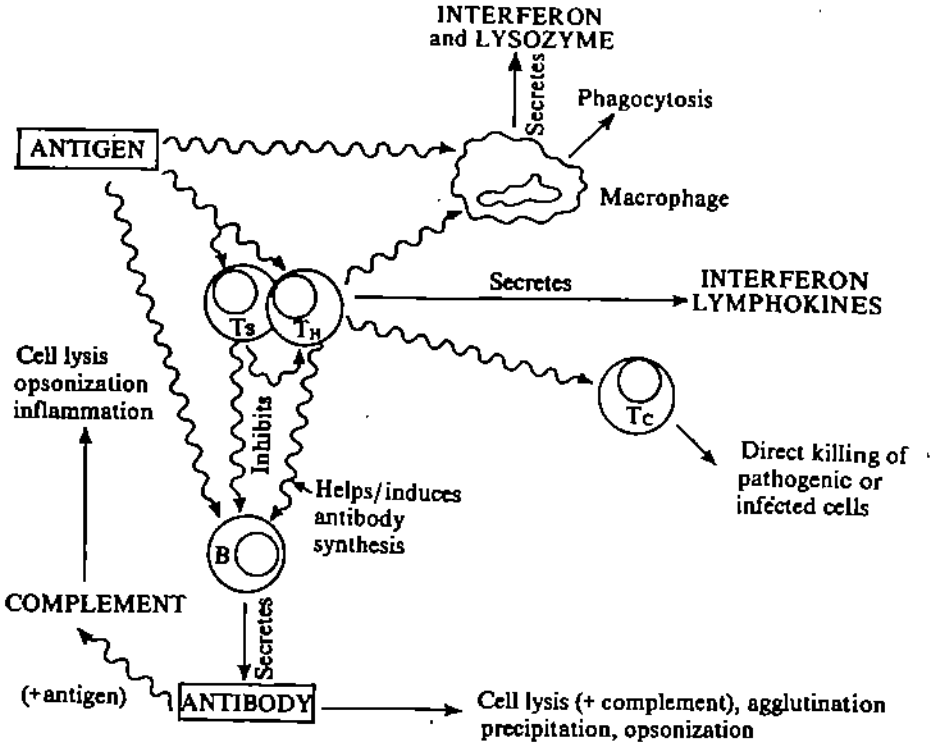


Fig. 18.5 : Schematic diagram showing the role of different types of T cells and the B cells in producing the immune response.

Another function of the T lymphocytes is to trigger the activation of B lymphocytes which are the primary cells involved in humoral mediated immunity. Two types of lymphokines are produced by T cell which trigger B cell differentiation and proliferation. One is B cell growth factor (BCGF) which induces the specific B cells to undergo proliferation, and the second is B cell differentiation factor (BCDF) which induces the newly dividing B cells to differentiate into antibody producing cells. The antibodies or immunoglobulins are the defense proteins which very specifically bind to the antigens of invading organisms and neutralise them.

Some of the lymphocytes are known as memory cells as they possess the ability to recall the previous contact with a particular antigen, and that yet another exposure to such an antigen leads to a more rapid and larger production of antibody. Finally once the immune reaction is complete and the attack by the invader is replied, a group of T lymphocytes, the supressor cells, deactivate the entire immune machinery so that further production of antibodies by B cells as well as the various functions T lymphocytes are suspended. You would observe from the above account that in eliciting an immune response the involvement of various types of cells is necessary and the various cell types act in conjunction to produce a desired immune response.

The various types of lymphocytes, the types of immunity they exhibit, the site of their differentiation and functions are presented in Table 18.3.

Tabel 18.3 : Different types of lymphocytes and their functions

Cell	Type of Immunity	Site of Differentiation	Functions
B Cell	Humoral	Primarily bone marrow and lymph rodes	Differentiates into plasma cells that secrete antibodies
Killer T cell (Tc)	Cell-mediated	Thymus	Destroys target cells upon contact
Helper T cell (T_h)	Cell-mediated	Thymus	Necessary for B cell activation by T-dependent antigens
Supressor T cell (Ts)	Cell-mediated	Thymus	Regulates immune responses and helps maintain tolerance
Memory T cell (Tm)	Cell-mediated	Thymus	Recalls the previous contact with a particular antigen and yet another exposure to the antigen leads to a more rapid and larger production of anbtibody.

18.3.2 Antigen-antibody Reactions

The antigen-antibody reactions which are of prime importance in immune defence may be categorised into two types: 1. Precipitin reaction, 2. Agglutination reaction. The precipitates or agglutinates formed as a result of such reactions are engulfed by the phagocytic cells of non-specific immune system. We shall briefly look into the two types of antigen-antibody reactions.

1. Precipitin Reaction

This reaction depends on the ability of antibodies to precipitate soluble antigens in solution. Most antigen molecules possess hydrophilic surfaces which interact with the surrounding water molecules resulting in their solubility. Antibodies are also soluble, and when an antibody combines with an antigen, an antigen-antibody complex is formed (Fig. 18.6). An antibody (IgG) is a bivalent molecule (refer to section 18.5.1) which means that it can combine with two antigens simultaneously (Fig. 18.6a). This property enables an antibody to cross link two or more molecules of antigen. Fig. 18.6b illustrates the way in which an antigen molecule with a bivalent IgG molecule. This cross linking essentially results in a network or lattice (Fig. 18.6c) and as the lattice grows it precipitates out of solution. The ability of antibodies to precipitate soluble antigens is known as precipitin reaction. The precipitin reaction is the basis for several immunoassy procedures performed in the laboratory.

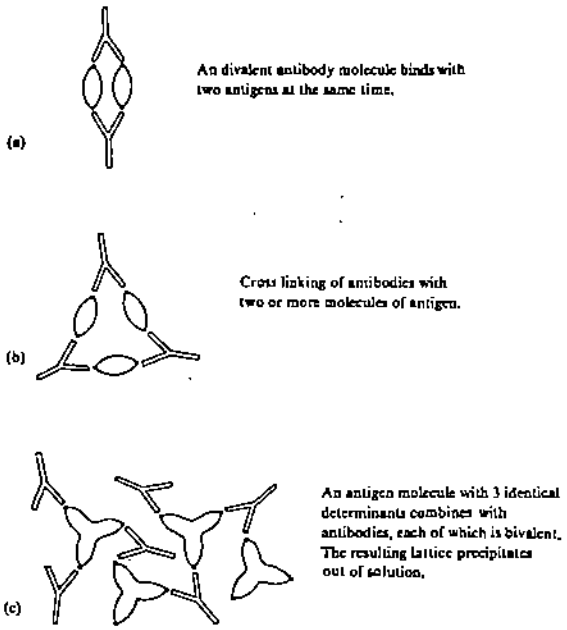


Fig. 18.6: Antigen – antibody reactions : a) A divalent antibody molecule binds with two *antigens at the same time.*

b) Cross linking of antibodies with two or more *molecules of antigen.*
c) An antign molecule with 3 identical determinants combines with antibodies, each of which is bivalent. The resulting lattice precipitates out of solution.

2. Agglutination Reaction

Agglutination reaction refers to the ability of the antibody molecules to cause foreign cells to stick together in clumps. Essentially the immungenic cells such as bacteria are agglutinated or clumped very similar to the way we described the precipitation reaction. IgM class are the first antibodies produced by a clone of differentiated lymphocytes. IgM have ten potential antigen combining sites that is suited for an agglutination reaction.

The function of antibodies, as we have seen so far is to agglutinate or precipilate the antigens. Either one of these two reactions is involved in eliminating antigens which could be soluble toxins released by bacteria responsible for such disease as tetanus or diphteria or viral particles such as common cold virus or infuenza virus or blood borne parasite.

The insoluble complexes thus formed, that is, the precipitate or the clumps, are engulfed and destroyed by specialised white blood cells or phagocytes. As we discussed earlier, the macrophage and the polymorphonuclear leucocytes, are the phagocytic agents which destroy the antigens once the antibodies have immobilised them. Besides their abililty to engulf the live pathogens, the phagocytic agents help in the removal of a lot of debris that accumulate as a result of antigen-antibody reactions. The phagocytic activity also interferes with bacterial cell divisions, aids in trapping viral particles and deactivates the toxins released by bacteria.

18.4 COMPLEMENT SYSTEM

Complement are a group of serum proteins which in the presence of antigen-antibody complexes exhibit certain specific biological properties, the chief among them being their ability to cause the lysis of the microorganisms or cells. They do not have the structure of immunoglobulins nor their numbers increase after immunisation. In mammalian system a group of nine proteins—C1 to C9—are known to form the complement system which are activated in a sequential manner, each reaction in the sequence being activated by an enzyme. The enzyme required for the activation arises as a result of a cascading effect which means that an enzyms required for the activation of a complement is synthesised in response to the activation of the previous complement. The final products of the pathway, the C8 and C9 are the ones which are directed against the microorgisms or cell and cause the lysis of the membrane.

SAQ 2

Match the items given in (A) to the description given in (B)

	A	B
1	Macrophage	a) Serum proteins capable of causing the lysis of microorganisms or cells in the presence of antigen-antibody complex.
2	B lymphocytes	b) Factors responsible for the differentiation of T lymphocytes to killer T cells.
3	Interleukin II	c) Cells which recall the previous contact with a particular antigen and further exposure to such an antigen leads to a more rapid and larger production of antibody.
4	BCGF	d) Factors which induce specific B lympho cytes to undergo proliferation and increase in number.
5	Memory T cells	e) Phagocytic cells which are the first line of defense in any immune response.
6	BCDF	F) Factor which induce the newly divided B cells to differentiate into antibody forming cells.
7	Complement	g) Cell which undergo differentiation in lymphoid tissue and synthesise antibodies against specific antigens.

18.5 GENETICS OF ANTIBODY DIVERSITY

To any understandilng of genes of immunoglobulins, it is important that you know the structure of immunoglobulins first. Cancerous plasma cells or lympocytes called

plasmacytoma or multiple myeloma cells often produce single type of immunoglobulins in large quantities. The proteins have been purified and their structure elucidated. R.R. Porter and G. Edelman shared the Nobel prize in 1972 for the deduction of the structure of the antibodies.

There are atleast five classes of antibodies. In this section you will learn the structure of one of the abundantly present immunoglobulins in mammalian system—the IgG molecule. The other classes of immunoglobulins would be described briefly later.

18.5.1 The Structure of an IgG Molecule

In its overall structure, an antibody resembles a two pronged fork or the alphabet 'Y'. Each IgG molecule is composed of four polypeptide chains interconnected by disulphide bridges (Fig.18.7). Two of the chains are longer hence called heavy chains (H) and the other two are shorter, the light chains (L). In any one immunoglobulin molecule the two light chains and the two heavy chains are identical. Each light chain or heavy chain is characterised by specific domains in it. Thus, a light chain has two domains, a variable domain (V_L) and a constant domain (C_L). The heavy chains similarily have a variable domain (V_H) but three constant domains (C_H1, C_H2, C_H3). The V_H and C_H1 are located in the prong of the fork and the C_H2 and C_H3 are located in the handle. The V domains are located at the tip of the prong of the fork and C domains of the molecule occupy the centre and handle positions.

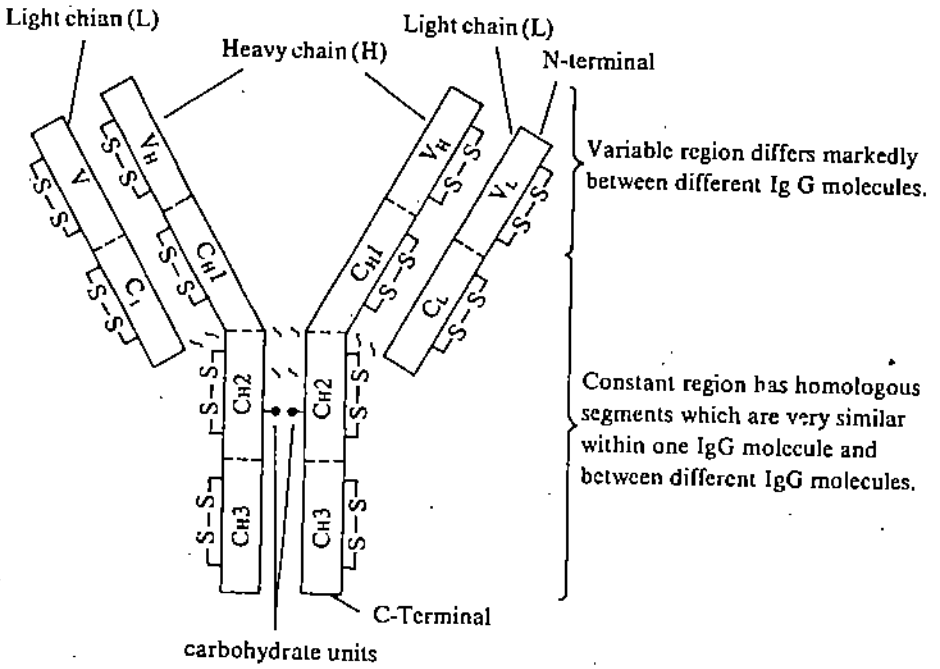


Fig. 18.7 : Structure of an IgG molecule. The molecule is a glycoprotein and the carbohydrate molety is attached to the constant domains of heavy chains.

The variable domains of lilght and heavy chains together constitue an antigen binding site. You were earlier told that each antibody binds specifically to an antigen and therefore, it is obvious that the antigen binding sites of different antibodies must be strikingly different from one another. There are 2 antigen binding sites for each antibody hence each antibody is a bivalent molecule. This means that an antibody can at a time bind to 2 antigens.

The constant domains play a significant role in eliciting the immune response. Once an antigen is recognised and bound to the variable portion of the antibody, the constant domains undergo a change in their conformation. Such a change in structure induces a chain reaction in which the macrophages and the T-lymphocytes, the cellular mediators of immune response migrate to the site of antigen to climate it from the body.

18.5.2 Structure of Variable and Constant Domains of the Light and Heavy Chains

The V_L domain of the light chain polypetide occupying the N terminus consists of the first 107 amino acids and the C_L domain comprises of amino acids 108 to 220 (Fig. 18.8). The constant domain is not as constant as the name implies. It may be composed of any one of the two subclasses, either kappa (K) or lambda (λ) polypeptide. The V_L domain as you could see from the Figure 18.8 is composed of four framework regions (F_1, F_2, F_3 , and F_4) and three hypervariabe regions (HVRI, HVR2, HVR3). When an IgG molecule folds to form a funtional molecule, the HVR regions come together to form the antigen binding site and the F regions provide the necessary support for the structure. The amino acid sequence in the HVR region tends to be very different from one antibody to another, and the framework regions exhibit less variability.

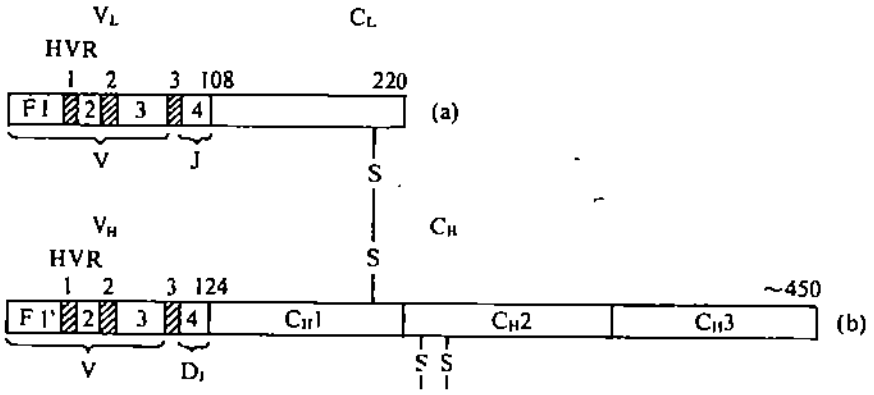


Fig. 18.8 : Detailed structure of a light Chain (a) and a heavy chain (b) in an IgG molecule. Numbers 1-220 in the light chain and 1-450 in the heavy chain refer to amino acid sequences. HVR-hypervariable region; F-framework region.

The V_H region occupying the N terminal end has 123 amino acids. Here also there are four framework regions (F) and three hypervariable regions (HVR). The HVR regions of V_H domain alongwith those V_L domain form the antigen binding site. The constant region of heavy chain comprises of amino acids 124 to about 450. In an IgG molecule the heavy chain is composed of a specific polypeptide chain the γ (gamma) chain, hence the molecule is called IgG. Besides IgG there are four other classes of immunoglobulins, IgA, IgD, IgE and IgM, each one of which carries a specific type of C_H chain. The IgA has an α (alpha) chain IgD has a δ (delta) chain, IgE has an ε (epsilon) chain and IgM has a μ (micron) chain. For a detailed classification you may refer to the Table 18.4.

Table 18.4: Classes of Human Immunoglobulins

	IgG	IgA	IgM	IgD	IgE
Heavy Chains :					
Class	γ	α	μ	δ	e
Subclass	$\gamma_1\gamma_2\gamma_2\gamma_3$	$\alpha_1\alpha_2$	--	--	--
Mol wt x 10^{-4}	5.5	~6	6-7	~6	~7.5
Light Chains :					
Class	K,λ	K,λ	K,λ	K,λ	K,λ
Mol wt x 10^{-4}	2.2-2.3	2.2-2.3	2.2-2.3	2.2-2.3	2.2-2.3
Whole molecule :					
Formula	$K_2\gamma_2$ or λ2γ2	$(K_2\alpha_2)n$ or $(\lambda_2\alpha_2)n$	$(K_2\mu_2)_5$ or $(\lambda_2\mu_2)_5$	$k_2\delta_2$ or $\lambda_2\delta_2$	$k_2\varepsilon_2$ or $\lambda_2\varepsilon_2$
		n = 1,2,3 . .			
Mol wt x 10^{-5}	~1.5	1.6	9-10	1.7-1.8	~2
Carbohydrate, %	2.9	7.5	7.7-10.7	12	10.7

18.5.3 Light Chain Gene of IgG of Embryonic and Differentiated Cells

From your studies on protein biosynthesis in Unit of Block 3 of Cell Biology course you may be aware that one gene codes for one polypeptide chain. The genetics of antibodies suggests a variation to this hypothesis. The genes for antibodies have to be created by

recombining different segments of DNA and such a constructed gene codes for a polypeptide. We shall study the sturcture of a light chain gene with a specific example. You were told earlier about the plasmacytoma or multiple myeloma cells. Several such cancer cells have been isolated both from mice as well as humans and maintained as separate cell lines. One such mouse cell line is MOPC-41 (MO-mouse and PC-plasmacytoma). The MOPC-41 cell line synthesises an IgG molecule. The IgG contains a kappa polypeptide in the C_L region. The aminocid sequence of the light chain of MOPC-41 IgG is known.

Leder and his associates studied the structure of the embryonic or germline MOPC-41 light chain gene and compared it with that of adult light chain gene.

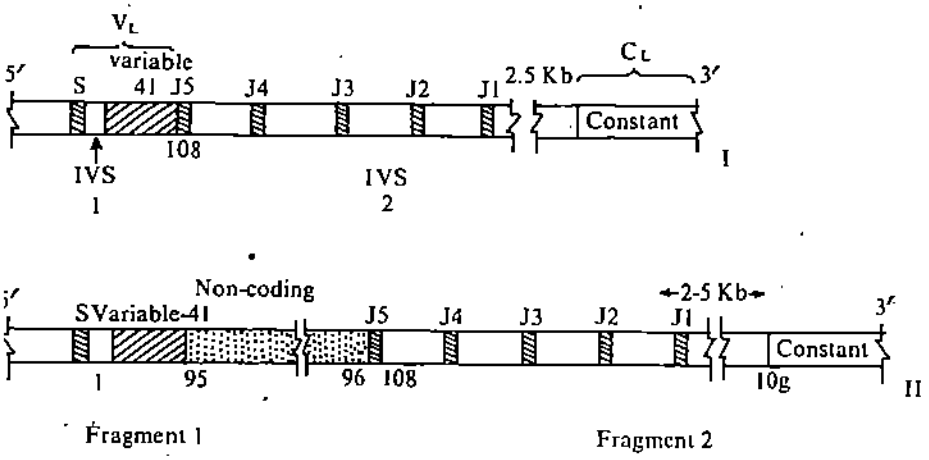


Fig. 18.9 : Diagram showing the somatic (I) and germline (II) sequences of a light chain gene of a mouse embryo. The explantion for abbreviations is provided in the text.

You may look at the structure of the gene in the Fig. 18.9 (I). The fig shows that the V_L domain of the gene is located at the 5' end of the DNA fragment and C_L domain at the 3' end. You may also see that the V_L coding region begins with a short 5' sequence known as signal sequence(S). This is followed by a short intervening sequence (IVS-1). Then follows V_{41} or variable-41 sequence which codes for amino acids 1 to 108 of the V_L domain. Finally, the sequence for C_L domain that is the kappa chain follows, coding from amino acid 109. It is obvious that the gene for light chain of MOPC-41 cell line has three exons (S, V_L and C_L) and two introns (IVS-1 and IVS-2).

5'-S-IVS1-V_L- IVS2- C_L-3'

At the time of transcription the intervening sequences are selectively removed (spliced out) and the mRNA has the following sequence.

5'-S-V_L-C_L-3'

The mRNA is translated into a polypeptide and soon after the commencement of translation, the 18 aminoacid signal sequence is also removed by a proteolyic enzyme. (Read about the post-translational changes in a polypeptide in Unit 14, of Block 3 of Cell Biology). The light chain polypeptide has the final structure as

V_L-C_L

The structure of the light chain gene obtained from MOPC-41 cell line when compared with that of embryonic cells shows (Fig. 18.9 (II)) certain remarkable differences:

- In the germline genome, the V_L segment and the C_L segment are located in 2 different DNA fragments. Earlier we saw that in myeloma genome they are located in the same fragment.
- More interestingly, the V_L sequence itself is split into two sequences, each one located in a different fragment.

Look at the Fig. 18.9 (II) carefully. You will find that the signal sequence (S) and the V_{41} sequence which codes for 1 to 95 aminoacids of the V_L domain of the light chain are located in fragment 1 whereas the sequence coding for the last 13 aminoacids (96-108) designated as J_5 sequence is located in a different fragment. The arrangement of the different sequences of the light chain gene in the embryonic cell can be represented as follows :

5-S-IVSI-V_{41}-noncoding DNA-J_5-IVS2-C_L-3'
fragment 1 fragment 2

EXON : The expressible region of DNA coding for a polypeptide or a segment of a polypeptide

Intron : A segment of DNA that does not code for a protein. These are intervening sequences of nucleotides located between coding sequences or exons.

Essentially at the time of differentiation of an embroyonic cell into a functional lymphocyte the V_{41} sequence located in fragment 1 joins with the J_5 sequence located in fragment 2 to create a functional V_L sequence and in the process the non-coding DNA found between the 2 fragments is discarded.

18.5.4 V_{41}-J_5 Joining Reaction

An understanding of the joining reaction between V_{41} and J_5 sequence requires knowledge about what are known as inverted repeat sequences or palindromic sequences in the DNA. A palindrome is a word which read from either end, reads the same. For example, MADAM-- The word 'madam' when read in both directions reads the same.

Similarly in DNA also there are such palindromic sequences or inverted repeat sequences or internally complementary sequences. Such sequences can reassociate themselves forming double strands. For example, look at the following nucleotide sequence

5' ..... AATGTACGXXXCGTACATT..... 3'

R R'

(x may denote any nucleotide)

It should be obvious to you that the sequence R and R' are complementary to each other. Under proper conditions, the regions R and R' can reassociate and fold back upon themselves like a hairpin.

5' .....AATGTACG X

X

3' .....TTACATGC X

Wherever joining reactions are to be carried out between two different sequences, the occurrence of inverted repeat sequences become very handy.

Let us now return to V-J joining.

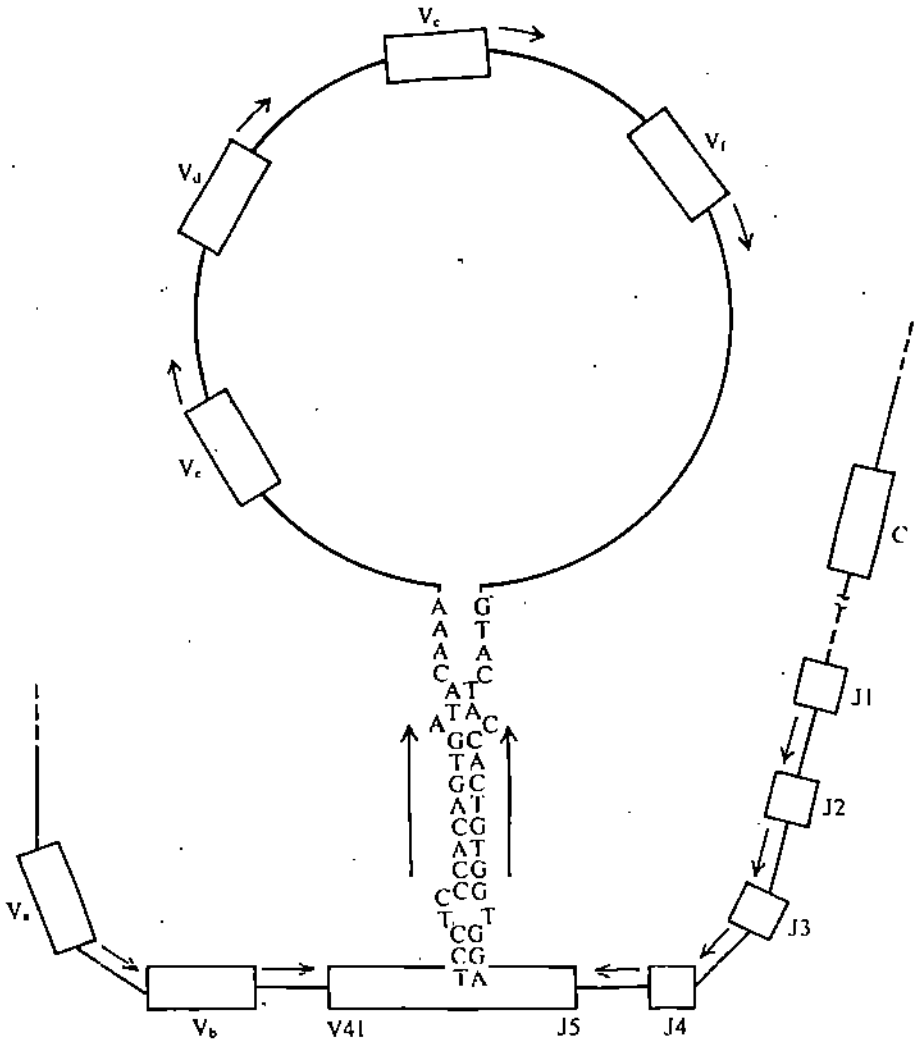


Fig. 18.10 : Diagram showing the coming together of V_{41} and J_5 sequence forming a stem and loop structure. In the stem the inverted repeat sequences are located and in the loop the sequences that are to be excised are located. The reconstruction takes place in the stem region.

The V_{41} sector for germline DNA has an inverted repeat sequence adjacent at its 3' end and the same such sequence is present at the 5' end of J sector as well. Probably at the time of differentiation, the V_{41} and J_5 sequences are brought close togeter by causing all the DNA between them to loop out in the form of a circle as shown in the Fig. 18.10. The sem of the loop i sthe site where the inverted repeat sequences of both V_{41} and J_5 sectors form complementary parirs. The stem is also the site where the recombination takes place and the enzyme for such proces jons V_{41} and J_5 into a continous strand. The DNA in the circle is excised and discarded.

The above joining reaction described for MOPC-41 immunoglobulin gene should hold good with the immunoglobulins produced by other lymphocyte clones. And what is more important is the diversity that is generated by V-J joining. It should be apparent to you from the Fig. 18.10 that there is more than one J region that could combine with V segment. In fact there are five J regions, J_1, J_2, J_3, J_4 and J_5. Similarly the germline DNA contains multiple copies of V sequence each arranged one behind the other in a linear sequence. In our example, the MOPC-41 cell line, it is only the V_{41} segment that combines specifically with J_5 segment. In fact, each lymphocyte clone resorts to specific V-J joining thereby creating not only diversity in the antibodies but much needed specificity for antigen binding. Several myeloma cell lines have been studied for V-J joining reaction and for each cell a unique V-J joining has occured. Thus, for MOPC-21 IgG, the first 96 amino acid specifying V_{21} region joins with the last 12 amino acid specifying J_4 sequence. In MOPC-149 cell line, a V_{149} sequence combines with a J_2 region. It is easy to explain that bringing together of different V and J combinations is one source of antibody diversity. It must also be clear to you that depending upon which V sequence and J sequence are chosen for a joining reaction, the rest of the non-coding regions become part of the circle of discarded DNA.

You may question as to how many such V regions could possibly be present in the mouse embryonic genome. It is believed that there are atleast 200 V sequences. You have earlier learnt that there are five J sequences. But J_3 sequence whenever it combines with any V segment, for some reason, always produces a non-functional gene. In effect there are only four J regions, any one of which can combine with anyone of the 200 V regions, to result in a functional gene. Obviously 4 x 200 = 800 different V_L domains can be created by the V-J joining reaction alone.

We can discuss yet another source of antibody diversity, the diversity being generated by the actual recombination event. You may look at Fig. 18.10 more carefully. The figure shows that during V_{41}-J_5 joining the recombination takes place between the 2nd and 3rd nucleotides of codon 95. For example,

	V_{41}			J_5		pro	trp
CC*T	CCC	x	TGG*	TGG		CCG	TGG
						95	96

The nucleotides marked with asteriks are the one that join together, and the recombined 95th and 96th codons are geneerated. More importantly, if there were to be a shift in the recombination frame, it would result in the appearance of a new 96th amino acid. For instance, if the recombination were to take place between the first and second nucleotides during the V_{41}-J_5 joining then,

	V_{41}			J_5		pro	arg
CCT	CCC	x	TGG	TGG		CCT	CGG
						95	96

Now, the IgG will carry an arginine instead of tryptophan in position 96. Aminoacid sequencing of light chains of MOPC-41 cell line has shown that the 96th position is in fact occupied by anyone of the following aminoacids: (trp, arg, pro, tyr and phe).

It is not without reason that 96th aminoacid of light chain is prone to variation. The 96th aminoacid is located at or close to antigen binding site and is responsible for contact between light and heavy chain. A change in the 96th place brings about a change in the antigen recognition properties of the antibody.

18.5.5 Construction of V_H Domains

You have just learnt that the variable domain of light chain gene can be created by V-D joining. This is true of heavy chain genes as well, except that in creating heavy chain genes an additional segment D is involved. In a heavy chain, the V segment codes upto F3

region (refer to Fig. 18.8), that is, from aminoacid 1 to 101; the D segment coding for a small sequence of five aminoacids from 102 to 106 is confined to a portion of HVR3 region, and J_H sequence codes for the remainder of HVR3 plus the F4 region accounting for amino acids 107 to 123. As in the case of light chain, the mammalian genome has about 200 different V_H sequences, 10 D sequences and 4 functional J_H sequences. Once again, the V, D, J, joining may create 200 x 4 x 10 = 8000 different V_H genes. Apart from the diversity caused by V_H D, J_H joining, the inverted repeat sequences that flank V,D and J regions in the embryonic genome while recombining can generate new aminoacids at V,D and D, J junctions.

You may see that both light chain and heavy chain genes utilise similar joining mechanisms. It must be mentioned that a V_L sequence does not join with any of the J_H sequences or a V_H sequence does not join with any of the J_L sequences. Such joinings are effectively prevented, possibly because genes for light chain and heavy chain are located in different chromosomes.

18.5.6 Combinatorial Associations

Earlier you were told that the light chain gene diversity is created by the joining of germ line sequences in all possible combinations to produce about 800 different V-J_L units. Similarly 8000 different V-D-J_H units can also be produced in the creation of heavy chain genes. Anyone of the 800 light chains can combine with any one of the 8000 heavy chain genes to creat a functional immunoglobulin molecule. Assuming that each lymphocyte clone creates its light chain gene independently of its heavy chain gene there is a possibility to create 800 L x 8000 H = 6.4 x 10^6 different antibody proteins by combinational associations. This number would become much higher when you take into account the junctional diversity created by the shifts in recombination frames.

Additionally, it has now come to be known that V sequences that are identical in the germ line carry mutations in the differentiated lymphocyte genome, a phenomenon known as sometic mutations. Such immunoglobulins have the potential for the recognition of different types of antigens.

The diversity thus created in different ways may account for the production of millions of antibodies in a small organism like mouse and much higher numbers in human being.

18.6 THE MAJOR HISTOCOMPATIBILITY COMPLEX

The mammalian's immune system is capable of distinguishing self from non-self due to a group of protein markers (also called antigen here) known as the **major histocompatibility complex (MHC).** In humans they are called human leucocyte antigen (HLA). Two organisms are said to be histocompatible if they can accept solid tissue transplants from each other and incompatible if they cannot. Compatible organisms have the same spectrum of histocompatibility antigens associated with their cell surfaces and therefore, lack the same spectrum of antibodies. The genes coding for such proteins are present as a complex (five different linked genes) on one chromosome. These MHC are essential for recognition and respond to foreign antigens. In humans they exist on chromosome 6, whereas in mice on chromosome 17.MHC antigens play a crucial role in transplant of an organ or a piece of skin from one person to another. The transplant is recognised as non-self and soon elicits an immune response by killing or rejecting the tissues. But if the transplant tissue is taken from a genetically identical person, the material is recognized as self and not rejected.

18.7 AUTOIMMUNE DISEASES

Self refers to tissue antigens of the individual himself. Non-self refers to the foreign antigens entering into his system.

The body usually possesses mechanisms to prevent antibody formation against self components. However, sometimes it loses its ability to distinguish its own cell surface proteins (antigens) and starts forming antibodies against self-cell and tissues antigens. This type of immune response is known as **autoimmunity.** Some of the diseases in this group are most difficult to treat and frequently lead to premature death. Autoimmune disease may be orgen specific such as glomerulonephritis - a degenerative disease of kidney, or they may be non-orgen specific such as rheumatoid arthritis which can attack connective tissues at different sites at the same time.

There are a number of other autoimmune diseases that result from T cells attacking the various tissues: for example **mutiple sclerosis** (inflammation and destruction of nerve cells). **Myastheni gravise** (a rare disease of neuromuscular junction) is caused by allergic

.eactions due to drugs, certain food and pollens of the plants and several such substances. Many autoimmune diseases are known to have genetic origin for example, multiple sclerosis, rheumatic fever and some of the cancers.

One of the most dreaded disease now prevelent in the world is acquired immuno deficiency syndrome (AIDS) that acts by suppressing the immune system of the victim. AIDS virus (a retrovirus) remains hidden within helper T-cells that are present in the blood or semen of an infected individual. The virus may be present in inactivated helper T-cell for years until some other infection triggers the T-cells to multiply, and express the disease.

Now, monoclonal antibodies are being developed for immunisation against malaria and influenza, hepatitis B. You have read about monocolonal antibodies in LSE-01. They hold promise for cancer treatment, and might also be used to overcome one of the major alternatives to chemotherapy that affects non-target cells. Recently drug molecules are being attached to monoclonal antibodies that may reach only target cells of cancer. Diagnostic kits using monoclonal antibodies for colon cancers are already available.

SAQ 3

a) Answer the following questions to review your understanding of the structure of the antibody.

i) How many domains are there in an IgG molecule? What are their names?

.....

.....

ii) How many antigen binding sites are there in an IgG?

.....

.....

iii) How many domains are there for an antigen binding site? What are their names?

.....

.....

iv) What are constant domains of an IgG?

.....

.....

v) What are the names of the constant light chain domains of an immuno golobulin molecules?

.....

.....

vi) What is the basis of classification of heavy chains of immunoglobulins? What are the classes?

.....

.....

vii) What are the subregions of the variable domains of a light chain? A heavy chain?

.....

b) Describe the following terms.
   i) Inverted repeat sequences.
   ii) Stem and loop structure

.....

.....

.....

.....

c) Write down the sequences of the light chain gene of an embryonic and differentiated lymphocytes.

.....

.....

.....

.....

18.8 SUMMARY

The following is the summary of various concepts that you have learnt in this unit.

- Non-specific and specific immune responses are distinguished. The immune system responds in two ways to contact invading agents. One, the cell mediated immunity in which the macrophages and T lymphocytes remove the foreign bodies by phagocytosis and two, the humoral mediated immunity in which the antibody proteins, the immunoglobulins synthesised and secreted by B lymhocytes causes the agglulination or precipitation of foreign antigens.
- Antigens are substances which evoke a specific immune response leading to the production of antibodies. They would also be called as immuogens since they stimulate a specific immune response. Haptens are small molecules (with the molecular weight less than 10,000) which by themselves cannot evolve antibody production but do elicit an immune response when conjugated with a larger carrier protein.
- Jenner is credited with the discovery of use of vaccines for combating diseases caused by microorganisms, and antennuated microorganisms are used as vaccine for inducing immunity against them.
- The sturcture of antibody molecule is discussed in detail. Antibodies, the negatively charged immunoglobulins are of five classes namely IgA, IgG, IgM, IgD and IgE. IgG resembles a two pronged fork with two light chains and two heavy chains. Each light chain is composed of a variable domain and a constant domain and each heavy chain is composed of a variable and three constant domains. The variable domain in turn is composed of framework (F) and hypervariable (HVR) regions. The constant domain of light chains is either kappa (K) chain or lambda (λ) chain. The classification of antibodies into five classes is essentially based on the type of heavy chain they have. IgA molecules have α chains, IgG molecules γ chain, IgD molecule δ chain, IgM molecule μ chain and IgE moelcules have ε chain.
- The various components of cell mediated immunity act in conjuction to bring about a defense response. Macrophages, helper T cells , killer T cells, memory T cells and supressor T cells are the various components of cell mediated immunity. Lymphokines are the chemical messengers that are involved in the chemical signalling of the activation of various components of cell and humoral mediated immune responses. Further, immune system comprises of a group of serum proteins called complements which in the presence of antigen-antibody complexes causes the lysis of microorganisms or other cells.
- The antibodies present an astonishing diversity in their structure which in turn makes each type of antibody specific to its binding antigen. An organism like mouse can

generate as many as 10^6 to 10^8 antibodies. Antibodies owe their diversity to their genes. An interesting feature of antibody genetics is that the different sequences of DNA are brought together and a functional gene is constructed. Such a construction occurs at the time of the differentiation of embryonic lymphocytes into adult ones. In the construction of an antibody gene, besides the different DNA sequences being brought together, phenomena such as junctional diversity (due to shifts in the recombination frames at the time of joining of different DNA sequences) and combinatorial associations are noticed. In other words, every aspect of antibody genetics is geared to generate diversity in the genes, which in turn accounts for antibody diversity.

- A brief discussion of major histocompatibility complex (MHC) and autoimmune diseases is also included in this unit.

18.9 TERMINAL QUESTIONS

1) Define and describe the following terms:
   a) antigen
   b) antibody
   c) cell mediated immunity
   d) humoral mediated immunity
   e) specific immune response

.....

.....

.....

.....

.....

.....

.....

.....

2) Briefly describe the sturcture of an IgG molecule.

.....

.....

.....

.....

.....

.....

.....

.....

.....

3) List the types of cells involved in the ceH mediated immunity and specify the functions of each one of them.

.....

.....

.....

.....

.....

.....

.....

.....

4) Explain the following terms briefly.
   a) Junctional diversity
   b) Combinatorial associations

.....

18.10 ANSWERS

Self Assessment Questions

1) i) False
   ii) True
   iii) False
   iv) Ture
   v) False
   vi) True
   vii) True

2) 1) e
   2) g
   3) b
   4) d
   5) c
   6) f
   7) a

3) a) i) Six domains. They are V_L, C_L, V_H, C_H1, C_H2, and C_H3.
   ii) Two
   iii) Two V_L and C_L
   iv) C_L, C_H1, C_H2, C_H3
   v) Kappa (k), lambda (λ)
   vi) The type of heavy chain determines the class of immunoglobulin molecules;
   IgG - γ
   Iga - α
   IgD - δ
   IgE - ε
   IgM-μ
   vii) The subregions of a variable domain of a light chain are F_1, F_2, F_3 and F_4; HVR1, HVR2 and HVR3.
   That of a heavy chain is similar to the light chain.

b) i) Inverted repeat sequences refer to sequences of nucleotides which are internally complementary. For instance the sequence

A G A T T A C X X X G T A A T C T

represents 2 sequences of nucleotides separated by 3 nucleotides (x). When necessary the entire sequence forms a hair pin like structure and a double strand, with complementary bases, is formed.

A G A T T A C

X
X
X

T C T A A T G

Inverted repeat sequences occur at sites where joining reaction between two DNA sequences is to be carried out.

ii) Stem and loop structure in an immunoglobulin gene is formed at the time when V and J sequences in a light chain and V_{+1}, D and J sequences in a heavy chain are to be joined. The stem structure is formed between inverted repeats located next to germline V and J region genes. (Refer to Fig 18.10). The loop contains DNA sequences that are spliced out. The stem is the site of recombination and

also is the site where shifts occur in recombination frame leading to the altered aminoacid in 96th position of the polypeptide chain.

Terminal Questions

1) a) **Antigen:** Antigen is a foreign substance that is not normally present in the cells or in the body fluids and is usually capable of eliciting an immune response.

b) **Antibody:** The principal humoral component of an immune system is antibody. It belongs to a family of specialised proteins and is synthesised by B lymphocytes in response to an exposure to an antigen. An antibody will bind specifically to the antigenic determinants on the surface of the antigen alone.

c) **Cell mediated immunity:** The contribution made by the specialised cells called T lymphocytes and macrophages to the immune response is the cell mediated immunity.

d) **Humoral mediated immunity:** The contribution made by the specialised cells called B lymphocytes to the immune response by producing globular protiens called antibodies is the humoral mediated immunity.

e) **Specific immune response:** Specific immune response refers to the ability to discriminate among different antigenic determinants and to respond to those that necessitate a response rather than to make a random, undiferentialted response.

2) IgG molecule resembles a two pronged fork consisting of 2 light chains and 2 heavy chains linked by disulphide bonds. Each light chain comprises of a variable domain and a constant domain. The variable domain broadly contains 108 aminoacids and is composed of 4 framework regions and 3 hypervariable regions. The constant domain could be formed of either one of the two polypeptide chains, the kappa chain or the lambda chains. Each heavy chains consists of one variable domain and three constant domains. The heavy chain in an IgG molecule is known as γ chain. The variable domain of a heavy chain also comprises four framework regions and three hypervariable regions. The variable domains of a light and heavy chain together form the antigen binding site of an IgG molecule (Refer to Fig 18).

3) The types of cells that are involved in cell mediated immunity are: i) Macrophages, ii) Helper T Cells, iii) Killer T cells, iv) Memory T cells, v) Supressor T cells.

i) **Macrophages:** Large phagocytes are amoeba like cells. The macrophages roam through the tissues engulfing bacteria, waste products and cell debris. These are the cells of non-specific immunity.

ii) **Helper T cells:** The helper T cells are the lymphocytes with a variety of functions. They themselves are phagocytic. Also they stimulate the cells of non specific immunity such as macrophages and polymorphs. Helper T cells activate B cells in contact with antigen. Further they activate the cytotoxic or killer T cells. Such activities are possible by the production and release of different kinds of lymphokines by them.

iii) **Killer T cells:** There are T lymhocytes with a particular immune specificity, They are capable of specifically killing its target cell after attachment to the target cell.

iv) **Memory T cells:** T cells which are in an active state of inmunity to a specific antigen, so that a second encounter with the same antigen leads to larger and more rapid response.

v) **Supressor T cells:** T cells when induced produced a specific state of immunologic unresponsiveness. This type of unresponsiveness is passively transferred by supressor T cells or their soluble products at the end of an immune response.

) a) **Junctional Diversity:** Junctional diversity is the mechanism of creating diversity in the antibody molecule at the time of V-J joining in the light chains and V-D-J joining in the heavy chains. In man there are at least 200 V sequences, 5 J sequences for light chains and 200 V sequences, 10 D sequences and 5 J sequences for heavy chains. Any one of the V sequences join with any one of the J sequences and this is true of heavy chains also. Diverse amino acid codons are generated at the sites of $V\text{-}J_L$ and $V\text{-}D\text{-}J_H$ recombination.

Combinatorial Association: In the synthesis of an immunoglobulin molecule any set of light chains presumably combine with any set of heavy chains to form a two pronged antibody molecule. A lymphocyte clone can create light chain genes independently of heavy chain genes. Approximately 800 light chain genes and 8000 heavy chain genes can be synthesised in the immune system of a mammal. This means 800L x 8000H = 64×10^6 different antibody proteins are possible by

with negative regulation, in which a regulatory element must be removed from DNA.

promoter : a specific DNA sequence at which RNA polymerase binds and initiates transcription.

promoter site (promoter sequence, promoter) : a specific regulatory nucleotide sequence in the DNA.

prophage : a temperate bacteriophage integrated into the chromosome of a lysogenic bacterium. It replicates with the replication of the host cell's chromosome.

regulated gene : a gene whose activity is controlled in response to the needs of a cell or organism.

regulatory gene : a gene with the primary function of controlling the rate of synthesis of products of one or more other genes.

respressible enzyme : an enzyme whose synthesis is diminished by a regulatory molecule.

retrovirus : a RNA virus that transcribes DNA using its RNA as the template and the enzyme reverse transcriptase.

structural gene : a gene that codes for a mRNA molecule and hence for a polypeptide chain.

tautomeric shift : shifts in the hydrogen atoms and double bonds in a DNA molecule that causes reversible changes of one isomer of a DNA base to another.

teratogen : an agent that causes birth defects during development.

transposons : a DNA unit capable of moving from one genomic location to another.

FURTHER READING

1. Freifelder, David. 1990. *Microbial Genetics.* Narosa Publishing House, New Delhi.
2. Freifelder, David. 1993. *Molecular Biology.* Second Edition. Narosa Publishing House, new Delhi.
3. Gardner, E.J..; Simmons, M.J. & Snustad, d.P. 1991. Eight Edition. *Principles of Genetics.* John Wiley & Sons, Inc. New York.
4. Gupta, P.K. 1987. Second Edition. *Genetics.* Rastogi Publications, Meerut.
5. Klug, W.S. & Cummings, M.R. 1986. *Concepts of Genetics.* Second Edition. Scott foresman & Co., Illinois.
6. Lewin, B. 1983 *Genes.* Wiley Eastern Ltd., New Delhi.
7. Rao, K.R. 1986. Third Edition. *Text Book of Biochemistry.* Prentice Hall of India (Pvt.) Ltd., New Delhi.
8. Lloyd. J.R. 1986. *Genes and Chromosomes.* English Language Book Society, Macmillan, Hong Kong.

Dear Student,

While studying these units you may have found certain portions of the text difficult to comprehend. We wish to know your difficulties and suggestions in order to improve the course. Therefore, we request you to fill and send us the following questionnaire which pertains to this block.

QUESTIONNAIRE

LSE-03
Block-3

Enrolment No. | | | | | | | | | |

1) How many hours did you need for studying the units ?

Unit Number	13	14	15	16	17	18
No. of hours						

2) How many hours (approximately) did you take to do the assigments pertaining to this block?

Assignment Number		
No. of hours		

3) In the following table we have listed 4 kinds of difficulties that we thought you might have come across. Kindly tick (✓) the type of difficulty and give the relevant page number in the appropriate columns.

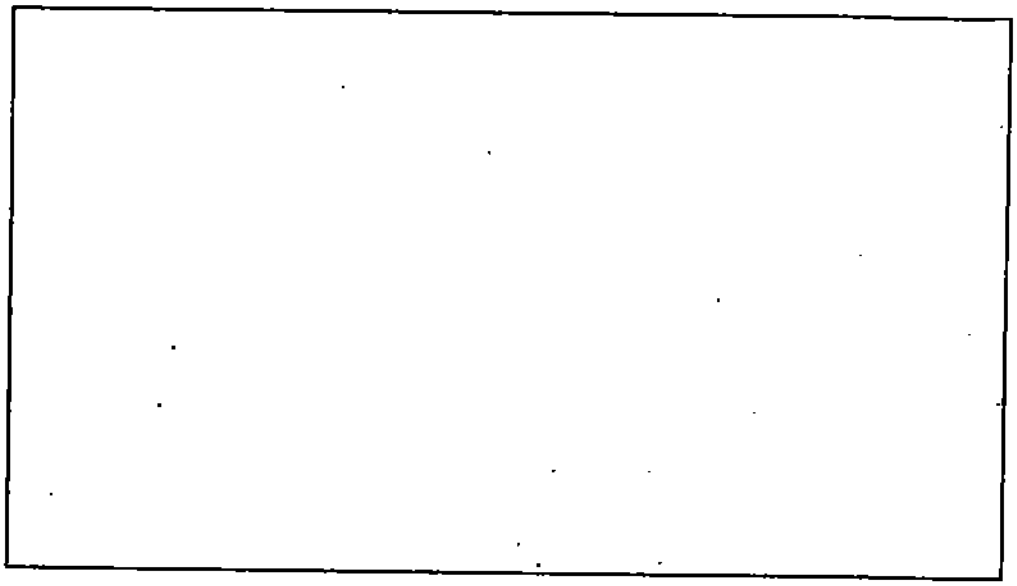
Page number and line number	Types of difficulties			
	Presentation is not clear	Language is difficult	Diagram is not clear	Words/Terms are not explained

4) It is possible that you could not attempt some SAQs and TQs. In the following table are listed the possible difficulties. Kindly tick (✓) the type of difficulty and the relevant unit and question numbers in the appropriate columns.

Unit No.	SAQ No.	TQ No.	Type of difficulty			
			Not clearly posed information	Cannot answer on the basis of information given	Answer given at end of unit not clear	Answer given is not sufficient

5) Were all the difficult terms included in the glossary? If not, please list in the space given below.

6) Any Other Suggestion(s)



To,

The Course Coordinator (LSE-03; Genetics)
School of Sciences
Indira Gandhi National Open University
Maidan Garhi
New Delhi - 110 068

NOTES

NOTES

Uttar Pradesh
Rajarshi Tandon Open University

UGZY/BY-07
Genetics

Block

4

POPULATION AND APPLIED GENETICS

BLOCK 4 POPULATION AND APPLIED GENETICS

In this last block of the course, the present and future applications of the science of genetics are dealt with. The block comprises of six units - Units 19 to 24. In Unit 19, we discuss the genetics of blood. More specifically, you will study the genetics of blood groups and that of haemoglobin. The study of genetics of blood groups assumes importance in blood transfusion as well as in tissue and organ transplantation. The unit also summarises the blood group related disorders in humans. You will also learn about the clustering of α and β chain genes of haemoglobin in chromosomes 16 and 11 respectively.

Unit 20 is concerned with one particular aspect of genetics namely the population genetics. In this unit, we shall study the behaviour of genes in a Mendelian population. We shall begin with the central theorem of population genetics - the Hardy-Weinberg theorem and then proceed to discuss how the frequencies of alleles in a population are altered by evolutionary forces such as mutation, selection and genetic drift. In short, this unit is concerned with the changes in the frequencies of genes in a population and such changes form the genetic basis of evolution.

Unit 21 discusses the traits that are controlled by more than one gene, known as polygenic inheritance. By and large such traits are quantitative traits. In this unit effects of environment on quantitative traits is also discussed. In addition, the concept of heritability in relation to quantitative indices is described in detail. Some salient aspects of genetics of twins are also hightlighted.

Unit 22 deals with the genetic factors that influence behaviour. Three different approaches for the study of behavioural patterns are discussed in detail. Special emphasis is laid on various studies related to intelligence quotient (I.Q.).

Unit 23, 'Genetics and Human Welfare' deals with applications of genetics to well-being of humans. This unit has two distinct parts in it. In the first part, we discuss one aspect of application of genetics namely genetic counselling. Genetic counselling enables any married couple to educate themselves of the probabilities of any of their children born with a genetic disorder. A genetic counsellor based on the family history of the couple and the pedigree analysis could predict the probability of any of the children to be born with a genetic abnormality. In the second part of the unit we discuss one of the techniques in genetics — the recombinant DNA technology that has opened up new vistas for promoting human welfare. The applications to the technology have far reaching effects in the field of agriculture, medicine as well as industry. We shall discuss these aspects under the caption applications of genetic engineering.

For the past fifty years, the science of genetics has been widely applied in the field of agriculture and crop improvement. The development of recombinant DNA technology in the recent years has enabled the geneticists and biotechnologists to engineer plants with desired characters. The transformatin of plants via the *Ti plasmid of Agrobacterium tumefaciens* has been an important breakthrough having, tremendous potential in the area of plant improvement. This methodology has been already tried successfully in some plants. You will study about them in this last unit (i.e., Unit 24). In addition, some other exciting instances of plant improvement are also discussed.

We hope, your study of this course has been enjoyable and fruitful. Do not forget to ...ur feedback on this block too.

Objectives

After studying this block you should be able to:

- discuss the genetics of blood group systems and cluster of haemoglobin genes;
- state the Hardy-Weinberg law and how the gene frequencies in a population could be altered by evolutionary forces such as mutation and selection;
- explain quantitative inheritance with appropriate examples and measure as heritability;
- explain heritability in human beings through twin studies;
- infer the genetic component in behaviour of living beings;
- explain the effect of environmental and genetic component on intelligence quotient (I.Q.);
- define the term genetic counselling and list the diagnostic procedures followed in genetic counselling;
- describe the recombinant DNA technology and enumerate its various applications;
- discuss the applications of the principles of modern genetics in agriculture and plant improvement.

Acknowledgements

We thank Dr P.S. Srivastava, Jamia Hamdard for providing photographs for Figures 24.1 and 24.2.; and Shri Rajbir Singh and Ms Usha Marwah for word processing.

UNIT 19 GENETICS OF BLOOD

19.1 INTRODUCTION

Blood, a "circulating tissue" nurtures the individual cells by supplying nutrients, removing their waste products and regulating their metabolic activities. It consists of red (erythrocytes) white (leukocytes) blood cells and the platelets (thrombocytes) suspended in a liquid medium, called plasma. Blood and plasma perform many functions that are absolutely necessary for the maintenance of health. The red blood cells are intimately associated with heart, lung and kidney functions. Hemoglobin is the main constitutent of erythrocytes that carries oxygen.

Blood group antigens found on the surface of erythrocytes, are responsible for immunological reactions of red blood cells. At least 20 blood groups have been recognized in man, of which ABO system is an extensively known system. The blood group substances are inherited according to simple Mendelian ratios through multiple alleles representing a single locus. About a dozen set of Rh blood types and a few others are inherited through other loci, independent of ABO system. Blood groups are most readily identified by means of specific antibodies present in the serum. Blood group systems have been used for clarifying disputed paternity. The ABO antigens are considered to be important in organ transplantation and blood transfusion.

Diseases of blood and blood forming organs are like a spectrum, at one end of which are diseases which are entirely genetic in origin such as chromosomal abnormalities due to single gene defects, and hemolytic anemia where environmental factors play little role. At the other end of the spectrum are those diseases which are entirely environmental in origin, such as infectious and nutritional deficiences and disorders of white blood cells. In the middle of the spectrum are many common disorders which are partly genetic and partly environmental in origin, called multifactorial disorders; for example most anemias are multifactoral. Since there is at present no effective treatment for most genetic disorders, such diseases can be mainly prevented through genetic counselling.

Objectives

After studying this unit you should be able to

- Explain blood group systems and Rh factor
- Describe hemoglobin genes
- Describe clinical application of blood groups and some associated disorders
- Explain transfusion reactions, and their hazards.

19.2 CHEMICAL STRUCTURE AND GENETICS OF BLOOD GROUP SYSTEMS

After the discovery of circulatory system in 17th century, the discovery of ABO blood group system that led to immunological basis of transfusion reaction was established for the first time in 1900 by Landsteiner, who was awarded Nobel Prize for his work. He pointed out that blood cells have unique cells surface markers, that function as antigens. The presence of specific antigens on the red blood cells are characteristic of different blood groups as shown in Table 19.1

Table 19.1: Established Blood Group Systems

System	Antigen	Number of antigens for which specific antibody is known	Chromosomal location
ABO	A, B, O	5	9(9)
MNS	M, N, S	30	4
P	P	2	?
Rhesus	Rh	44	1
Lutheran	Lu	16	19
Kell		21	?
Lewis	Le	2	19
Duffy	Fy	5	1
Kidd	Jk	3	2
Diego	Di	2	?
Cartwright	Yt	2	?
Auberger	Au	1	?
Sg		1	X
Scianna	Sc	3	1
Dombrock	Do	2	?
Cotton	Co	3	2
Landsteiner-Weiner	LW	3	19

Antigenic determinants of blood group antigens comprise of polysaccharides and proteins. The polysaccharide antigens occur as complex glycoproteins or glycolipid structures consisting of more than one sugar moiety. Sugars constituting the polysaccharides are glucose, galactose, fucose, N-acetylglucosamine, N-acetylgalactosamine and N-acetyl neuraminic acid (sialic acid) which contribute to the charge on cell surface (Table 19.2). The polysaccharide antigens are synthesised by the enzymes **glycosyl transferases**. These polysaccharides are attached to either proteins or lipids directly which are themselves attached to the membrane. The protein antigens of red blood cells are integral proteins that are

inside the hydrophobic lipid bilayers of the membrane. You have read about the membrane proteins in LSE-01.

Table 19.2: Terminal sugars of the antigenic groups for each blood type.

Blood Group A -N- acetylgalactosamine D- Galactose — N–acetylglucosamine | L–Fucose (Fig. 19.1a)
Blood Group O D Galactose — N–acetylglucosamine | (L) Fucose (Fig. 19.1b)
Blood Group B D- Galactose—D- Galactose —Nacetylglucosamine | (L) Fucose (Fig. 19.1c)

19.2.1: ABO Blood Group System:

Landsteiner divided human red blood cells into four groups designated A, B, AB and O. The four phenotypes are determined by three major alleles situated on chromosome 9 (Table 19.3). The I^A andI^B dominant alleles control the A and B blood group antigens and a third recessive I^O allele does not control synthesis of any antigen. Since then, at least twenty other blood groups have been discovered. But our discussion will be limited only to the well known ABO and Rh system. The ABO blood group system depends on the presence or absence of two polysaccharide antigens located on cell membranes of red blood cells. Those individuals with A type of blood possess antigen A on their red blood surface and type B individuals possess antigen B. Persons with type O blood lack both A and B surface antigens. The serum of persons with A type antigen contains antibodies against type B red cells and type B blood persons have antibodies against type A red cells. Persons with type O blood have antibodies against A and B cells. Individuals with AB type serum contain no antibodies against A and B antigens. Thus if the blood of type B is transfused into a person with type A blood, the antigens on type B blood cell will react with antibodies against B and antigen-antibody reaction will take place.

Table: 19.3: The genotypes of ABO Blood groups system

Genotype	Antigen	Phenotype
I^AI^A	A	A
I^AI^O	A	A
I^BI^B	B	B
I^BI^O	B	B
I^BI^B	A, B	AB
I^OI^O	Neither	O

$I^A$and I^B alleles behave dominantly to I^O allele.

Antibodies that agglutinate the cells bearing these antigens are called isohemagglutinins. The IgM antibodies are one of such naturally occurring isohemaglutinins against the red blood cell antigens. The ABO blood groups are presumed to have arisen as a result of immunization by the bacteria in respiratory and gastrointestinal tracts that bear determinants similar to the antigens of ABO groups. Since IgM can not pass through placenta, incompatibility of ABO blood group between fetus and mother does not occur. The ABO blood group antigens are encoded by single gene with three alleles I^A, I^B and I^O. The I^A and I^B alleles are dominant over I^O and codominant with respect to each other.

However, group O individuals have a glycoprotein on the surface of their erythrocytes called the H substance. H substance can be recognised by antisera from different animals. The glycoprotein is not the end product of the gene since it is present on RBC of people who are homozygous for A or B genes. H substance is synthesized just during the synthesis of blood group molecules. Substance H is present in all individuals irrespective of their blood group and no antibodies are produced against it. The sugar moiety contains L-fucose which is recognised by the antibodies. The I^A allele codes for an enzyme that converts the H substance into another glycoprotein i.e. substance A by a adding a terminal N-acetyl galactosamine group that forms the antigenic determinant of blood group A (Fig. 19.1 a).

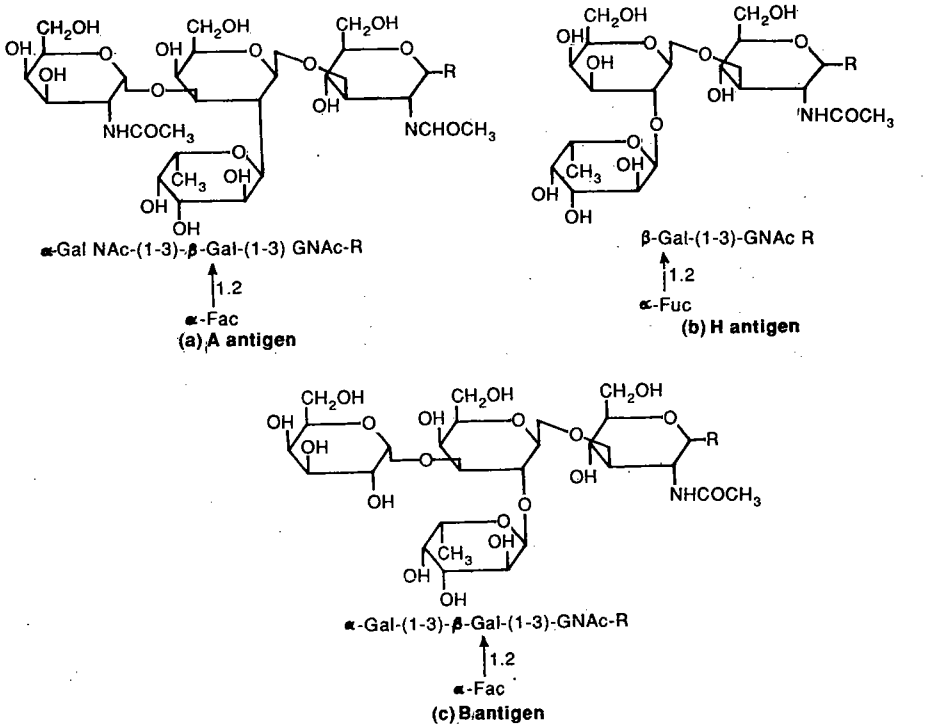


Fig. 19.1 Structure of carbohydrates in ABO Blood group system.

It is now known that the I^A and I^B alleles code for enzymes that attach carbohydrate molecule to the basic glycoprotein called H substance. The carbohydrate group that takes either D galactose or the modified N-acetyl -D-galactosamine at the terminal position is called the H substance and type O individuals (of genotype I^OI^O) possess the H substance on their red blood cell surface.(Fig. 19.1b)

The allele I^A codes for the enzyme N- acetylgalactosaminyl transferase. This enzyme adds -D-N- acetylgalactosamine to the H substance, thereby generating the A type antigenic determinant and A type antigenic structure. The I^B allele codes for the enzyme galactosyltransferase, and converts the H substance similarly into another glycoprotein called the B antigen. Thus the phenotype of a person whose genetic make up is I^AI^O will be group A, but he may pass on the I^O allele to his offspring.

In 1952, a woman in Bombay was found to lack the H substance altogether and she demonstrated a most interesting genetic history and blood type. She was found to lack both A and B antigen and was typed as O blood group. However, as shown in

the partial pedigree (Fig. 19.2) one of her parents was AB and the mother was the obvious donor of I^B allele to two of her offsprings. Thus she was genetically type B but phenotypically type O. She was homozygous for a rare recessive allele h, which prevented her from sysnthesising the complete substance H. The gene responsible for the formation of H-antigen, that serves as a precursor for the A and B antigens was absent in her. As a result the blood group genotypes can not find expression in individuals of hh. To distinguish such types from the rest of the population they were said to demonstrate the Bombay phenomenon. The frequency of the h allele is extremely low, and the vast majority of the human populations is of HH genotype and can synthesize the H substance.

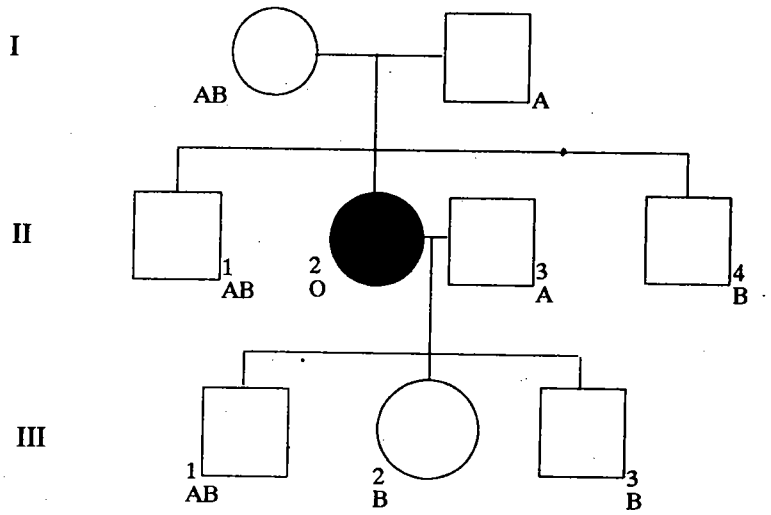


Fig.19.2: Partial pedigree that explains Bombay phenomenon.

Blood group A is further divided into sub groups A_1 and A_2; A_3, and Am on the basis of number of antigenic sites (epitopes) on the cell. The A_1 cells have high antigenic density. Table 19.4 summarises ABO blood group systems.

Table 19.4: Antigens and antibodies of ABO blood group system.

Group	Specific antigen	Quantity of H-antigen	Antibodies in serum	May receive blood from	May give blood to
O	None	High	Anti-A Anti-B	O	O,A,B,AB
A	A, A_1	Low	Anti-B	A,O	A, A_2, AB
A_2	A_2	More than A	Anti-B Anti-A_1	A (usually) A_2O (occasionally)	A, A_2, AB
B	B	Like A	Anti-A	B,O	B,AB
AB	A, A_1, B	Less than A or B	None	AB,A,B,O	AB
A_2B	A_2,B	More than AB	None	AB,A,B,O	AB

SAQ 1:

a) Fill in the blanks with appropriate words.

1). ABO blood group systems were discovered by ---------- in the year 1904.

2). The polysaccharide antigens are generated by ----------.

3). ---------- does not code for any enzyme in the ABO system.

4). Bombay phenomen is a unique case where mother's blood lacked antigens and two of her offsprings belonged to B blood group; genetically her blood group was ---------- but phenotypically ----------

b) Match the terminal sugars with the blood groups in column-I

	Sugars		Column-I
1)	D-galactose-(L)-fucose	(i)	Blood group B
2)	Nactyl glactosamino — D-galactose -L-fucose	(ii)	Blood group A
3)	D-galactose-D galacrose-(L)-fucose	(iii)	Blood group O

19.2.2 The Rh System

The Rh system was discovered by K. Landsteiner and A.S.Weiner from rabbits immunized with the blood of monkey *Macaca rhesus*. This blood group antigen is also found on the surface of human erythrocytes. About 85% of the population possesses Rh antigen on the surface of their red blood cells and are called Rh positive persons. Those individuals who do not possess Rh factor on their red blood cells are called Rh negative. The plasma of Rh negative persons does not contain antibodies (agglutinins) against Rh positive factor but such persons can produce antibodies if blood of Rh positive persons is transferred to them.

The Rh antigens are proteins and are of common occurrence in humans and so preformed antibodies are rare. They are the product of immunization either through pregnancy or by transfusion. The original antigen (D) now designated as Rh^o is present in about 85% of whites which means that they have antigen D on the surface of their red blood cells. In the remaining 15% Rh negative individuals who have no antigen D, will produce antibodies against that antigen when they are exposed to Rh positive blood. At least eight different kinds of Rh antigens each referred to as an Rh factor is present in Rh system. Only three genes residing at three separate but closely linked loci regulate the synthesis of Rh antigen. These exist as the allelic pair Cc, Dd and Ee. By far the most important of these is the allele coding for antigen D.

19.2.3 Secretors and Lewis Blood Group System

The ABO blood group substances are present as glycoproteins on the surface of erythrocytes and on the surface of many endothelial and epithelial cells. The genetically controlled A, B and H substances are present in about 80% humans as mucopolysaccharides in body secrations such as saliva, sweat, urine, seminal fluid and gastric juices. Such prople are termed as secretors. The secreted substances are immunologically identical to those present on their red blood cells. The inherited characteristics are controlled by allelic gene pairs denoted by *Sc* and *sc*. Both homozygous (*ScSc*) and heterozygous (*Scsc*) are secretors and sc gene is recessive; non-secretors possess *scsc* alleles.

In the Lewis (Le) system, there are two substances Lewis a(Le^a) and Lewis b(Le^b). These antigens are complex glycoproteins or glycolipds that are found free in serum and have a natural ability to be adsorbed on RBC surfaces. The Lewis antigen is produced from the same precursor as those of ABO(H) antigens. The H substance is a key intermediate in the pathway to the A,B and Le^b antigens synthesis. The Le^a gene activates fucosyl transferase which adds fucosyl residue to the precursor moiety and produces a Le antigen. The precursor substance is operated upon by the H gene which controls the fucosyl transferase. The Le^a and H genes are structural genes for transferase that have slightly different functions. To generate b(Le^b) antigen, a third type of fucosyl transferase adds a fucose molecule to precursor

substance. Thus, the structure of ABO(H) and Lewis antigens are closely related to one another.

19.2.4 MN Blood Groups

The MN System under the control of a locus on chromosome 4, consists of three blood groups M,N and MN phenotypes. Genotype L^ML^M represents blood group M. L^ML^N blood group MN and L^NL^N blood group N (see table 19.5). (L represents the name of the discoverer Lewis)

Table 19.5. Inheritance of MN blood groups.

Parental Phenotypes	Offspring Phenotypes
M × M	All M
N × N	All N
M × N	All MN
M × MN	1/2 M: 1/2 MN
N × MN	1/2 N: 1/2 MN
MN × MN	1/4 M: 1/2 MN: 1/4 N.

The M type elicits antibodies (anti M.serum) for M, which could agglutinate M antigens, while N red blood cells caused the production of antibodies specific for N (anti N serum). Both type of antibodies however could agglutinate the MN red blood cells. It is known now that MN system is inherited as a result of two alleles of a gene.

19.2.5 The Kelly & Duffy Systems

The Kelly and Duffy blood groups are minor blood groups, but can cause haemolysis in transfusion reactions. The kelly system consist of two allelic forms as K and k. The K antigens are found on the RBCs of about 10% of the Georgian population and are highly immunogenic. Exposure of K antigen during pregnancy or transfusion may lead to the formation of anti-K IgG which causes agglntination reactions. There are a number of other antigens belonging to Kell system (Kpa, Kpb, Jsa, Jsb, Uia) but antibodies other than anti-K are rare. About two thirds of the white population have the antigen Fya in Duffy system. Compared to other antigens in this system, anti-Fya is a relatively common cause of hemolytic transfusion reactions. The allelic form was named as Fyb . Approximately 60% of all black population lacks Fy antigens and lacks receptor for the malarial parasite Plasmodium vivax, hence resistant to P.vivax malaria. In another minor system called Kidd (JK) system two antigens have been described - JKa and JKb . The antibodies to these antigens are particularly unstable on storage and do not remain in the serum of sensitised patients.

SAQ 2:

1 a) Tick (√) the correct statement.

An individual is said to be a secretor when

a) soluble form of blood group antigens are found in body secretions.

b) insoluble form of blood group antigens are found in blood.

c) soluble form of blood group antigens are found in saliva.

d) soluble form of blood group antigens are found in urine.

b) Briefly describe the following.

Kelly System:

.....

.....

.....

.....

.....

Duffy System:

.....

.....

.....

.....

.....

19.3 HEMOGLOBIN GENE IN EUKARYOTES

Human hemoglobins are conjugated proteins in which a prosthetic group, heme, is attached to each of four polypeptide subunits. Adult hemoglobin (HbA) consists of two α and two β polypeptide chains The peptide chain of hemoglobin is responsible for the species specificity. The pigmentary property and respiratory functions are associated with heme, the iron containing pigment, but the globin fraction of hemoglobin functions to transport carbon dioxide. Hemoglobin has the property of combining reversibly with atmospheric oxygen, forming oxyhemoglobin. This occurs in the capillaries surrounding the alveoli of the lungs. The oxygen is then transported by the arterial blood to tissues and the blood depleted of its oxygen returns to lungs for oxygenation.

19.3.1 Chemistry of hemoglobin

As you know that in a complete hemoglobin molecule there are four individual polypeptide chains divided into two identical alpha (α) chains and two identical beta (β) chains. The alpha and beta chains are almost of the same chain length and of similar structural conformations but different in chemical and electrophoretic properties.

The four polypeptide chains of mammalian hemoglobin are individually linked with one heme group. Hemoglobin exists in several forms, made up of various combinations of six different peptide chains, designated as α, β, gamma (γ), delta (δ) and zeta (ζ) and epsilon (ε).

About 90% of the normal adult hemoglobin (HbA) consists of two α and two β chains and is designated as HbA. The subunits of various human hemoglobins are given in Table 19.6.

Table 19.6: Subunit composition of human hemoglobins.

Embryonic hemoglobin	**Subunits**
Hb Gower 1	$\zeta_2 \varepsilon_2$
Hb Gower 2	$\alpha_2 \varepsilon_2$
Hb Portland	$\zeta_2\varepsilon_2$
Fetal hemoglobins	
HbF	$\alpha_2\gamma_2$
Adult hemoglobins	
HbA	$\alpha_2 \beta_2$
HbA_2	$\alpha_2 \delta_2$ (a minor variant of adult-hemoglobin).

19.3.2 Hemoglobin gene Clusters

Different hemoglobins are synthesized in different stages of development: embryonic hemoglobins are followed by fetal hemoglobin present in the developing foetus and is normally displaced by adult haemoglobin within six months after birth.

Each of these hemoglobins consists of two alpha type (α) and two beta (β) type of chains. The α type genes are clustered on chromosome 16 in human and the β-type genes on chromosome 11 (Figure 19.3).

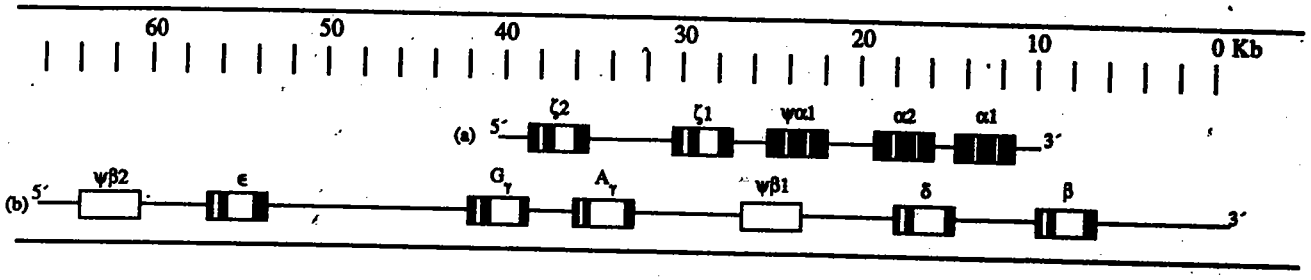


Fig. 19.3: Gene clusters of different chains of haemoglobin.

In the α - cluster, the gene for embryonic zeta(ζ) chain appears before two genes for α-chains, which are components of both fetal and adult hemoglobins. Both the α and β gene clusters contain pseudogenes represented by symbol psi (ψ). These sequences are homologous to adjoining genes but they do not code for functional products.

In the β - cluster, the gene for the embryonic epsilon (∈) chain is followed by two genes for fetal β chains, and then by the genes for the adult chains. Thus, the sequence of the human globin genes matches the order in which they are expressed during development.

A single pair of genes code for α chains and another pair for β chains. There are at least 2 different copies of the genes. One codes for a chain with glycine at position 136 while the other codes for a chain having alanine in the same position.

As is true of various other proteins, all the four hemoglobin polypeptide chains are synthesized separately prior to their union as tetramer. Generally, proteins formed from two or more polypeptide chains are controlled by adjacent genes. This is not the case with hemoglobins, since the two genes are not even on the same chromosomes. As noted earlier, the duplicate pair of alpha genes is on chromosome 16, and the beta gene cluster is on chromosome 11.

SAQ 3:

i) State whether the following statements are True (T) or False (F).

a) Human hemoglobins are proteins with four subunits. [ ]

b) Heme contributes to the pigmentary property and performs chief respiratory functions. [ ]

c) The α - type genes are clustered on chromosome 16 and β -type gene on chromosome 10. [ ]

d) The sequence of the human globin genes does not match the order in which they are expressed during development. [ ]

ii) Match the items in column I with those in column II

	Column I		Column II
a)	HbA	i)	$\alpha_2\gamma_2$
b)	HbA_2	ii)	$\zeta_2\varepsilon_2$
c)	Hb portland	iii)	$\alpha_2\beta_2$
d)	HbF	iv)	$\alpha_2\delta_2$

iii) Hemoglobin exists in several forms made out of six different peptide chains. The six different types of polypephide chains that go to form different forms of haemoglobin are:

19.4 BLOOD GROUPS AND MEDICINE

There are many applications of blood groups in medicine. Their involvement in blood transfusion and organ/tissue transplantation is well documented. Blood group markers in conjunction with other parameters are being emplolyed in paternity disputes. The complications of pregnancy and blood group-related disorders are some of the other examples of the role of blood groups in clinical sciences.

19.4.1 Blood transfusion

The discovery of blood types (A, B and O) and Rh factor along with many advances in immunohematology provided the immunological understanding required for modern and safe transfusion of blood. Parallel advances in the techniques of drawing, storage and administration of blood have also helped in the transfusion process.

Ideally a recipient should be transfused with exactly the same type of blood as his own. Table 19.4 shows the interchangeability of the blood groups in transfusion indicating clearly that group O persons are universal donors due to the absence of A and B antigens on their erythrocyte surface which could react with anti A or anti B antibodies in the serum of the recipient. Since both the A and B antigens are present as surface antigens and no antibody against A or B in their serum, the group AB persons are the universal recipients for blood group A, B, O. Similarly a person with blood type B could receive blood from O or B individual but would react to A or AB donor type blood as a consequence of circulating anti-A antibodies. Table 19.4 shows all the possible donor recipient relationships in blood transfusion reactions for ABO antibody system. One of the most important factors for transfusion therapy is choosing the right material. The development of blood banks has made available the stored whole blood. Its usefulness is limited because of changes that occur on storage. Most blood banks set a limit of 21 days storage for RBC, after which blood is considered unsuitable for transfusion. Leucocytes start

disintegrating even earlier and show very little bactericidal activity by the end of 4th day. Similar is the case with blood platelets unless special precaution is taken in drawing blood. If platelets are required it is best to administer blood within 24 hours after it is withdrawn.

Fresh whole blood is generally used when the recipient needs RBC in addition to platelets, leucocytes or labile coagulation factors. Plasma is used when the patient does not require any of the formed elements. Stored plasma is most useful in the treatment of shock and hypoproteinemia.

Recently plasma extenders and substitutes like dextran and methyl cellulose have been tried but each has its disadvantages. Also autologous (predonated) transfusion, the process of returning to a person his own blood when needed has been practiced. This has many advantages and eliminates transfusion hazards and possibility of transmitting disease.

Transfusion of blood and blood products represents one of the major advances in medicine. In spite of tremendous progress, a small but significant number of side reactions can be expected in transfusion. If a person whose blood belongs to one of the groups receives a blood transfusion from a donor of another group a hemolytic transfusion reaction can occur. This is due to the serum of the recipient that may agglutinate the cells of the donor or vice versa, which results in hemolysis of erythrocytes and liberation of free hemoglobin from the lysed red cells. Liberation of hemoglobin may cause secondary complications such as jaundice, fever and kidney function impairment. The most important causes of death related to blood transfusion before the advent of AIDS were hemolysis and hepatitis.

SAQ 4

a) Tick mark the correct answer from the alternatives provided:

(1) The Universal blood donor is

i) blood group A,

ii) blood group B,

iii) blood group AB,

iv) blood group Rh,

v) blood group O.

(2) The universal recipient is

i) AB blood group

ii) A blood group

iii) O blood group

iv) B blood group

b) Maternal antibodies capable of crossing the placenta are: (Indicate by (√) mark)

a)	IgA	[ ]
b)	IgG	[ ]
c)	IgD	[ ]
d)	IgE	[ ]
e)	IgM	[ ]

c) Fill in the blanks with appropriate words.

1) The Rh locus is located on ----------.

2) Isoantibody present in type O individuals are predominantly ----------.

3) Group B and O persons contain ---------- antibodies in their sera.

d) Tick mark the correct answers:

Fresh blood is generally used to suply,

1) White blood cells

2) Red blood cells

3) Platelets

4) Complements

19.4.2 Acquired Immune Deficiency Syndrome (AIDS)

Acquired immune deficiency syndorme or AIDS was first recognised in 1981 as a deadly disease and is spreading throughout the world at an alarming rate. The causative agent is human immunodeficiency virus (HTLV-III) that rapidly infects helper T cells resulting in irreversible defects in immune response. You may recall that helper T-cells stimulate the proliferation of B and T cells and as well as macrophages. Anti-HTLV III antibodies can be effectively measured by ELISA technique and hence screening of blood donors can be done. It is important that higher risk population groups (homo/bisexual men and drug addicts) should refrain from blood donation even if blood tested is anti- HTLV-III negative. The virus may circulate in blood in the absence of antibodies. Testing of blood before transfusion for other disease like syphilis, cytomegalovirus (CMV), malaria and other such infections have also been recommended.

19.4.3: Paternity Exclusion

A blood group antigen cannot be inherited by the child unless it is present in either of the parents. This is the basis of blood group tests in paternity disputes. The genotypes and phenotypes of the ABO system are given in Table.

Table 19.7: Inheritance of the ABO blood groups

Allele from one parent	Allele from other Parent	Genotypoe	Phenotype
I^A	I^B	I^AI^B	AB
I^A	I^A	I^AI^A	A
I^A	I^O	I^AI^O	A
I^B	I^B	I^BI^B	B
I^B	I^O	I^BI^O	B
I^O	I^O	I^OI^O	O

From the above table it is evident that an AB man cannot father a group O child. Similarly A group parents can not produce a B group child. An AB woman and an A_1 man can not give beirth to a A_2 child. Also O-Rh negative parents can not produce O-Rh positive chid. Testing of blood group can only exclude a putative father from fatherhood. This is now generally accepted in courts of law. Testing of

polymorphic antigens, such as HLA antigens is considered usually superior to blood group antigen in paternity disputes.

19.4.4: Maternal - foetal incompatibility and its prevention

Hemolytic disease of newborn (hemolytic anemia) results from the mother's antibodies against fetal red blood cells. Formation of Rh antibodies and antibodies to blood groups A and B may also cause hemolysis of fetal cells due to the maternal antibodies that are of IgG class and thus capable of crossing the placenta. You have read in Unit 18 of this course about the nature and functions of antibodies. In fact ABO immunization during pregnancy occurs more often than Rh immunization, but it seldom results in serious problems. This is because the maternal antibodies are neutralized before they can cause damage to red cells as A and B substances are present in the tissues of the mother including the placental endothelium.

Several types of maternal–fetal blood incompatibilities are known. Among them Rh incompatibility is the most important. If the mother is Rh negative and the father of the fetus is Rh positive, the child may also be positive having inherited the D-antigen from father. Ordinarily there is no mixing of maternal and fetal blood molecules and no exchange between the two circulatory systems by the placenta. However, late in pregnancy or during the birth at the time of parturition, a small quantity of blood from the fetus may pass through placenta. The fetal RBCs which bear antigen D sensitize the mother's WBCs inducing them to form antibodies against antigen D. When this mother becomes pregnant again, the anti Rh bodies in her may cross the placenta and enter the fetal blood causing the cells to clump together. In the extreme case of this situation, known as **erythroblastosis foetalis** the fetal red blood cells are destroyed and the fetus dies before birth.

The Rh locus is situated on chromosome 1. Rh antigenic determinants may be dependent on interaction between RBC membrane proteins and phospholipid molecule. Rh antigens are controlled by three closely linked allelic pairs of gene, which produce the antigenic determinants C or c, D or d, E or e respectively. The most important is the D atnigen, and RBCs possessing this antigen are Rh positive. Individuals with DD or Dd are Rh positive, whereas dd are Rh negative.

The risk of initial sensitization of Rh negative mothers has been reduced from 10-20% to less than 1% by intramuscular injection of human anti-D globulins within 72 hours of delivery or abortion. This destroys any Rh positive cells that have entered the mother's circulation, well before her own white blood cells could be sensetized. Also the antibodies that have been introduced are also soon eliminated from her body. As a result, when she becomes pregnant again her blood will not contain the anti-D that could harm her baby.

Hemolysis associated with ABO incompatibility is similar to Rh diseases wherein the maternal antibody enters fetal circulation and reacts with A or B antigen on erythrocyte surface. In type A and B individuals, naturally occurring anti-B and anti-A isoantibody are largely IgM, that do not cross placenta. In contrast, isoantibody present in type O individuals are predominantally IgG type and for this reason, ABO incompatibility is largely limited to type O mothers with type A or B fetus. The presence of IgG anti-A or anti-B in type O mothers frequently explains why hemolysis due to ABO incompatibility occurs in first pregnancy without prior sensitization. Incompatibility may also occur due to other minor blood groups.

SAQ 5:

a) Tick (√) the most appropriate answer.

i) Hemolytic disease of the new born occurs when

a) father is Rh-negative and mother is Rh-positive [ ]

b) both the parents are Rh-positive [ ]

c) both the parents are Rh-negative [ ]

d) mother is Rh-negative and father is Rh-positive [ ]

ii) Tick the most appropriate statement

i) The causative agent of AIDS is

a) HLV antigens [ ]

b) HTLV-III virus [ ]

c) EBV [ ]

d) Herpes simplex virus [ ]

b) Four babies were born in nursing home at one time. They had the blood groups B, AB, O and A. The four parents were B and B, AB and O, O and O,A and B. Assign the four babies to their correct parents.

	Babies		Parents	
1)	B	(A)	AB and O	( )
2)	AB	(B)	B and B	( )
3)	O	(C)	A and B	( )
4)	A	(D)	O and O.	( )

c) Which of the following statements relating to newborn hemolytic disease is/are irrelevant.

1) Injection of anti Rh globulins to an Rh mother soon after delivery of an Rh baby can suppress the formation of anti Rh globulins by the mother.

2) If mother is Rh^+ and newborn is Rh^-, the child becomes tolerant to Rh antigen.

3) The mother forms antibodies against Rh antigen of the foetus if D-antigen is not administered with in 72 hours of the birth of the first child.

19.4.5: Blood groups and diseases

Many years of extensive investigations to ascertain whether a person with a particular blood group is more likely to develop a particular disease has indicated some relationship between blood groups and diseases. The evidence have not been conclusive in view of the multifactorial origin of human ailments. The relationship could be summarized as follows:

1. Group A persons are more likely than those of group B or O to develop carcinoma of the stomach or pernicious anemia. Group A persons are more liable to thromboembolic disease. Such persons tend to have higher levels of anti-haemophilic globulin than others. Anti-A_1 in the serum of A_{2a} or A_{2b} persons can cause haemolytic transfusion reactions.

2. Group B or O persons whose sera normally contain anti-A antibodies would tend to have milder small pox than A or AB persons.

3. Non-secretors of group O are more likely to develop duodenal ulcer with increased liability to haemorrhage.

4. An association has been observed between rare Kell groups and some patients with chronic granulomatosus disease (CGD). It is an inherited X-linked defect in neutrophil function, in which there is a higher susceptibility to infection even by a low grade pathogen. Blood transfusion in CGD patients is a potential hazard since Kell system antibodies react with red cells of almost everyone else. Anti-Kidd antibodies may cause immediate or delayed haemolytic transfusion reactions or hemolytic disease of the newborn.

5. A possible relationship between Duffy groups and malaria has been observed. Persons of group Fy(a^-, b^-) appear to be resistant to *Plasmodium vivax* malaria. This accounts for evaluation of higher incidence of Fy ($a^- b^-$) in West Africa and resistance of its people to *P. vivax* malaria.

19.5 RACIAL DIFFERENCES

The incidence of blood group antigens varies from one race to another. Most of the surveys have been conducted on ABO system. Their distribution in some selected populations is shown in Table 19.8.

Table 19.8: Distribution of ABO blood group in selected populations.

Population	Blood group			
	O	A	B	AB
U.K	47	42	8	3
European Gypsies	31	27	35	7
Asiatic Indians	33	24	34	9
Japanese	30	39	22	9
Polynesians	40	56	3	1
Some South American Tribes	100	0	0	0

European gypsies have a blood distribution similar to that of Asiatic Indians, from whom they appear to have originated. Negroid and Mongoloid races have a very high incidence of Rh-positives, whereas Basques have an unusually higher incidence of Rh-negatives. At least three theories have been put forward to explain the racial differences in blood group distribution. These are:

1. Originally there are three human races of groups A, B and O. The present distribution is a consequence of migration and intermarriage.
2. Blood groups arose by mutation from one group, probably O. The exclusive O group of many south American tribes support this hypothesis. The origin of A and B gene has probably been from Europe and Asia respectively.
3. It is a reversal of the second hypothesis and speculates that the original blood group was AB and that A, B and eventually O arose from mutations.

SAQ 6:

i) Fill in the blanks:

1) Population having highest incidence of blood group O is --- ----------.

2) Origin of A and B genes is probably from --------- and --- ----------.

3) Blood groups A, B and O probably arose from ----------.

4) Negroid and Mongoloid races have a high incidence of -- ----------.

ii) Fill in the blanks:

1) Persons with blood groups ------- are more likely to suffer from carcinoma of stomach, pernicious anemia and thromboembolic diseases.

2) Persons of group Fy (a^-, b^-) are resistant to --- ----------malaria.

3) Non-secretors of group O are more likely to develop --- ----------.

4) Group B and O persons suffer mildly from small-pox due to ----------.

19.6 SUMMARY

- Blood consists of solid elements suspended in the plasma. Hemoglobin is the main functioning constituent of the erythrocytes. A group of antigenic substances responsible for major immunological reactions are found on the surface of the erythrocytes. There are at least twenty established blood group systems in man, the so called types.
- Genes coding for hemoglobin α and β chains are clustered on chromosomes 16 and 11 respectively.
- The clinical significance of a blood group system depends on two factors: the frequency of antibodies in the population and their relative potency. The antigens of the blood cells and their antibodies are important in transfusion of blood. Great care is to be taken to be certain that antibodies are not present in the serum of the recipient that might react with the antigens on the red cells of the donors, which would result in the destruction of the transfused cells.
- The emergence of AIDS as a dreadful disease in transfusion of blood, blood components and coagulation factor emphasizes that the infectious diseases are still the main cause of transfusion complications. Hemolysis and transfusion hepatitis were the most important causes of deaths related to blood transfusion before AIDS was discovered.
- The inheritance of a blood group by a child occurs only if it is present in either of the parents and that is the basis of blood group tests in paternity disputes generally accepted in courts of law. Hemolytic disease of the new born may be due to incompatibility of Rh, ABO and other minor blood groups.
- The most common cause of hemolysis of new born is transplacental transfer of maternal antibodies that destory fetal erythrocytes. Surveys conducted on the ABO system reveal that the distribution of blood group antigens varies from one race to another.

19.7 TERMINAL QUESTIONS

1. What is hemoglobin and what are its functions?
2. On which chromosome are the hemoglobin genes located?
3. What is autologous transfusion?
4. How will you detect severe hemolysis?
5. What will be the genotype and phenotype of a child with one parent having I^AI^A genotype and the other I^BI^B genotype?
6. What is ABO incompatibility?
7. What are the theories put forward to explain the racial differences in blood group distribution?

19.8 ANSWERS

SAQ 1:

a) 1) Landsteiner

2) Glycosyl-transferases

3) O allele

4) O group

b) (1) -iii; (2) - ii; (3) - i;

SAQ 2:

a) a.

b) i) Kelly system is one of the minor blood groups, sometime associated with transfusion risk and consists of two allelic forms K and k.

ii) Duffy System: Anti-Fy^a is a relatively of common cause of hemolytic transfusion reactions. The allelic form was named as Fy^b. Approximately 60% of West African population lacks Fy(a ,b) and receptor for malarial parasite *Plasmodium vivax*, hence resistant to *P. vivax* malaria.

SAQ 3:

i) a-F, b-T, c-F, d-F.

ii) a-iii, b-iv, c-ii, d-i

iii) $\alpha, \beta, \gamma, \partial, \varepsilon, \zeta$

SAQ 4:

a) 1) blood group O, 2) AB blood group.

b) IgG; c- (1), chromosome 1, (2). Anti-A

3) anti A and anti AB

d) 2 and 3

SAQ 5:

a) (i) d (ii) b

b) 1-B, 2-C, 3-D, 4-A.

c) 2.

SAQ 6:

i) 1- South American tribes, 2- Europe and Asia, 3 4- Rh^+ groups.

ii) 1-A, 2- *P.vivax*, 3-duodenal ulcer and haemorrhage 4- Antibody A.

Answers to Terminal Questions:

1. Hemoglobins are conjugated proteins in which a prosthetic group, heme, is attached to each of four subunits. The peptide chain of hemoglobin is responsible for the species specificity of the hemoglobin. Hence, the iron-containing pigment performs the chief respiratory functions and exhibits pigmentary property, but the globin fraction plays a role in carrying carbon dioxide.

2. Hemoglobin consists of two α- type chains and two β- type chains. The α-type genes are clustered on chromosome 16 and the β- type genes on chromosome 11.

3. It is the process of returning to a person his own blood when needed and is also known as predonated transfusion.

4. In severe transfusion reaction, billirubin, the breakdown product of heme, appears in the plasma, a few hours after the reaction. In acute hemolysis, hemoglobin is bound to albumin to make methemoglobin, which is brown in colour. The hemolysis could be detected by the colour of plasma and urine, determination of hemoglobin in plasma and urine and quantitation of billirubin, LDH and methemoglobin in plasma.

5.

Genotype	Phenotype
I^AI^B	AB

6. In type A and B individuals, naturally occurring anti-A and anti-B isoantibody are largely IgM, that does not cross placenta. Whereas, type O individuals predominately have IgG molecule. For this reason, ABO incompatibility is largely limited to type O mothers with type A or B fetus. The presence of IgG anti-A or anti-B in type O mothers causes hemolysis in first pregnancy without prior sensitization.

7. Theories are:

   1. The present distribution is the consequence of migration and intermarriage of the original human races of groups A,B and O.
   2. Mutation from one group, probably O, gave rise to other blood groups.
   3. The original blood group was AB and that A,B and O arose from mutation.

UNIT 20 BEHAVIOUR OF GENES IN POPULATIONS

Structure **Page No.**

20.1 INTRODUCTION

In the previous units of this course, you have studied about two major subdisciplines of science of Genetics–the transmission and molecular genetics. You are now aware that transmission genetics deals with the genetic processes that occur among individuals and the unit of study in transmission genetics is the individual. Essentially the study is concerned with the transmission of genetically controlled characters from individuals of one generation to another. Molecular genetics, the other subdiscipline deals with the chemical nature of hereditary material. Also molecular genetics explains the mechanisms of encoding of genetic information in DNA molecule as well as the transformation of genetic information into relevant phenotype within the cell. Naturally the unit of study of molecular genetics is the cell.

In this unit, we are to discuss about the third subdiscipline of genetics namely population genetics. The study deals with the genetics of groups of individuals or populations and how the genetically controlled variations evolve into adaptations. The unit of study is, therefore, a Mendelian population which means a group of interbreeding individuals who share a common gene pool.

Gene pool : The genes shared by group of interbreeding individuals or a Mendelian population.

The central theorem of population genetics was independently conceived by Godfrey H. Hardy, a British Mathematician and Wilhelm Weinberg, a German Physician. In this unit, we shall state and discuss Hardy-Weinberg law and that the behaviour of genes in a population does not generally conform to the specification of the law. In other words, behaviour of genes in populations is concerned with changes in the frequencies of genes in the population. Such changes in the frequencies of genes form the genetic basis of evolution. For this reason, population genetics is also referred to as evolutionary genetics.

Objectives

After studying this unit, you should be able to:

- state the Hardy-Weinberg law and justify the relationship between allelic and genotypic frequencies by applying the binomial equation.

- show that both allelic and genotypic frequencies are maintained constant in the absence of outside evolutionary forces such as mutation, selection, genetic drift and gene flow
- demonstrate that only if mutations were alone to occur in a population, allelic frequencies would change over a period of time and conclude that with both forward and reverse mutations occuring, an equilibrium in gene frequencies is reached over a period of time.
- define the terms adaptive value and selection coefficient and explain that natural selection can either reduce or maintain genetic variations depending on the extent of adaptation of a trait to the environment
- describe that genetic drift results in random changes in the frequency of genes in small populations and that it may also result in loss of genetic variation.
- discuss that the gene flow can either introduce new alleles into a population or modify the frequency of existing alleles.

20.2 HARDY WEINBERG LAW

Hardy-Weinberg law states that in a population, the frequencies of a pair of alleles remain constant generation after generation provided that

i) the population is infinitely large

ii) the population is randomly mating one

iii) there is no mutation, genetic recombination, gene migration and selection occurring in the population (in other words, the population is free from external evolutionary forces acting on it).

The term frequency used in this unit refers to the number of times an allele or a genetype occurs in a population.

The conditions (i), (ii) and (iii) listed above have to be satisfied to a large extent if the gene frequencies are to be maintained in an equilibrium in a population for several generations.

The first condition is that populations must be infinitely large. If populations were to be small then the gene frequencies tend to drift every generation due to chance deviation or sampling error (refer to Unit 13 of LSE-07 course). The word 'drift' would refer to large and sudden changes in gene frequencies in a population. You may say that genetic drift varies directly with the size of the population. That is smaller the population, greater will be the genetic drift. We may also mention here although in small populations drift is quite often common, the populations need not be "infinitely large" for Hardy-Weinberg law to be true.

The second condition requires that matings among the individuals of the population must be random. By random mating it is meant that any genotype should be able to mate with another genotype and there should not be any restriction in mating between genotypes. Also it means that matings occur between genotypes in proportion to the fequencies of genotypes in population. Stated differently, random mating means that the probability of mating between two genotypes is equal to the product of the two genotype frequencies. For instance, let us say that there are three genotypes AA, Aa and aa in a population occurring in a frequency of 0.36, 0.48 and 0.16 respectively. This means that assuming there are 1000 individuals in a population, 360 of them are AA, 480 of them are Aa and 160 of them are aa. The probability of mating occurring between genotypes AA and Aa should then be $0.36 \times 0.48 = 0.1728$ and that of AA and aa should be equal to $0.36 \times 0.16 = 0.0576$. You could observe from these examples that the probabilities of other possible matings are equal to the products of the genotypic frequencies when mating is random.

Hardy-Weinberg law also requires that populations must be free from outside evolutionary forces. Hardy-Weinberg law aims at defining the role of heredity alone

in changing the gene frequencies and the influence of reproduction on genotypic frequencies. Therefore, while applying Hardy-Weinberg law, we assume that there is no influence of other evolutionary forces such as mutation, natural selection, genetic recombination and gene migration. Later in this unit we shall discuss the role of other evolutionary forces in bringing about changes in gene frequencies. Before we proceed further attempt the following SAQ.

SAQ 1

Fill in the blanks with suitable words.

i) For Hardy-Weinberg equilibrium to be true, the population has to be ....

ii) .... and .... are certain factors which influence gene frequencies in a population.

iii) A group of interbreeding individuals who share a common gene pool is called ....

iv) Hardy-Weinberg law defines role of ..... in changing gene frequencies and the influence of ..... on genotypic frequencies.

v) Assuming the genotype BB has a frequency of 0.35 in a population, that bb is 0.25 and Bb is 0.4, the probability of the possible mating occurring between genotypes Bb and bb is equal to ..... .

20.2.1 Determination of Allelic and Genotypic Frequencies

In the previous section we stated the Hardy-Weinberg law. We shall verify the law based on certain arithematic calculations. Prior to that we shall learn to determine the frequencies of alleles and genotypes in a population. One can represent Hardy-Weinberg formulation by the simple algebric expression $p^2 + 2pq + r^2$, which is essentially the expansion of the binomial expression $(p + q)^2$. The two expressions can be effectively used to determine the frequencies of alleles of an ideal population. By ideal population we mean that it satisfies all the conditions required to maintain the frequencies of alleles at equilibrium generation after generation. To illustrate this idea let us consider a pair of alleles A and a, A being dominant to a. When the genotype AA mates with genotype aa, a heterozygous population of Aa is initially formed. Assuming the population is randomly mating, it will have three genotype classes AA, Aa and aa. Let us assume that the population consists of 1000 individuals 360 belonging to genotype AA, 480 to Aa and the remaining 160 to aa. Let us assume that frequency of A is p and a is q. Therefore, the frequency of genotypes AA, Aa and aa in the population is

$$\text{frequency of AA} = p^2 = 360/1000 = 36\% = 0.36$$

$$\text{frequency of Aa} = 2pq = 480/1000 = 48\% = 0.48$$

$$\text{frequency of aa} = q^2 = 160/1000 = 16\% = 0.16$$

Based on the above data it is possible to obtain the frequency of alleles A and a in the population.

$$\text{frequency of a} = q = \sqrt{\text{frequency of aa}} = \sqrt{0.16} = 0.4$$

$$\text{and the frequency of A} = p = (1 - q) = 1 - 0.4 = 0.6$$

frequency of A = 0.6

frequency of a = 0.4

It should be possible to arrive at genotypic frequencies in a population based on allelic frequencies.

$$\text{frequency of A} = p = 0.6$$
$$\text{frequency of a} = q = 0.4$$
$$(p+q) = 1$$
$$(p+q)^2 = (0.6+0.4)^2 = 1$$
$$p^2 + 2pq + q^2 = (0.6)^2 + 2(0.6)(0.4) + (0.4)^2 \text{ where}$$
$$p^2 = \text{frequency of AA,}$$
$$2pq = \text{frequency of Aa and}$$
$$q^2 = \text{frequency of aa}$$
$$= 0.36 + 0.48 + 0.16 = 1$$
$$\text{AA} \quad \text{Aa} \quad \text{aa}$$

You may observe here that the total allelic frequency $(p+q) = 1$ and the total genotypic frequency, $(p+q)^2$, is also equal to 1.

20.2.2 Verification of Hardy-Weinberg Law

In the last subsection, you have learnt to calculate the frequencies of alleles and genotypes, when the number of individuals under each category of genotypes is known. Let us use these frequencies to verify the Hardy-Weinberg law. The law requires that the population must be randomly mating and that there are no outside evolutionary forces acting on the population. The three genotypes (AA, Aa and aa) when they are randomly mating can form nine mating types.

AA × AA
AA × Aa
AA × aa
Aa × AA
Aa × Aa
Aa × aa
aa × AA
aa × Aa
aa × aa

We earlier said that the probability of each mating occurring is the product of the frequencies of mating genotypes.

Thus

AA × AA = 0.36 × 0.36 = 0.1296
AA × Aa = 0.36 × 0.48 = 0.1728
AA × aa = 0.36 × 0.16 = 0.0576
Aa × AA = 0.48 × 0.36 = 0.1728
Aa × Aa = 0.48 × 0.48 = 0.2304
Aa × aa = 0.48 × 0.16 = 0.0768
aa × AA = 0.16 × 0.36 = 0.0576
aa × Aa = 0.16 × 0.48 = 0.0768
aa × aa = 0.16 × 0.16 = 0.0256
1.0000

Let us calculate the frequencies of the genotypes of the offspring of the above matings.

Parents	Ratio of offspring			Frequencies of genotypes of offspring			Total frequency
				AA	Aa	aa	
AA × AA	1AA			0.1296			0.1296
AA × Aa	1AA	1Aa		0.0864	0.0864		0.1728
AA × aa		1Aa			0.0576		0.0576
Aa × AA	1AA	1Aa		0.0864	0.0864		0.1728
Aa × Aa	1Aa	2Aa	1aa	0.0576	0.1152	0.0576	0.2304
Aa × aa		1Aa	1aa		0.0384	0.0384	0.0768
aa × AA		1Aa			0.0576		0.0576
aa × Aa		1Aa	1aa		0.0384	0.0384	0.0768
aa × aa			1aa			0.0256	0.0256
				0.3600	0.4800	0.1600	1.0000

Therefore, at the end of the first generation the frequencies of genotypes have once again remained constant when all the conditions of Hardy-Weinberg law are met. It is needless to point out here that the frequencies of alleles A and a will also be 0.6 and 0.4 respectively, since there is no change in genotypic frequencies.

In the next section we shall start analysing one by one the role of outside evolutionary forces such as mutation and selection in influencing the Hardy-Weinberg law. Prior to that attempt the following SAQ.

SAQ 2

a) The frequency of the allele D in a population is 0.55. What is the frequency of the recessive allele in the same population?

.....

b) Two allele C and c have their frequencies as 0.7 and 0.3. What are the frequencies of the possible genotypes that these alleles could form?

.....

20.3 INFLUENCE OF OTHER EVOLUTIONARY FORCES ON GENE FREQUENCIES

In the previous section we stated the Hardy-Weinberg law and showed that the allelic and genotypic frequencies remained the same after a generation when there was random mating in the population and the other evolutionary forces are absent. You have also learnt that these evolutionary forces include mutation and selection. In this section we shall look into the influence of such evolutionary forces on allelic and genotypic frequencies. Besides the influence of evolutionary forces, it is true that populations are usually small and the matings among the individuals are non-random, all of which once again contribute to charge in gene frequencies. The net effect of the changes in gene frequencies over a number of generations in a population is the evolution of gene pool. The evolution is essentially due to the interplay of different evolutionary factors. We mentioned mutation and selection earlier. Genetic drift and gene migration are two other factors.

20.3.1 Mutation and Change in Gene Frequencies

Mutations are heritable changes in the genetic material occurring in a locus (refer to Unit 16 of LSE-03). Mutations are the raw materials for evolution. Mutations often convert one allelic form into another leading to a change in the phenotypic expression. The rate at which mutations occur is usually small. It varies between

10^{-4} to 10^{-6} in eukaryotes and 10^{-7} or 10^{-9} in prokaryotes. This means that in eukaryotes, it is only one out of 10,000 to 1,000,000 loci undergoes mutation. Mutation rate in effect determines the rate at which the gene frequencies change. If the mutations are lethal the rate of change in the frquencies of alleles is high as persons carrying lethal mutations do not survive.

We earlier said that mutations are the sources of all genetic variability. First mutations cause genetic variations and then in response to various evolutionary forces, the different alleles increase or decrease in frequency.

Many mutations are detrimental and therefore eliminated from the population. Some mutations confer adaptations on individuals who possess them and therefore are likely to spread in the population. Whether a mutation confers any advantage or disadvantage on individuals who possess them depends on the environment in which these individuals live. A mutation which is advantageous in one particular environment may prove to be otherwise in a changed environment (refer to units 11 of LSE-07 for a detailed discussion on mutations and evolution).

The mutation of a wild type allele to its mutant form is referred to as forward mutation. For instance mutation of A to a. If mutant locus undergoes mutation to its wild type, then it is referred to as reverse mutation. Reverse mutation rates are smaller than forward mutation rates. Now, let us look into the mechanism by which mutations alter the gene frequencies.

Let us consider a hypothetical population in which the allele A has the frequency of p and allele a has the frequency of q. Let us also say that A mutates to a (forward mutation) at a rate of u and a mutates to A (reverse mutation) at a rate of v. We assume that the population is large and there is no selection operating on the population. The change in frequency of A depends on the mutation rate u and on the initial frequency of A, p. For instance, assume that there are 100,000 alleles in a population and the mutation rate is 10^{-4}. If all the alleles in the population are A and $p = 1.0$, the number of alleles that would undergo mutation is $100{,}000 \times 10^{-4} = 10$. Contrarily if there are only 10,000 alleles then the number of mutant alleles in a generation would be $10{,}000 \times 10^{-4} = 1$. The decrease in the frequency of A as a result of mutation to a is up. Similarly the increase in frequency of a as a result of mutation from A is vq. The change in frequency $= vq - up$. As we pointed out earlier, when the frequency of A is high, the number of alleles undergoing mutation is also high. As the frequency of A decreases, the number of alleles undergoing mutation also decreases. Initially q will be small, but as more and more A alleles mutate to a, q increases and the number of alleles undergoing reverse mutation also increases. Essentially this would result in an equilibrium, a point at which the number of alleles undergoing forward mutation becomes equal to the number of alleles undergoing reverse mutation. Also, at this point there will be no further change in gene frequency although mutations continue to occur in both the directions. As the equilibrium is reached, the change in the frequency of A, will be equal to zero.

$$\Delta p = vq - up = 0$$

i.e., $vq = up.$

Since $p = (1 - q)$

$$vq = u(1 - q)$$

$$vq = u - uq$$

$$vq + uq = u$$

$$q(v + u) = u$$

$$\hat{q} = \frac{u}{(v + u)}$$

q is known as equilibrium value of q. Similarly the equilibrium value of p i.e., p can be obtained and this will be $\hat{p} = \frac{v}{v+u}$.

Now, we shall try to explain the above expressions in numerical terms. Let us say that in a population frequency of $A = p = 0.6$ and that of $a = q = 0.4$. Allele A mutates to a at a rate of $u = 4 \times 10^{-5}$ and a undergoes reverse mutation at a rate of $v = 2 \times 10^{-5}$. The change in gene frequency in the first generation is

$$\begin{aligned}\Delta p &= vq - up \\ &= (2 \times 10^{-5} \times 0.4) - (4 \times 10^{-5} \times 0.6) \\ &= -0.000016.\end{aligned}$$

The frequency of A in the 2nd generation is $0.6 - 0.00016 = 0.599984$ and that of $q = (1-p) = 0.400016$. At equilibrium, the frequency of allele a, q is equal to

$$q = \frac{u}{u+v} = \frac{(4 \times 10^{-5})}{(4 \times 10^{-5}) + (2 \times 10^{-5})} = 0.667.$$

You should remember that mutation rates are ordinarily very low and for the change in gene frequency to occur only due to the mutation pressure, it must take several generations. For instance if frequency of A is 1.00 then it might require 1000 generations to change the frequency to 0.99 at a mutation rate of 1×10^{-5}. It might take 2000 generations to change it from 0.5 to 0.49 and roughly 10,000 generations to change it from 0.1 to 0.09. Such calcualtions hold good only if there are no reverse mutations. But with reverse mutations occurring, the rate may be even slower.

To summarise it could be said that gene frequencies change over a period of time if mutations occur in the population and other evolutionary forces are not active. Further, with both forward and reverse mutations occurring, the gene frequencies tend to reach an equilibrium over a period of time.

SAQ 3

Match the following:

a)	$p =$	i)	$\left(\frac{u}{u+v}\right)$
b)	$\hat{q} =$	ii)	zero
c)	$\hat{p} =$	iii)	$(1-q)$
d)	At equilibrium $vq - up =$	iv)	one
e)	$(p+q) =$	v)	$\left(\frac{v}{v+u}\right)$

20.3.2 Selection and Change in Gene Frequencies

The theory of natural selection propounded by Charles Darwin explains the process of evolutionary change. (For a detailed discussion on natural selection refer to units 11 and 12 of LSE-07.) Natural selection is a process which has a direction and order. It aims to promote adaptations in populations. Adaptations refer to the phenotypic traits or expressions that make the organisms more suited to their immediate environment. Adaptations besides contributing to the increased survival of the organisms, also increase their reproductive efficiency. Adaptations arise from among the variations in the traits of the organisms. Variations are universal in nature. No single trait is fixed. All traits exhibit a measure of variability. Such of those variations which are useful to the organism in terms of its survival and

reproduction in a given environment are deemed as adaptations. The green colour of the insect living on a tree that helps it to escape predation by enemies is a clear case of adaptation. Other insects may possess slightly different colours but may not be able to camoflouge themselves among green leaves of the tree and therefore easily picked up predators such as birds. Insects which are well adapted to live in such an environment have better chances of survival and leave behind more number of offspring. In fact, natural selection is related to the notion of differential reproduction. Whereas survival *per se* is important, it is the reproduction that ensures that one's genes are passed on to the succeeding generation. What is more important is not the survival of individuals but the genes which they leave behind, and for this reproduction must occur. Therefore, natural selection is quantified by assessing reproduction. **The relative reproductive efficiency of genotypes is termed Darwinian fitness or adaptive value.**

Fitness is usually symbolised as w. Since it is a relative value, it is assigned the value of 1 to that genotype that produces maximum number of offspring. For instance, let us say that two alleles A and a form genotypes AA, Aa and aa. AA produces four offspring, Aa produces six offspring and aa two offspring. Here Aa produces maximum number of offspring and therefore, is assigned the fitness value of 1. The fitness value of other genotypes is calculated relative to the genotype with the fitness value of 1, in this instance, Aa. Therefore, AA has a fitness value (w) of $4/6 = 0.67$ and that of aa is $2/6 = 0.33$.

Fitness value is a measure of reproductive efficiency of an organism in terms of natural selection. The converse, namely the measure of relative intensity of selection against a genotype is called **selection coefficient**. Selection coefficient symbolised as $s = 1 - w$. In the example cited above, the genotype AA will have a selection coefficient of $1 - 0.67 = 0.33$; Aa has a s value of $1 - 1 = 0$ and that of a is $1 - 0.33 = 0.67$

Does natural selection act to change the allelic frequencies in a population? The answer appears to be yes. Depending on the type of environment in which it acts, selection may prefer to maintain a constant frequency of an allele by eliminating variations in a population. Or it may endeavour to change the frequency by maintaining variations. The relative fitness of the individuals in a population and the initial gene frequencies influence the action of natural selection.

In Table 20.1 we have computed fitness value and selection coefficient for three genotypes we have chosen as example. Table 20.2 explains the influence of selection on gene frequencies.

Table 20.1 Calcuation of fitness value (w) and selection coefficient (s) of three genotypes

Genotypes	AA	Aa	aa
No.of breeding adults in generation 1	12	20	10
No. of offspring produced by all adults in generation 2	48	120	20
Average number of offspring produced per adult	48/12 = 4	120/20 = 6	20/10 = 2
Fitness value (w) (relative number of offspring produced)	4/6 = 0.67	6/6 = 1	2/6 = 0.33
Selection coefficient (s)	1 – 0.67 = 0.33	1 – 1 = 0	1 – 0.33 = 0.67

Table 20.1 shows the method of arriving at the fitness value and selection coefficient, when the number of individuals belonging to various genotypes and number of offspring they produce are known. From Table 20.2 you could learn how the gene frequencies are altered under the influence of natural selction.

Table 20.2 Influence of Natural Selection on the Changes in Gene Frequencies in a Population

Genotypes	AA	Aa	aa
Initial genotype frequency	p^2	$2pq$	q^2
Fitness value (w)	w_{11}	w_{12}	w_{22}
Frequency of genotypes after selection	$p^2 w_{11}$	$2pqw_{12}$	$q^2 w_{22}$
Relative genotype frequency after selection	$P' = \frac{p^2 w_{11}}{\bar{w}}$	$H' = \frac{2pqw_{12}}{\bar{w}}$	$Q' = \frac{q^2 w_{22}}{\bar{w}}$
(where $\bar{w}$ = mean fitness value of population $= p^2 \bar{w}_{11} + 2pq\bar{w}_{12} + q^2 \bar{w}_{22}$)			

Gene frequency after selection:

frequency of A = $p' = P' + \frac{1}{2}(H)$

frequency of a = $q' = 1 - p'$

Change in gene frequency of A due to selection =

$\Delta p = p' - p$

Changes in gene frequency of a due to selection =

$\Delta q = q' - q$

In Table 20.2, the frequency of different genotypes after selection is calculated by multiplying the initial frequency (frequency before selection) with the fitness value of each genotype (w_{11} for AA, w_{12} for Aa and w_{22} for aa). The sum of the genotypic frequencies after selection = $\bar{w}$ and is known as mean fitness value of the population. The frequency of each genotype after selection is divided by the mean fitness value to obtain the relative frequency of each genotype. From these values the allelic frequencies can be obtained by using the following formula.

Frequency of allele A after selection = p′ = frequency of AA after selection + ½ (frequency of Aa after selection) and q′ = (1 – p′).

Let us substitute the numerical values for the above expressions as shown in Table 20.3. Let us consider two alleles A and a with initial gene frequencies of 0.6 and 0.4 respectively.

Table 20.3

Genotypes	AA	Aa	aa
Intial genotype frequencies	p^2 0.6×0.6 $= 0.36$	$2pq$ $2 \times 6.0 \times 0.4$ $= 0.48$	q^2 0.4×0.4 $= 0.16$
Fitness value (w)	w_{11} $= 0.67$	w_{12} $= 1$	w_{22} $= 0.33$
Frequency after selection	$p^2 w_{11} =$ 0.36×0.67 $= 0.240$	$2pqw_{12} =$ 0.48×1 $= 0.48$	$q^2 w_{22} =$ 0.16×0.33 $= 0.053$
Mean fitness value of the population $\bar{w}$ =	$p^2 w_{11} + 2pqw_{12} + q^2 w_{22}$ $0.24 + 0.48 + 0.053$ $= 0.773$		
Relative genotype frequency after selection	$P' = \frac{p^2 w_{11}}{\bar{w}}$ $H' = \frac{2pqw_{12}}{\bar{w}}$ $Q' = \frac{q^2 w_{22}}{\bar{w}}$		
	$= \frac{0.24}{0.773}$ $\frac{0.48}{0.773}$ $\frac{0.053}{0.773}$		
	0.31 0.62 0.07		

Frequency of A after selection = $p' = P' + \frac{1}{2}(H')$

$$= 0.31 + \frac{1}{2} \times 0.62 = 0.62$$

And the frequency of a after selection = $1 - p' = 1 - 0.62 = 0.38$

Change in frequency due to selection for allele A

$\Delta p = p' - p = 0.62 - 0.6 = 0.02$

and for allele a = $\Delta q = q' - q = 0.38 - 0.4 = -0.02$

In other words, after selection process, the frequency of A has increased from 0.6 to 0.62 and that of a has decreased from 0.4 to 0.38.

20.3.3 Selection against a Recessive Trait

Many new traits arise as recessive mutations and are subject to severe selection pressure. Obviously such recessive mutants have a reduced fitness value. Under such circumstances both the dominant homozygote and the heterozygote have a fitness value of 1. This means that there is no selection pressure on these genotypes. The fitness value of recessive genotype would be (1 - s) where s is the **selection coefficient**. Let us tabulate the frequency and the fitness values of these genotypes.

Genotype	Frequency	Fitness Value
AA	p^2	1
Aa	$2pq$	1
aa	q^2	$1-s$

Assuming the frequency of genotypes are in Hardy-Weinberg proportions, the contribution of each of genotype to the next generation will be the product of their frequency and fitness value. Thus, the frequency of genotypes in the next generation will be

AA $\quad p^2 \times 1 = p^2$

Aa $\quad 2pq \times 1 = 2pq$

aa $\quad q^2 \times (1-s) = (q^2 - sq^2)$

The mean fitness value of the population would then be

$$p^2 + 2pq + q^2 - sq^2$$

Since $p^2 + 2pq + q^2 = 1$,

the mean fitness vlaue = $(1 - sq^2)$.

The relative genotypic frequencies after selection will be

$$AA = \frac{p^2}{(1 - sq^2)}$$

$$Aa = \frac{2pq}{(1 - sq^2)}$$

$$aa = \frac{q^2 - sq^2}{1 - sq^2}$$

In order to obtain the frequency of allele a after selection, that is q' value, we have to add the frequency of homozygote (aa) and half of the frequency of heterozygote (Aa)

$$q' = \frac{q^2 - sq^2}{1 - sq^2} + \frac{1}{2} \times \frac{2pq}{1 - sq^2}$$

$$= \frac{q^2 - sq^2}{(1 - sq^2)} + \frac{pq}{(1 - sq^2)}$$

$$= \frac{q^2 - sq^2 + pq}{1 - sq^2}$$

$$q' = \frac{q^2 + pq - sq^2}{1 - sq^2}$$

$$= \frac{q(q + p) - sq^2}{1 - sq^2}$$

Since $(q + p) = 1$

$$q' = \frac{q - sq^2}{1 - sq^2}.$$

and

$$p' = (1 - q')$$

The above expression helps you to determine the frequencies of a pair of alleles when there is a selection pressure on the recessive genotype and the homozygous dominant and heterozygous genotypes have a fitness value of 1. At this point we end our discussion on the role of selection in changing the gene frequencies. Before we proceed to discuss the role of genetic drift and gene migration in changing the gene frequencies in populations, attempt the following SAQ.

SAQ 4

i) Define the terms adaptive value and selection coefficient.

.....

.....

.....

.....

ii) If allele has an adaptive value of 0.55, what is its selection coefficient?

.....

iii) Genotype AA has a value of $w = 0.5$ and its frequency in the population is 0.6, genotype Aa has a $w = 1.0$ and has an initial frequency of 0.3. Genotype aa has $w = 0.2$ and has an initial frequency of 0.2. What is the mean adaptive value of the population $(\overline{w})$?

.....

20.3.4 Genetic Drift

For gene frequencies to remain constant generation after generation, Hardy-Weinberg law requires that population to be infinitely large. But usually natural populations are not infintely large. At the same time, they are large enough that the gene frequencies do not change drastically. But there are several small populations in which there are random changes in gene frequencies. Such random changes in gene frequencies due to chance event are called **genetic drift**. The phenomenon was first brought to light by an American population geneticist Sewall Wright, hence come to be known as Sewall Wright effect.

Random changes in gene frequencies or genetic drift is an important evolutionary force in small populations. To illustrate genetic drift in small populations, let us

look into the following example. Imagine a small population of mice living in a barn of a farmer as four or five extended families. The farmer tries a variety of ways to get rid of the mice. He may try to set traps, use a shot gun, make surprise vists, employ cats and other means as well. Such acts of the farmer exert a severe selection pressure. Under such circumstances the traits that would be selected are the swiftness, short tails, hearing acuity and cautiousness. In other words, the frequencies of genes for such traits would be very high. Contrarily the gene for thick fur on the body will have a low frequency as thick fur is not the trait that will be chosen in the present environment. Assuming hearing acuity has a frequency of 0.9 in the population, thick fur may have a frequency of 0.2. Now, let us imagine a sudden environmental change. A severe winter sets in. The farmer confines himself to the fire place. Such of those traits which had a very high frequency when the farmer was hunting the mice will lose their significance in the altered environment. So, the next generation of mice with thick fur alone survive and produce offspring with thick fur. The frequency of the gene for thick fur increases in population and that of hearing acuity decreases. In the next generation the respective frequencies of genes for heaving acuity and fur thickness could be respectively 0.3 and 0.8. Since the population is small, severe winter eliminates from the environment mice with less fur in the body. Assuming the population consisted of 1000 individuals, it is unlikely that most individuals with less thick fur would be eliminated from the population by chance alone and the frequency for hearing acuiity would not decrease randomly from 0.9 to 0.3.

The change in environment causing a heavy mortality in natural populations is one of the several ways in which genetic drift occurs. Such a drift is essentially caused by sampling error. Sampling error would refer to chance deviations from expected ratios. For example, in an infinitely large population, infinitely large number of gametes are produced. The frequency of a pair of alleles in such a population, let us say, is p and q and the frequency of genotypes they form are p^2, $2pq$ and q^2. Such frequencies would remain same in the next generation as well. But, from such a large population, if you take a sample, then the number of progeny of this sample is limited. Just by chance or error, the sample may deviate from the large gene pool. The potential for deviation is much larger in smaller and smaller samples.

All genetic drift is due to sampling error. And the sampling error may occur in a number of ways. They are:

- Genetic drift arises when the population continuously remains small over many generations.

- Genetic drift may arise through **founder effect**. By founder effect it is meant that population initially consists of a small number of individuals but may subsequently grow in size and later may consist of large number of individuals. But the gene pool of the population is derived from the genes of individuals who originally founded the population. The founder population may not be the representative of the parent population from which it is formed.

An often cited example for founder effect in natural population is that of Dunker families living in United States. Dunkers are a religious sect who migrated from Germany to Pennsylvania in USA between 1719 and 1729. Fifty Dunker families emigrated, settled in USA and remained as an isolated group and formed a small population. They very rarely married outside their communities. During 1950s, geneticists studied the Dunker population consisting of 300 individuals. Studies showed that frequencies of some of genes of Dunker population were very different from the frequencies of genes of population of US where Dunkers have settled as well as from the frequencies of genes of population in Germany from where they migrated. Table 24.4 shows study of the frequencies of alleles ABO blood group system in Dunker population, American population and German population.

Table 20.4 Comparison of frequencies of blood group alleles and phenotypes of American, German and Dunker populations.

Population	Allelic frequencies			Phenotypic frequencies			
	I^A	I^B	I^O	A	B	AB	O
Dunker	0.38	0.03	0.59	0.593	0.036	0.023	0.348
USA	0.26	0.04	0.70	0.431	0.058	0.621	0.490
W.Germany	0.29	0.07	0.64	0.455	0.095	0.041	0.410

You could observe from the Table 20.4 that the frequencies of alleles controlling blood group system of Dunkers are different from those of Americans and Germans.

You should remember that Dunkers have largely remained isolated and did not mix with the American population. Under such circumstances only the phenomenon of drift could explain for producing random changes in the gene pool of Dunker population. To begin with, it was founder effect as only 50 families migrated from Germany to get settled in US in 1729. From then on drift had influenced gene frequencies in each generation, since the population size had always remained small.

- The third form of sampling error is due to **bottleneck effect**. Bottleneck effect occurs when the population size is drastically reduced in one generation. During such reduction in the size of the population, some genes may be lost from the gene pool by chance. We earlier pointed out that reduction in population size may occur due to several reasons including natural calamities. Such reduction in population sizes cause random changes in frequencies of gene and the resultant drift.

Genetic drift and the random changes in gene frequencies in small populations cause much changes in the genetic structure of the populations. Such effects can be summarised as follows:

- Drift causes changes in gene frequencies of a population over a period of time.
- Drift may also result in the loss of genetic divergence in populations. In some cases an allele may be completely lost when all the individuals possessing a particular allele in a small population are killed due to chance event.
- Drift also causes loss of genetic variation in population. With a loss of a particular allele from the population, the population may increasingly become homozygous. Decrease in heterozygosity results in the loss of genetic variation in the population.

We end our discussion on genetic drift here and proceed to discuss yet another factor that could bring out changes in gene frequencies, namely migration.

20.3.5 Migration

While describing Hardy-Weinberg law, we mentioned that for the law to hold true, the population should not be influenced by other evolutionary forces. The population should more or less remain a closed one. But most populations are not very isolated and exchange of genes with other populations do often occur. When new individuals enter into a population, they may either introduce a new allele into the population or alter the frequencies of the existing alleles of the population. **Migration** of individuals which in effect means migration of genes into a population may alter the Hardy-Weinberg equilibrium and can cause the evolution of gene

frequencies in a population. To be a force in evolution, migration should ensure that migrating individuals are reproductively active contributing genes to the recipient population. In such an event, the migration can also be described as **gene flow**.

The fact that gene flow causes changes in the frequencies of allele in a population can be explained by a simple model.

Let us consider two populations, population I and population II. The frequency of allele A in population I is p_i and that of population II is p_{ii} At each generation, let us say, some individuals randomly migrate from population I to II. At the end of migration population II will have two groups of individuals—the migrants and residents, with allelic frequencies of p_i and p_{ii} respectively. Let us designate the proportion of individuals who form migrants in population II as m

Frequency of allele A in population I before migration	**Frequency of allele A in population II after migration**
f(A) = p_i	(Residents = p_{ii} ; migrants = mpi) f (A) = $mp_i + (1 - m)\, p_{ii}$

The change in the frequency of A due to migration =

$$\Delta p = p'_{ii} - p_{ii}$$
$$\Delta p = mp_i + (1 - m)\, p_{ii} - p_{ii}$$
$$\Delta p = mp_i + p_{ii} - mp_{ii} - p_{ii}$$
$$\Delta p = mp_i - mp_{ii}$$
$$= m(p_i - p_{ii})$$

As you could observe from the above equation the change in frequency due to gene flow depends on two factors. (i) The proportion of migrants in the final population (m); (ii) the difference in the frequency between migrants and residents (mp_i–p_{ii}).

Migration causes certain significant changes in a population. One, it might increase the net effective size of the populations. Two, it might reduce the genetic drift in small population as well as genetic divergence.

SAQ 5

State whether the following statements are true or false.

i) For Hardy-Weinberg equilibrium to hold good population sizes should generally be small. (True/false)

ii) Change in gene frequencies due to chance events is called genetic drift. (True/false)

iii) Chance deviations from expected ratios is referred to as sampling error. (True/false)

iv) Founder effect and bottleneck effect refer to phenomena that are contrary to genetic drift. (True/false)

v) Drift increases genetic variations in small populations. (True/false)

vi) Migration of individuals in effect means the flow of genes into a population. (True/false)

vii) Gene flow causes an increase in genetic divergence in a population. (True/false)

20.4 SUMMARY

- Hardy-Weinberg equilibrium forms the central theorem of population genetics.
- The frequencies of a pair of alleles tends to remain constant generation after generation, if the population is infinitely large and randomly mating, and the outside evolutionary forces have no influence on the population.
- Outside evolutionary forces such as mutation and selection significantly alter the gene frequencies over a number of generations.
- Genetic drift caused by chance events tends to cause random and unpredictable changes in the frequencies of alleles.
- Migration of individuals and the subsequent gene flow cause alterations in gene frequencies. Gene flow may minimise the effects of genetic drift in small populations as it increases the net effective size of the population and decreases the genetic divergence.

20.5 TERMINAL QUESTIONS.

1. State the Hardy-Weinberg law.

2. ABO blood groups are controlled by a system of three alleles I^A, I^B, I^O. Assuming a population having these three alleles consists of genotypes which are randomly mating, what are the possible mating pairs?

.....

.....

.....

.....

.....

3. Discuss briefly that mutations are an evolutionary force.

.....

.....

.....

.....

.....

4. What is the role of selection in altering the gene frequencies in a population?

.....

.....

.....

.....

5. Define the term genetic drift. How does genetic drift affect gene frequencies in a population?

.....

.....

.....

.....

20.6 ANSWERS

SAQs

1. (i) Infinitely large

   (ii) Mutation, natural selection, migration and genetic drift

   (iii) Mendelian population

   (iv) heredity, reproduction

   (v) 0.1

2. (a) 0.45

   (b) 0.49, 0.42, 0.09

3. a) iii b) i c) v d) ii e) iv

4. (i) Adaptive value: Refers to the relative reproductive efficiency of a genotype.
   Selection coefficient: The amount of selection pressure on a genotype. It is related to adaptive value by the equation $s = (1 - w)$

   (ii) 0.45

   (iii) 0.64

5. (i) False (ii) True (iii) True (iv) False (v) False
   (vi) True (vii) True

Terminal Questions

1. Hardy-Weinberg laws states that frequencies of a pair of alleles moved remain constant in a population generation after generation, provided that the population is infinitely large, randomly waiting and that there are no outside evolutionary forces acting on the population.

2. $I^AI^A \times I^AI^A$
   $I^AI^A \times I^AI^O$
   $I^BI^B \times I^BI^B$
   $I^BI^B \times I^BI^O$
   $I^AI^B \times I^AI^A$
   $I^AI^B \times I^BI^B$
   $I^AI^B \times I^AI^O$
   $I^AI^B \times I^BI^O$
   $I^AI^B \times I^AI^B$
   $I^AI^A \times I^AI^O$
   $I^AI^B \times I^OI^O$
   $I^OI^O \times I^AI^A$
   $I^OI^O \times I^AI^O$
   $I^OI^O \times I^BI^B$
   $I^OI^O \times I^BI^O$
   $I^OI^O \times I^OI^O$

3. Mutations which cause heritable changes in the genetic material occuring in a locus are the raw maternals for evolution. They are responsible for generating variations in the population. Such of those mutations which are advantageous to the organisms are retained in the population and their frequencies increase in the population. Accumulation of adaptations is an evolutionary process. Since one of the definitions of evolution is that it is change in gene frequencies, mutations being responsible for change in gene frequencies, contribute to the evolutionary process.

4. Refer to Section 20.3.2

5. Refer to Section 20.3.4

UNIT 21 QUANTITATIVE TRAITS AND GENETICS OF TWINS

Structure **Page No.**

21.1 INTRODUCTION

In previous units you have studied phenotypic variation that is easily classified into distinct traits, such as, irregular or regular variation of coleus leaves; presence or absence of horns in cattle; coat colour in rabbits; blood group in humans. These phenotypes are examples of **discontinuous variation** where discrete phenotypic categories exist.

However, not all inherited traits are expressed in this discontinuous fashion. For example, height, skin and eye colour in humans, colour and food yield in various plants, size in many plants and animals, as well as degree of coat spotting in some animals are examples of **continuous phenotypic variation**. It is now known that traits exhibiting continuous variation are often controlled by two or more genes and are termed **polygenic**. In cases where several genes make additive contributions to the phenotype, the trait is said to exhibit quantitative or **continuous variation**.

In this Unit you will study about the phenotypic traits which are controlled by genes at two or more loci. You will also know how statistical tools are used by geneticists to study traits that exhibit continuous variation. Further, you will learn about the effect of nongenetic factors on gene expression; these include environmental influences. In addition, you will be introduced to the concept of **heritability**, which is used to estimate the degree of genetic and environmental influence on the expression of traits controlled by genes at many loci. You will also study about twins. This study is necessary to determine the relative role of heredity and environment in the differences existing between individuals. Before reading this unit you must read Block I and II of MTE-03. It will also be helpful in comprehending various statistical concepts used in this Unit.

Objectives

After studying this unit you will be able to:

- distinguish between continuous and discontinuous variations,
- distinguish between the role of genotype and environment in controlling the phenotype of a quantitative traits,
- estimate the possible number of genes involved in the expression of a given phenotype,
- identify the relative role of heredity and environment by the twin studies.

21. 2 CONTINUOUS VARIATION

Most of the genetic traits can be identified by their distinct phenotype. That means mutants can be easily distinguished from wild type because of a clear cut phenotypic difference. All individuals fall into a few phenotypic classes with respect to such traits. Such traits are called discontinuous traits, some examples are ABO blood groups of humans, coat colour of cattle, prototrophs and auxotrophs in bacteria. For discontinuous traits, the relationship between phenotype and genotype is clear and simple except under co-dominance and epistasis. Therefore, it is possible to infer about the genotype when the phenotype is known.

Certain characteristics such as birth weight, adult height, I.Q., and skin pigmentation in humans exhibit a wide range of possible phenotypes. Similarly in cattle a trait like milk production shows a continuous range in phenotype with no clear separation between one phenotype and the next. These traits can be measured numerically in single individuals and can fall anywhere on a continuous scale of measurement and the number of possible phenotypes is virtually unlimited.

After rediscovery of Mendel's laws in 1900, many controversies arose, one of them was whether all traits follow the classical Mendel's laws or there are differences in the inheritance patterns. During the latter part of 19th century, Francis Galton and his associate Karl Pearson studied a number of continuous traits in humans such as height, weight and I.Q., by developing various statistical techniques, they could show that these characters do not have a simple mode of transmission, i.e., the inheritance pattern is not controlled by a single gene. Continuous variation is determined by multiple genes, each of which is segregated independently in a classical Mendelian manner. These traits are also influenced by the environment.

21. 2.1 Origin of Continuous Variation

Continuous variation exists as a number of phenotypes among individuals of a group. This is due to the influence of many loci and alleles influencing the trait. For example, there are three genotypes for a gene at a single locus with two alleles. If a trait is controlled by two loci each with two alleles, the number of genotypes becomes 3^2, i.e. 9 (AABB, AABb, AAbb, AaBB, AaBb, Aabb, aaBB, aaBb and aabb). Therefore, the general formula calculating the number of genotypes is 3^n, where 'n' is the number of loci with two alleles. The number of genotypes also increases as the number of alleles at each locus increases (Fig 21.1).

Therefore, these traits are called polygenic traits as they are controlled by many gene loci. If each genotype of a polygenic trait encodes a separate phenotype, then many phenotypes with slight differences will be the result. For polygenic or multifactorial traits, no single relationship exists between genotype and phenotype. Therefore, the simple modes of inheritance predicted by Mendel do not provide information about genes involved in continuous traits. The genetic basis of these traits can be understood by special analytical procedures/methods, and cannot be studied by the usual pedigree analysis.

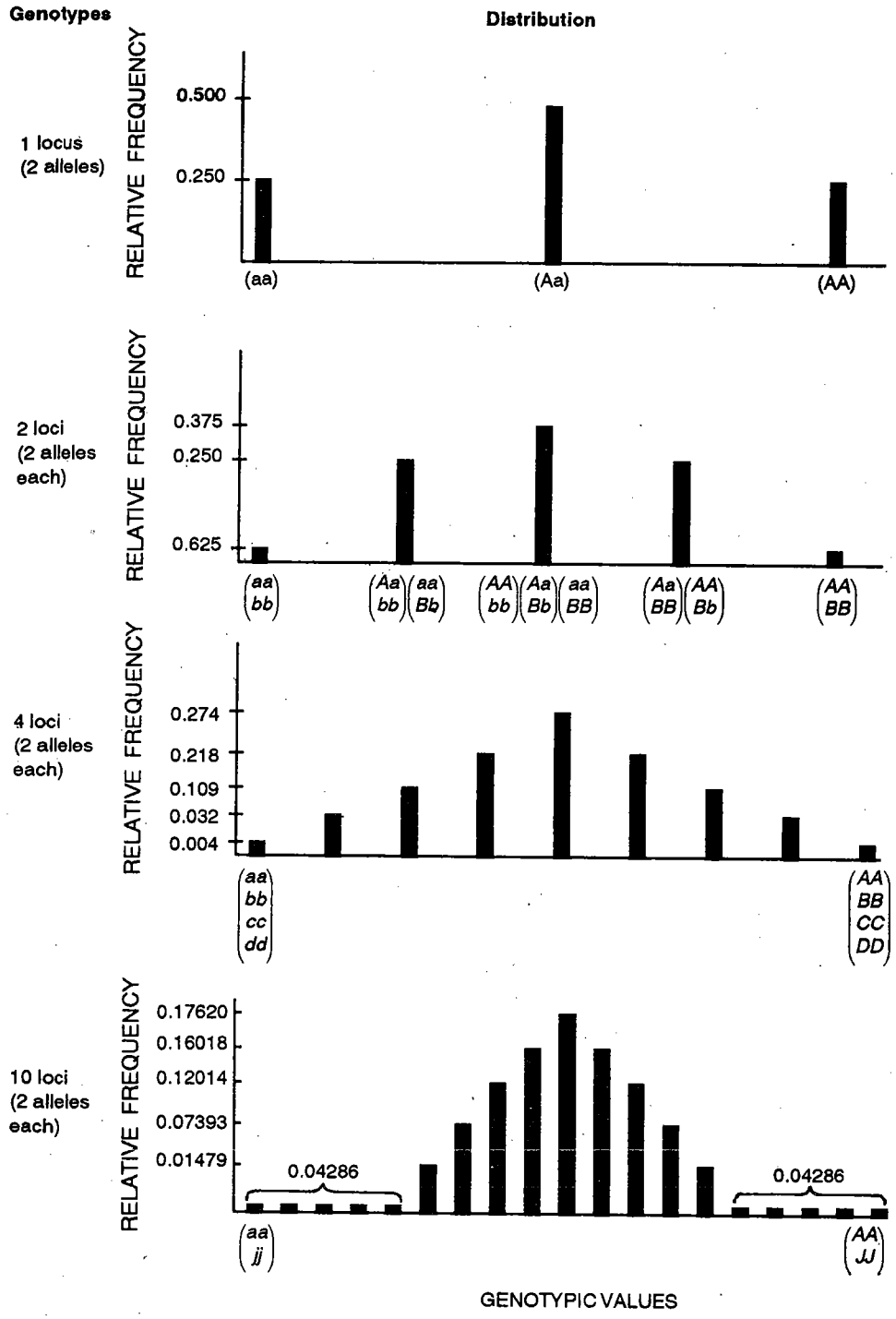


Fig 21.1: Distribution of Frequencies in Genotype.

When dominance is present at any of the loci, different genotypes may express the same phenotype. This could be the reason for lesser number of phenotypes observed than expected. The second important effect is that of the environment. When environmental factors influence the genotypes, each genotype can produce different phenotypes depending on the extent of interaction with the environment.

Phenotype = Genotype + Environment (P = G + E). One common example is yield in crop plants. This depends on the gene pool, as well as the rainfall, fertilisers, plant density and soil quality. Many applications of genetics are based on an understanding of continuous traits. In agriculture, yield from plants, in animal husbandry, milk production, egg-laying and fleece weight, and in humans, height, weight, IQ, serum cholesterol and lifespan are such quantitative traits.

SAQ 1

Let us assume i) that a trait is controlled by the loci on different chromosomes, A/a, B/b, C/c, ii) that A locus has average phenotypic effect of the three genotypes as AA = 4, Aa = 3, aa = 1, and that iii) similar effects exist for B + C loci. Now, calculate the distribution of phenotypes when two triple heterozygotes (AaBbCc) are crossed.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

21.3 QUANTITATIVE INHERITANCE

Many traits are not controlled by only alleles of two or more genes, but are also influenced by the effects of environment. Consequently, by the study of quantitative traits you will be able to comprehend the following:

1) The contribution of genetic and environmental factors to the phenotypic variation.

2) The number of genes involved and the contribution from each of them.

3) The extent of interaction between different alleles and the role of additive and epistatic interactions.

4) The effect of natural selection on phenotypic expression.

With the help of an example the inheritance of a quantitative trait is explained below.

21.3.1 Inheritance of Ear Length (Cob Size in Corn)

Corn has been the subject of genetic and cytological studies for many years. Ear length of corn is one of the characters that was studied by Emersson and East in 1913 to demonstrate a pattern of quantitative inheritance. A cross was made between two pure lines, i.e., Black Mexican sweet corn (which had short ears of mean length 6.63 cm) and Tom Thumb popcorn (which had long ears of mean length 16.80 cm). The F_1 plants had mean ear length of 12.12 cm which was approximately intermediate between ear length of the two parental lines. Assuming that the parental lines were homozygous, the F_1 plants are then heterozygous for all the genes and should theoretically have cobs of the same ear length. But the observations indicated a variability in the ear length from plant to plant. Therefore, the range of ear length phenotypes seen in F_1 plants must be due to factors other than genetic factors. These are environmental factors since it is impossible to grow all the plants in exactly identical conditions (Fig 21.2).

In F_2 the mean ear length was 12.89 cm. F_2 mean was about the same as the F_1 mean. But the F_2 population has a much larger variation around the mean than the F_1 population. Is this variation is the result of the effects of environmental factors? Certainly, if the environment was responsible for variation in the parental and the F_1 generations, we have every reason to believe that it would have a similar effect

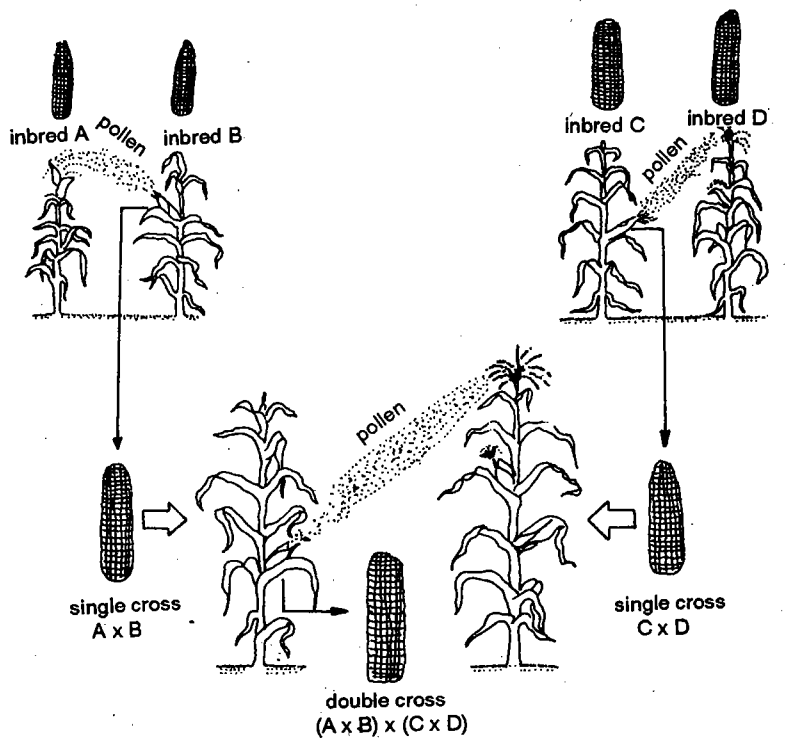


Fig 21.2: Quantitative inheritance in maize.

on the F_2 generation. However, there is no reason to believe that environment would have a greater genetic variation in the F_2 generation.

For a moment if we keep the environmental influence aside, from our data we can make the following four observations:

1) The mean value of the quantitative trait in the F_1 is approximately intermediate between the mean of the two true-breeding parental lines.

2) The mean value for the trait in the F_2 generation is approximately equal to the mean for theF_1 population.

3) The F_2 shows more variability around the mean than the F_1 does.

4) The extreme values for the quantitative trait in the F_2 extend further into the distribution of the two parental values than do the extreme values of the F_1.

21.3.2 Polygenic Hypothesis for Quantitative Trait

The simplest explanation that could be given for the data described above is that quantitative traits are controlled by many genes. This hypothesis is called **polygene** or **multiple gene hypothesis**. This hypothesis can be dated back to 1909 when Nilson-Ehle crossed two true breeding wheat plants, one with red kernels and the other with white kernels. The F_1 had grains of intermediate colour between red and white. When F_1 plants were interbred, F_2 progeny showed a ratio of 15 red (all shades) : 1 white kernel (Fig 21.3).

The polygene or multiple gene hypothesis for quantitative inheritance is regarded as one of the landmarks of genetic thought.

If we assume that there are two pairs of independently segregating alleles that control the red colour pigmentation i.e., 'R' and 'C' both would result in red colour and 'r' and 'c' result in white colour. From this variety of wheat, genetically pure

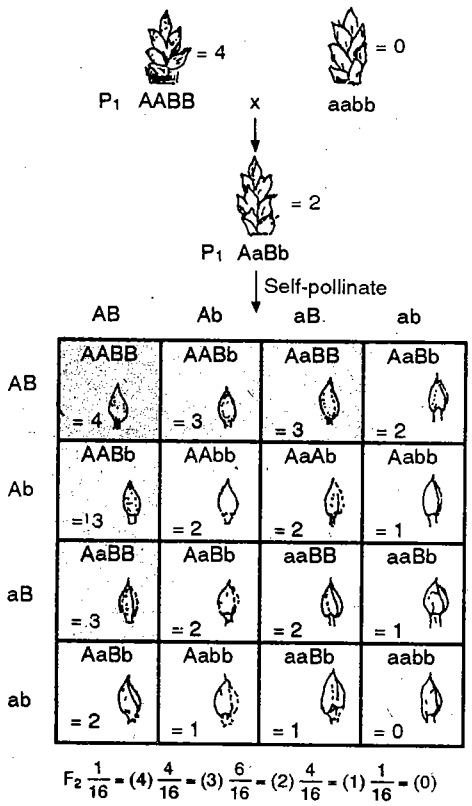


Fig 21.3: Quantitaive Inheritance. Diagrammatic representation of Nilsson-Ehle's crosses between a red variety (AABB) of wheat and a white variety (aabb). The 16 possible gamete combinations of the F_2 generation are shown as a normal distribution based on number of plus genes present in genotype. (The curve area of F_2 generation have been enlarged to accomodate all the genotype)

strains were developed by inbreeding and used as parents. If **R** and **C** are dominant over **r** and **c** we get a phenotypic ratio 1:4:6:4:1. The most plausible explanation is that each of the two genes controlled the production of the colour pigment. Therefore, the intensity of red colour is a function of the number of dominant **R** and **C** alleles. The alleles (like **R** and **C**) which contribute to the phenotype are called contribution of alleles while those (like **r** and **c**) which do not contribute to the phenotype are called the non-contributing alleles. In the present context the inheritance of red kernel colour in wheat is an example of polygene series of as many as 4 contributing alleles. Depending upon the number of contributory alleles involved different phenotypic ratios are obtained. In the case of 3:1 ratio one gene with two alleles is involved, while in 15:1 there are 2 genes with 2 alleles and in the 63:1 case, a polygenic series with 6 contributory alleles from 3 genes are involved.

Identification of the number of genes in a quantitative trait has not been accomplished as yet in most cases. You can see from the theoretical model (presented) in Fig 21.4 that the number of genotypes increases with the increase in number of loci each having two alleles.

In its basic form, multiple gene hypothesis proposes a number of attributes of quantitative inheritance. It can be explained on the basis of action and segregation of alleles at a number of loci that have identical additive effects on the phenotype without complete dominance. It can be summed up that quantitative traits are influenced by many genes, each of which contributes a small and additive effect on the the phenotype.

By now you will realise that in support of polygenic hypothesis a number of assumptions have been made. Amongst the genes involved in the expression of the trait:

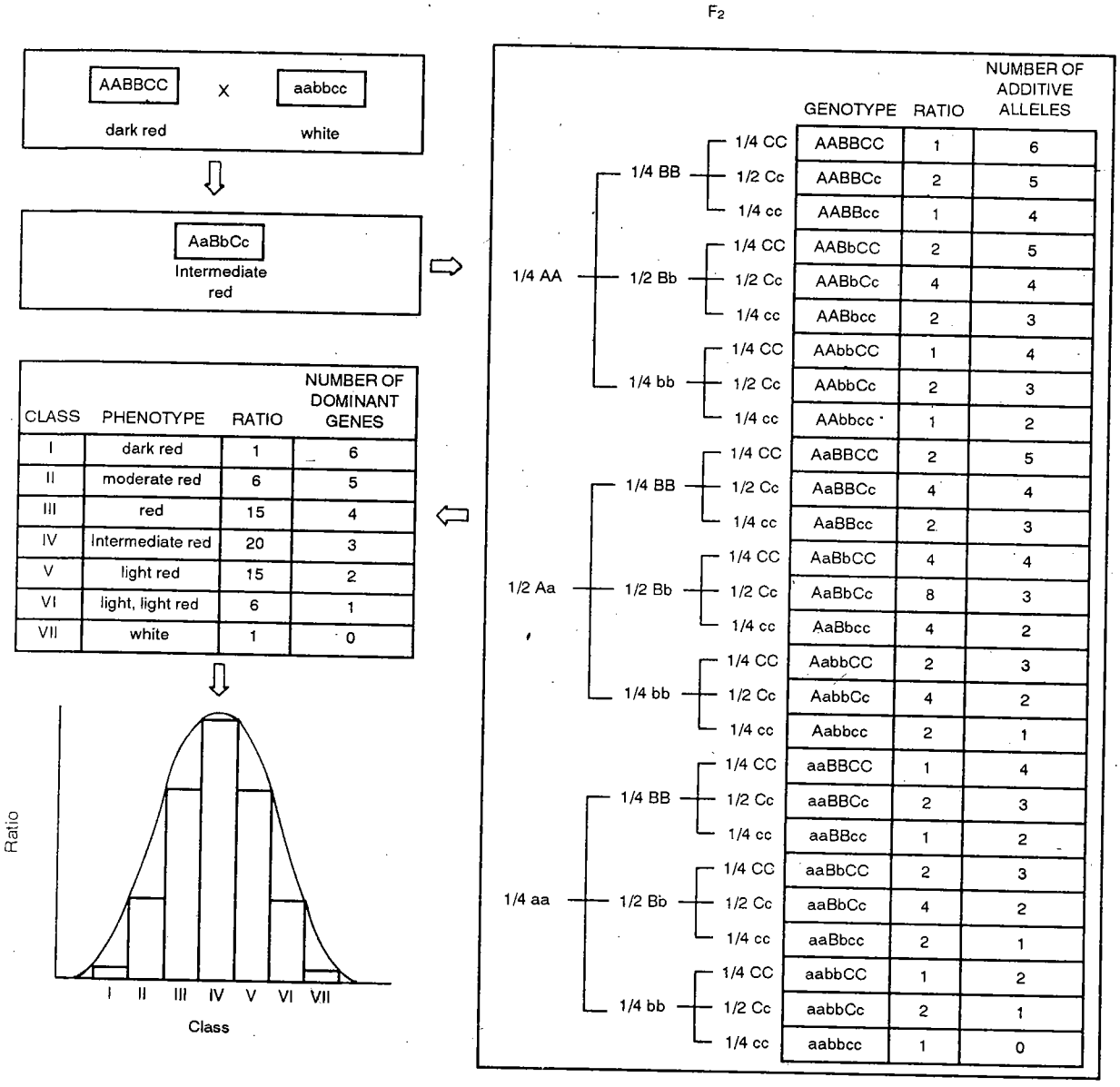


Fig. 21. 4: Polygenic inheritance where three genes are contributing to grain colour in wheat.

1) No allelic pairs exhibit dominance.
2) Only a series of contributory and non-contributory alleles are involved.
3) Each contributory allele has an equal effect.
4) The effect of each contributory allele is additive.
5) There is no genetic interaction between alleles of different loci.
6) There is no linkage between the loci involved.

However, in practice it is difficult to meet all these assumptions.

21.3.3 Some Examples of Quantitative Traits in Humans

i) **Stature in Man** : Height is a good example of a quantitative trait in humans. A number of loci with two alleles control the trait. Let us assume that there are 4 gene

Sometimes diseases like pitutary deficiency can lead to dwarfism. Individuals affected with such diseases tend to be short despite their carrying many genes for tallness.

loci with two alleles each (A/a, B/b, C/c and D/d). Then the stature of individuals will depend on the number of contributing alleles. For example, according to this hypothesis a very tall person would be AABBCCDD, a very short aabbccdd and person with medium height Aa Bb Cc Dd.

On an average the height of children falls between the range of the parents. For example, if the father is 170 cms and the mother 152 cms then the offspring height is usually between 170-152 cms.

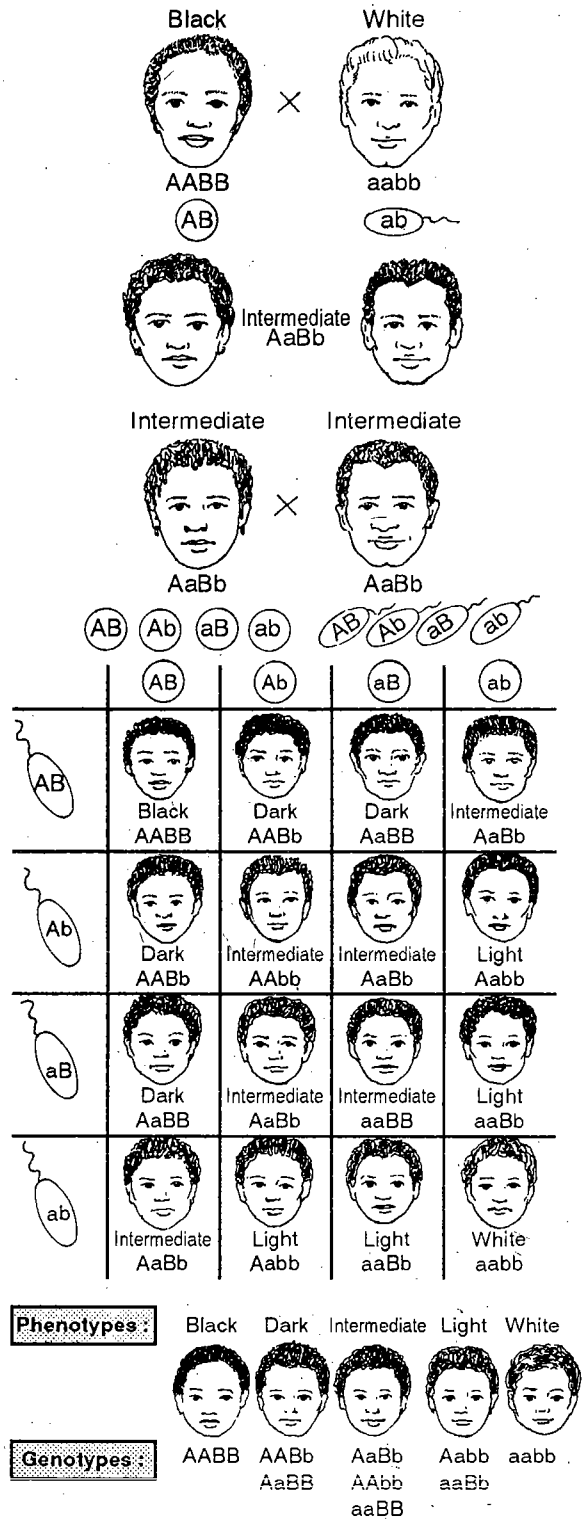
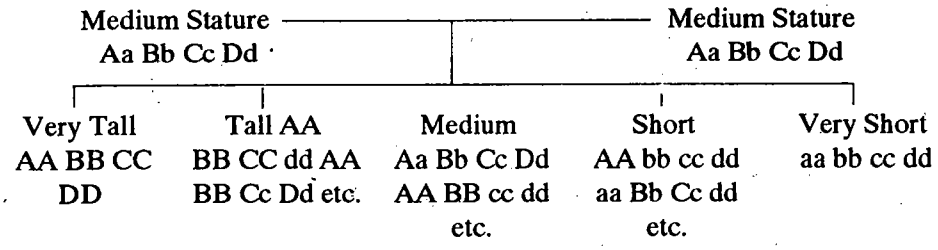


Fig. 21.5: Skin colour inheritance - the skin colour is attributed by two chief genes. They both contribute equally to melanin production. The black colour square represents such melanin producing gene. In white colour individual all four alleles produce minimum amount of melanin. But when one of the parent is black and other is white the children have intermediate brown skin colour. In such heterozygote children two alleles produce large amount of melanin. Since skin colour is additive effect such children are brown.

Now it would be demonstrated to you how medium sized parents can have a very tall or a very short child.



Apart from this, non- genetic factors, like diet, also play a major role in the development of human stature.

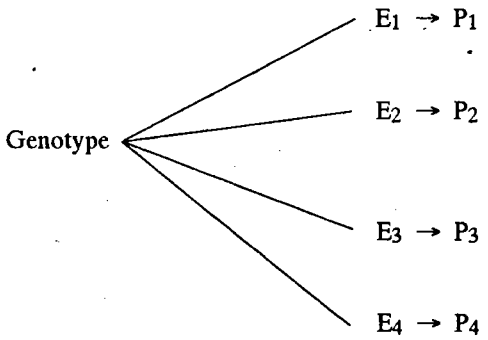
ii) **Skin Colour** : Skin colour is also dependent on a number of genes. We can see many gradations in human skin colour ranging from extremely fair to very dark. Colour of the skin is dependent on a pigment called **melanin**. The difference in skin colour of Negroes and Caucasians is very clear and about 4 to 7 genes are considered to be involved in controlling it.

The offspring of marriages between Negroes and caucasians have an intermediate skin colour, and are referred to as Mulattos (Fig 21.5). The progeny of mulattos may range from very fair to very dark. This variation may be explained on the same basis as that for height.

In this case in the presence of noncontributory alleles the individuals inherit white colour and the dark individuals possess alleles which contribute to colour formation.

21.4 EFFECT OF ENVIRONMENT ON QUANTITATIVE TRAITS

By now you know that the phenotypic variation observed has genotypic as well as environmental contributions. Johannsen (1909) studied the responses of a genotype to different environmental circumstances and results can be modelled as follows:



'E' denotes environment, 'P' denotes phenotype.

It indicates that the same genotype interacts with different environmental factors that influence the phenotypic expression of the traits. Environmental factors can be physical (nonbiological in origin) biological (other organisms and products of other genes within the organism) or both.

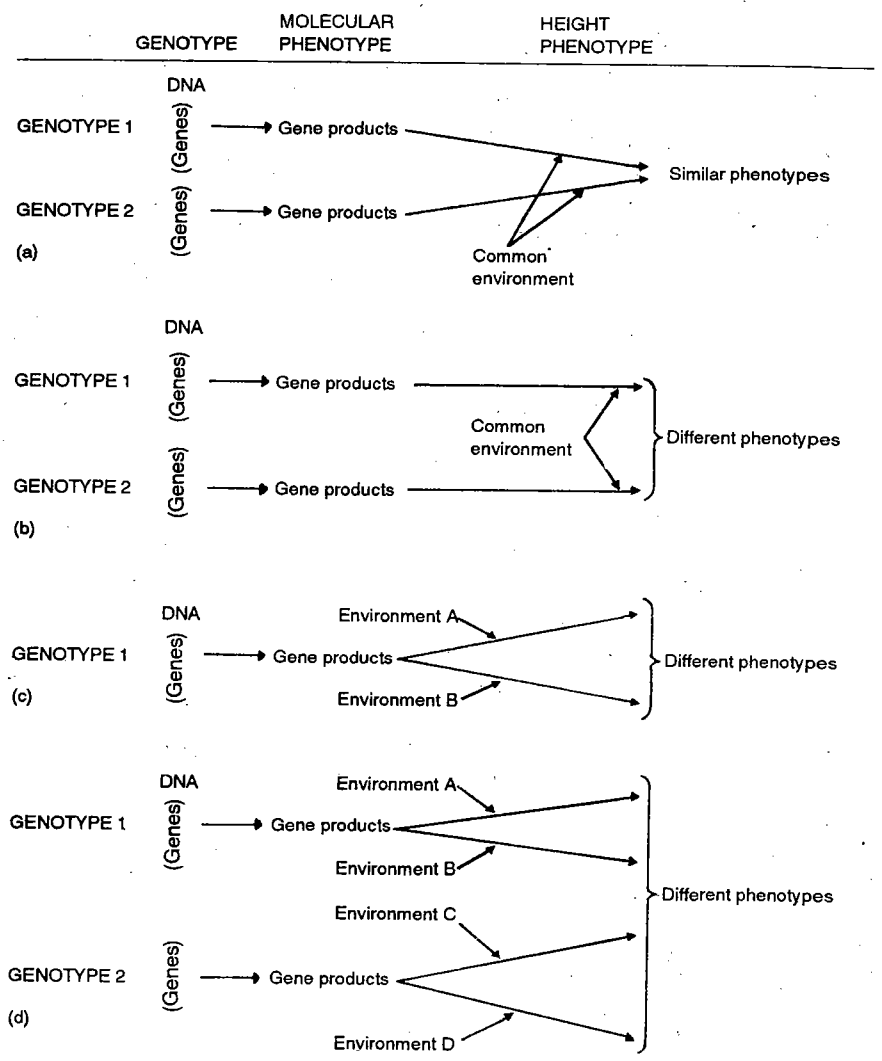


Fig 21.6: Quantitative phenotypes are affected by interplay between genotype and environment.

Three different types of interplay occur between genotype and environment to affect the phenotype. These are depicted in Fig 21.6. The interaction between genes and environment make the study of quantitative phenotypic variation rather complex.

21.5 HERITABILITY

The concept of heritability is used to examine the relative contributions of genes and environment to variation in a specific trait. As we have seen continuous traits are frequently influenced by multiple genes and by environmental factors. One of the most important question arises in quantitative genetics is; how much of the variation in phenotype is due to genetical differences and how much is due to environmental variation. For example, multifactorial traits such as weight of cattle, number of eggs laid by chickens, and fleece produced by sheep are important for breeding programs and agricultural management. Many ecologically important traits such as variation in body size, fecundity and developmental rate are also multifactorial and the genetic contribution to this variation is important for understanding how natural populations evolve. Thus we can define **heritability as the proportion of variability, attributable to genetic factors out of the total phenotypic variability existing in a population.**

21.5.1 Components of Phenotypic Variance

In this section we will learn how to identify and measure the relative contribution of the genetic effects to the total phenotypic variation of a trait.

The phenotypic variance is a measure of the variability of a trait. It is calculated as the total variance observed in a population with respect to the trait in question and is designated as V_P. Differences among individuals arise from several factors, and therefore, we can partition the phenotypic variance into several components attributable to different sources. Contribution of genetic differences among individuals to the phenotypic variation is called genetic variance and denoted by V_G.

The individuals experiencing different environment may contribute to the differences in their phenotypes. The environmental variance is symbolised by V_E and by definition, it includes nongenetic source of variation; temperature, nutrition and parental care are example of obvious environmental factors that may cause difference among individuals.

A third component of phenotypic variance is an interaction between genetic and environmental factors and is represented as V_{GE}. For example, in cold temperature, a plant with AA genotype may grow 40 cm tall, and that with Aa may grow 35 cm tall. But in warm temperature plants with the same genotypes may grow to 45 cm and 60 cm respectively (Fig 21.7).

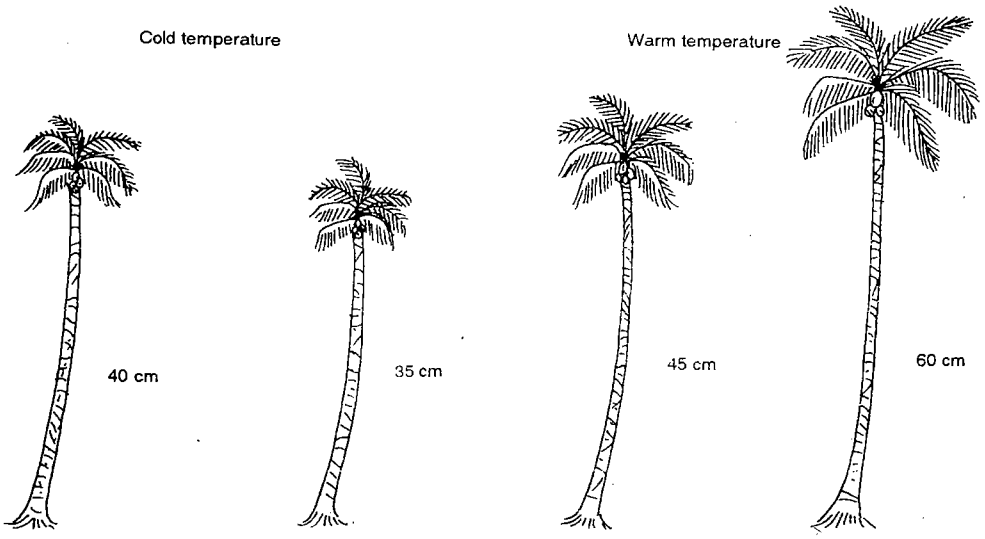


Fig 21.7: Hypothetical example of genetic and environmental interaction.

In this hypothetical example of genetic and environmental interaction, the plant with a genotype of Aa which grows slower at cold temperature outgrows AA at higher temperature. This suggests that genotypes express differentially in different environmental conditions. Plants with both types of genotypes grow taller at higher temperatures, but the effect varies with the type of genotype. Therefore, these differences contribute to the phenotypic variance but the effects of genetic and environmental factors cannot be added together. So an additional component of variance that accounts for interaction between genetic and environment effects (V_{GE}) must be considered.

Hence we come to the conclusion that the total phenotypic variance is composed of a number of factors: 1) Genetic composition of a population, 2) Specific environmental conditions, 3) The manner in which genes interact with the environment.

Variance also includes random factors that occur during development, which are sometimes referred to as developmental noise.

The genetic variance (V_G) can be further subdivided into components arising from different types of interaction between genes. Let us assume **'a'** allele contributes 2 cm to plant height, while **'A'** allele contributes 4 cm. Thus **aa** genotype would contribute 4 cm in height, while Aa contributes 6 cm and AA contributes 8 cm in height. Then to determine genetic contribution of these alleles to the phenotype the effects of alleles at this locus would be added to the effects of alleles at other loci which are known to influence the phenotype. Such genes are said to have additive effects. The variation that arises due to additive effects is called additive genetic variance, and symbolised as V_A.

Some genes exhibit **dominance** and this becomes a source of genetic variance. This is called dominance variance (V_D). Dominance of one allele masks the effect of the other allele at the same locus.

We must also consider the interactions of alleles at different loci. Epistasis (allele at one locus alters the expression of alleles at another locus) is another source of variance arising out of interaction of genes is designated as V_I genetic variation. So the genetic variance can be partitioned as:

$$V_G = V_A + V_D + V_I$$

Total phenotypic can finally be partitioned as:

$$V_P = V_A + V_D + V_I + V_E + V_{GE}$$

Table 21.1: Components of Variance

Variance Component	Symbol	Value whose variance is measured
Phenotypic	V_P	Phenotypic value
Genotypic	V_G	Genotypic value
Additive	V_A	Breeding value
Dominance	V_D	Dominance value
Interaction	V_I	Dominance deviation
Environmental	V_E	Environmental deviation

The components of variance are listed in Table 21.1. This partitioning of phenotypic variance is very important to understand the relative contribution of different factors towards the variation in phenotype.

21.5.2 Broad Sense and Narrow Sense Heritability

Geneticists frequently partition the phenotypic variance of a trait to determine the extent to which variation among individuals results from genetic differences. Thus, they are interested in how much of the phenotypic variance V_P can be attributed to genetic variance V_G. This quantity, the proportion of the phenotypic variance that consists of genetic variance is called the broad-sense heritability and is expressed as follows:

$$\text{Broad-sense heritability} = H^2 = \frac{V_G}{V_P} = \frac{V_G}{V_G + V_E}$$

Heritability of a trait can thus range from 0 to 1. If H^2 is '0' it means that differences found between individuals are not due to genes. And a heritability of '1' indicates that phenotypic variation is completely or 100 per cent genetic in origin. Heritability depends on magnitude of all the components of variance, a change in any of which will affect it. Therefore, it is a value for a population under specific conditions.

The proportion of phenotypic variance due to additive gene effects is of more interest because additive genes are those that allow the prediction of the phenotype of an offspring from the phenotype of the parents. Let us consider a cross involving a trait that results from effects of alleles at a single locus; one parent has a genotype of $A^1 A^1$ and is 10 cm tall, while the other parent $A^1 A^2$ is 20 cm tall. All the offsprings (F_1) from this cross will be $A^1 A^2$. If these alleles are additive and contribute equally to height, the offsprings should be 15 cm tall, exactly intermediate between the parents. Such an additive genetic variance helps in predicting the resemblance between offspring and parents. Therefore, quantitative geneticists frequently determine the proportion of phenotypic variance that results from additive genetic variance; this proportion is termed as narrow sense heritability and is represented as :.

$$h^2 = V_A/V_P$$

21.5.3 Estimation of Heritability

Heritability is an important property which expresses the total variance that is attributable to the average effects of genes and which determines the degree of resemblance between relatives. Estimation of heritability involves related and unrelated individuals or individuals with different degrees of relatedness (Table 21.2).

The estimate of heritability gives the relative importance of genetic versus environmental factors. If genes are more involved in determining phenotypic variance, then closely related individuals should be more similar in phenotype, as they have more common genes. Alternatively, if environmental factors are responsible for determining differences in the trait, then related individuals should be no more similar than unrelated individuals.

Table 21.2: Degree of Regression and Correlation

Relatives	Covariance	Regression (b) or correlation (t)
Offsprings and one parent	$\frac{1}{2} V_A$	$b = \frac{1}{2} h^2$
offspring amd mid-parent	$\frac{1}{2} V_A$	$b = h^2$
Half sibs	$\frac{1}{4} V_A$	$t = \frac{1}{4} h^2$
Full sibs	$\frac{1}{2} V_A + \frac{1}{4} V_D + V_{Ec}$	$t > \frac{1}{2} h^2$

Only the phenotypic values of individuals can be directly measured; however, it is the breeding value that determines their influence on the next generation. The degree of correspondence between phenotypic values and breeding values enables us to predict if it is possible to change the characteristics of the population. Heritability is the measure of such a degree of correspondence.

Quantitative genetics relies extensively on similarity among relatives to assess the importance of genetic factor. Therefore, narrow sense heritability, could be estimated from the degree of resemblance between relatives (correlation).

21.5.4 Variance versus Correlation

Genetic data about families is frequently collected as pairs of numbers like parent-offspring, sib-sib and twins. For example, height of three pairs of mother-daughter, is given in Table 21.3. To know if tall daughters are born to tall mothers, a correlation coefficient (r) between variables is calculated as follows:

$(x_1 y_1), (x_2 y_2), \ldots, (x_n y_n)$ are N pairs of measurements obtained. The x and y values correspond to the mother and daughter (parent-offspring) measurements, which are used to calculate covariance (C)

$$(C) = [1/(N-1)][(x_1 - x)(y_1 - y) + (x_2 - x)(y_2 - y) \ldots + (x_N - x)(y_N - y)]$$

The correlation coefficient (r) between x and y is then calculated from this covariance.

$r = C/(S_x S_y)$ in which S_x and S_y are the standard deviations of the variables. The correlation coefficient can range from – 1.0 to + 1.0. When $r = +1.0$ it indicates a perfect association. Thus, r is the usual measure of the precision of a relationship between two variables. Taking the values from Table 21.3.

Table 21.3 : Adult Height of three Mother-daughter Pairs

Pair	Height of Mother (inch)	Height of daughter (inch)
(x_1, y_1)	65.0	61.5
(x_2, y_2)	64.0	65.5
(x_3, y_3)	69.0	69.0

Mothers : mean (X) = 66.0, variance (S_x^2) = 7.0 Standard deviation (S_x) = 2.65

Daughters : mean (y) = 65.3, variance (S_y^2) = 14.1 Standard Deviation (S_y) = 3.75

$$C = [1/2](65-66)(61.5-65.3)+(64-66)(65.5-65.3)+(69.66)(69-65.3)]$$
$$= 7.25$$
$$r = 7.25/(2.65 \times 3.75) = 0.73$$

The r value of 0.73 which suggests that tall mothers do tend to have tall daughters.

SAQ 2

Correlations coefficient measured for three traits between mothers and daughters were as follows:

Trait	Corelation coefficient (r)
Blood Pressure	0.21
Height	0.49
Serum cholesterol	0.28

What can you comment about the heritability of these traits?

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

21.5.5 Uses of heritability estimates

Plant and animal breeders use information on heritability estimates in planning breeding programmes, for improving traits with economic value. Heritability estimates in humans provide useful information in predicting diseases caused by continuously varying phenotypes such as blood pressure. The identification of a gentic component in the causation of a disease can help in determining preclinical phenotypes. Individuals with risk for developing a disease can be identified and advised suitably.

21.6 TWIN STUDIES

One experiment of nature useful in the study of complex traits is the occurrence of multiple births — twins, triplets and so on. Because of greater prevalence twin briths provide good opportunities for observations. The utility of twin briths arises from the existence of two kind of twins: monozygotic (MZ or identical) or dizygotic (DZ or fraternal). MZ twins arise from a single zygote that forms two separate embryos very early in development. Since only one zygote is involved, the twins are genetically identical. DZ twins arise from two zygotes that are produced by fertilisation of two separate ova. Thus they have the same genetic relationship as ordinary sibs. They may be either belong to same sex (two boys or two girls) or opposite sexes (a boy and a girl). MZ twins are always of the same sex.

To determine the relative role of heredity and environment in the differences existing between individuals the phenotypic variation can be assessed in individuals with the same genotype. It is relatively easy to test the effect of nature and nurture in experimental animals or plants. Different strains can be produced, each of which is isogenic and thus environmental differences can be studied. In humans such isogenic strains are not available. Nevertheless, twins, seem to be ideal subjects to separate genotypic and environmental components of the observed phenotypic variance.

21.6.1 Frequency of Twinning

Frequency of twinning varies from one population to another. It is highest in Belgium with 1 in 55 births, result in twins. In the United States about 1 in 85 births is of twins, while among Japanese 1 in 145 births is of twins. Since the sample size in most twin studies is not very large the data should be interpreted with caution.

An empirical rule of thumb, known as Hellins's Law (see table 21.4) can be used to estimate the expected frequency of the higher multiple births. According to this law the frequency of twin births was 1/b, the frequency of triplets $1/b^2$ of quadruplets $1/b^3$, and so on. Actual data from 21 countries for 10 years consisted of 120,061,398 pregnancies. It showed 1,408,912 twins or 1/85.2, 15738 triplet or $1/(87.3)^2$ which is very close to the expected. By this rule, quintuplets would be very rare, expected in less than 1 in 50,000,000 pregnancies. However, it may be noted that widespread use of birth-control pills and fertility drugs may well have significant effects on these frequencies.

Table 21.4: Dr Hellin Rule:

Frequency of Birth	One birth in
Twins	89
Triplets	$(89)^2$
Quadruplets	$(89)^3$
Quintuplets	$(89)^4$

The frequency of monozygotic and dizygotic twins has also been found to vary in different populations. These frequencies can be estimated because of the fact that all unlike-sexed twins must be dizygotic. If we assume the probability of a boy to be *a* and of a girl *b*, and that the formation and development of the two zygotes are independent events, then the proportions of boy-boy, boy-girl/girl-boy, and girl-girl pairs can be estimated from the bionomical expansion of $(a + b)^2$, or a^2 (boy-boy), 2ab (boy- girl/girl-boy) and b^2 (girl-girl). Since $a = b = \frac{1}{2}$, p (boy – boy)= $\frac{1}{4}$, p (boy – girl/girl – boy) = $\frac{1}{2}$ and p (girl – girl = $\frac{1}{4}$.

Thus, the total for all dizygotic twins should be twice the observed number of boy-girl/girl-boy pairs. The number of monozygotic pairs is then determined by subtracting the estimates for the number of dizygotic twin pairs from the total number of all twin pairs.

This method of estimating the frequencies of monozygotic twins is known as Weinberg's differential method. When the method is applied, it revealed that in the USA about 1/3 of the sets of twins born are monozygotic but in Japan, nearly 2/3 of the twin pairs are monozygotic.

21.6.2 Diagnosis of Zygosity

In twin studies the question of whether each set of twins is monozygotic or dizygotic has been of great importance. If the sexes are similar they are identical but it is not necessary because fraternal twins can also be female-female or male-male pairs. Therefore, it is evident to find some method to determine the zygosity of monozygotic or dizygotic twins. Usually it is known that dizygotic twins that looked very much alike were misclassified as monozygotic twins. The reverse also occurred, however, because monozygotic twins formed by a relatively late separation of the cells of the original zygote can exhibit mirror imaging in some characteristics, that is, one may be right-handed and other left handed

A more objective diagnosis of zygosity has been based on the type of placentation. But this, too, is fraught with possibilities of error. The developing zygote is enveloped by two membranes, the inner delicate one is **amnion** and outer tougher **chorion** attached to the maternal tissue of the placenta. The most common arrangements of these membranes and the placentas in twins are indicated in Table 21.6 and Fig. 21.8 A.

Unequivocal diagnosis of zygosity is possible only when there is a monochorionic placenta (Fig. 21.8 B), a condition that prevails in about 70 per cent of monozygotic twins but not in dizygotic twins. Dizygotic twins, on the other hand, are always dichorionic, because dizygotic twins always implant separately into the uterine epithelium and always develop their own membranes. They need not, however, present separate placenta of the type (Fig. 21.8 A). In some 50 per cent of dizygotic twins the two placentas are sufficiently close that they become secondarily fused (Fig. 21.8 C). Such dizygotic twins of like sex are not infrequently misdiagnosed as monozygotic.

Most dependable method of diagnosing zygosity is by the skin graft test. A skin graft will always be accepted by monozygous twins because of their identical tissue antigens. Though the skin graft is often rejected by a dizygous twins this outcome is not certain. In addition, the method posses many practical problems.

However, the overall probability that dizygotic twins will be alike in all of the traits studied is almost always much smaller than the corresponding probability for monozygotic twins. In other words, the similarity approach is simply a method of arriving at degrees of probabilities as to the zygosity of a given pairs of twins.

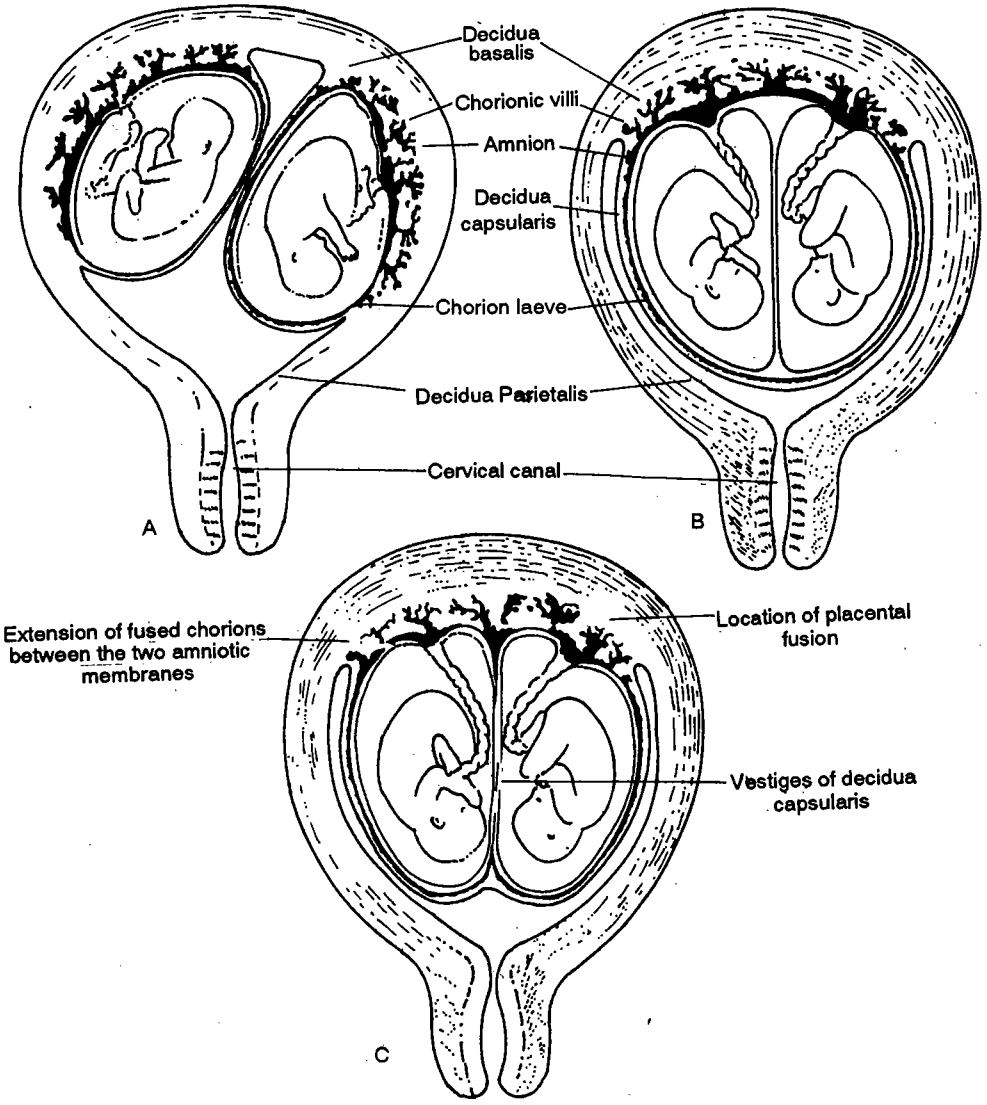


Fig. 21.8: Diagram showing the three most frequent relations of the fetus membranes of twins. A. Twins with entirely separate placenta. B. Twins with a single chorion and separate amnions. C. Twins implanted very close to each other with resultant secondary fusions of their membranes.

21.6.3 Use of Twin Studies

Francis Galton emphasised the importance of studying twins to obtain information on nature-nurture problem. Since monozygotic twins provide individuals of identical genotypes, the observed or measured phenotypic variance enables the estimation of the environmental components. Many studies have been conducted on human traits. They are of two main types.

i) Differences between identical and fraternal twins (Table 21.5).

Table 21.5 : Composition of the components of variance between and within pairs of twins

	Within Pairs	Between Pairs
Different	$1/2\ V_A + 3/4\ V_D$	$1/2\ V_A + 3/4\ V_D$
Fraternals	$1/2\ V_A + 3/4\ V_D + V_{EW}$	$1/2\ V_A + 1/4\ V_D + V_{EC}$
Identicals	V_{EW}	$V_A + V_D + V_{EC}$

ii) Comparison between identical twins reared together and reared apart. A simple way of scoring differences between twins is to evaluate traits which are either present or absent. Thus twins may be either concordant (i.e. both individuals of a pair possess the trait or are free of the trait in question) or discordant (i.e. only one member of the pair possess the trait).

Concordance is the percentage of cases in which both members of a twin pair have a particular trait. For example, 100 pairs of monozygotic twins were studied for a particular disorder, in 70 pairs only one member was affected while in 30 pairs both members were affected with the disorder. In this case the condordance would be 30/(30 + 70). Concordance varies with the degree of genetic determination of the trait. A 100 per cent concordance indicates that the trait is under complete genetic influence. An example of such a trait is blood group. Blood pressure and pulse rate, two physiological traits studied in twins, are represented in the Figure 21.9. Concordance in continuously varying traits like blood pressure, is defined as similarity within a specified range. Thus in the study of blood pressure, concordance meant agreement of the twins within a pressure difference of less than 5 mm mercury. Concordance of 63% indicates that there is an environmental influence inspite of the genetic similarity. Thirty six per cent concordance in non-identical twins also leads to the interpretation that heredity has a role in determining the differences in blood pressure.

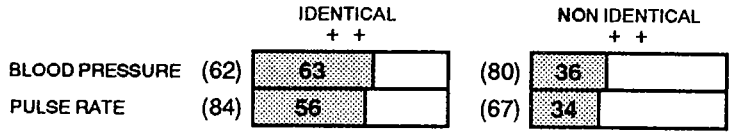


Fig. 21.9: Concordance and discordance in turns for blood pressure and pulse rate. The total width of bar is equal to 100%. And diagonally lined section denotes per cent of concordance. While white section represents discordance total number of twin pair investigated are given in parenthesis.

Congenital deformities and other pathological conditions affecting twins were also studied. These are represented in Figure 21.10. Members of identical pair of twins have been found to be extremely alike in most characters like facial expression even when they were reared apart.

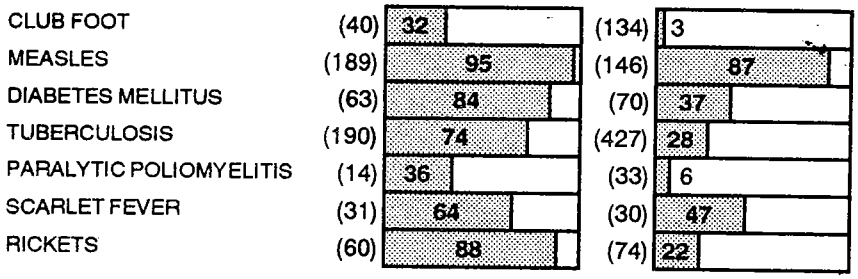


Fig. 21.10: Concordance and discordance in twins affected by various pathological conditions. Percentage of concordance based on affected pairs only.

In addition, numerous anatomical traits (e.g. height, weight, head lenght or width physiological conditions and pathological agent (e.g. clubfoot, diabetes and rickets) have been studied in twins. Non-identical twins and sibs have been found to be alike in distribution and degree of expression of traits. The similarity in these studies reflects the genetic correspondence between fraternal twins and sibs (Fig. 21.11 and Table 21.6).

Table 21.6: Average differences between the two members of identical twins nonidentical twins, pair of sibs reared together; and identical twins, reared apart

Difference in	Identical	Nonidentical	Sibs	Indentical reared apart
Height	1.7	4.4	4.5	1.8
Weight	4.1	10.0	10.4	9.9
Head length	2.9	6.2	-	2.20
Head width	2.8	4.2	-	2.89

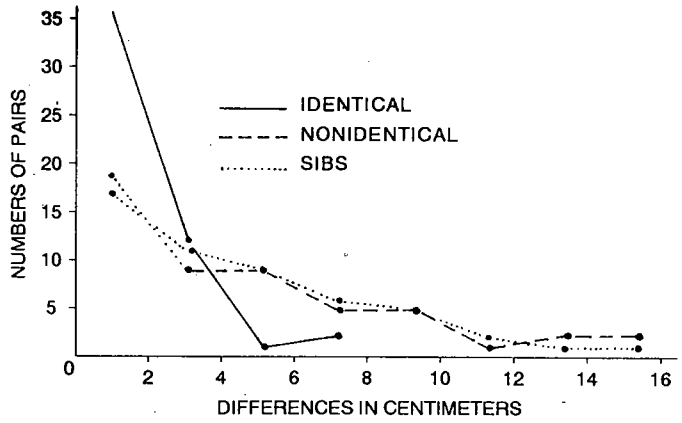


Fig. 21.11: Curves of distributions of differences in standing height of 50 identical twins, 52 nonidentical twins, 52 pairs of sibs.

21.6.4 Genetic Inference from Twin Studies

There have been many attempts to develop quantitative indices of heritability from twin studies. Heritability estimates based on twin studies for certain traits are represented in Table 21.7.

Table 21.7 : Resemblance between Twins

Character	Correlation coefficients		Differences	Hertability
	Identicals	Fraternals		
MAN				
Height	.93	.64	.29	0.81
Weight	.92	.63	.29	0.78
Intelligence	.88	.63	.25	0.68
Birth weight	.67	.58	.09	0.21
CATTLE				
Milk-yield, 1st lactation	.91	.65	.26	0.74
Butterfat-yield, 1st lactation	.90	.51	.39	0.80
Fat % in milk, 1st lactation	.95	.86	.09	0.64
Weight at 96 weeks	.83	.78	.05	0.23
Body length at 96 weeks	.75	.62	.13	0.34

Unfortunately for most characters the degree of genetic determination is very high indicating several important sources of error in twin studies. They are:

i) Genotype-environment interaction which will increase the variance in fraternal but not identical twins.

ii) Sharing of embryonic membrane making it mandatory for a similar intra-uterine environment.

iii) Similarity in the treatment of twins by parents, and teachers resulting in a decreased environmental variance in identical twins (Fig. 21.12). These errors can be overcome partly by comparison of traits between monozygotic and dizygotic twins. In Table 21.7 you can see the list of some traits estimated in identical and fraternal twins. If a trait has genetic component then the identical twins will be more alike with respect to the trait in question than will fraternal twins.

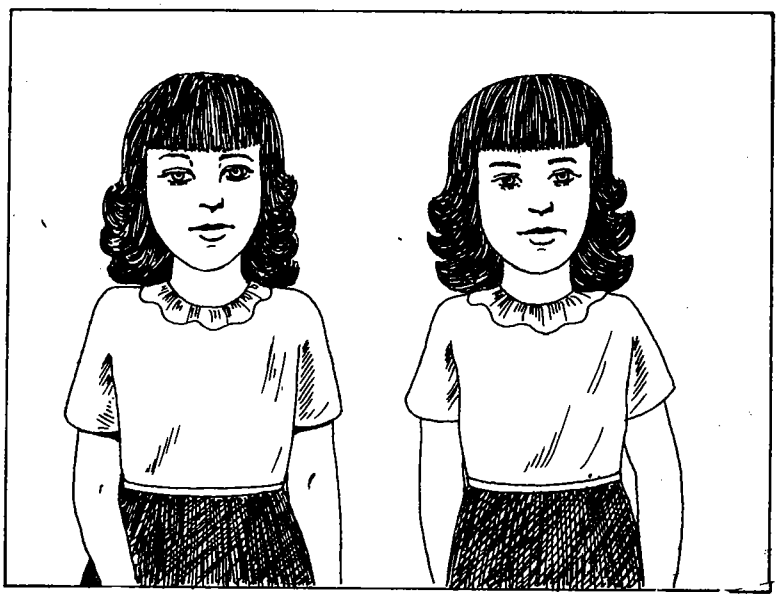


Fig. 21.12: Monozygotic Twins.

One of the earliest concepts was simply to divide the frequency of monozygotic concordance by the frequency of dizygotic concordance. Such a statement of relative frequencies can be misleading. Under it the point on the scale indicating "no heritability" is not zero but one, that is, when dizygotics are just as likely to be concordant as monozygotics. Further confusion comes from the fact that the bottom of the scale is 100 per cent. One of best improvements is Holzinger's formula under it.

$$\text{Heritability} = \frac{\text{Percent Monozygotic concordance–Percent Dizygotic concordance}}{100 - \text{Percent Dizygotic Concordance}}$$

Or, in short, it is

$$H = \frac{C_{MZ} - C_{DZ}}{100 - C_{DZ}}$$

The theoretical maximum would occur if monozygotics were 100 per cent concordant and dizygotics never, so that

$$\text{Heritability} = \frac{100 - 0}{100 - 0} = 1$$

Similarly the minimum would come when concordances in both types were equal and both were 30 per cent concordant, for instance

$$\text{Heritability} = \frac{30 - 30}{100 - 30} = 0$$

Thus, heritability could be scaled from zero to one.

Unfortunately, Holzinger's formula would not be applicable to the majority of quantitative traits that, as has been noted above, need some such method as twin studies to elucidate their heritability component. When variation is continuous,

difference in the genetic component does not take the form of all-or-none distinctions such as concordant or discordant. One cannot state, for instance, that concordance for intelligence means that both twins have identical IQ scores, any other result being labelled discordant. Clearly, a case in which twins have scores of 104 and 115 is less discordant than a case in which they have scores of 100 and 125.

So Holzinger himself modified his formula given above to use the correlation coefficient (r) as the measure of relative concordance and discordance; thus

$$H = \frac{r_{MZ} - r_{DZ}}{1 - r_{DZ}}$$

A further slight modification of this formula was the basic method of the classic study comparing monozygotic twins reared together (MZt) and similar twins reared apart (MZa) by Newman, Freeman and Holizinger:

$$H = \frac{r_{MZt} - r_{MZa}}{1 - r_{MZa}}$$

Another frequently used form is:

$$\text{Heritability} = \frac{\text{Variance of dizygotics} - \text{Variance of Monozygotics}}{\text{Variance of dizygotics}}$$

If the monozygotics were always to present identical phenotypes, their variance could be zero, and the heritability one; in that case heredity is all. If, on the other hand the two types of twins vary to the same extent, the heritability is zero; in that case all of the observed variability must be environmentally determined.

21.6.5 Problems of Twin Studies

In addition to questions concerning the reliability of zygosity diagnoses, twin studies have been subject to a number of other criticisms. One is that even if twin study has demonstrated that a trait is 100 per cent inherited it would not show how it is inherited. Is it due to a single locus? Is there dominance? If it appears associated with another trait, is this due to linkage, and if so, to what degree?

Also attempts to study specific traits in this way have encountered many obstacles of a technical nature. Many scientists disagree, for example, on the propriety of using the typical IQ tests as measures of inherited intelligence, especially across class and cultural boundaries. However, we can conclude that by studying twins we can obtain some evidence for the involvement or noninvolvement of genes in various traits. But we must be careful in interpreting the results of such a study since the variation in penetrance can complicate the results.

21.7 SUMMARY

In this Unit you have studied the quantitative traits and genetics of twins. More specifically we talked about the following:

- The traits which are determined by many genes are known as continuous traits or multifactorial traits and the study is known as quantitative genetics.
- The continuous traits like height, skin colour, and eye colour in human beings are expressed on a continuous scale.
- The trait exhibiting continuous variation, which are often controlled by two or more genes are termed as polygenic and where several genes, make additive contribution to the phenotype the trait is known as quantitative or continuous variation.

- To study the such traits statistical tools are used by geneticist.
- The effect of nongenetic factors on gene expression that include environmental effect and heritability is also considered i.e. $V_p = V_g + V_e$. The ratio of genotypic variance to the total phenotypic variance is called heritability. The heritability can be measured in several ways.
- The study of twins have played an important role in understanding genetics of traits and there have been several genetical inferences from twin studies. But there are also several problems in twin studies.

21.8 TERMINAL QUESTIONS

1) Kernal colour in wheat is determined by the action of two pairs of polygenes that produce colours varying from dark red to white. If AABB (dark red) and aabb (white) are crossed. What fraction of the F_2 generation can be expected to be like either parent?

2) Height in a certain plant species is controlled by two pairs of independently assorting alleles, with each participating allele A or B adding 5 cm to a base height of 5 cm. A cross is made between parents with genotype AABB and aabb. Disregarding environmental influences.

   a) What is the height of each parent?

   b) What is the expected height of the members of F_1 generation?

   c) What is the expected phenotypic ratios in the F_2 generation?

3) Put (√) mark on the correct answers.

   In a polygenic interaction:

   a) each contributing gene in a series produces an equal effect.

   b) there is no dominance

   c) there is no epistasis among genes of different loci.

   d) there is no linkage.

   e) all of the above

4) The cross between two plants of equal heights results in a progeny with five different phenotypes. How many pairs of polygenes were involved.

   a) 2

   b) 3

   c) 4

   d) 5

5) a) A _____ is a phenotype produced by the environment which simulates the effects of a known mutation.

   b) Intelligence is somewhat influenced by the environment, therfore, its heritability is _____

6) i) _____ is the measure of the degree to which a phenotype is genetically determined and environmentally influenced.

   ii) The study of human inheritance is not always possible because individual mating produces _____ number of offsprings.

   iii) The _____ curve is widely used to describe continuous variables.

7) Two homozygous varities of *Nicotiana longiflora* were crossed to produce F_1 hybrids. The average variance of corolla length for all three populations was 8.76. The variance of the F_2 was 40.96. Estimate the heritability of flower length in the F_2 population.

8) One thousand pairs of twins attend a twin convention in Minneapolis. Eight hundred pairs of these twins are of like sexes (Male/male and female/female). The remaining two hundred pairs are of unlike sexes (Male/female). What percentage of these twins is monozygotic and what percentage is dizygotic?

21.9 ANSWERS

Self Assessment Questions

1) Average phenotypic effects of A locus are given as AA = 4, Aa = 3 and aa = 1. Similarly average phenotypic effects of B or C loci are BB or CC = 4, Bb or Cc = 3 and bb or cc = 1. Using the branching method, phenotypes of a cross between two triple heterozygotes will be as follows:

					Genotype		Phenotype
				1CC	= 1AABBCC	=	1 (12)
		1	BB	2CC	= 2AABBCc	=	2 (11)
				1cc	= 1AABBcc	=	1 (9)
				1CC	= 2AABBCC	=	2 (11)
1	AA	2	Bb	2Cc	= 4AABbCc	=	4 (10)
				1cc	= 2AABbcc	=	2 (8)
				1CC	= 1AAbbCC	=	1 (9)
		1	bb	2Cc	= 2AAbbCc	=	2 (8)
				1cc	= 1AAbbcc	=	1 (6)
				1CC	= 2AaBBCC	=	2 (11)
		1	BB	2Cc	= 4AaBBCc	=	4 (10)
				1cc	= 2AaBBcc	=	2 (8)
				1CC	= 4AaBbCC	=	4 (10)
2	Aa	2	Bb	2Cc	= 8AaBbCc	=	8 (9)
				1cc	= 4AaBbcc	=	4 (7)
				1CC	= 2AabbCC	=	2 (8)
		1	bb	2Cc	= 4AabbCc	=	4 (7)
				1cc	= 2Aabbcc	=	2 (5)
				1CC	= 1aaBBCC	=	1 (9)
		1	BB	2Cc	= 2aaBBCc	=	2 (8)
				1cc	= 1aaBBcc	=	1 (6)
				1CC	= 2aaBbCC	=	2 (8)
1	aa	2	Bb	2Cc	= 4aaBbCc	=	4 (7)
				1cc	= 2aaBbcc	=	2 (5)
				1CC	= 1aabbCC	=	1 (6)
		1	bb	2Cc	= 2aabbCc	=	2 (5)
				1cc	= 1aabbcc	=	1 (3)

The results can be summarised as follows:

Category	Phenotypic value	No. of Individual
1	12	1
2	11	6
3	10	12
4	9	11
5	8	12
6	7	12
7	6	3
8	5	6
9	4	0
10	3	1
		64

2 A relatively higher correlation between mothers and daughters indicates a higher tendency for the daughters to resemble their mothers as compared to a lower correlation. A higher correlation of 0.49 for height as compared to that for blood pressure (0.21) and serum cholestrol (0.28) is indicative of the fact that there will be greater resemblance between the daughters and mothers for height than that for blood pressure and cholestrol. Similarly, daughters will have a somewhat higher tendency to resemble their mothers for serum cholestrol than that for blood pressure.

Terminal Questions

1) This problem is based on the principle of polygenic inheritance. Polygenic inheritance differs from the classical mendelian pattern in that the whole range of variation is covered in a graded series. In polygenic inheritance certain assumption are made.

   1) Each contributing gene in the series produces an equal effect.
   2) Effects of each contributing allele are cumulative.
   3) There is no dominance.
   4) There is no epistasis among genes of different loci.
   5) No linkage is involved.
   6) Environmental effects are either absent or may be ignored.

   a) If we symbolize the gene for red with capital letter A and B, and the alleles resulting in lack of pigment production by *a* and *b* the cross can be diagrammed like this:

P : AABB × aabb

darkred white

Gametes AB × ab

F_1 Aa Bb
intermediate red
F_2

	AB	Ab	aB	ab
AB	ABBB	AABb	AaBB	AaBb
Ab	AABb	AAbb	AaBb	Aabb
aB	AaBB	AaBb	aaBB	aaBb
ab	AaBb	Aabb	aaBb	aabb

Assuming each capital allele increases the depth of colour equally, we can classify the F_2 generation in this way:

Number of Genes for Red	Genotype	Phenotype	Fraction of F_2
4	AABB	dark red	1/16
3	AABb, AaBB	medium red	4/16
2	AAbb, aaBB, AaBb	intermediate red	6/16
1	aaBb, Aabb	light red	4/16
0	aabb	white	1/16

We can see that 2/16 of the F_2 generation resembles either parent from P generation, 1/16 of white and $\frac{1}{16}$ of dark red.

2) Base height = 5 cm

Since each allele contributes an additional 5 cm, we use the following formula:

Total height = (each effective allele × 5 cm + base height)

a) Height of AABB = (4 × 5 cm + 5 cm)

Height of AA BB = 25 cm

Height of aabb = 0 + 5 cm

Height of aabb = 5 cm

b) P : AABB × aabb

Gametes AA × ab

F_1 AaBb

Height of AaBb = (2 × 5) + 5

= 15

	AB	Ab	aB	ab
AB	ABBB	AABb	AaBB	AaBb
Ab	AABb	AAbb	AaBb	Aabb
aB	AaBB	AaBb	aaBB	aaBb
ab	AaBb	Aabb	aaBb	aabb

The genotypes and phenotypes of the F_2 can be arranged in tabular form.

Genotype	Number of genes for Height	Fraction of F_2	Height
AABB	4	1/16	25
AABb, AaBB	3	4/16	20
AAbb, aaBB, AaBb	2	6/16	15
aaBb, Aabb	1	4/16	10
aabb	0	1/16	5

3) e)

4) a)

5) a) phenocopy

b) high

6) i) Heritability

ii) small

iii) bell shaped

7) Since the two parental varieties and the F_1 are all genetically uniform, their average phenotypic variance of the F_2 (Vt) is partly genetic and partly environmental. The difference ($V_E - V_e$) is the genetic variance (V_g)

$$h^2 = \frac{V_g}{V_t} = \frac{Vt - Ve}{Vt} = \frac{40.96 - 8.76}{40.96}$$

$$= 0.79$$

8) To solve this problem, we must simply think logically—which may or may not be easy. First, remember that males and females are born with equal probability (1:1 ratio). So dizygotic twins stand a 50% chance of being of unlike sexes. Monozygotic twins, on the other hand, must be of the same sex. We know that 200 pairs of twins are of unlike sexes—these twins must be dizygotic. That is half of the dizygotic twins. The other half must be of like sex. Remember that there is a 50% chance of being of unlike sex so there must be a corresponding 50% chance of being of like sex. This gives us 200 extra pairs of dizygotic twins. Therefore, a total of 400 pairs of twins are dizygotic. This leaves 600 remaining pairs of twins as monozygotic. Dividing 400 dizygotes and 600 monozygotes by 1000 twins and then multiplying by 100% we have the answers. The twins at the minneaplis convention are 60% monozygotic and 40% dizygotic.

UNIT 22 BEHAVIOURAL GENETICS

Structure **Page No.**

22.1 INTRODUCTION

This unit deals with an extremely interesting as well as very controversial subject, that is, the role of genetics in determining behaviour. Since the time of Mendel, people have been divided in their opinions regarding the role of genetics in controlling different forms of social behaviour.

The distinction between human and animal behaviour genetics is more than a matter of species studied or techniques feasible in the two fields. The animal geneticists study the nature of variables which determine behaviour while, the reasearches involved in human genetics concentrate on problems of social significance. Studies have been carried out in humans on some socially unacceptable behaviour like mental defects and psychiatric diseases, alcoholism and criminality. Pedigree analyses have been done to study if these traits were transmitted in families. At present, most geneticists and psychologists are of the opinion that both heredity and environment are important in controlling human behaviour.

Research in human behavioural genetics has concentrated upto now on problems such as : (i) mode of inheritance of mental dysfunction or other traits regarded as unit characters, (ii) the determination of quantitative traits such as intelligence and personality, (iii) the interaction of nature and nurture in development.

Objectives

After reading this unit you would be able to:

- know the methods and difficulties which are faced by Geneticists in determining control of behaviour genetics,
- discuss in an unbiased manner the behavioural characters which are genetical and those which are purely environmental,
- give examples of different approaches used in behavioural genetics,

- describe various examples of human behaviour which are genetically determined,
- comprehend the influence of cultural effects and race on certain behavioural traits.

22.2 BEHAVIOUR AND GENETICS

In broad terms, every action, reaction, and response represents a type of behaviour. Animals run, remain still or counter-attack in the presence of a predator, birds build complex to simple nests, fruit flies execute intricate courtship rituals and phototropism in plants. Humans reflexively avoid painful stimuli as well as "behave" in a variety of ways as guided by their emotions, intellect and cultural environment. Thus behaviour is defined, generally, as a reaction to stimuli or environment.

Even though clear-cut cases of genetic influence on behaviour were known as early as 1900s, the study of behaviour was of greater interest to psychologists, who were concerned with learning and conditioning, while some traits were recognised as innate or instinctive behaviour that could be modified by prior experience received the most attention. Such traits or pattern of behaviour were thought to reflect the previous environmental setting to the exclusion of the organisms genotype. This philosophy served as the basis of the **behaviourist school**. In 1950s, the so-called nature-nurture controversy flourished and by that time it became clear that while certain behavioural patterns, particularly in less advanced animals seemed to be innate, other were the result of environmental modifications limited by genetic influences. The latter condition is particularly true in more advanced organisms with more complicated nervous system. The prevailing view is that all behaviour patterns are influenced, both genetically and environmentally. The genotype provides the physical basis and/or mental ability essential to execute behaviour and further determines the limitation of environmental influences.

Behaviour genetics has blossomed into a distinct speciality within the larger field of genetics, as more and more behaviour patterns have been found to be under genetic control. This unit will provide you an overview of the role of genes in behaviour.

22.3 STUDY OF BEHAVIOUR GENETICS : METHODS AND DIFFICULTIES

In 1965, David Merrell summarised the three major approaches used in studying behaviour genetics. The first involves the determination of behavioural differences between genetic strains of the same species or between closely related species. If such closely related organisms exist in similar environments and their survival needs are identical, any observed behavioural differences may be correlated with genetic differences. In the second approach, a modified behavioural trait is selected from a heterozygous population. If such a modified trait can be established in a new genetic strain, the positive influence of the genotype is established. The first two approaches identify behavioural patterns as being under the control of genes. Genetic investigation through controlled breeding experiments can then be performed in an attempt to establish the inheritance pattern.

The cases where these approaches have been successfully used, it has only been shown that the inheritance is not due to simple mendelian pattern. Instead, the behavioural traits have been attributed to polygenic inheritance or quantitative inheritance.

The third approach is infrequent but today it is extensively used. This comprises the study of the effects of single gene on behaviour. While a single complex behavioural trait may be controlled by many genes, disruption of the trait by a single mutation may occur and be analysed. This approach is more informative because it gives the most objective information concerning the role of genes in behaviour.

There are several difficulties which the geneticist faces in experimentation. Most important is the problem of environmental control. The study of the effect of genes on behaviour requires a constant environment in order to establish the role of genes and with human, there is no hope of precisely controlled studies.

Then comes the difficulty involved with the objective measurement. The quantitative standards of data collection possible in transmission genetics or molecular genetics, for example, cannot be approached in behaviour genetic studies. Therefore, the more simple the trait, the more objective and accurate is the analysis.

The most difficult part is study of any type of behaviour which involves learning and reasoning. Most of behavioural study in higher organisms, including humans, centers around socially significant traits. These traits are influenced greatly by environment and studied in organisms whose environments are most difficult to control. As a result, clear-cut single gene influences are very little known about these types of behaviour.

22.4 COMPARATIVE APPROACHES IN STUDYING BEHAVIOUR GENETICS

Here, you will study the two approaches to understand behavioural genetics. The first approach is a comparative study of behaviour in closely related strains. This will be explained by example of alcohol preference in mice and mate selection in baboons. The second study involves selection within heterozygous population for a modification of behaviour. This can be explained by study of learning in rats, geotaxis in **Drosophila**, and nest building in lovebirds.

22.4.1 Alcohol Preference in Mice

Several studies of **alcohol preference** in mice have been reported. Different inbred strains have been compared for preference or aversion to ethanol, with all studies indicating genetical control for this response.

For example, Rogers and McClearn compared four strains of inbred mice over a period of three weeks. Each strain was presented with seven vessels containing either pure water or alcohol varying in strength from 2.5 to 15.0 per cent. Daily consumption was measured. Table 22.1 shows the proportion of absolute alcohol to the total liquid consumed on weekly basis. When the data were examined, it was found that C 57 BL and C 3 H/c strains exhibit a preference for alcohol, while BALB/c and A/3 demonstrate an aversion towards it. Since the environment in raising the mice over many generations has been constant, the preference differences are attributed to the genotype of each strain. Presumably, these strains contain identical sets of genes and vary only by the fixation of different alleles at these loci.

Table 22.1: Alcohol Consumption in Mice

Strain	Week	Proportion of Absolute Alcohol to Total Liquid	X
C57BL	1	0.085	0.094
	2	0.093	
	3	0.104	
C3H/C	1	0.065	0.069
	2	0.066	
	3	0.075	
BALB/c	1	0.024	0.020
	2	0.019	
	3	0.018	
A/3	1	0.021	0.017
	2	0.016	
	3	0.015	

Source : From Rogers and McClearn, 1962.

In other similar studies, crosses between strain differing in alcohol preference have been performed. The results, have not established any simple inheritance. However. It appears that two or more genes may be involved. It has been suggested that alcohol preferences may be related to enzyme activity of the liver. While some differences between strains have been noted in the levels of alcohol dehydrogenase (ADH) and acetaldehyde dehydrogenase (ACDH), the results are still inconclusive. Both enzymes are involved in the metabolism of alcohol.

22.4.2 Baboon Mate Selection

Baboons are large quadrupedal terrestrial monkeys found in Africa. These primates exhibit a variety of social behavioural patterns and such patterns have been thoroughly investigated. They are classified into one genus, *Papio* and several species. *Papio anubis* is a savanna-dwelling group while *Papio hamadryas* is a desert dweller. Their ranges sometimes overlap.

P. hamadryas has a social system that includes harems. Each male is associated permanently with one or more females. *P. anubis* displays the more common social structure of baboons, in which permanent associations are not formed. Instead, matings are promiscuous but affected by dominance ranking of the males.

In the overlapping range of their habitation, which is an arid region resembling that of the *hamadryas* more than that of *anubis, hamadryas* male have been observed to kidnap juvenile *anubis* females. Then these females learn to cooperate with the herding behaviour of males in harem life. Hybrid males resulting from such mating were not so efficient in herding females as the pure *hamadryas.* Since both *hamadryas* and hybrid males appear to have the same opportunity to learn the herding behaviour, it has been concluded that herding and harem formation are at least partly genetically determined.

Now, you will study the second major approach in behaviour genetics, that is selection for behaviour modifications from a heterozygous population - leads to the production of inbreds.

22.4.3 Geotaxis in *Drosophila*

The investigation of geotaxis in *Drosophila* illustrates the general behaviour response as well as the selection technique. Before coming to experiment we must know what taxis means. In general, the movement of free living organism towards the source of an external stimulus is known as taxis. The stimuli may include chemicals, gravity and light. The geotaxis is response towards gravity.

Jerry Hirsch and his colleagues designed a mass screening device that allows about 200 flies to be tested per trial (Fig 22.1). The maze is placed vertically, a fluorescent light illuminates the side opposite the entry point, and flies are added. The flies those who continue to travel upward at each function finally reach the top and those who travel downward reach the bottom, and those making both decisions of going "up" and "down" along the way finally settle somewhere in the middle. It was found that flies could be selected for both positive and negative geotropism, establishing the genetic influence on the behavioural response.

As you can see in Figure 22.1, mean scores may vary from +4.0 to -6.0, corresponding to the number of T junctions the fly encounters going up and down. By observing the data, it is clear that extreme negative geotropism is stronger than the extreme positive geotropism. As the two selected lines were compared over many generations clear but fluctuating differences were observed. Such results indicate the additive effects of polygenic inheritance.

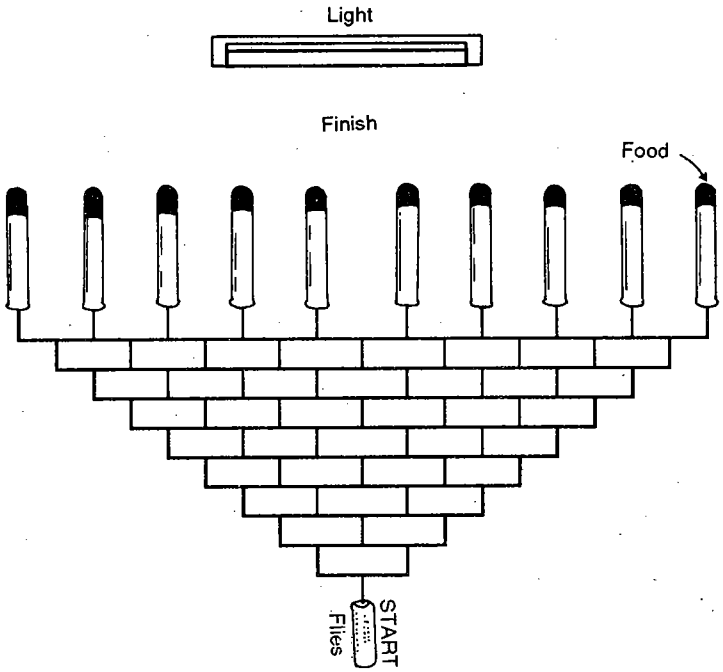


Fig 22.1: Schematic drawing of a maze used to study geotaxis in *Drosophila*

22.4.4 Nest Building in Lovebirds

One important example of the modification of behaviour in crosses between closely related organisms involves the genus Agapornis (lovebirds) in the African parrot family.

These birds can be easily trained and are often used in circus act.

In 1962, Willian C. Dilger reported on a nest building study involving hybrid birds resulting from a cross between two species *Agapornis fischeri* and *Agapornis roseicollis.*

The females of *A. fischeri* construct nests by materials tucked on back feathers and in addition twigs carried in beak while females of *A. roseicollis* only tuck material between their back and rump feathers to transport it to the nest. But the hybrid females behave in very confused manner. They always attempt to tuck the material but unsuccessfully.

This behaviour is something to behold. The hybrid birds may attempt to tuck material between their back but then fail to release the material in the tail feathers and sometimes dropping it. Sometimes inappropriate material may be tucked. These birds also have difficulty in determining how to get their bills back to their rump feathers and sometimes are seen running backwards in attempting this motion. Thus, their only real success at nest building is by carrying materials in their nests.

However, they continue to carry the material to their nests. In an initially only 6 per cent of the attempts are successful. But after two months the percentage grows upto 41 per cent. Eventually, after three years, the hybrid can carry the material to their nests successfully. This time span is relatively very short interval in which they can be trained for any other tasks. Thus, it is obvious that nest building behaviour is under genetic control, and the lack of specific genes is responsibles for the confused behaviour. Both types of instinctual behaviour are inherited, but incompletely. Over time, learning compensates for this confusion so that nests can be built.

Apparently, the F_1 hybrid birds are sterile, thus further generations are unavailable for study. On the basis of the F_1 data, it has been proposed that the nest building characteristics are under polygenic control. However, the data are insufficient to establish this hypothesis with absolute certainity.

22.5 SINGLE-GENE ON BEHAVIOUR

Most definitive information on the genetic influence on behaviour has come from the study of the effect of single gene on behaviour. In more recent studies often, induced mutations are analysed in order to infer general principles as to how normal behaviour is created and regulated.

There are obvious advantages of this approach. In the laboratory studies, by and large the environmental effect on behavioural response in minimised or eliminated. Thus, the genetic influence is more straight forward and easier to define than in studies where the environment is a major factor.

Therefore it is possible to dissect a behavioural pattern into its components. There are large number of examples of single gene-influences, but we will restrict ourselves to single example of nest-cleaning behaviour in honeybees.

22.5.1 Nest Cleaning Behaviour in Honeybees

Nests of honeybees are often infected with *Bacillus* larvae the agent causing American foulbrood disease. But the disease is counteracted by the **hygienic behaviour** of worker bees. The cells of infected combs containing afflicted larvae are opened, and the diseased organisms are removed from the hive. Hygienic hives are resistant to infection. While hives containing strains that do not display removal behaviour are susceptible to the diseases.

In 1964, Walter Rothenbubler crossed a hygienic (Brown) line with a nonhygienic (van Scoy) line. This work strongly favours the hypothesis that two recessive independently assorting genes (m and r) or a gene complex are responsible for hygienic behaviour.

The F_1 hybrids were all nonhygienic. However, when F_1 drones were back-crossed to hygienic queen four phenotypes were produced in roughly equal proportions as shown in Figure 22.2. While one group is hygienic other group is unhygienic. The other two group have interesting feature. The drones in one group could uncap cells but were unable to remove infected larvae. Drones in the other or the fourth group which was thought to be unhygienic, were able to remove larvae if the cells were artificially uncapped, but they were unable to uncap the cells.

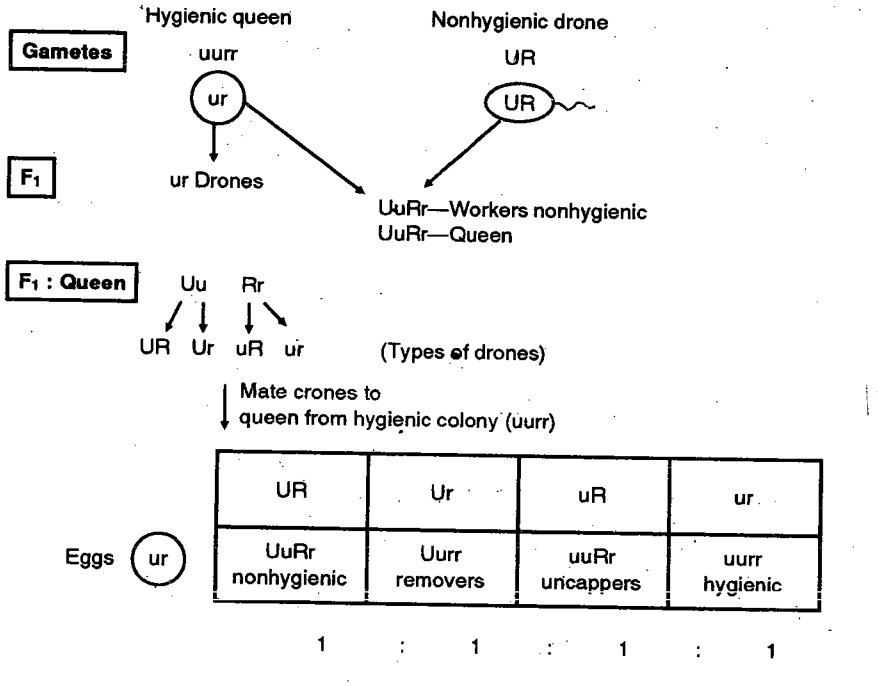


Fig 22.2: Result of a honeybee cross between hygienic diploid female and nonhygienic haploid male.

It seems that one gene pair (u/u) or a linked complex of genes determines uncapping behaviour and a second gene pair (r/r) or complex determines removal ability. Nest cleaning in honeybees is also one of the most striking examples of the far- reaching effects of genes on behavioural responses.

SAQ 1

A fully heterozygous ($u^+ u r r^+$) nonhygienic queen bee was mated with a hygenic drone (ur). Remember that queen bees are diploid and drones are haploid, what genotypes and phenotypes (uncapping of cells and removal of dead pupae) might be expected in the first generation and in what proportions.

.....

.....

.....

.....

.....

.....

.....

.....

22.6 GENETICS OF HUMAN BEHAVIOUR

The genetic aspect of behaviour in humans is much more difficult to characterise than that of any other organisms. The first constraint is the unavailability of humans as experimental subjects in genetic investigations. Secondly, the types of responses considered to be interesting behaviour are extremely difficult to study. Intelligence, language, personality or emotions are the most popular forms of behaviour to be studied. But there are two problems in examining such traits. First, all are difficult to define objectively and to measure quantitatively. Second, they are the traits most affected by the environment. In each case, while there is undoubtedly a genetic basis, it is a complex one. Furthermore, the environment is extremely important in shaping, limiting or facilitating the final phenotype for each trait.

Several studies of human behaviour have been performed by psychologists without proper input from the biologists or geneticists. The other hampering cause is that traits involving intelligence, personality and emotions have the greatest social and political significance. As such, these traits are more likely to be the subject of sensationalism when reported to the lay in public.

Darlington in 1963 wrote, "Human behaviour has thus become happy hunting ground for literary amateurs. And this is the reason that psychology and genetics, have failed to face the task together but since then great progress has been made in bridging the gap."

Because their study comes closet to infringing upon individual liberties such as the right to privacy, these traits are the basis of the most controversial investigations.

22.6.1 Genetical Basis of Mental Disorders

Even in humans, the genetic basis is known for certain types of behaviour. One of the most prominent example is Huntington's disease (sometimes known as Huntington's chorea). Inherited as an autosomal dominant disorder, it affects the nervous system, including the brain. Onset of the disease occurs usually in one's fifties with a loss of motor function and coordination. Degeneration of the nervous system is progressive, and personality changes occur. The affected individual is soon unable to care for himself. Since onset is usually after a family has been started, all children of an affected person must live with the knowledge that they face a 50 per cent probability of developing the disorder. Only recently molecular genetic techniques have made possible an early diagnosis of the disease.

One of the most bizarre behaviour disturbances in the humans is due to a single gene defect, which has a sex linked recessive mode of inheritance. This disease is a type of cerebral palsy, and is known as the Lesch-Nyhan syndrome in which afflicted children—all male—develop a compulsion for self-mutilation. It is characterised by physical and mental retardation. This disease developes within the first year of life and is most often fatal in early childhood.

Other metabolic disorders are also known to affect mental functions. For example, Tay-Sachs disease, an autosomal recessive disorder also known as amaurotic idiocy, involves severe mental retardation among other phenotypic characteristics. Another autosomal recessive disease, **phenylketonuria**, unless detected and treated early, results in mental retardation. **Porphyria**, which is under the control of an autosomal dominant gene, is marked by recurring periods of dementia. All these disorders alter the normal biochemistry of the affected individuals.

Chromosome abnormalities also produce syndromes with behavioural components. Down Syndrome (trisomy 21) result in mental retardation. Both Klinefelter syndrome (XXY) and Turner Syndrome (XO) may also result in diminished mental capacity.

SAQ 2

Choose the right option.

i) All males develop a compulsion for self mutilation in

   a) Amaurotic idiocy
   b) Phenylketonuria
   c) Lesch-Nyhan syndrome
   d) Down Syndrome

ii) Usually in one's fifties there is a loss of motor function and coordination and degeneration of the nervous system is progressive in

   a) Down syndrome
   b) Porphyria
   c) Huntington's disease
   d) Turner syndrome

iii) Tay-Sachs disease an autosomal recessive disorders also known as

   a) Lesch-Nyhan syndrome
   b) Down syndrome
   c) Amaurotic idiocy
   d) Klinefelter syndrome

iv) Another autosomal recessive disease, unless detected and treated early, results in mental retardation. It is known as

   a) Amaurotic idiocy
   b) Turner syndrome
   c) Klinefelter syndrome
   d) Phenylketonuria

22.6.2 Human Behaviour Traits with Less Defined Genetical Basis

Some other aspects of human behaviour, notably schizophrenia and manic depressive illness have been extensively studied. Investigations indicate that genetics has a role to play in mental disorder but the environment also plays a substantial role.

Here, we will discuss schizophrenia only. This mental disorder is characterised by withdrawn, bizarre and sometimes delusioned behaviour. It has a familial occurrence, with first degree relatives of schizophrenics having a much higher incidence of this disorder than the general population. The use of twin studies has improved the accuracy of our understanding of the problem. The concordance of schizophrenia seen to be higher in monozygotic twins than in dizygotic twins. There is enough evidence to suggest that the condition is influenced by polygenes.

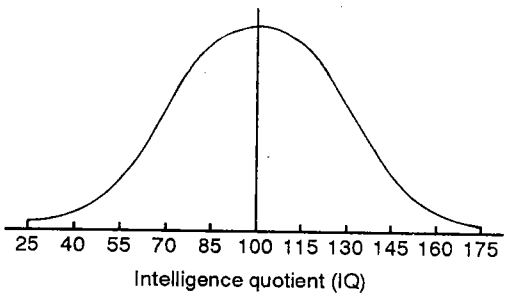
Biochemical analysis os Schizophrenics have shown that the patient have less than normal level of the enzyme **monoamineoxidase**.

It can be concluded that an individual may be endowed with a genetic predisposition for abnormal behaviour like schizopherinia and that environmental factors can serve to tigger its onset.

22.7 GENETICS AND I.Q.

Various studies indicate that there is relation between genetics and intelligence. The environment undoubtedly has a profound effect on the IQ. There is a long standing controversy about difference in intelligence between races.

But first we must define the I.Q. **Intelligence is defined as the ability to learn or understand.** It is the requirement that is most important for achievement in the schools of the Western world the IQ (**intelligence quotient**) (Fig. 22.3) tests have



Range	Descriptive classifications
86-115	Normal, average
71-85	Slow learner, borderline, moron
56-70	Mildly subnormal, mildly retarded, feebleminded
41-55	Retarded
26-50	Severely retarded, imbecile
0-25	Profoundly retarded, idiot
116-130	Bright
131-145	Gifted
146-200	Genius

Fig 22.3 : The IQ curve is based on half the population being above and half below mean IQ of 100. For the test question of varying difficulties are selected so that a large tested population (e.g. a rural area, an ethinic group, a poverty area) can be used to make a normal distribution curve. The 15 point partion are used on either side of the mean to define many categories of mental ability.

been used as a basis to measure intelligence. Wheather or not intelligence can be measured fairly, or even at all, in all populations by such tests is debatable. Attempts have been made to develop tests that measure intellectual ability cross-culturally. The analysis of various test scores of different groups of people and the comparision of the performances of siblings and twins have shown that the abilities measured by these tests have a hereditary basis.

The environment in which a person is raised and studied had profound effect on the test scores. The human-beings can never be totally separated from environmental effects and can never be placed in controlled experimental studies. Thus it is extremely difficult to determine the exact genetic basis of intelligence. However, several studies clearly state that both the environment and genetics can influence the IQ performance.

Twin studies have indicated that inheritance is important in the acquisition of intelligence in a child. Several studies have shown that closer the degree of

relationship, the higher the correlation coefficient. The importance of genetic component was realised, because twins (monozygotic that were reared apart as well as those that were reared together) showed a high correlation coefficient (a high coefficient indicates heritability).

Studies on adopted children show that the environment has a large role in the development of the type of intelligence that is measured by tests. IQ scores of adopted children show a higher correlation with those of their biological mother than with those of their adoptive mother.

In 1965, R.B. Cattell defined two components of general ability. These components encompass both hereditary and environmental influences. The first he called **fluid intelligence**, which is a function of brain development. This type of intelligence is completed by the age of 14 when the brain stops growing. It can theoretically be measured by tests that are not culturally and theoretically biased. The second component is **crystallised intelligence**. This component is a function of training and education. Unlike fluid intelligence, it can change after the completion of brain growth.

Arthur Jensen concluded in his studies that the white people do better than the black people on the IQ test because of a hereditary difference. He concluded that one half to three quarters of the difference between IQ was due to genetic factors. He was unable to prove his theory, but it was also not disproved because the experimental situation needed is unavailable and probably unobtainable.

But several reports have been published to dispute his claims. For instance, the difference in the IQ scores of northern groups of the whites and blacks is less than the difference between southern populations. Another main cause of the difference in IQ is that generally whites and blacks segregate into separate residential areas and hence schools. The schools that teach the black student are less adequate than those that teach white students, because mostly the black live in economically deprived area. Thus, the same number of years in school does not amount to the same level of training. Studies also show that blacks score higher than the whites when tested by a black. Some other variables such as diet, cultural beliefs and societal attitudes, interplay to such an extent that Jensen's hypothesis has been almost completely discredited.

This problem would have been resolved long ago if these studies were done on laboratory animals, because their genotypes can be replicated and their environment can be controlled. Human population have been morally barred from such studies hence concrete scientific proof is not available. However, it appear that environmental factors influence the extent to which the full genetic potential regarding intelligence is reached.

SAQ 3

Why Jensen's hypothesis has been completely discredited?

.....

.....

.....

.....

.....

.....

.....

22.8 SUMMARY

- Genetic basis of behaviour is a controversial subject. Both genetic and environmental components are involved in the expression of behavioural phenotype.
- Three major approaches have been used in behaviour genetics. First, studies of behavioural differences betweeen closely related organisms in similar environments, whose survival needs appear to be identical.
- The second approach is to identify in the laboratory for a modified behavioural trait and demonstrate that it is inherited.
- The third approach is a more direct appraisal of the role of genetics in behaviour, that is, to study the effects of single genes on behavioural patterns.
- Study of genetical influence on human behavioural traits is difficult, because environment also plays an important role in the development of these traits. But the difficulty is reduced when an abnormal behaviour such as Schizopherenia is often seen to run in families.
- I.Q. is a good example of a behavioural trait which can be regarded as a quantitative character with a wide dispersion.

22.9 TERMINAL QUESTIONS

1) Fill in the blanks with appropriate words.

   i) The movement by an organisms toward or away from a chemical is called.....

   ii) ..... is movement–oriented.

   iii) Flies that cannot orient themselves over vertical planes exhibit ..... geotaxis.

   iv) The *A. Fischeri* love birds carry nest materials on their back feathers and also carry twigs in their .....

   v) The specific cleaning behaviour of bees involves the ..... of the cells and the ..... of their content.

   vi) Different inbred strain of mice have been compared for ..... or ..... to ethanol with all study indicating ..... control.

   vii) ..... male resulting from the mating of *P. anubis* and *P. hamadryas* were not so efficient in herding females as are pure .....

   viii) Schizophrenia is believed to be ..... influenced.

2) State whether the following statement is true or false.

   i) Flies that orient themselves on vertical planes exhibit positive geotaxis. [ ]

   ii) *Agapormis roseicollis* love birds build their nests by carrying strips of materials on their rump and in their beak. [ ]

   iii) Hygienic behaviour in bees is controlled by two linked genes. [ ]

iv) Regardless of one's culture, the same I.Q. test is a reliable measure of intelligence. [ ]

v) Most of the expression of intelligence is controlled by the environment. [ ]

vi) The hereditary component of intelligence can be determined by comparing I.Q. scores of two separate populations with different cultural background. [ ]

vii) Huntington's disease is inherited as an autosomal dominant disorder, and affects the nervous system including brain. [ ]

viii) Phenylketonuria, autosomal recessive disease unless detected and treated early results in mental retardation. [ ]

3 i) If a haploid honeybee drone of the genotype uR is mated to a queen of genotype UuRr, what ratio of behavioural phenotypes will be observed in the hive by the offspring with respect to American foulbrood disease?

ii) For uR x Uurr?

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

22.10 ANSWERS

Self Assessment Questions

1)

	u^+r	u^+r^+	ur	ur^+
ur	u^+urr	uu^+rr^+	.uurr	$uurr^+$
	nonhygenic	nonhygenic	hygenic	nonhygenic

3 : 1

u^+ u r r^+

u^+r u^+r^+ ur ur^+

a) There will be four types of genotype u^+urr, uu^+rr^+, uurr, $uurr^+$

b) There will be two types of phenotypes-3 nonhygienic and 1 hygienic type

2) i) c

ii) c

iii) c

iv) d

3) Jensen's hypothesis was completely discredited because, for instance,

i) Difference in I.Q. scores of the northern groups of the whites and blacks is less than the difference between the southern populations.

ii) The white and black populations are generally segregated into separate residental areas and hence schools. The schools of the blacks are less adequate than those of the whites.

iii) Studies have shown that when tested by the blacks they score higher than when tested by the whites.

iv) There are some other variables such as diet, cultural beliefs and societal attitudes.

Terminal Quetions

1) i) Chemotaxis

ii) Geotaxis

iii) Negative

iv) Beaks

v) Uncapping, Removal

vi) Preference, aversion, genetical

vii) Hybrid, hamadryas

viii) Polygenically

2) i) True

ii) False

iii) False

iv) False

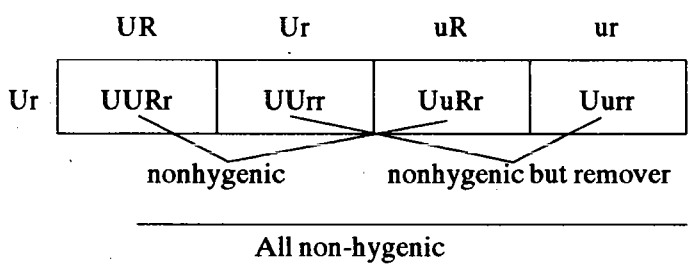
v) False

vi) False

vii) True

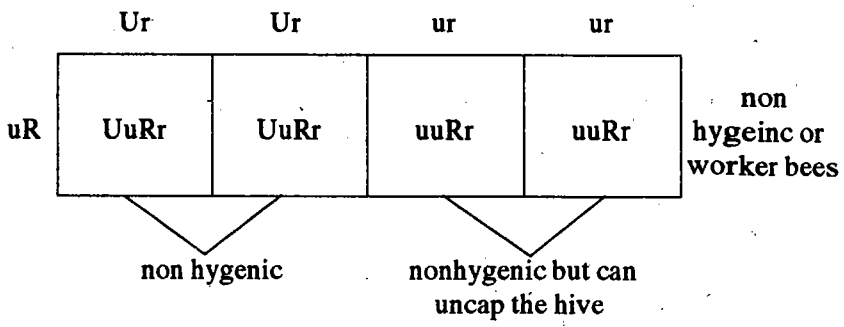
viii) True

3) i)



Phenotypically all will be nonhygenic

ii)



UNIT 23 GENETICS AND HUMAN WELFARE

Structure **Page No.**

23.1 INTRODUCTION

Genetics, one of the modern branches of biology can be effectively applied for identifying the various inherited disorders and more specifically for preventing the recurrence of the disorder in the human populations. Nature has contrived it in such a way that most the lethal genetic disorders do not surface because the affected embryos or foetuses are naturally aborted before the complete development takes place. This is true of many of the chromosomal disorders. It is also true that several genetic disorders, especially arising out of the abnormalities in the structure and number of chromosomes result in the sterility of the individuals concerned so that the disorders are not transmitted to the subsequent generations. We discussed in Units 9 and 10 of Block 2 the abnormalities in chromosomal structure and number, and said that many abnormalities result in mental retardation and decreased I. Q. of the individuals. Such individuals are less likely to be married and therefore have no chance of contributing their defective genes to the gene pool. Even if they were to marry, they tend to raise fewer children. You could observe from these statements that nature has devised a number of ways by which it could selectively prevent the spreading of the lethal and sublethal genes in the population. Nevertheless, the genetic disorders do occur and there is a need to diagnose and arrest their transmission to the subsequent generations. In other words, ways and means should be found to prevent the burdening of the population with unwanted genes, a phenomenon come to be known as **genetic load**. The study of genetics must be used as an effective tool to promote the human genetic welfare.

This unit aims at the study of the genetic counselling phenomenon for the promotion of human welfare. Further we also discuss at length one of the recent advances in molecular genetics, namely the recombinant DNA technology. With the advent of this technique the science of Genetics has opened new vistas for promoting human welfare. The technology has wide possibilities in areas of

medicine, agriculture as well as industry. We shall discuss the applications of the recombinant DNA technology, otherwise known as genetic engineering technique in various fields related to human welfare.

Objectives:

After studying this unit you should be able to:

- define the term genetic counselling
- distinguish between informative counsellng and supportive counselling
- discuss the concept of genetic diagnosis with special reference to amniocentesis and chorionic villus biopsy
- explain the principle behind the recombinant DNA technology and
- list the applications of genetic engineering technique.

23.2 WHAT IS GENETIC COUNSELLING

One of the recent trends in health care and human welfare is genetic counselling. The main aim of the counselling process is to identify individuals or families who may run a high risk for genetic disorders. Such genetic disorders may arise because of the exposure of individuals to certain special circumstances (such as radiation effects) or due to the occurrence of the disorders among the members of the family or the relatives. In short **genetic counselling can be defined as a communication process which deals with the human problems associated with the occurrence or the risk of occurrence of a genetic disorder in a family**. (Definition provided by a group of genetic counsellors—Committee on genetic counselling 1975. Genetic counselling, *American Journal of Human Genetics 27*, 240-242). By this process a trained person tries to help an individual or a family to understand the various aspects of the genetic disorder, the contribution of heredity to the disorder, the probability of its future occurrence in the family and the possible steps to be taken for avoiding the recurrence of the disorder in the family.

Genetic counselling has several components in it. The first component is the diagnosis of the genetic disorder. The diagnosis itself is done by the medical specialists after performing certain special tests that are available. Once the diagnosis is done, the next step in the counselling process is the preparation of the pedigree of the affected individual. Once the pedigree chart is made and analysed it would throw light on the type of disorder—either it is an autosomal or sex linked trait or whether it is a dominant or a recessive trait. The pedigree chart is drawn essentially based on the information obtained by way of interviews with the family members and the information obtained from the laboratory tests.

On the basis of the conclusion arrived by the above methods, the counsellor would advise the individual or the family on the risks involved in their children inheriting the genetic disorder. They are told of the probability that a child will not be affected and other courses of action if any. Such a counselling process is designated as **informative counselling**. Another aspect of the counselling process, the **supportive counselling** would provide the affected individuals professional help that may be needed to understand the intensity of the problem. For this purpose the counsellor should have had training in more specific areas of genetics.

23.2.1 Diagnosis through Genetic Counselling

For the diagnosis of any possible genetic disorder in unborn babies the extensively used method is the **prenatal diagnosis**. This diagnosis is done on samples of amniotic fluid (Fig. 23.1) collected from the mother and the process of collection is

called **amniocentesis**. The collection of the sample is done after the 14th week of pregnancy and before the 20th week. The fluid contains the foetal cells which are cultured in the laboratory and subjected to the karyotype analysis. Amniocentesis and subsequent karyotyping are the most commonly used tests in prenatal

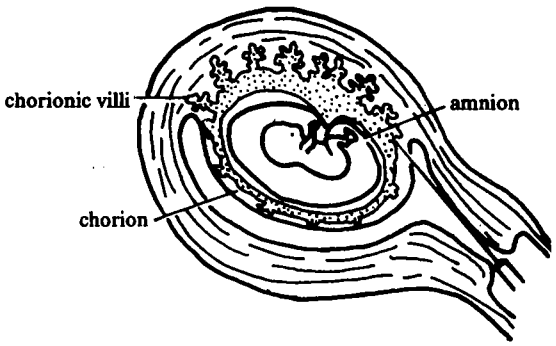


Fig. 23.1: Figure showing the chorionic villi, the amnion and the amniotic fluid surrounding a human foetus.

diagnosis. The karyotype analysis will reveal the structural or numerical chromosomal abnormalities if any that the developing foetus has. In case an abnormality is diagnosed (such as Down syndrome) the counsellor may advise the couple to have the pregnancy terminated.

A relatively less commonly used technique for the diagnosis is the **chorionic villus biopsy**. In this technique a sample of chorionic villi surrounding the foetus is removed and the karyotype of the cells is determined. In such a diagnosis there are possibilities for an error because certain differences may arise postzygotically between the foetal cells and the placental cells. It is said that in certain cases normal embryos have been aborted based on the abnormal karyotype of the chorion. Contrarily a normal chorionic karyotype may be associated with an abnormal foetal karyotype.

In general the prenatal diagnosis is done only in high risk pregnancies. For chromosomal abnormalities the high risk indicators are the advanced maternal age, prior birth of a child with a genetic disorder and the information that one of the parents has abnormal chromosomal rearrangement that can generate meiotic mispairing.

SAQ 1:

State whether the following statements are true or false.

1. Pedigree charts throw light on the type of genetic disorder–either it is an autosomal dominant or recessive trait or a sex linked dominant or recessive trait. True/False.
2. Supportive counselling provide the affected individual professional help needed to understand the intensity of genetic disorder. True/False.
3. In aminocentesis there are possibilities for an error because certain differences may arise postzygotically between foetal cells and the placental cells. True/False.
4. Prenatal diagnosis is done only in low risk pregnancies. True/False

23.2.2 Genetic Screening

Genetic screening has been defined as a "systematic search in a population for persons of certain genotypes". Genetic screening is a useful concept in that, the occurrence of many of the inherited diseases could be prevented, if parents at risk could be diagnosed and treated before they plan to have the child or if the affected new borns are identified and treated before degenerative changes occur. Recessive

genotypes are the ones that need to be screened rigorously since these alleles lie usually hidden and unsuspected.

Genetic screening is usually done at three levels: i) identifying the potential parents ii) identification by prenatal screening iii) For identification of the affected new borns.

Screening Potential Parents:

The genetic diseases caused by the recessive alleles essentially arise because of the marriage between the heterozygous parents. The only solution for such a problem is that by suitable pedigree charts the heterozygous individuals need to be identified. Marriage between such heterozygous individuals can be avoided. Alternatively, if the heterozygous individuals marry then they should avoid producing children. The third possibility is that they could have the pregnancies monitored to detect affected homozygotes. It should be made clear that the children do not run into the risk of inheriting a genetic disease unless both the parents are heterozygous for mutant alleles at the same locus.

Prenatal Screening:

There are several methods by which the genetic disorders could be detected in embryos and foetuses. Malformation if any could be detected by ultrasound techniques. The non-closing of the spinal chord, **spina bifida**, a high risk neural tube defect can be identified by high levels of α-foetoprotein (AFP) circulating in the maternal blood. We earlier mentioned the amniocentesis and chorionic villi biopsy as effective tools for identifying the structural and numerical chromosomal abnormalities.

Screening of New Borns:

Once the new borns have inherited the genetic disorder then every attempt must be made to prevent the disorder assuming serious dimensions. Any therapy should be done before irreversible physiological damage is done to the individual. In USA and Europe it is a legal requirement that the new borns be screened for hypothyroidism and phenylketonuria, the two diseases that could be treated effectively if detected very early.

23.2.3 Therapy for Inherited Diseases

Therapy for inherited diseases has taken various forms. Inherited defects in vision are treated by the use of various prosthetic devices such as eye glasses. New born children suffering from the inherited metabolic disorder galactosemia need to have milk omitted from their diet. Some diseases such as enzyme deficiences in liver may require drastic treatment such as liver transplants. Bone marrow transplants have become a common treatment for certain inherited diseases and the transplanted cells express the normal gene products.

It could be generalised that the transplantation of a genetically normal tissue into a patient with a defective gene function could be termed as gene therapy. But the term gene therapy is now come to be used in a more restricted sense namely it refers to the introduction of a specific gene into a patient to correct the genetic defect he suffers from. The technique of gene therapy revolves around the introduction of a vector carrying the normal gene into the cells of the affected individual. Many vectors such as viruses or plasmids are available for use with *E. Coli*, yeast, mamalian cells and plants. The entire technique of introducing a foreign gene into a host through a vector is come to be known as the recombinant DNA technique and is dealt with in the subsequent sections. Another technique known as **transfection** refers to the direct injection of DNA into mouse eggs resulting in the incorporation of some of the DNA into the host chromosome. The incorporated DNA then replicates alongwith the host genome.

Any type of gene therapy corrects the genetic deficiency only in a clone of somatic cells. There is no alteration in the genotype of the germ cells. This means that the defective gene would continue to be transmitted on the mendelian pattern. But the gene therapy does correct the defect permanently in the individual who was given such a therapy. Like any physical disorder the occurrence of genetic disorder is better prevented than cured.

SAQ 2:

I. Match the following:

a)	Screening potential parents	i)	Screening of infants for purpose of therapy
b)	Prenatal screening	ii)	Avoidance of marriage between heterozygous individuals
c)	Screening of newborns.	iii)	Amniocentesis and chorionic villus biopsy

II. Fill in the blanks:

i) Therapy for inherited defects on vision are treated by the use of ....

ii) New born infants suffering from ..... should avoid taking milk as a diet.

iii) Transplantation of a genetically normal tissue into a patient with a defective gene function is ....

23.3 RECOMBINANT DNA TECHNIQUE

The recombinant DNA technology widely known as genetic engineering or gene cloning aims at the isolation of required DNA fragment and recombining of the

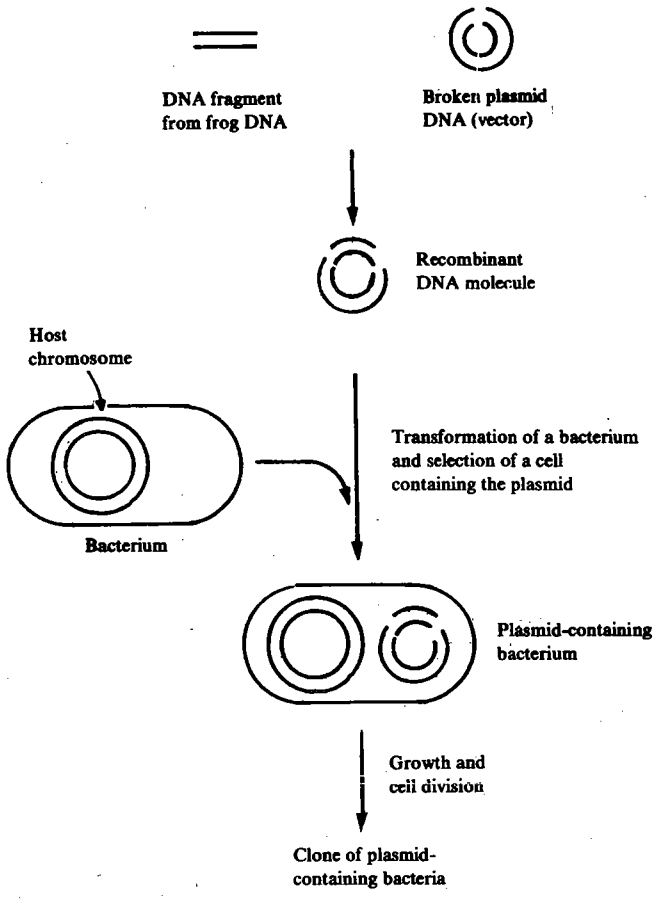


Fig. 23.2: Recombining and cloning of a DNA fragment.

isolated fragment by suitable procedures. More specifically two DNA molecules are isolated and cut into fragments by specific enzymes and then joined together in a desired combination. Such a reconstructed gene is restored to a cell by suitable procedures for replication and reproduction (Fig. 23.2). In this section you will learn in detail the methods adopted in the creation of recombinant DNA molecule and the process by which such a molecule is amplified in a bacterial cell–the gene cloning process.

In order to create a recombinant DNA molecule the required DNA fragment has to be incorporated into a suitable genome. Such a genome is known as a **vehicle** or a **vector.** To be useful in the cloning process a vector or a cloning vehicle should have the following properties.

- It must a small and a well characterised molecule.
- It must have a replication origin, enabling self-replication as well as the replication of inserted segment.
- It should be amenable for the selection of the hybrid molecules.

There are many vectors currently used for cloning. They are the small circular DNA molecules—the plasmids and phage particles such as the λ phage with 15 to 20 kilobases (kb) of DNA. We shall first look into the isolation of the DNA fragment by appropriate enzymes known as restriction enzymes and then the joining of the isolated DNA with the vector.

23.3.1 Restriction Enzymes

Restriction enzymes also known as restriction endonucleases are those enzymes which recognise specific base sequences in a DNA molecule and make two cuts, one in each strand. Such an action results in generating a 3' hydroxyl and a 5' phosphate termini.

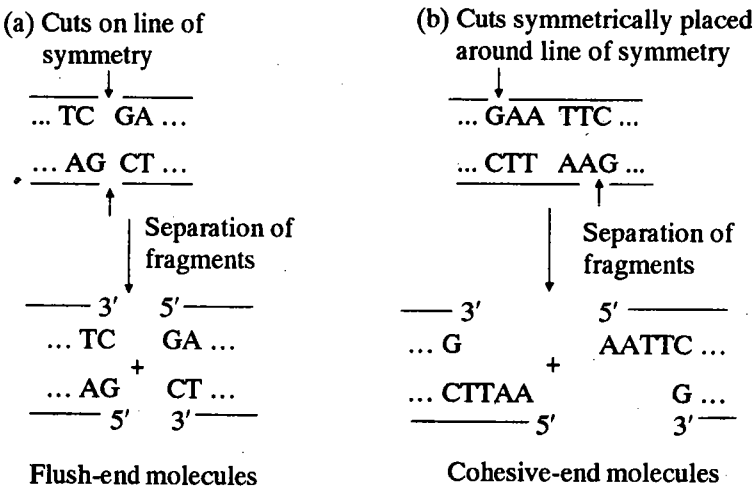


Fig. 23.3 The types of cuts made by restriction enzymes.

Restriction enzymes used in recombinant DNA are capable of making breaks in the DNA molecules by two distinct arrangements.

One type of cleavage results in blunt ended DNA fragment and the other type of cleavage generates cohesive ended molecules. The blunt ended molecules are produced when a restriction enzymes acts on the line of symmetry and the cohesive ended molecules are produced when a restriction enzyme acts around the line of symmetry (Fig. 23.3).

An important property of restriction enzymes is that a particular enzyme recognises

a unique sequence of DNA bases. (Table 23.1) This essentially means that any given restriction enzyme can make only a limited number of cuts in the DNA. For instance a restriction enzyme can create only a few hundred to few thousand DNA fragments from a bacterial chromosome which contains around 3 million base pairs. The genome of a plasmid which is a much smaller DNA molecule may have less than ten cutting sites. Such a specificity of restriction enzymes generates unique family of fragments from a DNA molecule. The DNA fragments produced by the action of a restriction enzyme are called restriction fragments.

Table 23.1 Some restriction endonucleases and their cleavage sites

Microorganism	Name of enzyme	Target sequence and cleavage sites
Generates cohesive ends		
E. coli	EcoRI	G A A ¦ T T C C T T ¦ A A G
Bacillus amyloliquefaciens H	BamHI	G G A ¦ T C C C C T ¦ A G G
B. globigii	BglII	A G A ¦ T C T T C T ¦ A G A
Haemophilus aegyptius	HaeII	Py G C ¦ G C Py Py C G ¦ C G Py
Haemophilus influenza	HindIII	A A G ¦ C T T T T C ¦ G A A
Providencia stuartii	PstI	C T G ¦ C A G G A C ¦ G T C
Streptococcus albus G	SalI	G T C ¦ G A C C A G ¦ C T G
Thermus aquaticus	TaqI	T C ¦ G A A G ¦ C T
Generates flush ends		
Brevibacterium albidum	BalI	T G G ¦ C C A A C C ¦ G G T
Haemophilus aegyptius	HaeI	(A/T) G G ¦ C C (T/A) C C (T/A) ¦ G G (A/T)
Serratia marcescens	SmaI	C C C ¦ G G G G G G ¦ C C C

Note: The vertical dashed / indicates the axis of dyad symmetry in each sequence; Arrows indicate the sites of cutting.

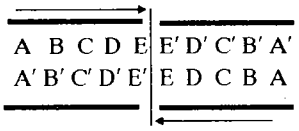


Table 23.4: A general form of nucleotide palindrome sequence

We earlier said that each restriction enzyme recognises only an unique sequence. All such unique sequences have been shown to be palindromic sequences or inverted repeat sequences. Such a sequence consists of four to six complementary bases (Fig. 23.4). The dashed or interrupted vertical line in Fig. 23.4 is the line of symmetry. The double stranded segment to the right of the line of symmetry can be superimposed on the one to the left.

23.3.2 Joining of the Cohesive Ended Restriction Fragment with the Vector

In the previous subsection you have learnt about the generation of cohesive and blunt

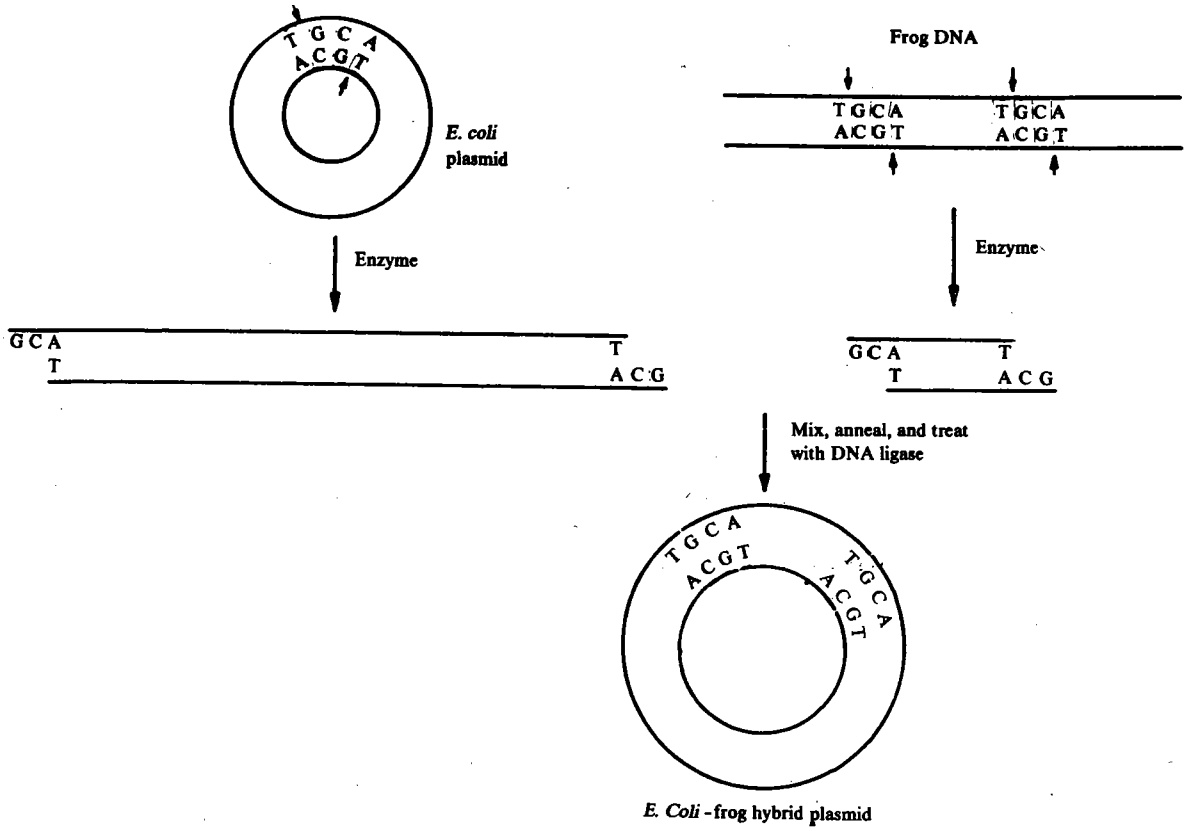


Fig. 23.5 The mechanism of consturction of a hybrid DNA between a bacterial plasmid and frog DNA

ended restriction fragments by the action of restriction enzymes. The restriction fragments thus generated have to be joined with the vector, either a plasmid or a phage before it is cloned into a bacterium where the gene could be amplified.

Cohesive ended restriction fragments, as shown in Fig 23.3 have single stranded terminal at each end that are complementary. The joining procedure takes advantage of the complementary single stranded termini. You may also recall that the fragments produced by a particular enzyme acting on two different DNA molecules have the same set of single stranded ends. This is because both the DNA molecules have the same restriction sequence recognised by the enzyme. Let us now look into the technique that enables the joining of DNA molecules from two different sources such as an *E. coli* plasmid and a frog (Fig. 23.5).

You can see from Fig. 23.5 that both the plasmid DNA molecule and the frog DNA have similar restriction sequences (TGCA). Moreover the plasmid has only one cleavage site for the restriction enzyme. Once the enzyme has independently cleaved the plasmid and the frog DNA, the two sources of DNA are mixed, annealed and treated with DNA ligase to obtain a permanent joining. The interspecific hybrid plasmid thus produced is also called a **chimera**. The hybrid plasmid can then be introduced into a bacterium where the DNA fragment would replicate as a part of the plasmid.

23.3.3 Joining of blunt ended restriction fragments with the vector

The joining of the blunt ended restriction fragments is achieved by the addition of a homopolymer tail to the 3′ hydroxyl group to produce an extended single stranded segment of a DNA chain. The homopolymer tail is essentially a sequence of a few similar nucleotides. For examples it could be a poly A or a Poly T tail (Fig. 23.6).

The molecules to be joined are first treated with a 5′ specific exonuclease that

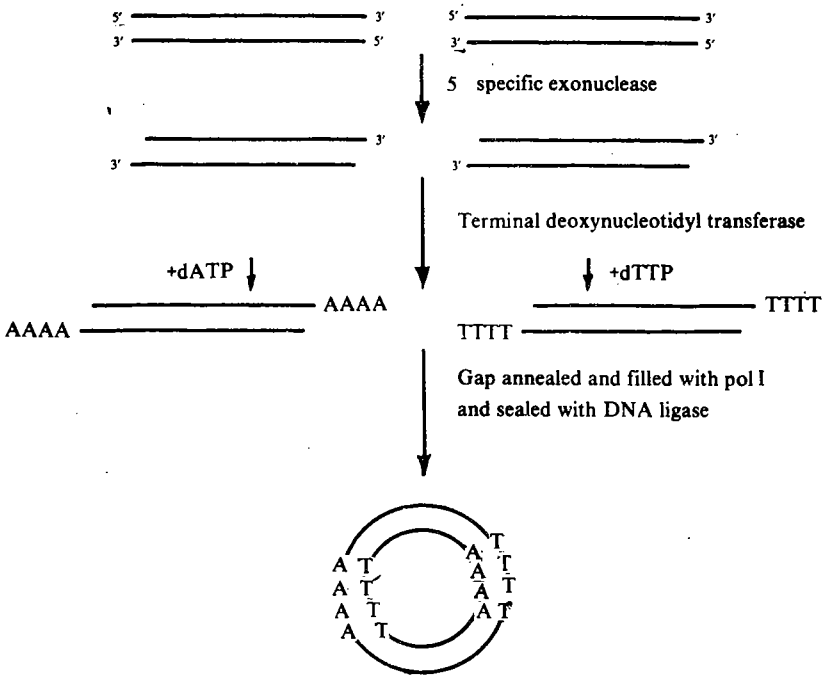


Fig. 23.6: Joining of two DNA fragments with complementary homopolymer tails.

removes a few terminal nucleotides from each strand at the 5′ ends of the molecule. Such an exonuclease treated DNA is mixed with dATP (deoxy adenosine triphosphate) and a specific enzyme, the terminal nucleotidyl transferase. This unusual polymerase obtained from animal tissues adds the nucleotides to the 3′ end of a single stranded DNA molecule. We call this enzyme an unusual polymerase in the sense that unlike the normal DNA polymerase it does not require a template strand for the addition of the nucleotides. In the reaction mixture a poly A tail will be formed at the 3′ends of a double stranded DNA molecule. If instead of dATP, dTTP (deoxy thymidine triphosphate) is added, then the second molecule will have a homopolymer tail of dT. Once the complementary homopolymer tails are attached to the two different DNA fragments, they are joined by the gap filling DNA polymerase-I and sealed with DNA ligase.

23.4 THE CLONING OF THE HYBRID DNA INTO A HOST

So far we described the methods of generating restriction DNA fragments, the vector systems and the joining mechanisms to produce recombinant DNA molecules. The recombinant DNA has to be introduced into a suitable host for its propagation. There are two approaches followed to cloning a specific gene. In the first approach all the fragments from a restriction digest are cloned non-selectively into a host and then screening for a desired gene is done. This non-selective method of cloning random DNA fragments into a host is called shot gunning method (Fig.23.7). This results in the creation of a gene bank or a gene library from which specific gene could be screened. In this method the hybrid plasmids or the phages containing the gene library represent the entire genome of the organism. Each plasmid or phage may carry a different small fragment of the genome. The second approach consists of using a purified probe for the gene for selecting the appropriate restriction fragment and then cloning that specific fragment.

The introduction of the hybrid plasmid or phage is done by transformation. And bacteria such as *E. coli* can be made permeable to DNA by treatment with calcium

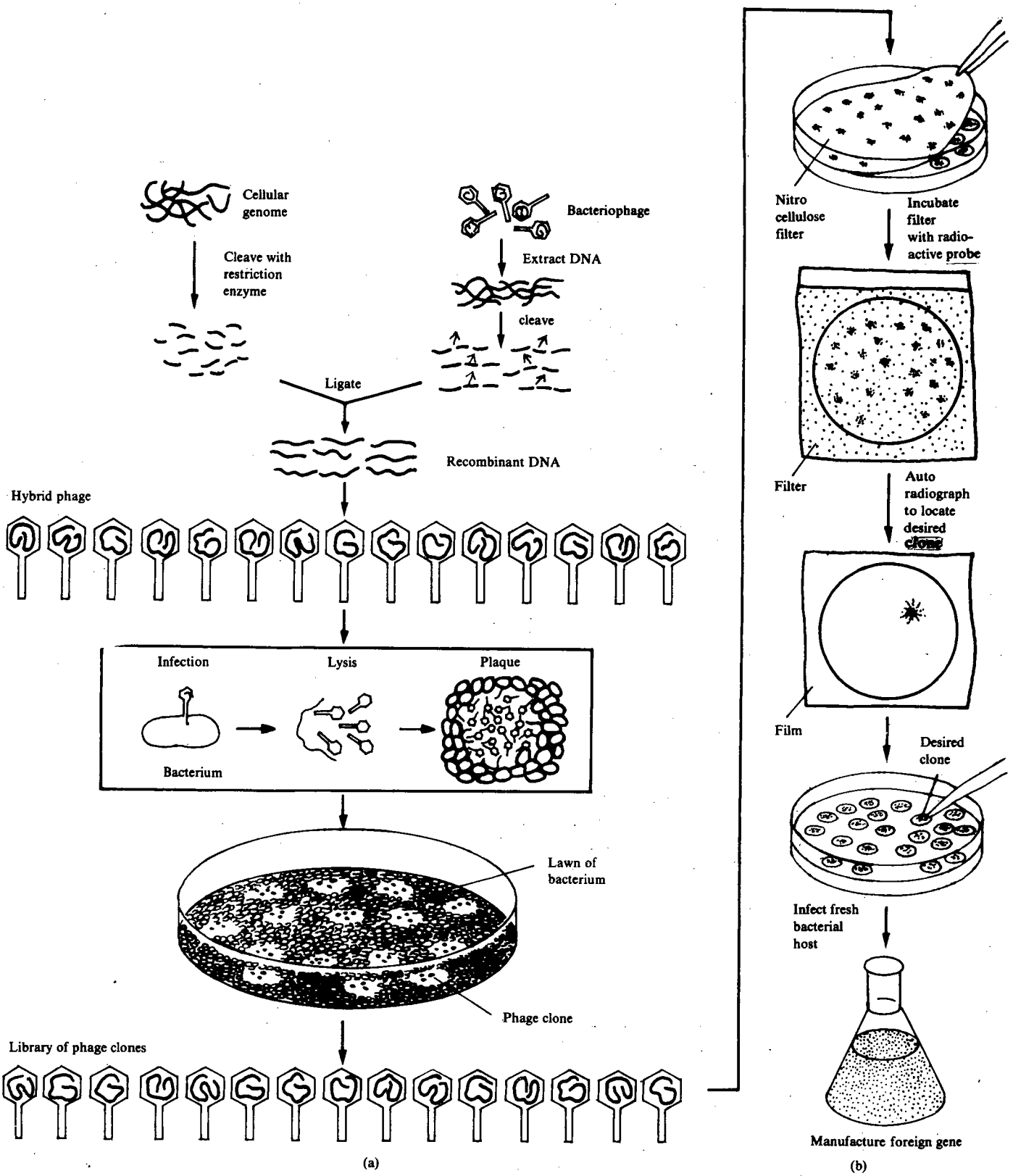


Fig. 23.7: The method of shotgunning, creating a gene library and selecting for the desired gene. a) The creation of recombinant DNA genome from an eukaryotic genome and a bacteriophage genome b) hybrid phages. c) Each hybrid phage is allowed to infect a bacterium and multiply. d) The plaque pattern is transferred to a nitrocellulose filter and the phage protein is dissolved leaving only the recombinant DNA. e) The filter is incubated with the radioactively labelled probe. The probe is actually a DNA copy of the messenger RNA representing the desired gene. f) The autoradiography reveals the position of the clone having the desired gene. g) The desired clone is then selected and transferred to a fresh bacterial host to obtain multiple copies recombinant DNA.

chloride (for detailed discussion on transformation refer to Unit 12 of LSE-03). The foreign DNA after it is cloned into a bacterium becomes part of the replication mechanism of the bacteria and the gene product is elaborated.

SAQ 3:

State whether the following statements are true or false.

1. A vector to be of use as a cloning vehicle must have a replication origin, capable of self replication and replication of a inserted segment.

   True/False.

2. Blunt ended restriction fragments are produced when a restriction enzyme acts around the line of symmetry.

   True/False.

3. Specificity of restriction enzymes generates unique family of fragments from a DNA molecule.

   True/False.

4. Homopolymer tail is essentially a sequence of four different nucleotides.

   True/False.

5. Terminal nucleotide transferase is a unique polymerase because it requires a template strand for the addition of nucleotides.

   True/False.

6. Cloning on the introduction of hybrid plasmid or phage is done by transformation.

   True/False.

23.5 APPLICATIONS OF GENETIC ENGINEERING

Genetic engineering has many practical applications. The following are the major benefits of the application of genetic engineering technique now come to be widely known as biotechnology.

- Production of a number of biochemicals such as enzymes and drugs as well as commercially important chemicals.
- Production of specific RNA and protein molecules in abundant quantities.
- Creation of new varieties of plants with particular desirable characteristics such as disease resistance or less fertiliser requiring.
- Isolation of a particular gene or a region of genome and its characterization.
- Creation of organisms with economically important features.

We shall look into a few chosen examples of the applications of the genetic engineering technique and the benefits derived therefrom.

23.5.1 Applications to Agriculture

An important application of recombinant DNA technology is to alter the genotype of plants for purposes of high yield, disease resistance and less fertilizer consumption. The first genetically engineered plant of commercial value was produced in 1985. Glyphosate is a commonly used weed killer and this it does by

inhibiting a particular essential enzyme in many plants. But glyphosate is not specific towards weeds alone but kills the useful crops also. It was discovered that the target gene of glyphosate is present in a bacterium *Salmonella typhimurium*. A mutant of *S. typhimurium* is resistant to glyphosate. The mutant gene was cloned to *E. coli* and then recloned to *Agrobacterium tumefaciens* through its plasmid T_i. Infections of plants with purified T_i containing the glyphosate resistant gene has yielded a variety of crops such as maize, cotton and tobacco, all of which are resistant to glyphosate. This makes possible to spray the crop fields with glyphosate to kill the weeds specifically.

23.5.2 Applications to Medicine

Clinical medicine is one of the areas where genetic engineering technique will be employed extensively in future. The technique has been quite useful in developing microorganisms that would overproduce antibiotics which in turn would reduce the production cost. Also some of the biologically active compounds useful in therapy such as insulin, somatostatin and α-interferon are produced by recombinant DNA technique. Somatostatin–a fourteen amino acid polypeptide hormone synthesised by the hypothalamus was earlier obtained from the human cadavers and that too in very small amounts. Somatostatin administered as a drug for certain growth related abnormalities appears to be species specific and that obtained from other mammals has no effect on humans; hence the extraction from the hypothalamus of cadavers. With the advent of genetic engineering technique the gene is chemically synthesised, joined to the pBR-322 plasmid DNA and then cloned into a bacterium. The bacteria is converted into a somatostatin synthesising factory. A similar story could be told with reference to insulin and genetically engineered insulin is now commercially available as humilin.

Another example of the product of biotechnology is the antiviral agent α interferon. The substance is presumed to act by supressing the infection duration of a virus and is found to be effective against herpes virus, infection of the eye, multiple sclerosis, atherosclerosis in rats and may possess the antitumor activity. Genetically engineered interleukin-II, a substance that stimulates the multiplication of B-lymphocytes is also available and is being currently tested on AIDS patients.

Biotechnology has also attempted to make vaccines for disease prevention. Although there are certain problems in cloning and purifying the genes for the viral antigens, such as their poor antigenecity and thermal instability attempts are underway to develop vaccines using the anti-smallpox agent, the *Vaccinia virus* as a carrier. The method consists of genetically engineering the viral antigens present on the surface of the viral particles onto the coat of *Vaccinia*. *Vaccinia* hybrids with surface antigens of hepatitis B, influenza virus and vesicular stomatitis virus which kills cattle, horses and pigs have been prepared and used in antimal tests. Attempts are also underway to obtain an antimalaria vaccine against the surface antigens of the *Plasmodium falciparum*.

23.5.3 Applications to Industry

Yet another area where the genetic engineering techniques are being extensively used is industry. The aim is to produce bacteria with novel phenotypes by combining the features of several bacteria. For instance several genes from different bacteria have been introduced to a simple plasmid and then cloned into a bacterium, resulting in an organism that metabolises oil spills in the oceans, thus getting the ocean water cleaned.

Genetically designed bacteria are put into use for synthesising industrial chemicals. Creation of organisms that are capable of composting waste very efficiently, fix nitrogen to reduce the fertilizer input but yet increase the fertility of the soil and convert biological waste into alcohol are some of the attempts that are made by people who are actively involved in biotechnological research.

Biotechnology can cut down on expenditure on pestcides and improve the quality of the environment as well. Maize and soyabean are extensively damaged by black cut worm. *Pseudomonas flourescens* lives in association with maize and soyabean. *Bacillus thuringiensis* contains a gene pathogenic to the pest. The pest has over the years not only become dangerous to the crops but has developed resistance to a number of pesticides. In preliminary studies the pathogenic gene from *B. thuringiensis* is cloned into *Ps. flourescens* and inoculated into the soil. It is found that the genetically engineered *Ps. flourescens* could cause the death of cut worms.

23.5.4 Production of Proteins from Cloned Genes

Genetic engineering aims at the production of large quantity of a simple proteins that may not be otherwise possible to obtain. Theoretically it is possible to insert a foreign gene into a bacterium adjacent to a promotor and get the gene transcribed. Assuming a high copy number hybrid plasmid is involved in the cloning process, as high as 5% gene product could be harvested. But there are several practical constraints in the expression of a eukaryotic gene in a prokaryotic system. The major problems are:

- The non-recognisition of eukaryotic promoter by bacterial RNA polymerase.
- Lack of a specific nucleotide sequence in the eukaryotic mRNA for binding to bacterial ribosomes.
- Absence of mechanisms in prokaryotic systems for eukaryotic mRNA processing.
- Inability of the prokaryotic system to bring about the post-translational changes of a nascent eukaryotic polypeptide.
- Finally eukaryotic proteins are often cleaved by the bacterial proteinases and being regarded as foreign proteins.

In this unit we have briefly discussed two different aspects of modern genetics that have far reaching implications for human welfare. Under genetic engineering not only we discussed the methods of constructing a recombinant DNA molecule and cloning it into a suitable organism but also the applications of the technique to agriculture, medicine and industry. In the last one decade the genetic engineering techniques have invited the attention of many a researchers. Over 250 biotechnology companies have focussed their attention on manufacturing products through genetic engineering process. There are many a journals and books devoted exclusively to the study of genetic engineering worldwide. It is multicrore rupee project and would remain so far many years to come.

SAQ 4:

Match the following:

1. *Salmonella typhimurium*	a) anti-smallpox agent
2. *Vaccinia virus*	b) pathogenic gene for cutworms
3. *Pseudomonas flourescens*	c) anti-cancer agent
4. α interferon	d) glyphosate
5. Somatostatin	e) pBV 322 plasmids

23.6 SUMMARY

In this unit you have studied that:

- Genetic counselling is one of the important aspects of health care. Counselling

helps to assess the genetic risks involved in giving birth to children with inherited disorders and to assist the parents in dealing with the risk.

- Whereas informative counselling aims at imparting the parents the knowledge to understand the risks and the probability that a child will not be affected, supportive counselling offers professionalto help the parents to effectively deal with the risks.
- Genetic diagnosis means the identification of genetic disorders either prior to the marriage of the couple, prior to the birth of the child or in the new borns.
- Amniocentesis and chorionic villus biopsy are the two recognised tests for prenatal diagnosis with the former being more effective than the latter.
- The recombinant DNA technology involves the introduction of a foreign gene into a plasmid or a phage genome known as a vector and the amplification of the recombined DNA in a suitable host such as bacteria.
- The restriction endonucleases are molecular scissors that cleave a DNA molecule at a specific sequence to generate cohesive ended or blunt ended restriction fragments. Each restriction enzyme can only specifically act on a given four to six palindromic nucleotide sequence to produce the restriction fragments.
- The plasmids or the phages are the efficient vectors in which a DNA fragment could be recombined to generate a new DNA molecule. The amplification or the replication of the recombined DNA can be done in a suitable bacterium after the introduction of the DNA into the host by the transformation process.
- The recombinant DNA technique has a wide range of applications in the fields of agriculture, medicine, industry as well as the genetic research.

23.7 TERMINAL QUESTIONS

1. Define the term genetic counselling? What are the components of genetic counselling process?

.....

.....

.....

.....

.....

.....

2. How can amniocentesis and chorionic villus biopsy be advantageously used in prenatal diagnosis?

.....

.....

.....

.....

.....

.....

3. What do you understand by genetic screening?

.....

.....

.....

.....

.....

.....

4. Give a brief account of restriction enzymes.

.....

.....

.....

.....

.....

.....

5. Describe any two applications of genetic engineering technique.

.....

.....

.....

.....

.....

.....

23.8 ANSWERS

Self-assessment Questions

1) 1) True 2) False 3) False 4) False

2) I. a) (ii) b) (iii) c) (i)

II. i) Prosthetic devices ii) Galactosemia iii) gene therapy.

3) 1) True 2) False 3) True 4) False 5) False

6) True.

4) 1) (d) 2) (a) 3) (b) 4) (c) 5) (e)

TERMINAL QUESTIONS

1. Genetic counselling can be defined as a communication process which deals with human problems associated with the occurrence or the risk of occurrence of a genetic disorder in a family. The different components of genetic counselling are i) diagnosis of genetic disorder, ii) counselling process leading to preparation of pedigree chart iii) drawing appropriate conclusion based on pedigree analysis.

2. Both amniocentesis and chorionic villus biopsy are useful prenatal diagnostic tests for detecting possible genetic disorder in the newborn babies. Both techniques depend on obtaining foetal cells (either from amniotic fluid or from chorionic villus surrounding foetus) and culturing them *in vitro*. The cultured

cells are then analysed for their chromosomes and karyotyping is done. Chromosomal abnormalities, if any, can be identified and suitable measures evolved to combat the problem.

3. Genetic screening would refer to systematic search in a population for persons of certain genotypes. The screening would enable to prevent occurrence in future of many of the genetic disorders in the population. The screening is done at three levels i) identification of potential parents ii) identification by prenatal screening iii) identification of affected new borns.

4. Restriction enzymes are enzymes which recognise specific base sequences in a DNA molecule and make two cuts in each strand. This generates DNA fragments with a 3′ OH and 5′ phosphate termini. There are two types of restriction enzymes. i) enzymes which act on the line of symmetry of the cleavage site ii) enzymes that act around the line of symmetry of the cleavage site.

5. Refer to Section 23.5

UNIT 24 GENETICS IN AGRICULTURE, AND PLANT IMPROVEMENT

Structure **Page No.**

24.1 INTRODUCTION

You might have tried growing plants either in small pots or in your backyard at some time or the other. You probably tried your hand at vegetable gardening too. In this experience, did you encounter some common pests such as bugs and other insects that destroyed a large portion of your plants? If yes, then quite natually you might have either thought of writing off your gardening experience completely, or you might have tried some type of pest control. And, if you chose a toxic pesticide, you would worry about its possible toxic effects on you and your family as well as to the environment. Perhaps you chose an organic preparation, but it might have been more difficult and expensive to use than its synthetic counterpart. In such a situation, you probably wished that the vegetables or other plants had a built-in pest-repellent. This wish is something that the genetic engineers and biotechnologists have been thinking of for many years. They have tried to find solutions to such problems by applying the science of genetics. Besides building pest resistance, they have worked in many areas of plant improvement. Many important breakthroughs have been made. In the following sections you are going to study some important developments in the improvement of plants.

Objectives

After studying this unit, you should be able to:

- explain with examples, the various kinds of improvements made in plants by the application of the principles of genetics (Section 24.2).

24.2 GENETIC ENGINEERING OF PLANTS

Though plant breeding is an ancient art, it is based on sound scientific principles that aim at improving the economically important plants for the benefit of mankind. The traditional methods of plant impovement involve *selective breeding* in which parents with superior characteristics are used. The resultant progeny or hybrid is superior too, as it demonstrates the traits of both the parents. This phenomenon known as *hybrid vigour* or *heterosis* forms the basis of the improvement of many crop plants of which corn is a well-known example.

Ever since Mendel published the results of his famed breeding experiments in the garden pea, in 1866, the science of genetics developed at an unprecedented pace. And during the course it also developed several new tools and techniques. The application of modern methods such as the recombinant DNA technology has given a new impetus to plant improvement. Given below are some examples of plants, whose various qualities have been altered or modified to the desired levels by the application of the recombinant DNA technology.

24.2.1 Regneration of Plants from Callus Tissues and Protoplasts

Plants have a unique property, that is, *totipotency* which means that their cells are able to regenerate complete new plants. The differentiated plant cells are able to dedifferentiate to the embryonic state and subsequently to redifferentiate into new cell types. Thus, there is no separation of germ line cells from somatic cells as in higher animals.

When excised tissues from mature plants, known as *explants*, are placed in the appropriate sterile tissue culture conditions, notably in the presence of the plant hormone 2,4- dichlorophenoxyacetic acid (2,4-D), cells in these tissue explants often dedifferentiate and grow into highly unorganised cell masses (see Fig. 24.1)

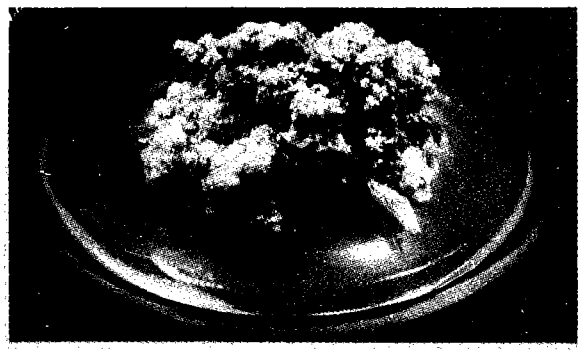


Fig.24.1: Undifferentiated callus growing on synthetic culture medium.

called calli (or calluses, the singular form is callus) . If these undifferentiated callus cell-clumps are transferred to a growth medium lacking 2,4-D, but containing growth hormones such as kinetin, in many species plantlets will regenerate (see Fig. 24.2).

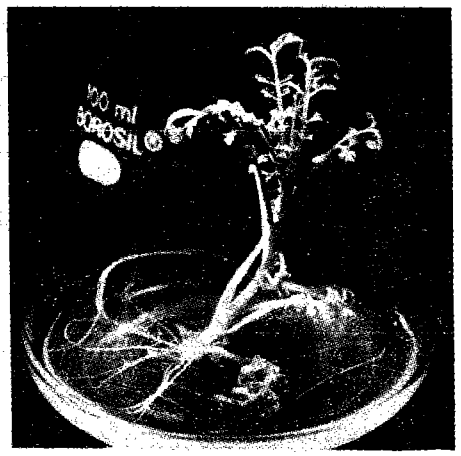


Fig.24.2: Plantlets regenerated from callus tissues growing on sterile culture medium.

Moreover, with some plant species such as petunias, tobacco, tomato and potato, one can regenerate plants from even isolated protoplasts. These are single cells from which the walls have beeen removed by using appropriate enzymes. Protoplasts can be used in two different ways. Firstly, protoplasts of two different plants can be used to produce a hybrid cell from which a hybrid plant can be produced. This hybrid plant may be subjected to conventional plant breeding techniques if desired. Secondly, the protoplasts can be used as recipient cells for the introduction of specific genes being carried on vectors, that is, they can be transformed just as the *E.coli* cells. In this manner a range of genetically engineered plants, including ones with increased yield, pest resistance and herbicide resistance can be generated.

The totipotency of plant cells is thus a major advantage for genetic engineering since it permits one to regenerate entire plants from individual modified somatic cells.

24.2.2 The *Ti Plasmid* of *Agrobacterium tumefaciens*

Till date, the most important tool in the genetic engineering of plants has been the *Ti plasmid* of the soil bacterium *Agrobacterium tumefaciens*. This bacterium causes, crown gall disease of dicotyledonous plants (see Fig.24.3). This disease is characterised by development of tumours or galls at the wounding sites which are often at the crown (hence the name crown gall). Crown is the junction between the root and the stem. Since the crown of the plant is usually located at the soil surface, this is where a plant is most likely to be wounded due to abrasion by soil particles or from the plant blowing in strong wind, and infected by the bacterium. However, *A. tumefaciens* can infect any part of the plant and induce formation of tumour.

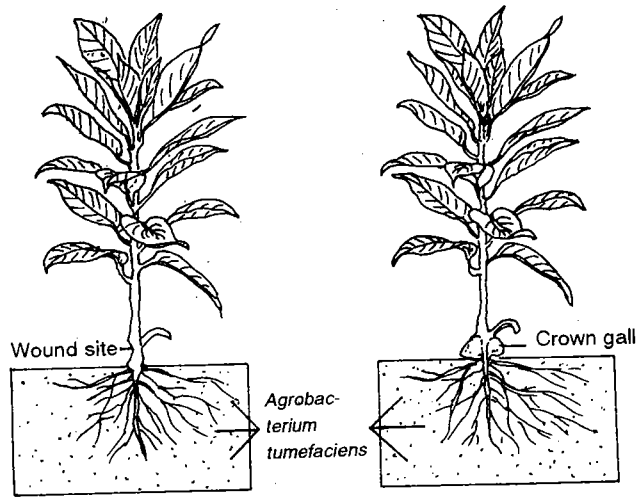
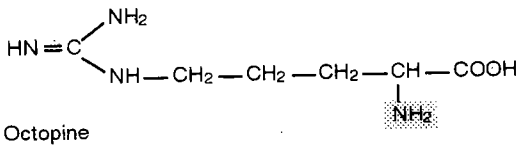


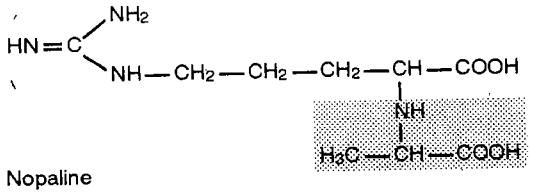
Fig.24.3: Cown gall formation on dicotyledonous plants after infection with the bacterium *Agrobacterium tumefaciens*.

When a wound site in a dicotyledonous plant is infected by *A.tumefaciens*, two key events occur. *One*, the plant cells begin to proliferate and form tumours. *Two*, they begin to synthesise an arginine derivative called an *opine*. The opine synthesised is usually either *nopaline* or *octopine* (Fig.24.4) depending on the *A. tumefaciens* strain involved. These opines are catabolised and used as energy sources by the infecting bacteria.The *A. tumefaciens* strains that induce the synthesis of nopaline can grow on nopaline, but not on octopine and vice versa. Thus the particular bacterial strain is able to divert the metabolic resources of the host plant to the synthesis of opines, which are of no known benfit to the host plant, but provides sustenance to the bacterium.

(a) Arginine



(b) Octopine



(c) Nopaline

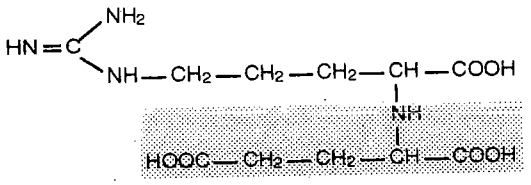


Fig.24.4: a) Chemical structure of the amino acid arginine, b) octopine, and c) nopaline

The genetic information for inducing the crown gall disease in plants is carried on a large, about 200,000 bp plasmid present in the *A.tumefaciens* cell (Fig.24.5). This plasmid is called the *Ti plasmid*, for its tumour-inducing capacity. Two components of the *Ti plasmid* are essential for the transformation of plant cells, and these are the *T-DNA* and the *vir region*. During the process of transformation, the *T-DNA* (refers

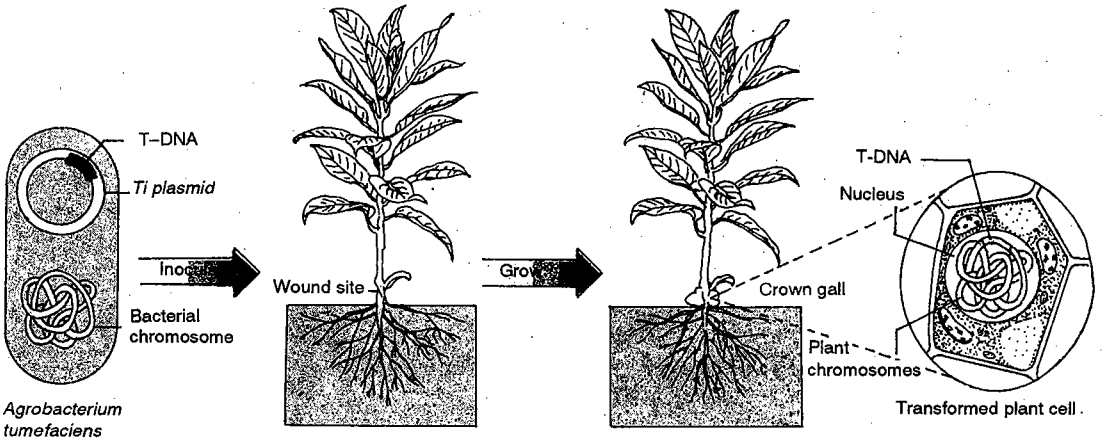


Fig.24.5: ***Agrobacterium tumefaciens*, bearing the *Ti plasmid*. Cell of the plant tumour (contain the *T-DNA* segment of the *Ti plasmid* integrated into the DNA of the plant chromosome.**

to transferred DNA) is excised from the *Ti plasmid*, and is transferred to a plant cell. Once in the plant cell, the *T-DNA* integrates into the DNA of the plant cell. The integration of *T-DNA* occurs at random chromosomal sites. The *T-DNA* in the *nopaline-type Ti plasmid* is about 23,000 bp having 13 genes. In *octopine type Ti plasmids*, there are two separate *T-DNA* segments. For convenience, we shall restrict to the *nopaline-type Ti plasmids* whose structure is shown in Fig. 24.6.

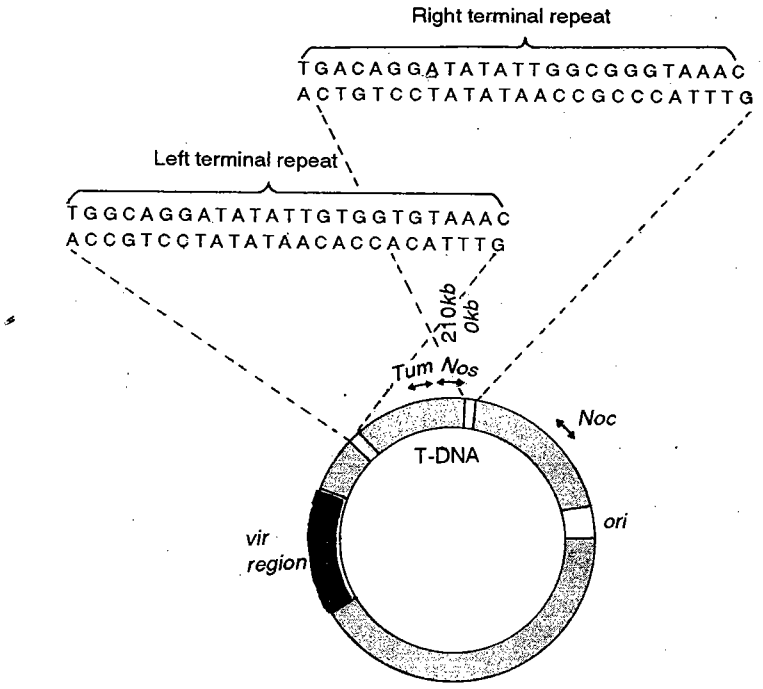


Fig. 24.6 : Diagrammatic representation of the structure of *nopaline Ti plasmid*. The description of symbols used in the figure are as follows: *Ori*, plasmid orgin of replication; *Tum*, region containing genes responsible for tumour formation or the genes coding for the biosynthesis of phytohormones; *Nos* region contains genes involved in nopaline biosynthesis; *Noc* region containing genes involved in catabolism of nopaline; *vir* or the virulence region containing the genes required for *T-DNA* transfer (From M.W. Bevan and M.D. Chilton, *"T-DNA of the Agrobacterium Ti and Ri plasmids"*, Annu. Rev. Gent., Vol. 16, pp 357-384. 1982).

Some of the genes on the *T-DNA* segment of the *Ti plasmid* encode enzymes that catalyse the synthesis of phytohormones such as the auxin indoleacetic acid and the cytokinin isopentenyl adenosine. These phytohormones result in the development of tumourous growth of cells forming crown galls. The *T-DNA* region is bordered by two repeated 25 bp sequences (see Fig.24.6), which are required for *T-DNA* excision and transfer. The deletion of either border sequence completely blocks the transfer of *T-DNA* to plant cells. The *vir* (virulence) region of the *Ti plasmid* contains the genes required for the *T-DNA* transfer process. These genes encode the DNA processing enzymes required for excision, transfer and integration of the *T-DNA* segment. Once the *T-DNA* region of the *Ti-plasmid* of *Agrobacterium tumefaciens* is transferred to plant cells it becomes integrated in the plant chromosomes. When this whole process of transformation became clear to the scientists, the potential use of *A.tumefaciens* in plant genetic engineering was obvious. Now, it was possible to introduce foreign genes into the *T-DNA* region and these genes would be transferred to the plant with the rest of the *T-DNA* segment.

In 1986, a research team led by Roger Beachy, at Washington University used *Agrobacterium* transformation to induce cross-protection against tobacco mosaic virus (TMV). The gene encoding the TMV coat protein was fused to the cauliflower mosaic virus promoter, and the fused gene was inserted into a *Ti plasmid.* Transgenic tobacco plants created by infection with the recombinant *Ti plasmid* expressed TMV coat protein. The coat protein gene was stably integrated into the genome of the transformed plants and was inherited in a Mendelian fashion by their progeny. The transgenic offspring showed delayed onset of symptoms following inoculation with live TMV, and up to 60% of transgenic plants showed no symptoms at all. This research finding shows the tremendous potential of the use of *Ti plasmid* of *Agrobacterium tumefaciens.*

24.2.3 Gene Transfer in Monocotyledonous Plants

In the earlier section you have seen how the *Ti plasmid* of *A.tumefaciens* can be employed as a vehicle for transferring foreign genes into the genome of a particular plant. This bacterium can work on in a limited range of plants, that is, the ones that are dicotyledonous. As so many of the economically important plants are monocotyledonous, there was a need to develop gene transfer technology that could be applicable to these plants. First, attempts were made to introduce DNA into plant protoplasts by transiently perturbing the cell membranes so as to make them permeable to macromolecules such as DNA. Based on this a successful technique called *direct gene transfer* was developed. It involved the addition of selected marker genes together with polyethylene glycol, which seems to stimulate membrane fusions. A second method with slight variation was also developed and it was known as *electroporation.* It employs a short pulse of high-intensity electrical current to disrupt cell membranes and render them temporarily permeable to DNA molecules. Although by using these procedures one could get transformed protoplasts of monocotyledonous plants but a poor percentage of these protoplasts could regenerate into plants.

Recently a novel procedure has been developed to carry on the transfer of DNA into intact plant cells by a microprojectile gun (Fig.24.7). This gun is used to shoot DNA-coated microprojectiles, the tungsten particles of 1 to 4 μm diameter, into plant cells. This method can be used both in culture cells as well as directly on the growing plants. This technique holds great promise for the future because of being technically simple, and also being applicable to all the plant species, including fungi and algae.

24.2.4 Herbicide Tolerant Plants

Herbicides are chemical agents that when applied to various types of plants, result in the death of the plant. These are used primarily to kill non-desirable plants like weeds, that compete with the desired plants say crops, for space, water and

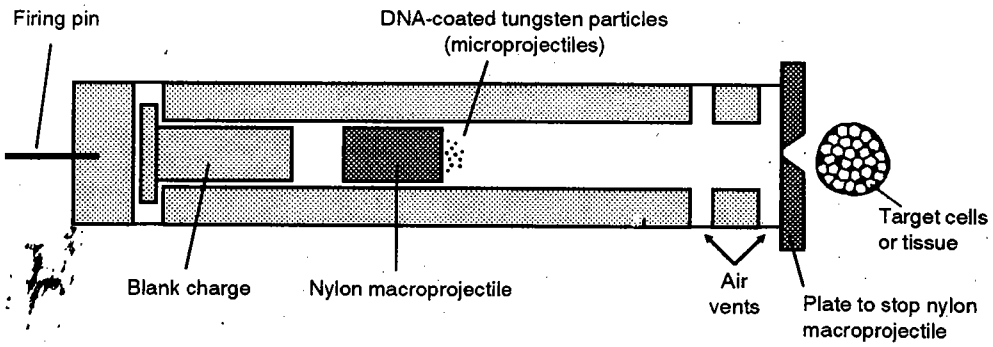


Fig.24.7: Diagrammatic representation of the structure of a microprojectile gun. (After T.M. Klein, E.D. Wolf, R.Wu, and J.G. Sanford. Nature. Volume *327*, pp 70-73, 1987).

nutrients. Unfortunately, the available herbicides seldom provide the degree of specificity that is desired, and most herbicides control only certain classes of weeds and not others. Broad-spectrum herbicides give good weed control, but, in so doing, usually have deleterious effects on the growth of the crop plant as well. As a result, scientists are now evaluating alternate approaches to weed control. The most promising approach, being considered presently is the development of herbicide-tolerant plant varieties for use with broad spectrum or totally nonspecific herbicides.

Though herbicides kill or inhibit the growth of the plants, but do not have any deleterious effects on animals. They act on plants by disabling the target enzymes in the metabolic pathways unique to plants such as those involved in photosynthesis or biosynthesis of amino acids. Glyphosate (see Fig. 24.8) is one of the most potent broad spectrum herbicides known, inhibits 5- enolpyruvylshikimate-3 phosphate synthase (EPSP synthase), an enzyme necessary in the biosynthesis of the aromatic amino acids tyrosine, phenylalanine and tryptophan in the chloroplast.

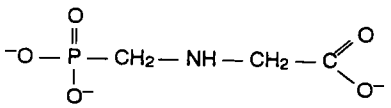


Fig. 24.8: The chemical structure of glyphosate, an active ingredient of many herbicides.

Presently, the major thrust is to identify and transfer herbicide resistance genes into major crop plants. The following three strategies are used for attaining herbicide resistance.

i) Stimulate overexpression of target protein of the herbicide, so that enough of it will escape disablement by the herbicide.

ii) Insert a genetically altered form of target protein that is less sensitive to the herbicide.

iii) Insert an enzyme that disables the herbicide.

The first and the second approaches have been used to produce plants that are resistant to glyphosate. A resistant petunia, in which the EPSP gene is highly amplified, was isolated by growing cultured cells in increasing levels of the herbicide. Petunias transformed with a fusion gene linking the viral promoter to a wild-type EPSP gene also show resistance. A mutant EPSP gene (*aro A*) isolated from *Escherichia coli* is 6,000 times less sensitive to glyphosate than to the wild-type enzyme, but it lacks the plant 'transit peptide' that is needed to transport EPSP into the chloroplast. To overcome this, a fusion gene was constructed, linking the *aro A* coding region with a dicot transit peptide sequence. Petunias and tobacco

plants transformed with this fusion gene express the *aro A* protein in the chloroplast, in addition to their own EPSP.

The third approach to glyphosate resistance has yet to be achieved. However, metabolic pathways have been defined in species of *Pseudomonas* and other bacteria that allow the use of glyphosate as a sole phosphate source. Efforts are under way to clone the genes that encode these glyphosate-degrading enzymes.

24.2.5 Resistance to Pests

Pesticides and insecticides are chemical agents that kill various kinds of insects and other pests, while leaving plants and most animals relatively unharmed. Unfortunately, many pesticides are non-selective. They kill insects which are beneficial to plants and the environment along with those which are not. In addition, pesticides can be toxic to humans and other non-target animals and they contaminate soil and water supplies. This environmental contamination may then lead to a second level of adverse effects on both plant and animal life. Furthermore, prolonged use of insecticides can lead to the development of insect strains which are no longer susceptible to the toxic effects of those chemicals. Such a development has already resulted in the appearance of DDT resistant mosquitoes. You may recall that in the past few decades, DDT was widely used for the control of malaria-causing mosquitoes. Because of all these problems a major goal of genetic engineers has been to provide plants with natural endogenous resistance to a variety of pests through techniques of genetic engineering.

The best known examples of the use of natural gene products to control plant pests are the insect toxins of *Bacillus thuringiensis*. Each of the *B.thuringiensis* toxin genes codes a large protein that aggregates to form protein crystals in spores, and those protein crystals are highly toxic to certain insects. To exploit this lethal characteristic, *B. thuringiensis* spores are combined with water to form a mixture which can be sprayed over an insect-infested area. When insect larvae ingest the bacterial spores, they get killed due to the toxic protein crystals. This treatment was effective but its insecticidal effects were transient due to limited field survival of spores, so long-term insecticidal activity required repeated applications of spores.

Genetic engineers tried to find solution to this problem by applying the techniques of recombinant DNA technology. In one line of research, scientists isolated genes which encode the *B. thuringiensis* toxin and inserted those genes into other organisms such as *E.coli* and the bacterium *Pseudomans fluorescens* which are better suited for survival in field. *P.fluorescens* lives on roots of many different types of plants, including corn. When these transformed bacteria (*P.fluorescens*) were sprayed on corn plants, these bacterial cells colonised the root area. As a result the *B. thuringiensis* toxin is synthesised at the site of plant itself, this helped the plant to ward off their pests.

A second line of research involved linking the *B.thuringiensis* toxin gene to a constitutively expressed promoter and introducing it via the *Ti plasmid*, directly into the cells of the plant which is to be protected. One such study which took place in Belgium involved the insertion of the toxin gene into cells of tobacco plant. Analysis of tissues of the mature transformed plants showed that the tobacco plant indeed synthesised *B. thuringiensis* toxin. When the treated plants were infected with tobacco hornworm larvae, very few of them survived, and the damage was negligible. In contrast, the control tobacco plants which were not transformed with the engineered *Ti plasmid*, all died within approximately two weeks. The inserted gene clearly provided protection to the tobacco plants.

Based on similar lines, work is underway to identify genes for salt and drought tolerance in plants.

SAQ 1

Match the items of coloumns I and II. Write the correct choice in the space provided.

I			II	
i)	Explants	[ ]	a)	affects target enzymes of plant's metabolic pathways
ii)	Crown gall	[ ]	b)	tissues culture
iii)	Electroporation	[ ]	c)	*Bacillus thuringiensis*
iv)	Herbicide	[ ]	d)	direct transfer of DNA to protoplasm
v)	Natural pesticide	[ ]	e)	*Ti plasmid*

24.2.6 Resistance to Frost Formation

If the temperature falls below 32°F (O°C) ice crystals form and water freezes. Ice crystals, forming in the interior of the cell often destroy the cell wall with the result that cell death soon follows. This phenomenon can be illustrated by freezing a firm tomato or even some peas. As the temperature drops, ice crystals form and disrupt the delicate cells of tomato. When the frozen tomato is later allowed to thaw, only a soggy, limp fruit remains which was once quite firm. A similar effect can be seen on growing plants which are subjected to freezing temperatures.

The presence of the common bacterium *Pseudomonas syringae* facilitates ice crystals or frost formation on a plant. *P. syringae*, found on many types of plants, contains a protein which acts as a site of nucleation for the formation of ice crystals at temperature of 32°F (O°C). In contrast, when *P.syringae* or other nucleating agents are absent, ice crystals do not form until the temperature drops to approximately 20°F (–7°C).

Scientists found that the nucleation ability of *P. syringae* results from a single gene in the bacterial genome. By removing this gene, scientists created a strain of bacteria which, when applied to growing plants does not provide the necessary site of nucleation for frost formation even when temperature drops briefly as low as 23°F (– 5°C). This genetically engineered bacterial strain called *ice minus* was developed at University of California in Berkley. It has a great potential for future use in the prevention of frost damage to a variety of crops.

24.2.7 Enhanced Nitrogen Utilisation

As you know, plants can only utilise nitrogen that has been incorporated into chemical compounds such as ammonia, urea or nitrates. No green plants is capable of utilising diatomic nitrogen (N_2) directly from the atmosphere. In other words, a continuous supply of nitrogen in usable form or the *fixed nitrogen* must be available for the growth and development of plants.

For obtaining optimal yields of crops, farmers usually supplement the soil with fertilisers, nitrogen being its important constituent. Because the purchase of nitrogen fertilisers represents one of the major expenses incurred with current agricultural production methods, a major effort has been and continues to be devoted to the development of enhanced methods of biological nitrogen fixation.

Certain free-living soil bacteria such as *Azotobacter vinelandii* and *Klebsiella pneumoniae* directly convert atmospheric nitrogen to ammonia. These bacteria are valuable research material for conducting investigations on the mechanism of nitrogen fixation. In *Klebsiella*, there are 17 *nif* genes (nitrogen

fixation genes) that are organised in seven operons. The situation in nitrogen fixation is very different from the examples discussed above, such as herbicide tolerance etc. You may recall in the instances discussed above, a single gene was constructed and transferred to plants. But in this case, it is quite difficult, if not impossible, to engineer 17 different genes and transfer all of them to the same recipient plant, and to coordinate their expression as well, so that all the components of the complex nitrogen fixing enzymatic machinery are synthesised in proper amounts and proper cells. Presently, we are still away by a few years from our goal.

Besides the free-living nitrogen fixers, the symbiotic nitrogen fixers too are very important sources of biologically fixed nitrogen. Can you guess we are talking about which organisms? It is, about the relationship between the genus *Rhizobium* and plants of family *Leguminosae*, i.e., peanuts, pea, soyabean, alfa-alfa etc. This kind of nitrogen fixation takes place in root nodules that develop due to the interaction of *Rhizobium* with roots of legumes. Thus nodule formation is dependent on the genetic information of both the plant and the bacteria. The nitrogenase that brings about reduction of nitrogen is encoded by the bacterial genome, but the fixed nitrogen is utilised for growth of both the bacteria and host legume plant. Once the mechanism responsible for establishing this symbiotic relationship, and for nodule formation are known and the genes that control these processes have been identified, it might be possible to use genetic engineering to modify nonlegume plants so that they too acquire the nitrogen-fixing capability. Presently, many research laboratories are working on this aspect to turn this dream to reality.

24.2.8 Improving Nutritional Value

Worldwide, the seeds of legumes and cereal grains are estimated to provide humans with 70 per cent of their dietry protein requirements. Protein molecules are composed of varying arrays of twenty different amino acids, of which human body can synthesise twelve. The remaining eight amino acids, called *essential amino acids* must be provided to the body by ingestion. This means that people must eat foods containing these eight amino acids to provide the complete proteins necessary for growth. All of these eight essential amino acids are present in a variety of animal products, including red meat, poultry and milk. In contrast, however, no source of plant food contains adequate supplies of all eight of the amino acids – they all lack at least one. Consider, for example, the beans. While the beans have more than enough of the essential amino acid lysine, they lack the amino acid methionine. Wheat and rice, on the other hand, contain suboptimal levels of lysine while containing a sufficient amount of methionine. When we consider that the majority of the world's population exists on a diet which rarely, if ever, contains meat, we realise how important it is for people to be able to get sufficient quantities of each amino acid from a plant-based diet. Therefore, it is desirable that plants contain complete proteins with all the amino acids required for human body.

The scientists are currently using the knowledge gained through genetic engineering in an attempt to alter the genes of a variety of plant proteins. For example, the genes of the beans could be altered so as to encode a protein which contains sufficient quantities of methionine. If the gene for phaseolin, the primary protein molecule of a beans could be altered to contain codons which specify the amino acid methionine without altering the overall structure or growth pattern of the plant, phaseolin could become a complete protein.

In addition to beans, the 'high lysine corn' is also worth mentioning. In cereals, the major seed storage proteins are called prolamines (zeins in corn) and they have been found to have a lower lysine content. Therefore, the diets based on cereal grains are deficient in lysine. In the case of corn, the seed proteins are also deficient in tryptophan, and to a lesser extent, methionine. Because of the importance of cereals to man and animals, scientists have been working for several decades to develop varieties with increased lysine, tryptophan and methionine content. Considerable progress has been made, but the successes obtained had little

agricultural value. Some corn mutants such as *sugary-1, floury-2* and *opaque-2* have increased lysine and/or methionine content, but these strains have soft kernels and lower yields. These high lysine strains also have lower prolamine (zein) content.

Several corn genes encoding zeins have been cloned and sequenced. After this success, a blue-print for engineering 'high-lysine' corn by site-specific mutagenesis was prepared. Then these high-lysine coding sequences could be joined to strong promoters and reintroduced into corn plants by means of electroporation or a microprojectile gun. Recently, B.A. Larkins and colleagues introduced new lysine and tryptophan codons into a zein cDNA by site specific mutagenesis. And when RNA transcripts of these modified cDNAs were injected into Xenopus laevis oocytes, RNAs were translated efficiently and the 'high- lysine' zein products were found to self-aggregate into dense structures as seen in corn. These results give positive indication that 'high-lysine' corn might indeed be produced by means of genetic engineering. In fact, by using the same approach, one can tailor other plant storage proteins to the optimal amino acid compositions. Thus, by employing the genetic engineering techniques, the nutritional quality of seed storage proteins can be desirably improved.

24.2.9 Manipulating Gene Expression Using Antisense RNA

An unusual approach to the control of gene expression has scored preliminary success in changing an important commercial trait of tomatoes. Fresh tomatoes must be shipped or transported while still green, because ripe fruit is too soft and is bruised easily. The enzyme polygalacturonase, which breaks down plant cell walls, is primarily responsible for fruit softening. It has been found that transformation of an 'antisense' copy of the polygalacturonase gene in tomato plants decreases expression of the softening protein as much as by 90%.

The antisense gene was made by fusing a cDNA clone of the polygalacturonase in reverse orientation, to a constitutive promoter. This reverse or antisense gene was then linked to the *T-DNA* and transferred to the tomato plants. Each transformed plant thus carried an antisense gene, as well as a normal copy of the polygalacturonase gene. During gene transcription, antisense messenger RNA (mRNA) molecules are produced and these are complementary to normal mRNA molecules (Fig. 24.9). It is believed that antisense mRNAs bind to a proportion of normal RNAs, making them unavailable for translation into protein.

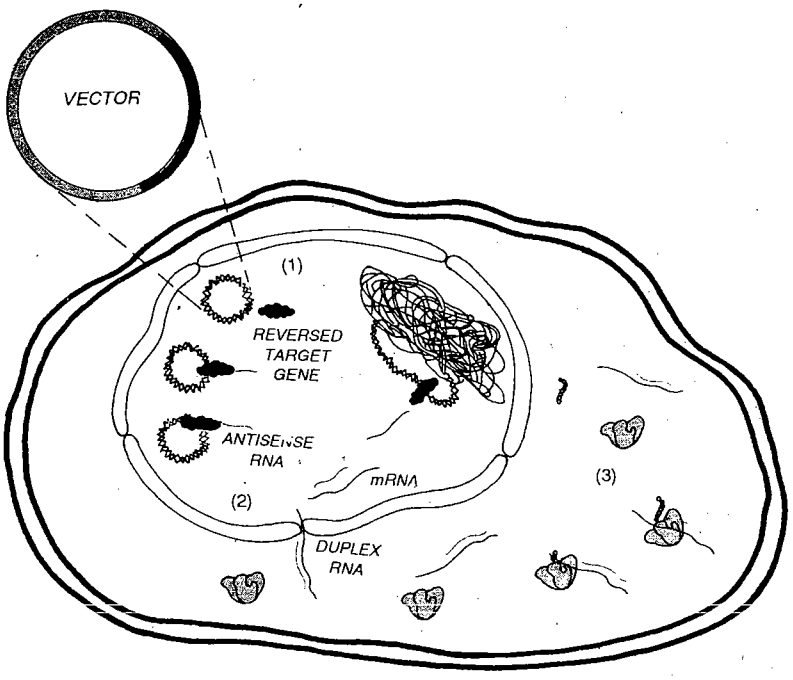


Fig.24.9: Antisense RNA. a) The target gene in the plant chromosome expresses normal mRNA chromosome, and plasmid vector expresses complementary antisense RNA. b) Antisense RNA hybridises with the normal mRNA. c) When the ribosomes encounter the duplex RNA, translation is halted.

SAQ 2

Fill in the blanks with appropriate words:

i) Resistance to frost formation was developed by ..... a gene with nucleation ability in the *Pseudomonas syringae*.

ii) *Klebsiella*, a free-living bacterium having 17 ..... genes, organised into 7 operons, is a valuable material for research on enhanced ..... utilisation in plants.

iii) Beans have sufficient amount of amino acid ..... , but usually they are deficient in ..... .

iv) In cereals, that constitute the staple diet of a large section of people, the major seed storage proteins are .....

v) The result of transformation of an antisense copy of the ..... gene in tomatoes was better keeping quality of the fruit.

24.3 SUMMARY

This unit provides you an overview with the help of some representative examples, of some of the ongoing and anticipated applications of genetics in plant improvement. With the development of recombinant DNA and gene cloning techniques, biologists are now able to isolate and dissect almost any gene and any chromosomal segment of an organism. Transformation of plants with vectors based on *Ti plasmid* have opened the way for the genetic engineering of plants using recombinant DNA technology. It is expected that many types of desirably improved plants will result from the application of this new technology.

24.4 TERMINAL QUESTIONS

1) Write short notes on the following:
   i) Significance of regeneration of plantlets from protoplasts,
   ii) Microprojectile gun,
   iii) Herbicide-tolerant plants,
   iv) Ice-minus bacterium,
   v) Antisense RNA
2) Highlight the potential use of tissue culture in crop improvement.
3) Starting from the infection of a wound in a dicotyledonous plant, by *Agrobacterium tumefaciens*, outline the molecular events involved in the development of crown gall.
4) Comment on the following statement. 'The *Ti plasmid* of *Agrobacterium tumefaciens* – one of the most valuable tools in genetic engineering, has tremendous potential in plant improvement.' Support your argument with suitable examples.

5) Discuss the strategies adopted for imparting herbicide resistance to plants.

6) Explain with the help of a suitable example, the use of *Ti plasmid* in developing pest resistance in plants.

24.5 ANSWERS

Self Assessment Questions

1) i) b
   ii) e
   iii) d
   iv) a
   v) c

2) i) removing
   ii) *nif*, nitrogen
   iii) lysine, methionine
   iv) prolamines
   v) polygalacturonase

Terminal Questions

1) i) See Subsection 24.2.1
   ii) See Subsection 24.2.3
   iii) See Subsection 24.2.4
   iv) See Subsection 24.2.6
   v) See Subsection 24.2.9

2) See Subsection 24.2.1
3) See Subsection 24.2.2
4) See Subsection 24.2.2
5) See Subsection 24.2.4
6) See Subsection 24.2.5

GLOSSARY

adaptive value : refers to relative reproductive success of a genotype measured in terms of fecundity or the number of offspring left behind. Adaptive value is synonymous with Darwanian fitness and is symbolised as ω.

agglutinin : antibodies produced by agglutinogen that specifically react with the same agglutinogen (antigen). A test for agglutination makes it possible to classify people into four groups.

agglutinogen/Isoagglutinogen are the antigens that agglutinate or clump the red blood cells of the same species. Agglutination is caused by the interaction of the antigen on the surface of the cell with an antibody present in the serum.

allele : one of the forms of the gene at a particular locus that are members of the same gene pair, each kind of allele affecting a particular character somewhat differently than the others.

allelic frequency : the proportion of alleles in a population that belong to a given type.

allergy : originally defined as altered reactivity on second contact with antigen now usually refers to a type of hypersensitivity reaction.

alloantigen : antigen that differ between individuals of the same species.

allogenic : refers to intraspecies genetic variations.

amorph : a gene which has no demonstrable product like O gene.

antibody : a defense protein synthesised by mature B-cells (plasma cells) of a higher organism; it binds specifically to the foreign molecule that induced its synthesis.

antigen :a molecule which induces the formation of antibody.

antisense RNA : RNA that is made from the DNA strand that is complementary to the sense strand of the DNA.

bactericidal : chemotherapeutic agent, i.e., compounds that have an irreversible lethal action.

basque : a member of a group of people of unknown origin inhabiting in France and Spain.

bilirubin : a degradation product of haeme.

bottleneck effect : random changes in the gene frequencies observed in a population where the population size is drastically reduced to one generation.

complement : a group of serum proteins involved in the control of inflammation, activation of phagocytes and the lytic attack on cell membranes; the system can be activated by the interaction with the immune system.

electroporation : introduction of DNA fragments into a cell by means of an electric field.

epitopes : (antigenic determinants) small regions on the antigen that bind the antigen (paratope) to antibody.

erythroblast : one of the intermediate cells in the biosysthesis of erythrocytes.

estrogen : a female sex hormone.

exon : a protein coding region of a eukaryotic gene, the RNA transcribed form such a region.

founder effect : random changes in the gene frequencies observed in a small population founded as a non-representative of a larger population.

gall : a tumourous growth in plants.

gene flow : refers to movement of individuals from one population to another that results in either the formation of new alleles into the population or alteration in the frequencies of existing alleles.

genetic cloning : a process by which a large number of a DNA segment is produced after introducing the segment into a plasmid or other suitable vector. The replication of the cell or the phage vector results in clones.

genetic counselling : a communication process which explores the risks of producing a genetically defective child by a couple, especially when family history

of genetic disease exists.

genetic drift : random changes in allelic frequency that result from sampling of gametes from generation to generation; characterised of small population.

genetic engineering : linking two DNA molecules by in vitro manipulations for the purpose of creating a novel organism with desired characteristics.

genetic equilibrium : refers to a state where there is no change in the allelic frequenies of the population.

genotype : The genetic material inherited from parents, not all of it is necessarily expressed in the individual.

genotypic frequency : relative proportion of individuals in a population with a designated genotype.

heme : the iron-porphyrin prosthetic group of heme proteins.

hemolysis : disintegration of RBC membranes with subsequent release of haemoglobin.

heterozygote : a heterozygous individual with unlike members of any given pair or series of alleles that consequently produces unlike gametes.

HLA : the human leucocyte antigen also known as major histocompatibility complex (HMC).

homozygote : homozygote is an individual possessing a pair of identical alleles at corresponding loci on the homologous chromosome.

hybrid vigour (heterosis) : unusual growth, strength and health of heterozygous hybrids from two less vigourous homozygous parents.

hydrophobic : "water hating" non-polar molecules or groups that are insoluble in water.

immune response : the capacity of a vertebrate to generate antibodies to an antigen, a macromolecule foreign to it.

immunoglobulin : an antibody protein generated by a specific antigen.

locus : the site on a chromosome where a gene is located.

Mendelian population : a group of organisms of same species showing a common gene pool and capable of interbreeding.

mutation : a change in the nucleotide base pairs of a gene, or a rearrangement of genes within chromosomes so that their interactions produce different effects; a change in the chromosomes themselves.

natural selection : a process in nature whereby one genotype leaves behind more offspring than another genotype because of species adaption to the environment.

phenotype : the physical chemical expression of an organism's gene.

plasmid : an extrachromosomal genetic element consisting of double-stranded DNA that replicates autonomously from the host chromosome.

platelets : cell fragments also called thormbocytes found in the blood, synthesised in bones marrow that help in clotting.

pseudogene : a non-coding sequence of genomic DNA with homology to the given gene being probed.

randomly mating population : population consisting of individuals in which the probability of members mating with individuals of particular genotype is equal to their frequency in the population.

recombinant DNA technology : refers to techniques of gene cloning. The term recombinant DNA refers to the hybrid of foreign vector DNA.

restriction enzymes : These are endonucleases capable of recognising specific DNA sequences which they cleave. Useful enzymes in recombinant DNA technology.

restriction fragment : a DNA segment excised from a larger DNA by restriction enzymes.

selection coefficient : the reduction in the relative fitness of a genotype. The genotype may be less well adapted to the environment and this may cause reduction in fecundity. Selection coefficient is related to adaptive value by the equation s (1-w).

totipotency : the capability of any cell of an organism to differentiate and develop into a complete organism.

FURTHER READING

1. Carlson, E. A. 1985. *Human Genetics*. Tata McGraw Hill Publishing Co. Limited, New Delhi.
2. Freifelder, David. 1993. *Molecular Biology*. Second Edition. Narosa Publishing House, New Delhi.
3. Gardner, E. J.; Simmons, M. J. & Snustad, D. P. 1991. Eight Edition. *Principles of Genetics*. John Wiley & Sons, Inc. New York.
4. Gupta, P.K. 1987. Second Edition. *Genetics*. Rastogi Publications, Meerut.
5. Klug, W.S. & Cummings, M.R. 1986. *Concepts of Genetics*. Second Edition. Scott Foresman & Co., Illinois.
6. Rao, K.R. 1986. Third Edition. *Text Book of Biochemistry*. Prentice Hall of India (Pvt) Ltd., New Delhi.
7. Lloyd, J.R. 1986. *Genes and Chromosomes*. English Language Book Society, Macmillan, Hong Kong.

Dear Student,

While studying these units you may have found certain portions of the text difficult to comprehend. We wish to know your difficulties and suggestions in order to improve the course. Therefore, we request you to fill and send us the following questionnaire which pertains to this block.

QUESTIONNAIRE

LSE-03
Block-4

Enrolment No. | | | | | | | | | | |

1) How many hours did you need for studying the units?

Unit Number	19	20	21	22	23	24
No. of hours						

2) How Many hours (approximately) did you take to do the assignments pertaining to this block?

Assignment Number	
No. of hours	

3) In the following table we have listed 4 kinds of difficulties that we thought you might have come across. Kindly tick (√) the type of difficulty and give the relevant page number in the appropriate columns.

Page Number	Types of difficulties			
	Presentation is not clear	Language is difficult	Diagram is not clear	Terms are not explained

4) It is possible that you could not attempt some SAQs and TQs.

In the following table are listed the possible difficulties. Kindly tick (√) the type of difficulty and the relevant unit and question numbers in the appropriate columns.

Unit No.	SAQ No.	TQ No.	Types of difficulties			
			Not clearly posed	Cannot answer on basis of information given	Answer given (at end of Unit) not clear	Answer given is not sufficient

5) Were all the difficult terms included in the glossary. If not, please list in the space given below.

6) Any Other Suggestion(s)

To

The Course Coordinator (LSE-03; Genetics)
School of Sciences
Indira Gandhi National Open University
Maidan Garhi
New Dehli-110 068.