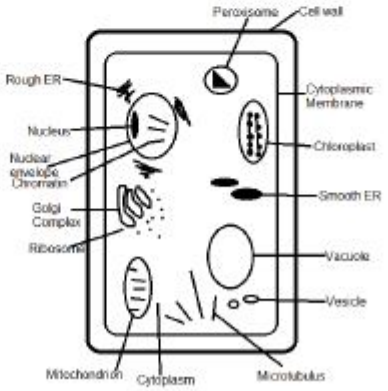
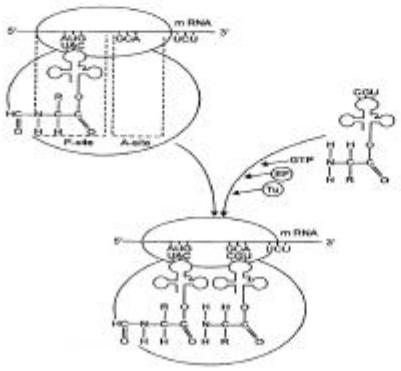
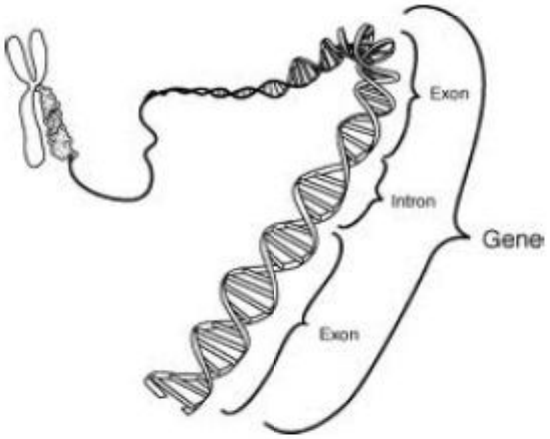
MBO-03

Vardhman Mahaveer Open University, Kota







Cell Biology, Genetics, Biostatistics and Computational Biology

Course Development Committee

Chair Person

Prof. Vinay Kumar Pathak
Vice-Chancellor
Vardhman Mahaveer Open University, Kota

Coordinator and Members

Convener

Dr. Anuradha Dubey
Department of Botany
School of Science & Technology
Vardhman Mahaveer Open University, Kota

Members

- **Prof. L.R.Gurjar**
  Director (Academic)
  Vardhman Mahaveer Open University, Kota
- **Dr. Arvind Pareek**
  Director (Regional Centre)
  Vardhman Mahaveer Open University, Kota
- **Prof. B.L. Choudhary**
  Former Vice-Chancellor,
  MohanLal Sukhadia University, Udaipur
- **Prof. S.L. Kothari**
  Department of Botany
  University of Rajasthan, Jaipur
- **Dr. G.P. Singh**
  Department of Botany
  University of Rajasthan, Jaipur
- **Prof. T.N. Bhardwaj**
  (Special Invite)
  Former Vice-Chancellor,
  Vardhman Mahaveer Open University, Kota
- **Dr. P.K. Sharma**
  Department of Botany
  MSJ College, Bharatpur
- **Dr. P.P. Paliwal**
  Department of Botany
  Govt. PG College, Banswara
- **Dr. Ekta Menghani**
  Department of Botany
  JECRC University, Jaipur

Editing and Course Writing

Editor

Dr. P.K. Sharma
Department of Botany
M.S.J. College, Bharatpur

Writers:

- **Dr. Shalini Malhotra** 1,2
  Department of Botany
  Tagore Biotech College,
  Jaipur
- **Dr. Aparna Pareek** 3,9,10
  Department of Botany
  University of Rajasthan,
  Jaipur
- **Dr. Vandana Sharma** 4,5
  Department of Botany
  Govt. College, Kota
- **Dr. Jang Bahadur Khan** 6,7,8
  Department of Botany
  Govt. College, Churu
- **Dr. Rishikesh Meena** 11,12,13
  Department of Botany
  Govt. College, Kota
  University of Rajasthan,
  Jaipur
- **Dr. Neerja Srivastava** 14
  Department of Botany
  Govt. College, Kota
- **Dr. Arvind Pareek** 15
  Director (Regional Centre)
  Vardhman Mahaveer Open
  University, Kota

Production : June 2015

Printed and Published on behalf of the Registrar, V.M. Open University, Kota.

Printed by : Pragya Publication Pvt. Ltd., Mathura

Vardhman Mahaveer Open University, Kota

Index

Preface

The present book entitled "**Cell Biology, Genetics, Biostatistics and Computational Biology**" has been designed so as to cover the unit-wise syllabus of MBO-03 course for M.Sc. Botany (Previous) students of Vardhman Mahaveer Open University, Kota. The basic principles and theory have been explained in simple, concise and lucid manner. Adequate examples, diagrammes, photographs and self-learning exercises have also been included to enable the students to grasp the subject easily. The unit writers have consulted various standard books on the subject and they are thankful to the authors of these reference books.

Unit-1

The Dynamic Cell

NOTES

Structure of the unit

1.0 Objectives

After going through this unit you will be able to understand:

- Structural organization of plant cell and various cell organelles
- Different types of specialized plant cells
- Chemical foundation of cell, different types of bonding and nonbonding interactions

- Biochemical energetic and various energy transactions of cell

1.1 Structural Organization of Plant Cell

1.1.1 Introduction

NOTES

Each cell is a community of subcellular components. Each type of component has its own particular set of functions. The individual parts could not survive for long outside the cell, but within the cellular environment they support each other so effectively that the cell as a whole is a viable entity. This subcellular cooperation not only ensures survival, but also provides for growth and multiplication of the cell (if given the necessary nutrients) and ultimately differentiation for a particular function. Looking at a higher level of organization, multicellular organisms are cooperative communities of cells, tissues and organs, all analogous to the subcellular components in that each contributes in a specialized way to the life of the system of which it is a part. No matter how complex the system, however, it is the cell that is the simplest, indivisible, unit which is viable hence the common statement that the cell is the unit of life.

All living systems have the structural and functional similarities of plant and animal cells. There are, in addition, cellular features in which the two kingdoms differ, mostly deriving from two major events in the evolution of living organisms - the development of a cell wall and the acquisition of photosynthetic capabilities.

Structurally, plant and animal cells are very similar because they are both eukaryotic cells.They both contain membrane-bound organelles such as the Nucleus, Mitochondria, endoplasmic Reticulum, Golgi apparatus, Lysosomes, and Peroxisomes. Both also contain similar membranes, cytosol, and cytoskeletal elements. The functions of these organelles are extremely similar between the two classes of cells (peroxisomes perform additional complex functions in plant cells having to do with cellular respiration). However, the few differences that exist between plant and animals are very significant and reflect a difference in the functions of each cell. Plant cells can be larger than animal cells. The normal range for an animal cell varies from 10 to 30 micrometers while that for a plant cell stretches from 10 to 100 micrometers. Beyond size, the main structural differences between plant and animal cells lie

in a few additional structures found in plant cells. These structures include: chloroplasts, the cell wall and vacuoles.

Plant cells vary in the extent to which different functions are developed, for, as with most multi-cellular organisms, plants exhibit division of labour. As a result of the varied requirements of maintaining life and supporting growth and development, specialized cells develop for protection, mechanical support, synthesis or storage of food reserves, transport, absorption and secretion, meristematic activity, reproduction, and the vital role of interconnecting the more specialized tissues.

NOTES

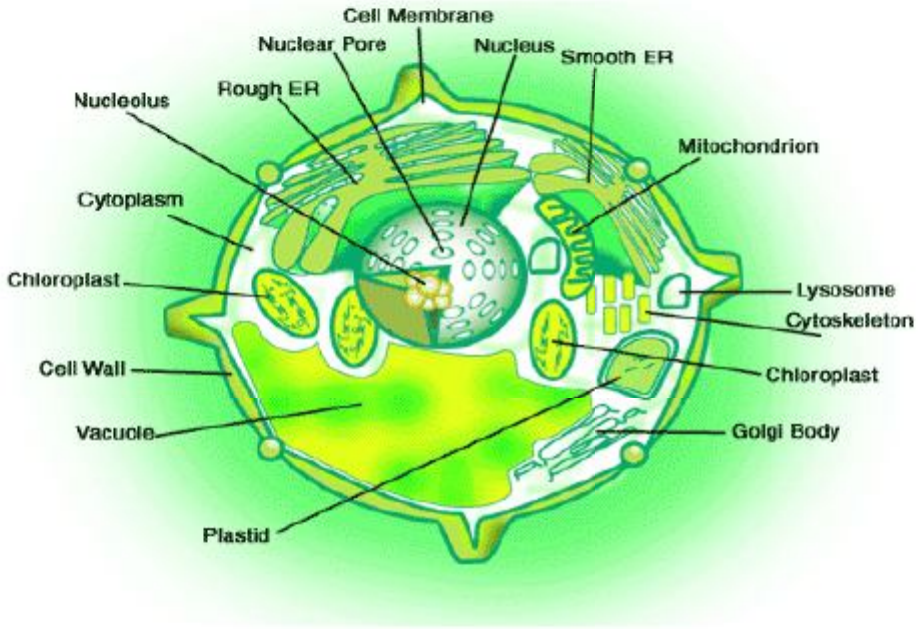


Fig. 1.1 : Plant Cell

1.1.2 Various Plant Cell Organelles

Plant tissues are composed of the non-living extracellular region and the living protoplasm of the cells proper. The former consists of intercellular spaces and cell walls. Each protoplast consists of a nucleus (or sometimes several nuclei) and the cytoplasm. Within these are the various membranous and non-membranous components which are:

1. Plasma membrane

The bounding membrane of the protoplast normally is in close contact with the inner face of the cell wall.

NOTES

2. Cytoplasm

A collective term for everything outside the nucleus, out to and including the plasma membrane. Includes membranous and other inclusions, and also the general matrix, or cytosol, in which the cytoplasmic components reside.

3. Nucleus

This is bounded by the nuclear envelope and contains genetic material in the form of chromatin, and the nucleolus (or, if more than one, nucleoli in a matrix of nucleoplasm.The cell nucleus is one of the largest organelles found in cells and also plays an important biological role. It composes about 10% of the total volume of the cell and is found near thecenter of eukaryotic cells. Its importance lies in its function as a storage site for DNA, our genetic material. The cell nucleus is composed of two membranes that form a porous nuclear envelope, which allows only select molecules in and out of the cell.The DNA that is found in the cell nucleus is packaged into structures called chromosomes.Chromosomes contain DNA and proteins and carry all the genetic information of an organism. The nucleus gains support from intermediate filaments that both form the surrounding nuclear lamina and makes direct contact with the endoplasmic reticulum.The nucleus is also the site of DNA and RNA synthesis.

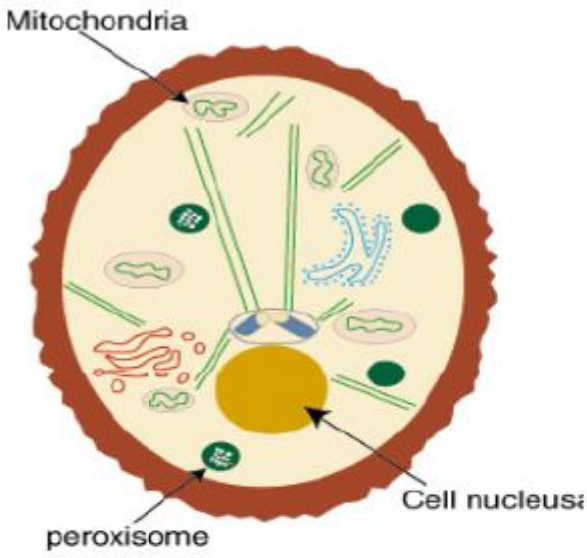


Fig. 1.2 : Mitochondria and Peroxisomes

4. Nucleoplasm

Everything enclosed by the nuclear envelope falls in the category of nucleoplasm, just as objects outside it are constituents of the cytoplasm. The word is often, however, used to denote the ground substance in which the chromatin and nucleolus lie.

NOTES

5. Chromatin

Chromatin contains the genetic material of the cell, i.e. information in the form of DNA that is passed from parent cell to daughter cell during the multiplication of cells and reproduction of the organism. It can exist in less dense (euchromatin) and more dense (heterochromatin) forms. During division of nuclei it is condensed into discrete units, chromosomes.

6. Nucleolus

A mass of filaments and particles, largely a sequence of identical repeating units of specialized genetic material together with precursors of ribosomes produced from that genetic information.

7. Nuclear envelope

A cisterna (a general term meaning a membrane-bound sac) wrapped around the contents of the nucleus. The space between the two membranous faces of the cisterna is the peri-nuclear space.

8. Nuclear envelope pores

Elaborate perforations in the nuclear envelope, involved in transport between nucleus and cytoplasm and in the processing of messenger and ribosomal RNA molecules that are being exported from the nucleus.

9. Endoplasmic reticulum

Membranous cisternae that ramify through the cytoplasm, occasionally connected to the outer membrane of the nuclear envelope. The bounding membrane segregates the contents of the cisterna from the cytoplasm. The outer face frequently bears attached ribosomes and polyribosomes. Endoplasmic reticulum is described as rough, or granular, and forms that lack ribosomes as smooth, or agranular. A special form that lies just inside the

plasma membrane is called cortical ER. ER cisternae may or may not have visible contents, which distend the cisternae when present in bulk.

NOTES

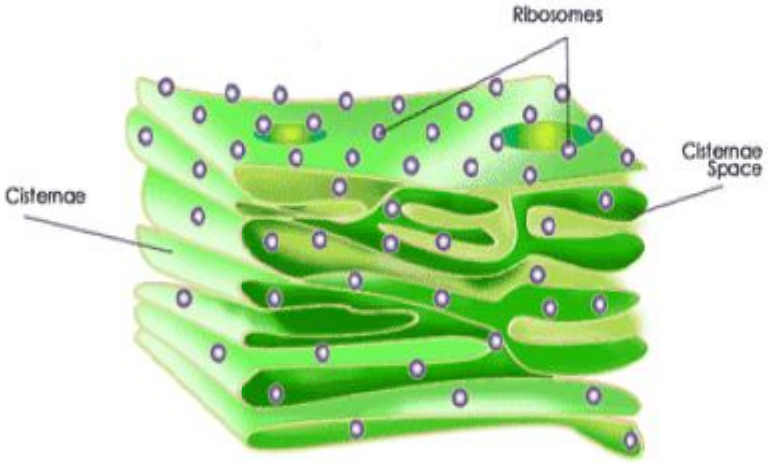


Fig. 1.3 : Position of Ribosomes on RER

The endoplasmic reticulum, or ER, is a very important cellular structure because of its function in protein synthesis and lipid synthesis. For example, the ER is the site of production of all transmembrane proteins. Since nearly all proteins that are secreted from a cell pass through it, the ER is also important in cellular trafficking. In addition to these major roles, the ER plays a role in a number of other biological processes. There are two different types of ER: smooth ER and rough ER. The rough ER has its name because it is coated with ribosomes, the structures most directly responsible for carrying out protein synthesis. Smooth ER lacks these ribosomes and is more abundant in cells that are specific for lipid synthesis and metabolism.

In addition to protein and lipid synthesis, the ER also conducts post-synthesis modifications. One such modification involves the addition of carbohydrate chains to the proteins, though the function of this addition is unknown. Another major modification is called protein folding, whose name is rather self-explanatory. Another role of the ER is to capture calcium for the cell from the cytosol. Finally, the ER can secrete proteins into the cell that are usually destined for the Golgi apparatus.

10. Ribosomes

Small particles of RNA and protein lying free in the cytoplasm or else attached to the endoplasmic reticulum. They aggregate in clusters, chains, spirals, or other polyribosome configurations when they are engaged in protein synthesis.

NOTES

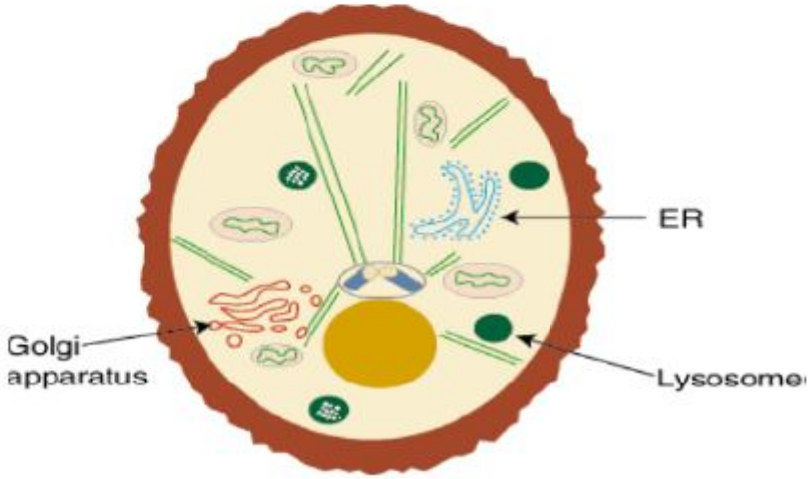


Fig. 1.4 : Golgibody and Endoplasmic Reticulum

11. Golgi bodies

Each Golgi body (Golgi stack) consists of layered cisternae together with many small vesicles that are involved in traffic to and from the Golgiapparatus and between its constituent cisternae.The Golgi apparatus is usually located near the cell nucleus. It is composed of a series of layers called Golgi stacks. Proteins from the ER always enter and exit the Golgi apparatus from the same location. The cis face of the Golgi is where proteins enter. A protein will make its way through the Golgi stacks to the other end called the trans face where it is secreted to other parts of the cell.

NOTES

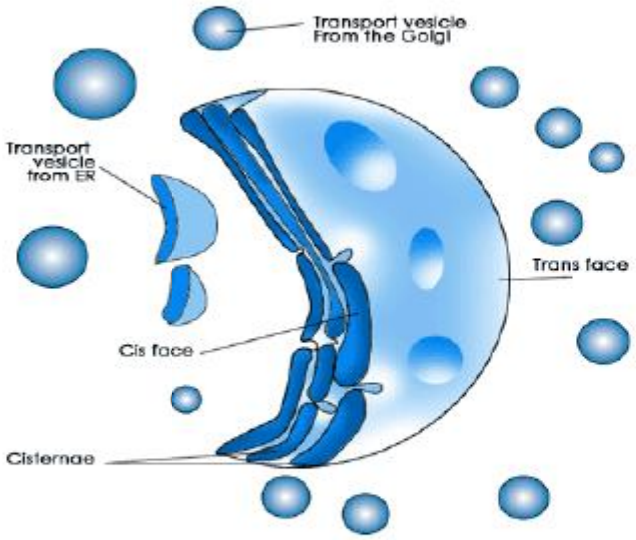


Fig. 1.5 : Vesicles formation from Golgibody

In the Golgi apparatus, more carbohydrate chains are added to the protein while other chains are removed. The Golgi stacks also sort proteins for secretion. After sorting, the membrane of the Golgi buds off, forming secretory vesicles that transport proteins to their specific destination in the cell. A protein's destination is often signaled with a specific amino acid sequence at its end. A protein secretion most often travels back to the ER or to the plasma membrane or to the lysosomes.

12. Vacuole

Compared with the surrounding cytoplasm, these are usually empty looking spaces, spherical when small. They are often very large, and can occupy 90% or more of the volume of the cell in mature tissues. Vacuoles are large, liquid-filled organelles found only in plant cells. Vacuoles can occupy up to 90% of a cell's volume and have a single membrane. Their main function is as a space-filler in the cell, but they can also fill digestive functions similar to lysosomes (which are also present in plant cells). Vacuoles contain a number of enzymes that perform diverse functions, and their interiors can be used as storage for nutrients or provide a place to degrade unwanted substances.

13. Mitochondria

These pleiomorphic bodies consist of a compartment, the matrix, surrounded by two membrane barriers, a double envelope. The outer membrane of the

double envelope is more or less smooth, but the inner is thrown into many folds - mitochondrial cristae - that project into the matrix.

14. Tonoplast

It is the membrane which bounds a vacuole. Except for its position in the cell it looks very like the plasma membrane.

NOTES

15. Plastids

This is a group name for a whole family of cell components. In the young root-tip cells of the first three plates this group is represented by the structurally simplest member, which is called the proplastid. Proplastids are usually larger than mitochondria, but, like them, have a double membrane envelope surrounding (in these examples) a fairly dense ground substance the stroma. Starch grains may be present in them. Other members of the plastid family are: chloroplasts etioplasts amyloplasts, and chromoplasts.

16. Chloroplast

In animal cells, the mitochondria produce the majority of the cells energy from food. It does not have the same function in plant cells. Plant cells use sunlight as their energy source; the sunlight must be converted into energy inside the cell in a process called photosynthesis. Chloroplasts are the structures that perform this function. They are rather large, double membrane bound structures (about 5 micrometers across) that contain the substance chlorophyll, which absorbs sunlight. Additional membranes within the chloroplast contain the structures that actually carry out photosynthesis.

Chloroplasts carry out energy conversion through a complex set of reactions similar to those performed by mitochondria in animals. The double membrane structure of chloroplasts is also reminiscent of mitochondria. The inner membrane encloses an area called the stoma, which is analogous to the matrix in mitochondria and houses DNA, RNA, ribosomes, and different enzymes. Chloroplasts, however, contain a third membrane and are generally larger than mitochondria.

17. Microbodies

These are bounded by a single membrane, and are distinguished from vesicles by their size and dense contents (sometimes including a crystal).

NOTES

18. The Cytoskeleton

The cytoskeleton is similar to the lipid bilayer in that it helps provide the interior structure of the cell the way the lipid bilayer provides the structure of the cell membrane.The cytoskeleton also allows the cell to adapt. Often, a cell will reorganize its intracellular components, leading to a change in its shape. The cytoskeleton is responsible formediating these changes. By providing "tracks" with its protein filaments, the cytoskeleton allows organelles to move around within the cell. In addition to facilitatingintracellular organelle movement, by moving itself the cytoskeleton can move the entire cells in multi-cellular organisms. In this way, the cytoskeleton is involved in intercellular communication.

The cytoskeleton is composed of three different types of protein filaments: actin, microtubules, and intermediate filaments.

1 Actin

Actin is the main component of actin filaments, which are double-stranded, thin and flexible structures. They have a diameter of about 5 to 9 nanometers. Actin is the most abundant protein in most eukaryotic cells. Most actin molecules work together to give support and structure to the plasma membrane and are therefore found near the cellmembrane.

2 Microtubules

Microtubules are long, cylindrical structures composed of the protein tubulin and organized around a centrosome, an organelle usually found in the center of the cell near the cell nucleus. Unlike actin molecules, microtubules work separately to provide tracks on which organelles can travel from the center of the cell outward. Microtubules are much more rigid than actin molecules and have a larger diameter: 25 nanometers. One end of each microtubule is embedded in the centrosome; the microtubule grows outward from there. Microtubules are relatively unstable and go through a process of continuous growth and decay. Centrioles are small arrays of microtubules that are found in the center of a centrosome. Certain proteins will use microtubules as tracks for laying out organelles in a cell.

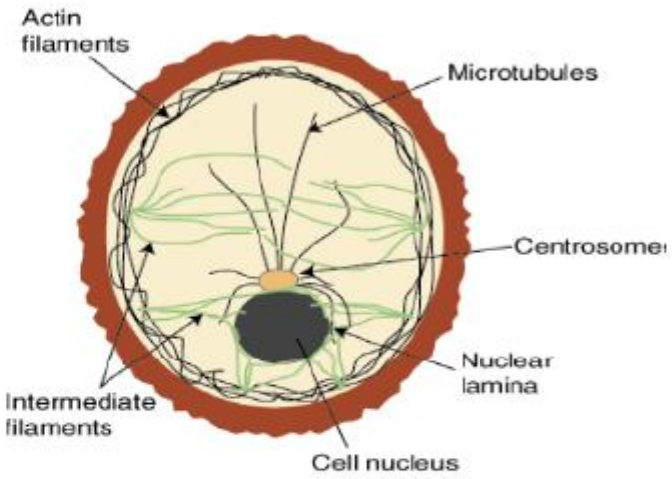


Fig. 1.6 : Cytoskeleton in Cell

3 Intermediate filaments

Intermediate filaments are the final class of proteins that compose the cytoskeleton. These structures are rope-like and fibrous, with a diameter of approximately 10 nanometers. They are not found in all animal cells, but in those in which they are present they form a network surrounding the nucleus often called the nuclear lamina. The filaments help to resist stress and increase cellular stability.

19. Peroxisomes

Peroxisomes are single-membrane structures found in all eukaryotic cells. They are small, membrane-bound structures that use molecular oxygen to oxidize organic molecules. The structure is one of the major oxygen utilizing organelles, the other being the mitochondria. Peroxisomes contain oxidative enzymes and other enzymes that help produce and degrade hydrogen peroxide. Because of their varying enzymatic compositions, peroxisomes are diverse structures. Their main function is to help breakdown fatty acids. They perform specific functions in plant cells.

20. Cell wall

This is a thin structure in meristematic cells, but it can be very massive and elaborate in mature cells. It is external to the living protoplast, but nevertheless

contributes very significantly to the life of the plant cell; indeed, along with plastids, it is the major determinant of the lifestyle of plants. One of its main constituents is microfibrillar cellulose - the most abundant macromolecule on Earth.

NOTES

1.2 Specialized Plant Cell Types

The major classes of cells differentiate from undifferentiated meristematic cells (analogous to the stem cells of animals) to form the tissue structures of roots, stems, leaves, flowers and reproductive structures.All plant cells have several common features, such as chloroplasts, a cell wall, and a large vacuole. In addition, a number of specialized cells are found only in vascular plants.

1.2.1 Meristematic cells

They give rise to all three fundamental mature cell types. Their major function is cell division and so their cell cycle indeed cycles. The walls are thin, the vacuole is largely missing, the plastids are immature, etc.

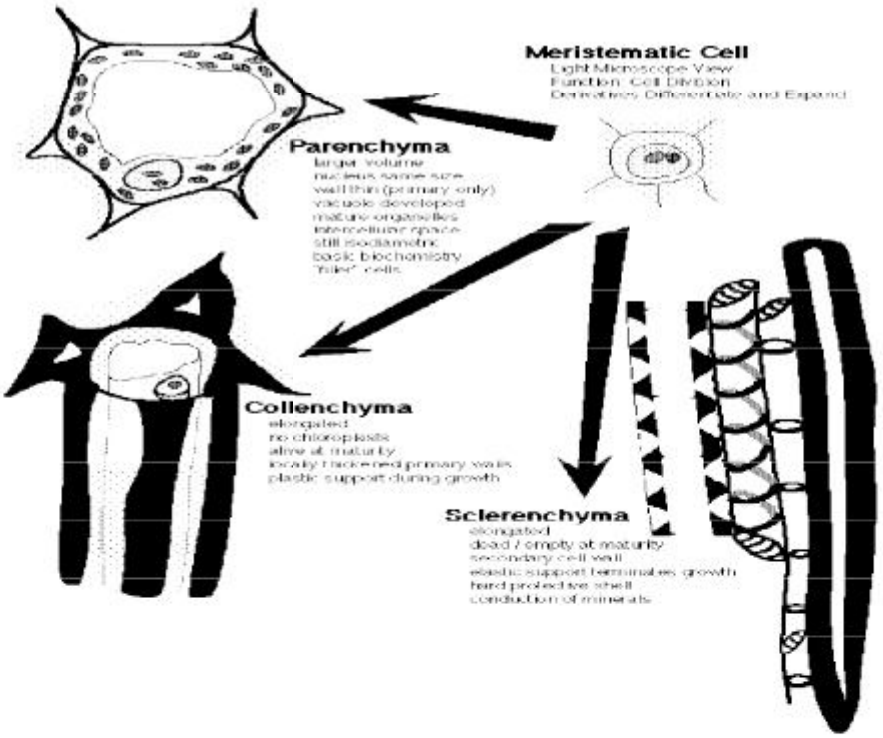


Fig. 1.7 : Specialized Plant Cell Types

1.2.2 Parenchyma cells

Parenchyma (from the Greek para, meaning "beside," and en + chein, meaning "to pour in") cells are the most common cells found in leaves, stems and roots. They are often spherical in shape with only primary cell walls. Parenchyma cells play a role in food storage, photosynthesis, and aerobic respiration. They are living cells that have functions ranging from storage and support to photosynthesis and phloem loading (transfer cells). Apart from the xylem and phloem in their vascular bundles, leaves are composed mainly of parenchyma cells. Some parenchyma cells, as in the epidermis, are specialized for light penetration and focusing or regulation of gas exchange, but others are among the least specialized cells in plant tissue and may remain totipotent, capable of dividing to produce new populations of undifferentiated cells, throughout their lives. Parenchyma cells have thin, permeable primary walls enabling the transport of small molecules between them and their cytoplasm is responsible for a wide range of biochemical functions such as nectar secretion, or the manufacture of secondary products that discourage herbivores. Parenchyma cells that contain many chloroplasts and are concerned primarily with photosynthesis are called chlorenchyma cells. Others, such as the majority of the parenchyma cells in potato tubers and the seed cotyledons of legumes, have a storage function.Other than support functions; this cell type is the basis for all plant function.

NOTES

1.2.3 Collenchyma cells

Collenchyma (from the Greek term kola, meaning "glue") cells have thickened primary cell walls and lack secondary cell walls. They form strands or continuous cylinders just below the surfaces of stems or leaf stalks. The most common function of collenchyma cells is to provide support for parts of the plant that are still growing, such as the stem. Similar to parenchyma cells, collenchyma cells are living cells at maturity and have only a primary wall. These cells mature from meristem derivatives that initially resemble parenchyma, but differences quickly become apparent. Plastids do not develop, and the secretory apparatus (ER and Golgi) proliferates to secrete additional primary wall. The wall is most commonly thickest at the corners, where three

NOTES

or more cells come in contact and thinnest where only two cells come in contact, though other arrangements of the wall thickening are possible.

Pectin and hemicellulose are the dominant constituents of collenchyma cell walls of dicot angiosperms. Collenchyma cells are typically quite elongated, and may divide transversely to give a septate appearance. The role of this cell type is to support the plant in axis still growing in length, and to confer flexibility and tensile strength on tissues. The primary wall lacks lignin that would make it tough and rigid, so this cell type provides what could be called plastic support – support that can hold a young stem or petiole into the air, but in cells that can be stretched as the cells around them elongate. Stretchable support (without elastic snap-back) is a good way to describe what collenchyma does. Parts of the strings in celery are collenchyma.

Collenchyma cells are also alive at maturity and have only a primary wall. These cells mature from meristem derivatives. They pass briefly through a stage resembling parenchyma, however they are determined to differentiate into collenchyma, and this fact is quite obvious from the very earliest stages. Plastids do not develop and secretory apparatus (ER and Golgi) proliferates to assist in the accumulation of additional primary wall. This is laid down where three or more cells come in contact. Areas of wall where only two cells come in contact remain as thin as those of parenchyma cells.

1.2.4 Sclerenchyma cells

Sclerenchyma cells (from the Greek skleros, hard) are hard and tough cells with a function in mechanical support. They are of two broad types – sclereids or stone cells and fibers. The cells develop an extensive secondary cell wall that is laid down on the inside of the primary cell wall. The secondary wall is impregnated with lignin, making it hard and impermeable to water. Thus, these cells cannot survive for long' as they cannot exchange sufficient material to maintain active metabolism. Sclerenchyma cells are typically dead at functional maturity, and the cytoplasm is missing, leaving an empty central cavity.

Functions for sclereid cells (hard cells that give leaves or fruits a gritty texture) include discouraging herbivory, by damaging digestive passages in small insect larval stages, and physical protection (a solid tissue of hard sclereid cells form the pit wall in a peach and many other fruits). Functions of fibers include

provision of load-bearing support and tensile strength to the leaves and stems of herbaceous plants. Sclerenchyma fibers are not involved in conduction, either of water and nutrients (as in the xylem) or of carbon compounds (as in the phloem), but it is likely that they may have evolved as modifications of xylem and phloem initials in early land plants.

NOTES

These cells are hard and brittle. The cells develop an extensive secondary cell wall (laid down on the inside of the primary wall). This wall is invested with lignin, making it extremely hard. Lignin, plus suberin and/or cutin make the wall waterproof as well. Thus, these cells cannot survive for long as they cannot exchange materials well enough for active (or even maintaining) metabolism. They are typically dead at functional maturity. The cytoplasm is missing by the time the cell can begin to carry out its function.

Functions for sclerenchyma cells include discouraging herbivory (hard cells that rip open digestive passages in small insect larval stages, hard cells forming a pit wall in a peach fruit), support (the wood in a tree trunk, fibers in large herbs), and conduction (hollow cells lined end-to-end in xylem with cytoplasm and end walls missing).

Sclerenchyma includes the fibers used for making thread and fabric.Particularly the fibers from flax those are spun and woven into linen. Sclerenchyma cells provide rigid support for the plant. There are two types of sclerenchyma cells-fiber and sclereid. Fiber cells are long, slender cells that usually form strands or bundles. Sclereid cells, sometimes called stone cells, occur singly or in groups and have various forms. They have a thick, very hard secondary cell wall. Most sclerenchyma cells are dead cells at maturity.

1.2.5 Xylem

Xylem (from the Greek term xylos, meaning "wood") is the main water-conducting tissue of plants and consists of dead, hollow, tubular cells arranged end to end. The water transported in xylem replaces that lost via evaporation through stomata. The two types of water-conducting cells are tracheids and vessel elements. Water flows from the roots of a plant up through the shoot via pits in the secondary walls of the tracheids. Vessel elements have perforations in their end walls to allow the water to flow between cells.

1.2.6 Phloem

NOTES

The two kinds of cells in the food-conducting tissue of plants, the phloem (from the Greek term phloios, meaning "bark"), are sieve cells and sieve-tube members. Sieve cells are found in seedless vascular plants and gymnosperms, while sieve-tube members are found in angiosperms. Both types of cells are elongated, slender, tube-like cells arranged end to end with clusters of pores at each cell junction. Sugars (especially sucrose), other compounds, and some mineral ions move between adjacent food-conducting cells. Sieve-tube members have thin primary cell walls but lack secondary cell walls. They are living cells at maturity.

1.2.7 Epidermis

Several types of specialized cells occur in the epidermis including guard cells, trichomes, and root hairs. Flattened epidermal cells, one layer thick and coated by a thick layer of cuticle, cover all parts of the primary plant body.

1.3 Chemical Foundation of Cell

The life of a cell depends on thousands of chemical interactions and reactions exquisitely coordinated with one another in time and space and under the influence of the cell's genetic instructions and its environment.

Life first arose in a watery environment, and the properties of this ubiquitous substance have a profound influence on the chemistry of life. Constituting 70–80 percent by weight of most cells, water is the most abundant molecule in biological systems. About 7 percent of the weight of living matter is composed of inorganic ions and small molecules such as amino acids (the building blocks of proteins), nucleotides (the building blocks of DNA and RNA), lipids (the building blocks of bio-membranes), and sugars (the building blocks of starches and cellulose), the remainder being the macromolecules and macromolecular aggregates composed of these building blocks.

Many biomolecules (e.g., sugars) readily dissolve in water; these water-liking molecules are described as hydrophilic. Other biomolecules (e.g., fats like triacylglycerols) shun water; these are said to be hydrophobic (water-fearing). Still other biomolecules (e.g., phospholipids), referred to as amphipathic, are a bit schizophrenic, containing both hydrophilic and hydrophobic regions. These are used to build the membranes that surround cells and their internal

organelles. The smooth functioning of cells, tissues, and organisms depends on all these molecules, from the smallest to the largest. Indeed, the chemistry of the simple proton (H) with a mass of 1 dalton (Da) can be as important to the survival of a human cell as that of each gigantic DNA molecule with a mass as large as 8.6×10^{10} Da (single strand of DNA from human chromosome 1).

NOTES

1.3.1 Atomic Bonds and Molecular Interactions

Strong and weak attractive forces between atoms are the glue that holds them together in individual molecules and permits interactions between different biological molecules. Strong forces form a covalent bond when two atoms share one pair of electrons ("single" bond) or multiple pairs of electrons ("double" bond, "triple" bond, etc.). The weak attractive forces of non-covalent interactions are equally important in determining the properties and functions of biomolecules such as proteins, nucleic acids, carbohydrates, and lipids. There are four major types of non-covalent interactions: ionic interactions, hydrogen bonds, van der Waals interactions, and the hydrophobic effect.

In many molecules, the bonded atoms exert different attractions for the electrons of the covalent bond, resulting in unequal sharing of the electrons. The extent of an atom's ability to attract an electron is called its electronegativity. A bond between atoms with identical or similar electronegativities is said to be nonpolar. In a nonpolar bond, the bonding electrons are essentially shared equally between the two atoms, as is the case for most C—C and C—H bonds. However, if two atoms differ in their electronegativity, the bond between them is said to be polar.

One end of a polar bond has a partial negative charge, and the other end has a partial positive charge. In an O—H bond, for example, the greater electronegativity of the oxygen atom relative to hydrogen results in the electrons spending more time around the oxygen atom than the hydrogen. Thus O—H bond possesses an electric dipole, a positive charge separated from an equal but opposite negative charge. We can think of the oxygen atom of the O—H bond as having, on average, a charge of 25 percent of an electron, with the H atom having an equivalent positive charge. Because of its two O—H bonds, water molecules (H_2O) are dipoles that form electrostatic, non-covalent interactions with one another and with other molecules. These interactions play

a critical role in almost every biochemical interaction and are thus fundamental to cell biology. The polarity of the O=P double bond in H_3PO_4 results in a "resonance hybrid".

NOTES

1. Covalent bonds

These are much stronger and more stable than non-covalent interactions because the energies required to break them are much greater than the thermal energy available at room temperature (25 ºC) or body temperature (37 ºC). For example, the thermal energy at 25 ºC is approximately 0.6 kcal/mol, whereas the energy required to break the carbon-carbon single bond in ethane is about 140 times larger. Covalent single bonds in biological molecules have energies similar to that of the C—C bond in ethane. Because more electrons are shared between atoms in double bonds, they require more energy to break than single bonds.

2. Ionic interactions

These are attractions between oppositely charged ions result into the attraction of a positively charged ion, a cation, for a negatively charged ion, an anion. In sodium chloride (NaCl), for example, the bonding electron contributed by the sodium atom is completely transferred to the chlorine atom. Unlike covalent bonds, ionic interactions do not have fixed or specific geometric orientations, because the electrostatic field around an ion, its attraction for an opposite charge is uniform in all directions. In aqueous solutions, simple ions of biological significance, such as Na^+, K^+, Ca^{2+}, Mg^{2+}, and Cl^-, do not exist as free, isolated entities. Instead, each is hydrated, surrounded by a stable shell of water molecules, which are held in place by ionic interactions between the central ion and the oppositely charged end of the water dipole. Most ionic compounds dissolve readily in water because the energy of hydration, the energy released when ions tightly bind water molecules, is greater than the lattice energy that stabilizes the crystal structure. Parts or all of the aqueous hydration shell must be removed from ions when they directly interact with proteins. For example, water of hydration is lost when ions pass through protein pores in the cell membrane during nerve conduction.

.

3. Hydrogen bonds

They determine water solubility of uncharged molecules. Hydrogen bond is the interaction of a partially positively charged hydrogen atom in a molecular dipole (e.g., water) with unpaired electrons from another atom, either in the same (intramolecular) or in a different (intermolecular) molecule. Normally, a hydrogen atom forms a covalent bond with only one other atom. However, a hydrogen atom covalently bonded to an electronegative donor atom D may form an additional weak association, the hydrogen bond, with an acceptor atom A, which must have a nonbonding pair of electrons available for the interaction: The length of the covalent D—H bond is a bit longer than it would be if there were no hydrogen bond, because the acceptor "pulls" the hydrogen away from the donor. An important feature of all hydrogen bonds is directionality. In the strongest hydrogen bonds, the donor atom, the hydrogen atom, and the acceptor atom all lie in a straight line. Nonlinear hydrogen bonds are weaker than linear ones; still, multiple nonlinear hydrogen bonds help to stabilize the three-dimensional structures of many proteins.

NOTES

4. Van der Waals interactions

These are caused by transient dipoles, when any two atoms approach each other closely; they create a weak, nonspecific attractive force called a van der Waals interaction. These nonspecific interactions result from the momentary random fluctuations in the distribution of the electrons of any atom, which give rise to a transient unequal distribution of electrons. If two non-covalently bonded atoms are close enough together, electrons of one atom will perturb the electrons of the other. This perturbation generates a transient dipole in the second atom, and the two dipoles will attract each other weakly. Similarly, a polar covalent bond in one molecule will attract an oppositely oriented dipole in another. Van der Waals interactions, involving either transiently induced or permanent electric dipoles, occur in all types of molecules, both polar and nonpolar. In particular, van der Waals interactions are responsible for the cohesion between molecules of nonpolar liquids and solids, such as heptane, that cannot form hydrogen bonds or ionic interactions with other molecules. The strength of van der Waals interactions decreases rapidly with increasing distance; thus these non-covalent bonds can form only when atoms are quite close to one another. However, if

NOTES

atoms get too close together, they become repelled by the negative charges of their electrons. When the van der Waals attraction between two atoms exactly balances the repulsion between their two electron clouds, the atoms are said to be in van der Waals contact. The strength of the van der Waals interaction is about 1 kcal/mol, weaker than typical hydrogen bonds and only slightly higher than the average thermal energy of molecules at 25 °C. Thus multiple van der Waals interactions, a van der Waals interaction in conjunction with other non-covalent interactions, or both are required to significantly influence intermolecular contacts.

5. Hydrophobic effects

They cause nonpolar molecules to adhere to one another because nonpolar molecules do not contain charged groups, possess a dipole moment, or become hydrated, they are insoluble or almost insoluble in water and they are hydrophobic. The covalent bonds between two carbon atoms and between carbon and hydrogen atoms are the most common nonpolar bonds in biological systems. Hydrocarbons are molecules made up only of carbon and hydrogen and are virtually insoluble in water. Large triacylglycerol (or triglycerides), which comprise animal fats and vegetable oils, also are insoluble in water. Nonpolar molecules or nonpolar portions of molecules tend to aggregate in water owing to a phenomenon called the hydrophobic effect. Because water molecules cannot form hydrogen bonds with nonpolar substances, they tend to form "cages" of relatively rigid hydrogen-bonded pentagons and hexagons around nonpolar molecules. This state is energetically unfavorable because it decreases the randomness (entropy) of the population of water molecules. (The role of entropy in chemical systems is discussed in a later section.) If nonpolar molecules in an aqueous environment aggregate with their hydrophobic surfaces facing each other, there is a reduction in the hydrophobic surface area exposed to water. As a consequence, less water is needed to form the cages surrounding the nonpolar molecules, and entropy increases (an energetically more favorable state) relative to the unaggregated state. In a sense, then, water squeezes the nonpolar molecules into spontaneously forming aggregates. Rather than constituting an attractive force such as in hydrogen bonds, the hydrophobic effect results from an avoidance of an unstable state (extensive water cages around individual nonpolar molecules).

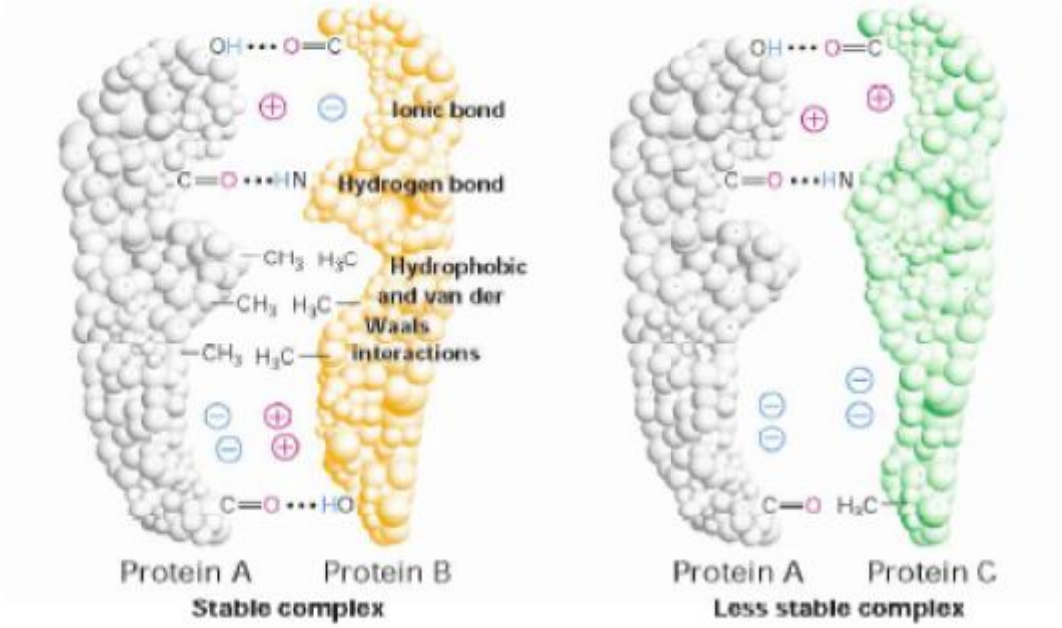


Fig.1.8 : Interaction of Different Bonds

1.3.2 Chemical Building Blocks of Cells

The three most abundant biological macromolecules are proteins, nucleic acids, and polysaccharides. They are all polymers composed of multiple covalently linked identical or nearly identical small molecules, or monomers.The covalent bonds between monomer molecules usually are formed by dehydration reactions in which a water molecule is lost.

Proteins are linear polymers containing ten to several thousand amino acids linked by peptide bonds. Nucleic acids are linear polymers containing hundreds to millions of nucleotides linked by phosphodiester bonds. Polysaccharides are linear or branched polymers of monosaccharides (sugars) such as glucose linked by glycosidic bonds.

1. Proteins

Amino acids differing only in their side chains compose proteins. The monomeric building blocks of proteins are 20 amino acids, all of which have a characteristic structure consisting of a central carbon atom (C) bonded to four different chemical groups: an amino (NH_2) group, a carboxyl (COOH) group, a hydrogen (H) atom, and one variable group, called a side chain, or R group. Because the carbon in all amino acids except glycine is asymmetric, these

NOTES

molecules can exist in two mirror-image forms called by convention the D (dextro) and the L (levo) isomers.

NOTES

The two isomers cannot be interconverted (one made identical with the other) without breaking and then re-forming a chemical bond in one of them. With rare exceptions, only the L forms of amino acids are found in proteins.Some amino acids are more abundant in proteins than other amino acids. Cysteine, tryptophan, and methionine are rare amino acids; together they constitute approximately 5 percent of the amino acids in a protein. Four amino acids leucine, serine, lysine, and glutamic acid, are the most abundant amino acids, totaling 32 percent of all the amino acid residues in a typical protein. However, the amino acid composition of proteins can vary widely from these values.

2. Nucleic Acids

Two types of chemically similar nucleic acids, DNA (deoxyribonucleic acid) and RNA (ribonucleic acid), are the principal information-carrying molecules of the cell. The monomers from which DNA and RNA are built, called nucleotides, all have a common structure: a phosphate group linked by a phosphoester bond to a pentose (a five-carbon sugar molecule) that in turn is linked to a nitrogen- and carbon-containing ring structure commonly referred to as a "base". In RNA, the pentose is ribose; in DNA, it is deoxyribose. The bases adenine, guanine, and cytosine are found in both DNA and RNA; thymine is found only in DNA, and uracil is found only in RNA.Adenine and guanine are purines, which contain a pair of fused rings; cytosine, thymine and uracil are pyrimidines, which contain a single ring. The bases are often abbreviated A, G, C, T and U respectively.

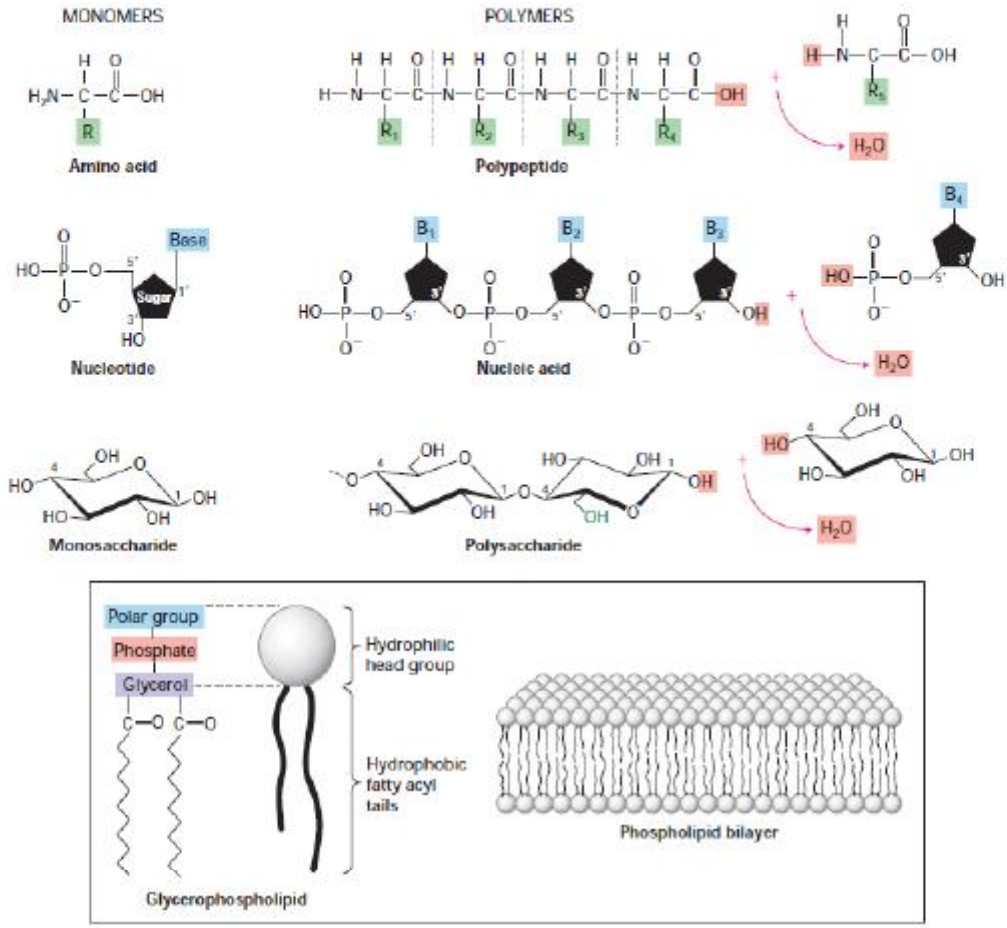


Fig. 1.9 : Different Monomers forming their Biopolymers

NOTES

3. Carbohydrates

Monosaccharides joined by glycosidic bonds form linear and branched polysaccharides. The building blocks of the polysaccharides are the simple sugars, or monosaccharides. Monosaccharides are carbohydrates, which are literally covalently bonded combinations ofcarbon and water in a one-to-one ratio $(CH_2O)_n$, where n equals 3, 4, 5, 6, or 7. Hexoses (n= 6) and pentoses (n = 5) are the most common monosaccharides. All monosaccharides contain hydroxyl (O—H) groups and either an aldehyde or a keto group: D-Glucose ($C_6H_{12}O_6$) is the principal external source of energy for most cells in higher organisms and can exist in three different forms: a linear structure and two differenthemiacetal ring structures. If the aldehyde group on carbon 1 reacts with the hydroxyl group on carbon 5, the resulting hemiacetal, D-glucopyranose, contains a six memberring. In theanomer of D-glucopyranose, the hydroxyl group attached to carbon 1 points "downward" fromthe ring; in

NOTES

theanomer, this hydroxylpoints "upward." In aqueous solutionanomers readily interconvert spontaneously; at equilibriumthere is about one-thirdanomer and two-thirds, withvery little of the open-chain form. Because enzymes can distinguishbetween theandanomers of D-glucose, theseforms have distinct biological roles. Condensation of the hydroxyl group on carbon 4 of the linear glucose with its aldehyde group results in the formation of D-glucofuranose, ahemiacetal containing a five-member ring. Although all three forms of D-glucose exist in biological systems, the pyranose form is by far the most abundant.

Disaccharides, formed from two monosaccharides, are the simplest polysaccharides. The disaccharide lactose, composed of galactose and glucose, is the major sugar in milk; the disaccharide sucrose, composed of glucose and fructose, is a principal product of plant photosynthesis and is refined into common table sugar.

Larger polysaccharides, containing dozens to hundreds of monosaccharide units, can function as reservoirs for glucose, as structural components, or as adhesives that help hold cells together in tissues. The most common storage carbohydrate in animal cells is glycogen, a very long, highly branched polymer of glucose. As much as 10 percent byweight of the liver can be glycogen. The primary storage carbohydrate in plant cells, starch, also is a glucose polymer. It occurs in an unbranched form (amylose) and lightly branched form (amylopectin). Both glycogen and starch are composed of the anomer of glucose. In contrast, cellulose, the major constituent of plant cell walls, is an unbranchedpolymer of the anomer of glucose. Human digestive enzyme scan hydrolyze the glycosidic bonds in starch, but not the glycosidic bonds in cellulose. Many species of plants, bacteria, and molds produce cellulose-degrading enzymes.

4. Lipids

Fatty acids are precursors for many cellular lipids. Fatty acids consist of a hydrocarbon chain attached to acarboxyl group (OCOOH). They differ in length; although the predominant fatty acids in cells have an even number of carbon atoms, usually 14, 16, 18, or 20. Fatty acids often are designated by the abbreviation Cx:y, where x is the number of carbons in the chain and y is the

number of double bonds. Fatty acids containing 12 or more carbon atoms are nearly insoluble in aqueous solutions because of their long hydrophobic hydrocarbon chains.Fatty acids with no carbon-carbon double bonds are said to be saturated; those with at least one double bond are unsaturated.

NOTES

Unsaturated fatty acids with more than one carbon-carbon double bond are referred to as polyunsaturated. Two "essential" polyunsaturated fatty acids, linoleic acid (C18:2)

and linolenic acid (C18:3), cannot be synthesized by mammalsand must be supplied in their diet. Mammals can synthesize other common fatty acids. Two stereoisomeric configurations, cis and trans, are possible around each carbon-carbon double bond.

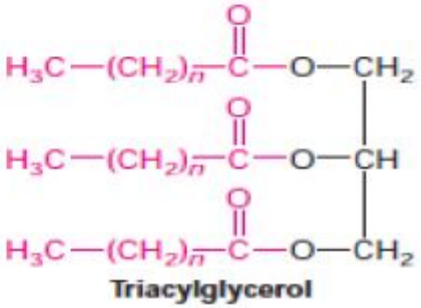


Fig. 1.10 : Triacyglycerol

Fatty acids can be covalently attached to another molecule by a type of dehydration reaction called esterification, in whichthe OH from the carboxyl group of the fatty acid and a H from a hydroxyl group on the other molecule are lost. In the combined molecule formed by this reaction, the portion derived from the fatty acid is called an acyl group, or fatty acyl group.

1.4 Biochemical Energetics

The production of energy, its storage, and its use are central to the economy of the cell. Energy may be defined as the ability to do work; a concept applicable to automobile engines and electric power plants in our physical world and to cellular engines in the biological world. The energy associated with chemical bonds can be harnessed to support chemical work and the physical movements of cells.

1.4.1 Biological Energy Transformations

NOTES

Thefirst law is the principle of the conservation of energy: for any physical or chemical change, the total amount of energy in the universe remains constant; energy may change form or it may be transported from one region to another, but it can not be created or destroyed. The second law of thermodynamics, which can be stated in several forms, says that the universe always tends toward increasing disorder: in all natural processes, the entropy of the universe increases.

1. Gibbs free energy G

It expresses the amount of energy that capable of doing work during a reaction at constant temperature and pressure. When are action proceeds with the release of free energy (that is, when the system changes so as topossess less free energy), the free-energy change,G, has a negative value and the reaction is saidto be exergonic. In endergonic reactions, the system gains free energy and G is positive.The units of G are joules/mole or calories/mole.

2. Enthalpy H

It is the heat content of the reacting system. It reflects the number and kinds ofchemical bonds in the reactants and products.When a chemical reaction releases heat, it is said to be exothermic; the heat content of the products is less than that of the reactants and H has, by convention, a negative value. Reacting systems that take up heat from their surroundings are endothermic and have positive values of H. The units of H are joules/mole or calories/mole.

3. Entropy S

It is a quantitative expression for therandomness or disorder in a system.When the products of a reaction are less complex and more disordered than the reactants, the reaction is said to proceed with a gain in entropy. The units of S are joules/mole Kelvin (J/mol K).

Under the conditions existing in biological systems (including constant temperature and pressure), changes in free energy, enthalpy, and entropy are related to each other quantitatively by the equation.

$$\Delta G = \Delta H - T\Delta S$$

Direction of a chemical reaction, the change in free energy determines the direction of a chemical reaction. Because biological systems are generally held

at constant temperature and pressure, it is possible to predict the direction of a chemical reaction from the change in the free energy G, named after J. W. Gibbs, showed that "all systems change in such a way that free energy G is minimized." In the case of a chemical reaction, reactants products, the change in free energy G is given by

NOTES

$$\Delta G = G_{products} - G_{reactants}$$

The relation of G to the direction of any chemical reaction can be summarized in three statements:

- If G is negative, the forward reaction (from left to right as written) will tend to occur spontaneously.
- If G is positive, the reverse reaction (from right to left as written) will tend to occur.
- If G is zero, both forward and reverse reactions occur at equal rates; the reaction is at equilibrium.

An unfavorable chemical reaction can proceed if it is coupled with an energetically favorable reaction. Many processes in cells are energetically unfavorable ($\Delta G>0$) and will not proceed spontaneously. Examples include the synthesis of DNA from nucleotides and transport of a substance across the plasma membrane from a lower to a higher concentration. Cells can carry out an energy-requiring reaction by coupling it to an energy-releasing reaction if the sum of the two reactions has a net negative ΔG.

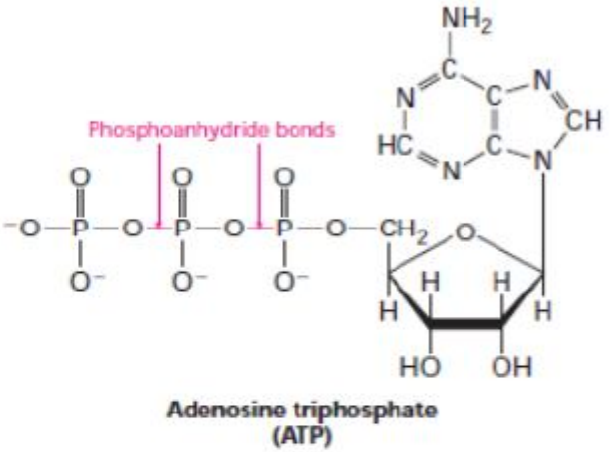


Fig. 1.11 : ATP

NOTES

4. Coupling Reaction

Hydrolysis of ATP releases substantial free energy and drives many cellular processes. In almost all organisms, adenosine triphosphate, or ATP, is the most important molecule for capturing, transiently storing, and subsequently transferring energy to perform work (e.g., biosynthesis, mechanical motion). The useful energy in an ATP molecule is contained in phosphoanhydride bonds, which are covalent bonds formed from the condensation of two molecules of phosphate by the loss of water. An ATP molecule has two key phosphoanhydride bonds. Hydrolysis of a phosphoanhydride bond in each of the following reactions has a highly negative free energy of about 7.3 kcal/mol.

ATP is generated during photosynthesis and respiration- In photosynthesis, plants and certain microorganisms can trap the energy in light and use it to synthesize ATP from ADP and Pi. Much of the ATP produced in photosynthesis is hydrolyzed to provide energy for the conversion of carbon dioxide to six-carbon sugars, a process called carbon fixation.

$$6\,CO_2 + 6\,H_2O \xrightarrow{\text{ATP} \;\rightarrow\; \text{ADP} + P_i} C_6H_{12}O_6 + 6\,O_2$$

In animals, the free energy in sugars and other molecules derived from food is released in the process of respiration. All synthesis of ATP in animal cells and in non-photosynthetic microorganisms results from the chemical transformation of energy-rich compounds in the diet (e.g., glucose, starch).

$$C_6H_{12}O_6 + 6\,O_2 \longrightarrow 6\,CO_2 + 6\,H_2O$$

1.5 Summary

Plant cells are complex cells in comparison to animal cells as they possess some special organelles such as chloroplast, vacuoles etc. All organelles work in unison to perform various complex tasks such as photosynthesis.

There are different specialized plant cell types too which perform their unique functions such as support to the plant, carrying food and water etc.

With the help of several bonding and nonbonding interactions different types of biomolecules interact with each other and perform different tasks such as working as building blocks for various bio-polymers.

Plant cell systems also follow thermodynamic laws and drive so many complex reactions such as photosynthesis and respiration according to these laws.

1.6 Glossary

NOTES

- **Actin :** A very abundant protein in eukaryotic cells that is the main component of actin filaments.
- **Actin Filaments :** They are approximately 5-9 nanometers in diameter. Provide structural support to the plasma membrane. As a cytoskeletal protein provides for movement of organelles within cells.
- **Cell wall :** The thick and rigid layer that covers the plasma membrane in plant cells.Composed of fat and sugar molecules in a matrix.
- **Centromere :** A round structure that holds together sister chromatids.
- **Centrosome :** A region of the cell near the nucleus from which microtubules sprout. Centrosomes are not found in all cells. Centrosomes are comprised of two centrioles.
- **Chlorophyll :** A pigment located within a chloroplast that absorbs light in plant cells, helping to convert light energy into biological energy in the process of photosynthesis.
- **Chloroplast :** A double membrane-bound organelle found in plant cells that contains chlorophyll and is responsible for mediating photosynthesis.
- **Chromosome :** A structure composed of DNA and proteins containing all the genetic material of a cell. Found in the cell nucleus.
- **Cytoplasm :** A fluid found in the main compartment of eukaryotic cells. Includes everything outside the cell nucleus but the organelles and the cytoskeleton. The main component is cytosol.
- **Cytoskeleton :** A system of protein filaments found throughout the cytoplasm of eukaryotic cells that help provide for cell structure. Composed of actin, intermediate filaments, and microtubules.
- **Cytosol :** The main component of the cytoplasm that fills the main compartment of eukaryotic cells.

NOTES

- **Endoplasmic Reticulum :** A membrane-bound organelle found in eukaryotic cells.Makes direct contact with the cell nucleus and, since it is dotted with ribosomes, is the site of lipid and protein synthesis. Comes in two forms, smooth and rough.
- **Endosome :** A membrane-bound organelle found in eukaryotic cells. Responsible for delivering molecules to the lysosome for digestion.
- **Eukaryote :** An organism composed of one or more cells with defined intracellular components including a nucleus and cytosol. Includes all organisms except bacteria and viruses.
- **Golgi apparatus :** A membrane-bound organelle found near the cell nucleus in eukaryotic cells. Responsible for sorting and packaging proteins for secretion to various destinations in the cell.
- **Intermediate Filament :** One of three protein components of the cytoskeleton.A fibrous protein filament approximately 10 nanometers in diameter.Forms the nuclear lamina that helps protect the cell nucleus.
- **Intermembrane Space :** The space between the outer and inner membrane in a mitochondria.
- **Lysosome :** A membrane-bound organelle found in eukaryotic cells. Contain acids and enzymes that degrade unwanted molecules.
- **Matrix :** The space inside the inner membrane of mitochondria.
- **Microtubule :** One of three protein components of the cytoskeleton, long, cylindrical structures approximately 25 nanometers in diameter. Extend from the centrosome to all parts of the cell, forming tracks on which organelles can travel within the cell.
- **Mitochondria :** An organelle within the cell. Much of cell respiration is carried out within its bounds.
- **Nucleus :** A large, double membrane-bound organelle found in eukaryotic cells. Contains DNA and RNA.
- **Peroxisome :** A small, membrane-bound organelle found in eukaryotic cells. Contains oxidizing enzymes that oxidize organic molecules and process hydrogen peroxide in the cell.

- **Photosynthesis :** A process in which plants convert sunlight into energy sources that can be used inside the cell to sustain life.
- **Prokaryote :** An organism composed of usually one, but occasionally more, cells that lack defined sub-cellular compartments. All essential material is enclosed within the cell membrane. Includes all bacteria and close relatives.
- **Ribosome :** A molecule composed of ribosomal RNA and proteins, and located on the endoplasmicreticulum. Responsible for mediating protein synthesis.
- **Rough endoplasmic reticulum :** Endoplasmic reticulum that is coated with ribosomes and involved in protein synthesis.
- **Smooth endoplasmic reticulum :** Naked endoplasmic reticulum that lacks ribosomes and is more involved in lipid synthesis.
- **Vacuole :** Membrane-bound, fluid-filled organelle found only in plant cells. Can compose up to 90% of a cell's volume and performs diverse functions in plant cells, including digestion of intracellular molecules.

NOTES

1.7 Self-Learning Exercise

Section -A (Very Short Answer Type Questions)

1. What is chromatin?
2. What is the significance of plant vacuoles?
3. What is the role of nucleolus?
4. What are collenchyma cells?
5. What is plant cell wall?
6. What is the main function of phloem?
7. Write down the equation of change in free energy.

Section -B (Short Answer Type Questions)

NOTES

1. Write down a short note on meristematic cells.
2. Differentiate between xylem and phloem.
3. Write down a short note on free energy.
4. What do understand by the term enthalpy?
5. Explain the role of amino acids as building blocks of cell.

Section -C (Long Answer Type Questions)

1. Write down an essay on plant cell organization.
2. Explain various kinds of specialized plant cell types.
3. Describe different types of chemical building blocks of cell.
4. Explain different forms of energies in biological systems.

Answer Key of Section-A

1. Chromatin is protein clustered with DNA, present in nucleus, carry genetic information.
2. They act as a space-filler in the cell, but they can also fill digestive functions similar to lysosomes.
3. Formation of rRNA.
4. They have thickened primary cell walls and lack secondary cell walls. They form strands or continuous cylinders just below the surfaces of stems or leaf stalks. They provide support for parts of the plant that are still growing,
5. Each cell in a plant is completely enclosed by, an elaborate extracellular matrix called the plant cell wall.

6. Transport of food.

7. $\Delta G = \Delta H - T\Delta S$.

1.8 References

NOTES

- Alberts, Bruce et al. Molecular Biology of the Cell. Third edition. Garland Publishing, Inc. New York, 1994.
- Grifiths, JF et al. Modern Genetic Analysis. Freeman Publishing, Inc. New York, 1999.
- Losick, Richard. Biological Sciences 10: Introduction to Molecular Biology. Harvard University, Department of Molecular and Cellular Biology, 1999.
- Voet, Donald and Judith Voet. Biochemistry. John Wiley & Sons, Inc, New York, 1995.
- Gerald Karp. Cell and Molecular Biology. Fifth edition. John Wiley & Sons, Inc. 2008.

Unit - 2

Cell Wall and Plasma Membrane

NOTES

Structure of the unit

2.0 Objectives

After going through this unit you will be able to understand:

- Structure and functions of Cell Wall along with its biogenesis
- Structure & functions of Plasma membrane and movement of various solutes across it

- Different types of receptors and their role
- Structure of Plasmodesmata, its functions, differences between plasmodesmata and gap junctions

2.1 Cell Wall

NOTES

Each cell in a plant is completely enclosed by, an elaborate extracellular matrix called the plant cell wall. It was the thick cell walls of cork, visible in a primitive microscope that in 1663 enabled Robert Hooke to distinguish and name cells for the first time. The walls of neighboring plant cells, cemented together to form the intact plant, are generally thicker stronger, and, most important of all, more rigid than the extracellular matrix produced by animal cells. In evolving relatively rigid wails, which can be up to many micrometers thick.

The composition of the cell wall depends on the celltype; all cell walls in plants have their origin in dividing cells, as the cell plate forms during cytokinesis to create a new partition wall between the daughter cells. The new cells are usually produced in special regions called meristems, and they are generally small in comparison with their final size. To accommodate subsequent cell growth, the walls of the newborn cells, called primary cell walls, are thin and extensible, although tough. Once growth stops, the wall no longer needs to be extensible: sometimes the primary wall is retained without major modification, but, more commonly, a rigid, secondary cell wall is produced by depositing new layers of matrix inside the old ones. These new layers generally have a composition that is significantly different from that of the primarywall. The most common additional polymer in secondary walls is lignin, a complex network of covalently linked phenolic compounds found in the walls of the xylem vessels and fiber cells of woody tissues.

Although the cell walls of higher plants vary in both composition and organization, they are all constructed, like animal extracellular matrices, using a structural principle common to all fiber-composites, including fiberglass and reinforced concrete. One component provides tensile strength, while another, inwhich the first is embedded, provides resistance to compression. While the principle is same in plants and animals, the chemistry is different. Unlike the

NOTES

animalextracellular matrix, which is rich in protein and other nitrogen-containing polymers, the plant cell wall is made almost entirely of polymers that contain no nitrogen, including cellulose and lignin. For a sedentary organism that depends on CO_2, H_2O, and sunlight, these two abundant biopolymers represent, cheap, carbon-based, structural materials, helping to conserve the scarce fixed nitrogen available in the soil that generally limits plant growth. Thus trees, for example, make a huge investment in the cellulose and lignin that comprise the bulk of their biomass.

In the cell walls of higher plants, the tensile fibers are made from the polysaccharide cellulose, the most abundant organic macromolecule on Earth, tightly linked into a network by cross-linking glycans. In primary cell walls, the matrix in which the cross-linked cellulose network is embedded is composed of pectin, a highly hydrated network of polysaccharides rich in galacturonic acid. Secondary cell walls contain additional molecules to make them rigid and permanent; lignin, in particular forms hard, waterproof filler in the interstices between the other components. All of these molecules are held together by a combination of covalent and noncovalent bonds to form a highly complex structure.

The plant cell wall thus has a "skeletal" role in supporting the structure of the plant as a whole, a protective role as an enclosure for each cell individually, and a transport role, helping to form channels for the movement of fluid in the plant. When plant cells become specialized, they generally adopt a specific shape and produce specially adapted types of walls, according to which the different types of cells in a plant can be recognized and classified.

2.1.1 Structure of Cell Wall

There are three major regions of the wall:

1. Middle Lamella

Middle lamella is outermost layer, glue that binds adjacent cells, composed primarily of pectin polysaccharides.

2. Primary Cell Wall

Primary cell wall deposited by cells before and during active growth. The primary walls of cultured cells are comprised of pectic polysaccharides, cross-

linking glycans (hemicellulose), cellulose and protein. The actual content of the wall components varies with species and age.All plant cells have a middle lamella and primary wall.

Growing plant cells are surrounded by a polysaccharide-rich primary wall. This wall is part of the apoplast which itself is largely self-contiguous and contains everything that is located between the plasma membrane and the cuticle. The primary wall and middle lamella account for most of the apoplast in growing tissue. The symplast is another unique feature of plant tissues. This self-contiguous phase exists because tube-like structures known as plasmodesmata connect the cytoplasm of different cells.

NOTES

(A) Primary Wall Composition

Primary walls isolated form higher plant tissues and cells are predominantly composed of polysaccharides together with lesser amounts of structural glycoproteins (hydroxyproline-rich extensins), phenolic esters (ferulic and coumaric acids), ionically and covalently bound minerals (e.g. calcium and boron), and enzymes. In addition walls contain proteins (expansins) that are believed to have a role in regulating wall expansion.

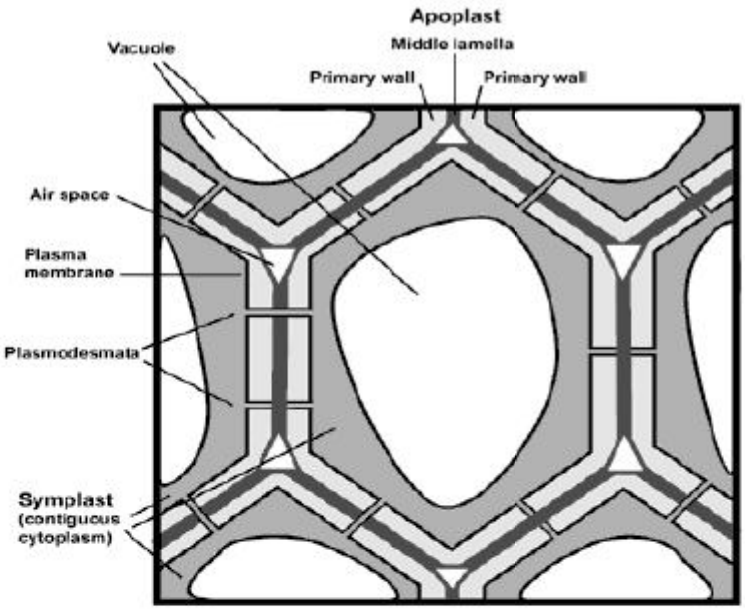


Fig. 2.1 : Primary Cell Wall

(B) The major Polysaccharides in the Primary Wall Cellulose

A polysaccharide composed of 1,4-linked β-D-glucose residues.

NOTES

Hemicellulose

Branched polysaccharides that are structurally homolgous to cellulose because they have a backbone composed of 1,4-linked β-D-hexosyl residues. The predominant hemicellulose in many primary walls is xyloglucan. Other hemicelluloses found in primary and secondary walls include glucuronoxylan, arabinoxylan, glucomannan, and galactomannan.

Pectin

A family of complex polysaccharides that contain 1,4-linked α-D-galacturonic acid. To date three classes of pectic polysaccharides have been characterized: Homo-galacturonans, rhamnogalacturo-nans, and substituted galacturonans.

3. Secondary Cell Wall

Some cells deposit additional layers inside the primary wall. This occurs after growth stops or when the cell begins to differentiate (specializes). The secondary wall is mainly for support and is primarily comprised of cellulose and lignin. Often it can distinguish distinct layers, S1, S2 and S3 - which differ in the orientation, or direction, of the cellulose micro fibrils.

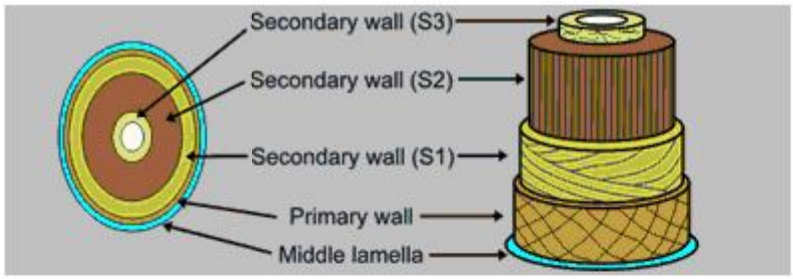


Fig. 2.2 : Secondary Cell Wall

Plants form two types of cell wall that differ in function and in composition. Primary walls surround growing and dividing plant cells. These walls provide mechanical strength but must also expand to allow the cell to grow and divide. The much thicker and stronger secondary wall, which accounts for most of the carbohydrate in biomass, is deposited once the cell has ceased to grow. The secondary walls of xylem fibers, tracheids, and sclereids are further strengthened by the incorporation of lignin.

The evolution of conducting tissues with rigid secondary cell walls was a critical adaptive event in the history of land plants, as it facilitated the transport

of water and nutrients and allowed extensive upright growth. Secondary walls also have a major impact on human life, as they are a major component of wood and are a source of nutrition for livestock. In addition, secondary walls may help to reduce our dependence on petroleum, as they account for the bulk of renewable biomass that can be converted to fuel. Nevertheless, numerous technical challenges must be overcome to enable the efficient utilization of secondary walls for energy production and for agriculture.

NOTES

2.1.2 Cell Wall Components

The main ingredients in cell walls are polysaccharides (complex carbohydrates/complex sugars) which are built from monosaccharides (simple sugars).Eleven different monosaccharides are common in these polysaccharides including glucose and galactose. Carbohydrates are good building blocks because they can produce a nearly infinite variety of structures. There are a variety of other components in the wall including protein, and lignin. These wall components in more detail are:

1. Cellulose

Cellulose is β1,4-glucan . Made up of as many as 25,000 individual glucose molecules. Every other molecule (called residues) is "upside down". Cellobiose (glucose-glucose disaccharide) is the basic building block. Cellulose readily forms hydrogen bonds with itself (intra-molecular H bonds) and with other cellulose chains (inter-molecular H-bonds). A cellulose chain will form hydrogen bonds with about 36 other chains to yield a micro fibril. This is somewhat analogous to the formation of a thick rope from thin fibers. Micro fibrils are 5-12 nm wide and give the wall strength - they have a tensile strength equivalent to steel. Some regions of the micro fibrils are highly crystalline while others are more "amorphous".

2. Cross-linking Glycans (Hemicellulose)

This is a diverse group of carbohydrates that used to be called hemicellulose.They are characterized by being soluble in strong alkali. They are linear (straight), flat, with a β-1,4 backbone and relatively short side chains. Two common types include xyloglucans and glucuronarabinoxylans. Other less common ones include glucomannans, galactoglucomannans, and

NOTES

galactomannans. The main feature of this group is that they don't aggregate with themselves - in other words, they don't form micro fibrils. However, they form hydrogen bonds with cellulose and hence the reason they are called "cross linking glycans". There may be a fucose sugar at the end of the side chains which may help keep the molecules planar by interacting with other regions of the chain.

3. Pectic Polysaccharides

These are extracted from the wall with hot water or dilute acid or calcium chelators (like EDTA).They are the easiest constituents to remove from the wall. They form gels (i.e., used in jelly making). They are also a diverse group of polysaccharides and are particularly rich in galacturonic acid (galacturonans= pectic acids).They are polymers of primarily β1,4 galacturonans are called homogalacturons (HGA) and are particularly common. These are helical in shape. Divalent cations, like calcium, also form cross-linkages to join adjacent polymers creating a gel.Pectic polysaccharides can also be cross-linked by dihydrocinnamic or diferulic acids.The HGA's (galacturonans) are initially secreted from the Golgi as methylated polymers; the methyl groups are removed by pectin methylesterase to initiate calcium binding.

Although most pectic polysaccharides are acidic, others are composed of neutral sugars including arabinans and galactans.The pectic polysaccharides serve a variety of functions including determining wall porosity, providing a charged wall surface for cell-cell adhesion - or in other words gluing cells together (i.e,. middle lamella), cell-cell recognition, pathogen recognition and others.

4. Protein

Wall proteins are typically glycoproteins (polypeptide backbone with carbohydrate side chains). The proteins are particularly rich in the amino acids hydroxyproline (hydroxyproline-rich glycoprotein, HPRG), proline (proline-rich protein, PRP), and glycine (glycine-rich protein, GRP). These proteins form rods (HRGP, PRP) or beta-pleated sheets (GRP).Extensin is a well-studied HRGP. HRGP is induced by wounding and pathogen attack. The wall proteins also have a structural role since: (1) the amino acids are characteristic

of other structural proteins such as collagen; and to extract the protein from the wall requires destructive conditions. Protein appears to be cross-linked to pectic substances and may have sites for lignification. The proteins may serve as the scaffolding used to construct the other wall components.Another group of wall proteins are heavily glycosylated with arabinose and galactose. These arabinogalactan proteins, or AGP's, seem to be tissue specific and may function in cell signaling. They may be important in embryogenesis and growth and guidance of the pollen tube.

NOTES

5. Lignin

These are polymer of phenolic compounds, especially phenylpropanoids. Lignin is primarily a strengthening agent in the wall. It also resists fungal/pathogen attack.

6. Suberin, Wax, Cutin

A variety of lipids are associated with the wall for strength and waterproofing.

7. Water

The wall is largely hydrated and comprised of between 75-80% water. This is responsible for some of the wall properties. For example, hydrated walls have greater flexibility and extensibility than non-hydrated walls.

Table 2.1 : Composition and Functions of Biomolecules in Cell Wall

POLYMER	COMPOSITION	FUNCTIONS
Cellulose	linear polymer of glucose	fibrils confer tensile strength on all walls
Cross-linking glycans	xyloglucan, glucuronoarabinoxylan, and mannan	cross-link cellulose fibrils into robust network
Pectin	homogalacturonans and rhamnogalacturonans	forms negatively charged, hydrophilic network that gives compressive strength to primary walls; cell–cell adhesion
Lignin	cross-linked coumaryl, coniferyl, and sinapyl alcohols	forms strong waterproof polymer that reinforces secondary cell walls
Proteins and glycoproteins	enzymes, hydroxyproline-rich proteins	responsible for wall turnover and remodeling; help defend against pathogens

2.1.3 Functions of the Cell Wall

The cell wall serves a variety of purposes including:

1. Since protoplasts are invariably rounded, this is good evidence that the wall ultimately determines the shape of plant cells.

NOTES

2. Support and mechanical strength (allows plants to get tall, hold out thin leaves to obtain light)
3. Prevents the cell membrane from bursting in a hypotonic medium (*i.e.,* resists water pressure)
4. Controls the rate and direction of cell growth and regulates cell volume
5. Ultimately responsible for the plant architectural design and controlling plant morphogenesis since the wall dictates that plants develop by cell addition (not cell migration)
6. It has a metabolic role (*i.e.,* some of the proteins in the wall are enzymes for transport, secretion)
7. Cell plays a role of physical barrier to pathogens and water in suberized cells.However, the wall is very porous and allows the free passage of small molecules, including proteins up to 60 KD. The pores are about 4 nm.
8. Carbohydrate storage - the components of the wall can be reused in other metabolic processes (especially in seeds).Thus, in one sense the wall serves as a storage repository for carbohydrates.
9. Signaling- fragments of wall, called oligosaccharins, act as hormones. Oligosaccharins, which can result from normal development or pathogen attack, serve a variety of functions including:(a) stimulate ethylene synthesis; (b) induce phytoalexin (defense chemicals produced in response to a fungal/bacterial infection) synthesis; (c) induce chitinase and other enzymes; (d) increase cytoplasmic calcium levels and (d) cause an "oxidative burst".This burst produces hydrogen peroxide, superoxide and other active oxygen species that attack the pathogen directly or cause increased cross-links in the wall making the wall harder to penetrate.
10. Recognition responses - for example: (a) the wall of roots of legumes is important in the nitrogen-fixing bacteria colonizing the root to form nodules; and (b) pollen-style interactions are mediated by wall chemistry.
11. Economic products - cell walls are important for products such as paper, wood, fiber, energy, shelter, and even roughage in our diet.

2.1.4 Biogenesis of Cell Wall

Cell wall is made during cell division when the cell plate is formed between daughter cell nuclei. The cell plate forms from a series of vesicles produced by the Golg iapparatus. The vesicles migrate along the cytoskeleton and move to

the cell equator. The vesicles coalesce and dump their contents. The membranes of the vesicle become the new cell membrane. The Golgi synthesizes the non-cellulosic polysaccharides. At first, the Golgi vesicles contain mostly pectin polysaccharides that are used to build the middle lamella. As the wall is deposited, other non-cellulosic polysaccharides are made in the Golgi and transported to the growing wall.

NOTES

Cellulose is made at the cell surface. The process is catalyzed by the enzyme cellulose synthase that occurs in a rosette complex in the membrane. Cellulose synthase, which is initially made in by the ribosomes (rough ER) and move from the ER → vesicles →Golgi →vesicle →cell membrane.The enzyme apparently has two catalytic sites that transfer two glucoses at a time (*i.e.*, cellobiose) from UDP-glucose to the growing cellulose chain.Sucrose may supply the glucose that binds to the UDP.Wall protein is presumably incorporated into the wall in a similar fashion.

Exactly how the wall components join together to form the wall once they are in place is not completely understood. Two methods seem likely:

- Self-assembly- This means that the wall components spontaneously aggregate.
- Enzymatic assembly- various enzymatic reactions are designed for wall assembly. For example, one group of enzymes "stitches" xylans together in the wall to form long chains.Oxidases may catalyze additional cross-linking between wall components and pectin methyl esterase may play an important role.

The primary cell walls are built from cellulose micro fibrils, interwoven with a network of pectin polysaccharides. Cellulose gives the primary cell wall tensile strength. Each cellulose molecule consists of a linear chain of at least 500 glucose residues that are covalently linked to one another to form a ribbonlike structure, which is stabilized by hydrogen bonds within the chain. In addition, hydrogen bondsbetween adjacent cellulose molecules cause them to stick together in overlappingparallel arrays, forming bundles of about 40 cellulose chains, all of which have the same polarity. These highly ordered crystalline aggregates, many micrometers long, are called cellulose micro fibrils, and they

NOTES

have a tensile strength comparable to steel. Sets of micro fibrils arearranged in layers, or lamellae, with each micro fibril about 20-40 nm from itsneighbors and connected to them by long cross-linking glycan molecules that are attached by hydrogen bonds to the surface of the micro fibrils. The primary cell wall consists of several such lamellae arranged in plywoodlike network.

The cross-linking glycans are a heterogeneous group of branched polysaccharides that bind tightly to the surface of each cellulose micro fibril and thereby help to cross-link the micro fibrils into a complex network. Their function is analogous to that ofthe fibril-associated collagens discussed earlier. There are many classes of cross-linking glycans, but they all have a long linear backbone composed of one type of sugar (glucose, xylose, or mannose) from which short side chains of other sugars protrude. It is the backbone sugar molecules that form hydrogen bonds with the surface of cellulose micro fibrils, cross-linking them in the process. Both the backbone and the side-chain sugars vary according to the plant species and its stage of development. Coextensive with this network of cellulose micro fibrils and cross-linking glycans is another cross-linked polysaccharide network based on pectin. Pectins are a heterogeneous group of branched polysaccharides that contain many negatively charged galacturonic acid units. Because of their negative charge, pectins are highly hydrated and associated with a cloud of cations, resembling the glycosaminoglycans of animal cells in the large amount of space they occupy. When $Ca^{2+}$is added to a solution of pectin molecules, it cross-links them to produce a semi rigid gel (it is pectin that is added to fruit juice to make jam set). Certain pectins are particularly abundant in the middle lamella, the specialized region that cements together the walls of adjacent cells; here, ca^{2+} cross-links are thought to help hold cell-wall components together. Although covalent bonds also play a part in linking the components, very little is known about their nature. Regulated separation of cells at the middle lamella underlies such processes as the ripening oftomatoes and the abscission (detachment) of leaves in the fall.

2.2 Plasma Membrane

NOTES

The cell membrane (also known as the plasma membrane or cytoplasmic membrane) is a biological membrane that separates the interior of all cells from the outside environment. The cell membrane is selectively permeable to ions and organic molecules and controls the movement of substances in and out of cells.The basic function of the cell membrane is to protect the cell from its surroundings. It consists of the phospholipid bilayer with embedded proteins. Cell membranes are involved in a variety of cellular processes such as cell adhesion, ion conductivity and cell signaling and serve as the attachment surface for several extracellular structures, including the cell wall, glycocalyx, and intracellular cytoskeleton. Cell membranes can be artificially reassembled.

The plasma membrane is a structure of the plant cell that forms a semipermeable, or selective, barrier between the interior of the cell and the external environment; they also function in transport of molecules into and out of the cell. In addition to forming the structural barrier between the internal contents of a cell and the external environment, plasma membranes contain proteins involved in the transport of molecules and other substances into and out of the cell, and they contain proteins and other molecules that are essential for receiving signals from the environment and from plant hormones that direct growth and division.

Like all other cellular membranes, the plasma membrane consists of both lipids and proteins. The fundamental structure of the membrane is the phospholipid bilayer, which forms a stable barrier between two aqueous compartments. In the case of the plasma membrane, these compartments are the inside and the outside of the cell. Proteins embedded within the phospholipid bilayer carry out the specific functions of the plasma membrane, including selective transport of molecules and cell-cell recognition.

2.2.1 Structure of Plasma Membrane

Membranes are lipid–protein assemblies in which the components are held together in a thin sheet by non-covalentbonds. The core of the membrane consists of a sheet of lipids arranged in a bimolecular layer. The lipid bilayer serves primarily as a structural backbone of the membrane and provides the

NOTES

barrier that prevents random movements of water-soluble materials into and out of the cell. The proteins of the membrane, on the other hand, carry out most of the specific functions. Each type of differentiated cell contains a unique complement of membrane proteins, which contributes to the specialized activities of that cell type.

The ratio of lipid to protein in a membrane varies, depending on the type of cellular membrane (plasma vs. endoplasmic reticulum vs. Golgi), the type of organism (bacterium vs. plant vs. animal), and the type of cell (cartilage vs. muscle vs. liver). For example, the inner mitochondrial membrane has a very high ratio of protein/lipid in comparison to the red blood cell plasma membrane, which is high in comparison to the membranes of themyelin sheath that form a multilayered wrapping around a nerve cell. To a large degree these differences can be correlated with the basic functions of these membranes. The inner mitochondrial membrane contains the protein carriers of the electron-transport chain, and relative to other membranes, lipid is diminished. In contrast, the myelin sheath acts primarily as electrical insulation for the nerve cell it encloses, a function that is best carried out by a thick lipid layer of high electrical resistance with a minimal content of protein.

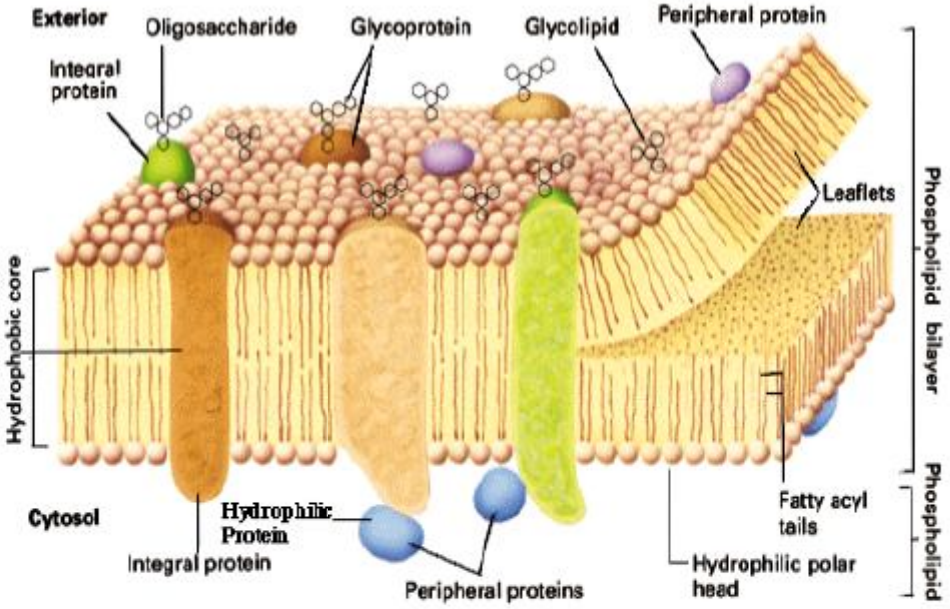


Fig. 2.3 : Structure of Plasma Membrane

1. Membrane Lipids

Membranes contain a wide diversity of lipids, all of which are amphipathic; that is, they contain both hydrophilic and hydrophobic regions. There are three main types of membrane lipids: phosphoglycerides, sphingolipids, and cholesterol.

NOTES

(A) Phosphoglycerides

Most membrane lipids contain a phosphate group, which makes them phospholipids. Because most membrane phospholipids are built on a glycerol backbone, they are called phosphoglycerides. Unlike triglycerides, which have three fatty acids and are not amphipathic, membrane glycerides are diglycerides only two of the hydroxyl groups of the glycerol are esterified to fatty acids; the third is esterified to a hydrophilic phosphate group. Without any additional substitutions beyond the phosphate and the two fatty acyl chains, the molecule is called phosphatidic acid, which is virtually absent in most membranes. Instead, membrane phosphoglycerides have an additional group linked to the phosphate, most commonly either choline (forming phosphatidylcholine, PC), ethanolamine (forming phosphatidylethanolamine, PE), serine (forming phosphatidylserine, PS), or inositol (forming phosphatidylinositol, PI). Each of these groups is small and hydrophilic and, together with the negatively charged phosphate to which it is attached, forms a highly water-soluble domain at one end of the molecule, called the head group.

At physiologic pH, the head groups of PS and PI have an overall negative charge, whereas those of PC and PE are neutral. In contrast, the fatty acyl chains are hydrophobic, unbranched hydrocarbons approximately 16 to 20 carbons in length. A membrane fatty acid may be fully saturated (i.e., lack double bonds), monounsaturated (i.e., possess one double bond), or polyunsaturated (i.e., possessmore than one double bond). Phosphoglycerides often contain one unsaturated and one saturated fatty acyl chain. With fatty acid chains at one end of the molecule and a polar head group at the other end, phosphoglycerides exhibit a distinct amphipathic character.

(B) Sphingolipids

A less abundant class of membrane lipids, are derivatives of sphingosine, an amino alcohol that contains a long hydrocarbon chain. Sphingolipids consist of

NOTES

sphingosine linked to a fatty acid by its amino group. This molecule is a ceramide. The various sphingosine-based lipids have additional groups esterified to the terminal alcohol of the sphingosine moiety. If the substitution is phosphorylcholine, the molecule is sphingomyelin, which is the only phospholipid of the membrane that is not built with a glycerol backbone. If the substitution is a carbohydrate, the molecule is a glycolipid. If the carbohydrate is a simple sugar, the glycolipid is called a cerebroside; if it is an oligosaccharide, the glycolipid is called a ganglioside. Since all sphingolipids have two long, hydrophobic hydrocarbon chains at one end and a hydrophilic region at the other, they are also amphipathic and basically similar in overall structure to the phosphoglycerides. Glycolipids are interesting membrane components.

(C) Cholesterol

Another lipid component of certain membranes is the sterol cholesterol, which incertain animal cells may constitute up to 50 percent of the lipid molecules in the plasma membrane. Cholesterol is absent from the plasma membranes of most plant and all bacterial cells. Cholesterol is smaller than the other lipids of the membrane and less amphipathic. Cholesterol molecules are oriented with their small hydrophilic hydroxyl group toward the membrane surface and the remainder of the molecule embedded in the lipid bilayer. The hydrophobic rings of a cholesterol molecule are flat and rigid, and they interfere with the movements of the fatty acid tails of the phospholipids.

2. Membrane Carbohydrates

The plasma membranes of eukaryotic cells contain carbohydrates which are covalently linked to both lipid and protein components. Depending on the species and cell type, the carbohydrate content of the plasma membrane ranges between 2 and 10 percent by weight. More than 90 percent of the membrane's carbohydrate is covalently linked to proteins to form glycoproteins; the remaining carbohydrate is covalently linked to lipids to form glycolipids, all of the carbohydrate of the plasma membrane faces outward into the extracellular space. The carbohydrate of internal cellular membranes also faces away from the cytosol.

The carbohydrate of glycoproteins is present as short, branched oligosaccharides, typically having fewer than about 15 sugars per chain. In contrast to most high-molecular-weight carbohydrates (such as glycogen, starch, or cellulose), which are polymers of a single sugar, the oligosaccharides attached to membrane proteins and lipids can display considerable variability in composition and structure. Oligosaccharides may be attached to several different amino acids by two major types of linkages. These carbohydrate projections play a role in mediating the interactions of a cell with its environment and sorting of membrane proteins to different cellular compartments.

NOTES

3. Membrane Proteins

Depending on the cell type and the particular organelle within that cell, a membrane may contain hundreds of different proteins. Each membrane protein has a defined orientation relative to the cytoplasm, so that the properties of one surface of a membrane are very different from those of the other surface. This asymmetry is referred to as membrane "sidedness." In the plasma membrane, for example, those parts of membrane proteins that interact with other cells or with extracellular ligands, such as hormones or growth factors, project outward into the extracellular space, whereas those parts of membrane proteins that interact with cytoplasmic molecules, such as G proteins or protein kinases, project into the cytosol.

Membrane proteins can be grouped into three distinct classes distinguished by the intimacy of their relationship to the lipid bilayer. These are-

1. Integral proteins that penetrate the lipid bilayer. Integral proteins are transmembrane proteins; that is, they pass entirely through the lipid bilayer and thus have domains that protrude from both the extracellular and cytoplasmic sides of the membrane. Some integral proteins have only one membrane-spanning segment, whereas others are multi-spanning. Genome sequencing studies suggest that integral membrane proteins constitute about 30 percent of all encoded proteins.
2. Peripheral proteins that are located entirely outside of the lipid bilayer, on the cytoplasmic or extracellular side, yet are associated with the surface of the membrane by non-covalent bonds.

3. Lipid-anchored proteins that are located outside the lipid bilayer, on either the extracellular or cytoplasmic surface, but are covalently linked to a lipid molecule that is situated within the bilayer.

NOTES

2.2.2 Models given for Plasma Membrane

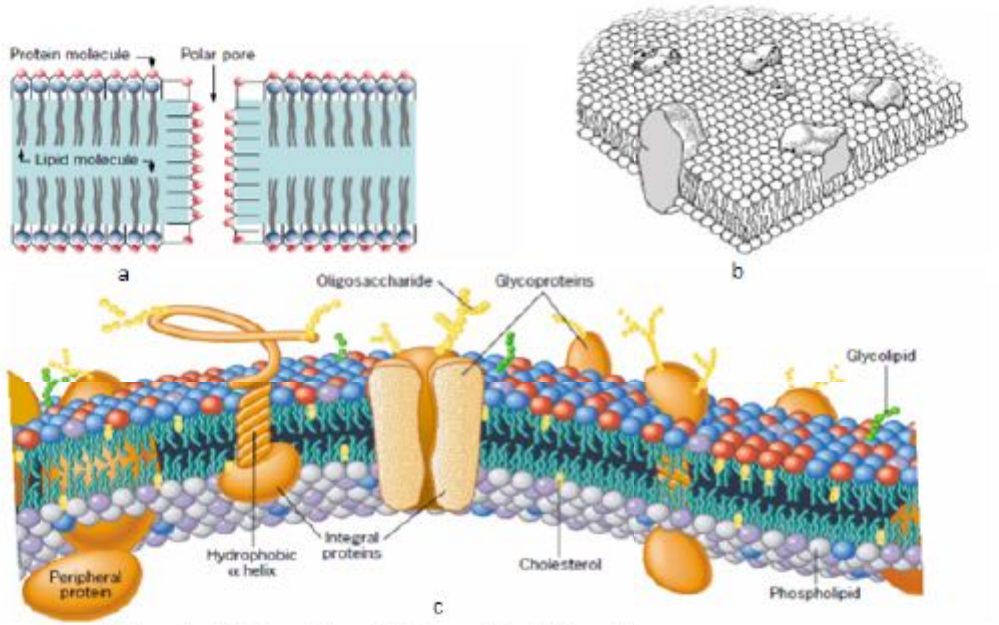
The first insights into the chemical nature of the outer boundary layer of a cell were obtained by Ernst Overton of the University of Zürich during the 1890s. Overton knew that nonpolar solutes dissolved more readily in nonpolar solvents than in polar solvents, and that polar solutes had the opposite solubility. Overton reasoned that a substan ceentering a cell from the medium would first have to dissolve in the outer boundary layer of that cell. To test the permeability of the outer boundary layer, Overton placed plant root hairs into hundreds of different solutions containing a diverse array of solutes. He discovered that the more lipid soluble the solute, the more rapidly it would enter the root hair cells. He concluded that the dissolving power of the outer boundary layer of the cell matched that of fatty oil.

The first proposal that cellular membranes might contain a lipid bilayer was made in 1925 by two Dutch scientists, E. Gorter and F. Grendel. These researchers extracted the lipid from human red blood cells and measured the amount of surface area the lipid would cover when spread over the surface of water. Since mature mammalian red blood cells lack both nuclei and cytoplasmic organelles, the plasma membrane is the only lipid-containing structure and all of the lipids extracted from the cells can be assumed to have resided in the cell'splasma membranes. The ratio of the surface area of water covered by the extracted lipid to the surface area calculated for the red blood cells from which the lipid was extracted varied between 1.8 to 1 and 2.2 to 1. Gorter and Grendel speculated that the actual ratio was 21 and concluded that the plasma membrane contained a bimolecular layer of lipids, that is, a lipid bilayer. They also suggested that the polar groups of each molecular layer (or leaflet) were directed outward toward the aqueous environment. This would be the thermodynamically favored arrangement because the polar head groups of the lipids could interact with surrounding water molecules, just as the hydrophobic fatty acyl chains would be protected from contact with the

aqueous environment. Thus, the polar head groups would face the cytoplasm on one edge and the blood plasma on the other.

In the 1920s and 1930s, cell physiologists obtained evidence that there must be more to the structure of membranes than simply a lipid bilayer. It was found, for example, that lipid solubility was not the sole determining factor as to whether or not a substance could penetrate the plasma membrane. Similarly, the surface tensions of membranes were calculated to be much lower than those of pure lipid structures. This decrease in surface tension could be explained by the presence of protein in the membrane.

NOTES



(*a*) A revised 1954 version of the Davson-Danielli model
(*b*) The fluid-mosaic model as initially proposed by Singer and Nicolson
(*c*) A current representation of the plasma membrane (by Singer and Nicolson)

Fig. 2.4 : Plasma Membrane Models

Davson and Danielli Model

In 1935, Hugh Davson and James Danielli proposed that the plasma membrane was composed of a lipid bilayer that was lined on both its inner and outer surface by a layer of globular proteins. They revised their model in the early 1950s to account for the selective permeability of the membranes they had studied. In the revised version, Davson and Danielli suggested that-

NOTES

- Cell membrane is made of a phospholipid bilayer sandwiched between two layers of globular protein.
- The polar (hydrophilic) heads of phospholipids are oriented towards the protein layers forming a hydrophilic zone.
- The nonpolar (hydrophobic) tails of phospholipids are oriented in between polar heads forming a hydrophobic zone.
- Membrane proteins are not soluble in water, and, like phospholipid, they are amphipathic.

Protein layer not likely because its hydrophobic regions would be in an aqueous environment, and it would also separate the hydrophilic phospholipid heads from water.

Fluid-Mosaic Model

Experiments conducted in the late 1960s led to a new concept of membrane structure, as detailed in the fluidmosaic model proposed in 1972 by S. Jonathan Singer and Garth Nicolson of the University of California, San Diego. In the fluid-mosaic model, which has served as the "central dogma" of membrane biology for three decades, the lipid bilayer remains the core of the membrane, but attention is focused on the physical state of the lipid.

Unlike previous models, the bilayer of a fluid-mosaic membrane is present in a fluid state, and individual lipid molecules can move laterally within the plane of the membrane. The structure and arrangement of membrane proteins in the fluid-mosaic model differ from that of previous models in that they occur as a "mosaic" of discontinuous particles that penetrate the lipid sheet.

2.2.3 Functions of Plasma Membrane

1. **Compartmentalization:** Membranes are continuous, unbroken sheets and, as such, inevitably enclose compartments. The plasma membrane encloses the contents of the entire cell, whereas the nuclear and cytoplasmic membranes enclose diverse intracellular spaces. The various membrane-bound compartments of a cell possess markedly different contents. Membrane compartmentalization allows specialized activities to proceed without external interference and enables cellular activities to be regulated independently of one another.

2. **Scaffold for Biochemical Activities:** Membranes not only enclose compartments but are also a distinct compartment themselves. As long as reactants are present in solution, their relative positions cannot be stabilized and their interactions are dependent on random collisions. Because of their construction, membranes provide the cell with an extensive framework or scaffolding within which components can be ordered for effective interaction.

NOTES

3. **A Selectively Permeable Barrier:** Membranes prevent the unrestricted exchange of molecules from one side to the other. At the same time, membranes provide the means of communication between the compartments they separate. The plasma membrane, which encircles a cell, can be compared to a moat around a castle: both serve as a general barrier, yet both have gated "bridges" that promote the movement of select elements into and out of the enclosed living space.

4. **Transporting Solutes:** The plasma membrane contains the machinery for physically transporting substances from one side of the membrane to another, often from a region of amino acids that are necessary to fuel its metabolism and build its macromolecules. The plasma membrane is also able to transport specific ions, thereby establishing ionic gradients across it. This capability is especially critical for nerve and muscle cells.

5. **Responding to External Signals:** The plasma membrane plays a critical role in the response of a cell to external stimuli, a process known as signal transduction. Membranes possess receptors that combine with specific molecules (or ligands) having a complementary structure.Different types of cells have membranes with different receptors and are, therefore, capable of recognizing and responding to different ligands in their environment. The interaction of a plasma membrane receptor with an external ligand may cause the membrane to generate a signal that stimulates or inhibits internal activities. For example, signals generated at the plasma membrane may tell a cell to manufacture more glycogen, to prepare for cell division, to move toward a higher concentration of a particular compound, to release calcium from internal stores, or possibly to commit suicide.

NOTES

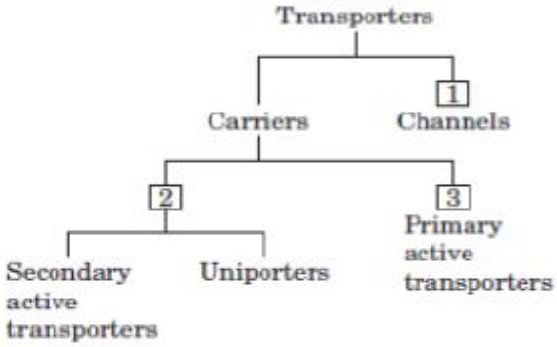
6. **Intercellular Interaction:** Situated at the outer edge of every living cell, the plasma membrane of multicellular organisms mediates the interactions between a cell and its neighbors. The plasma membrane allows cells to recognize and signal one another, to adhere when appropriate, and to exchange materials and information.

7. **Energy Transduction:** Membranes are intimately involved in the processes by which one type of energy is converted to another type (energy transduction). The most fundamental energy transduction occurs during photosynthesis when energy in sunlight is absorbed by membrane-bound pigments, converted into chemical energy, and stored in carbohydrates. Membranes are also involved in the transfer of chemical energy from carbohydrates and fats to ATP. In eukaryotes, the machinery for these energy conversions is contained within membranes of chloroplasts and mitochondria.

2.2.4 Movement across Plasma Membrane

Transporters constitute a significant fraction of all proteins encoded in the genomes of both simple and complex organisms. There are probably a thousand or more different transporters in the human genome. A few hundred transporters from various species have been studied with biochemical, genetic, and electrophysiological tools, but investigators have determined the three-dimensional structures for only a handful of these. Examination of the many transporter genes reveals obvious sequence similarities among subsets of transporters. And as experience has shown, similar amino acid sequences in proteins generally reflect similar three-dimensional structures and, often, similar mechanisms of action. It is reasonable to hope that by determining the structure and mechanism of action of at least one member of each transporter family, we can learn much about the other members of the family—about their structures, substrate specificities, transport rates, and mechanisms of energy coupling.

Transporters are of two general classes-Carriers and channels



NOTES

1. Carriers

Carrier proteins cycle between conformations in which a solute binding site is accessible on one side of the membrane or the other. There may be an intermediate conformation in which a bound substrate is inaccessible to either aqueous phase. With carrier proteins there is never an open channel all the way through the membrane.

The transport rate mediated by carriers is faster than in the absence of a catalyst, but slower than with channels. A carrier transports only one or a few solute molecules per conformational cycle, whereas a single channel opening event may allow flux of many thousands of ions.

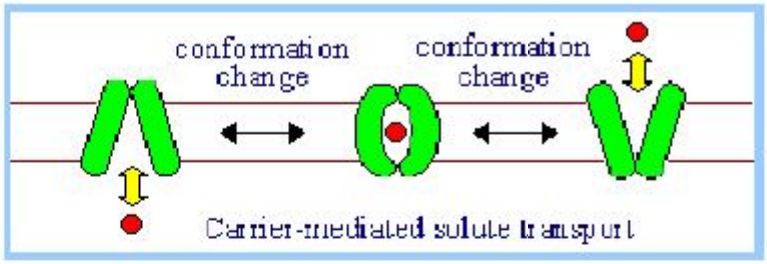


Fig. 2.5 : Carrier Mediated Solute Transport

Valinomycin is a carrier for potassium. Puckering of the ring, stabilized by H-bonds, allows valinomycin to closely surround a single unhydrated K^+ ion. Six oxygen atoms of the ionophore interact with the bound K^+, replacing oxygen atoms of waters of hydration.Valinomycin is highly selective for K^+ relative to Na^+. The smaller Na^+ ion cannot simultaneously interact with all six oxygen atoms within valinomycin. Thus it is energetically less favorable for Na^+ to shed its waters of hydration to form a complex with valinomycin. Whereas the interior of the valinomycin-K^+ complex is polar, the surface of the complex is

hydrophobic. This allows valinomycin to enter the lipid core of the bilayer, to solubilize K^+ within this hydrophobic milieu.

NOTES

Valinomycin is a passive carrier for K^+. It can bind or release $K^+$when it encounters the membrane surface. Valinomycin can catalyze net K^+ transport because it can translocate either in the complexed or uncomplexed state. The direction of net flux depends on the electrochemical K^+ gradient. Proteins that act as carriers are too large to move across the membrane. They are transmembrane proteins with fixed topology. An example is the GLUT1 glucose carrier, in plasma membranes of various cells, including erythrocytes. GLUT1 is a large integral protein, predicted via hydropathy plots to include 12 transmembrane a-helices.

Classes of Carrier Proteins:

- **Uniport-**(facilitated diffusion) carriers mediate transport of a single solute. An example is the GLUT1 glucose carrier. The ionophore valinomycin is also a uniport carrier.
- **Symport-**(cotransport) carriers bind two dissimilar solutes (substrates) and transport them together across a membrane. Transport of the two solutes is obligatorily coupled. A gradient of one substrate, usually an ion, may drive uphill (against the gradient) transport of a co-substrate. It is sometimes referred to as secondary active transport. e.g. the glucose-Na^+ symport found in plasma membranes of some epithelial cells, the bacterial lactose permease, a H^+symport carrier.
- **Antiport-**(exchange diffusion) carriers exchange one solute for another across a membrane. Usually anti-porters exhibit "ping pong" kinetics that is when a substrate binds and is transported across the membrane. Then another substrate binds and is transported in the other direction. Onlyexchange is catalyzed, not net transport, because the carrier protein cannot undergo the conformational transition in the absence of bound substrate.
- **Antiport-**(exchange diffusion) carriers exchange one solute for another across a membrane. Usually anti-porters exhibit "ping pong" kinetics that is when a substrate binds and is transported across the membrane.

Then another substrate binds and is transported in the other direction. Onlyexchange is catalyzed, not net transport, because the carrier protein cannot undergo the conformational transition in the absence of bound substrate.

NOTES

2. Channels

Channels are integral membrane proteins that enclose a central aqueous pore. Most ion channels are highly selective in allowing only one particular type of ion to pass through the pore. As with the passive diffusion of other types of solutes across membranes, the diffusion of ions through a channel is always downhill, that is, from a state of higher energy to a state of lower energy. Most of the ion channels that have been identified can exist in either an open or a closed conformation; such channels are said to begated. The opening and closing of the gates are subject to complex physiologic regulation and can be induced by a variety of factors depending on the particular channel. Two major categories of gated channels are:

- Voltage-gated channels are those channels whose conformational state depends on the difference in ionic charge on the two sides of the membrane.
- Ligand-gated channels are those channels whose conformational state depends on the binding of a specific molecule (the ligand), which is usually not the solute that passes through the channel.

Some ligand-gated channels are opened (or closed) following the binding of a molecule to the outer surface of the channel; others are opened (or closed) following the binding of a ligand to the inner surface of the channel. For example, neurotransmitters, such as acetylcholine, act on the outer surface of certain cation channels, while cyclic nucleotides, such as cAMP, act on the inner surface of certain calcium ion channels.

3. Pumps

These are those transporter molecules which utilize energy (ATP) to drive movement of various molecules against the gradient and this process is called active transport as it utilizes energy. Active transport enzymes couple net solute movement across a membrane to hydrolysis of ATP. An active transport pump

NOTES

may be a uniporter or an antiporter. ATP-dependent ion pumps are grouped into classes, based on transport mechanism as well as genetic and structural homology. e.g.

(A) P-class pumps-These pumps phosphorylate the transported molecule first. P-class ion pumps are a gene family exhibiting sequence homology. They include:

- Na^+- K^+ATPase are present in plasma membranes of most animal cells, and these are antiport ion pumps. They catalyze ATP-dependent transport of Na^+ out of a cell in exchange for K^+ entering the cell.
- H^+- K^+ ATPase are involved in acid secretion in the stomach, and are antiport pumps. They catalyze ATP-dependent transport of H^+ out of the gastric parietal cell (toward the stomach lumen) in exchange for K^+ entering the cell.
- Ca^{2+} ATPases are present in endoplasmic reticulum (ER) and plasma membranes of many cells; catalyze ATP-dependent transport of Ca^{2+} away from the cytosol, into the ER lumen or out of the cell. Some evidence indicates that these pumps are antiporters, transporting protons in the opposite direction. Ca^{2+}-ATPase pumps function to keep cytosolic Ca^{2+} low, allowing Ca^{2+} to serve as a signal.

(B) ABC (ATP binding cassette) transporters- catalyze trans-membrane movements of various organic compounds including amphipathic lipids and drugs.

2.3 Receptors

Cell surface receptors (membrane receptors, transmembrane receptors) are specialized integral membrane proteins that take part in communication between the cell and the outside world. Extracellular signaling molecules (usually hormones, neurotransmitters, cytokines, growth factors or cell recognition molecules) attach to the receptor, triggering changes in the function of the cell. This process is called signal transduction: The binding initiates a chemical change on the intracellular side of the membrane. In this way the receptors play a unique and important role in cellular communications and signal transduction.

2.3.1 Structure of Receptors

Many transmembrane receptors are composed of two or more protein subunits which operate collectively and may dissociate when ligands bind, fall off, or at another stage of their "activation" cycles. They are often classified based on their molecular structure, or because the structure is unknown in any detail for all but a few receptors. The polypeptide chains of the simplest ones cross the lipid bilayer only once, while others cross as many as seven times (for example, the so-called G-protein coupled receptors).

NOTES

There are various kinds, such as glycoprotein and lipoprotein. Almost all known membrane receptors are transmembrane proteins. A certain cell membrane can have several membrane receptors with various amounts on its surface. A certain receptor may also exist at varying concentrations on different membrane surfaces, depending on the membrane and cell function. Since receptors usually form “clusters” on the membrane surface, the distribution of receptors on membrane surface is mostly heterogeneous.Like any integral membrane protein, a transmembrane receptor may be subdivided into three parts or domains.

1. Extracellular domain

The extracellular domain is the part of the receptor that sticks out of the membrane on the outside of the cell or organelle. If the polypeptide chain of the receptor crosses the bilayer several times, the external domain can comprise several "loops" sticking out of the membrane. A receptor's main function is to recognize and respond to a specific ligand, for example, a neurotransmitter or hormone (although certain receptors respond also to changes in transmembrane potential), and in many receptors these ligands bind to the extracellular domain.

2. Trans-membrane domain

In the majority of receptors for which structural evidence exists, transmembrane alpha helices make up most of the transmembrane domain. In certain receptors, such as the nicotinic acetylcholine receptor, the transmembrane domain forms a protein-lined pore through the membrane, or ion channel. Upon activation of an extracellular domain by binding of the appropriate ligand, the pore becomes accessible to ions, which then pass through. In other receptors, the transmembrane domains are presumed to

undergo a conformational change upon binding, which exerts an effect intracellularly. In some receptors, such as members of the 7TM superfamily, the transmembrane domain may contain the ligand binding pocket.

NOTES

3. Intracellular domain

The intracellular (or cytoplasmic) domain of the receptor interacts with the interior of the cell or organelle, relaying the signal. There are two fundamentally different ways for this interaction:

- The intracellular domain communicates via specific protein-protein-interactions with effector proteins, which in turn send the signal along a signal chain to its destination.
- With enzyme-linked receptors, the intracellular domain has enzymatic activity. Often, this is a tyrosine kinase activity. The enzymatic activity can also be located on an enzyme associated with the intracellular domain.

2.3.2 Types of Membrane Receptors

1. Ion Channel Linked Receptors

These are ion-channels (including cation-channels and anion-channels) themselves and constitute a large family of multipass transmembrane proteins. They are involved in rapid signaling events most generally found in electrically excitable cells such as neurons and are also called ligand-gated ion channels. Opening and closing of Ion channels are controlled by neurotransmitters.

In the signal transduction event in a neuron, the neurotransmitter binds with the receptor and alters the conformation of the protein, which opens the ion-channel, allowing extracellular ions to go into the cell. The ion permeability of the plasma membrane is altered, and this will instantaneously convert the extracellular chemical signal into intracellular electric signal, which will alter the excitability of the cell.

Acetylcholine receptor is a kind of cation-channel linked receptor. The protein consists of 4 subunits: α, β, γ, and δ subunits. There are two α subunits, containing one acetylcholine binding site each. This receptor can exist in three different conformations. The unoccupied-closed state is the protein at its original conformation. After two molecules of acetylcholine bind

simultaneously to the binding sites on α subunits, the conformation of the receptor is altered and the gate is opened, allowing for the penetration of many ions and small molecules. However, this occupied-open state can only last for a very short period of time and then the gate is closed again, forming the occupied-closed state. The two molecules of acetylcholine will quickly dissociate from the receptor and the receptor will returns to its unoccupied-closed state and is ready for next transduction cycle again.

NOTES

2. Enzyme-Linked Receptors

These are either enzymes themselves, or are directly associated with the enzymes that they activate. These are usually single-pass transmembrane receptors, with the enzymatic portion of the receptor being intracellular. The majorities of enzyme-linked receptors are protein kinases, or associate with protein kinases.

There are 6 known types of enzyme-linked receptors: Receptor tyrosine kinases; Tyrosine kinase associated receptors; Receptor-like tyrosine phosphatases; Receptor serine/threonine kinases; Receptor Guanylyl cyclases and Histidine kinase associated receptors. Receptor tyrosine kinases is the one kind with the largest population and most widely application. The majority of these molecules are receptors for growth factors and hormones like epidermal growth factor (EGF), platelet derived growth factor (PDGF), fibroblast growth factor (FGF), hepatocyte growth factor (HGF), insulin, nerve growth factor (NGF) etc.

Most of these receptors will dimerize after binding with their ligands in order to activate further signal transductions. For example, after the epidermal growth factor (EGF) receptor binds with its ligand EGF, two receptors dimerize and then undergo phosphorylation of the tyrosine residues in the enzyme portion of each receptor molecule, which will activate the tyrosine protein kinase and catalyze further intracellular reactions.

3. G Protein-Coupled Receptors

These are integral membrane proteins that possess seven membrane-spanning domains or trans-membrane helices. These receptors activate a G protein ligand binding. G-protein is a trimeric protein. The 3 subunits are called α, β and γ.

NOTES

The αsubunit can bind with guanosine diphosphate, GDP. This causes phosphorylation of the GDP to guanosine triphosphate, GTP, and activates the α subunit, which then dissociates from the β and γ subunits. The activated α subunit can further affect intracellular signaling proteins or target functional proteins directly.

G protein-coupled receptors are found only in eukaryotes.The ligands that bind and activate these receptors include light sensitive compounds, odors, pheromones, hormones, and neurotransmitters, and vary in size from small molecules to peptides to large proteins. There are two principal signal transduction pathways involving the G-protein coupled receptors: cAMP signal pathway and Phosphatidylinositol signal pathway.Both activate a G protein ligand binding.

2.4 Plasmodesmata

Plasmodesmata (singular: plasmodesma) are microscopic channels which traverse the cell walls of plant cells and some algal cells, enabling transport and communication between them. Plasmodesmata evolved independently in several lineages,and species that have these structures include members of the Charophyceae, Charales, Coleochaetales and Phaeophyceae (which are all algae), as well as all embryophytes, better known as land plants.Unlike animal cells, every plant cell is surrounded by a polysaccharide cell wall. Neighbouring plant cells are therefore separated by a pair of cell walls and the intervening middle lamella, forming an extracellular domain known as the apoplast. Although cell walls are permeable to small soluble proteins and other solutes, plasmodesmata enable direct, regulated, symplastic intercellular transport of substances between cells. There are two forms of plasmodesmata: primary plasmodesmata, which are formed during cell division, and secondary plasmodesmata, which can form between mature cells.

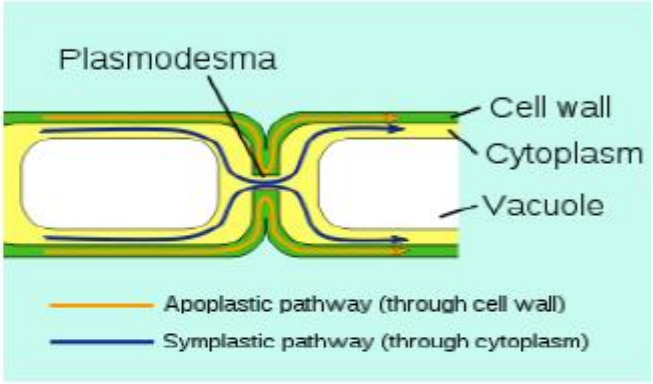


Fig. 2.6 : Plasmodesmata

Plasmodesmata are narrow channels that act as intercellular cytoplasmic bridges to facilitate communication and transport of materials between plant cells. The plasmodesmata serve to connect the symplastic space in the plant and are extremely specialized channels that allow for intercellular movement of water, various nutrients, and other molecules (including signalling molecules). Plasmodesmata are located in narrow areas of cell walls called primary pit fields, and they are so dense in these areas (up to one million per square millimeter) that they make up one percent of the entire area of the cell wall.Similar structures, called gap junctions and membrane nanotubes, interconnect animal cells and stromules form between plastids in plant cells.

2.4.1 Structure of Plasmodesmata

It has been demonstrated that the plasma membrane is continuous between cells, the outer leaflet contiguous with the cell wall and the inner leaflet contiguous with the plasmodesmal pore. Within the plasmodesmal pore, a tightly wound cylinder of membrane termed the desmotubule runs the length of the plasmodesma. Thus, the desmotubule is essentially a tube within a larger tube bordered by the plasma membrane. The structure of the desmotubule and how it relates to the overall structure of plasmodesmata was studied by Tilney, by using plasmolysis, Triton X detergent extraction, and protease digestion. This investigation utilized fern (Onoclea sensibilis) gametophytes by cutting them in half, exposing the cut surfaces to Triton X 100, and then fixing the gametophytes. This detergent solubilized the plasma membrane component of the plasmodesmata of the gametophytes, but the desmotubule was not affected by the detergent. However, when the cut gametophyte surfaces were exposed to

NOTES

NOTES

papain, the desmotubule is destroyed but the plasma membrane remained intact (yet swollen). Finally, the gametophytes were plasmolyzed, and it was found that plasmodesmata remained intact as long as the desmotubule stayed in its normal, fixed position as the cells detached from the cell walls. Thus, Tilney suggested that the desmotubule provides a rigid stability to plasmodesmata and confers a fixed diameter and pore size to the plasmodesmal canal, much like a cytoskeletal structure. However, one should be cautious in mistaking the desmotubule as a completely rigid structure, since the desmotubule is linked to the endoplasmic reticulum in each of the adjacent cells, forming a dynamic endomembrane continuum in the symplastic space.

The space between the plasmalemma and the desmotubule is the cytoplasmic sleeve or cytoplasmic annulus, and transport through plasmodesmata has been proposed to occur either through the lipid portions of the desmotubule or this cytoplasmic sleeve (or both).

A typical plant cell may have between 103 and 105 plasmodesmata connecting it with adjacent cellsequating to between 1 and 10 per μm^2.Plasmodesmata are approximately 50-60 nm in diameter at the midpoint and are constructed of three main layers, the plasma membrane, the cytoplasmic sleeve, and the desmotubule.They can transverse cell walls that are up to 90 nm thick.

1. Plasmodesmatal Plasma Membrane

The plasma membrane portion of the plasmodesma is a continuous extension of the cell membrane or plasmalemma. It is similar in structure to the cellular phospholipid bilayers.

2. Cytoplasmic Sleeve

The cytoplasmic sleeve is a fluid-filled space enclosed by the plasmalemma and a continuous extension of the cytosol. Trafficking of molecules and ions through plasmodesmata occurs through this passage. Smaller molecules (e.g. sugars and amino acids) and ions can easily pass through plasmodesmata by diffusion without the need for additional chemical energy. Larger molecules, including proteins (for example Green fluorescent protein) and RNA, can also pass through the cytoplasmic sleeve diffusively.Plasmodesmal transport of some larger molecules is facilitated by unknown mechanisms. One main

mechanism of regulation of plasmodesmal transport is the accumulation of the polysaccharide callose accumulates around the neck region of plasmodesmata to form a collar, reducing their diameter and thereby controlling permeability to substances in the cytoplasm.

NOTES

3. Desmotubule

The desmotubule is a tube of appressed endoplasmic reticulum that runs between two adjacent cells. Some molecules are known to be transported through this channel,but it is not thought to be the main route for plasmodesmatal transport. Around the desmotubule and the plasma membrane areas of an electron dense material have been seen, often joined together by spoke-like structures that seem to split the plasmodesma into smaller channels.These structures may be composed of myosin and actin, which are part of the cell's cytoskeleton.

2.4.2 Role of Plasmodesmata in Movement of Molecules

Classical studies on transport through plasmodesmata have utilized microinjection of small, fluorescently-labeled probes to examine the passive transport mechanisms of plasmodesmata. These probes were first used to describe the size-exclusion limits, or the maximum size of a molecule that can pass through plasmodesmata passively, of plasmodesmata in various plant species. Looking at the cytoplasmic sleeve region and the spokes that constrict the neck region plasmodesmata, it was determined that the average width of channels allowing molecules to passively travel through a plasmodesmal annulus is approximately 3 nm, and thus the average size exclusion limit is molecules weighing approximately 800-1000 Da.

1. Variations on Passive Transport

However, it has recently been shown that size exclusion limits vary between species and even cell types during passive transport. In tobacco mesophyll cells, the size exclusion limit of fluorescently labeled dextrans is only 1 kDa, whereas in trichome cells, the size-exclusion limit is as large as 7 kDa. Another study by Wang and Fisher employed normally apoplastic probes such as Lucifer Yellow and incubated sections of crease tissue from developing wheat grains in solutions of these probes. The dye moved into the symplastic space

NOTES

through plasmodesmata, and it was determined that in all cells except the pericarp, the diameter of the plasmodesmata in this tissue was 6.2 nm (twice the normal exclusion limit). The microinjection technique used to exogenously add small probe molecules may be triggering a wound response in the plant cell and a consequent partial closure of plasmodesmata, perhaps by callose plugging up the transport pathways. The varying size exclusion limits may also be due to the specific spoke molecules allowing passive transport through the neck region of plasmodesmata; different species of plants may have different neck region compositions, accounting for the discrepancies in size exclusion limits for passive transport.

2. Active Transport of Macromolecules

Other macromolecules destined for intercellular transport are shuttled through plasmodesmata via active transport mechanisms. Plasmodesmata can alter their dimensions such that they expand/dilate outward into an electron-lucent sleeve surrounding the normal-sized plasmodesmata. This expansion would allow for the transport of larger molecules through the cytoplasmic sleeve. The mechanism or trigger of this plasmodesmal dilation is virtually unknown, but it has been recently postulated that proteins in the neck region may be implicated in this phenomenon.

Another major recent finding concerning active transport through plasmodesmata involves a possible association of cytoskeletal elements with both trafficking/targeting molecules to the plasmodesmata and assisting in the energetics of active transport. The endoplasmic reticulum is closely associated with microtubules, and viral movement proteins track through the plant cell along these microtubules. Actin and tubulin bind to viral movement protein, which transports itself through the plant via plasmodesmata. Thus, cytoskeletal elements are helping to guide various macromolecules to plasmodesmata, and also actin has been thought to help loosen the sphincter elements of the neck region to allow for the transport of these larger molecules. In addition, myosin may also act as a cytoskeletal motor to provide the energetics for active plasmodesmal transport. Myosin has been found in several species of plants and is an ATP-dependent protein that generates directional movement (as in flagellar motion). High ATPase activity has been demonstrated in

plasmodesmata, so myosin is a prime candidate for aiding active transport and may be the spokes, sometimes seen in electron micrographs, connecting the desmotubule to the cytoplasmic annulus throughout the plasmodesmata.

2.4.3 Comparison of Plasmodesmata with Gap Junctions

NOTES

A multicellular organism can be composed of hundreds to trillions of cells. Each of these cells contains the necessary organelles and other components to carry out all its vital functions. However, cells do not function in isolation. Cells must be able to bind together. It is essential for the organism that its cells be able to communicate with, and respond to, each other, too. Gap junctions and plasmodesmata help to fulfill this goal in animal and plant cells, respectively.

In animal cells, several different types of cell connections, or junctions, exist. These junctions have different and sometimes opposing roles. Some junctions serve to seal cells together and prevent the exchange of material between them. Other junctions help cells attach to each other or to the extracellular matrix. Yet other junctions actually allow for the exchange of chemical and electrical signals directly between cells. Gap junctions are such communicating junctions. The plasmodesmata of plant cells serve much the same role.

Gap Junctions-A gap junction is a protein channel that forms between the plasma membranes of two cells. One subunit of the protein, called a connexon, penetrates the plasma membrane of each cell, and the connexons then join together. A continuous channel then exists between the cytoplasm of one cell and the cytoplasm of the other. This protein channel is large enough for ions and small molecules to pass directly from one cell to the other.

Plasmodesmata-Plasmodesmata are structurally different from gap junctions but serve essentially the same function. Each plasmodesma is a channel between two plant cells that passes through the cell wall. Unlike in gap junctions, in a plasmodesma the plasma membrane of one cell is continuous with the plasma membrane of the other, forming a narrow passageway between them. The cytoplasm of both cells is thus connected and small molecules can pass from one cell to the other. A structure called a desmotubule is also usually found in a plasmodesma - it connects, and it continuous with, the smooth endoplasmic reticula of both cells.

NOTES

Although plasmodesmata and gap junctions resemble each other functionally, their structures differ in two significant ways. The plasma membranes of the adjacent plant cells merge to form a continuous channel, the annulus, at each plasmodesma, whereas the membranes of cells at a gap junction are not continuous with each other. In addition, an extension of the endoplasmic reticulum called a desmotubule passes through the annulus, which connects the cytosols of adjacent plant cells. Many types of molecules spread from cell to cell through plasmodesmata, including proteins, nucleic acids, metabolic products, and plant viruses. Soluble molecules pass through the cytosolic annulus, whereas membrane bound molecules can pass from cell to cell through the desmotubule.

2.5 Summary

Each plant cell possesses an outermost covering which is extracellular matrix and is known as cell wall. Cell wall is a rigid structure which plays a role in support and movement of some molecules across it.

Plasma membrane is lipid bilayer which separates cell's internal environment from external environment. Plasma membrane also facilitates selective movement of various molecules across it.

Receptors play a role of a connecting link between a ligand and its response.

Plasmodesmata are channels which enables movement of various molecules between plant cells.

2.6 Glossary

- **Active transport :** The transport of molecules across a membrane and against their natural flow; mediated by carrier proteins and requiring outside energy.
- **Carrier protein :** A protein responsible for mediating the active transport of molecules from one side of the lipid bilayer to the other. Transport is carried out by a conformational change that occurs within the protein that forms an opening for specific molecules to pass through.
- **Channel protein :** A protein responsible for mediating the passive transport of molecules from one side of the lipid bilayer to the other.

Transport is carried out by its membrane-spanning hydrophilic structure which, when open, allows molecules to pass through.

- **Diffusion :** The transport process in which molecules naturally travel from an area of higher concentration to an area of lower concentration.
- **Glycocalyx :** A layer of carbohydrates that coats the exterior of higher-ordered cells.Functions in protecting the cell from damage.
- **Hydrophilic :** A polar molecule that selectively associates itself with water through hydrogen bonds.
- **Hydrophobic :** A nonpolar molecule that does not readily associate with water through hydrogen bonds.
- **Integral protein :** A membrane protein that cannot be easily removed from the lipid bilayer and is integrated in lipid bilayer.
- **Ionophore :** A class of membrane transport proteins. Small, hydrophobic molecules that increase membrane permeability to certain ions.
- **Lipid bilayer :** A thin double layer of phospholipid molecules.Provides the structure of a cell membrane. Structure is a result of hydrophobic and hydrophilic forces.
- **Lipid-bound protein :** Membrane proteins that are located entirely within the lipid bilayer, having no part touching either the inside or outside of the cell.
- **Multi-pass protein :** Transmembrane proteins that cross the lipid bilayer more than one time.
- **Passive transport :** Transport mediated by channel proteins. The movement of molecules across a membrane according to the natural flow.
- **Peripheral protein :** A membrane protein that can be easily removed from the lipid bilayer and present on periphery of plasma membrane.

NOTES

- **Single-pass protein :** A transmembrane protein that only crosses the lipid bilayer one time.
- **Transmembrane protein :** A membrane protein that spans the lipid bilayer having portions in contact with both the inside and outside of the cell. Area within the lipid bilayer forms an alpha-helix.

NOTES

2.7 Self-Learning Exercise

Section -A (Very Short Answer Type Questions)

1. What is the main function of the lipid bilayer?
2. Why is the structure called a lipid bilayer?
3. What are the names of the two main classes of membrane proteins?
4. What is the name of the natural process by which molecules flow from an area of higher concentration to one of lower concentration?
5. What is plant cell wall?
6. What are the major regions of cell wall?
7. What is antiport?
8. What is plasmodesma?

Section -B (Short Answer Type Questions)

1. Write down a short note on different cell wall components.
2. Explain some major functions of cell wall.
3. Describe different types of lipids present in plasma membrane.
4. Differentiate between carriers and channels.
5. Write a short note on comparison of plasmodesmata with gap junctions.

Section -C (Long Answer Type Questions)

1. Explain the structure of cell wall in detail with well labeled diagram.
2. Explain the process of formation of cell wall in detail.
3. Describe the structure of plasma membrane with well labeled diagram.

4. Explain the structure of receptors and their different types.

5. Describe the structure of plasmodesmata along with its role in movement of molecules.

Answer Key of Section-A

NOTES

1. The lipid bilayer acts as a barrier between the inside and outside of the cell. It is highly impermeable and does not allow most molecules to freely pass through it into or out of the cell.

2. It is called a lipid bilayer because it is composed of two layers of fat, or lipid, molecules.

3. The two main classes of membrane proteins are integral proteins and peripheral proteins.

4. Diffusion.

5. Each cell in a plant is completely enclosed by, an elaborate extracellular matrix called the plant cell wall.

6. Middle lamella, primary cell wall, secondary cell wall.

7. It is carrier mediated exchange of one solute against another across plasma membrane.

8. Plasmodesma is microscopic channels which traverse the cell walls of plant cells enabling transport and communication between them.

2.8 References

- Alberts, Bruce et al. Molecular Biology of the Cell. Third edition. Garland Publishing, Inc. New York, 1994.
- Brett, C. & K. Waldron. 1996. Physiology and Biochemistry of Plant Cell Walls. 2nd edn. Chapman & Hall, NY.
- Gerald Karp. Cell and Molecular Biology. Fifth edition. John Wiley & Sons, Inc. 2008.
- Grifiths, JF et al. Modern Genetic Analysis. Freeman Publishing, Inc. New York, 1999.

NOTES

- Keegstra et al (1973) The structure of plant cell walls. III. A model of the walls of suspension-cultured sycamore cells based on the interconnections of the macromolecular components. *Plant Physiol.*, 51, 188-196.
- Singer SJ, Nicolson GL (Feb 1972). "The fluid mosaic model of the structure of cell membranes". *Science* 175 (4023): 720 31. doi:10.1126/science.175.4023.720.PMID 4333397.
- Voet, Donald and Judith Voet. Biochemistry. John Wiley & Sons, Inc, New York, 1995.
- http://www.ncbi.nlm.nih.gov., adapted from Plant Cell Biology - William V. Dashek

Unit - 3

Chloroplast, Mitochondria and Plant Vacoule

NOTES

Structure of the Unit

3.0 Objectives

After studying this unit you will be able to understand:

- Structure of Chloroplast, genome organization, gene structure
- RNA editing, Nucleochloroplastic Interactions
- Structure of Mitochondria, Genome Organization, Biogenesis
- Structure of Plant Vacuole, Tonoplant Membrance, ATPase, Transporters, as storage organelle

NOTES

3.1 Introduction

Chloroplasts are organelles, specialized subunits, in plant and algal cells. Their main role is to conduct photosynthesis, where the photosynthetic pigment chlorophyll captures the energy from sunlight, and stores it in the energy storage molecules ATP and NADPH while freeing oxygen from water.

Mitochondria are unusual organelles. They act as the power plants of the cell, are surrounded by two membranes, and have their own genome. They also divide independently of the cell in which they reside, meaning mitochondrial replication is not coupled to cell division. Some of these features are holdovers from the ancient ancestors of mitochondria, which were likely free-living prokaryotes.

A vacuole is a membrane-bound organelle which is present inall plant and fungal cells and some protist, animal and bacterial cells. Vacuoles are essentially enclosed compartments which are filled with water containing inorganic and organic molecules including enzymes in solution, though in certain cases they may contain solids which have been engulfed. Vacuoles are formed by the fusion of multiple membrane vesicles and are effectively just larger forms of these. The organelle has no basic shape or size, its structure varies according to the needs of the cell. The function and importance of vacuoles varies greatly according to the type of cell in which they are present, having much greater prominence in the cells of plants, fungi and certain protists than those of animals and bacteria.

3.2 Chloroplast: Structure

Chloroplasts are organelles, specialized subunits, in plant and algal cells. Their main role is to conduct photosynthesis, where the photosynthetic pigment chlorophyll captures the energy from sunlight, and stores it in the energy storage molecules ATP and NADPH while freeing oxygen from water. They then use the ATP and NADPH to make organic molecules from carbon dioxide in a process known as the Calvin cycle. Chloroplasts carry out a number of other functions, including fatty acid synthesis, much amino acid synthesis, and the immune response in plants.

A chloroplast is one of three types of plastid, characterized by its high concentration of chlorophyll. (The other two types, the leucoplast and the chromoplast, contain little chlorophyll and do not carry out photosynthesis.) Chloroplasts are highly dynamic—they circulate and are moved around within plant cells, and occasionally pinch in two to reproduce. Their behavior is strongly influenced by environmental factors like light color and intensity. Chloroplasts, like mitochondria, contain their own DNA, which is thought to be inherited from their ancestor—a photosynthetic cyanobacterium that was engulfed by an early eukaryotic cell. Chloroplasts cannot be made by the plant cell, and must be inherited by each daughter cell during cell division.

NOTES

With one exception (a member of the genus Paulinella), all chloroplasts can probably be traced back to a single endosymbiotic event (the cyanobacterium being engulfed by the eukaryote). Despite this, chloroplasts can be found in an extremely wide set of organisms, some not even directly related to each other—a consequence of many and even tertiary endosymbiotic events.

3.3 Genome Organization, Gene Structure

Chloroplast genomes are relatively large, usually 140kb in higher plants&<200kb in lower eukaryotes. This is comparable to size of a large bacteriophage,e.g. T4 at ~65kb.There are multiple copies of the genome per organelle, typically 20-40 in a higher plant multiple copies of the organelles per cell, again 20-40.

The chloroplast genome codes for all the rRNA& tRNA species needed for protein synthesis. The ribosomes include two small rRNAs in addition to the major species. The tRNA set resembles that of Mitochondria in including fewer spp. than would suffice in the cytoplasm. The chloroplast genome codes for ~50 proteins, including RNA polymerase &some ribosomal proteins. Again the rule is that organelle genes are transcribed& translated the apparatus of the organelle.

The chloroplast genome of the higher plants varies in length, but displays a characteristic landmark. It has a lengthing sequence, 10-24kb depending on the plants, that is present in two identical copies as an inverted repeat(Gene that are

coded within the inverted repeats are present in two copies per genome &include the rRNA genes.

NOTES

Chloroplasts contain a genome that is a relic of the endosymbiont that gave rise to the organelle. These genomes typically contain 100–200 genes and encode proteins important for photosynthesis and other chloroplast functions, although dinoflagellate algae have a greatly reduced and fragmented chloroplast genome. Many nonphotosynthetic taxa retain a remnant chloroplast genome. In some organisms the chloroplast genome may be retained to allow redox-mediated control of gene expression, whereas in others it may continue to exist because transfer of essential genes to the nucleus is no longer possible. Multiple ribonucleic acid (RNA) polymerases are involved in chloroplast transcription in land plants, and post-transcriptional processing involving nuclear-encoded RNA-binding proteins also plays an important part in the expression of the chloroplast genome. Gene expression may be regulated at a number of levels, and redox poise in the organelle may be particularly important for this. Chloroplasts originated in the ancestor of plants and red and green algae by endosymbiotic acquisition of a cyanobacterium, and then spread to many other eukaryotic lineages. There are other examples of stable acquisition of a cyanobacterium by a nonphotosynthetic host. Chloroplast genomes are typically 100–200 kbp in size, and include a set of genes for proteins essential to photosynthesis. Other genes present in the ancestral symbiont have been lost or relocated to the nucleus. Many nonphotosynthetic organisms retain remnant chloroplasts, with genomes, reflecting the fact that photosynthesis is not the only biochemical process that takes place in the chloroplast. In many organisms, gene transfer from chloroplast to nucleus can still take place, at an unexpectedly high frequency.Establishment of the chloroplast has resulted in the development of nuclear-encoded RNA-binding protein families and, in land plants, additional RNA polymerases that play a central role in chloroplast gene expression.

Post-transcriptional RNA processing plays an important role in chloroplast gene expression. In photosynthetic organisms, chloroplast gene expression is closely linked to the organelle's redox poise. The chloroplast can also influence the expression of nuclear genes.

3.4 RNA Editing

NOTES

The Central Dogma of Molecular Genetics states that the information that is found in DNA is used to produce mRNA molecules that are instrumental in the production of proteins. Therefore, the information flows directly from DNA to protein, via the RNA intermediate molecule. Recently it has been discovered that the information that is contained in the DNA is not always found in the RNA products used to make proteins.

It has now been demonstrated that mitochondria and chloroplast contain the biochemical machinery to alter the sequence of the final transcription product. This process is called RNA editing. This process was identified in the following manner. Sequence analysis of a number of cytochrome c oxidase subunit II genes from non-plant species revealed that a tryptophan residue was invariant at several locations in the final protein product. But sequence analysis of this gene in several plant species revealed arginine at those positions. This amino acid change would cause a radical alteration in protein structure because an acidic amino acid would replace a neutral, hydrophobic amino acid.

Position 1

Plants: GLU ILE LEU ARG THR ILE PHE PRO

Bovine: GLU THR ILE TRP THR ILE PHE PRO

Position 2

Plants: GLN TRP TYR ARG THR TYR

Bovine: GLN TRP TYR TRP SER TYR

Since a single base pair change in the codons for the two amino acids could generate this change (CGG for UGG), it was suggested that CGG encoded for tryptophan and not arginine in plant mitochondria. (This is the only change in codon usage that has been suggested for plants and has been postulated for several other genes as well.) But this change in codon usage was not universal, that is some CGG codons actually specified arginine in the final protein product. Furthermore, no amino acyl tRNA that recognized CGG was found to be charged with tryptophan, a prerequisite if this codon specification was actually real.

NOTES

The solution to this dilemma was found by sequencing the mRNA products for cytochrome oxidase subunit II genes. It was found that in the mRNA the cytosine residue had been changed (edited) to uridine at the sequence location where the invariant tryptophan residue is found. This changed the codon at that location to UGG which is recognized by a tRNA that carries the amino acid tryptophan. An analysis of three other plant mitochondrial genes where the same altered codon usage was predicted suggested that mRNA editing was also occurring at the codon and that a cytosine residue was edited to uridine. This editing process has also been detected in protozoa and it remains to be determined if RNA editing is a widespread function in mitochondria. A final point that this editing function highlights is that the sequence that is found in the DNA is not entirely and faithfully represented in the final protein product.

Specific Features of RNA Editing

1. Editing can occur in both mitochondria and chloroplasts.
2. To date more than 300 different editing events have been detected in plant mitochondria.
3. The vast majority of the events involve a C to U tansition. A few cases of U to C transitions have been reported. This suggests that the editing mach nery can also carry out the reverse modification.
4. RNA editing can modify from 0.8% to 5.8% of the nucleotides of a specific transcript.
5. The RNA editing events appear to occur at random in the transcript.
6. Both 5' and 3' non-coding regions of mRNAs have also been shown to be edited.
7. Structural RNAs such as tRNAs and rRNAs do appear to be affeceted.
8. Editing can convert a tryptophan codon to a arginine codon (CGG to UGG).
9. Start AUG codons can be created from ACG threonine codons
10. Stop codons can be created by editing CAG, CAA and CGA codons.
11. The most frequent amino acid substitions derived from RNA editing are Pro to Leu, Ser to Leu and Ser to Phe.

12. Plant mitochondria do not use the universal genetic code.

The primary benefit of RNA editing could be evolutionary conservation of protein structure. For example, bound copper is required for the funciton of cytochrome c oxidase subunit II (coxII). After editing, all amino acids #228 are converted to cysteine, an amino acid required for copper to bind. In all species except for plants, the coxII gene encodes for methionine at codon #235. In plants, this methionine is generated by RNA editing. These events suggest that this protein is under very strong structural and functional constraints.

NOTES

3.5 Nucleo-Chloroplastic Interactions

In common with all other eukaryotes, nucleocytoplasmic interaction in plants takes the form of a traffic of signals across the nuclear envelope. The signalling molecules involved range from nuclear gene transcripts to small polypeptides often possessing specific amino acid targeting sequences. Whilst the nuclear pore is clearly the main route taken by these signals, there is some evidence that certain classes of molecule can pass through the two elements of the nuclear envelope. The response of the nucleus to exogeneous signals depends on a number of factors, amongst which is its position in the cell cycle, for it is only during the GI phase that the nuclei are competent to receive signals, and the stage in the life cycle of the plant.

The cytoskeleton and the DNA containing organelles are the cytoplasmic components most involved in interaction with the nucleus. New evidence suggests that, in addition to being involved in the division of chromosomes, the cytoskeleton may also play a part in the pairing of chromatids before meiosis, and in the establishment of polarity of the new nucleus following cytokinesis. The relationship between the nucleus and cortical families of microtubules is not clear, but they must be linked in some way to enable the pre-phophase band to so accurately presage the polarity of the subsequent nuclear division. Organelles are under hybrid genetical control, with the majority of their structure and function being regulated by nuclear genes. Thus, most nucleo-organellar interaction consists the synthesis of organellar proteins by translation of nuclear message on cytoplasmic ribosomes. Exciting discoveries are presently being made in the way these organellar proteins are targeted to their final destinations, and in the mechanism by which they become inserted into

NOTES

the correct domain within the organelle. No specific families of signalling molecules have been discovered to be involved in the regulation of organellar activities, but physical contact, particularly between mitochondria and the nucleus, occurs at particular developmental stages. The composition of the three genomes within the plant cell is far from static, and characteristic sequences can be found to be shared both between chloroplast and mitochondria, and between these organelles and the nucleus. This apparent genetic fluidity presumably provides the basis for the transfer of organellar genes to the nucleus that has taken place over the course of evolution. At certain stages during the life cycle of a plant, emission of nuclear signals decreases to a low level and, at these points; cells have been shown to be particularly developmentally labile. This lack of developmental commitment can result in striking switches in pathways of differentiation, and, is probably responsible for the development of young microspores into haploid embryoids rather than pollen grains.

3.6 Mitochondria: Structure

Mitochondria are well-defined cytoplasmic organelles of the cell which take part in a variety of cellular metabolic functions. Survival of the cells requires energy to perform different functions. The mitochondria are important as the fact that these organelles supply all the necessary biological energy of the cell, and they obtain this energy by oxidizing the substrates of the Krebs cycle. Energy of the cell is got from the enzymatic oxidation of chemical compounds in the mitochondria. Hence, the mitochondria re referred to as the 'power houses' of the cell. Almost all the eukaryotic cell has mitochondria, though they are lost in the later stages of development of cell like in the red blood cells or in elements of phloem sieve tube.

In 1890, mitochondrion was first described by Richard Altmann and he called them as bioblasts. Benda in the year 1897 coined the term mitochondrion. In the 1920s, a biochemist Warburg found that oxidative reactions takes place in most tissues in small parts of the cell.

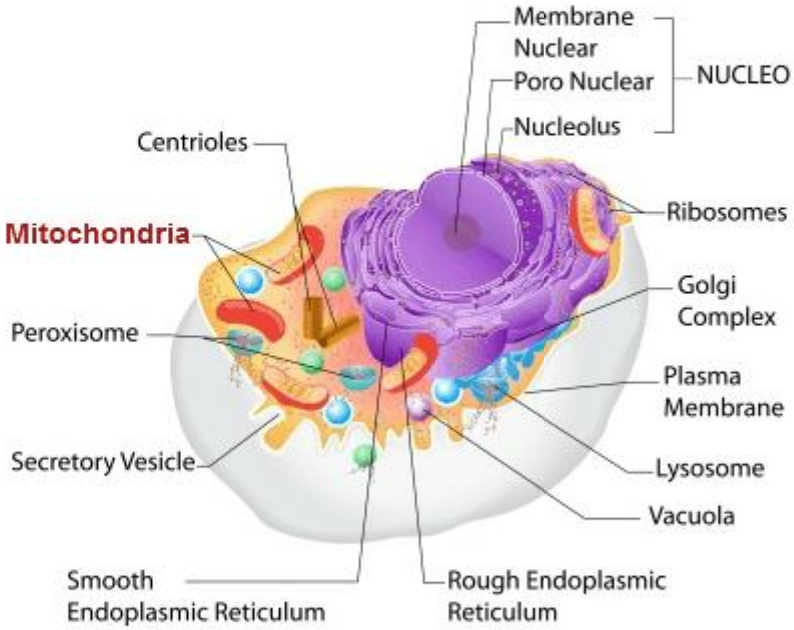


Fig. 3.1 : Mitochondria in Cell

Mitochondrion is a membrane bound cellular structure and is found in most of the eukaryotic cells. The mitochondrion ranges from 0.5 to 1.0 micrometer in diameter. The mitochondria are sometimes described as power plants of the cells. These organelles generate most of the energy of the cell in the form of adenosine triphosphate (ATP) and it is used a source of chemical energy. The mitochondria also involved in other cellular activities like signaling, cellular differentiation, cell senescence and also control of cell cycle and cell growth. Mitochondria also affect human health, like mitochondrial disorder and cardiac dysfunction and they also play important role in the aging process. The term 'mitochondrion' is derived from a Greek word 'mitos' which means 'thread' and 'chondrion' which means 'granule'.

Structure

Mitochondria are rod shaped structure found in both animal and plant cells. It is a double membrane bound organelle. It has the outer membrane and the inner membrane. The membranes are made up of phospholipids and proteins.

NOTES

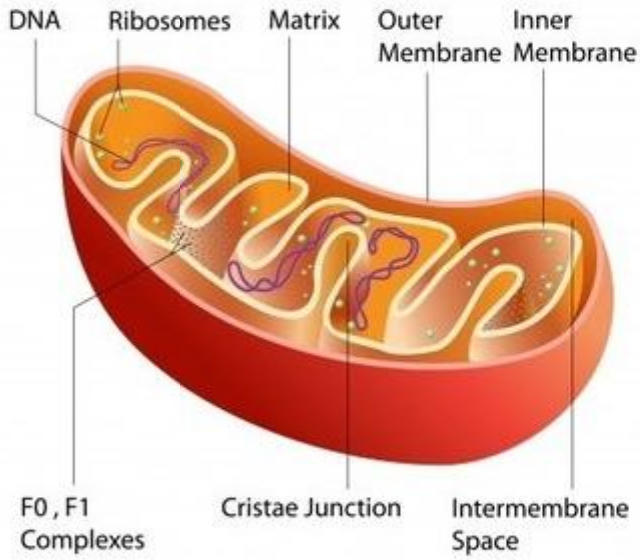


Fig.3.2 : Mitocondria

The components of mitochondria are as follows:

Outer membrane

It is smooth and is composed of equal amounts of phospholipids and proteins. It has a large number of special proteins known as the porins.The porins are integral membrane proteins and they allow the movement of molecules that are of 5000 daltons or less in weight to pass through it. The outer membrane is freely permeable to nutrient molecules,ions, energy molecules like the ATP and ADP molecules.

Inner membrane

The inner membrane of mitochondria is more complex in structure. It is folded into a number of folds many times and is known as the cristae. This folding help to increases the surface ares inside the organelle. The cristae and the proteins of the inner membrane aid in the production of ATP molecules.

Various chemical reactions take place in the inner membrane of the mitochondria.

Unlike the outer membrane, the inner membrane is strictly permeable, it is permeable only to oxygen, ATP and it also helps in regulating transfer of metabolites across the membrane.

Intermembrane space

It is the space between the outer and inner membrane of the mitochondria, it has the same composition as that of the cell's cytoplasm. There is a difference in the protein content in the intermembrane space.

NOTES

Matrix

The matrix of the mitochondria is a complex mixture of proteins and enzymes. These enzymes are important for the synthesis of ATP molecules, mitochondrial ribosomes, tRNAs and mitochondrial DNA.

Function of Mitochondria

Functions of mitochondria depend on the cell type in which they are present.

The most important function of the mitochondria is to produce energy. The simpler molecules of nutrition are sent to the mitochondria to be processed and to produce charged molecules. These charged molecules combine with oxygen and produce ATP molecules. This process is known as oxidative phosphorylation.

Mitochondria help the cells to maintain proper concentration of calcium ions within the compartments of the cell.

The mitochondria also help in building certain parts of blood and hormones like testosterone and estrogen.

The liver cells mitochondria have enzymes that detoxify ammonia.

The mitochondria also play important role in the process of apoptosis or programmed cell death. Abnormal death of cells due to the dysfunction of mitochondria can affect the function of organ.

Mitochondrial DNA

Mitochondrial DNA or mt-DNA or mDNA is the DNA in the mitochondria, rest of the DNA present in the eukaryotic cells is in the nucleus, in plants DNA is also found in chloroplasts.

The mitochondria have a small amount of DNA of their own. Human mitochondrial DNA spans about 16,500 DNA base pairs, it represents a small fraction of the total DNA in cells. The mt-DNA contains 37 genes. All these genes are essential for normal function of the mitochondria.

NOTES

These DNA help the mitochondria divide independently from the cell. mt-DNA is maternally inherited. The fact that mt -DNA is maternally inherited enables to trace the maternal lineage far back in time.

The mt-DNA in most multicellualr organisms is circular, covalently closed, double-stranded DNA. mt-DNA is susceptible to free oxygen radicals. Mutations in the mitochondrial DNA leads to a number of illnesses like exercise intolerance.

Mitochondrial Disease

Disease of mitochondria results due to the failure of mitochondria. Dysfunction in the mitochondria fails to produce energy that is needed for the sustainment of life and growth of an organism. Injury in the cell or even cell death results in the production of less energy. If the process happens throughout the body, the whole system begins to fail. The disease primarily affects young. The mitochondrial disease causes most of the damage to the cells of brain, heart, liver, muscles, kidney, respiratory and the endocrine systems. The symptoms may be as follows depending upon the cells that are affected:

- Loss of motor control,
- Muscle weakness and pain,
- Gastro-intestinal disorders,
- Swallowing difficulties,
- Poor growth,
- Cardiac disease,
- Liver disease,
- Respiratory illness,
- Seizures,
- Visual/hearing problems,
- Lactic acidosis,
- Developmental delays and
- Susceptibility to infection.

Plant Cell Mitochondria

Like in other eukaryotic cells, the mitochondria in plants play an important role in the production of ATP via the process of oxidative phosphorylation. Mitochondria also play essential roles in other aspects of plant development and performance. It also has various properties which allow the mitochondria to interact with special features of metabolism in plant cell.

NOTES

Animal Cell Mitochondria

Mitochondria are known as "power houses" of the cells; they are unusual organelles and are surrounded by a double membrane. These organelles have their own small genome. They divide independently by simple fission. The division of the mitochondria is a result of the energy demand, so the cells with high need of energy have greater number of mitochondria.

The process creating energy for the cell is known as cellular respiration. Most of the chemical reactions of this process happen in the mitochondria.

A typical animal cell may have about 1000 to 2000 mitochondria.

3.7 Mitochondrial Genome Organization

Mitochondrial DNA (mt-DNA) encodes essential components of the cellular energy-producing apparatus, and lesions in mt-DNA and mitochondrial dysfunction contribute to numerous human diseases. Understanding mt-DNA organization and inheritance is therefore an important goal. Recent studies have revealed that mitochondria use diverse metabolic enzymes to organize and protect mt-DNA, drive the segregation of the organellar genome, and couple the inheritance of mt-DNA with cellular metabolism. In addition, components of a membrane-associated mt-DNA segregation apparatus that might link mt-DNA transmission to mitochondrial movements are beginning to be identified. These findings provide new insights into the mechanisms of mt-DNA maintenance and inheritance.

3.8 Mitochondrial Biogenesis

Mitochondrial biogenesis is the process by which new mitochondria are formed in the cell. Mitochondrial biogenesis is activated by numerous different signals during times of cellular stress or in response to environmental stimuli. The

mitochondrion is a key regulator of the metabolic activity of the cell, and is also an important organelle in both production and degradation of free radicals. It is reckoned that higher mitochodrial copy number (or higher mitochondrial mass) is protective for the cell.

NOTES

Mitochondria are produced from the transcription and translation of genes both in the nuclear genome and in the mitochondrial genome. The majority of mitochondrial protein comes from the nuclear genome, while the mitochondrial genome encodes most parts of the electron transport chain along with mitochondrial rRNA and tRNA. A major adaptation to mitochondrial biogenesis results in more mitochondrial tissues which increases metabolic enzymes for glycolysis, oxidative phosphorylation and ultimately a greater mitochondrial metabolic capacity.

The master regulators of **mitochondrial** biogenesis appear to be the peroxisome proliferator-activated receptor gamma (PGC) family of transcriptional coactivators, including PGC-1α, PGC-1β, and the PGC-related coactivator, PRC. PGC-1α, in particular, is thought to be a master regulator. It is known to co-activate nuclear respiratory factor 2 (NRF2/GABPA), and together with NRF-2 coactivates nuclear respiratory factor 1 (NRF1). The NRFs, in turn, activate the mitochondrial transcription factor A (tfam), which is directly responsible for transcribing nuclear-encoded mitochondrial proteins. This includes both structural mitochondrial proteins as well as those involved in mt-DNA transcription, translation, and repair.

3.9 Plant Vacuole: Structure

The vacuoles of plant cells are multifunctional organelles that are central to cellular strategies of plant development. They share some of their basic properties with the vacuoles of algae and yeast and the lysosomes of animal cells. They are lytic compartments, function as reservoirs for ions and metabolites, including pigments, and are crucial to processes of detoxification and general cell homeostasis. They are involved in cellular responses to environmental and biotic factors that provoke stress. In the vegetative organs of the plant, they act in combination with the cell wall to generate turgor, the driving force for hydraulic stiffness and growth. In seeds and specialized

storage tissues, they serve as sites for storing reserve proteins and soluble carbohydrates. In this way, vacuoles serve physical and metabolic functions that are essential to plant life.

Plant cell vacuoles were discovered with the early microscope and, as indicated in the etymology of the word, originally defined as a cell space empty of cytoplasmic matter. Technical progress has variously altered the operating definition of the plant vacuole over time. Today, definitions continue to be colored by the tools and concepts brought to bear in any given study. Indeed, the combination of microscopy, biochemistry, genetics, and molecular biology is fundamental to research into the plant vacuole.

NOTES

There are millions of cells in every multi-cellular living organism. These cells are different in their structure and function. Furthermore, the cells found in different living species also vary. There are many differences between an animal cell, and a plant cell. A plant cell is photosynthetic in nature, so some of its organelles are different from those that are found in an animal cell. One such plant cell organelle which is unique to plants is the vacuole.

In plant cells, the vacuole takes up a large amount of space. At times, this could be more than 90% of the plant cell space. It is said that they are usually formed by the fusion of many membrane vesicles. Due to this reason, a vacuole does not have any specific size or shape. Its structure is such that it complements its function. Many mature and grown plant cells usually have a single large vacuole, surrounded by a structure known as a tonoplast. This is said to be a very active and dynamic membrane of this all important part of plant cell structure.

The vacuole in its central, hollow region contains a fluid known as cell sap. This fluid contains different compounds, some of which are secretory, and some are excretory in nature. Also, the one in the middle, depending on the cell type and requirement, contains various concentrations of salts, sugars and different kinds of soluble pigments. The cell sap, which is a part of the central vacuole structure, also contains various enzymes that are even capable of digesting the cell itself. Although most mature plant cells contain a large single vacuole, when studying the cell biology in young plant cells, there are many vacuoles which slowly enlarge and eventually coalesce together. This

eventually pushes the cytoplasm, nucleus and other such structures against the plasma membrane and the cell wall

NOTES

The vacuole structure is designed to aid this important cell organelle, which is one of the key plant and animal cell differences. Its membrane, the tonoplast, helps to separate its contents from those floating in the cytoplasm. Thus, this membrane does not allow harmful substances present in here from entering and harming the rest of the cell. As the tonoplast is selectively permeable in nature, it also tends to maintain the pH and ionic concentration of the cell, by regulating what travels in and out of the vacuole. Also, due to the structure being so large, it pushes the contents of the cell to the borders, near the cell wall and cell membrane and thus, helps maintain the turgor pressure of the cell. Furthermore, the vacuole sap contains many different digestive enzymes that are capable of destroying the cell (which come handy when there is a need for apoptosis). The tonoplast also aids in, along with maintaining turgor pressure, supporting the structures of the leaves and flowers of the plant.

This structure has been designed to aid in the vacuole function. There is usually a slightly acidic, that is, a low pH maintained inside this structure, because this helps in the functioning of the degradative enzymes present inside the it. In some cases, its structure may differ slightly depending on the type of cell it is present in and its function, for example, in yeast cells, the vacuole is a very dynamic structure that can change its morphology. However, in general, it usually contains a tonoplast and the cell sap within it. In general, the functions of the vacuole include:

- Isolates materials that might be harmful or a threat to the cell
- Containes waste products
- Containes water in plant cells
- Maintaines internal hydrostatic pressure or turgor within the cell
- Maintaines an acidic internal pH
- Containes small molecules
- Exportes unwanted substances from the cell

- Allows plants to support structures such as leaves and flowers due to the pressure of the central vacuole
- In seeds, stored proteins needed for germination are kept in 'protein bodies', which are modified vacuoles.

NOTES

3.10 Tonoplast Membrane

The vacuole plays an important role in the homeostasis of the plant cell. It is involved in the control of cell volume and cell turgor; the regulation of cytoplasmic ions and pH; the storage of amino acids, sugars, and CO_2; and the sequestration of toxic ions and xenobiotics. These activities are driven by specific proteins present in the tonoplast. According to the chemiosmotic model for energy-dependent solute transport, the proton-motive force generated by either the V-ATPase or the H^+-translocating inorganic pyrophosphatase (V-PPase) can be used to drive secondary solute transports. Movement of ions and water down their thermodynamic potentials is achieved by specific ion channels and water channels (aquaporins). The resulting ion, water, and metabolite fluxes across the vacuolar membrane are crucial to the diverse functions of the vacuole in plant cells, such as cell enlargement and plant growth, signal transduction, protoplasmic homeostasis, and regulation of metabolic pathways .

Recent studies have demonstrated the existence of a group of organic solute transporters, belonging to the ABC superfamily, which is directly energized by MgATP. These pumps are competent in the transport of a broad range of substances, including sugars, peptides, alkaloids, and inorganic anions. Belonging to the ABC family, the multidrug resistance–associated proteins (MRPs) identified in plants are considered to participate in the transport of exogenous and endogenous amphipathic anions and glutathionated compounds from the cytoplasm to the vacuole. They function in herbicide detoxification, cell pigmentation, storage of antimicrobial compounds, and alleviation of oxidative damage. A role for plant MRPs is also suspected in channel regulation and transport of heavy metal chelates.

3.11 ATPase, transporters, as storage organelle

NOTES

Most plant cells contain one or several vacuoles, which may occupy up to 95% of the cellular space. The vacuoles often appear empty under a light microscope (hence their name), except when they contain pigments or precipitated substances. The vacuole is delimited from the cytosol by the vacuolar membrane, which is also called tonoplast. Vacuoles are compartments of the secretory pathway derived from the endoplasmic reticulum and the Golgi apparatus. Their function includes the storage of ions, sugars, proteins and xenobiotics. They also participate in volume changes during growth and development, movements such as stomata opening and closing and the maintenance of internal turgor pressure for the mechanical stiffness of green tissues. Although most plant cells have a single central vacuole, some cells have two different vacuoles with different contents and functions. Vacuole biogenesis is a complicated process involving several intermediate compartments, vesicle trafficking and fusion.

The vacuolar system of eukaryotic cells consists of the entire internal membrane network of the cell excluding semiautonomous organelles, such as mitochondria and chloroplasts. This chapter provides an overview of some of the recent advancements in understanding the structure, function, molecular biology, biogenesis, and evolution of vacuolar-ATPases. Fundamental enzymes are highly conserved and maintain a high degree of amino acid sequence homology through evolution. Sequence conservation reveals not only the evolution of the given enzyme, but also indicates events that led to the evolution of eubacteria, archaebacteria, and eukaryota. Fundamental enzymes function in every known living cell and, therefore, nature did not permit eliminating them. Consequently, a mutation leading to their inactivation is not permitted, and if it occurs, it causes lethality. This limits the diseases in which they may be involved and makes them less appealing for the less informed public.

During vegetative and embryonic developmental transitions, plant cells are massively reorganized to support the activities that will take place during the subsequent developmental phase. Studying cellular and subcellular changes that occur during these short transitional periods can sometimes present

challenges, especially when dealing with *Arabidopsis thaliana* embryo and seed tissues. As a complementary approach, cellular reprogramming can be used as a tool to study these cellular changes in another, more easily accessible, tissue type. To reprogram cells, genetic manipulation of particular regulatory factors that play critical roles in establishing or repressing the seed developmental program can be used to bring about a change of cell fate. During different developmental phases, vacuoles assume different functions and morphologies to respond to the changing needs of the cell. Lytic vacuoles (LVs) and protein storage vacuoles (PSVs) are the two main vacuole types found in flowering plants such as *Arabidopsis*. Although both are morphologically distinct and carry out unique functions, they also share some similar activities. As the co-existence of the two vacuole types is short-lived in plant cells, how they replace each other has been a long-standing curiosity. To study the LV to PSV transition, leafy cotyledons a key transcriptional regulator of seed development, was overexpressed in vegetative cells to activate the seed developmental program. At the cellular level, Arabidopsis leaf LVs were observed to convert to PSV-like organelles. This presents the opportunity for further research to elucidate the mechanism of LV to PSV transitions. Overall, this example demonstrates the potential usefulness of cellular reprogramming as a method to study cellular processes that occur during developmental transitions.

NOTES

3.12 Summary

Chloroplasts occur in the cytoplasm of all the green cells of plants except autotrophic prokaryots. Chloroplasts are self-duplicating, cellular organelles, where the complete photosynthesis occurs. They exhibit a certain degree of functional autonomy. Chloroplasts thus have three different membrane types, the outer membrane, the inner membrane and the thylakoid membrane. The thylakoid membrane contains all the enzymatic components required for photosynthesis. Interaction between chlorophyll, electron carriers, coupling factors and other components takes place within the thylakoid membrane. Light energy is used for spitting water to oxygen. Reduction of NADP+ and ATP take place simultaneously.

NOTES

Mitochondria arose once in evolution, and their origin entailed an endosymbiosis accompanied by gene transfers from the endosymbiont to the host. Anaerobic mitochondria pose a puzzle for traditional views on mitochondrial origins but fit nicely in newer theories on mitochondrial evolution that were formulated specifically to take the common ancestry of mitochondria and hydrogenosomes into account. The presence of mitochondria in the eukaryote common ancestor continues to change the way we look at eukaryote origins, with endosymbiosis playing a more central role in considerations on the matter now than it did twenty years ago. The integral part that mitochondria play in many aspects of eukaryote biology might well reflect their role in the origin of eukaryotes themselves.

Evolutionary perspectives place vacuoles at a central position in the physiological strategies of plants in their environment. In the vast majority of cells from the plant body, vacuoles provide the true *milieu intérieur*. They are responsible for the high cell surface–to–protoplasmic volume ratio required for extensive exchanges of material and information between cells and their environment. In cooperation with the cell wall, they create turgor, which is basic to cell hydraulic stiffness and plant growth. In specialized cells, pigment- and allelochemical-accumulating vacuoles serve as mediators of plant–plant, plant–microorganism, and plant–herbivore interactions. In seeds, vacuoles store proteins to be used for anabolism during seedling growth. The diversity of vacuolar functions parallels diversity in morphology, biochemistry, and biogenesis.

3.13 Glossary

- **Chloroplast :** A chlorophyll-containing plastid found in algal and green plant cells.
- **Cytoplasm :** All the contents of a cell, including the plasma membrane,but not including the nucleus.
- **Plastid :** Any of several pigmented cytoplasmic organelles found in plant cells and other organisms, having various physiological functions, such as the synthesis and storage of food.

- **Photosynthesis :** A means of acquiring energy for metabolism which involves trapping radiant energy in chloroplasts, the use of that energy to break up water molecules (hydrolysis) and to convert released energy into an accessible form - such as the molecule ATP. The only form of autotrophy in eukaryotic cells. Some heterotrophic protists have symbiotic algae which allow them to exploit photosynthesis.
- **Photosynthetic Pigments :** Large molecules in chloroplasts that absorb radiant energy (hence they have colour), mostly chlorophylls and carotenes and, occasionally phycobilins.
- **DNA (Deoxyribonucleic Acid) :** A two-stranded molecule that contains the genes, or blueprints, for the makeup of a person's body.
- **mt-DNA:** Mitochondrial DNA, the blueprint for the makeup of the mitochondria.
- **Mitochondrion:** The part of the cell that is responsible for converting nutrients into energy as well as many other specialized tasks.
- **Krebs cycle (Citric acid cycle) :** The sequential (and cyclic) oxidation of substrates and conservation of energy by enzymes of the mitochondrial matrix.
- **Matrix (of Mitochondria) :** The compartment surrounded by the inner membrane of mitochondria.
- **Oxidation :** 1. The combination of oxygen with a molecule with liberation of heat; combustion. 2. Removal of hydrogen from a molecule. 3. Loss of electrons, increasing the positive valence. Oxidation is associated with a loss of energy.
- **Reduction (Chemical) :** A reaction involving the gain of electrons, addition of hydrogen atoms, or loss of oxygen atoms. Reduction is associated with a gain of energy.

NOTES

3.14 Self-Learning Exercise

NOTES

Section-A (Very Short Answer Type Questions)

1. What important pigments are found in the chloroplasts?
2. Is a tree a good example of chloroplasts?
3. What reaction takes place in the chloroplast?
4. In what part of a chloroplast does glucose production occur?
5. What specialized structures are found inside the chloroplast?
6. What is the meaning of the biology acronym NADPH?
7. Which organelle is called as powerhouse of cell?
8. Who coined the term Mitochondria?

Section-B (Short Answer Type Questions)

1. Why are chloroplast found only in plant cells?
2. What is the flattened sac in a chloroplast? And what is its function?
3. Is mitochondrion a double membrane organelle?
4. Where is mitochondrion located in a cell?
5. Write short note on function of vacuole.
6. Give function of ATPase transporters.
7. What is tonoplast?

Section-C (Long Answer Type Questions)

1. Describe in detail the structure and function of chloroplast?
2. Write a note on nucleocytoplasmic interactions?
3. What is the Origin of the Mitochondria?
4. Why is the Mitochondria Important?
5. What are Differences between Nuclear and Mitochondria DNA?
6. Describe the structure and function of vacuole.
7. What are ATPase transporters?

Answers Key of Section- A

1. Chlorophyll causes plants to turn green.
2. Yes, trees are full of chloroplasts in their leaves.
3. Light and dark reactions

4. Stroma
5. chlorophyll and other structures in a cell
6. Nicotinamide Adenine Dinucleotide Phosphate.
7. Mitochondria
8. Benda

NOTES

3.15 References

- Burgess, Jeremy (1989). *An Introduction to Plant Cell Development*. Cambridge: Cambridge University press. p. 62. ISBN 0-521-31611-1.
- www. Bendich, A. J. (2004). "Circular Chloroplast Chromosomes: The Grand Illusion". *The Plant Cell Online* **16** (7): 1661–6. doi:10.1105/tpc.160771. PMC 514151. PMID 15235123.
- Henze K, Martin W; Martin, William (2003). "Evolutionary biology: essence of mitochondria". Nature 426 (6963): 127–8. doi:10.1038/426127a. PMID 14614484.
- www. "mitochondria". Online Etymology Dictionary.
- www. Campbell, Neil A.; Brad Williamson; Robin J. Heyden (2006). Biology: Exploring Life. Boston, Massachusetts: Pearson Prentice Hall. ISBN 0-13-250882-6.
- www. McBride HM, Neuspiel M, Wasiak S (2006). "Mitochondria: more than just a powerhouse". Curr. Biol. 16 (14): R551–60. doi:10.1016/j.cub.2006.06.054. PMID 16860735.
- Brooker, Robert J, et al. (2007). *Biology (First Edition)*, (McGraw-Hill, New York), p. 79 ISBN 0-07-326807-0
- Plant cells vs. Animal cells. Biology-Online.org

Unit - 4

Nucleus

NOTES

Structure of the Unit

4.0 Objectives

This unit will make student understand the structure of nucleus which determines the characters of the cell and character of individual.

- DNA structure & different forms of DNA
- Plant promoters and transcription factors
- mRNA & r RNA and its biosynthesis

4.1 Introduction

By going through this unit students will come to know about the nucleus which is most essential part of the cell. The role of nucleus is to direct & control the function of the cell and ultimately characters of the individuals. Besides, you will know about nucleopore and nucleosome also.

NOTES

DNA (de-oxyribose nucleic acid) is found in all living organisms. DNA is the genetic material which transmits hereditary informations from one generation to another. DNA is the most stable biological compound which can be considered as genetic materials. It has potentiality to carry out all kinds of biological activities of cells.

DNA has the power of self replication and repair. It controls and regulates the biosynthesis of protein by the processes of transcription and translation.

In cellular organism, RNA is also present with DNA in good quantity and different forms which play a very important role in the synthesis of cellular proteins which are called non genetic RNA (t RNA, m RNA, r RNA)

4.2 Structure of Nucleus

Nucleus is most essential and prominent part of the cell which directs and controls all activities of the cell. A typical nucleus (Karyon) is a largest and conspicuous organelle which is bound with double membrane. It contains thread like structure called chromosomes which are made up of genetic material i.e. highly coiled DNA molecules. Nucleus encloses all genetic informations and transmits them to next generation. A true nucleus is present in all eukaryotic cells. The nucleus of plant cell and animal cell has the similar structure and function.

Nucleus was originally detected by **Leeuwenhoek** in 1700 in blood corpuscles of salmon blood. **Fontana** (1781) observed similar bodies in the skin cells of eel.

Robert brown (1831) detailed out distribution of nucleus in orchid cells and showed the presence of genetic material in the cell.

The nucleus is present in all eukaryotic cells except prokaryotes but some mature plant and animal cells are without nucleus e.g. sieve tubes of phloem

cells vascular plants and mammalian RBCS. The cells which are deprived of nuclei, fail to devide and differentiate. Cell division takes place only in those cells which possess nuclei.

NOTES

Position of nucleus is characteristic of each cell type. Generally it occupies the central position in the cell but in mature cells, its position becomes peripheral due to the presence of a large vacuole.

In plant cell, it is held in position by cytoplastic strand. In animal cell, it is free floating in the cell cytoplasm. Its position is based in glandular cells and in Acetabularia (an alga)

Shape of nucleus varies in different types of cells. Shape of the nucleus is variable depending upon the internal environment of the cell and state of activity. Generally the nucleus is spherical and located centrally in young meristematic cells but becomes flat, oval or irregular in mature and fully differentiated cell.

In human neutrophils nucleus is trilobed. It is branched in silk spinning cells. Nucleus is lanceolate in sperm cell and its shape is irregular in leucocytes and cartilage cells

(Fig.4.1)

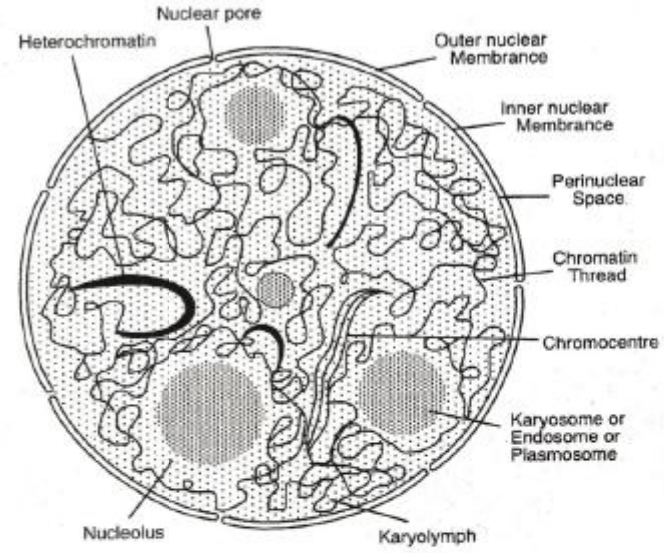


Fig. 4.1 : Ultra Structure of Nucleus

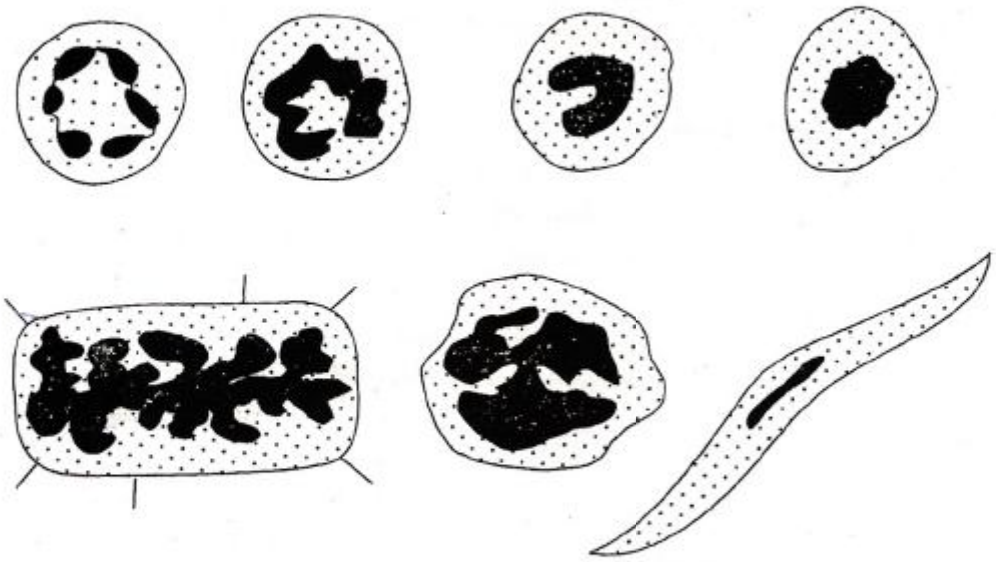


Fig. 4.2 : Different Shapes of Nuclei

Shape of Nucleus depends upon following

i. **Nucleoplasmic Index (NP) :** Hertwig (1906) suggested that there is a definite ratio between the volume of cytoplasm to that of nucleus means if volume of cytoplasm increases the volume of nucleus will also increase. It was expressed by nucleoplasmic index (NP) as shown below–

$$NP = \frac{V_n}{V_c - V_n}$$

Where

NP = Nucleoplasmic index

V_n = Volume of nucleus

V_c = Volume of cytoplasm

Every cell has a definite nycleoplasmic index, any change in this will result in cell division.

ii. **Chromosome number** : size of nucleus depends upon the number of chromosomes it possesses e.g. size of haploid nucleus is small as compared to diploid nucleus. The polyploid cell has still larger nucleus.

iii. **Metabolic activity of cell** : The nuclei are larger in those cells which are metabolically active as composed to those which are metabolically inactive.

NOTES

Number

Generally there is a single nucleus in each cell. Such cells are called uninuceleate.

On the basis of presence, absence and number of nuclei, cells are of following types –

1. **A nucleate :** The cells without nucleus are called anucleate e.g. Prokaryotes, sieve tube of phloem of plant and RBC of mammal's.
2. **Mononucleate :** Most common type containing single nucleus.
3. **Binucleate :** Such cells contain two nuclei per cell e.g. liver cells and paramaecium.
4. **Polynucleate :** Cells with many nuclei are called polynucleated cells e.g. multinucleated animal cells (cell of plasmodium) are called **syncytia** and in plants they are called **coenocytes**.

Chemical composition

Chemically, nucleus is composed of 12% DNA, 5% RNA, 80% protein 3 % lipid, inorganic salts and other enzymes.

Proteins : The protein occurs as structural proteins –

1. **Basic protein** (low molecular weight proteins) : Basic protein are of two types – the histones and protamines. The protamines participate in the maintenance and reproduction of the chromosomes.
2. **Acidic protein** (High molecular weight protein) : They are non histone protein present outside the chromosomes. They take part in the metabolic functions of the nucleus.
3. **Enzymes** : There are numerous enzymes in the nucleus e.g. Diphosphopyridine nucleotide synthetase, DNA & RNA Polymerases (essential for synthesis of DNA), NAD-synthetase, diaminase, guanase, aldolase etc.

Inorganic salt : The inorganic salts of sodium, calcium, Magnesium, iron and phosphorus are associated with nuclear proteins and enzymes, Inorganic salts are present in traces in the nucleoplasm but they are of great significance.

Structure of Nucleus

NOTES

There are five components of nucleus.

1. Nuclear envelop (Karyotheca)
2. Nucleolus
3. Nuclear sap (Karyolymph)
4. Chromatin
5. Nucleolus

Nuclear Envelope or Nuclear Membrane or Karyotheca

Nuclear envelope is a double membraned envelope which separates the nuclear contents from cytoplasm.

1. Outer membrane (ectokaryotheca)
2. Inner membrane (endokaryotheca)

1. Outer Membrane

This membrane is about 75 $\overset{0}{A}$ thick. It is thicker than inner membrane.it is rough due to the presence of ribosomes on its surface. it is continuous with membrane of ER at some point.

2. Inner Membrane

It is also 75 $\overset{0}{A}$ thick. It is smooth lacking ribosomes. Its inner or nuclear surface is coated with a homogenous filamentous fibrous lamina which is 300 $\overset{0}{A}$ thick.

Perinuclear Space or Cisternae: It is about 100-300 $\overset{0}{A}$ wide. This space is filled with fluid which may contain fibres antibodies, crystalline bodies or lipid droplets. This space is continuous with the channel of endoplasmic reticulum.

4.2.1 Nuclear Pores

Nuclear membrane or Karyotheca is perforated by pores called **nuclear pores**. Nucleus occupies a central position among the cellular structures.

NOTES

Nuclear pores were first observed by **Callan & Tomlin (1950)**. Generally the pore looks as octagonal areas. They are formed by fusing the two nuclear membrane at certain places. Theses pores may be of 400-1200 $\overset{0}{A}$ in diameter. Pores are occupied by a ring like arrangement, called annulus of fibrillar material which appears to be continuous with the inner fibrous lamina. The annulus and pore organized to form a pore complex. The annulus around the pores regular exchanges the macromolecules in relation to their size and chemical nature.

The pores remain open and close under different circumstances. The nuclear pores are concerned with the exchange of materials between nucleus and cytoplasm. Macromolecules like ribonucleo protein, gold particles and some ion Na^+, K^+, Cl^- can pass through the nuclear pores from nucleoplasm to the cytoplasm.

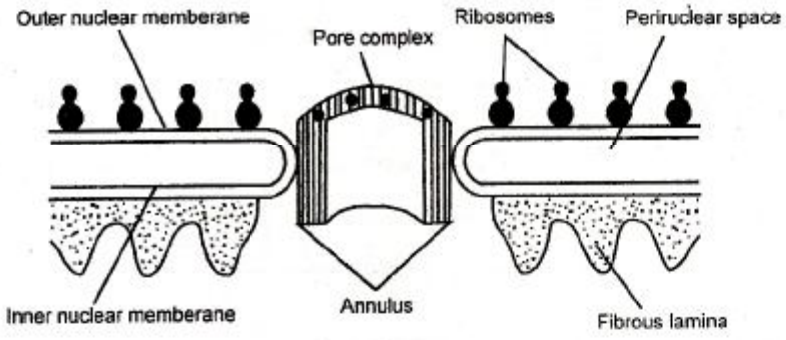


Fig. 4.3 : Nuclear Pore

The nuclear pore is a large complex structure of 125 million daltons or 30 times the size if a Eukaryotic ribosome. The pore is 120nm in diameter and 50 nm in thickness.

It is consist of flour elements:

1. The Scaffold : It is stack of closely opposed cytoplasmic ring, nucleoplasmic ring and a central ring of thick spokes, the spoke of central ring are attached to transporter on the inner side and to the nucleoplasmic and cytoplasmic rings on the outer side.

Inter space between the spoke are aqueous channels which (9nm wide) allows diffusion of protein and metabolites between the nucleus and the cytoplasm.

2. Transporter : Transporter is a (36-38nm in diameter) protinaceous ring. It is consist of two irises of eight arms each. Two irises are stacked on top of one another. It carries out active transport (import and export) of protein and RNAS. They open like diaphragm of a camera to let a nuclear protein or RNA pass through from the nucleus to cytoplasm.

NOTES

3. Thich Filament : They are attached to the cytoplasmic side of the pore.

4. Basket : A large basket like structure is found on the nuclear side, which is consist of 8 filaments (each 100nm long) extending from nucleoplasmic ring of the pore and meeting a smaller ring with in the nucleus. The basket may play an important role in RNA export.

4.2.2 Nucleosome Organization

Nucleosome (Sub unit of chromatin)

Eukaryotic chromosomes are composed of DNA, Protein and a very small amount of RNA. Proteins are histone proteins which contain high percentage of basic amino acids (arginine & lysine).

The chromatin fibre was studied further under electrone microscope and in 1974 R.D. Kornberg and J.O. Thomas proposed model for basic chromatin fibre as a flexibly jointed chain resembling with beads on a string forming a number of repeating units the nucleosomes. This model explains that nucleosome is the basic structure of chromatin. Each nucleosome is consist of a spiral of DNA wrapped around an octamer of histone protein molecules forming core particle known as platysome. The double helical DNA contains 200 base pairs twines around a set of 8 histone molecules. This DNA is called core DNA.

Each nucleosome is consist of a core particle and a linker DNA or interrnucleosomal DNA that can be folded or packed in to tiny volume. Nucleosome has a diameter of $200\overset{0}{A}$.

The Core Particle : The core particles is consist of the octamer of histones which has two copies of H_2A, H_2B, H_3 and H_4 proteins. A strand of DNA is wrapped around this core. This DNA has 200 base pairs.

NOTES

The Spacer DNA : It is a small segment of DNA which is consist of 4 base pairs. H_1 Histone protein is associated with this DNA. In linker DNA, number of base pairs varies from 1 to 80.

Nucleosome Packing : In nucleus of Interphase, thin chromatin fibre is a linear array of nucleosomes like beads on a string. Spiral coiling of thin chromatin fibres give rise to thick chromatin fibres. This shows solenoid type of ultrastrcture which 6 or 7 nucleosomes per turn.

At the stage of metaphase or anaphase supersolenoid (400 $\overset{0}{A}$ diameter) is condensed to produce define shape and dimensions of the chromosome.

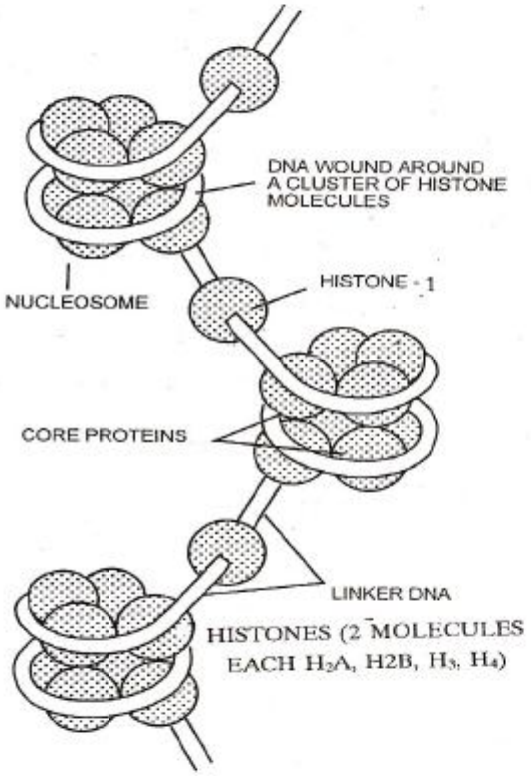


Fig. 4.4 : Nucleosome Model of Chromatin Structure

(Core Particles Linker DNA & Histone Protein))

4.3 Structure of DNA

Deoxyribose nucleic acid (DNA) is a macromolecule forming the genetic material (the major store house of genetic information) of all Eukaryotic and Prokaryotic cells, some intra celluler organelles and many viruses, where RNA acts as genetic material. It is a polymer of nucleotide (Polynucleotide) which is

present in chromosomes as a major constituent and in plasmid as a sole constituent.

DNA is the most complex and heaviest macro molecule of the cell having molecular weight of 10^8 to 10^{11} or in other words molecular weight in millions.

NOTES

It is the key molecule of living world and acts as a hereditary material. On hydrolysis, these molecules yield a large number of nucleotides (monomers). Each nucleotide is composed of nucleoside and a phosphate group.

The nucleoside consists of a pentose sugar and a nitrogenous base so nucleotide unit is consist of

I. A pentose sugar : The 5 carbon de -oxyribose sugar

II. Nitrogenous base : A heterocyclic ring containing nitrogen

III. Phosphoric Acid

I. **A pentose sugar molecule :** DNA molecules contain deoxyribose sugar. Pentose sugar as a main constituent of nucleic acid, was identified by **lewine(1909)** Deoxyribose sugar has one oxygen atom less than ribose sugar molecule.

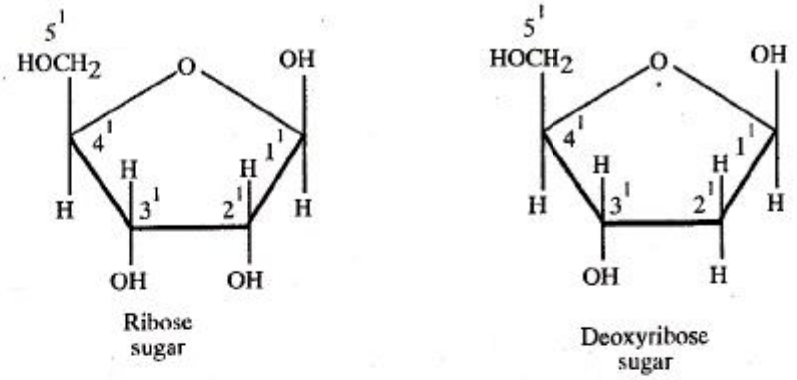


Fig. 4.5 : Pentose Sugar Moleccules

The deoxyribose sugar has a pentagonal ring with 5 carbon atoms (1^1, 2^1, 3^1, 4^1, 5^1). Among 5 carbon 1^1 carbon is attached to the nitrogenous base and 3^1 and 5^1 carbons are attached to Phosphoric acid. At carbon 1^1 OH group is present and it is the point of attachment of the base and it is linked to the 1- nitrogen atom of the pyrimidine bases and to a nitrogen atom of Purine bases.

NOTES

II Nitrogen Base

Nitrogen bases are of two types-

i. **Purines :** Purines are double ringed nitrogen compounds in which two benzene rings are fused. In DNA two types of purines Adenine (A) and Guanine (G)are found. Their structural formulae have been mentioned here.

ii. **Pyrimidines :** Pyrimidines are 6 membered single benzene ringed compound . Thymine (T) and cytosine (C) are pyrimidines. But in RNA instead of thymine, Uracil is present.

Structure formulae of purine and pyrimidine

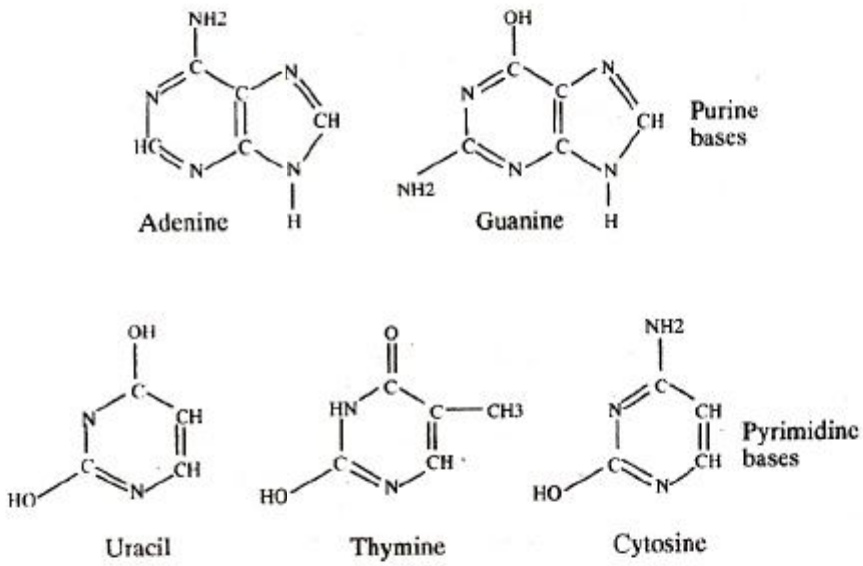


Fig. 4.6 : Different type of Nitrogenous bases

Nucleosides

A molecule of pentose sugar (deoxyribose and ribose) attached with a molecule of nitrogenous base devoid of phosphate group is known as **nucleoside**. The linkage between base and sugar occurs at N-3 in pyrimidine or N-9 in Purine and Carbon-1 of pentose sugar with the elimination of one water molecule.

The nucleosides are named as follows –

Adenine + Deoxyribose : Adenosine

Guanine + Deoxyribose : Guanosine

Cytosine + Deoxyribose : Cytodine

Thymine + Deoxyribose : Thymidadine

NOTES

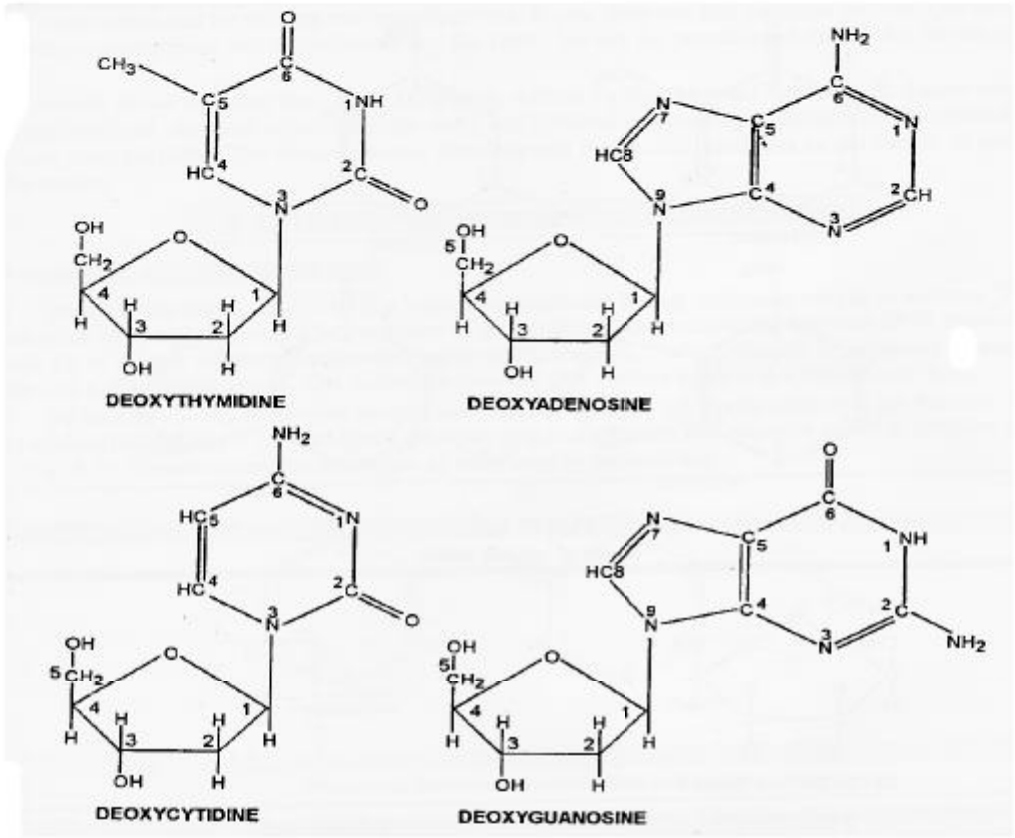


Fig. 4.7 : Nucleosides of DNA

Nucleotide : The nucleotides are monomeric units of nucleic acid. Each nucleotide is consisting of one molecule of deoxyribase sugar, one mol. of phosphoric acid and one of the four nitrogenous bases. Each nucleotide consists of a nucleoside linked with phosphoric acid.

The molecule of phosphoric acid joins with the pentose sugar of nucleoside at carbon-5 or at carbon. The base and sugar are joined by glycosidic or N ribosidic bond and sugar and phosphate are joined by ester bond.

NOTES

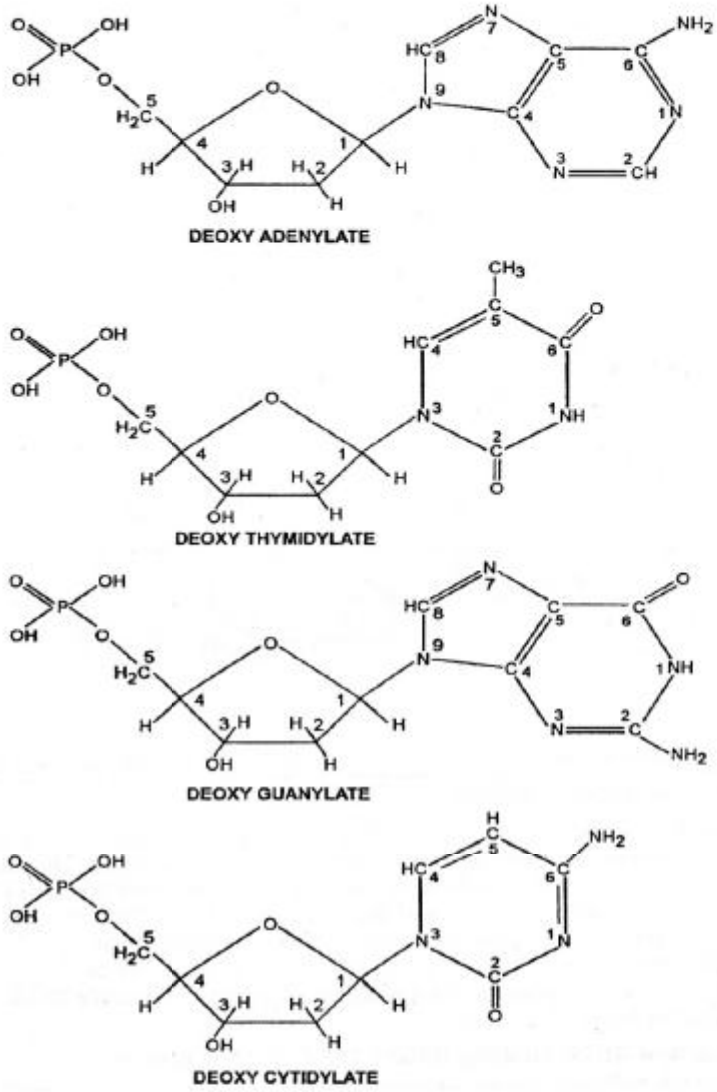


Fig. 4.8 : Nucleotides of DNA

Nucleotides are named as follows

Deoxycytidine + phosphate = deoxycytidylic acid / deoxycytidylate

DeoxyThymidine + phosphate = deoxythymidylic acid deoxythymidylate

Deoxyguanosine + phosphate = deoxyguanylic acid or deoxythymidylate

Deoxyadenosine + phosphate = deoxyadenylic acid / deoxyadenylate

Besides, these nucleotide, there are some other nucleotides which play important role in cell metabolism.

e.g.

FAD (Flavin adenine dinucleotide)

NAD (Nicotinamide adenine dinucleotide)

ADP (Adenosine-5^1- diphosphate)

ATP (Adenosine-5^1 triphosphate)

NADP (Nicotinamide adenine dinucleotide Phosphate)

Poly nucleotide structure

DNA is a macromolecule and consists of several thousand of monomer units which are known as deoxyribonucleotides. Nucleotides are linked together with the help of phosphodiester bonds in a order to make a polynucleotide chain. The back bone of a Polynuceotide chain consists of alternating sugar and phosphate units. The nitrogenous base is attached to the first carbon atom (1^1 Carbon) of the deoxyribose.

They are directed at right angles to the long axis of the Polynucleotide chain and they are stacked one above the other.

A polynucleotide chain has 3^1 and 5^1 ends. In polynucleotide chain a covalent linkage between the phosphoric acid residue of one nucleotide and 3^1 Carbon of the sugar on the next nucleotide takes place. This linkage is called as 3^1 5^1 phosphodiester bond (because phosphate is doubly esterified to two OH groups, one is linked 3^1 Carbon and other one is attached to Carbon 5^1

A polynucleotide chain has 3^1 and 5^1 ends, it means one end of the chain has a sugar residue with C3 carbon atom which is not attached to another nucleotide.

4.3.1 Types of DNA

Mainly there are two types of DNA

1. Linear DNA
2. Circular DNA

1. **Linear DNA :** Linear DNA is found in the nuclei of eukaryotic cells and in some DNA viruses. Each linear DNA is thread like, long and has two free ends 3^1 end and 5^1 end. The eukaryotic linear DNA is associated with protein in highly condensed form. It is organized in to a number of chromosomes. Every chromosome has single long DNA helix. But viral

NOTES

DNA molecules are linear but not tightly associated with protein. They are packed in an envelope of protein.

NOTES

2. **Circular DNA :** This form of DNA molecule has no free ends so it is called **circular DNA**. This circular DNA may be single stranded (e.g. few small bacteriophages) or double stranded helical (e.g. large bacteriophages, Prokaryotes, mitochondria and chloroplasts). The Mitochondria DNA (mitDNA) is double stranded circular DNA which is different from nuclear DNA. In prokaryotes the DNA is circular so there are no free 3^1 and 5^1 ends. Circular DNA is not associated with protein but they occur in compacted form like linear DNA molecules.

Different forms of linear DNA

There are 5 forms of linear DNA A-, B-, C-, D- and Z DNA. The difference between these forms is due to direction of coiling, distance between successive base pairs and number of nucleotide per turn.

1. **A- DNA :** the double helical structure of DNA (described by **Watson and Crick**) have right handed helical coiling. In this type of DNA, the length of helix is reduced to $28,15\overset{0}{A}$ as compared to $34\overset{0}{A}$ length of helix in 'B' form DNA given by **Watson and crick**. There are 11 base pairs per turn of helix. There is a tilting at base pairs from the axis of the helix with the axial rise per base pair is about $2.36\overset{0}{A}$. The diameter of helix is $25.5\overset{0}{A}$.

2. **B-DNA** : The double helical structure of DNA which has right handed helical coiling is called **B-DNA**. The structural model for this type of DNA was proposed by **Watson and crick**.

3. It is most common form of DNA. This form of DNA occurs naturally in living organisms. The helix is right handed with the length of about 34 $\overset{0}{A}$. Each turn contains 10 pairs of mononucleotide units. The Diameter of DNA molecules is $20\overset{0}{A}$. The rise of helix per base pair is $3.4\overset{0}{A}$ and the tilt of base pair is 6.3. The repeating unit is mono nucleotide and the

orientation of sugar molecule is not alternating. Each successive base pair in the stack turns 36^0 .

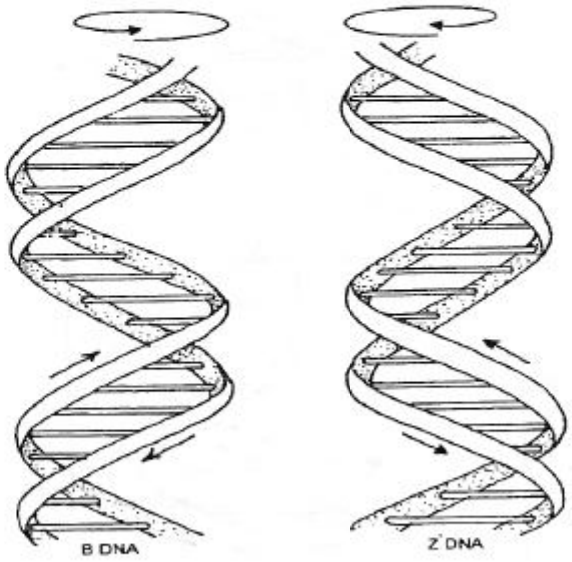


Fig. 4.9 : Right handed DNA(B-DNA) and Left handed DNA (Z-DNA)

Z DNA or Left handed DNA

Rich and his coworkers (1979) have proposed a left handed double helical model for DNA molecule which is called **Z DNA**. This form of DNA has been artificially produced in the laboratory. The Z DNA has following characteristics:

1. Sugar molecules have an alternating orientation so that a repeating unit is a di nuceotide and phosphate groups pursue a zig-zag course.
2. One complete helix i.e. a twist through 360^0 has 12 base pairs.
3. Every turn is 45 $\overset{0}{A}$ long/or say length of one complete helix is 45 $\overset{0}{A}$.
4. The angle of the twist per dinucleotide repeating unit is 60^0.
5. The diameter of Z DNA molecule is 18 $\overset{0}{A}$
6. Z DNA functions in the regulation of gene transcription and this form of DNA is biologically significant.

Similarities between Z-DNA and B-DNA

1. In both forms, DNA molecule is double helical.

2. both forms Z-DNA and B-DNA exhibit guanine cytosine pairing by 3-H bonds (G ≡ C)

NOTES

Differences between Z-DNA and B-DNA

S.No.	Z-DNA	B-DNA
1	Z-DNA helical is left handed	B-DNA helical is right handed
2	Sugar phosphate backbone follows a zig-zag pattern in polynucleotide chain in Z-DNA.	It is regular polynucleotide chain of B-DNA
3	In Z-DNA, sugar molecules have alternating orientation so that repeating unit is dinucleotide	In B-DNA, sugar molecules do not have alternating orientation so that repeating unit is mononucleotide
4	In Z DNA one complete helix has 12 base pair	one complete helix has 10 base pairs in B-DNA
5	Length of one complete helix is $45\,\overset{0}{A}$	Length of one complete helix is $34\,\overset{0}{A}$ in B-DNA
6	The diameter of Z-DNA molecules is $18\,\overset{0}{A}$	The diameter is $20\,\overset{0}{A}$ in B-DNA
7	The angle of twist per repeating unit (dinucleotide) is 60^0	The angle is 36^0 per repeating unit i.e. mononucleotide

4.3.2 DNA Replication

Watson and Crick proposed a detailed molecular structure of DNA. By this structure we can conveniently understand the most important and inevitable process of DNA replication.

One of the most characteristic function of DNA is the self direct reproduction of itself during the interphase just before cell division. So that each daughter cell receives an exact copy of polynucleotide sequence contained in the parent. The processes by which the parent DNA molecule gives rise to two excitedly

similar sister DNA molecules, called replication of DNA. The Watson - Crick Model of DNA provides a replication mechanism.

In the DNA structure, the two Polynucleoride chains are complementary to each other. They follow specific base pairing rule i.e. purine base pair with pyrimidine base (A=T T=A and G≡C or C≡G) and they are linked by weak H-bonds

NOTES

For DNA replication each chain acts as a template on which new complimentary chains are synthesized.

In replication, the DNA duplicates exactly into two chains. Both chains of double stranded DNA move apart from each other by dissolution of H bonds and uncoiling of chain occurs. Due to specificity of base pairing, each nucleotide of the separated chain attracts its complimentary nucleotide from cytoplasm.

Like this new complimentary chains are formed on the templates and newly formed daughter chains are the mirror images of their parallel chains.

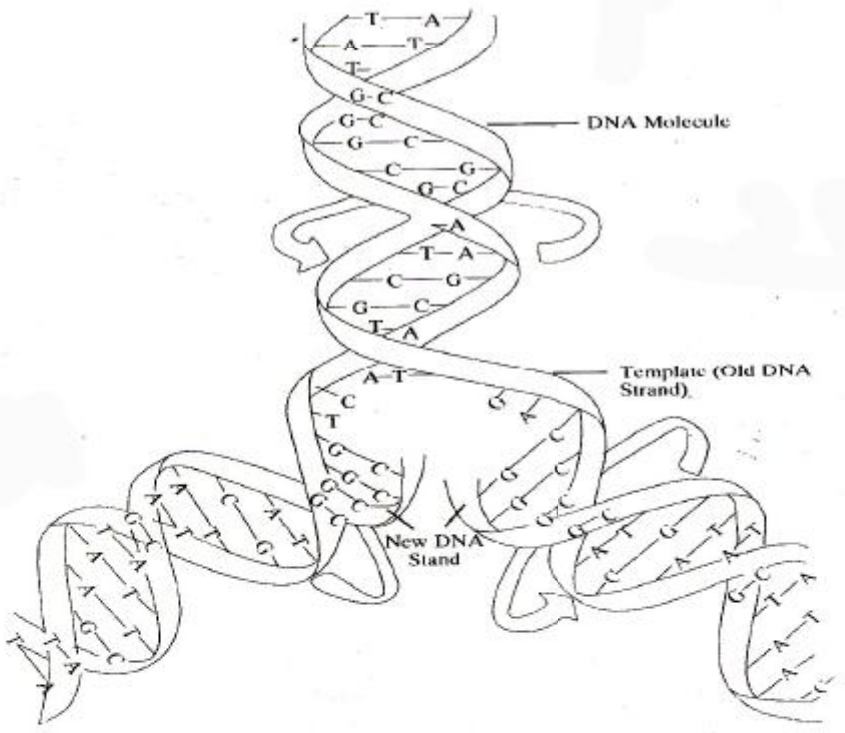


Fig. 4.10 : Semi-Conservative Replication of DNA

Types of DNA Replication

Following are three possible mechanism by which double stranded DNA molecule can replicate –

NOTES

(i) Semi conservative (ii) Conservative (iii) Dispersive

a. **Semi Conservative Replication**: In this method, two strands at DNA molecule separate from each other. Each strand then gets nucleotide from the pool and synthesizes its complimentary strands. Thus one strand at each daughter DNA molecule is derived from the parent molecule and the other strand is formed new.

b. **Conservative Replication** : In this type of replication, one daughter molecule contains both parental polynucleotide where as the other molecules contains both newly synthesized chains.

Or we can say in this conservative method, original DNA molecules remains as such and a brand new copy of DNA molecular is synthesized from the old molecule.

c. **Dispersive Replication** : In this method parent DNA molecule undergoes disintegration and breaks up in to its components nucleotides. These nucleotides along with newly synthesized nucleotides synthesize two double stranded DNA molecules.

4.3.3 DNA Damage & Repair

DNA is generally stable. Its fidelity is maintained through inherent stability of the molecule. There is always a risk of getting DNA damaged by physical and chemical agents; in other words DNA is under constant threat. Therefore maintaining chromosomal integrity, an effective repairing mechanism is essentially required.

In living cells , there is an existence of DNA repair mechanism that finds and corrects the damaged DNA and maintains genetic continuity.

Damage of DNA

Various environmental and internal factors that affect DNA of living cells and cause damage to it are following.

i. Radiation ii. Physical stimulus iii. Chemicals iv. Abnormalities during DNA replication like wrong and irregular base pairing and intra molecular rearrangement in the distribution of H bonds etc.

Theses factors change the basic structure of the DNA molecule. The damaged DNA molecule affects the functioning and vitality of the cell.

NOTES

Types of DNA Damage

1. Mono adduct (single base changes)
2. Di adduct (structural distortions)

1. **Mono adduct :** When only single base gets damaged by chemical mutagen in a DNA molecule.

For example nitrous acid deaminates the bases G C and A with decreasing frequency in DNA and with equal frequencies in RNA.

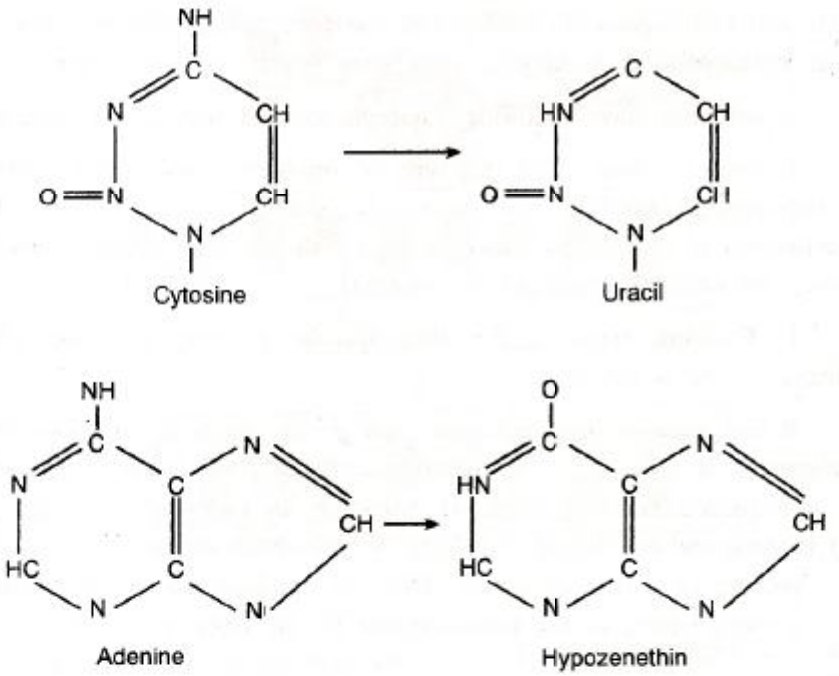


Fig. 4.11 : Damage due to Deamination

2. **Diadduct :** In this type of DNA damage, structural changes in DNA occurs.Exposure to damaging radiations like UV rays results in structural distortion of DNA molecule which takes place due to removal of a base or introduction of a covalent link between bases of same or

NOTES

different strands. For example : Formation of thymine dimer due to UV rays.

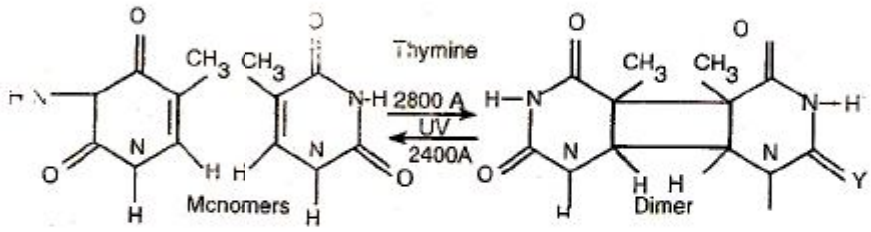


Fig. 4.12 : Formation of A Dimer of Thymine

Repairs of DNA :

All living cells have an inherent capacity of repairing the damaged parts of their DNA. For repairing DNA, there are some special feature by which damaged DNA gets repaired itself. These features are (i) Molecular structure of DNA (ii) self replication power of DNA (iii) ability to recognize damaged part (iv) altered base requires.

Because of these features, cells regain their vitality and stand normal functioning. Prokaryotic and Eukaryotic cells have their repair enzyme system.

A cell has below mentioned system to deal with DNA damage.

1. **Direct repair** : In this type of DNA repairing, no fresh DNA synthesis is needed. This type of DNA repair involves reversal of the damage. For example photo reactivation in pyrimidine dimmer involves removal of covalent bonds so that DNA may have original structure (monomers).
2. **Excision repair** : In this repairing system ,specific enzyme is required which can cleave covalent bond . This occurs in dark and many enzymes are required for this –
   (i) First of all recognition of a damaged or altered base takes place followed by incision of a sequence including damaged bases by an endonuclease.
   (ii) Excision and broadening of the gap by an exonuclease (DNA Polymerase)
   (iii) Synthesis of a new stretch of DNA to replace excised material.
   (iv) By the ligase enzyme covalent joining of the Polynucleotide.

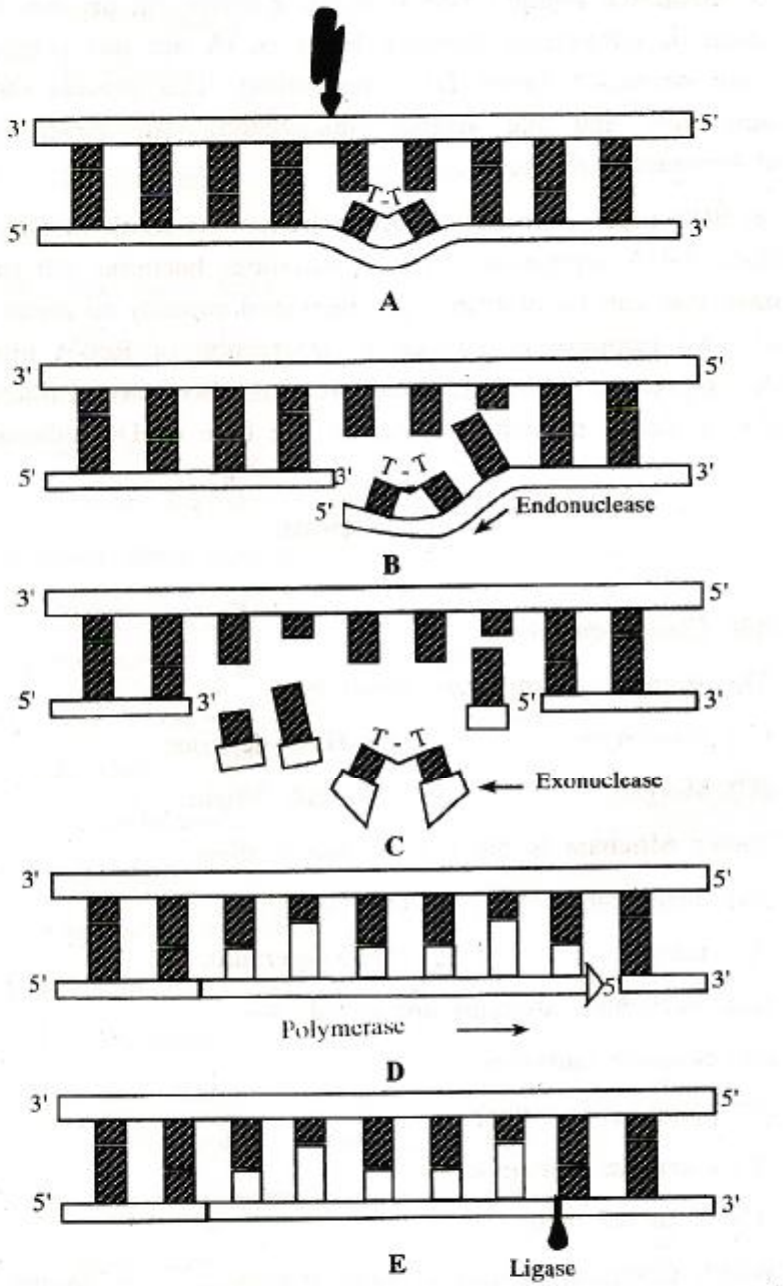


Fig. 4.13 : Diagram showing the effect of uitra violet light UV on DNA molecule

& the DNA repair mechanism

3. **Mismatch repair** : This process is an error correction process. It involves correction of mismatches between bases which are not

NOTES

complementary. This correction process occured during DNA replication. This process helps in recognizing "new" and "old" strands (Methyled). The newly synthesized DNA remains error free.

4. **SOS repair :** In E-coli, few treatments lead to DNA damages or say they inhibit DNA replication. In theses situations bacterial cell shows SOS response that can be described as "Increased capacity to repair damaged DNA". This response is triggered by interaction of Rec-A protein with Lex A repressor. At the time of DNA damage, more than 15 protein start synthesizing and an amount of Rec-A protein increases at the time of DNA damage.

4.4 Transcription

In eukaryotes, there are different promoter sequences for three different RNA polymerases to initiate transcription. There are also enhancer and silencer sequences for RNA polymerase II (RNAPII). The promoter in eukaryotes for RNAPII has a TATA box-PSE or Inr at start point, a CAAT box at -25bp and one or more GC boxes at -100 position RNAPll promoter thus may be $TATA^{+}Inr$, .

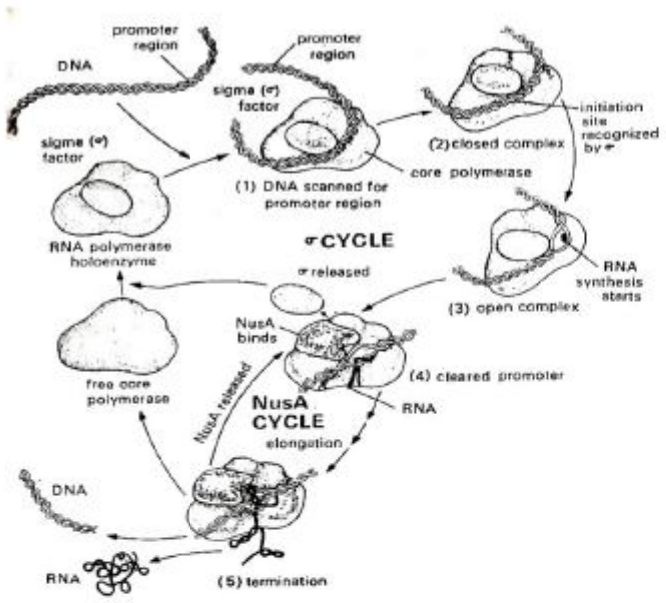


Fig. 4.14 : Role of sigma factor and core enzyme of RNA polymerase during transcription

TATA˙Inr+, TATA Inr or TATA˙Inr˙. Enhancers also called upstream activation sites (UAS) are also present for RNAPII either at -100/200bp position or at any other position, and enhance transcription. Silencers can also be present for RNAPII at long distances and repress transcription (silencers are known in yeast at HML and HMR loci). Promoters for RNAPI and RNAPIII differ from those for RNA~PII .A brief account of transcription by RNAPII is presented here, since >90% of transcription makes use of RNAPII only.

NOTES

The initiation of transcription in eukaryotes involves the formation of a pre-initiation 'transcription complex' or 'transcriptosome'. Seven 'transcription factors' (TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH, TFIIJ) take part in the formation of transcription complex for RNA polymerase II, at the promoter sites of 'house keeping genes' (genes functional in all cells at all times for essential activities).

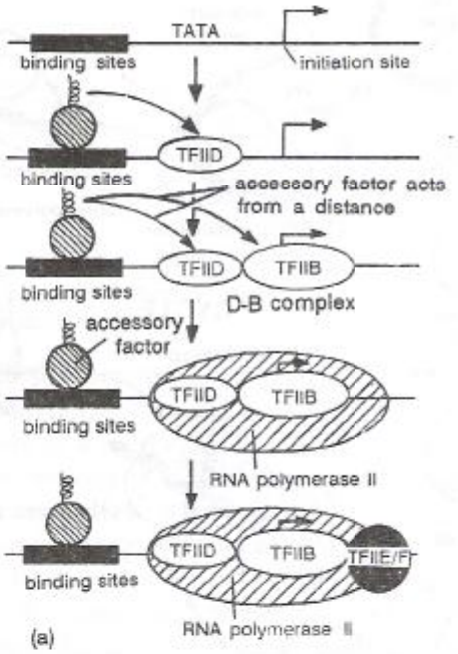


Fig. 4.15 : Formation of Transcription complex at the promoter for RNA polymerase

These promoters are called **'generic promoters'**. There are other genes called **'smart genes'**, whose expression is regulated by additional transcription factors. Hundreds of these transcriptional factors have now been isolated from higher plants and animals. Most of these transcriptional factors have each a

NOTES

'DNA binding domain' and an 'activation domain. On the basis of similarity, DNA binding domains have been grouped into the following main families: (i) helix-turn-helix (HTH); (ii) helix-loop-helix (HLH); (ill) Zinc fingers; (iv) leucine zippers. Similarly activation domains have been grouped into (i) acidic domains, (ii) glutamine rich domains and (iii) proline rich domains. Detailed structure and function of a number of transcription factors have been studied in recent years.

Transcription by RNAP II may be (i) basal transcription, when it involves house keeping genes and makes use of only general transcription factors (GTFs) or may be (ii) activator dependent transcription, when regulated expression requires activator(s) and/or coactivator(s), in addition of GTFs. Basal transcription requires the action of several transcription factors in the following defined order (Fig: 14 (i) TFIlD, consisting of TATA binding protein (TBP) and TBP associated factors (TAFs), binds at the TATA box; (ii) TFIIA and TFIIB bind to TFIID-DNA complex to form TFIID-TFIIA-TFIIB-DNA complex (also called. DAB complex); an intermediate DA complex or DB complex may be involved in the formation of DAB complete (TFIIA is involved only in activator dependent transcription and TFIIB provides a link between promoter and RNAPII; (iii) RNAPII is escorted to the promoter site by TFIIF, forming a transcription complex; (iv) orderly addition of TFIIE, TFIIH and .TFIIJ help in the initiation. The carboxy-terminal domain (CTD) of RNAP II gets phosphorylated with the help of TFIIH (Fig. 13 and 14)which is an essential requirement for transcription by RNAII after initiation of transcription, RNAPII leaves promoter but TFIID remains bound to start another cycle of transcription initiation. The RNAPII after initiation keeps on synthesizing RNA, thus elongating the RNA chain. Atleast two elongation factors are used. (i)TFIIF accelerates RNA

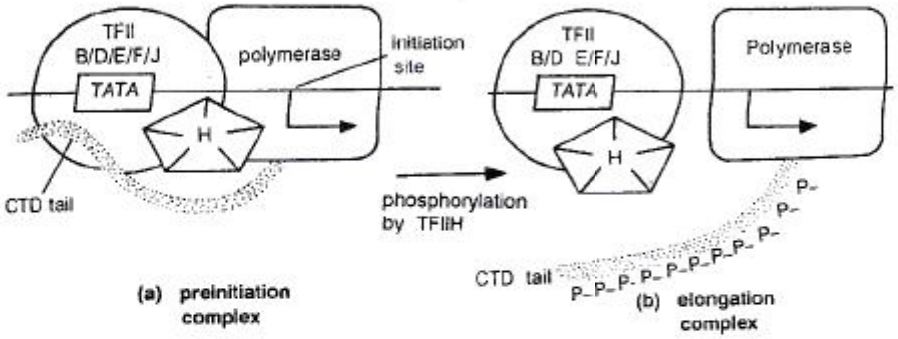


Fig. 4.16 : Picture showing the action of transcription factor TFIIH.

chain growth uniformly and (ii) TFIIS removes obstructions in the path of elongation. Termination of RNA chain synthesis is brought about by termination site similar to those in prokaryotes.

4.5 Plant Promoters

In Eukaryotes there are three RNA polymerases I or RNAPI for synthesis of pre-rRNA; RNA polymerase II or RNAPII for synthesis of pre-mRNA or hnRNA heterogenous nuclear RNA) and several snRNAs, and RNA polymerase III or RNAPIII for synthesis of 5S RNA, tRNAs, U6 snRNA, 7SL, EBER2 and VA (7S RNA is present in a small RNA protein complex called SRP; EBER RNA, is specified by Epstein-Barr virus; VA RNA is specified by adenovirus. Different promoter sequences have been identified for different RNA polymerases.

Promoters for RNA Polymerase

I. Promoters for RNA polymerase I have at least two elements: (i) a GC-rich upstream (-180 to -107) control element (UCE) and (ii) a core, region that overlaps the transcription start site (-45 to +20)

Promoters, enhancers and silencers for RNA polymerase II. The promoters sequences for eukaryotic RNAPII include core promoters for basal transcription and other regulatory sequences for activator-dependent transcription. A core promoter is generally a sequence, ~40 base pairs long, that is sufficient to direct accurate transcription by RNAPII with the help of basal transcription factors, which is also described as general transcription factors (GTFs). There are several known core promoter motifs, which include the following : (i) TATA box (an A/T rich sequence) is located ~ 30 bp upstream

NOTES

NOTES

of transcription start site, at which TATA binding protein (TBP) subunit of TFIID binds; (ii) BRE(TFIIB recognition element) is located immediately upstream of the TATA box of some TATA-containing promoters an increases the affinity of TFIIB for core promoter;(iii) Inr (initiator) is a conserved sequence encompassing the start site, and functions to direct accurate initiation of transcription either by itself or in conjunction with a TATA or DPE-motif, and (iv) DFE(downstream promoter element), which is a conserved sequence (as common as TATA box), located ~30 bp downstream of the start site. A typical DFE-containing promoter has Inr and DPF motifs and lacks a TATA box. TFIID binds to lnr and DPE motifs through its TAF (TBP-associated factor) subunits. Since loss of TATA box can be compensated by DPE the eukaryotic promoter can be classified into **TATA-driven promoters** and **DPE-driven promoters**. In Drosophila, DPE-driven promoters are used by DSTF (DPF-specify transcription factor), which is now described as dNC2(in view of its homology to the human transcription repressor NC2) and consists of two subunits (dNC2$\alpha$and dNC2β; prefix 'd' stands for Drosophila). dNC2 is a bifunctional transcription factor, which represses TATA-driven promoter and activates DPF-driven promoters(Fig. ----). TATA box is compared with Pribnow box of prokaryotes and further upstream of this box are present GC box(-60 or -100 bp) and CAAT box (-80 bp) as shown in figure 16.

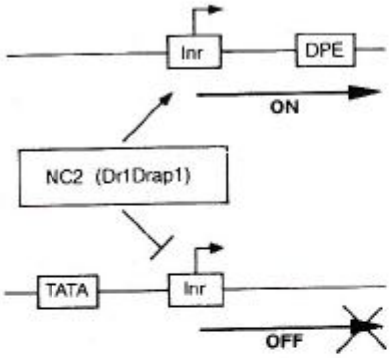


Fig. 4.17 : Two types of promoters (inr-DPE and TATA-Inr) in eukaryotes and the Contrasting effects of transription factor NC2 on theses two promoters.

Eukaryotic promoters often also possess elements located 100 or 200 base pairs upstream, which interact with proteins other than RNA polymerase. These

elements, called enhancers or upstream activation sites (UAS) can be moved several hundred or even thousands of base pairs upstream without alteration of their activities. These are enhancers may cause upto a 200 fold increase in transcription rate of an affected gene, hence the name. These are other regulatory elements known as silencers, which repress gene expression. Silencers, like enhancers, can function at great distance from genes they repress. They are sites for binding of proteins just like enhancers. The location of different promoters and an enhancer is shown in Fig. .....

NOTES

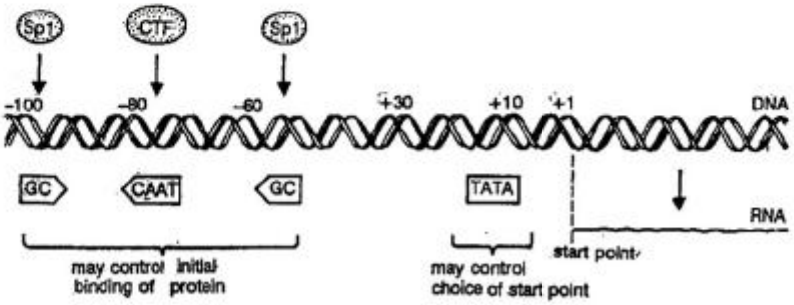


Fig. 4.18 : A DNA segment in a Eukaryote showing promoters sites (TATA box and CAAT bo) and the enhancer site

Promoters for RNA **polymerase** III. Three classes of promoters have been described for RNA polymerase III: (i) tRNA gene promoters have two elements, each 10 bp long, spaced 30-120 bp apart, and located downstream of the transcription start site within the transcribed region; theses are called box A and box B. (ii) U6 snRNA gene promoters differ in humans and yeast. In humans U6 snRNA gene is devoid of intragenic promoter elements. Instead, they have promoter elements likes those for RNAPII, (TATA motif, PSE and OCT) In yeast, on the other hand, intragenic elements, box A and box B are found along with TATA box (but no PSE). (iii) 5S RNA gene promoters have intragenic control region called box C, which lies well within the transcription unit 50 bp downstream rather than upstream from the start point however, the location of downstream promoters for RNA polymerase III may vary in different genes at different positons in 5S and tRNA genes.

.

NOTES

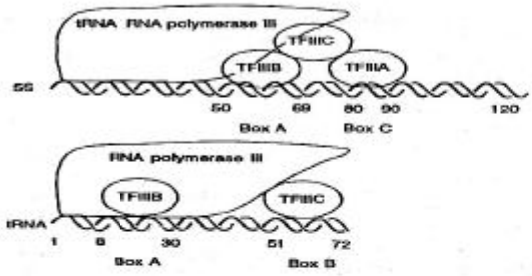


Fig. 4.19 : Binding of TFIIIB to downstream promoters (box A) located

It is interesting to know - how RNA polymerase III attached to +55 to+80 bp sequence can start transcription from the start point. It is believed that the enzyme is big enough to occupy start point and the promoter region simultaneously and thus may start transcribing at the start point without difficulty.

4.6 Transcription Factors

Transcription factors are one of the groups of protein that read and interpret the genetic blue print in the DNA. They bind to the DNA and help to initiate a program of increased or decreased gene transcription. They are also vital for many important cellular processes.

In molecular biology and genetics a transcription factor (some times called a sequence DNA –binding factor) is a protein that binds to specific DNA sequences there by controlling the rate of transcription by genetic information from DNA to messenger RNA. Transcription Factor performs this function alone or with other proteins in a complex, by promoting (as an activator) or blocking (as a repressor) the requirement of RNA polymerase (the enzyme that performs the transcription of genetic information from DNA to RNA to specific genes).

A defining feature of Transcription factor is that they contain one or more DNA binding domains (DBDs) which attach to specific sequences of DNA adjacent to the genes that they regulate. Additional proteins such as co-activators, chromatin remodelers, histone acetylases, deacetylases, Kinases and methylases while also playing crucial roles in gene regulation, lack DNA binding domains and there fore are not classified as transcriptional factors.

Transcription factor are essential for the regulation of gene expression and found in all living organisms. The number of transcription factor found in organism increases with genome and larger genomes have more transcription factor per gene. There are about 2600 proteins in the human genome that contain DNA binding domains, and they are presumed to function as transcription factors about 10% of genes in the genome code for transcriptional factors so this family is the single largest family of human proteins. Genes are often flanked by several binding sites for distinct transcription factors and each gene expression requires the co-operative action of several different transcriptional factors.

NOTES

Hence use of subset of about 2000 human transcription factor in combination accounts for the unique regulation of each gene in the human genome during development.

4.7 Messenger – RNA (mRNA)

It is also known as complementary RNA or template RNA or unstable RNA or information RNA. mRNA acts as a template RNA that carries information for protein synthesis from nuclear DNA to ribosomes in the cytoplasm. Because of this feature Jacob and Monad (1961) called it messenger RNA.

Life span of messenger RNA (mRNA) is very short. Its amount is very less 5 to 10% of total cellular RNA. It is synthesized in nuclear region and then after it gets migrated to cytoplasm where ribosomes are active in protein synthesis (in Eukaryotes). There it acts as a template for protein synthesis.

mRNA is single stranded. The strand shows complementary base sequences like the DNA template from which it is transcribed. It only differs from DNA template in presence of base uracil (U) in place of thyamine (T).

In Eukaryotes, mRNA are transcribed on single gene so they are called Mono cistronic where as in prokaryotes, it gets transcribed from several adjacent gene so it is called polycistronic mRNA

Molecular weight and size at mRNA varies so they are heterogenous in nature. Molecular weight of mRNA ranges from 5 to 20 lakhs daltons.

Life span of mRNA is very short after performing its function (translation) it degenerates. In prokaryotes, for example in (bacteria), life span of mRNA is

only 2 minutes where as in Eukaryotes, mRNA exhibits life span of 1 to 4 hours or upto some days.

Structure of mRNA

NOTES

mRNA has following structural features –

1. **Cap** : This part of mRNA influences the rate of protein synthesis. At 5^1 end of mRNA strand there occurs a cap which is blocked by any of the 4 methylated nucleotides. In case of absence of cap, mRNA shows poor binding of ribosomes.
2. **Non coding region-I (NC-I)** : After cap regions, non coding region-I comes. This region is at 10-100 nucleotides which is rich in A & U bases and does not translate protein.
3. **Initiation codon :** After non coding region-I 'AUG' codon is present which initiates the polypeptide chain.
4. **The coding region :** This coding region is consist of about 1500 nucleotides. It translates a particulars protein molecule.
5. **Termination coding :** It is present after coding region. This codon terminates protein synthesis on mRNA. In Eukaryotes termination codons of mRNA are UAG, UGA and UAA.
6. **Non-coding region-II (NC-II) :** This region is consist of 5O-150 nucleotides. This region contains AAU sequence and this region does not translate protein.
7. **Poly-A sequence :** Poly-A sequence is consist of 200-250 nucleotides. Poly A sequence is added in the nucleus before mRNA is released in to the cytoplasm. At the 3^1 end of mRNA, polyadenylate sequence is present.

4.8 Biosynthesis of mRNA

For biosynthesis of mRNA following steps are involved –

1. For biosynthesis of mRNA, one strand of double stranded DNA acts as a template.
2. The defined units of genetic informations are copied in to mRNA molecules.

3. For this, the enzyme DNA dependent RNA polymerase binds it self at the initiation site at structural gene or cistron DNA catalyses the synthesis of mRNA chain.
4. The ribo nucleoside triphosphates act as substrate.
5. The polymerisation takes place in a 5^1 to 3^1 direction. This process of mRNA synthesis is called "**transcription**".

NOTES

In Eukaryotic cells, synthesis of mRNA takes place in the form of heterogenous RNA(Hn-RNA) in the nucleus, at the 3^1 end of Hn RNA about 200 poly A^1 nucleotides are added. After this 5^1 end at Hererogenous RNA degenerates and releases poly A^+ mRNA molecule which diffuses out in to cytoplasm.

Studies show that mRNA as poly-A sequences at its 3^1 and 7 methyl guanosine at its 5^1 end.

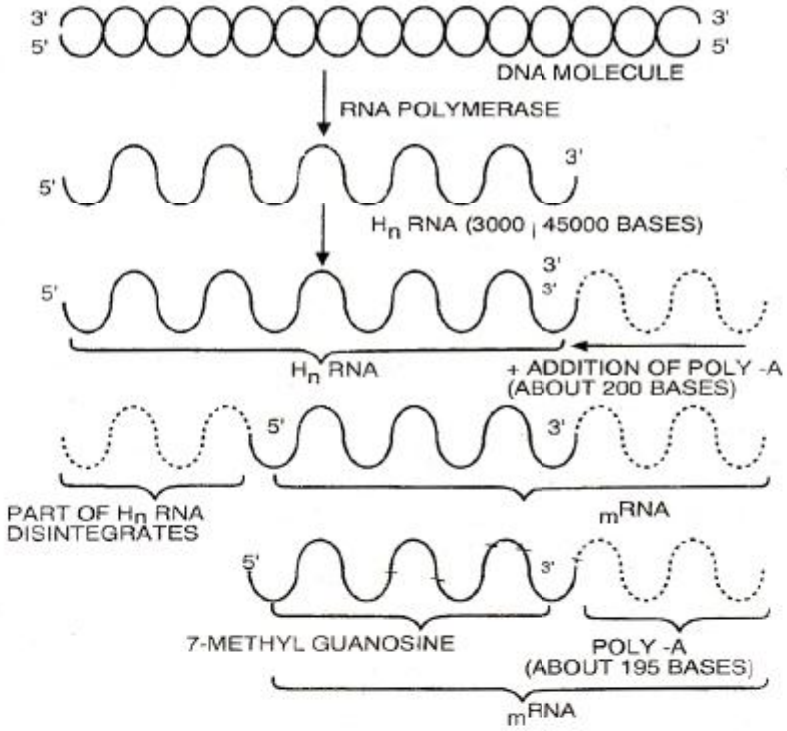


Fig. 4.20 : Gene Expression

Informosomes

In certain Eukaryotic cells, the mRNA does not enter the cytoplasm as a naked RNA strand but it remains ensheathed by certain protein and becomes more

stable. Such mRNA + protein complete have been called informosomes by spirin.

Informosomes is inactive form of mRNA. Proteins of informosomes protect mRNA from degrading action of ribonuclease enzyme.

NOTES

Informosomes are more stable. They can survive in the cytoplasm upto period of 1-1 week.

4.9 Nucleolus

Nucleolus (Plasmasomes - True Nucleolus)

Nucleolus is a dense, spherical colloidal body filled with acidophilic material which remains embedded in the nuclear matrix. It can be stained with acidic dyes (e.g. Pyronine stain) and disappears during mitosis. It was first described by **Fontona** in 1781. It has no membrane of its own but it is highly duse body as, compared to its surrounding nucleoplasm and distinctly visible.

Size: The size of nucleolus is related with the synthetic activity of the cell. The cells which are active in synthesis of proteins and other substances possess large nucleoli whereas cells with little or no. synthetic activity contain smaller or no nucleoli. The secretory cells, neurons and oocytes possess large nucleoli whereas muscle cells and sperm cells do not contain nucleoli.

Number: The number of nucleoli depends, upon the number of chromosomes in a cell. The haploid cells of both plants & animals have only one nucleous but in man, there are two pairs of nucleoli in each diploid nucleus.

Position : In the nucleus ,the heterochromatic regions of certain chromosomes are found to be associated with the nucleolus which are called as nucleolar organizing regions of the chromosomes. These regions are mainly concerned with the constitution of nucleolus, although other chromosomes also contribute in the formation of nucleolus. The nucleolar organizing regions contain genes that code for 18S, 28S, 5.8S and other RNAs which are synthesized in the nucleolus. The 5S RNA is synthesized on the chromosomes outside the nucleolus. The ribosomal proteins are synthesized in the cytoplasm. All these components migrate to the nucleolus and assemble to form ribosomes which are transported to the cytoplasm.

Chemical composition of nucleolus

Generally the chemical composition of nucleolus includes about 3-20% RNA of the total RNA content of the nucleus. A high concentration of phosphoprotein has been observed but histone poteins are absent. Similarly high concentrations of orthophosphates which are the precursor of the RNA phosphorous have been reported. Besides these different types of enzymes, viz., acid phosphatase, nucleoside phosphorylase, NAD synthesizing enzymes etc. are also found in the nucleolus which participate in the synthesis of coenzymes, nucleotides and rRNA. Some nucleoli also contain methylase enzyme which transfers methyl group to the base of RNA. Although DNA is absent in nucleolus but it is surrounded by DNA ring which is a heterochromatin part of the chromosome associated with it.

NOTES

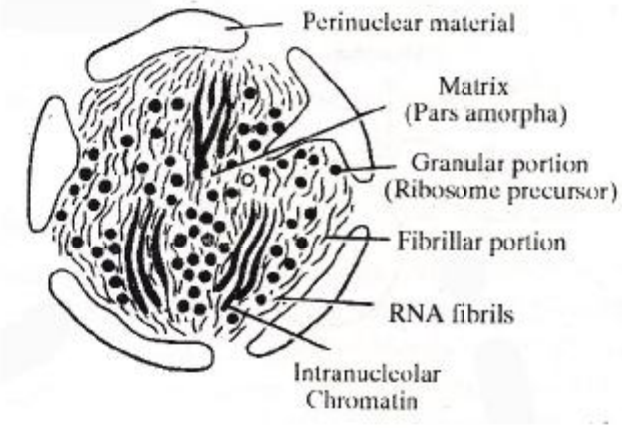


Fig. 4.21 : Ultra-structure of Nucleolus

Electron microscopic examination of nucleolus reveals following four structural components:

(1) Granular portion : The peripheral part of the nucleolus is composed of dense granules with a diameter averaging 150-200 $\overset{0}{A}$, which are more or less numerous along and between the strands of the fibrillar network. These granules are made up of ribonucleoproteins and they are precursors of ribosomes. .

(2) Fibrillar portion: It is present inner to the granular zone and is made up of many fibrils of ribonucleoproteins, which are about 50-80 $\overset{0}{A}$ in diameter and 300-400 $\overset{0}{A}$ in length. This region is also called as **nucleolonema**.

(3) Amorphous zone: This region is of low electron density, found in only some nucleoli. It contains certain proteins.

(4) Nucleolar-associated chromatin: These are 100 $\overset{0}{A}$ thick fibres situated around the nucleolus which have intra nucleolar components also. They contain DNA.

NOTES

Nucleolar cycle

In the nucleus, the nucleolus as an organized body is not always a permanent figure but it disappears at the beginning of cell division in prophase stage and reappears at the end of cell division in telophase stage. This is called as **nucleolar cycle** which is controlled by a chromosome called nucleolar organizer chromosome. The nucleolus is organized at nucleolar zone, a specific region of this chromosome. This chromosome is also marked by secondary constriction. The nucleolus is found to be associated with the nucleolar zone which carries genes for 18S, 28S and 5.8S rRNA synthesis.

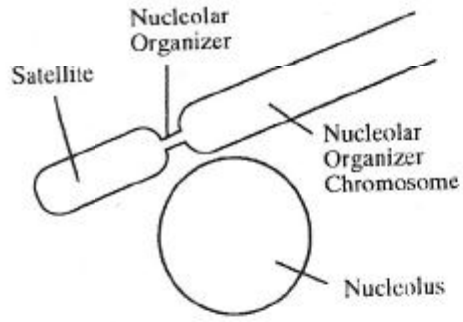


Fig. 4.22 : Nucleolar organizer chromosome

1. Nucleoli are the site of ribosomes synthesis as they store proteins and rRNA.
2. In these nuclear components only the 45S rRNA breaks into two subunits (18S rRNA and 28S rRNA) and their methylation occurs.
3. Histone proteins are also synthesized in the nucleolus.
4. Nucleolus collects freshly synthesized mRNA for the cellular protein synthesis.
5. It provides sufficient energy for nuclear activity.

6. It plays an important role in cell division. If it is damaged somehow, the mitosis may cease for ever.

4.10 Ribosomal-RNA (rRNA)

NOTES

According to **Kurland (1960)** the RNA which is extracted from ribosomes is called **ribosomal-RNA**. As compared to other types of RNA, its amount is very high up to 80% of total cellular-RNA. The rRNA molecules twist upon themselves and generate a surface that associate with particular proteins to form the ribosomes. Thus ribosomes are the ribonucleo-protein complex in which 60% of r-RNA and 40% proteins are present. They are the sites of protein synthesis.

It is the most stable RNA. The molecule of rRNA consists of single polynucleotide strand twisted upon itself in some regions called **helicals**. In the helical region most of the base pairs are complementary and are joined with hydrogen bonds. In the unfolded single strand regions the bases have no complements. Its molecular weight, varies from 0.5 to 2.0 x 10^6 daltons. rRNA differs from other types of RNAs in size and base contents. They are much large & molecules and are rich in guanine and cytosine contents. Besides, these some methylated bases may also be present. It is produced by nucleolar organizer and stored in nucleolus. There it combines with some special proteins and diffuses out into the cytoplasm.

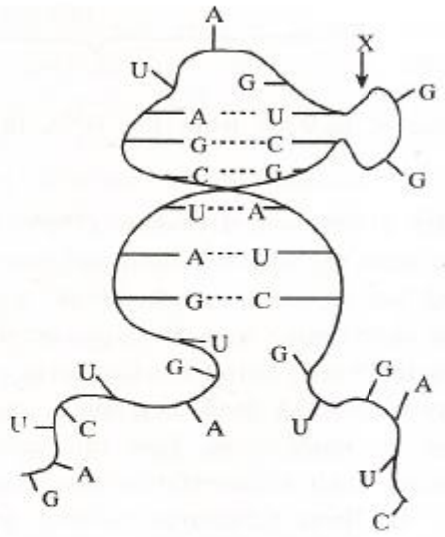


Fig. 4.23 : Molecular Structure of RNA

NOTES

From the large subunit of Eukaryotic ribosomes, three types of rRNAs have been identified, viz., 28S, 5.8S and 5S rRNAs and from small ribosomal subunit a single rRNA molecule, i.e., 18S rRNA is identified. From Prokaryotic ribosomes only two kinds of rRNAs, 23S and 5S can be extracted from the large subunit of ribosomes. The small subunit contains a single type 16S rRNA molecule. Thus Prokaryotic rRNA molecules' are smaller than that of the Eukaryotic rRNAs.

Biosynthesis of rRNA

In Eukaryotes rRNAs are transcribed on rDNA which is located in the core of nucleolar organizer region of a particular chromosome to which nucleolus is attached.

45S RNA first it synthesizes which is a large initial precursor molecule for 58S rRNA and 18S rRNA. After entering nucleolus 45S RNA is combined with methyl group and proteins to form a complex. This methylated protein complex is cleaved into 32S and 18S pieces. Now the protein composition changes and change in the configuration of two particles takes place.

After this 18S pieces fuse with basic proteins and form 40S submit of ribosomes which diffuses in to cytoplasm. After that 32S pieces also cleaved into 28S rRNA particles and combines with proteins. It gets incorporated with 5S rRNA particles of non nuclear origin to form 60S subunit of ribosomes which also diffuses into cytoplasm.

Both sub units of ribosomes combine with each other at very low concentration of Mg^{++} ions in the cytoplasm and form 80S type of ribosomes.

Function of rRNA

In the molecule of rRNA, the unpaired bases may bind mRNA and tRNA to ribosomes by Mg^{++} linkages between phosphate groups on the two molecules.

4.11 Summary

Nucleus is a important cell organelle which controls all metabolic activity of cells. Nucleus controls the production of materials that enters the cytoplasm and participates in the control of growth and morphology. Nuclear pores have important role in transfer of ribo nucleoprotein particles in to cytoplasm.

It preserves the genetic material and controls the cell reproduction.

DNA has the potentiality to carry out all kinds of biological activities of cells. It is the genetic material which is responsible for transferring characters from one generation to another.

NOTES

4.12 Glossary

- **Transcription :** The process of making RNA from a DNA template by RNA polymerase.
- **Transcription Factor :** A substance such as protein that contributes to the cause of a specific bio chemical reaction.
- **Upregulation :** Increase the rate of gene transcription.
- **Down regulation :** Decrease the rate of gene transcription .
- **Co-activator :** A protein that works with transcription factor to increase the rate of gene transcription.
- **Co-repressor:** A protein that works with transcription factor to decrease the rate of gene transcription.
- **Transcription regulation :** Controlling the rate of gene transcription by helping or hindering RNA polymerase binding to DNA.
- **Co-repressor :** A protein that works with transcription factors to decrease the rate of gene transcription
- **Co-activator :** A protein that works with transcription factors to increase the rate of gene transcription.
- **DNA replication :** The process by which DNA molecule gives rise to two similar DNA molecules is called replication of DNA.

NOTES

- **Nucleosome :** Repeating units of chromatin fibres which is flexibly joined chain resembling beads on a string is nucleosome.

- **Chromatin Fibre :** A complex of DNA(35%) protein (60%) and DNA (5%) in the form of a supercoiled and folded fibre is known as chromatin fibre.

- **Nucleosome :** A fundamental organizational packing unit of Eukaryotic nuclear chromatid is **nucleosome**.

- **Histone :** Histones are small basic proteins containing 20 to 30% basic amino acids lysine and arginine which are positively charged.The positive charge enable histone moleculer to bind to DNA.

- **Nuclear pore :** Karyotheca is perforated by pores at intervals is called **nucleopore**.

- **Nucleoplasm :** It is also called karyolymph The interphasic nucleus contains a transparent semi solid granular and homogenous ground substance or matrix called **nuclear sap**.

4.13 Self-Learning Excercise

Section A (Very Short Answer Type Questions)

1. Name those Scientists who were involved in discovery of DNA as chemical basis of heredity -------?
2. The most important component of the nucleolus is .....?
3. Who discovered nucleolus .....?
4. Who first discovered mRNA from Bacteria & viruses?
5. Name the scientist who has given nucleoplastic.....?
6. Histone protein are synthesized by ..... ?

7. Transcription involves .....?

8. The role of nucleus in regulating the morphology of the plant was discovered in .

NOTES

Section B (Short Answer Type Questions)

1. Write short note on nucleolus?
2. Write two functions of DNA?
3. What are nucleosome or PS particles?
4. Write the difference between DNA and RNA?
5. What do you understand by transcription?
6. Write the role of histone protein in the nucleus?

Section C (Long Answer Type Questions)

1. "DNA is the genetic material" explain it with experimental evidences.
2. Write notes on -
    a. Double helix
    b. Types of RNA
3. Describe the mechanism of DNA replication what enzymes are necessary during this process.
4. Describe the ultra structure of nucleus with suitable diagram.

Answer key of Section – A

1. Avery, Macleod and Mccarty
2. RNA
3. Fontana

4. Huxley

5. Hertwig (1906)

NOTES

6. Nucleus

7. synthesis of RNA over DNA

8. Acetabularia

4.14 References

- "Cell biology genetics and plant breeding" by Trivedi & Sharma RBD (2006)
- "Cell biology & Genetics" by Srivastava, H.N. Vol. II PP 1-184 Pradeep Publications, Jalandhar.
- "Cytology genetics biotechnology and biostatistics" by Gupta,P.K. Rastogi publications. Merrut .

Unit - 5

Cellular Organelles

NOTES

Structure of the Unit

5.0 Objectives

After going through this unit you all will be knowing about the –

- Cell organelles and their structure and functions.
- Protein synthesis and sorting

5.1 Introduction

By going through this unit student will learn about the various cell organelles or microbodies of cell.

These microbodies includes Spherosomes, Peroxisomes Glyoxysomes, Ribosomes, Lysosomes and endoplasmic veticulum.

They are very important organelle of the cell and are responsible for various activities of cell. Different microbodies perform different functions. Few of them act as phagosomes. Few of them like vacuoles are responsible for mechanical support, storage and osmoregulation. Ribosomes are the site of protein synthesis.

NOTES

5.2 Other Cellular organelle

5.2.1 Microbodies of the Cell

Recent studies of plant cells with the help of electron microscope have confirmed the presence of a large number of tiny small bags which are membrane bound are called microbodies. Some microbodies are Spherosomes, Peroxisomes Glyoxysomes, Vacuoles.

Spherosomes : Higher plant cells do not contain well defined lysosomes as found in animal cells but in these cells unit membrane bounded cytoplasmic bodies of lipids (98%) and hydrolytic enzymes and digestive enzymes are found which are known as **spherosomes**

History : Spherosomes were first described by B.perner(1953) and Muhlethaler(1955). They named as **spherosomes**. Spherosomes were isolated from the seedling cells of maize by Samadini (1967).

Structure : They are identical with lysosomes of animal cells. They are small, spherical, single membrane bounded vacuole like microbodies which are 0.5μ - 2.5μ in diameter.

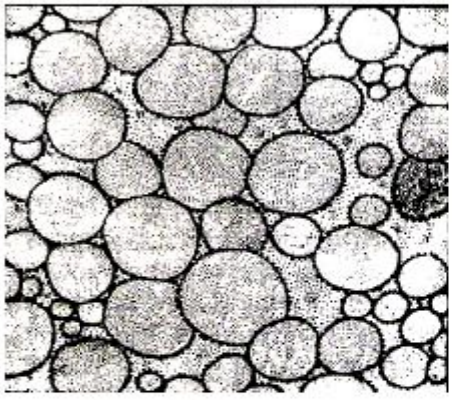


Fig. 5.1 : Spherosomes

The outer most boundary of these organelles is single membrane made up of Lipoprotein. They are filled with digestive enzymes, lipids and proteins.

Spherical vesical which orginate from the endoplasmic reticulum by pinching off a very small vesicle. These vesicles enlarge to form a spherosome. They mainly contain enzymes like phosphatase, esterase, ribonuclease, hydrolase etc. Present enzymes help in synthesis of fats.

NOTES

Function : Main function of spherosomes is digestion of ingested materials in the form of **phagosomes**.

Spherosomes are also involved in lipid melabolism i.e. synthesis, storage and transport of lipids.

Protein bodies / Aleurone grains : They are unit membrane bounded storage granules which are also known as protein bodies. They are found in the cells of cotyledons and endosperm. These bodies store proteins as globulins and phosphate in the form of phytin.

Main function of aleurone grains is the storage of proteins in the mature seeds and their release as enzyme at the time of seed germination for the growth and development of seedlings in to plant.

Aleurone grains also originate from ER. They are composed of two parts –

1. **Crystalloid** : Composed of proteins and amides.
2. **Globoid** : Composed of proteins & minerals

5.2.2 Peroxisomes

Peroxisomes are small spherical membrane bound microbodies which are found in the cells of green leaves. Theses microbodies are rich in enzymes like Peroxidase, catalase and specially the enzyme urate oxidase. The diameter of these bodies is 0.5 µm to 1.0 µm. These microbodies contain fine granular substances which may condense in the center to form nucleoid.

History : First of all they were isolated by Tolbert and his coworkers (1968). The term peroxisomes was coined by de-Duve(1969)

Occurence : they are found in various animal and plants cell. Besides they have been found in Protozoa. In plants they are found in the cells of Fungi, brown algae, bryophyta, ferns and higher plants. In higher plants they are present in the various tissues like cells of leaves, coleoptile, hypocotyle, and ripened fruits.

In animals, they are found in liver and kidney cells of rodents and mammal.They are present near the ER, mitochondria & chloroplast.

NOTES

Structure : These granules are ovoid in shape and bounded by a single lipo - protinaceaus membrane. They contain fine granular substance which may condense in center to form nucleoid. Nucleoid may or may not be present in all peroxisomes. The number of peroxisomes per cell varied between 70 to 100.

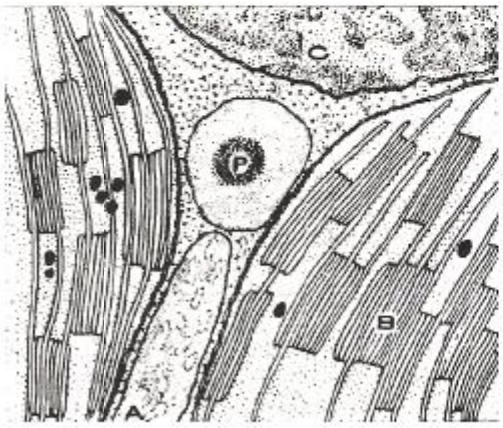


Fig. 5.2 : Peroxisomes

Origin : Peroxisomes originate from the out growths of the membranes of ER in which granular substances and enzyme are collected. These enzymes are synthesized in the ribosomes which are attached to granular ER. Life period of peroxisomes is very short. They get destroyed by autolysis only after 4or5 days.

Functions : Main function of peroxisomes is two biochemical reactions:

1. **Glycolate Cycle** or **Photorespiration** : In plants peroxisomes of green leaves cells contain additional enzyme glycolic acid oxidase which carries out photorespiration.
2. **Hydrogen peroxide melabolism**

5.2.3 Glyoxysomes

Glyoxysomes are membrane bound spherical particles which were first isolated from caster bean endosperm. They measure 0.5 to 1.0 µm in diameter. These microbodies are common in micro organisms and in higher plants. They are usually present in germinating fatty seeds where fat is being converted in to carbohydrates.

Glyoxysomes contain and oxidase enzyme catalase and oxidase enzyme and enzymes of glyoxylate cycle e.g. glutamate, glyoxalate, serina glyoxylate, malate dehydrogenase and certain co-enzymes like FAD, NAD and NADP.

During glyoxylate cycle conversion of lipid in to carbohydrate takes place and H2O2 is produced, which is decomposed in to water and oxygen by the enzyme catalase. Glyoxysomes also originate from ER. They are consist of an amorphous protein matrix bounded by a membrane.

NOTES

5.2.4 Lomasomes

They are membranous, vesicular structure usually present between the cell wall and the plasma membrane of plant cells. They appear to be originated from Golgi bodies. Their function is not known but they help in cell wall elaboration.

5.2.5 Vacuoles

The intracellular spaces which are enclosed by membranes are known as **vacuoles**. Porter(1945) first observed cytoplasmic vacuoler system.

In Prokaryotes generally vacuoles are absent. In RBC's, eggs and undifferentiated embyonic cells also vacuoles are absent.

In plant cells most part of cell lumen is occupied with a large vacuole due to which cell cytoplasm appears towards the periphery of the cell in the form of primordial utricle.

The membrane enclosing the vacuole is called **tonoplast** and the fluid filled in the vacuoles is called cell-sap.

In vacuoles the cell sap contains different types of food material, minerals, salts and acids in soluble form. This sap also contains anthocyanin pigment which gives colour to leaves, flowers and fruits.

Function

The functions of vacuoles are following –

NOTES

i. Mechanical support : Vacuoles divide the fluid content of cell in to compartments thus provide mechanical support to the cytoplasm.

ii. Storage : The primary function of vacuoles is the storage of synthetic products. In animal cell, vacuole store glycogen, enzymatic and structural proteins etc.

In plant cell, they contain water, phenol, flavanols, alkaloids, fats, sugar, anthocyanins etc. Some times they also contain waste products of various metabolic reactions.

iii. Osmo regulation : They regulate the osmotic pressure of cell in unicelluar organisms. In Blue green algae, gas vacuoles are found which provide buoyancy to cells.

5.2.6 Ribosome

Ribosomes were reported in plant cells by **Robinson** and **Brown** in 1953 for the first time in bean root cells. In animal cell, they were reported first time by **George Palade** in 1955. The name "**Ribosome**" was proposed by Robert (1958). **Tissiers** and **Watson** isolated these ribosomes from the cell of the bacterium *E.Coli.*

Occurence

Ribosomes are present in almost all living cells, Prokaryotic as well as Eukaryotic. In Eukaryotic cells they are seen attached on endoplasmic reticulum(ER) or freely distributed in the cytoplasmic matrix, chloroplast and mitochondria. Ribosomes are also found attached with outer surface of outer nuclear membrane.

In Prokaryotes, they may occur in free form through out the cytoplasm and they are called monosomes and some times ribosomes are joined with MRNA forming linear or helical groups called Polyribosome and Polysome.

Ribosomes are absent in Viruses, mature red blood cells and sperm cells.

Shape, size and number

The shape of the ribosomes is small dense and round bodies. They are minute particles of the average diameter of 150 to 250A^0. They are visible only in electrone microscope.

NOTES

On the basis of shape, size and sedimentation coefficient, ribosomes are of two types :

a. 70S Ribosomes :

b. 80S Ribosomes

Difference between two type of ribosomes

70S ribosomes	80S ribosomes
Small size	Slightly larger in size than 70S types of ribosomes
They are found in bacterial cells and in the mitochondria and chloroplasts of Eukaryotic cell	They are found in all Eukaryotic cells
Their molecular weight is $2.7X10^6$ Daltons	Their molecular weight is $4X10^6$ Daltons

Structure of ribosomes

1. All type of ribosomes exhibit uniformity in their structure and composition.
2. Each ribosome is spherical, porous, hydrated and composed of two sub-units.
3. One unit is larger in size and dome shaped where as other sub units is smaller in size which is attached with larger subunit and forms a cap like structure.
4. Both sub units are separable and bear sub protruberances and ridges.

NOTES

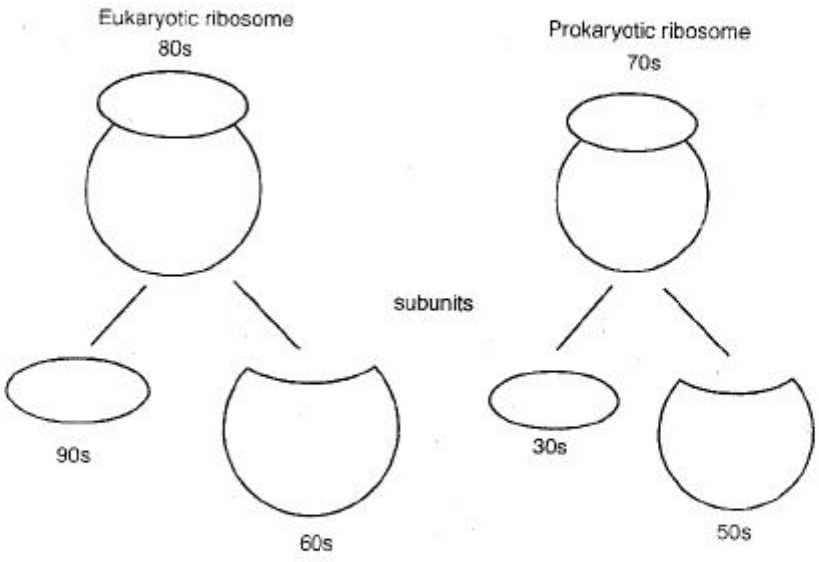


Fig. 5.3 : Subunits of Eukaryotic & Prokaryotic ribosomes

The ribosomes of Prokaryotic cells are composed of 50S and 30S sub units, where as ribosomes of Eukaryotic cells are made up of 60S and 40S sub units. Recent studies confirms that association and dissociation of ribosomes sub units depends on Mg^{++}ion concentration. Each sub unit has specific parts.

The small subunit has a head, a base and a platform where as larger sub unit has a ridge, a central protruberance and a stalk.

The outer boundary of each ribosomes is an envelope of protein which surrounds rRNA. The ribosomes consist of 60% rRNA and 40% basic proteins.

Function of Ribosomes

1. Ribosomes are important cell organelle which remains involved in protein synthesis.
2. They are called "Protein manufacturing Factories" because they are the sites of protein synthesis. In Eukaryotes, they get assembled on the Endoplasmic reticulum and provide platform for protein synthesis.

   In Prokaryotes, they get assembled in Polysome before the synthesis of protein take place.

3. Ribosomes play important role in protecting the newly formed polypeptides from degradation by cytoplasmic enzymes in a tunnel of their larger subunit before releasing them in to lumen of endoplasmic reticulum.

NOTES

4. At the time of protein synthesis, several ribosomes become attached to m RNA with the help of smaller unit. They are called **polyribosomes**.

   Larger sub unit contains peptidyl transperase enzyme which catalyses the synthesis of peptide bond between amino acids during protein synthesis.

5.2.7 Lysosomes

Lysosomes are single membrane bounded small spherical vacuolar bodies found in the cytoplasm of most of the animal cell and few plant cells

Discovery : Lysosomes were discovered by C-deDove-in 1949 and named them lysosomes in 1955.

Lysosomes is a Greek word, which means lysis=dissolution; soma=body; since then they contain lytic enzyme so they are named as lysosomes.

Structure : Lysosomes are single membrane bounded vacuolar bodies which are filled with dense material containing 40 different hydroytic enzymes especially tissue dissolving enzymes acid phosphatase, acid ribonuclease, acid deoxy-ribonuclease, cathepsin, glycosidases etc so these organelles are responsible for autolysis of dead cell parts, cells and tissues.

Lysosomes are made up of two parts-

1. **Limiting Membrane :** Lysosomes are bounded with a single unit membrane which is lipoprotinaceous. This membrane retains enzymes.
2. **Inner Dense Mass :** Inner mass of lysosome contains different type of contents. In some lysosomes, this mass is solid and dense where as in others outer zone is dense and core is less dense and they have cavities or vacuoles with in the granular material structure of Lysosomes.

   The major types of lysosomes are as follows –

NOTES

i. **Primary Lysosomes :** They are formed from Golgi complexes. They are small sac like bodies formed by budding of Golgi Complexes. They are also called storage granules. Their enzymatic contents are synthesized from ribosomes of rough endoplasmic reticulum and accumulated in ER. From ER they are transported to Golgi region. There acid phosphotase reaction takes place and lysosomes are produced called as **Primary Lysosomes**.

ii. **Secondary Lysosomes :** They are also called heterophagosomes or digestive vacuoles. They are formed due to the fusion of phagosomes / pinosomes and primary Lysosomes.

Pinosomes are the extra celluler substances containing bodies enclosed by plasma membrane. They are involved in the process of digestion. They are of two types.

1. **Heterophagosomes :** If the substances for digestion come from out side the cell by endocytosis and the digested products are diffused out in to cytosol through lysosomal membrane, then the lysosome is called heterophagosome.
2. **Autophagosome** : They originate from ER and digest the organell of the cell in which they occur. **Hence they are known as suicide bags of the cell.** \

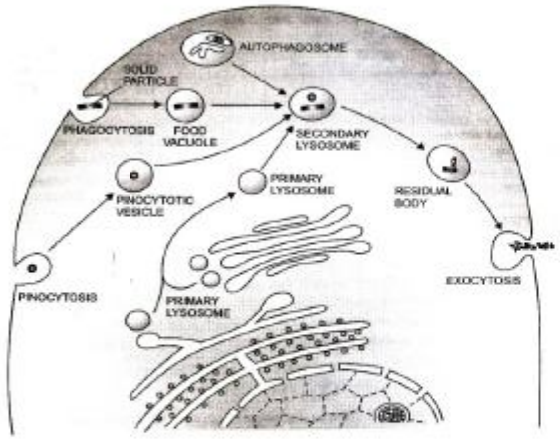


Fig. 5.4 : Formation of Lysosomes & Intra Cellular Digestion

Chemical nature of lysosomes

In lysosomes, approximately 40 hydrolytic enzymes are present. But oxidative enzymes which are found in mitochondria are totally absent.

NOTES

Following are few lysosomal enzymes :

1.	Phosphotases :	Phospho protein phosphatase, acid pyrophosphatase and acid phosphatase.
2.	Proteases :	Cathepsins, collaginase, peptidase
3.	Nucleases :	Acid ribonuclease, acid deoxy ribonuclease
4.	Glycosidases :	α and β Glucosidases, β Galactosidases, α Fucosidases, Lysozyme
5.	Sulphatases :	Aryl sulphatase
6.	Lipases :	Phospholipase, Esterase, Triglyceride lipase.

Origin of Lysosomes

Depending upon functions and occurences, lysosomes have different origin.

1. **Extra celluler origin** : In this type of origin, primary lysosomes originate due to invagination of cell wall in to cytoplasm by forming pinosomes and phagosomes which are absorbed by the cell. These vacuoles becomes the part of cytoplasm but they are with no enzymes.

   When they come in contact with Golgi complex, enzymes enter in them and they become primary lysosomes.

ii. **Origin of Endoplasmic reticulum** : According to Novikoff(1965) the lysosomes originate directly from the enlargements of granular ER containing protein granules which are pinched off as lysosomes.

iii. **Origin from Golgi complex** : It has been proved that lysosomes originate from the membranes of Golgi bodies during the process of secretion and represent zymogen granules.

Function of lysosomes

1. Digestion of intra cellular particles

2. Digestion of extra cellular particles
3. Removal of deal cells
4. Role in fertilization
5. Role during cell division
6. Role in chromosomal break
7. Helps in seed germination
8. Protection

NOTES

5.2.8 Endoplasmic Reticulum (ER)

Historical Background

Endoplasmic reticulum was first observed by **Porter, Claude and Fullan (1945)**. It is a network of strands in the cytoplasm of thinly-spread tissue culture cells. Further studies have shown that, in majority of cases, the network of vesicle-like bodies was confined to the endoplasmic (or inner) part of the cytoplasm. So **Porter and Kallman** (1953) termed it as endoplasmic reticulum(ER).

The details of endoplasmic reticulum in living cells were studied by **Fawcett** and **Ito (1958)** and **Rose** and Pomerat (1960) with the aid of a phase-contrast microscope.

Definition

The semiviscous fluid of the entire cytoplasm (cytosol) of eukaryotic cells is traversed by a membrane-bound network system called **endoplasmic reticulum**. It is a network of cisternae, tubules and vesicles extending from nuclear membrane to the cell surface and even to the neighbouring cells.

Occurrence

The ER occurs in almost all the eukaryotic cells. It is absent in prokaryotes and in some embryonic cells and mature etythrocytes. It is more prominent in young and meristematic cells as compared to older cells where it is represented by apparently discontinuous vesicles. The amount of ER varies from cell to cell. ER present in abundance in those cells which are synthesizing proteins

actively. On an average, in a cell ER forms about 30 to 60% of the total membranes increasing the internal surface many fold.

Type of Endoplasmic Reticulum

NOTES

There are two types of ER - (i) Smooth or agranular Endoplasmic Reticulum (SER).

This type of ER has smooth endo membranes due to lack of ribosomes. The SER is more abundant near the peripheral part of the cytoplasm where it may be attached to plasma membrane. It forms the transport vesicles in which the large molecules (proteins and lipids) are transported from one cell to another via plasmodesmata or within the cells. Often the transport vesicles move towards plasma membrane or to the Golgi complex. (ii) Rough or Granular Endoplasmic Reticulum (RER). The type of ER bears a large number of ribosomes attached over their cytoplasmic surface. Generally the rough ER is more abundant in the deeper part of cytoplasm near the nucleus where it is connected with the outer membrane of the nuclear envelope. Rough ER is engaged in the synthesis and transport of proteins. Generally one form of ER gets changed into another form within a few minutes.

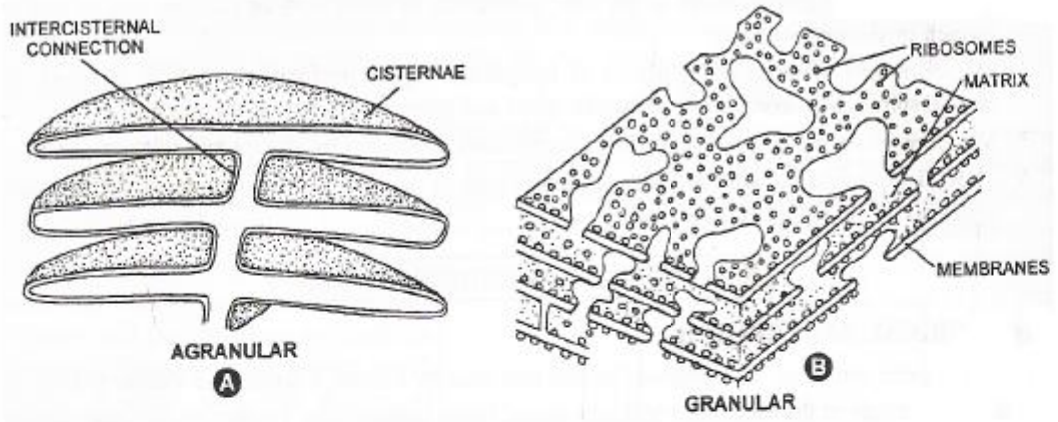


Fig. 5.6 : Endoplasmic Reticulum: (A) - the interconnecting vesicles of a granular ER, (B) granular ER enlarged in three dimentional view.

Structure

There are three major forms of elements which constitute the extensive network of ER. These are -cisternae, tubules and vesicles.

NOTES

(i) Cisternae. These are flattened, unbranched, collapsed sacs which lie parallel to each other. These sacs are interconnected with each other. Generally they bear ribosomes on their cytoplasmic face so they appear rough. The ribosomes are attached on specific sites which are rich in trans membrane glycoproteins called ribophorin I and ribophorin II.

(ii) Tubules. These are tube like extensions generally connected to cisternae. They are smooth and often remain free of ribosomes. The tubules are irregular branching elements which form network mostly in those cells which synthesize and transport cholesterol, glycerides and steroids.

(iii) Vesicles. These are spherical or oval vacuole-like elements which remain freely in the cytoplasmic matrix. They are also free of ribosomes.

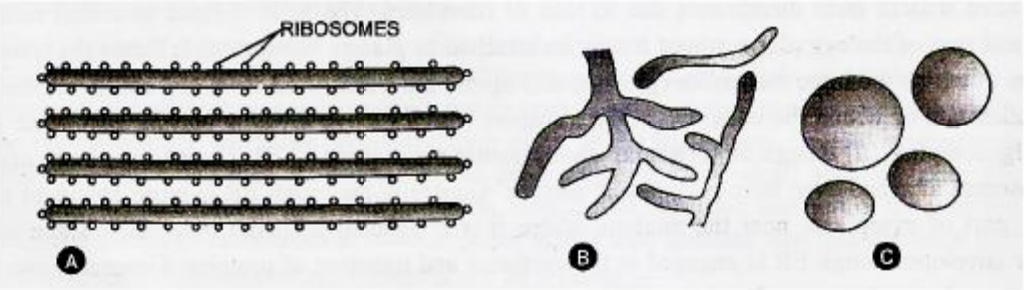


Fig. 5.7 : Elements of Endoplasmic Reticulum: A. Cisternae; B. Tubules ; C. Vesicles.

The endoplasmic reticulum is a hollow system enclosing a single internal space filled with endoplasmic matrix. The membranes of this system are lipo-protein in nature, somewhat analogous to that of the plasma membrane. They are 50 - 60 $\overset{0}{A}$ in thickness.

The various enzymes associated with ER are stearases, NADH cytochrome C - reductase, MgH activated ATPase and many other enzymes of glycerides, fatty acids, L-ascorbic acid and steroid metabolism. The enzyme glycosyl transferase is located in the ER membrane.

Functions of Endoplasmic Reticulum

1. The network of endoplasmic reticulum separates cytoplasm of the cell into many small compartments. This compartmentalization of the cytoplasm helps a cell to perform specific activities by providing enzymes and metabolites within specific chambers excluding others.

NOTES

2. The ER gives mechanical support to the cytoplasm, provides a kind of intracellular cytoskeleton to maintain the form of cell and keeps the cell organelles distributed properly in relation to one another.
3. The ER has osmotic properties and therefore, it is involved in intracellular exchange of materials, particularly between the cytoplasm and the nucleus.
4. ER is usually concerned with the production of membrane in eukaryotic cells by incorporating membrane proteins and phospholipids from precursors in the cytosol. ER plays important role in the synthesis of new nuclear membrane during cell division.
5. The enzyme glycosyl transferase, located in the ER membrane, causes glycosylation of newly synthesized proteins within the cisternae.
6. The ER is usually concerned with the formation of lecithin from fatty acid, glycerol phosphate and choline. The protein components are directly obtained from the sites of protein synthesis and fed into the ER lumen.
7. The ER serves as circulatory system for the transportation of cellular molecules and helps in the storage of synthesized molecules.
8. ER is involved in metabolism of cholesterol and steroid hormones. Smooth ER contains enzymes of lipid metabolism so these are found abundant in the cells which are involved in the synthesis of cholesterol and their conversion into steroid hormone.
9. The RER provides active sites for the attachment of ribosomes where protein synthesis takes place. Some specific proteins, which are synthesized on ribosomes, enter into the ER lumen where they get modified, i.e., subsequent folding and disulphide bond formation. These modified proteins are then packed in membrane bound transport

vesicles which are budded off to carry protein and lipid to other parts of the cell. Some transport vesicles fuse with the Golgi complex for further processing and packaging into other vesicles to be exported by exocytosis.

NOTES

10. The ER causes detoxification of many endogenous and exogenous compounds. The large amount of SER, present in the liver cells, contains enzymes of cytochrome P_{450} family which detoxify lipid soluble drugs converting them into water soluble and excretable products. The additional SER is then removed by autophagosomes.

5.2.9 Golgi apparatus Or Dictyosome or Lipochondrion

Golgi apparatus is consist of membranous structure made of Parallely arranged flattened sacs which has branched tubules, vacuoles and vesicles filled with liquid secretory materials.

Discovery

The Golgi complex was discovered by **Camillo Golgi** in 1898 in nerve cells of barn owl.

Barker(1963) suggested that, this organelle shows morphological variations **distyosome** was coined by Dalton & Felix for Golgi appratus.

Baker coined term "**Lipochondria**" because of presence of lipid content in it.

Electron microscopic studies demonstrated that dictyosomes are universally present in plant cells.

Occurence The Golgi complex occurs almost in all Eukaryotic cells. They are absent in Prokaryotes.

In plant cells, they are scattered through out the cytoplasm. But they are absent in sperm cells of bryophytes and pteridophytes, nature sieve tubes of vascular plants and certain fungi.

There are found scattered or diffused throughout the cytoplasm but their position is fixed for each cell type e.g. the position of Golgi apparatus is Polar between nucleus and periphery of the cell in ectodermal cells. In neurons they are found around nucleus.

The presence of well maintained Golgi complex depends upon the presence of nucleus. In young metabolically active cells they are well developed but they get reduced in mature cells and in those cells which are under starvation.

Size, Shape number of Golgi apparatus

NOTES

The shape and size of Golgi complex is variable when depend upon the active state of the cell. The shape of Golgi apparatus varies at different functional stages. The shape may be from a compact mass to disperse filamentous network.

Number of Golgi apparatus varies one to many depending on the cell type for example–

S.No.	Type of Cells	Number of Golgi body
1.	Animal cell	3-7
2.	Liver cells	Upto 50
3.	In plant cells	From 10 to 20
4.	In algal rhizoids	Exceptionally very high upto 25000
5.	In paramecium	Two Golgi complex are found

Structure : Golgi apparatus looks like disc shaped and comprises of the morphological components.

i. Flattened discs : They remain filled with fluid. Each cisternae has cavity of about 150 $\overset{0}{A}$ wide whose edges are dilated. In some cells cisternae are held together by a thin layer of an opaque dense material called 'Node".

They are usually arranged in parallel rows separated from each other by thin layers of inter cisternae cytoplasm.

Each cisternae has two faces

a. Cis face / convex forming : This face lies away from the nucleus.

b. Trans face / Concave forming : This face lies toward the nucleus.

Both the ends of cisternae Possess rounded swelling.

NOTES

ii. Clusters off tubules : From the peripheral area of cisternae, flat network of tubule arise. Diameter of these tubules is 300 to 500 $\overset{0}{A}$.

iii. Vesicles : Golgi complex contain 3 type of vesicles.

a. Smooth vesicles b. Rough vesicles c. Golgian vesicles

iv. Large vacuoles

Smooth vesicles	Rough vesicles	Golgian vesicles
They are called secretory vesicles as they contain secretory materials	They are found at the periphery of the Golgi complex	They are large rounded sacs
They are smooth surfaced vesicle	They are rough surfaced spherical protuberances	They are formed by fusion of secretory vesicles
They arise from the ends of cisternae tubules		They are filled with amorphous substances

Chemical composition : Following are the chemical constituents of Golgi complex:

a. Phospholipid b. Protein & enzymes c. Carbohydrates

iv. Vitamin C

Function of Golgi Complex

1. Golgi bodies synthesize carbohydrate of cell wall matrix (e.g. pectic compounds, hemicelluloses, mucopoly saccharides)
2. Vesicles of Golgi apparatus fuse with plasma membrane and helps in increasing surface area during growth. Vesicles also help in forming lysosomes.

3. They help in secretion and transport of materials out of the cell.

4. In plants, cell plate is formed in the middle of dividing cells by fusion of vesicles produced by Golgi apparatus.

5.3 Protein Synthesis and Sorting

NOTES

Proteins

Proteins are ubiquitous organic nitrogenous components of all living organisms - plants, animals and bacteria. These macromolecules contain carbon, hydrogen, oxygen, nitrogen, sulphur and sometimes phosphorus and iron. The proteins can be broken in to simpler substances known as aminoacids – which are building blocks of proteins. There are 20 different kinds of aminoacids which are arranged in a definite sequences and number in different protein molecules. The aminoacids are linked together by peptide bonds (-CO-NH-) which are formed by the condensation of an aminogroup ($-NH_2$) of one aminoacid and carboxyl group (-COOH) of another.

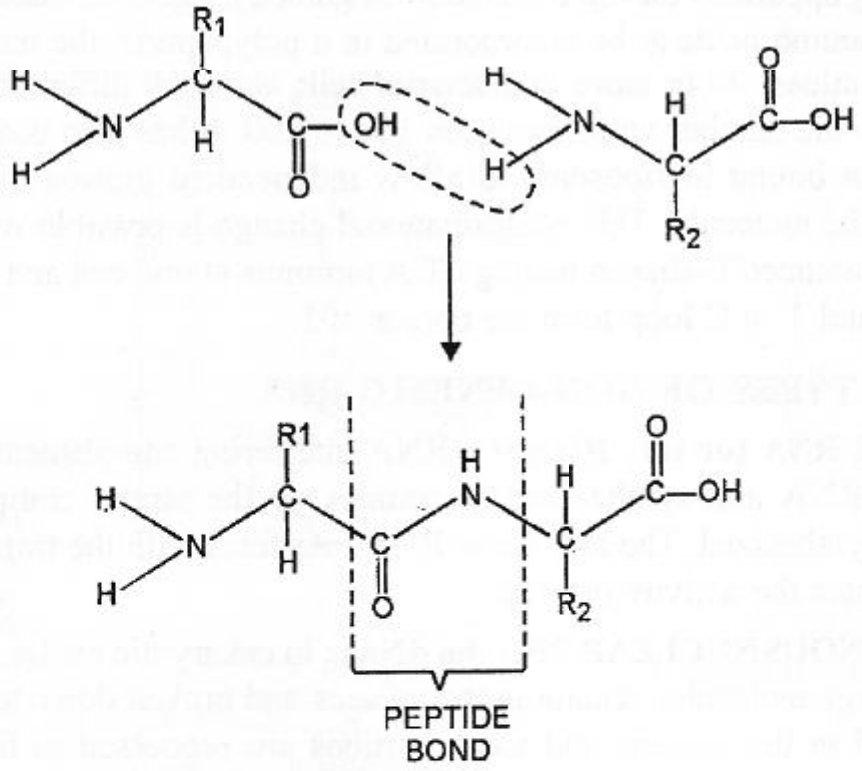


Fig. 5.8 : Mechanism of Protein Synthesis

NOTES

The proteins are macromolecules consisting of a series of polypeptide chains, made up of varying amounts of 20 different an aminoacids. The protein molecules differ from each other in their length and definite sequence of aminoacids. The protein synthesis is directly under genetic control.

Various steps of protein synthesis have been given below:

1. Transcription

The important information for the specific sequence and the total number of aminoacids, to be incorporated into the protein, are stored in the master molecule DNA in the form of sequence of bases. It transfers this information to the protein synthesizing apparatus through the special messenger called **mRNA**. The transfer of information from DNA to RNA is called **transcription**. The portion of DNA responsible to synthesize mRNA leads to the production of one polypeptide chain is known as **gene**.

The process starts with the uncoiling and separation of the two strands of gene DNA by breaking of hydrogen bonds between the bases. Out of the two strands of DNA, one being gene strand acts as template and the other its complement strand, which do not transcribe. The single stranded mRNA – molecule is synthesized on DNA template. The enzyme responsible to synthesize mRNA is called **DNA dependent RNA polymerase**. The enzyme bears a sigma factor which recognizes the specific site on DNA template. The complementary ribonucleotides (i.e., ATP, GTP, UTP and CTP) are joined together by the enzyme and a molecule of mRNA is synthesized. The mRNA is usually synthesized in a 5' and 3' direction. Finally, the growing chain of mRNA gets detached from the template and again the two strands of DNA join together by hydrogen bonds.The mRNA which is synthesized on DNA, bears the similar complimentary sequence of bases present on the gene-DNA For example, if the DNA strand read as :

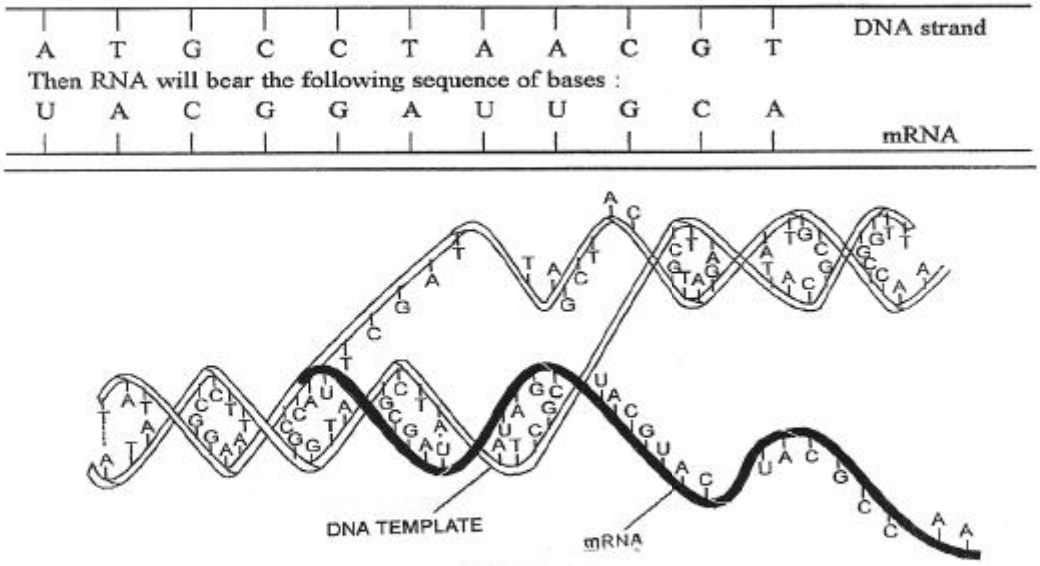


Fig 5.9 : A Model of Transcription - On DNA template – Syynthesis of mRNA

In prokaryotic organism (e.g., bacteria, cyanobacteria, etc.), the mRNA molecule are short lived and generally polycistronic (i.e., carry sequences of bases coding for several proteins). Thus single mRNA is transcribed from a group of adjacent genes.

In eukaryotic organisms, the mRNA molecule is generally transcribed as a large precursor molecule (i.e., heterogenous nuclear RNA or hn RNA) which undergoes post-transcription modification in the nucleus and then moves into the cytoplasm where it is translated. During modification (or processing), the precursor RNA undergoes enzymatic cleavage of message sequence (splicing), capping of 5' end and polyadenylation of 3' end. After enzymatic cleavage, a cup of methylated guanine (G) is added at 5' end. The G is added in the reverse orientation from all other nucleotides. This structure is called **cap**. A long chain of about 100 - 200 adenylic acid (A) is added to the 3' end of mRNA and called **poly (A) tail**. The poly (A) tail help in transporting the transport of mRNA to the cytoplams.

2. **Amino Acid Activation**

The transfer of amino acid, for the formation of a polypeptide chain, takes place with the help of a specific RNA called **tRNA**. The molecule of tRNA, is single stranded Clover leaf shaped structure. It has two endings - CCA - GH

ending and G - ending. The specific amino acid is picked by its CCA-OH ending whereas the function of G - ending is still un known.

NOTES

There is a large number of different kinds of aminoacids. about 20 amino acids take part in the synthesis of protein. Each amino acid is picked by its specific tRNA hence the number of different kinds of tRNA is more than 20 (approx. 64 or more). They differ from each other in the presence of specific anticodon,(i.e., a triplet of three unpaired bases) on the anticodon loop (or loop II). The bases of anticodon form hydrogen bonding with complementary bases of codon present in mRNA, at the time of protein synthesis.

Thus the attachment of specific aminoacid to its specific tRNA is called **aminoacid activation**. This process is independent of transcription and occurs in the cytosol.

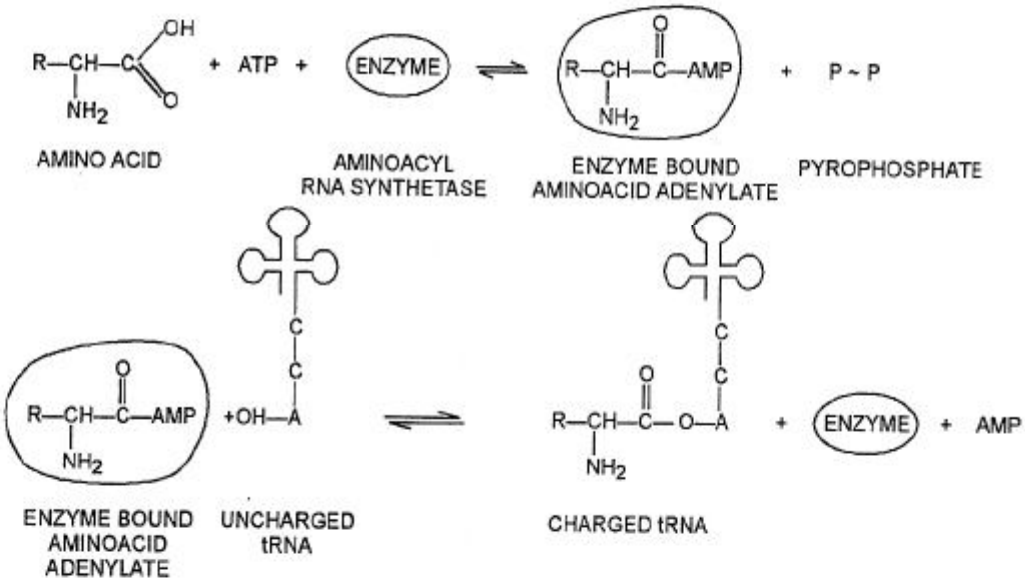


Fig 5.10 : Aminoacid Activation

The aminoacids first react with ATP in presence of enzyme aminoacid tRNA - synthetase and form enzyme-bound aminoacid adenylate by releasing pyrophosphate.

Aminoacid + ATP + Amino acyl-RNA synthetase == Enzymebound amino-acid adenylate + Pyrophosphate.

The adenylate remains bound to the enzyme till it is transferred to tRNA. The bound adenylate reacts only with tRNA molecule specific for its particular

aminoacid and produce amino-acyl-tRNA (i.e., charged tRNA). The enzyme and AMP are evolved in the reaction.

Enzyme bond Aminoacid adenylate + t-RNA == Aminoacyl t-RNA + AMP + Enzymei

NOTES

The process of aminoacid activation and formation of charged tRNA is not dependent on transcription.

3. **Ribosome Template (The Protein Synthesizing Apparatus)** The actual site for protein synthesis is small cellular organelles, called **ribosomes**. In eukaryotes, they remain attached to endoplasmic reticulum or occur freely in the cytoplasmic matrix, chloroplast and mitochondria. In prokaryotes, they occur only in free form. Sometimes many ribosomes are joined with mRNA forming linear or helical groups, called **polyribosomes** or **polysomes**. The ribosomes which occur freely in the cytoplasmic matrix act as platforms for the synthesis of structural proteins whereas those bound to endoplasmic reticulum synthesize globular proteins and transfer them to cisternae for transport to other parts of the cell.

The ribosomes found in the cytoplasmic matrix of eukaryotes are 80 S whereas those of prokaryotes and chloroplast and mitochondria are 70 S. Each ribosome consists of two sub units of unequal size. The composition of these two sub units is as follows.

(i) 80S ribosome (found in cytoplasmic matrix of eukaryotes)

- (a) Larger sub unit (60S) : consists of 5S RNA, 28S RNA and 40-45 protein molecules:
- (b) Smaller sub unit (40S) : Consists of 18S RNA and 30-:-35 protein molecules.

(ii) 70S ribosome (found in prokaryotes and chloroplast and mitochondria of eukaryotes)

- (a) Larger sub unit (50S) : Consists of 5S RNA, 23S RNA and 30-35 protein molecules.
- (b) Smaller sub unit (30S) : Consists of 16 S RNA and 20-25 protein molecules.

NOTES

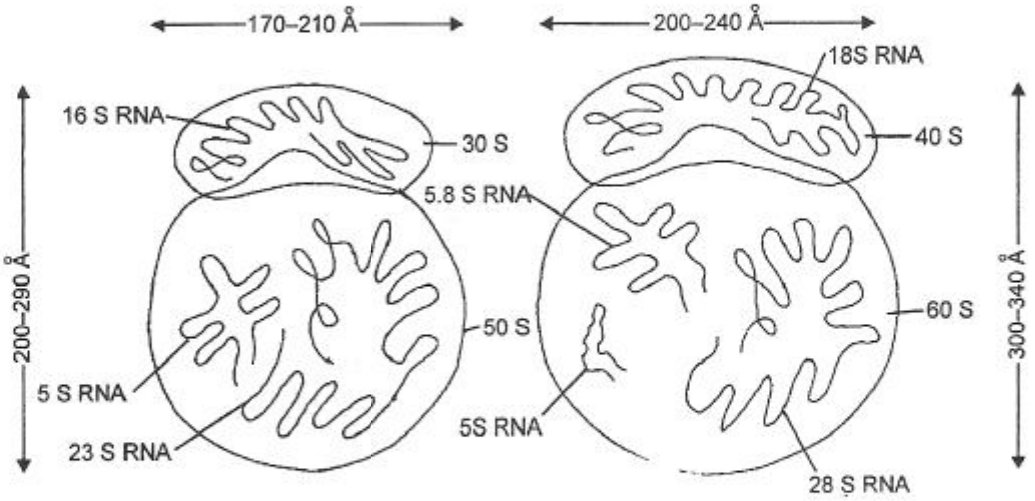


Fig. 5.11 : Structures of 70S (A) and 80S (B) Ribosomes

The ribosomes remain dissociated into their sub units (i.e., 50S and 30S in prokaryotes and 60S and 40S in eukaryotes). However, the two subunits of each ribosome join to form a protein synthesizing apparatus (ribosome template) as soon as the initiation of protein synthesis starts.

Formation of ribosome template requires - amino acyl tRNA complex, mRNA molecule, 40S and 60S sub units (in case of eukaryotes) or 30S and 50S sub units (in case of prokaryotes), ATP, GEP, and some specific proteins designated as 'initiation factors' (IF). The ribosome template ready for protein synthesis has four active sites -

I. **P-site (peptidyl site)** : This site located in the ribosome where amino acyl tRNA can be inserted. It is called peptidyl site because the free carboxyl group (-COOH) of the peptidyl tRNA takes part in the peptide bond formation.

II. **A-site (amino** acyl site) : It is the site located in the ribosome where free amino group ($-NH_2$) is present with amino acyl tRNA available for peptide bond formation.

The position of P-site is not fixed in the ribosome. They keep on changing with elongation of polypeptide.

(i) **Peptidyl transferase centre**: It is located on 50 S or 60S sub unit of ribosome and concerned with the formation of peptide bond between the two amino acids.

(ii) **GTPase Site**: It is mainly located on 50S or 60S sub unit of ribosome and concerned with the hydrolysis of GTP to GDP and pyrophosphate during elongation of polypeptide.

4. Translation

NOTES

Transporter of genetic information from the nucleotide sequence of mRNA to amino acid sequence of protein is called **translation**. This process involves three important steps - (i) **Initiation**, (ii) **Elongation** and (iii) **Termination**.

(i) Initiation of protein synthesis:

In Prokaryotes

The important events leading to the formation of protein synthesizing complex and the initiation of protein synthesis in prokaryotes are as follows -

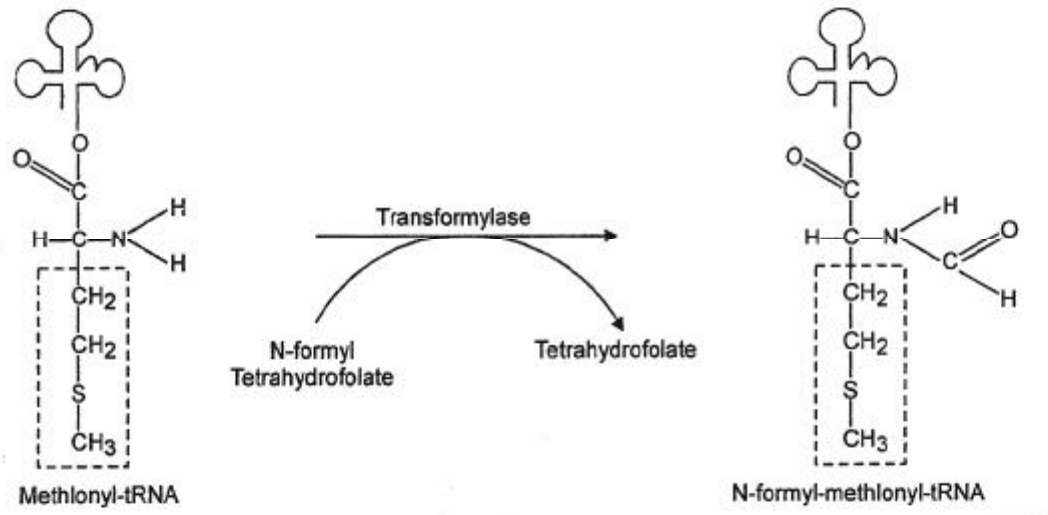


Fig. 5.12 : N-formyl Tetrahydrofolate Tetrahydrofolate

(a) In the beginning, an amino acyl tRNA molecule having attached a molecule of methionine (i.e., methionyl - tRNA) is formylated in presence of enzyme transformylase to form N-formyl-methionyl tRNA.

(b) The mRNA, (transcribed from the DNA), has two ends 5' end and 3' end. The 5' end of mRNA has nucleotide sequence complementary to those present at the 3' end of 16S rRNA (ribosomal RNA) present in the smaller sub unit (30S) of dissociated ribosome. The mRNA gets attached to 30S subunit of ribosome in presence of initiation factor (1F3).

NOTES

(c) The first t-RNA carrying formyl methionine, called N-formylmethionyl-tRNA is selected by initiation factor (1F2) to form a ternary complex in presence of GTP. The ternary complex then binds to P-site (peptidyl site or donor site) of 30S sub unit of ribosome to which IF3 is attached. This binding involves attachment of anticodon of tRNA to the initiation codon of mRNA. The tRNA, to which amino acid - formyl methionine is attached, has anticodon UAC. The mRNA has initiation codon AUG at 5' end. The codon AUG of mRNA and anticodon UAC of formyl methionine - tRNA are complementary to each other. They come in line and get attached by the formation of hydrogen bonds between the bases.

(d) At this stage, the complex consists of 30S sub unit of ribosome mRNA N-formyl methionyl tRNA, IF2 and IF3.

(e) Finally, this complex joins the larger sub unit (50S) to form 70S ribosomal complex. This process occurs in presence of initiation factor (IFI). Once the 70S ribosome initiation complex is formed, the IFI, IF2 and IF3 is released and GTP is hydrolysed to GDP. Now this initiation complex is ready for translation of mRNA.

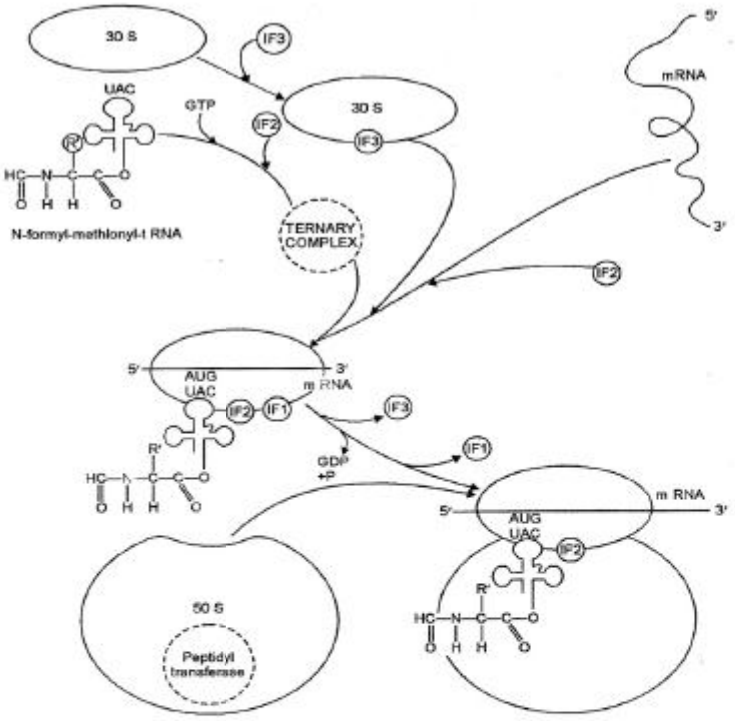


Fig. 5.13 : Initiation of a Polypeptide Chain in Prokaryotes

In Eukaryotes

In eukaryotes, the process of the formation of protein synthesizing complex and the initiation of protein synthesis is more complicated and characterized by the following

NOTES

(a) The eukaryotic mRNA has a methylated cap at its 5' terminus and 100 - 200 bases of polyadenylic acid (i.e., poly-A) at 3' terminus. . ,

(b) The process involves about 9 initiation factors (i.e., IF1, IF2, IF3, IF4A, IF4 B, IF4 C, IF4D, IF5 and IF6).

(c) The first amino acid-methionine is not formylated.

(d) In eukaryotes, the methionyl tRNA gets attached to 40S sub unit of ribosome first and then the mRNA gets attached to it.

(e) The 40 S sub unit then gets attached to 60S subunit to form 80S ribosome complex.

(ii) Elongation

Elongation of polypeptide chain begins as soon as the ribosome template is formed. It involves three major steps –

(a) Binding of codon directed amino acyl- tRNA at the vacant A-site of ribosome. During this process, a molecule of GTP is hydrolysed to release GDP and a phosphate group. The IF 1 is released.

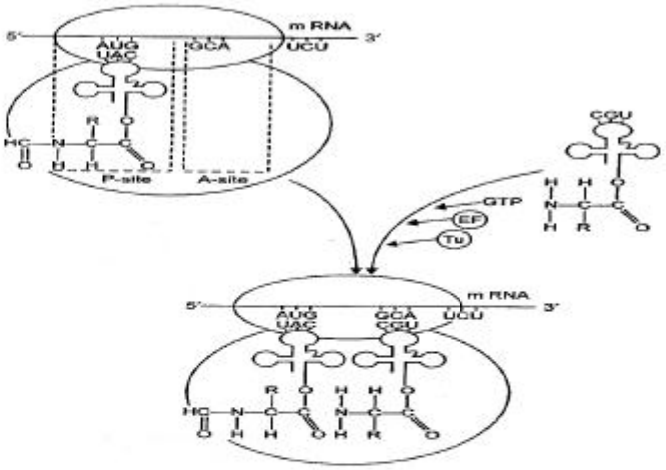


Fig. 5.14 : Role of Protein Synthesizing Apparatus (Ribosome) in the Elongation of a Polypeptide Chain in Prokaryotes

NOTES

(b) Formation of peptide bond between the first amino acid(attached to peptidyl tRNA located at the P-site) and the next amino acid (attached to amino acyl tRNA located at the A-site) occurs. The energy released during conversion of GT'P to GDP is used up in the peptide bond formation. The reaction is catalysed by enzyme pentidyl transferase.

The peptide bond is formed between the carboxylic group (-COOH) of first amino acid and free amino group (-NH_2) of second amino acid so that the two amino acids (di peptide) get attached to second tRNA and the first tRNA becomes free.

(c) The ribosome-mRNA complex then moves by one codon in 5'→3' direction so that the empty tRNA is released from the P-site and the tRNA having a di peptide translocates from A-site to P-site site leaving A-site open for the next aminoacyl tRNA. The new amino acyl tRNA, specific for the next codon of mRNA, gets attached to open A-site and the peptide bond formation takes place between the second and third amino acids. This process is repeated many times and the elongation of polypeptide continues until the termination codon is reached.

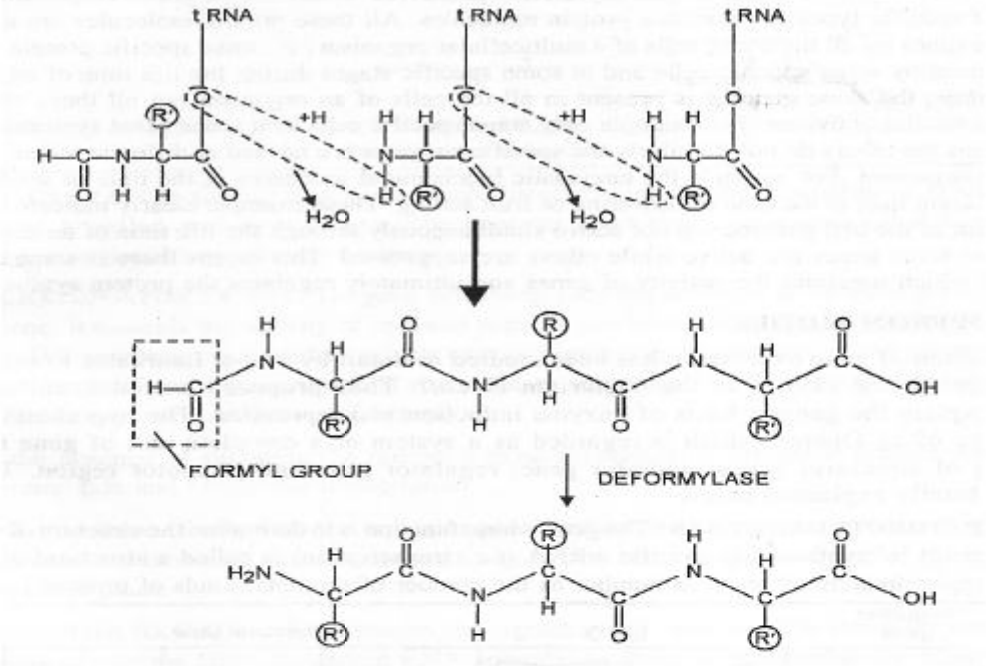


Fig. 5.15 : A-Formation of Peptide Bonds in the Process of Protein Synthesis

The process of elongation and peptide bond formation requires several elongation factors viz., EF - Tu, EF - Ts and EF - G in prokaryotes and EF - I and, EF - 2 in eukaryotes.

NOTES

(iii) **Termination**. The synthesis of polypeptide chain terminates when mRNA comes across one or more of the chain termination codons i.e., UAA, AGA or UGA. These codons are called **stop triplets**. The cells do not possess such tRNA having anticodons which would fit with these codons. Thus the stop triplets do not attach to any of the amino acyl tRNA molecules. During the process of termination the enzyme peptidyl synthetase catalyses the transfer of polypeptide chain to water instead of tRNA. The polypeptide tRNA linkage is cleaved and the polypeptide chain is released from the ribosome template. This step requires at least two termination factors R_1 and R_2 . The role of these factors is still not clear.

Atlast the ribosome is released from mRNA and dissociates into its sub-units. The mRNA is degraded rapidly into its nucleotides which can be reused. It has been shown that each mRNA molecule can serve as a template for about 15-20 polypeptide chains.Finally by specific folding and refolding,the polypeptide chains are converted in to protein molecules..

Regulation of protein synthesis

The protein synthesis is regulated by the genetic material of the cell and the external chemical environment. The genetic material (mostly DNA) possesses thousand of genes which responsible for the synthesis of specific types of numerous protein molecules. All these protein molecules are not synthesized all the times by all the living cells of a multicellular organism i.e., some specific protein molecules are synthesized by some specific cells and at some specific stages during the life time of an organism.

In other words, the same genome is present in all the cells of an organism but all these cells do not perform the similar activities. For example only some specific cells of a green plant synthesize chlorophyll whereas the others do not. Similarly the specific enzymes are needed at different stages in the life cycle of an organism. For example the enzymatic biochemical processes at the time of seed germination are different than at the time of flowering or fruit setting.

NOTES

These example clearly indicate that all the genes present in the cell genome are not active simultaneously through the life time of an organism. At a given time, some genes are active while others are suppressed. This means there is some regulatory mechanism which regulates the activity of genes and ultimately regulates the protein, synthesis.

The Operon Model

Regulation of gene expression has been studied in detail by Nobel Laureates **Francois Jacob and Jacuqes Monod (l961)** in the bacterium E. coli. They proposed a model, called **Operon model** to explain the genetic basis of enzyme induction and repression. The hypothesis proposes the presence of an Operon which is regarded as a system or a complete unit of gene regulation comprising of structural genes, operator gene, regulator gene and promotor region. These elements are explained below:

1. **The Structural Genes:** The gene whose function is to determine the structure of a polypeptide (or protein) by synthesizing specific mRNA (i.e., transcription) is called a structural gene. These genes are approximately as many in number as the number of proteins (kinds of proteins).

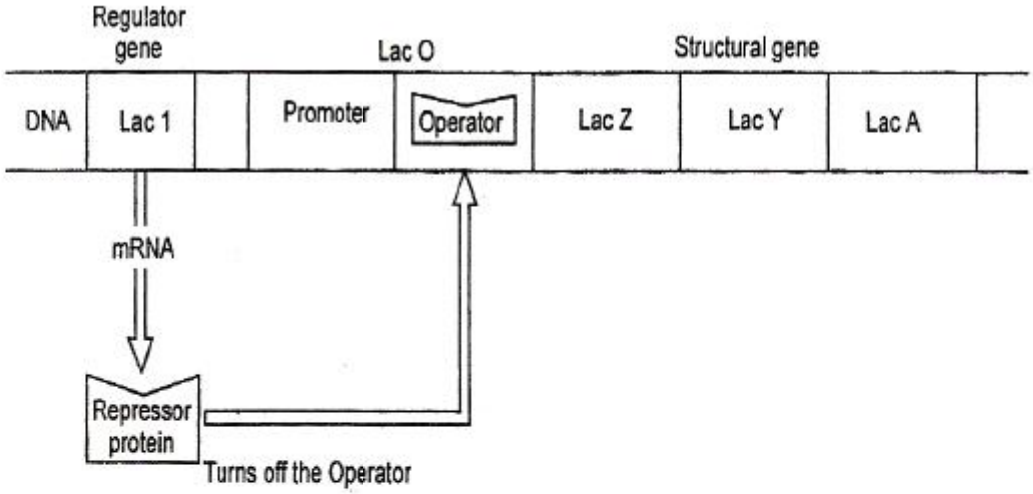


Fig. 5.15 : B Lac Operon - No Transcription and no Translation of Enzyme

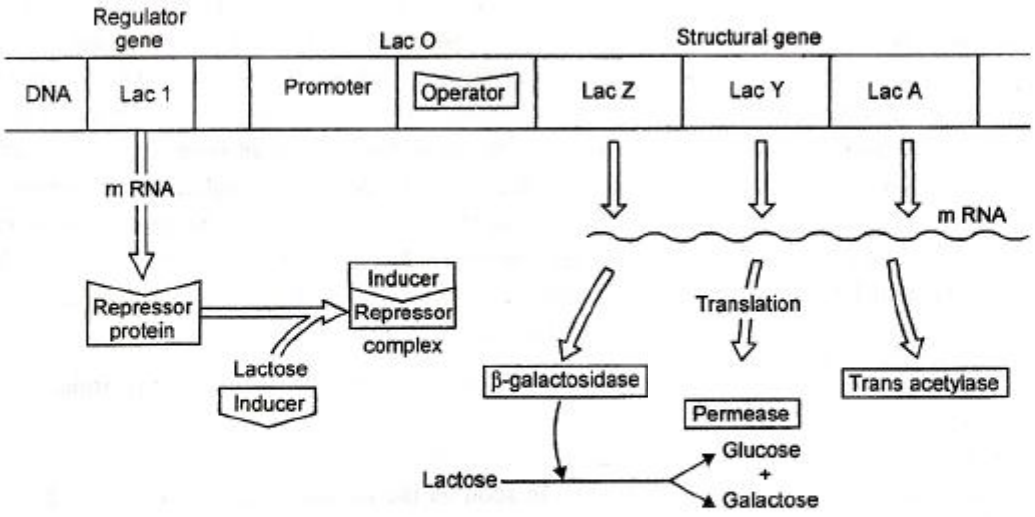


Fig. 5.16 : Lac Operon- Induction

2. **The Operator Gene**. The gene which directly regulates the formation of mRNA of the structural gene is called an **operator gene**. The operator gene acts as a switch either blocking the structural genes or permitting them to function.

3. **The Regullatory Gene.** The gene which regulates the action of an operator gene is called a **regulatory gene**. It controls the activity of operator gene by producing inhibitors of gene activity called **repressors**. The repressor is a protein which causes repression (repression is the phenomenon in which the synthesis of particular enzyme is repressed or stopped). The function of regulator gene is affected by the presence or absence of substrate -

   (a) In the absence of substrate the redressor gene synthesizes. a repressor protein which binds itself to operator side and blocks the transcription.

   (b) In the presence of substrate the regulator gene synthesizes a repressor protein which forms a complex with the substrate so that the operator site is not blocked. Consequently, the process of transcription takes place.

4. **Promotor Region.** It lies between the regulator and operator. The Promoter region is the site for binding of enzyme DNA-dependent RNA polymerase which is responsible for transcription.

NOTES

NOTES

Originally the model was proposed for the genetic regulation in the synthesis of an inducible enzyme β-galactosidase in bacterium E. coli. The model was called **lac operon**. Now a large number of operons have been discovered so far – (1) Lac operon, (2) Tryptophan operon, (3) Leucine operon, (4) Arabinose operon, (5) Galactose operon, (6) Histidine operon, (7) Arginine operon, (8) Glycerol operon (9) Lambda Phase operon.

Lac Operon (Or Lactose Operon) : The model proposes regulation of genes for lactose metabolism in Escherichia coli. Degradation of lactose involves three enzymes - β-galactosidase, permease and transacetylase. The enzyme permease helps penetration of lactose molecules into the bacterial cell whereas the enzyme β-galactosidase breaks down lactose into glucose and galactose. However, the exact role of transacetylase is still to be known. The enzyme β-galactosidase is an inducible enzyme and synthesized only when lactose is supplied to the cell. The other two enzymes are also synthesized along with β-galactosidase. Thus the substrate, lactose acts as an inducer for the synthesis of all the three enzymes. According to Lac operon model, the genes responsible for the synthesis of these three enzymes occur adjacent to each other and thus are linked. These are designated as structural genes - Lac Z (which codes for enzyme β-galactosidase), Lac Y (which codes for β-galactoside permease) and Lac A (which codes for enzyme transacetylase). These three structural genes are regulated by a single switch referred to as the Operator (or the gene Lac O).

The operator switch (Lac O) is turned on or off by a protein called repressor protein, which is coded by another gene, called regulator gene (or Lac i). The repressor protein, when binds to the Operator of operon and blocks, it results inactivation of the structural genes so that the transcription and protein synthesis is stopped. As soon as the lactose (the inducer) is added to the culture of *E. coli*, a few molecules of lactose are carried into the bacterial cell with the help of permease enzyme. A small amount of pemease is present even under repressed conditions. These lactose molecules are converted to an active form (allo-lactose) which binds to the repressor. It results inactivation of repressor which fails to bind the operator switch. The operator switch is then turned on so that the structural genes are transcribed. In this way the operon is induced by the inductor substrate-lactose.

5.4 Summary

Different micorbodies / cell organelles are responsible for various functions e.g. vacuoles are responsible for mechanical support, storage, osmoregulation, and ribosome are responsible for protein synthesis.

NOTES

5.5 Glossary

- **Cell Biology :** Branch which deals with structure and function of cells and cellular components, cell chemistry, cell physiology, cell reproduction and cell genetics.
- **Cytoplasmic Organelles :** Cellular organelle which are found in cytoplasm are called cytoplasmic organelle. Nucleus is excluded here.
- **Polyribosome / Polysome:** Some times many ribosomes are joined with mRNA forming linear or helical group called Polysomes.
- **Suicide bags :** Also called "autophagosome". These Lysosomes originate from endoplasmic reticulum and digest the organelle of the cell in which they occur. So they are known as "suicidal bags" of the cell.
- **Nucleus :** It is the most important part of the cell which direct and controls and all activities of the cell. It is also called 'Karyon'.
- **Nucleoid :** In Prokaryotes(Bacterium Cynobacteria, Mycoplasma), a well organised nucleus with distinct nuclear membrane is absent but the chromatin material (A highly folded circular DNA molecule) lies in the cell cytoplasm known as "Nucleoid". (Prokaryon /genophore)
- **Endosomes :** Small chromatin bodies present in the nucleoplasm of nucleus.
- **Nucleolus** : It is a dense spherical body filled with acido phillic material which remain embedded in the matrix.
- **Chromatin** : They appear as a viscous gelatinous substance which contain DNA, RNA basic (Histones) & acidic proteins.

NOTES

- **Ribosomes** : They are the actual site for the synthesis of polypeptides (i.e. Protein synthesis)
- **Protein** : Protein are ubiquitous organic mitrogenous components of all living organism (plant, animals and bacteria or protein are mecromolecules consist of a sereis of polypeptide chains, made up of 20 different amino acids.
- **Transcription** : The transfer genetic information from DNA to RNA is called transcription. It is I stage of gene expression during which geneti information for protein synthesis are transferred to mRNA. In this process coding stand of DNA acts as a template for synthesis of complementary strand of mRNA in the presence of enzyme RNA polymerase.
- The process is transcription and RNA molecule synthesis is called mRNA transcript.
- **Translation** : Transfer of genetic information from the nucleotide sequence of mRNA to amino acid sequence of protein. It involves 3 steps.

  i. Initiation ii. Elongation iii. Termination

- **Genetic RNA** : In some plant and animal viruses, genetic material is RNA instead of DNA. It is called genetic RNA
- **Non genetic RNA** : In systems where genetic material is DNA, RNA molecule functions as non genetic materials e.g. tRNA, rRNA, mRNA
- **Central dogma** : Flow of genetic information from of N.A. (DNA) to another N.A. (RNA) and to protein is called central dogma
- In order words the information encoded in Deoxyribonuclie acid flows to ribonucleic acid and to protein is called **central dogma**.

5.6 Self-Learning Excercise

NOTES

Section A (Very Short Answer Type Questions)

1. The largest cell organelle is .....
2. Lysosomes contain .....
3. Ribosomal RNA is synthesized in .....
4. The site of synthesis and degradation of H_2O_2 in a plant cell is .....
5. Maximum number of Golgi bodies found .....
6. What is the name f a membrane which surrounds the vacuole ?
7. Which cell organelle is called suicidal bags?
8. The smallest organelle in the cell is .....

Section B (Short Answer Type Questions)

1. Compare the 705 and 805 ribosome?
2. Write the structure and function of peroxisome?
3. What is translation?
4. Write note on spherosomes?
5. What is the role of lysosomes for the cell?

Section C (Long Answer Type Questions)

2. Write notes on followings –
   a. Chemical constitution of ribosomes
   b. Vacuoles
3. Write short notes on –
   a. Golgi Body
   b. Sphaerosomes
4. Describe the structure and functions of ribosomes.
5. Describe the structure of Endoplasmic reticulum and its function.

Answer key of Section – A

NOTES

1. Plastid
2. Hydrolytic enzymes
3. Nucleolus
4. Peroxisome
5. Root apices
6. Tonoplast
7. Lysosomes
8. Ribosomes

5.7 References

- "Cell biology genetics and plant breeding" by Trivedi & Sharma RBD (2006)
- "Cell biology & Genetics" by Srivastava, H.N. Vol. II PP 1-184 Pradeep Publications, Jalandhar.
- "Cytology genetics biotechnology and biostatistics" by Gupta,P.K. Rastogi publications. Merrut .

Unit – 6

Cell Shape and Motility

NOTES

Structure of the Unit

6.0 Objectives

After going through this unit you will be able to understand:

- Cell Shape and Motility
- Cytoskeleton
- Microtubules
- Microfilaments
- Intermediate Filaments (IF)

6.1 The Cytoskeleton

NOTES

Introduction and Organization

Nearly all cells contain some form of filamentous protein polymer in their interior. Eukaryotic cells often possess an intricate three-dimensional network of interpenetrating and partly cross linked filaments whose length, concentration, and attachments to each other or to the cell membrane are remodelled in response to specific signals. This reversible assembly and spatial organization of biopolymers is intimately related to the cell's ability to maintain its shape or to alter it as it moves or is otherwise activated. The network of filamentous proteins, often operationally defined as that part of a cell which remains insoluble when the membrane is disrupted by detergents, is generally called the cytoskeleton. The filaments in this skeleton are composed of three types of protein polymer–microfilaments, intermediate filaments and microtubules – and the many accessory proteins that bind to them.

The cytoskeleton of a cell is responsible for structure, strength, and motility within and between the cells. It provides cellular scaffolding upon which the cellular organization is arranged. Cytoskeleton present in a wide variety of prokaryotic and eukaryotic cells both plants and animals. The cell membrane (plasma membrane) is also associated with cytoskeleton and form 'membrane skeleton'.

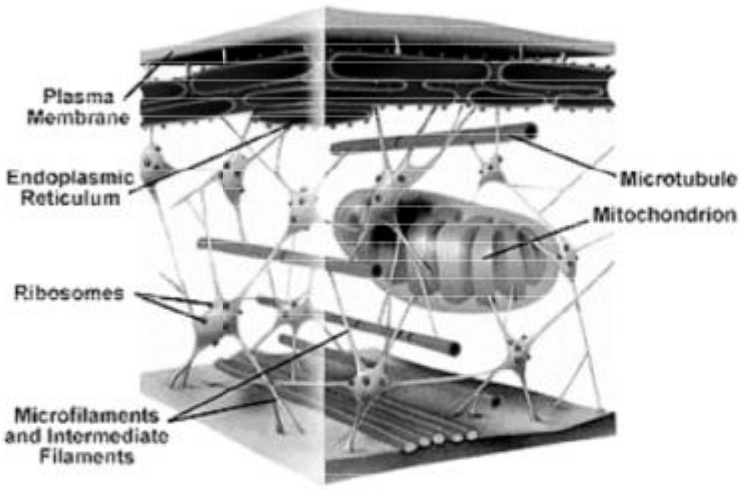


Fig. 6.1 : Cytoskeleton in Cell

Cytoskeleton composes by three types of components:1. Microfilaments - These are the thinnest cytoskeleton component and are composed of the globular protein 2. Microtubules- These are the thickest cytoskeleton component and are constructed of globular Tubulin proteins. 3. Intermediate fibres (IF) - IFs are intermediate in thickness as compared to microtubules and microfilaments. Keratin is an example of an IF.

NOTES

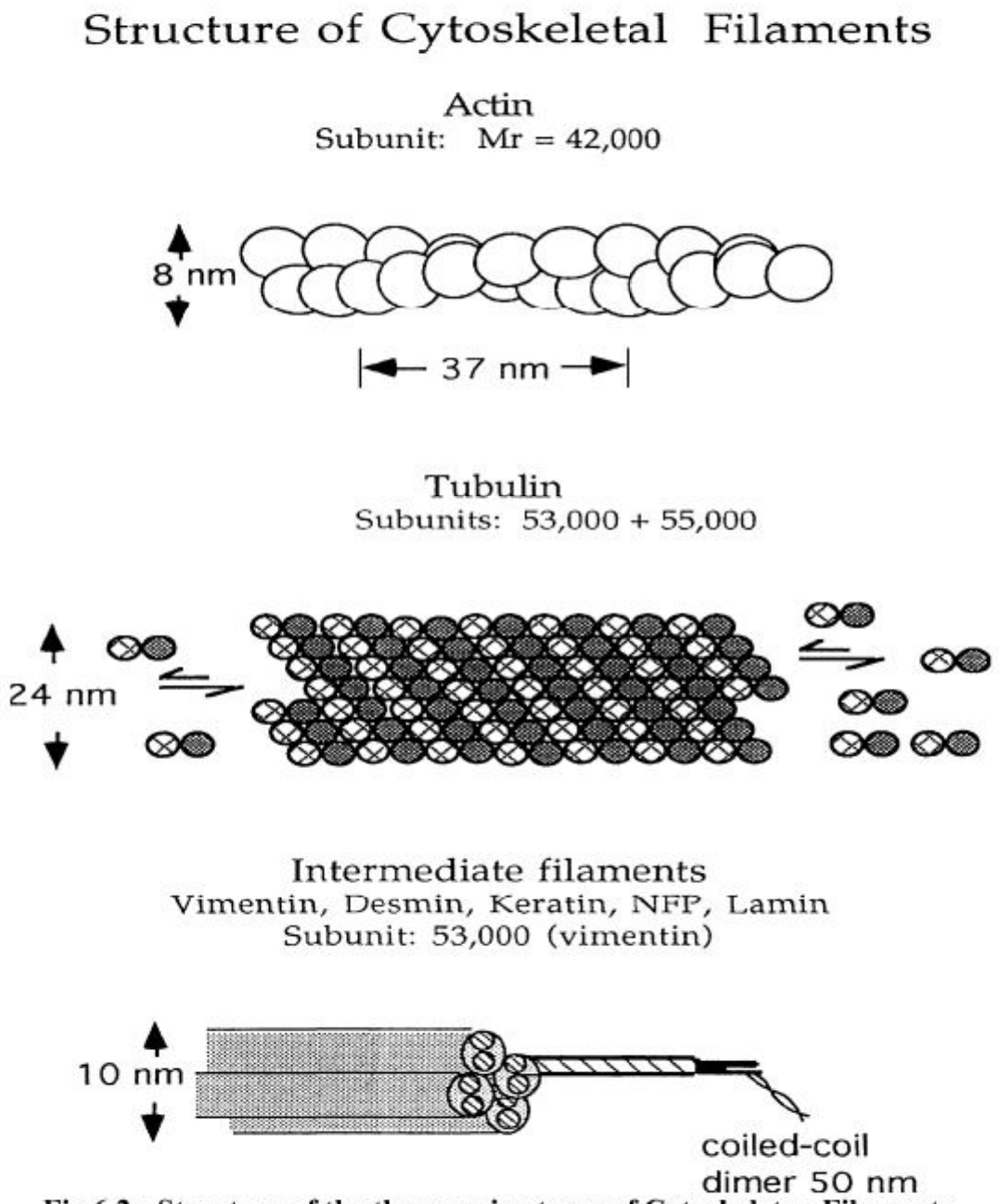


Fig.6.2 : Structure of the three major types of Cytoskeleton Filaments

6.2 Microtubules

NOTES

Microtubules are first observed by DeRoberties and Franchi (1953) in nerve fibres and the term microtubule was introduced by Slautterbsck (1963). Microtubules are filamentous intracellular structures that are responsible for various kinds of movements in all eukaryotic cells. Microtubules are involved in nucleic and cell division, organization of intracellular structure, and intracellular transport, as well as cilliary and flagellar motility. Because the functions of microtubules are so critical to the existence of eukaryotic cells (including our own), it is important that we understand their composition, how they are assembled and disassembled, and how their assembly/disassembly and functions are regulated by cells.

For the sake of brevity, only the very basic and universal concepts about microtubules and their organization into flagella will be presented here, leaving many questions unanswered. You will find that unit provide more complete descriptions of microtubules and their structures and functions, but they also leave many questions unanswered. Unit seldom tell us is how much science knows and does not know about them, and of course they cannot be up to date with the latest discoveries. To fully understand a subject it is important to go to multiple sources. If the subject is especially important to you, you should seek the primary literature, namely original research reports. Microtubules shows polarity, with a plus (+) end (for rapid growth and polymerization) and minus (-) end (for depolymerisation). The minus ends remain embedded in centrosome and plus ends are free to grow by adding tubulin molecules. Microtubules always arrange towards centre with their minus ends.

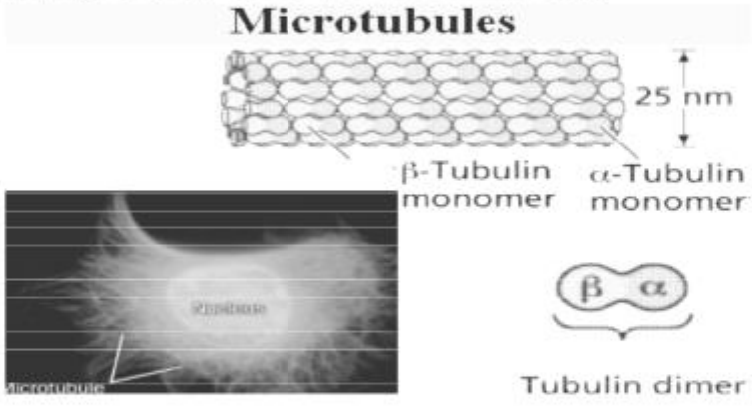
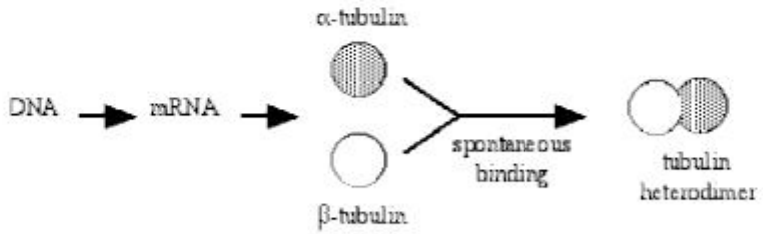


Fig. 6.3 : Microtubules and their Components

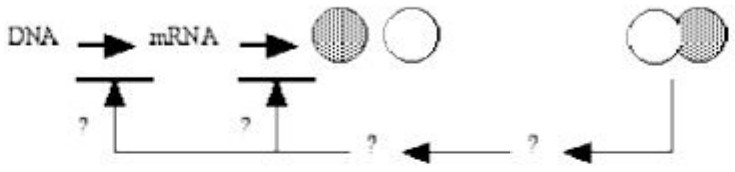
6.2.1 Components of Microtubules

All eukaryotic cells produce the protein tubulin, in the usual way. The usual way, of course, is by transcription of genes coding for tubulin to produce messenger RNA, followed by the translation of mRNA by the ribosomes in order to produce protein. Cells maintain at least two types of tubulin, which we call alpha tubulin and beta tubulin. However, it is doubtful that the two types can found in cells as individual proteins.

NOTES



Alpha and beta tubulin spontaneously bind one another to form a functional subunit that we call a heterodimer. A heterodimer is a protein that consists of two different gene products. The term is entirely descriptive - the prefix *hetero-* means "different," the prefix *di-* means "two," and the suffix *-mer* refers to a unit, in this case a single polypeptide. Obviously, cells do not continue to make tubulin (or any other protein) until they run out of resources. Some process must regulate the synthesis of tubulin. A common regulatory mechanism is feedback inhibition.



The figure illustrates the inhibition of tubulin synthesis by the presence of heterodimers in the system. Exactly how that inhibition takes place is irrelevant to this discussion. More about the important concept of feedback inhibition can be found elsewhere.

6.2.2 Aggregation of Microtubules

When intracellular conditions favor assembly, tubulin heterodimers assemble into linear protofilaments. Protofilaments in turn assemble into microtubules. All such assembly is subject to regulation by the cell.

NOTES

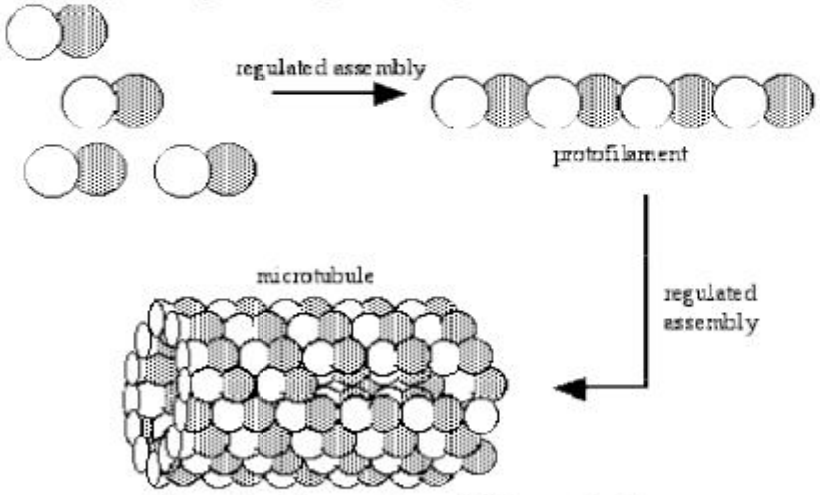


Fig. 6.4 : Formation of Microtubules

Microtubules form a framework for structures such as the spindle apparatus that appears during cell division, or the whiplike organelles known as cilia and flagella. Cilia and flagella are the well-studied models for microtubule structure and assembly, and are often used by textbooks to introduce microtubules.

6.2.3 Dynamic Instability of Microtubules

Under steady state conditions a microtubule may appear to be completely stable, however there is action taking place constantly. Populations of microtubules usually consist of some that are shrinking and some that are growing. A single microtubule can oscillate between growth and shortening phases. During growth, heterodimers are added on to the end of a microtubule, and during shrinkage they come off as intact subunits. The same heterodimer can come off and go back on.

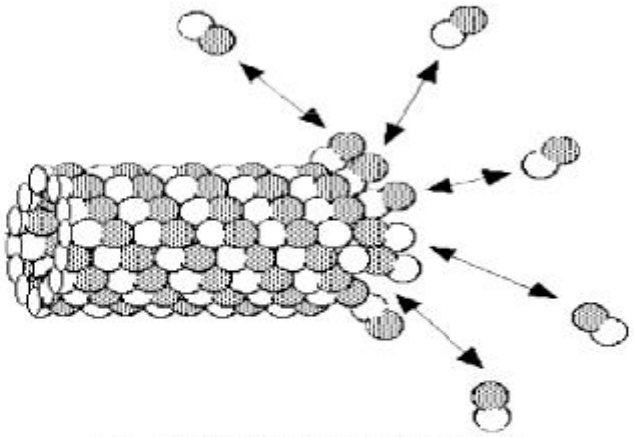


Fig. 6.5 : Microtubule Stability

Since even apparently stable microtubular structures have an intrinsic instability, they are considered to be in a dynamic equilibrium, or steady state. Look here to learn about the difference between a steady state and a true equilibrium.

6.2.4 Microtubules in Cilia and Flagella

To understand the regulation of microtubule assembly and function in any organism is a difficult task. To study microtubules in cells as complex vertebrate (e.g., human) cells is a nearly impossible task, without a few "hints" as to how to proceed. The basic mechanisms can be worked out using a much less complex biological model such as a flagellate. For example, the flagella of the photosynthetic protist *Chlamydomonas* are composed of microtubules, as are all flagella and cilia.

Cilia and flagella have the same basic structure. They are attached to structures known as basal bodies, which in turn are anchored to the cytoplasmic side of the plasma membrane. From the basal bodies the microtubule "backbone" extends, pushing the plasma membrane out with it.

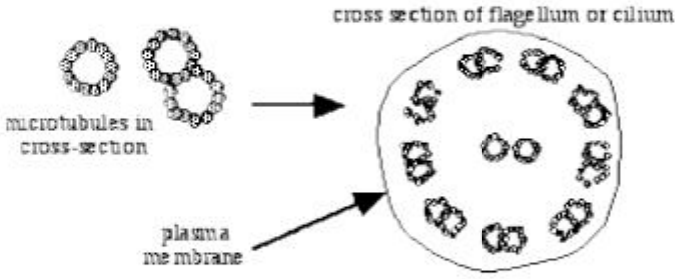


Fig. 6.6 : Microtubules in Flagella

NOTES

To form cilia or flagella, microtubules arrange themselves in a "9 + 2" array. Each of the two central microtubules consists of a single microtubule with 13 protofilaments arranged to form the wall of a circular tube. Each of the outer nine consists of a pair of microtubules that share a common wall (see the cross sections of microtubules in the figure). Look at the complete cross section carefully. The hair-like appearance of flagella and cilia in a light microscope is misleading. The entire structure lies within the cytoplasm of the cell.

The treatment given here to the structure of microtubules ignores their true complexity. Functional microtubules include associated proteins, anchoring sites and organizing centers, sites for enzyme activity, etc. In cilia and flagella, tubulin forms a core structure to which other proteins contribute structures called dynein arms, radial spokes, and nexin links. The arms, spokes, and links hold microtubules together and allow interaction between microtubules that is superficially similar to the sliding of actin and myosin filaments in muscle contraction.

6.2.5 Ciliary and Flagellar Motion

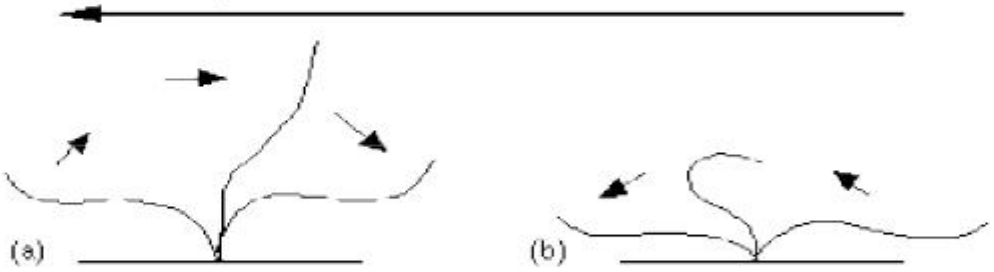
One might appreciate the complexity of microtubular organelles by looking at the motion of cilia and flagella. Despite the similarities in structure, the difference in nature of motility by flagella versus cilia is profound, as one can see by comparing representatives of the groups Ciliophora (the ciliates) and Mastigophora (the flagellates). Ciliates and flagellates behave differently, live in different habitats and occupy different niches, and likely represent two different evolutionary lineages. The main difference in function is in how they are organized.

Flagella are much longer than cilia and are usually present singly or in pairs. A single flagellum may propel the cell with a whip-like motion. A pair of flagella may move in a synchronized manner to pull the organism through the water, in a way similar to the breast stroke of a human swimmer.

Cilia tend to cover the surface area of a cell. Both cilia and flagella bend as the microtubules slide past one another. The arrangement of cilia permits their coordinated movement in response to signals from the cytoplasm. A small ciliate may have hundreds of individual cilia, all beating in a coordinated manner. How is all of the sliding and bending coordinated? How does the organism "decide" in what direction to move, or how to turn, rotate, or feed?

How does it convey the information to hundreds of cilia to bend in a certain way? Questions of that nature are fascinating to cell biologists. They are very difficult to address, because each system is so complex. Nevertheless, with a genome about a hundred times smaller than that of a human, a typical protist is much easier to study than a human cell.

NOTES



The "oar-like" motion of cilia involves a *power stroke* (a) followed by a *recovery stroke* (b). The large arrow indicates the direction of motion of the cell

Fig. 6.7 : Movement Presentation

The motion of an individual cilium or flagellum superficially resembles that of an oar, in that it sweeps through the medium with a power stroke that propels the cell. Each power stroke and return stroke involves perhaps thousands of chemical reactions. There may be dozens of strokes per second, and one action may involve thousands of cilia. You may notice that ciliates respond very quickly to obstacles or changes in their environment. It is fascinating to speculate on just how they receive information, process it, and deliver the signals to the cilia to produce precise movement. From the perspective of our relatively slow world, it is also difficult to comprehend how so much can go on in such a short time. Think of the effect on your perception of the universe if you were to shrink by several orders of magnitude.

6.2.6 Microtubules in Cell Division

Microtubules play very important role during cell division. When a cell is ready for division, the network present in cytoplasm disappears. The network disassembles and much of the resulting tubulin reassembles in the spindles of the dividing cell through organization of the microtubules. The mitotic spindle is a bipolar, self-organizing microtubule based machine that uses energy liberated from nucleotide hydrolysis to segregate sister chromatids accurately into the daughter cells during cell division.

NOTES

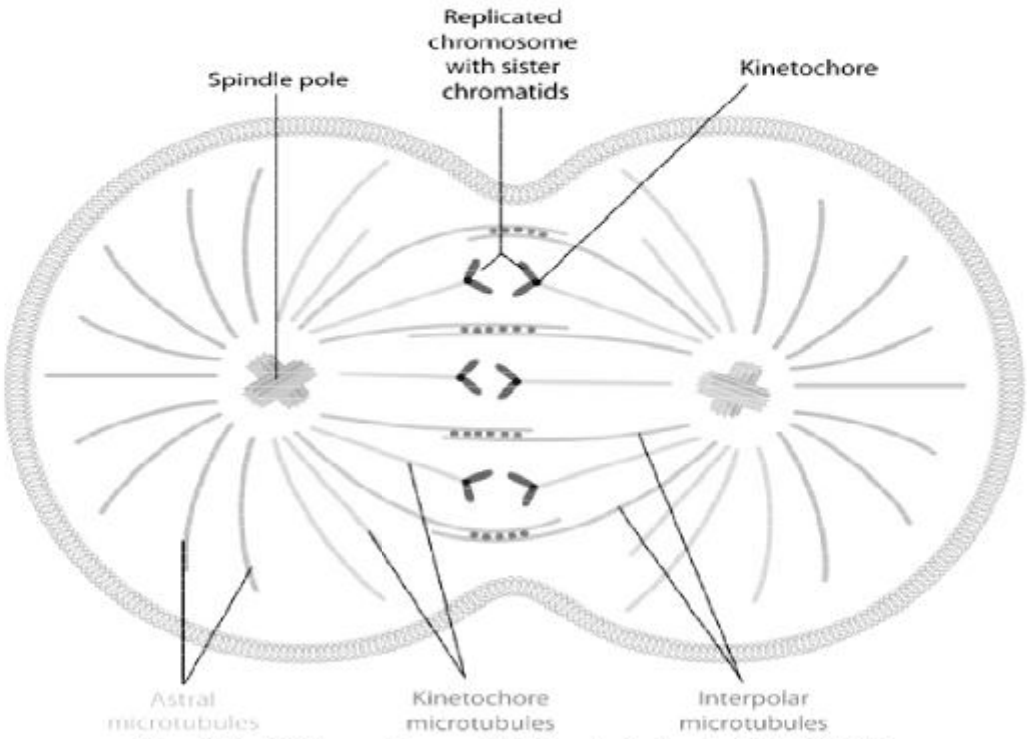


Fig. 6.8 : Different type of Microtubules in Cell Division

There are several theories proposed to explain chromosome movement. All these mechanism assume assembly and disassembly of microtubules during formation of spindle and their function. Assembly of microtubules at the pole may create a force sufficient for chromosome movement. This force may create by interaction of parallel microtubules on the spindles and cross bridging between parallel microtubules. The rate of depolymerisation of microtubules at the poles determines by the chromosome movement.

6.2.7 Microtubules and Motor Proteins

The cytoskeleton is a dynamic 3-dimensional structure that fills the cytoplasm, and is present in both eukaryotic and prokaryotic cells. The cytoskeleton acts as both muscle and skeleton, and plays a role in cell protection, cell motility (migration), cytokinesis, intracellular transport, cell division and the organization of the organelles within the cell. Motor proteins are responsible for the active transport of most proteins and vesicles in the cytoplasm. They are a class of molecular motors that are able to move along the surface of a suitable substrate, powered by the hydrolysis of ATP. There are three superfamilies of cytoskeleton motor proteins. Myosin motors act upon actin filaments to generate cell surface contractions and other morphological changes, as well as

vesicle motility, cytoplasmic streaming and muscle cell contraction. The kinesin/kinectin and dynein/dynectin microtubule based motor superfamilies move vesicles and organelles within cells, cause the beating of flagella and cilia, and act within the mitotic and meiotic spindles to segregate replicated chromosomes.

NOTES

The general theories of how microtubules and their associated proteins help in the distribution of intracellular structures seem to be fairly well understood. For example, it is known that microtubules are polar polymers, with their minus ends and plus ends. The centre of the spindle are located the plus ends and the two poles of spindles are located the minus ends. Thus the microtubule provides a directional track for the movement of chromosomes and other cellular organelles, for which the necessary force is derived from the motor protein.

6.3 Microfilaments

Microfilaments are the thinnest cytoskeleton component and are composed of the globular protein. Actin proteins associate in a head to tail fashion to form long chains called "microfilaments". When microfilaments associate they form a twisted double chain. When these chains associate in parallel they are referred to as Actin stress fibre. Microfilaments are linear polymers of the globular protein actin. The globular monomer (G-actin) polymerizes into filamentous (F-) actin which appears in electron micrographs as two right-handed helices wound around each other with a repeat distance of approximately 36 nm. The polyelectrolyte nature of F-actin is likely to be relevant to its interaction with membrane phospholipids. Actin filaments have a diameter of 7 nm and can reach lengths of 30–100 _m *in vitro*, and at least several microns *in vivo*. Because these filaments are so long, they form semi-dilute solutions at extremely low volume fraction (< 0.05%) in which rotational diffusion of the filaments is greatly retarded due to solute-solute interactions.

Actin participates in many important cellular processes, including muscle contraction, cell motility, cell division and cytokinesis, vesicle and organelle movement, cell signalling, and the establishment and maintenance of cell junctions and cell shape. Many of these processes are mediated by extensive and intimate interactions of actin with cellular membranes. In vertebrates, three

main groups of actin isoforms, alpha, beta, and gamma have been identified. The alpha actins, found in muscle tissues, are a major constituent of the contractile apparatus. The beta and gamma actins coexist in most cell types as components of the cytoskeleton, and as mediatorsof internal cell motility.

NOTES

A cell's ability to dynamically form microfilaments provides the scaffolding that allows it to rapidly remodel itself in response to its environment or to the organism's internal signals, for example, to increase cell membrane absorption or increase cell adhesion in order to form cell tissue. Other enzymes or organelles such as cillia can be anchored to this scaffolding in order to control the deformation of the external cell membrane, which allows endocytosis and cytokinesis. It can also produce movement either by itself or with the help of molecular motors. Actin therefore contributes to processes such as the intracellular transport of vesicles and organelles as well as muscular contraction and cellular migration. It therefore plays an important role in embryogenesis, the healing of wounds and the invasivity of cancer cells. The evolutionary origin of actin can be traced to prokaryotic cells, which have equivalent proteins. Lastly, actin plays an important role in the control of gene expression.

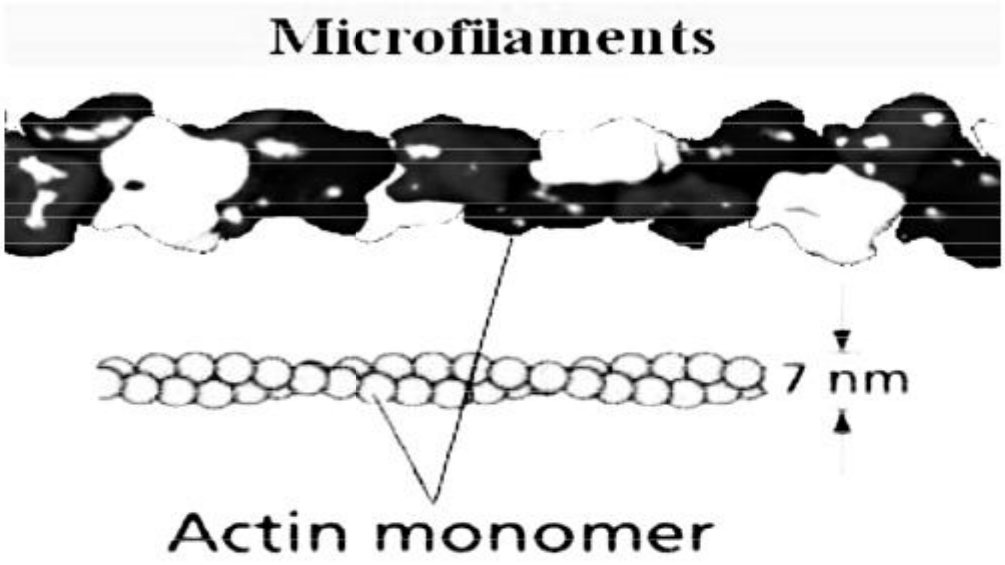


Fig. 6.9 A: Microfilaments

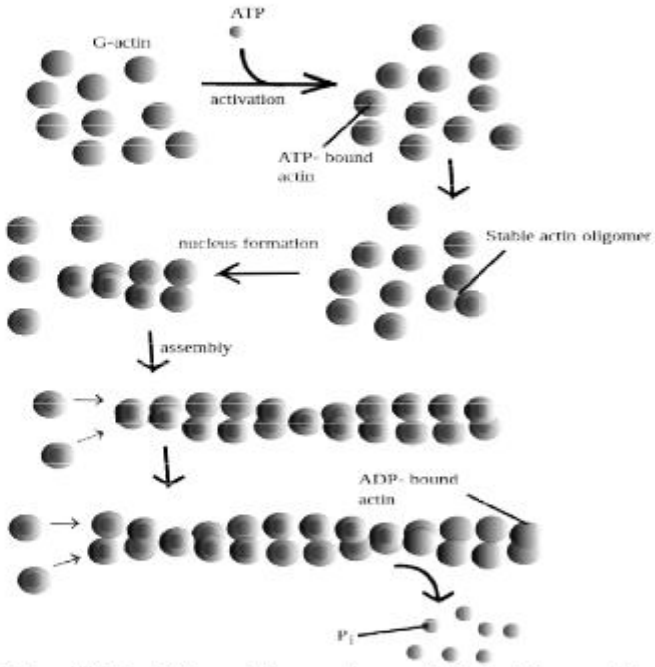


Fig. 6.9 B: Microfilaments and their Formation

6.3.1 Actin filament in muscle cells

Our muscles work in a similar fashion. Muscles are composed of two major protein filaments: a thick filament composed of the protein myosin and a thin filament composed of the protein actin. Muscle contraction occurs when these filaments slide over one another in a series of repetitive events. Let's see how myosin molecules play a role similar to the oars of a rower. Myosin is a motor protein that generates the force in a muscle contraction much like the stroke of an oar. It consists of a head and a tail region. Together, the tails of approximately three hundred myosin molecules form the shaft of the thick filament. The myosin heads of these molecules project outward toward the thin filaments like the oars of a rowboat. Actin is a spherical protein that forms, among other things, the thin filament in muscle cells. Thin filaments are composed of two long chains of these actin molecules that are twisted around one another. Each actin molecule has a myosin-binding site where a myosin head can bind. Let's consider the organization of myosin and actin in skeletal muscle, the muscles responsible for voluntary movements. Skeletal muscle is composed of a repeating structure of myosin and actin fibers. Each myosin thick filament is surrounded by actin thin filaments, and each thin filament is surrounded by thick filaments. Several of these filament bundles make up the

NOTES

functional portion of a muscle cell. It is also important to note that the thin filaments are attached to a very dense protein plate called a Z disc.

NOTES

So how do the thick and thin filaments generate muscle contraction? The attraction between the myosin head and the myosin-binding site of actin are strong enough that the bond can form spontaneously. Once the two proteins are bound, the myosin protein undergoes a conformational change, or a change in protein shape, that 'cocks' the head. Like the oar stroke of a rower, the movement of the myosin head causes the thin filament to move. This is where the Z disc comes into play. Without the anchoring the Z disc provides, the thick filaments would simply slide past the thin filaments in opposite directions. By anchoring the thin filaments in place, movement generated by the myosin heads instead causes the muscle cell to contract and, by extension, cause our body to move.

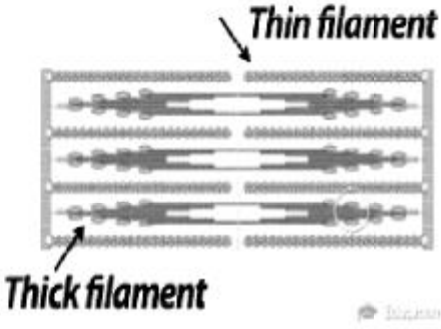


Fig. : 6.10 Thick and Thin Filaments in a Muscle Contraction

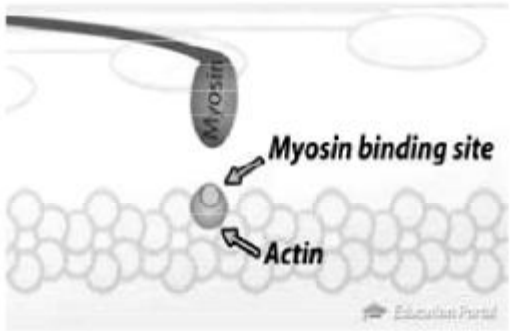


Fig. 6.11 : Illustration of the Myosin Binding Site

6.4 Intermediate Filaments (IF)

Intermediate filaments are major fibrous proteins that are found in the nucleoplasm and the cytoplasm of most types of animal cells. In the cytoplasm, they are usually organized into complex arrays of 10-nm diameter cytoskeleton IF that are prevalent in the perinuclear region, where they frequently form a cage that surrounds and appears to position the nucleus. In this region, individual IF appears to be attached to the outer nuclear envelope membrane or to nuclear pore complexes. Continuous with this perinuclear array, IF extend radially through the cytoplasm, eventually forming close associations with the cell surface. Frequently, these membrane associations are concentrated in

regions containing desmosomes, hemidesmosomes, and other types of adhesion sites.

The overall organization of cytoskeletal IF networks, as well as their *in vitro* properties, suggests that they link the nuclear and cell surfaces and that they are involved in numerous cell functions, including the maintenance of the overall integrity of cytoplasm and cell shape. These latter roles are supported by the finding that some blistering skin diseases are caused by point mutations in keratins that have been related to structural alterations in IF, changes in the shape of individual keratinocytes, and the loss of mechanical properties of skin. The resistance to break- age of vimentin IF under conditions of high strain in vitro also supports their role in providing cells with resilience to mechanical stress in vivo. Based on their mechanical properties, as well as their in-solubility *in vitro*, IF have long been thought to be very stable cytoskeletal elements relative to other cytoskeletal systems such as microtubules or microfilaments. However, it has recently been shown that, at least in actively growing cultured cells, IF are in fact very dynamic structures *in vivo*. For example, at the posttranslational level, polymerized vimentin and keratin IF can incorporate microinjected unpolymerized subunits in living fibroblasts and epithelial cells. Experiments also demonstrate that steady state equilibrium exists between soluble IF subunits and polymerized IF in several cell types. These in vivo studies of IF dynamics have also been supported by fluorescence energy transfer experiments carried out in vitro. Furthermore, it has been shown that IF exhibit rapid organizational and biochemical changes following various stimuli such as heat shock, phosphorylation after exposure to growth factors, and the transient association of IF with protein kinase C in response to a signal at the cell surface. These findings have implicated IF in dynamic roles in several physiological activities, including signal transduction. Numerous other properties of IF also reflect their dynamic nature. For example, biochemical- and metabolic labelling studies suggest that IF protein synthesis occurs continuously and that newly synthesized subunits are incorporated into IF at both the co- and posttranslational levels. However, the pools of soluble subunits, whether they are newly synthesized or not, are very small, as indicated by the fact that the vast majority of IF protein is palatable in cell-free preparations. Other evidence supporting the dynamic nature of IF assemblies in

NOTES

NOTES

vivo come from transient transfection studies that demonstrate that the insertion of mutant polypeptide chains into endogenous polymerized networks have disruptive effects leading to their disassembly. There is considerable evidence in support of a role for phosphorylation in regulating at least some of the dynamic properties of IF in vivo. This evidence includes the finding that the hyperphosphorylation of IF proteins is accompanied frequently by their disassembly. For example, during mitosis in many types of eukaryotic cells, nuclear breakdown is accompanied by the hyper- phosphorylation and coincident disassembly of the Type V IF polymer, the nuclear lamins.

In cells, it has been shown that vimentin IF disassemble during mitosis and that this alteration in the state of polymerization is regulated by phosphorylation. In other cell types, a more localized disassembly of cytoskeletal IF due to phosphory- lation appears to take place in the cleavage furrow during cytokinesis. In addition, it has been shown that vimentin IF are rapidly hyperphosphory- lated and disassembled in cells exposed to phosphatase inhibitors, suggesting that IF assembly states in interphase cells are regulated by kinase/phosphatase equilibria. In vitro studies support these in vivo ob- servations, as numerous kinases phosphorylate and disas- semble IF in cell-free preparations.

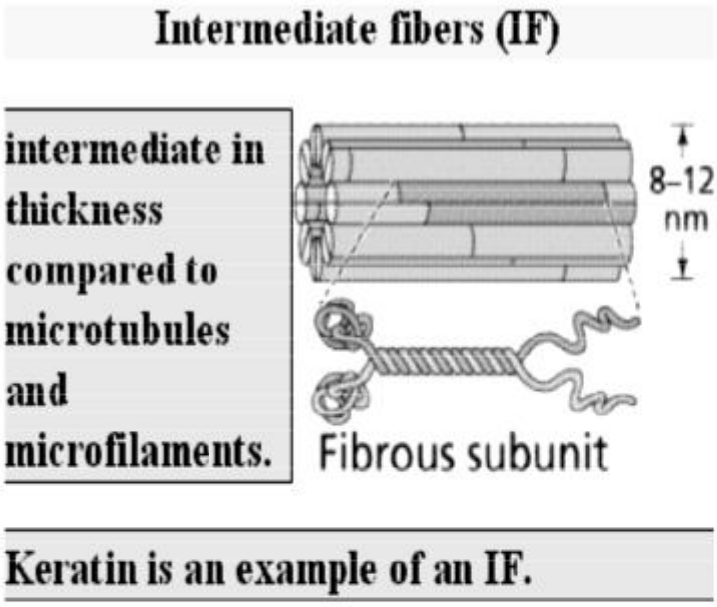


Fig. 6.12: Intermediate Filaments

6.5 Summary

This unit will attempt to summarize current understanding of the biophysical properties of the cytoskeleton. The interplay between the three-dimensional protein networks called the cytoskeleton. Nearly all cells contain some form of filamentous protein polymer in their interior. Eukaryotic cells often possess an intricate three-dimensional network of interpenetrating and partly crosslinked filaments whose length, concentration, and attachments to each other or to the cell membrane are remodelled in response to specific signals. The cytoskeleton of a cell is responsible for structure, strength, and motility within and between the cells. It provides cellular scaffolding upon which the cellular organization is arranged. The filaments in this skeleton are composed of three types of protein polymer–microfilaments, intermediate filaments and microtubules – and the many accessory proteins that bind to them. Cytoskeleton composes by three types of components: 1. Microfilaments 2. Microtubules and 3. Intermediate fibres (IF).

Microtubules are filamentous intracellular structures that are responsible for various kinds of movements in all eukaryotic cells. Microtubules are involved in nucleic and cell division, organization of intracellular structure, and intracellular transport, as well as cilliary and flagellar motility. Microfilaments are the thinnest cytoskeleton component and are composed of the globular protein. Actin proteins associate in a head to tail fashion to form long chains called "microfilaments". When microfilaments associate they form a twisted double chain. When these chains associate in parallel they are referred to as Actin stress fibre. Intermediate filaments are major fibrous proteins that are found in the nucleoplasm and the cytoplasm of most types of animal cells. In the cytoplasm, they are usually organized into complex arrays of 10-nm diameter cytoskeleton IF that are prevalent in the perinuclear region, where they frequently form a cage that surrounds and appears to position the nucleus.

NOTES

6.6 Glossary

- **Basal bodies** :Short cylindrical array of flagella/cilia.
- **Cytoplasm** : Gel like substance enclosed by cell membrane.
- **Desmosome** : Cell structure specialized for cell to cell adhesion.

NOTES

- **Endocytosis**: Process by which cell absorbs molecules.
- **Eukaryotic**: Cell contain nucleus and other membrane bound organelles.
- **In-vitro** : Biological molecule studies in artificial culture medium.
- **Meiosis**: Cell division in which chromosome number reduced to half of original number.
- **Mitosis**: Cell division in which chromosome number remain to same of original number.
- **Photosynthesis**: Process to convert light energy into chemical energy by the green plants
- **Tinsel Flagella**:Flagella having numerous hairlike projections.
- **Whiplash Flagella**: Flagellum are smooth without any projections.

6.7 Self-Learning Exercise

Section – A (Very Short Answer Type Questions)

1. Give the name of cytoskeleton components.
2. Who maintain the shape of cell?
3. Give the name of thinnest filament of cytoskeleton.
4. Give the example of intermediate filament.

Section – B (Short Answer Type Questions)

1. Write the components of microtubules.
2. Describe the role of actin filaments in muscle cells.
3. What is a correlation between microtubule and motor protein?

Section – C (Long Answer Type Questions)

1. Give the detail structure and organization of microtubules.
2. Describe microfilaments.
3. Give the illustrative structure of intermediate filaments.

Section – A (Answer Key)

1. Microtubule, Microfilament and intermediate filament.
2. Cytoskeleton.
3. Microfilaments.
4. Keratin.

NOTES

6.8 References

- Amos, L.A. and W.B. Amos, 1991, Molecules of the Cytoskeleton (Guilford Press, New York).
- Bershadsky, A.D. and J.M. Vasiliev, 1988. Cytoskeleton (Plenum, New York).
- Carlier, M., 1989, Role of nucleotide hydrolysis in the dynamics of actin filaments and microtubules, Int. Rev. Cytol. **115**, 139–170.
- Carraway, V. and C.A.C. Carraway, 1992, The Cytoskeleton: A Practical Approach (IRL Press,
- Foisner, R. and T. Wiche, 1991, Intermediate filament-associated proteins, Curr. Opinion in Cell Biol. **3**, 75–81.
- Gittes, F., B. Mickey, J. Nettleton and J. Howard, 1993, Flexural rigidity of microtubules and actin filaments measured from thermal fluctuations in shape, J. Cell Biol. **120**, 923–934.
- Goldman, R. and P. Steiner, 1990, Cellular and Molecular Biology of Intermediate Filaments (Plenum, New York).
- Holmes, K., D. Popp, W. Gebhard and W. Kabsch, 1990, Atomic model of the actin filament, Nature **347**, 44–49.
- Janmey, P., C. Cunningham, G. Oster and T. Stossel, 1992, Cytoskeletal networks and osmotic pressure in relation to cell structure and motility, in: Swelling Mechanics: From Clays to Living Cells and Tissues, ed. T. K. Karalis (Springer, Heidelberg) pp. 333–346.

NOTES

- Janmey, P.A., 1991, Mechanical properties of cytoskeletal polymers, Curr. Opinion Cell Biol. **3**, 4–11.
- Janmey, P.A., J. Peetermans, K.S. Zaner, T.P. Stossel and T. Tanaka, 1986, Structure and mobility of actin filaments as measured by quasielastic light scattering, viscometry, and electron microscopy, J. Biol. Chem. **261**, 8357–8362.
- Kleinfeld, A., 1987, Current views of membrane structure, in: Membrane Structure and Function, Vol. 29, eds R. Klausner, C. Kempf and J. van Renswoude (Academic Press, San Diego) pp. 1–25.
- Luna, E.J. and A.L. Hitt, 1992, Cytoskeleton-plasma membrane interactions, Science **258**, 955–964.
- Malev, V., D. Gromov, Y.Y. Komissarchik and M. Brudnaya, 1992, Interfacial models of nerve fiber cytoskeleton, Biophys. J. **63**, 1101–1117.
- McLaughlin, S., 1989, The elctrostatic properties of membranes, Annu. Rev. Biophys. Biophys. Chem. **18**, 113–136.
- Miyamoto, H. and H. Hotani, 1988, Polymerization of microtubules within liposomes produces morpholigical change of their shapes, in: Dynamics of Microtubules, ed. H. Hotani (Taniguchi Foundation).
- Mooseker, M. and J. Morrow, 1991, Ordering the Membrane-Cytoskeleton Trilayer (Academic Press, San Diego) p. 245.
- Olmsted, J., 1986, Microtubule-associated proteins, Annu. Rev. Cell Biol. **2**, 421–458.

Unit-7

Cell Cycle and Apoptosis

NOTES

Structure of the Unit

NOTES

7.0 Objectives

After going through this unit you will be able to understand:

- Cell Cycle and Apoptosis
- Control Mechanism
- Cyclin-Cyclin Dependent Kinase
- Retinoblastoma
- E2F Protein
- Programmed Cell Death

7.1 Introduction

Cell proliferation is the process whereby cells reproduce themselves by growing and then dividing into two equal copies. All organisms, even the largest, start their life from a single cell. Growth and reproduction are characteristics of cells, indeed of all living organisms. All cells reproduce by dividing into two, with each parental cell giving rise to two daughter cells each time they divide. These newly formed daughter cells can themselves grow and divide; giving rise to a new cell population that is formed by the growth and division of a single parental cell and its progeny. In other words, such cycles of growth and division allow a single cell to form a structure consisting of millions of cells.

Cell division occurs by an elaborate series of events, whereby chromosomes and other components are duplicated and evenly distributed into two daughter cells. It is a highly ordered and tightly regulated process that causes an irreversible and unidirectional change in cell state. Under appropriate environmental conditions, unicellular organisms like yeast exist perpetually in the cell cycle, constantly growing and dividing. Although the same occurs with some cancer cells, most cells in multicellular organisms are not cycling. A

significant fraction of these cells, including those that have already differentiated, are no longer (generally) capable of proliferation.

7.2 Cell Cycle

NOTES

Cell division is a very important process in all living organisms. During the division of a cell, DNA replication and cell growth also take place. All these processes, i.e., cell division, DNA replication, and cell growth, hence, have to take place in a coordinated way to ensure correct division and formation of progeny cells containing intact genomes. The sequence of events by which a cell duplicates its genome, synthesizes the other constituents of the cell and eventually divides into two daughter cells is termed cell cycle. Although cell growth (in terms of cytoplasmic increase) is a continuous process, DNA synthesis occurs only during one specific stage in the cell cycle. The replicated chromosomes (DNA) are then distributed to daughter nuclei by a complex series of events during cell division. These events are themselves under genetic control.

Fig. 7.1 : Organization of Cell Cycle

The eucaryotic cell cycle is divided into four stages: G_1, S, G_2, and M. G_1 is the gap phase during which cells prepare for the process of DNA replication. It is during the G_1 phase that the cell integrates mitogenic and growth inhibitory signals and makes the decision to proceed, pause, or exit the cell cycle. An important checkpoint in G_1 has been identified in both yeast and mammalian cells. Referred to as start in yeast and the

NOTES

restriction point in mammalian cells, this is the point at which the cell becomes committed to DNA replication and completing a cell cycle (1–3). S phase is defined as the stage in which DNA synthesis occurs. G_2 is the second gap phase during which the cell prepares for the process of division. M stands for mitosis, the phase in which the replicated chromosomes are segregated into separate nuclei and cytokinesis occurs to form two daughter cells. In addition to G_1, S, G_2, and M, the term G_0 is used to describe cells that have exited the cell cycle and become quiescent.

7.2.1 Mitosis

A typical eukaryotic cell divide once in approximately every 24 hours. However, this duration of cell cycle can vary from organism to organism and also from cell type to cell type. Yeast for example, can progress through the cell cycle in only about 90 minutes. The cell cycle is divided into two basic phases: Interphase and M Phase (Mitosis phase). The M Phase represents the phase when the actual cell division or mitosis occurs and the interphase represents the phase between two successive M phases. It is significant to note that in the 24 hour average duration of cell cycle, cell division proper lasts for only about an hour. The interphase lasts more than 95% of the duration of cell cycle. The M Phase starts with the nuclear division, corresponding to the separation of daughter chromosomes (karyokinesis) and usually ends with division of cytoplasm (cytokinesis).

7.2.2 Interphase

The interphase, though called the resting phase, is the time during which the cell is preparing for division by undergoing both cell growth and DNA replication in an orderly manner. The interphase is divided into three further phases: G_1 phase (Gap_1), S phase (Synthesis), and G_2 phase (Gap_2).

G_1 phase corresponds to the interval between mitosis and initiation of DNA replication. During G_1 phase the cell is metabolically active and continuously grows but does not replicate its DNA. S or synthesis phase marks the period during which DNA synthesis or replication takes place. During this time the amount of DNA per cell doubles. If the initial amount of DNA is denoted as 2C then it increases to 4C. However, there is no increase in the chromosome

number; if the cell had diploid or 2n number of chromosomes at G_1, even after S phase the number of chromosomes remains the same, i.e., 2n. In animal cells, during the S phase, DNA replication begins in the nucleus, and the centriole duplicates in the cytoplasm. During the G_2 phase, proteins are synthesized in preparation for mitosis while cell growth continues. Some cells in the adult animals do not appear to exhibit division (e.g., heart cells) and many other cells divide only occasionally, as needed to replace cells that have been lost because of injury or cell death. These cells that do not divide further exit G_1 phase to enter an inactive stage called quiescent stage (G_0) of the cell cycle. Cells in this stage remain metabolically active but no longer proliferate unless called on to do so depending on the requirement of the organism. In animals, mitotic cell division is only seen in the diploid somatic cells. Against this, the plants can show mitotic divisions in both haploid and diploid cells. From your recollection of examples of alternation of generations in plants identify plant species and stages at which mitosis is seen in haploid cells.

NOTES

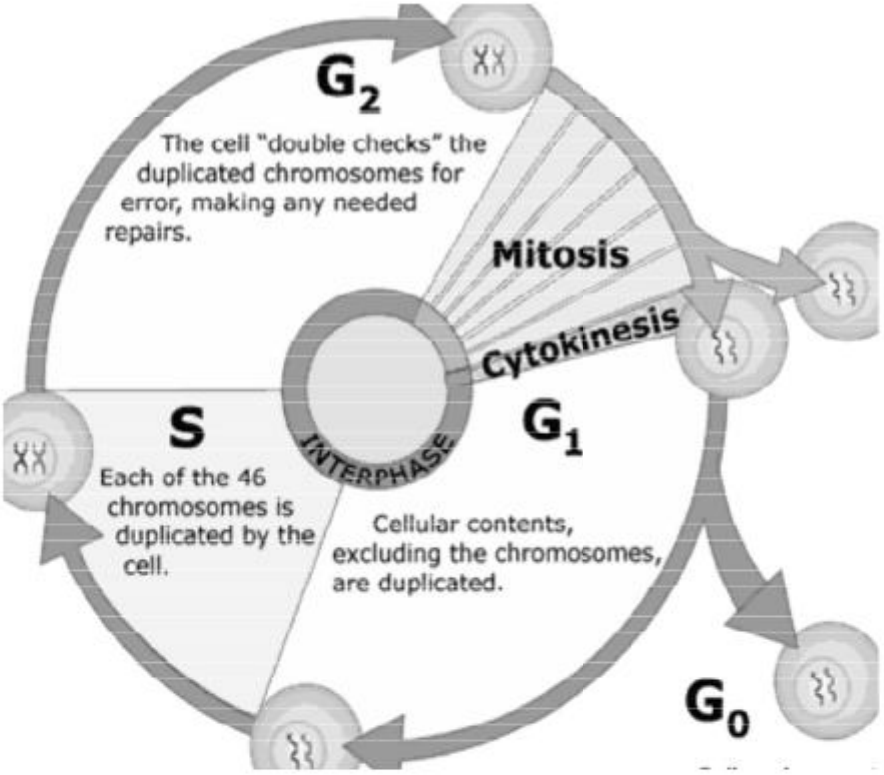


Fig. 7.2 : Different Phases of Cell Cycle

7.2.3 M Phase

NOTES

This is the most dramatic period of the cell cycle, involving a major reorganization of virtually all components of the cell. Since the number of chromosomes in the parent and progeny cells is the same, it is also called as equational division. Though for convenience mitosis has been divided into four stages of nuclear division, it is very essential to understand that cell division is a progressive process and very clear-cut lines cannot be drawn between various stages. Mitosis is divided into the following four stages:

Prophase

Prophase occupies over half of mitosis. The nuclear membrane breaks down to form a number of small vesicles and the nucleolus disintegrates. A structure known as the centrosome duplicates itself to form two daughter centrosomes that migrate to opposite ends of the cell. The centrosomes organise the production of microtubules that form the spindle fibres that constitute the mitotic spindle. The chromosomes condense into compact structures. Each replicated chromosome can now be seen to consist of two identical chromatids (or sister chromatids) held together by a structure known as the centromere.

Prometaphase

The chromosomes, led by their centromeres, migrate to the equatorial plane in the midline of cell - at right-angles to the axis formed by the centrosomes. This region of the mitotic spindle is known as the metaphase plate. The spindle fibres bind to a structure associated with the centromere of each chromosome called a kinetochore. Individual spindle fibres bind to a kinetochore structure on each side of the centromere. The chromosomes continue to condense.

Metaphase

The chromosomes align themselves along the metaphase plate of the spindle apparatus.

Anaphase

The shortest stage of mitosis. The centromeres divide, and the sister chromatids of each chromosome are pulled apart - or 'disjoin' - and move to the opposite ends of the cell, pulled by spindle fibres attached to the kinetochore regions.

The separated sister chromatids are now referred to as daughter chromosomes. (It is the alignment and separation in metaphase and anaphase that is important in ensuring that each daughter cell receives a copy of every chromosome.)

Telophase

NOTES

The final stage of mitosis and a reversal of many of the processes observed during prophase. The nuclear membrane reforms around the chromosomes grouped at either pole of the cell, the chromosomes uncoil and become diffuse, and the spindle fibres disappear.

Cytokinesis

Mitosis accomplishes not only the segregation of duplicated chromosomes into daughter nuclei (karyokinesis), but the cell itself is divided into two daughter cells by a separate process called cytokinesis at the end of which cell division is complete. In an animal cell, this is achieved by the appearance of a furrow in the plasma membrane. The furrow gradually deepens and ultimately joins in the centre dividing the cell cytoplasm into two. Plant cells however, are enclosed by a relatively inextensible cell wall, thererfore they undergo cytokinesis by a different mechanism. In plant cells, wall formation starts in the centre of the cell and grows outward to meet the existing lateral walls. The formation of the new cell wall begins with the formation of a simple precursor, called the cell-plate that represents the middle lamella between the walls of two adjacent cells. At the time of cytoplasmic division, organelles like mitochondria and plastids get distributed between the two daughter cells. In some organisms karyokinesis is not followed by cytokinesis as a result of which multinucleate.

7.3 Significance of Mitosis

Mitosis or the equational division is usually restricted to the diploid cells only. However, in some lower plants and in some social insects haploid cells also divide by mitosis. It is very essential to understand the significance of this division in the life of an organism. Are you aware of some examples where you have studied about haploid and diploid insects? Mitosis results in the production of diploid daughter cells with identical genetic complement usually. The growth of multicellular organisms is due to mitosis. Cell growth results in

disturbing the ratio between the nucleus and the cytoplasm. It therefore becomes essential for the cell to divide to restore the nucleo-cytoplasmic ratio. A very significant contribution of mitosis is cell repair. The cells of the upper layer of the epidermis, cells of the lining of the gut, and blood cells are being constantly replaced. Mitotic divisions in the meristematic tissues – the apical and the lateral cambium, result in a continuous growth of plants throughout their life. condition arises leading to the formation of syncytium

NOTES

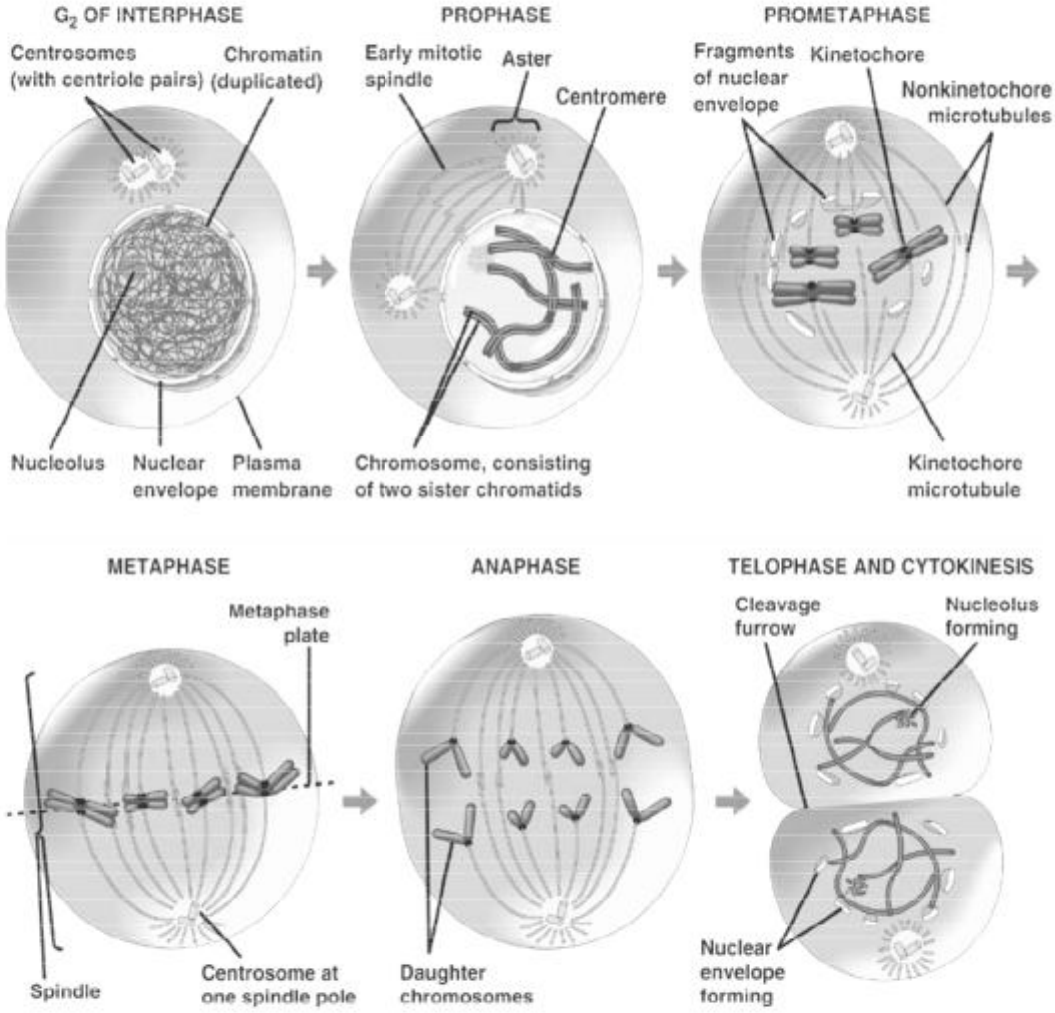


Fig. 7.3 : Cell Division- Mitosis

7.4 Meiosis

NOTES

The production of offspring by sexual reproduction includes the fusion of two gametes, each with a complete haploid set of chromosomes. Gametes are formed from specialized diploid cells. This specialized kind of cell division reduces the chromosome number by half results in the production of haploid daughter cells. This kind of division is called meiosis. Meiosis ensures the production of haploid phase in the life cycle of sexually reproducing organisms whereas fertilization restores the diploid phase. We come across meiosis during gametogenesis in plants and animals. This leads to the formation of haploid gametes. The key features of meiosis are involves two sequential cycles of nuclear and cell division called meiosis I and meiosis II but only a single cycle of DNA replication. Meiosis I is initiated after the parental chromosomes have replicated to produce identical sister chromatids at the S phase. Meiosis involves pairing of homologous chromosomes and recombination between them. Four haploid cells are formed at the end of meiosis II.

7.4.1 Meiosis I

Prophase I

Prophase of the first meiotic division is typically longer and more complex when compared to prophase of mitosis. It has been further subdivided into the following five phases based on chromosomal behaviour, i.e., Leptotene, Zygotene, Pachytene, Diplotene and Diakinesis. During leptotene stage the chromosomes become gradually visible under the light microscope. The compaction of chromosomes continues throughout leptotene. This is followed by the second stage of prophase I called zygotene. During this stage chromosomes start pairing together and this process of association is called synapsis. Such paired chromosomes are called homologous chromosomes. Electron micrographs of this stage indicate that chromosome synapsis is accompanied by the formation of complex structure called synaptonemal complex. The complex formed by a pair of synapsed homologous chromosomes is called a bivalent or a tetrad. However, these are more clearly visible at the next stage. The first two stages of prophase I are relatively short-lived compared to the next stage that is pachytene. During these stage bivalent

NOTES

chromosomes now clearly appears as tetrads. This stage is characterised by the appearance of recombination nodules, the sites at which crossing over occurs between non-sister chromatids of the homologous chromosomes. Crossing over is the exchange of genetic material between two homologous chromosomes. Crossing over is also an enzyme-mediated process and the enzyme involved is called recombinase. Crossing over leads to recombination of genetic material on the two chromosomes. Recombination between homologous chromosomes is completed by the end of pachytene, leaving the chromosomes linked at the sites of crossing over. The beginning of diplotene is recognised by the dissolution of the synaptonemal complex and the tendency of the recombined homologous chromosomes of the bivalents to separate from each other except at the sites of crossovers. These X-shaped structures, are called chiasmata. In oocytes of some vertebrates, diplotene can last for months or years. The final stage of meiotic prophase I is diakinesis. This is marked by terminalisation of chiasmata. During this phase the chromosomes are fully condensed and the meiotic spindle is assembled to prepare the homologous chromosomes for separation. By the end of diakinesis, the nucleolus disappears and the nuclear envelope also breaks down. Diakinesis represents transition to metaphase.

Metaphase I

The bivalent chromosomes align on the equatorial plate. The microtubules from the opposite poles of the spindle attach to the pair of homologous chromosomes.

Anaphase I

The homologous chromosomes separate, while sister chromatids remain associated at their centromeres.

Telophase I

The nuclear membrane and nucleolus reappear, cytokinesis follows and this is called as diad of cells. Although in many cases the chromosomes do undergo some dispersion, they do not reach the extremely extended state of the interphase nucleus. The stage between the two meiotic divisions is called interkinesis and is generally short lived. Interkinesis is followed by prophase II, a much simpler prophase than prophase I.

7.4.2 Meiosis II

Prophase II

Meiosis II is initiated immediately after cytokinesis, usually before the chromosomes have fully elongated. In contrast to meiosis I, meiosis II resembles a normal mitosis. The nuclear membrane disappears by the end of prophase II. The chromosomes again become compact.

NOTES

Metaphase II

At this stage the chromosomes align at the equator and the microtubules from opposite poles of the spindle get attached to the kinetochores of sister chromatids.

Anaphase II

It begins with the simultaneous splitting of the centromere of each chromosome (which was holding the sister chromatids together), allowing them to move toward opposite poles of the cell.

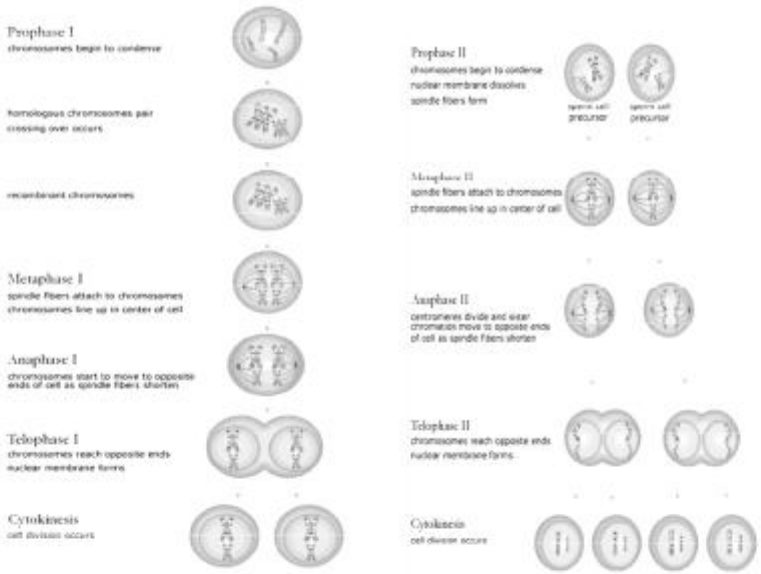


Fig. 7.4 : Cell Division- Meiosis I & II

Telophase II

Meiosis ends with telophase II, in which the two groups of chromosomes once again get enclosed by a nuclear envelope; cytokinesis follows resulting in the formation of tetrad of cells i.e., four haploid daughter cells.

7.5 Significance of Meiosis

NOTES

Meiosis is the mechanism, by which conservation of specific chromosome number of each species is achieved across generations in sexually reproducing organisms, even though the process, per se, paradoxically, results in reduction of chromosome number by half. It also increases the genetic variability in the population of organisms from one generation to the next. Variations are very important for the process of evolution.

7.6 Cell Cycle Control

The cell cycle is controlled by numerous mechanisms ensuring correct cell division. Strict cell cycle regulation is vital to ensure faithful segregation of genetic material and thereby allow normal development and maintenance of multicellular organisms. Failure to coordinate such processes leads to genome instability, often associated with birth defects and cancer. Accordingly, cells have checkpoint controls ensuring that the correct sequence of events is firmly maintained. Checkpoints consist of at least three components: a sensor, that detects the error, a signal, generated by the sensor via a signal transduction pathway and, finally, a response element in the cell cycle machinery to block cell cycle progression . Three checkpoints have been amply documented: the DNA damage checkpoint, which arrests cells in G_1, S phase, G_2 or even mitosis in case of DNA lesions; the DNA replication checkpoint, which ensures that mitosis is not initiated until DNA replication is complete and also that no DNA is replicated twice; and, the Spindle Assembly Checkpoint (SAC), which delays anaphase onset until all chromosomes are properly attached to the mitotic spindle.

7.6.1 Cyclin-Cyclin-Dependent Kinases

The core components of the cell cycle control system are a family of enzymes, cyclin-dependent kinases (Cdks) and the corresponding cyclin regulatory subunits. Like other kinases, Cdks act by catalysing the covalent attachment of phosphate groups derived from ATP (adenosine triphosphate) to protein substrates, therefore, inducing changes in the enzymatic activity or capacity of the substrate to interact with other proteins. Cdks are activated by direct binding to its regulatory subunits cyclins. Cyclins were discovered as proteins

whose levels oscillate in synchrony with the cell cycle. The first cyclins to be identified were cyclin A and B, since they showed progressive accumulation through interphase, reaching its maximum during mitosis and abruptly disappearing upon mitotic exit. The original *mitotic cyclins* were first described in *Xenopus* egg extracts as one of the components of the maturation-promoting factor (MPF), later found to be an heterodimer composed of Cdk1 and cyclin B. The cyclin-Cdk complexes and other regulators that drive the cell cycle are assembled into a highly interconnected regulatory system, whose effectiveness is enhanced by a number of important features. Firstly, the cell-cycle control system includes feedback loops and other regulatory interactions that lead to irreversible, switch-like activation/inactivation of cyclin-Cdk complexes. Therefore, cell cycle events are triggered in an all-or-nothing fashion, avoiding damages resulting from events only partly initiated. Secondly, regulatory interactions between different cyclin-Cdk complexes guarantee that the proper order is followed and coordinated with each other. Thirdly, the control system is highly robust: the activation or inactivation of any cyclin-Cdk switch is governed by several mechanisms, conferring redundancy to the system and allowing it to operate well under a variety of conditions even if some components fail. As a result, the system becomes quite adaptable and allows the timing of the major regulatory switches to be adjusted by regulatory inputs from various sources. Additionally, different types of cyclins are produced at different points of cell cycle progression, establishing a periodicity that rules the formation of different cyclin-Cdk complexes, which in turn set off different cell cycle events. The control system works in such way that each cyclin-Cdk complex promotes the activation of the next in sequence, ensuring that the cell cycle follows an ordered trend.

NOTES

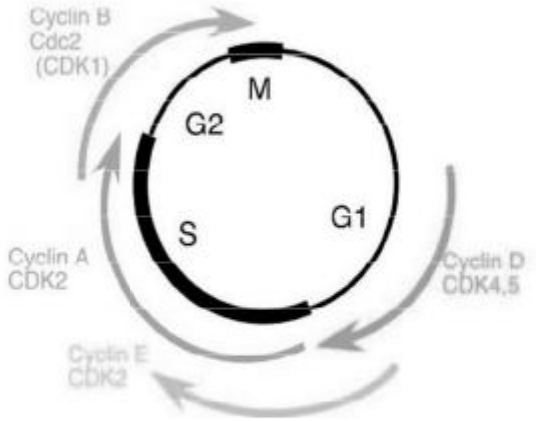


Fig. 7.5 : Regulation of Cyclin levels throughout the Cell Cycle

7.6.2 Cell Cycle Progression

Entry into a new cell cycle normally begins when signals from outside of the cell (mitogens or appropriate environmental conditions) trigger a combination of events that prompts cells to overcome the so-called *restriction point* and commit to a new round of division. The restriction point corresponds to a late event in G_1. Concomitantly, G_1/S- and S-cyclin gene expression increases, activating G_1/S-Cdk complexes and launching the S phase. DNA replication causes S-cyclins- Cdk inactivation, in order to prevent a second round of genome duplication. Towards the end of S-phase, M-cyclins (Cyclins-A/-B) gene expression is switched on and their relative concentration within the cell increases, leading to the accumulation of M-cyclins-Cdk complexes during G_2. Unlike cyclin A, which also has a noteworthy role in progression through S phase, cyclin B function is restricted to mitosis. Accordingly, cyclin B is both synthesised and degraded later than cyclin A. While cyclin A starts to accumulate during G_1, through S phase peaking in G_2, it begins to be degraded in early prometaphase and is completely absent by metaphase; whereas cyclin B degradation starts only at the metaphase to anaphase transition.

In most cell types, M-cyclins-Cdk complexes are initially kept inactive due to phosphorylation of two specific and conserved sites on the Cdk subunit that are catalyzed by Myt1 and Wee1 kinases. At G_2/M transition, removal of inhibitory phosphate groups from Cdk, by members of the Cdc25 family, combined with

cyclin binding, abruptly activates cyclin B-Cdk1, which in turn phosphorylates several downstream targets resulting in: induction of chromosome condensation, NEBD, assembly of the mitotic spindle apparatus and alignment of fully resolved and condensed chromosomes at the metaphase plate. In addition to Cdk1, several other mitotic kinases are known to influence cell division (e.g. Polo, Aurora B, Mps1, BubR1) and are mostly involved in controlling the spatial-temporal order of later mitotic events, like: removal of cohesion complexes from chromosomal arms during prophase (Polo), correction of improper interactions between spindle microtubules and kinetochores (Aurora B, BubR1) or setting off sister chromatid separation during metaphase, by alleviating APC/C inhibition (BubR1, Mps1). Some of these proteins are activated by cyclin B-Cdk1 complexes and further activate it, creating a positive feedback loop that rapidly amplifies to the fullest cyclin B-cdk1 activity.

NOTES

M-cyclins-Cdks pave their own way to destruction, since once activated they will also phosphorylate core subunits of the anaphase-promoting complex or cyclosome (APC/C) (Cdc16, Cdc27, Cdc23) and also Cdc20, main co-activator of the APC/C during mitosis, thus stimulating its E3 ubiquitin-ligase activity, resulting in mitotic cyclins degradation and Cdk1 inactivation in a negative feedback loop vital to ensure mitotic exit. APC/CCdc20 is also responsible for securin degradation, thereby releasing separase from inhibition, which in turn triggers the sudden loss of cohesion between sister chromatids. Both cyclin B degradation and sister chromatid separation are strictly dependent on APC/CCdc20 activity; in this way the cell ensures the synchrony of events at anaphase onset, imperative for successful partioning of the duplicated genome between daughter nuclei. In addition, the activity of the APC/CCdc20 complex is further regulated by SAC signalling, which monitors chromosome alignment at the metaphase plate by targeting Cdc20, creating an intricate network with different levels of regulation that influence each other, ultimately leading to a controlled mitotic exit. When each sister chromatid of a chromosome binds microtubules from opposite spindle poles biorientantion is achieved, the checkpoint is satisfied, and the APC/C is activated. At this point, the APC/C triggers sister chromatid separation and Cdk1 inactivation, resulting in anaphase onset and mitotic exit.

NOTES

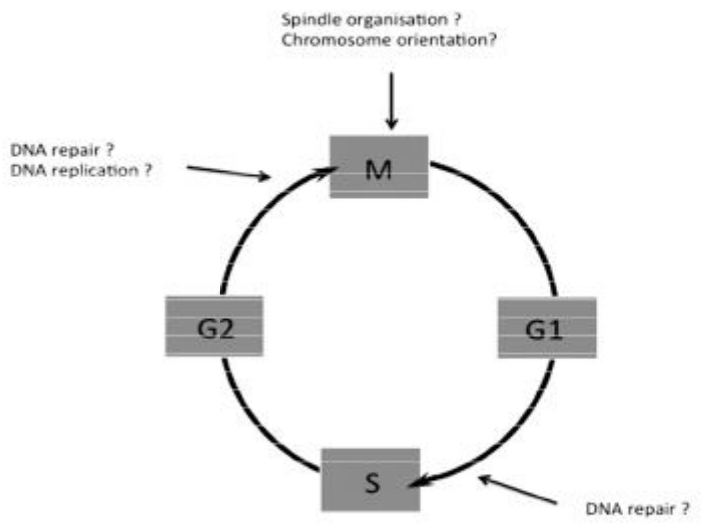


Fig. 7.6 : Cell Cycle Checkpoints

7.7 Retinoblastoma

The retinoblastoma protein (protein name abbreviated Rb; gene name abbreviated RB or *RB1*) is a tumor suppressor protein that is dysfunctional in several major cancers. One function of pRb is to prevent excessive cell growth by inhibiting cell cycle progression until a cell is ready to divide. It is also a recruiter of several chromatin remodeling enzymes such as methylases and acetylases. Rb belongs to the pocket protein family, whose members have a pocket for the functional binding of other proteins. Should an oncogenic protein, such as those produced by cells infected by high-risk types of human papilloma viruses, bind and inactivate pRb, this can lead to cancer.

Two forms of retinoblastoma were noticed: a bilateral, familial form and a unilateral, sporadic form. Sufferers of the former were 6 times more likely to develop other types of cancer later in life. This highlighted the fact that mutated Rb could be inherited and lent support to the two-hit hypothesis. This states that only one working allele of a tumour suppressor gene is necessary for its function (the mutated gene is recessive), and so both need to be mutated before the cancer phenotype will appear. In the familial form, a mutated allele is inherited along with a normal allele. In this case, should a cell sustain

only *one* mutation in the other RB gene, all Rb in that cell would be ineffective at inhibiting cell cycle progression, allowing cells to divide uncontrollably and eventually become cancerous. Furthermore, as one allele is already mutated in all other somatic cells, the future incidence of cancers in these individuals is observed with linear kinetics. The working allele need not undergo a mutation *per se*, as loss of heterozygosity is frequently observed in such tumors.

NOTES

However, in the sporadic form, both alleles would need to sustain a mutation before the cell can become cancerous. This explains why sufferers of sporadic retinoblastoma are not at increased risk of cancers later in life, as both alleles are functional in all their other cells. Future cancer incidence in sporadic Rb cases is observed with polynomial kinetics, not exactly quadratic as expected because the first mutation must arise through normal mechanisms, and then can be duplicated by LOH to result in a tumor progenitor.

7.7.1 Cell Cycle Suppression

Rb prevents the cell from replicating damaged DNA by preventing its progression along the cell cycle through G_1 (first gap phase) into S (synthesis phase). Rb binds and inhibits transcription factors of the E2F family, which are composed of dimers of an E2F protein and a dimerization partner (DP) protein. The transcription activating complexes of E2 promoter-binding–protein-dimerization partners (E2F-DP) can push a cell into S phase. As long as E2F-DP is inactivated, the cell remains stalled in the G_1 phase. When Rb is bound to E2F, the complex acts as a growth suppressor and prevents progression through the cell cycle. The Rb-E2F/DP complex also attracts a histone deacetylase (HDAC) protein to the chromatin, reducing transcription of S phase promoting factors, further suppressing DNA synthesis.

Rb is phosphorylated to pRb by certain Cyclin Dependent Kinases (CDKs). pRb is described as being hyperphosphorylated and when in this state, it is unable to complex E2F and therefore, unable to restrict progression from the G_1 phase to the S phase of the cell cycle. During the M-to-G_1 transition, pRb is progressively dephosphorylated by PP1, returning to its growth-suppressive hypophosphorylated state Rb.

NOTES

When it is time for a cell to enter S phase, complexes of cyclin-dependent kinases (CDK) and cyclins phosphorylate Rb to pRb, inhibiting its activity. The initial phosphorylation is performed by Cyclin D/CDK4/CDK6 and followed by additional phosphorylation by Cyclin E/CDK2. pRb remains phosphorylated throughout S, G_2 and M phases. Phosphorylation of pRb allows E2F-DP to dissociate from pRb and become active. When E2F is free it activates factors like cyclins (e.g. Cyclin E and A), which push the cell through the cell cycle by activating cyclin-dependent kinases, and a molecule called proliferating cell nuclear antigen, or PCNA, which speeds DNA replication and repair by helping to attach polymerase to DNA.

Rb family proteins are components of the DREAM complex (also named LINC complex), which is composed of LIN9, LIN54, LIN37, MYBL2, RBL1, RBL2, RBBP4, TFDP1, TFDP2, E2F4 and E2F5. There is a testis-specific version of the complex, where LIN54, MYBL2 and RBBP4 are replaced by MTL5, MYBL1 and RBBP7, respectively. In Drosophila both DREAM versions also exist, the components being mip130 (lin9 homolog, replaced by aly in testes), mip120 (lin54 homolog, replaced by tomb in testes), and Myb, Caf1p55, DP, Mip40, E2F2, Rbf and Rbf2. The DREAM complex exists in quiescent cells in association with MuvB (consisting of HDAC1 or HDAC2, LIN52 and L3mbtl1, L3mbtl3 or L3mbtl4) where it represses cell cycle-dependent genes. DREAM dissociates in S phase from MuvB and gets recruited by MYB.

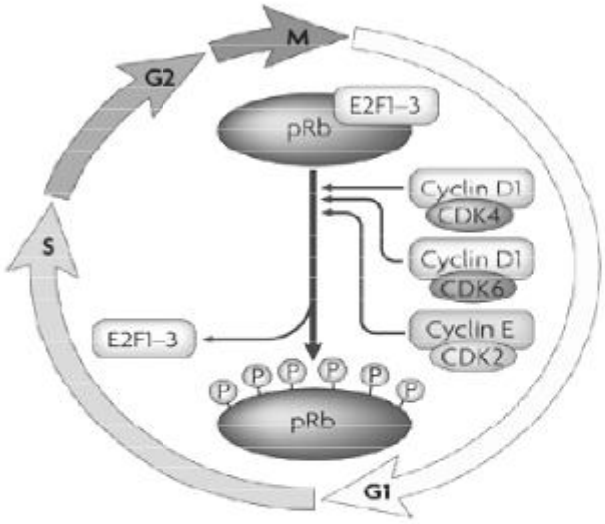


Fig. 7.7 : Retinoblastoma Protein and Cell Cycle Progression

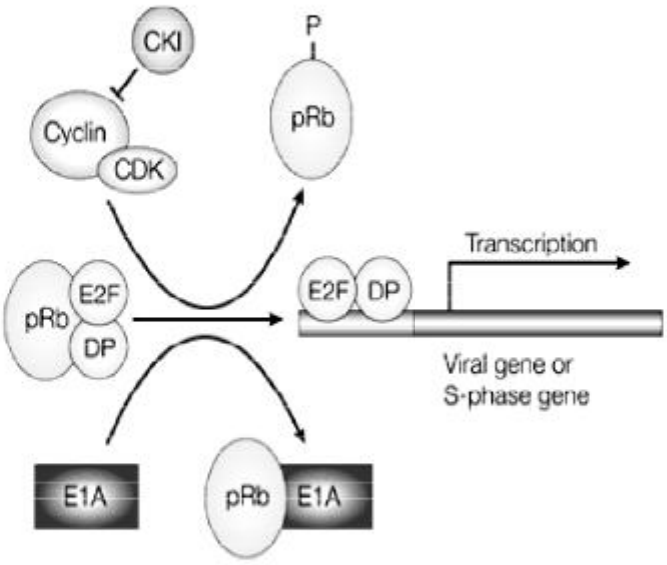


Fig. 7.8 : Retinoblastoma Dependent Cell Cycle

7.8 E2F Protein

E2F is a group of genes that codifies a family of transcription factors (TF) in higher eukaryotes. Three of them are activators: E2F1, 2 and E2F3a. Six others act as suppressors: E2F3b, E2F4-8. All of them are involved in the cell cycle regulation and synthesis of DNA in mammalian cells. E2Fs as TFs bind to the TTTCCCGC (or slight variations of this sequence) consensus binding site in the target promoter sequence.

7.8.1 Role in the cell cycle

E2F family members play a major role during the G_1/S transition in mammalian and plant cell cycle. DNA microarray analysis reveals unique sets of target promoters among E2F family members suggesting that each protein has a unique role in the cell cycle. Among E2F transcriptional targets are cyclins, cdks, checkpoints regulators, DNA repair and replication proteins. Nonetheless, there is a great deal of redundancy among the family members. Mouse embryos lacking E2F1, E2F2, and one of the E2F3 isoforms, can develop normally when either E2F3a or E2F3b, is expressed.

The E2F family is generally split by function into two groups: transcription activators and repressors. Activators such as E2F1, E2F2, E2F3a promote and

NOTES

help carry out the cell cycle, while repressors inhibit the cell cycle. Yet, both sets of E2F have similar domains. E2F1-6 have DP1,2 heterodimerization domain which allows them to bind to DP1 or DP2, proteins distantly related to E2F. Binding with DP1,2 provides a second DNA binding site, increasing E2F binding stability. Most E2F have a pocket protein binding domain. Pocket proteins such as pRB and related proteins p107 and p130, can bind to E2F when hypophosphorylated. In activators, E2F binding with pRB has been shown to mask the transactivation domain responsible for transcription activation. In repressors E2F4 and E2F5, pocket protein binding (more often p107 and p130 than pRB) mediates recruitment of repression complexes to silence target genes. E2F6, E2F7, and E2F8 do not have pocket protein binding sites and their mechanism for gene silencing is unclear. Cdk4 (6)/cyclin D and cdk2/cyclin E phosphorylate pRB and related pocket proteins allowing them to disassociate from E2F. Activator E2F proteins can then transcribe S phase promoting genes. In REF52 cells, overexpression of activator E2F1 is able to push quiescent cells into S phase. While repressors E2F4 and 5 do not alter cell proliferation, they mediate G_1 arrest.

E2F activator levels are cyclic, with maximal expression during G_1/S. In contrast, E2F repressors stay constant, especially since they are often expressed in quiescent cells. Specifically, E2F5 is only expressed in terminally differentiated cells in mice. The balance between repressor and activator E2F regulate cell cycle progression. When activator E2F family proteins are knocked out, repressors become active to inhibit E2F target genes.

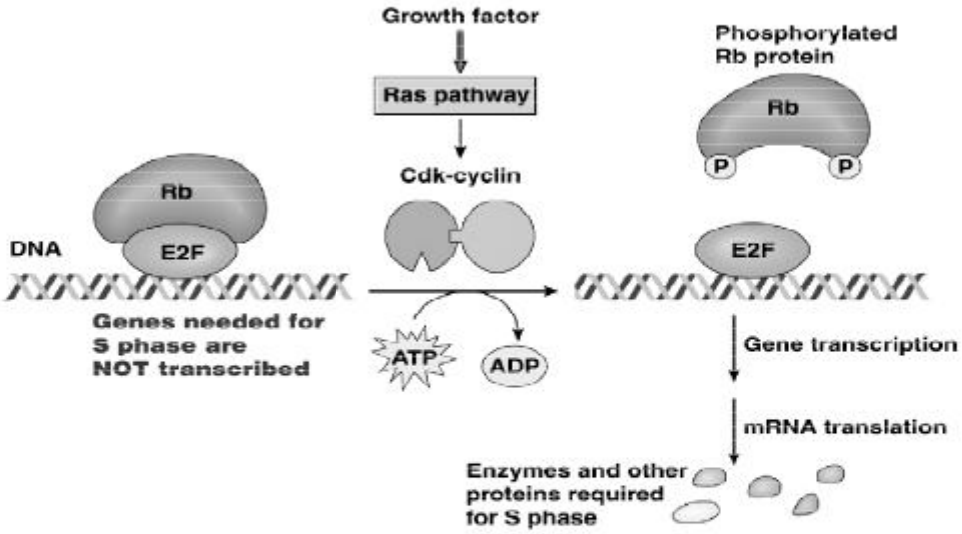


Fig. 7.9 : E2F Dependent Cell Cycle

7.9 E2F/pRb complexes

NOTES

The Rb tumor suppressor protein (pRb) binds to the E2F1 transcription factor preventing it from interacting with the cell's transcription machinery. In the absence of pRb, E2F1 (along with its binding partner DP1) mediates the trans-activation of E2F1 target genes that facilitate the G_1/S transition and S-phase. E2F target genes encode proteins involved in DNA replication, and chromosomal replication. When cells are not proliferating, E2F DNA binding sites contribute to transcriptional repression. In vivo footprinting experiments obtained on Cdc_2 and B-myb promoters demonstrated E2F DNA binding site occupation during G_0 and early G_1, when E2F is in transcriptional repressive complexes with the pocket proteins.

pRb is one of the targets of the oncogenic human papilloma virus protein E7, and human adenovirus protein E1A. By binding to pRB, they stop the regulation of E2F transcription factors and drive the cell cycle to enable virus genome replication.

7.9.1 Activators: E2F1, E2F2, E2F3a

Activators are maximally expressed late in G_1 and can be found in association with E2F regulated promoters during the G_1/S transition. The activation of E2F-3a genes follows upon the growth factor stimulation and the subsequent phosphorylation of the E2F inhibitor retinoblastoma protein, pRB. The phosphorylation of pRB is initiated by cyclin D/cdk4,cdk6 complex and continued by cyclin E/cdk2. Cyclin D/cdk4,6 itself is activated by the MAPK signaling pathway. When bound to E2F-3a, pRb can directly repress E2F-3a target genes by recruiting chromatin remodeling complexes and histone modifying activities (e.g. histone deacetylase,HDAC) to the promoter.

7.9.2 Inhibitors: E2F3b, E2F4, E2F5, E2F6, E2F7, E2F8

NOTES

E2F3b, E2F4, E2F5 are expressed in quiescent cells and can be found associated with E2F-binding elements on E2F-target promoters during G_0-phase. E2F-4 and 5 preferentially bind to p107/p130.

- E2F-6 acts as a transcriptional repressor, but through a distinct, pocket protein independent manner. E2F-6 mediates repression by direct binding to polycomb-group proteinsor via the formation of a large multimeric complex containing Mga and Max proteins.
- The repressor genes E2F7/E2F8, located on chromosome 7, are transcription factors responsible for protein coding cell cycle regulation. Together, they are essential for the development of an intact, organized, and functional placental structure during embryonic development.

7.10 Programmed Cell Death

Programmed cell-death (or PCD) is death of a cell in any form, mediated by an intracellular program. PCD is carried out in a regulated process, which usually confers advantage during an organism's life-cycle. For example, the differentiation of fingers and toes in a developing human embryo occurs because cells between the fingers apoptose; the result is that the digits are separate. PCD serves fundamental functions during both plant and metazoa (multicellular animals) tissue development. Apoptosis and autophagyare both forms of programmed cell death, but necrosis is a non-physiological process that occurs as a result of infection or injury.

Necrosis is the death of a cell caused by external factors such as trauma or infection and occurs in several different forms. Recently a form of programmed necrosis, callednecroptosis, has been recognized as an alternate form of programmed cell death. It is hypothesized that necroptosis can serve as a cell-death backup to apoptosis when the apoptosis signaling is blocked by endogenous or exogenous factors such as viruses or mutations.

The concept of "programmed cell-death" was used by Lockshin & Williams in 1964 in relation to insect tissue development, around eight years before

"apoptosis" was coined. Since then, PCD has become the more general of these two terms.

The first insight into the mechanism came from studying BCL2, the product of a putative oncogene activated by chromosome translocations often found in follicular lymphoma. Unlike other cancer genes, which promote cancer by stimulating cell proliferation, BCL2 promoted cancer by stopping lymphoma cells from being able to kill themselves.

NOTES

The program cell death can be divided in two types- apoptosis and autophagy.

7.10.1 Apoptosis

Apoptosis is the process of programmed cell death (PCD) that may occur in multicellular organisms. Biochemical events lead to characteristic cell changes and death. These changes include blebbing, cell shrinkage, nuclear fragmentation, chromatin condensation, and chromosomal DNA fragmentation. It is now thought that- in a developmental context- cells are induced to positively commit suicide whilst in a homeostatic context; the absence of certain survival factors may provide the impetus for suicide. There appears to be some variation in the morphology and indeed the biochemistry of these suicide pathways; some treading the path of "apoptosis", others following a more generalized pathway to deletion, but both usually being genetically and synthetically motivated. There is some evidence that certain symptoms of "apoptosis" such as endonuclease activation can be spuriously induced without engaging a genetic cascade, however, presumably true apoptosis and programmed cell death must be genetically mediated. It is also becoming clear that mitosis and apoptosis are toggled or linked in some way and that the balance achieved depends on signals received from appropriate growth or survival factors.

NOTES

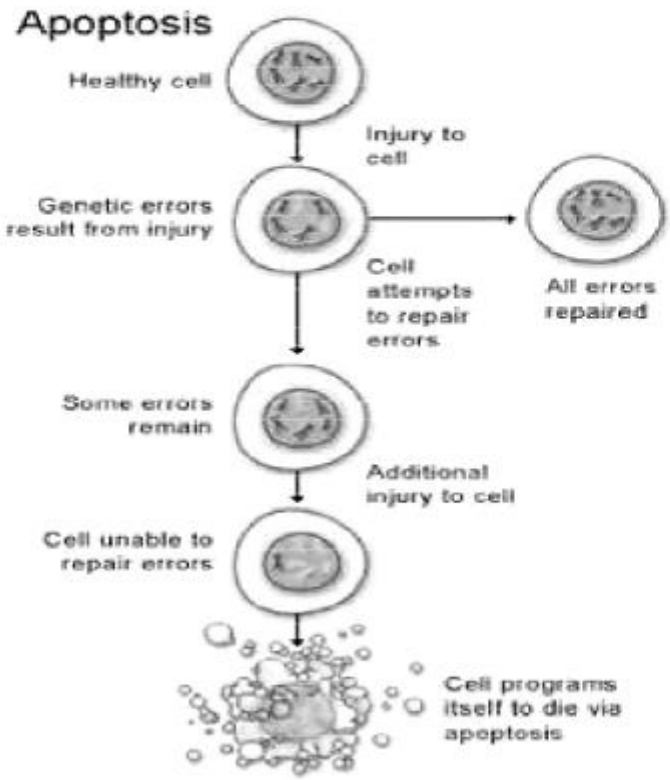


Fig. 7.10 : Program Med Cell Death (Apoptosis)

7.10.2 Autophagy

Macroautophagy, often referred to as autophagy, is a catabolic process that results in the autophagosomic-lysosomal degradation of bulk cytoplasmic contents, abnormal protein aggregates, and excess or damaged organelles.

Autophagy is generally activated by conditions of nutrient deprivation but has also been associated with physiological as well as pathological processes such as development, differentiation, neurodegenerative diseases, stress physiology, infection and cancer.

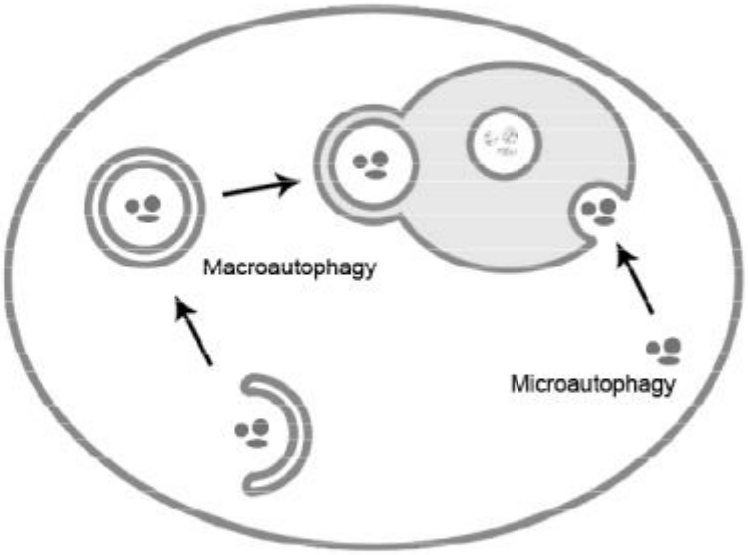


Fig. 7.11 : Program Med Cell Death (Autophagy)

7.10.3 Mechanism

A critical regulator of autophagy induction is the kinase mTOR, which when activated, suppresses autophagy and when not activated promotes it. Three related serine/threoninekinases, UNC-51-like kinase -1, -2, and -3 (ULK1, ULK2, UKL3), which play a similar role as the yeast Atg1, act downstream of the mTOR complex. ULK1 and ULK2 form a large complex with the mammalian homolog of an autophagy-related (Atg) gene product (mAtg13) and the scaffold protein FIP200. Class III PI3K complex, containing hVps34, Beclin-1, p150 and Atg14-like protein or ultraviolet irradiation resistance-associated gene (UVRAG), is required for the induction of autophagy.

The ATG genes control the autophagosome formation through ATG12-ATG5 and LC3-II (ATG8-II) complexes. ATG12 is conjugated to ATG5 in an ubiquitin-like reaction that requires ATG7 and ATG10. The Atg12–Atg5 conjugate then interacts non-covalently with ATG16 to form a large complex. LC3/ATG8 is cleaved at its C terminus by ATG4protease to generate the cytosolic LC3-I. LC3-I is conjugated to phosphatidylethanolamine (PE) also in a ubiquitin-like reaction that requires Atg7 and Atg3. The lipidated form of LC3, known as LC3-II, is attached to the autophagosome membrane.

NOTES

Autophagy and apoptosis are connected both positively and negatively, and extensive crosstalk exists between the two. During nutrient deficiency, autophagy functions as a pro-survival mechanism, however, excessive autophagy may lead to cell death, a process morphologically distinct from apoptosis. Several pro-apoptotic signals, such as TNF, TRAIL, and FADD, also induce autophagy. Additionally, Bcl-2 inhibits Beclin-1-dependent autophagy, thereby functioning both as a pro-survival and as an anti-autophagic regulator.

7.11 Summary

According to the cell theory, cells arise from preexisting cells. The process by which this occurs is called cell division. Any sexually reproducing organism starts its life cycle from a single-celled zygote. Cell division does not stop with the formation of the mature organism but continues throughout its life cycle. The stages through which a cell passes from one division to the next is called the cell cycle. Cell cycle is divided into two phases called (i) Interphase – a period of preparation for cell division, and (ii) Mitosis (M phase) – the actual period of cell division. Interphase is further subdivided into G1, S and G2. G1 phase is the period when the cell grows and carries out normal metabolism.

Most of the organelle duplication also occurs during this phase. S phase marks the phase of DNA replication and chromosome duplication. G2 phase is the period of cytoplasmic growth. Mitosis is also divided into four stages namely prophase, metaphase, anaphase and telophase. Chromosome condensation occurs during prophase. Simultaneously, the centrioles move to the opposite poles. The nuclear envelope and the nucleolus disappear and the spindle fibres start appearing. Metaphase is marked by the alignment of chromosomes at the equatorial plate. During anaphase the centromeres divide and the chromatids start moving towards the two opposite poles. Once the chromatids reach the two poles, the chromosomal elongation starts, nucleolus and the nuclear membrane reappear. This stage is called the telophase. Nuclear division is then followed by the cytoplasmic division and is called cytokinesis. Mitosis thus, is the equational division in which the chromosome number of the parent is conserved in the daughter cell.

In contrast to mitosis, meiosis occurs in the diploid cells, which are destined to form gametes. It is called the reduction division since it reduces the chromosome number by half while making the gametes. In sexual reproduction when the two gametes fuse the chromosome number is restored to the value in the parent. Meiosis is divided into two phases – meiosis I and meiosis II. In the first meiotic division the homologous chromosomes pair to form bivalents, and undergo crossing over. Meiosis I have a long prophase, which is divided further into five phases. These are leptotene, zygotene, pachytene, diplotene and diakinesis. During metaphase I the bivalents arrange on the equatorial plate. This is followed by anaphase I in which homologous chromosomes move to the opposite poles with both their chromatids. Each pole receives half the chromosome number of the parent cell. In telophase I, the nuclear membrane and nucleolus reappear. Meiosis II is similar to mitosis. During anaphase II the sister chromatids separate. Thus at the end of meiosis four haploid cells are formed.

NOTES

7.12 Glossary

- **Anaphase :** The stage of mitosis or meiosis during which centromeres split and chromatids separate and chromatids move to opposite poles.
- **Bivalent/Tetrad :** A homologous pair of chromosomes in the synapsed, or paired, state during prophase I of the meiotic division and it refer to the fact that the structure contains 4 chromatids.
- **Cell Cycle :** The cell cycle is the series of events that take place in a cell leading to its replication. These events have interphase-during which the cell grows, accumulating nutrients needed for mitosis and duplicating its DNA—and the mitotic (M) phase, during which the cell splits itself into two distinct cells, often called "daughter cells".
- **Centromere :** It is the primary constriction in chromosome to which the spindle fibres attach during mitotic and meiotic division. It appears as a constriction when chromosomes contract during cell division. After chromosomal duplication, which occurs at the beginning of every mitotic and meiotic division, the two resultant chromatids are joined at the centromere.

NOTES

- **Chiasmata :** X-shaped observable regions in diplotene in which nonsister chromatids of homologous chromosomes cross-over each other are called chiasmata.
- **Chromatids :** The copied arm of a chromosome, joined together at the centromere, that separate during cell division.
- **Chromatin :** Chromatin is the complex of DNA and protein that makes up chromosomes. It is found inside the nuclei of eukaryotic cells, and within the nucleoid in prokaryotes. The functions of chromatin are to package DNA into a smaller volume to fit in the cell, to strengthen the DNA to allow mitosis and meiosis, and to serve as a mechanism to control expression.
- **Chromosomes :** Thread like strands of DNA and associated proteins in the nucleus of cells that carry the genes and functions in the transmission of hereditary information.
- **Crossing over :** Crossing over is a process in which homologous chromosomes exchange genetic material through the breakage and reunion of two chromatids with the help of enzyme recombinase. This process can result in an exchange of alleles between chromosomes.
- **Homologous :** Chromosomes - Homologous chromosomes are chromosomes in a biological cell that pair (synapse) during meiosis and contain the same genes at the same loci but possibly different genetic information, called alleles, at those genes.
- **Synapsis :** The pairing of homologous chromosomes along their length; synapsis usually occurs during prophase I of meiosis, but it can also occur in somatic cells of some organisms.

7.13 Self-Learning Exercises

Section – A (Very Short Answer Type Questions)

5. A typical eukaryotic cell divides once in?
6. How many phases in a cell cycle?
7. Abbreviation of Cdks.

8. The concept of "programmed cell-death" was first used by?

Section – B (Short Answer Type Questions)

4. Write about interphase of mitosis.
5. Describe the role of Cdks in cell cycle control.
6. What is significance of mitosis?
7. Why mitosis called equational division?

Section – C (Long Answer Type Questions)

4. Give the detail process of meiosis.
5. Describe cell cycle control.
6. Describe the programmed cell death.
7. Give detail account about E2F/pRB complexes.

Section – A (Answer Key)

5. In every 24 hours.
6. Two – Interphase and M phase.
7. Cyclin dependent kinases.
8. Lockshin & Williams in 1964.

NOTES

7.14 References

- Alushin GM, Ramey VH, Pasqualato S, Ball DA, Grigorieff N, Musacchio A & Nogales E (2010) The Ndc80 kinetochore complex forms oligomeric arrays along microtubules. *Nature* 467, 805-810.
- Baker DJ, Dawlaty MM, Galardy P & van Deursen JM (2007) Mitotic regulation of the anaphase-promoting complex. *Cell Mol Life Sci* 64, 589-600.
- Cheeseman IM, Chappie JS, Wilson-Kubalek EM & Desai A (2006) The conserved KMN network constitutes the core microtubule-binding site of the kinetochore. *Cell* 127, 983-997.
- Cohen J (2002) Sorting out chromosome errors. *Science* 296, 2164-2166.

NOTES

- Dienemann A & Sprenger F (2004) Requirements of cyclin a for mitosis are independent of its subcellular localization. *Curr Biol* 14, 1117-1123.
- Hoyt MA, Totis L & Roberts BT (1991) S. cerevisiae genes required for cell cycle arrest in response to loss of microtubule function. *Cell* 66, 507-517.
- Karess R (2005) Rod-Zw10-Zwilch: a key player in the spindle checkpoint. *Trends Cell Biol* 15, 386-392.
- King RW, Glotzer M & Kirschner MW (1996) Mutagenic analysis of the destruction signal of mitotic cyclins and structural characterization of ubiquitinated intermediates. *Mol Biol Cell* 7, 1343-1357.
- Laycock JE, Savoian MS & Glover DM (2006) Antagonistic activities of Klp10A and Orbit regulate spindle length, bipolarity and function in vivo. *J Cell Sci* 119, 2354-2361.
- Lemos CL, Sampaio P, Maiato H, Costa M, Omel'yanchuk LV, Liberal V & Sunkel CE (2000) Mast, a conserved microtubule-associated protein required for bipolar mitotic spindle organization. *EMBO J* 19, 3668-3682.
- Lengauer C, Kinzler KW & Vogelstein B (1998) Genetic instabilities in human cancers. *Nature* 396, 643-649.
- Li R & Murray AW (1991) Feedback control of mitosis in budding yeast. *Cell* 66, 519-531.
- Li X & Nicklas RB (1995) Mitotic forces control a cell-cycle checkpoint. *Nature* 373, 630-632.
- Maia AF, Lopes CS & Sunkel CE (2007) BubR1 and CENP-E have antagonistic effects upon the stability of microtubule-kinetochore attachments in Drosophila S2 cell mitosis. *Cell Cycle* 6, 1367-1378.
- Maiato H, Khodjakov A & Rieder CL (2005) Drosophila CLASP is required for the incorporation of microtubule subunits into fluxing kinetochore fibres. *Nat Cell Biol* 7, 42-47.
- Maiato H, Sampaio P & Sunkel CE (2004) Microtubule-associated proteins and their essential roles during mitosis. *Int Rev Cytol* 241, 53-153.

- Mao Y, Abrieu A & Cleveland DW (2003) Activating and silencing the mitotic checkpoint through CENP-E-dependent activation/inactivation of BubR1. *Cell* 114, 87-98.
- Musacchio A & Salmon ED (2007) The spindle-assembly checkpoint in space and time. *Nat Rev Mol Cell Biol* 8, 379-393.
- Peters JM (2002) The anaphase-promoting complex: proteolysis in mitosis and beyond. *Mol Cell* 9, 931-943.
- Reimann JD, Freed E, Hsu JY, Kramer ER, Peters JM & Jackson PK (2001) Emi1 is a mitotic regulator that interacts with Cdc20 and inhibits the anaphase promoting complex. *Cell* 105, 645-655.
- Rieder CL & Salmon ED (1994) Motile kinetochores and polar ejection forces dictate chromosome position on the vertebrate mitotic spindle. *J Cell Biol* 124, 223- 233.
- Wei RR, Schnell JR, Larsen NA, Sorger PK, Chou JJ & Harrison SC (2006) Structure of a central component of the yeast kinetochore: the Spc24p/Spc25p globular domain. *Structure* 14, 1003-1009.
- Weiss E & Winey M (1996) The Saccharomyces cerevisiae spindle pole body duplication gene MPS1 is part of a mitotic checkpoint. *J Cell Biol* 132, 111-123.
- Yen TJ (2002) The complexity of APC/C regulation: location, location, location. *Cell Cycle* 1, 260-261.

NOTES

NOTES

Unit – 8

Techniques in Cell Biology

Structure of the Unit

8.0 Objectives

After going through this unit you will be able to understand:

- Immunotechniques
- In situ Hybridization
- Transcripts
- FISH
- GISH
- Confocal Microscopy

NOTES

8.1 Immunotechniques: Introduction

The immune system comprises of cells that have the ability to recognize, respond to and bring about inactivation of microbial infections, neutralization of pathogen-derived toxins and provide protection against diseased states such as cancers. These cells are located in primary and secondary lymphoid organs which include bone marrow, thymus, spleen and lymph nodes and are also found in circulation in blood. There are two components of the immune system, the humoral and the cellular. The former mediates its function through antibodies, soluble molecules that are produced by bone marrow derived cells called B lymphocytes. On the other hand the cellular component is largely defined by cells which are dependent on thymus for their development and maturation. These cells are called T-lymphocytes (T for thymus derived). Even though these two arms of the immune system perform independent functions, their functioning is dependent on mutual collaboration commonly referred to as T-B collaboration.

B-cells can produce antibodies against virtually any molecule. This ability to generate a large diversity of antibodies is achieved through a process of recombination that involves shuffling of immunoglobulin genes. B-cells recognize an antigen in its native state through specific B-cell receptors. Each B-cell receptor has a recognition module made of immunoglobulin heavy and light chains which are non-covalently associated with molecules that enable it to transduce signals inside the cell. The help provided by T-cells enables B-cells proliferate and differentiate into antibody-secreting cells called plasma

cells. During this process antibodies also undergo somatic hypermutation which increases their affinities for their cognate antigens. The signals provided by T-cells also enable B-lymphocytes to produce antibodies with different functional capabilities by a process called isotype switching.

NOTES

Unlike B-cells, T cells recognize an antigen only after one of its fragments is presented in association with molecules of the major histocompatibility complex (MHC). This job of antigen presentation is carried out by antigen presenting cells which include B-cells or specialized immune cells that can take up microorganisms. The two main cell types in the latter category are macrophages and dendritic cells; these cells internalize microorganisms by phagocytosis. Like B-cells, T-cells also express specific receptors for their cognate ligands, the peptide-MHC complexes, and respond by secreting soluble mediators called cytokines.

Upon encountering a pathogenic microorganism, the immune system first senses the presence of a pathogen. The pathogen is phagocytosed by macrophages or dendritic cells and its molecules are broken down into smaller fragments. This process of fragmentation called antigen processing is carried out in a compartment which is enriched in degradatory enzymes and is also identified by the presence of MHC class II molecules. Optimal sized peptides are loaded onto MHC class II molecules, transported to the membrane and presented to antigen-specific T cells. Upon activation, T cells increase in number and perform different functions. Infection of a macrophage or a dendritic cell with a pathogen or its activation with a pathogen-derived entity such as bacterial lipopolysaccharide also leads to induction and upregulation of molecules called costimulatory molecules which provide a second signal for activation of antigen-specific T cells. As mentioned earlier, antigenic peptides derived from infectious agents can also be presented to cytotoxic T cells in association with MHC class I molecules. These T cells can be identified by the presence of cell surface associated co-receptor molecule, CD8 while helper T cells are identified by the CD4 molecule.

8.2 Antigens

An antigen is a substance/molecule that when introduced into the body triggers the production of an antibody by the immune system which will then kills or

neutralize the antigen that is recognized as a foreign and potentially harmful invader. These invaders can be molecules such as pollen or cells such as bacteria. Originally the term came from antibody generator and was a molecule that binds specifically to an antibody, but the term now also refers to any molecule or molecular fragment that can be bound by a major histocompatibility complex (MHC) and presented to a T-cell receptor"Self" antigens are usually tolerated by the immune system; whereas "Non-self" antigens are identified as intruders and attacked by the immune system. Autoimmune disorders arise from the immune system reacting to its own antigens.

NOTES

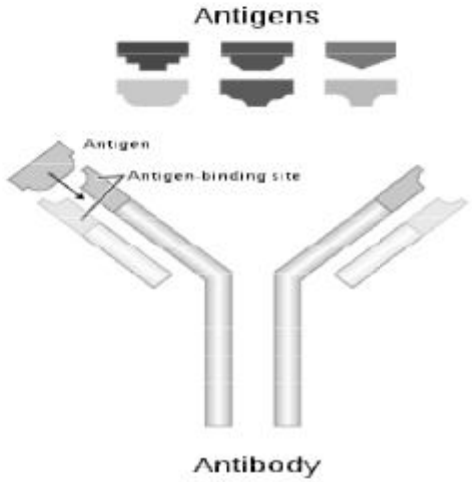


Fig. 8.1 : Antigen-Antibody Binding

Each antibody binds to a specific antigen; an interaction similar to a lock and key. Similarly, an immunogen is a specific type of antigen. An immunogen is defined as a substance that is able to provoke an adaptive immune response if injected on its own. Said another way, an immunogen is able to induce an immune response, while an antigen is able to combine with the products of an immune response once they are made. The overlapping concepts of immunogenicity and antigenicity are thereby subtly different.

Immunogenicity is the ability to induce a humoral and/or cell-mediated immune response Antigenicity is the ability to combine specifically with the final products of the (i.e. secreted antibodies or surface receptors on T-cells). Although all molecules that have the property of immunogenicity also have the property of antigenicity, the reverse is not true."

NOTES

At the molecular level, an antigen is characterized by its ability to be "bound" at the antigen-binding site of an antibody. Note also that antibodies tend to discriminate between the specific molecular structures presented on the surface of the antigen. Antigens are usually proteins or polysaccharides. This includes parts (coats, capsules, cell walls, flagella, fimbrae, and toxins) of bacteria, viruses, and other microorganisms. Lipids and nucleic acids are antigenic only when combined with proteins and polysaccharides. Non-microbial exogenous (non-self) antigens can include pollen, egg white, and proteins from transplanted tissues and organs or on the surface of transfused blood cells. Vaccines are examples of immunogenic antigens intentionally administered to induce acquired immunity in the recipient. Cells present their immunogenic-antigens to the immune system via a histocompatibility molecule. Depending on the antigen presented and the type of the histocompatibility molecule, several types of immune cells can become activated.

8.2.1 Origin of the term Antigen

In 1899 Ladislas Deutsch named the hypothetical substances halfway between bacterial constituents and antibodies "substances immunogenes or antigenes". He originally believed those substances to be precursors of antibodies, just like zymogen is a precursor of zymase. But by 1903 he understood that an antigen induces the production of immune bodies (antibodies) and wrote that the word antigen was a contraction of "Antisomatogen". The Oxford English Dictionary indicates that the logical construction should be "anti(body)-gen".

8.2.2 Classification of Antigens

Antigens can be classified in order of their class:

Exogenous Antigens

Exogenous antigens are antigens that have entered the body from the outside, for example by inhalation, ingestion, or injection. The immune system's response to exogenous antigens is often subclinical. By endocytosis or phagocytosis, exogenous antigens are taken into the antigen-presenting cells (APCs) and processed into fragments. APCs then present the fragments to T helper cells ($CD4^+$) by the use of class II histocompatibility molecules on their surface. Some T cells are specific for the peptide: MHC complex. They become

activated and start to secrete cytokines. Cytokines are substances that can activate cytotoxic T lymphocytes (CTL), antibody-secreting B cells, macrophages, and other particles. Some antigens start out as exogenous antigens, and later become endogenous (for example, intracellular viruses). Intracellular antigens can be released back into circulation upon the destruction of the infected cell, again.

NOTES

Endogenous antigens

Endogenous antigens are antigens that have been generated within previously normal cells as a result of normal cell metabolism, or because of viral or intracellular bacterial infection. The fragments are then presented on the cell surface in the complex with MHC class I molecules. If activated cytotoxic $CD8^+$ T cells recognize them, the T cells begin to secrete various toxins that cause the lysis or apoptosis of the infected cell. In order to keep the cytotoxic cells from killing cells just for presenting self-proteins, self-reactive T cells are deleted from the repertoire as a result of tolerance (also known as negative selection). Endogenous antigens include xenogenic (heterologous), autologous and idiotypic or allogenic (homologous) antigens.

Autoantigens

An autoantigen is usually a normal protein or complex of proteins (and sometimes DNA or RNA) that is recognized by the immune system of patients suffering from a specific autoimmune disease. These antigens should, under normal conditions, not be the target of the immune system, but, due to mainly genetic and environmental factors, the normal immunological tolerance for such an antigen has been lost in these patients.

Tumor antigens

Tumor antigens or neoantigens are those antigens that are presented by MHC I or MHC II molecules on the surface of tumor cells. These antigens can sometimes be presented by tumor cells and never by the normal ones. In this case, they are called tumor-specific antigens (TSAs) and, in general, result from a tumor-specific mutation. More common are antigens that are presented by tumor cells and normal cells, and they are called tumor-associated antigens (TAAs). Cytotoxic T lymphocytes that recognize these antigens may be able to destroy the tumor cells before they proliferate or metastasize. Tumor antigens

can also be on the surface of the tumor in the form of, for example, a mutated receptor, in which case they will be recognized by B cells.

NOTES

Nativity

A native antigen is an antigen that is not yet processed by an APC to smaller parts. T cells cannot bind native antigens, but require that they be processed by APCs, whereas B cells can be activated by native ones.

Antigenic specificity

Antigen(ic) specificity is the ability of the host cells to recognize an antigen specifically as a unique molecular entity and distinguish it from another with exquisite precision. Antigen specificity is due primarily to the side-chain conformations of the antigen. It is a measurement, although the degree of specificity may not be easy to measure, and need not be linear or of the nature of a rate-limited step or equation.

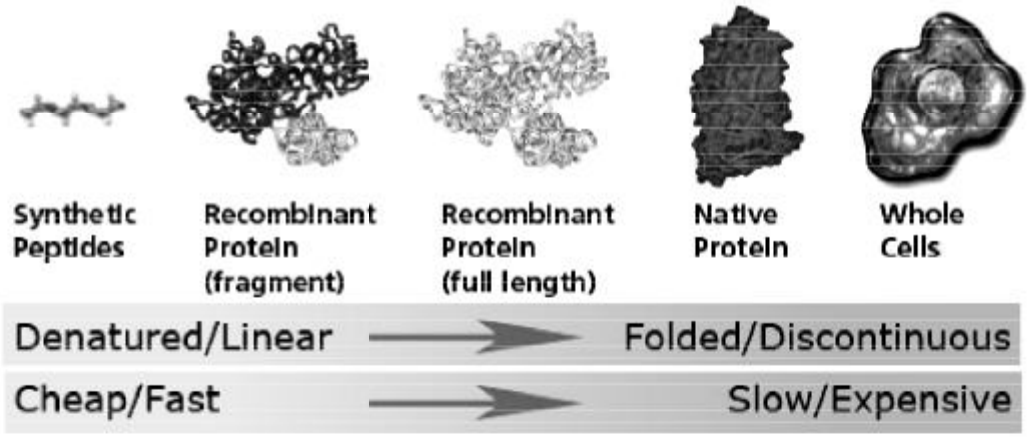


Fig. 8.2 Structure of Antigen

8.3 Antibodies

Polyclonal antibodies (or antisera) are antibodies that are obtained from different B cell resources. They are a combination of immunoglobulin molecules secreted against a specific antigen, each identifying a different epitope.

8.3.1 Analysis of Antibody Responses

The antibody response to an antigen is determined by mixing the antigen with the cognate antibody followed by detection of the antigen-antibody complex. The latter can be achieved either by a precipitation technique or by an immunoassay. The precipitation assays are based on the formation of a precipitin line when an antibody complexes with an antigen, whereas immunoassays use a labelled antigen or a labelled antibody to detect antigen-antibody complex.

NOTES

Precipitation assays: Double immunodiffusion

This is one of the simplest assays that can be employed to analyze antibody responses to an antigen without a need for any sophisticated instrument. This assay was first reported by Swedish scientist Orjan Ouchterlony 30 years ago. In this technique, an antigen and an antibody are allowed to migrate towards each other in a gel and the presence of antigen-antibody complex is revealed by the appearance of a line of precipitation. This line is formed at the point of equivalence; the precipitate is soluble in excess antigen or excess antibody. The rate of formation and the intensity of the precipitin line depend on the concentration and the molecular size of the antigen. Moreover, multiple precipitation lines are produced if the antigen preparation contains a mixture of several molecular species.

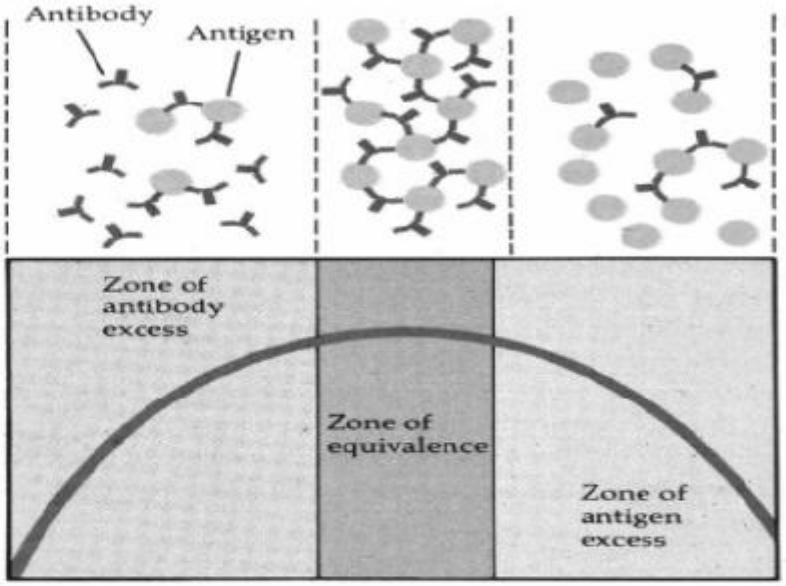


Fig. 8.3 : Antigen Antibody Reaction

NOTES

Immunoelectrophoresis

Immunoelectrophoresis combines electrophoresis with immunodiffusion. The combination of these two techniques increases resolution and enables to analyze interaction of antibodies with multiple antigens. A mixture of proteins is first allowed to separate into its components on the basis of charge by electrophoresis and following this, the separated components and the antibodies are made to interact with each other by allowing them to diffuse through the gel. The binding of an antigen with its complementary antibody results in the formation of a precipitin arc.

The sensitivity of immunoelectrophoresis can be increased by incorporating the antiserum in the gel. The formation of a precipitin line in this method is seen in the form of a rocket when an electric field is applied. This modified version is therefore named as "Rocket Immunoelectrophoresis". The precipitation assays while simple to perform are not very sensitive. Therefore, these assays have been replaced with more sensitive immunoassays.

Immunoassays

Immunoassays involve mixing of an antibody with an antigen followed by detection of the antigen-antibody complex with the help of a labelled antigen or a labelled antibody. These assays can be employed not only to analyze antibody responses during infection with a pathogen or upon vaccination with an antigen but also to detect antigens in biological fluids. Antibodies have over the years provided extremely powerful reagents in designing assays for laboratory diagnosis of infections and other diseased states. In 1960, Yalow and Berson first reported use of radiolabel insulin for the detection of insulin in biological fluids. The presence of insulin in a biological sample was detected by allowing the sample to compete with radiolabel insulin for binding to anti-insulin antibody. By using a set of standards, it was possible to determine the concentration of insulin in the unknown sample. The radioimmunoassay in effect laid the foundation for developing immunoassays. In recent years, this technique has been largely replaced by enzyme linked immunosorbent assay (ELISA) which uses an enzyme instead of a radioisotope. ELISAs combine the specificity of antibodies with the sensitivity of simple enzyme assays. Coupled with monoclonal antibodies, enzyme immunoassays have literally revolutionized laboratory diagnosis of infectious diseases. These assays can be

classified as either competitive or non-competitive assays. In the competitive ELISA, unlabeled and labelled antigens compete for a limited number of binding sites on an antibody. In the non-competitive type of ELISA, the antigen (or antibody) to be measured is first allowed to react with antibody (or antigen) on a solid phase followed by measurement of the binding of enzyme-labelled antibody. One of the most commonly used non-competitive ELISA is the "sandwich ELISA." In this assay, immobilized antibody is incubated with either the standard or the test antigen. After washing, the immobilized antibody-antigen complex is incubated with an excess of enzyme-labelled antibody which binds to remaining antigenic sites. After washing away excess unbound labelled antibody, the bound labelled antibody is detected by an enzymatic reaction that converts a colourless substrate into a coloured reaction product.

NOTES

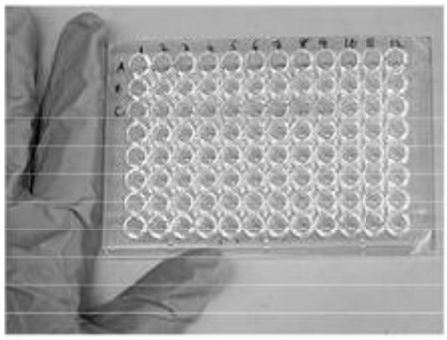


Fig. 8.4 A 96-well Microtiter Plate being used for ELISA

Indirect ELISA for the detection of specific antibody

This assay is extensively used to screen immune sera or B cell hybridomas for specific antibodies. Antigen is coated into the wells of a microtiter ELISA plate, followed by incubation with antibodies. The unbound antibody is washed and the bound antibody is detected by adding enzyme-labelled antibodies against the test antibody. This is followed by the determination of enzymatic activity.

Competitive ELISA

A third use of ELISA is through competitive binding. The steps for this ELISA are somewhat different than the first two examples:

1. Unlabelled antibody is incubated in the presence of its antigen (Sample)
2. These bound antibody/antigen complexes are then added to an antigen coated well.

NOTES

3. The plate is washed, so that unbound antibody is removed. (The more antigen in the sample, the less antibody will be able to bind to the antigen in the well, hence "competition.")
4. The secondary antibody, specific to the primary antibody is added. This second antibody is coupled to the enzyme.
5. A substrate is added, and remaining enzymes elicit a chromogenic or fluorescent signal.

For competitive ELISA, the higher the sample antigen concentration, the weaker the eventual signal. The major advantage of a competitive ELISA is the ability to use crude or impure samples and still selectively bind any antigen that may be present.

The labeled antigen competes for primary antibody binding sites with your sample antigen (unlabelled). The more antigen in the sample, the less labeled antigen is retained in the well and the weaker the signal).Commonly the antigen is not first positioned in the well.

Multiple and Portable ELISA (M&P ELISA)

A new technique uses a solid phase made up of an immunosorbent polystyrene rod with 8-12 protruding ogives. The entire device is immersed in a test tube containing the collected sample and the following steps (washing, incubation in conjugate and incubation in chromogenous) are carried out by dipping the ogives in microwells of standard microplates pre-filled with reagents.

The advantages of this technique are as follows:

1. The ogives can each be sensitized to a different reagent, allowing the simultaneous detection of different antibodies and / or different antigens for multi-target assays;
2. The sample volume can be increased to improve the test sensitivity in clinical (saliva, urine), food (bulk milk, pooled eggs) and environmental (water) samples;
3. One ogive is left unsensitized to measure the non-specific reactions of the sample;
4. The use of laboratory supplies for dispensing sample aliquots, washing solution and reagents in microwells is not required, facilitating the development of ready-to-use lab-kits and on-site kits.

Common applications of ELISA

ELISA is routinely used for

- Analyzing antibody responses during infection or after vaccination
- Laboratory diagnosis of different diseases, detection of infectious agents and pathogen-derived molecules in sera and other biological fluids.
- Determining antibody production by B cell hybridomas.
- Estimating cytokine levels during an immune response or during diseased conditions.

Immunofluorescence

Immunofluorescence is a technique that allows visualization of a cell or a tissue section with the help of a specific antibody. Visualization is achieved either by directly labeling the antibody with a fluorescent dye such as fluorescein isothiocyanate (direct immunofluorescence) or by using a secondary antibody that has been labeled with the dye (indirect immunofluorescence). Fluorescent samples are examined under a fluorescence microscope.

The most common fluorescent dyes are fluorescein, which emits green light, Texas Red, which emits red light, and rhodamine and phycoerythrin (PE) which emit orange/red light. By using selective filters, light coming from only the fluorochrome is detected in the fluorescence microscope.

Flow Cytometry

Flow cytometry is a technique for detecting, counting and sorting light signals generated by microscopic particles as they flow in a fluid stream past a beam of laser light. It allows simultaneous multiparametric analysis of the physical and chemical characteristics of single cells flowing through an optical/electronic detection apparatus.

Cytometers use lasers as the source of illumination for their ability to provide intense focused illumination. Numerous detectors are aimed at the point where the stream of cells passes through the light beam; one in line with the light beam (Forward Scatter or FSC) and several perpendicular to it (Side Scatter or SSC) and one or more fluorescent detectors. Each particle passing through the beam scatters the light in some way, and fluorescent chemicals in the particle may be excited into emitting light at a lower frequency than the light source. This combination of scatters and fluorescent light is detected and fluctuations

NOTES

in brightness at each detector are analyzed. This gives information about the physical and chemical structure of the particle. For example FSC correlates with cell volume and SSC with inner complexity of the particle (like granularity or membrane roughness). Major applications of flow cytometry include analysis of specific cell populations in a mixture, determination of expression of cell surface or intracellular molecules, analysis of cell cycle and cell death. Flow cytometry is also used for cell sorting.

8.3.2 Production of Antibodies

Polyclonal antibodies

Antibodies can be produced by injecting experimental animals with an antigen of interest. The most widely used animal for generating polyclonal antibodies in a laboratory set-up is the rabbit. Large amounts of antibodies can be produced by immunizing sheep, goats or donkeys. For monoclonal antibodies, mice are the most extensively used animals. Small molecules (haptens) that can't generate antibodies on their own need to be covalently linked to a carrier protein that provides T cell help. The ability of an antigen to produce an antibody response is considerably increased when it is injected in an insoluble/particulate form. This is achieved by mixing it with immune potentiators/ immunostimulators called adjuvants; particulate antigens such as bacteria can be injected without an adjuvant.

Monoclonal antibodies

Polyclonal antibodies can be produced in large amounts with much ease and without a need for any specialized experimental facility. However, lack of reproducible characteristics and ill defined specificity are major problems with polyclonal anti-sera. These problems were overcome in 1975 when George Kohler and Cesar Milstein reported that antigen-specific B cells could be fused with cancerous myeloma cell lines that would lead to generation of immortal cell lines with an ability to secrete antibodies indefinitely. This technique now routinely called hybridoma technique, for which Kohler and Milstein were awarded Nobel Prize in Physiology and Medicine in 1984, has made a revolutionary impact not only in the field of immunology but in biomedical research in general. This technique has been extensively used to produce mouse

monoclonal antibodies but it is also possible to make rat, hamster, human and rabbit monoclonal antibodies as well.

Before outlining the procedure of this technique, a brief account of the principle of the technique is described. Spleen cells from an immunized mouse are fused with a mouse myeloma cell line which is a cancerous cell line of B cell origin. This cell line does not produce any antibody on its own. Another characteristic feature of this cell line is that it lacks the enzyme hypoxanthine guanosine phosphoribosyl transferase (HGPRT) and therefore cannot use salvage pathway of DNA synthesis. If the *de novo* pathway of DNA synthesis is inhibited in these cells, these cells die. It is this property which has been exploited to generate hybridomas. After fusion, three different types of hybrid cells are generated: spleen x spleen, spleen x myeloma and myeloma x myeloma. Fused cells are grown in a medium which inhibits *de novo* protein synthesis but can support salvage pathway. Spleen x spleen hybrids die because these cells have a short half life, myeloma x myeloma die in HAT (Hypoxanthine Aminopterin Thymidine) selection medium because these are HGPRT negative and aminopterin blocks their *de novo* pathway of DNA synthesis but spleen x myeloma hybrids survive because these have HGPRT and can therefore use the salvage pathway for growth and survival.

NOTES

8.3.3 Purification of Antibodies

Monoclonal antibodies can be purified from tissue culture fluids or ascitic fluids by a number of methods including salt fractionation, ion exchange chromatography, gel filtration or affinity chromatography. However, affinity chromatography using Protein A- Sepharose or Protein G- Sepharose offers the most convenient method and is widely used for purifying IgG class of antibodies. Protein A and Protein G bind to the Fc region of the antibody. These proteins have different affinities for antibodies from different animal species as well as for different isotypes from the same animal species.

8.3.4 Applications of Monoclonal Antibodies

Diagnostic tests

Once monoclonal antibodies for a given substance have been produced, they can be used to detect the presence of this substance. The Western blot test and immuno dot blot tests detect the protein on a membrane. They are also very useful in immunohistochemistry, which detect antigen in fixed tissue sections

and immunofluorescence test, which detect the substance in a frozen tissue section or in live cells.

Therapeutic treatment

NOTES

One possible treatment for cancer involves monoclonal antibodies that bind only to cancer cell-specific antigens and induce an immunological response against the target cancer cell. Such mAb could also be modified for delivery of a toxin, radioisotope, cytokine or other active conjugate; it is also possible to design bispecific antibodies that can bind with their Fab regions both to target antigen and to a conjugate or effector cell. In fact, every intact antibody can bind to cell receptors or other proteins with its Fc region.

Monoclonal Antibodies for Cancer

ADEPT, antibody directed enzyme prodrug therapy; ADCC, antibody dependent cell-mediated cytotoxicity; CDC, complement dependent cytotoxicity; MAb, monoclonal antibody; scFv, single-chain Fv fragment.

The illustration below shows all these possibilities:

MAbs approved by the FDA include

- Bevacizumab
- Cetuximab
- Panitumumab
- Traztuzumab
- Pertuzumab

Autoimmune diseases

Monoclonal antibodies used for autoimmune diseases include infliximab and adalimumab, which are effective in rheumatoid arthritis, Crohn's disease and ulcerative Colitis by their ability to bind to and inhibit TNF-α. Basiliximab and daclizumab inhibit IL-2 on activated T cells and thereby help preventing acute rejection of kidney transplants. Omalizumab inhibits human immunoglobulin E (IgE) and is useful in moderate-to-severe allergic asthma.

Monoclonal antibodies have not only contributed immensely to biological research but have also revolutionized the field of laboratory diagnosis. Some of their important applications are in

- The measurement of proteins, hormones and drugs in biological fluids.
- Tissue and blood typing.

- The identification of infectious agents.
- The analysis of cell surface antigens on leukaemia's, lymphomas and other kinds of cancers.
- Imaging tumors.
- Immunotherapy of cancers.

NOTES

8.4 Adjuvants

There are many commercially available immunologic adjuvants. Selection of specific adjuvants or types varies depending upon whether they are to be used for research and antibody production or in vaccine development. Adjuvants for vaccine use only need to produce protective antibodies and good systemic memory while those for antiserum production need to rapidly induce high titer, high avidity antibodies. No single adjuvant is ideal for all purposes and all have advantages and disadvantages. Adjuvant use generally is accompanied by undesirable side effects of varying severity and duration. Research on new adjuvants focuses on substances which have minimal toxicity while retaining maximum immunostimulation. Investigators should always be aware of potential pain and distress associated with adjuvant use in laboratory animals. The most frequently used adjuvants for antibody production are Freund's, Alum, the Ribi Adjuvant System and Titermax.

Freund's adjuvants

There are two basic types of Freund's adjuvants: Freund's Complete Adjuvant (FCA) and Freund's Incomplete Adjuvant (FIA). FCA is a water-in-oil emulsion that localizes antigen for release periods up to 6 months. It is formulated with mineral oil, the surfactant mannide monoleate and heat killed *Mycobacterium tuberculosis*, *Mycobacterium butyricum* or their extracts (for aggregation of macrophages at the inoculation site). This potent adjuvant stimulates both cell mediated and humoral immunity with preferential induction of antibody against epitopes of denatured proteins. Although FCA has historically been the most widely used adjuvant, it is one of the more toxic agents due to non-metabolizable mineral oil and it induces granulomatous reactions. Its use is limited to laboratory animals and it should be used only with weak antigens. It should not be used more than once in a single animal

NOTES

since multiple FCA inoculations can cause severe systemic reactions and decreased immune responses. Freund's Incomplete Adjuvant has the same formulation as FCA but does not contain mycobacterium or its components. FIA usually is limited to booster doses of antigen since it normally much less effective than FCA for primary antibody induction. Freund's adjuvants are normally mixed with equal parts of antigen preparations to form stable emulsions.

Ribi Adjuvant System

Ribi adjuvants are oil-in-water emulsions where antigens are mixed with small volumes of metabolizable oil (squalene) which are then emulsified with saline containing the surfactant Tween 80. This system also contains refined mycobacterial products (cord factor, cell wall skeleton) as immunostimulants and bacterial monophosphoryl lipid A. Three different species oriented formulations of the adjuvant system are available. These adjuvants interact with membranes of immune cells resulting in cytokine induction, which enhances antigen uptake, processing and presentation. This adjuvant system is much less toxic and less potent than FCA but generally induces satisfactory amounts of high avidity antibodies against protein antigens.

Titermax

Titermax represents a newer generation of adjuvants that are less toxic and contain no biologically derived materials. It is based upon mixtures of surfactant acting, linear, blocks or chains of nonionic copolymers polyoxypropylene (POP) and polyoxyethylene (POE). These copolymers are less toxic than many other surfactant materials and have potent adjuvant properties which favor chemotaxis, complement activation and antibody production. Titermax adjuvant forms a microparticulate water-in-oil emulsion with a copolymer and metabolizable squalene oil. The copolymer is coated with emulsion stabilizing silica particles which allows for incorporation of large amounts of a wide variety of antigenic materials. The adjuvant active copolymer forms hydrophilic surfaces, which activate complement, immune cells and increased expression of class II major histocompatibility molecules on macrophages. Titermax presents antigen in a highly concentrated form to the

immune system, which often results in antibody titers comparable to or higher than FCA.

Specol

Specol is water in oil adjuvant made of purified mineral oil. It has been reported to induce immune response comparable to Freund's adjuvant in rabbit and other research animal while producing fewer histological lesions.

NOTES

8.4.1 Adjuvants and their modes of action

The trend towards the use of peptides and subunit proteins in modern vaccine design has necessitated the use of immunological adjuvants to achieve effective immunity. Aluminium hydroxide, a component of the diphtheria, tetanus and hepatitis B vaccines, was first described as an adjuvant over 60 years ago and is the only adjuvant currently approved for use in humans. It is also a common component of many veterinary vaccines. While this adjuvant is effective at enhancing antibody titres to antigens, the effectiveness of aluminium hydroxide is limited due to its inability to promote cell mediated immunity. Freund's Complete Adjuvant (FCA) has been used experimentally and does stimulate cellular immunity, but is unsuitable for human and veterinary use as it promotes, amongst other toxic side effects, local inflammation and granuloma formation at the site of injection. Thus, in recent years there has been a great deal of interest in developing novel, cheap, effective and safe adjuvants which stimulate cellular, as well as humoral immunity to be used with medical and veterinary vaccines. In addition, the recent unravelling of numerous immunological pathways has facilitated the rational development of new adjuvants and allowed a better understanding of the modes of action of traditional adjuvants.

8.5 In Situ Hybridization

In situ hybridization techniques allow specific nucleic acid sequences to be detected in morphologically preserved chromosomes, cells or tissue sections. In combination with immunocytochemistry, *in situ* hybridization can relate microscopic topological information to gene activity at the DNA, mRNA, and protein level. The technique was originally developed by Pardue and Gall (1969) and (independently) by John et al. (1969). At this time radioisotopes were the only labels available for nucleic acids, and autoradiography was the

NOTES

only means of detecting hybridized sequences. Furthermore, as molecular cloning was not possible in those days, *in situ* hybridization was restricted to those sequences that could be purified and isolated by conventional biochemical methods (e.g., mouse satellite DNA, viral DNA, ribosomal RNAs). Molecular cloning of nucleic acids and improved radiolabelling techniques has changed this picture dramatically. For example, DNA sequences a few hundred base pairs long can be detected in metaphase chromosomes by autoradiography. Also radioactive *in situ* techniques can detect low copy number mRNA molecules in individual cells. Some years ago, chemically synthesized, radioactively labelled oligonucleotides began to be used, especially for *in situ* mRNA detection.

In spite of the high sensitivity and wide applicability of *in situ* hybridization techniques, their use has been limited to research laboratories. This is probably due to the problems associated with radioactive probes, such as the safety measures required, limited shelf life, and extensive time required for autoradiography. In addition, the scatter inherent in radioactive decay limits the spatial resolution of the technique. However, preparing nucleic acid probes with a stable nonradioactive label removes the major obstacles which hinder the general application of *in situ* hybridization. Furthermore, it opens new opportunities for combining different labels in one experiment. The many sensitive antibody detection systems available for such probes further enhance the flexibility of this method. In this manual, therefore, we describe nonradioactive alternatives for *in situ* hybridization.

Direct and indirect methods

There are two types of nonradioactive hybridization methods: direct and indirect. In the direct method, the detectable molecule (reporter) is bound directly to the nucleic acid probe so that probe-target hybrids can be visualized under a microscope immediately after the hybridization reaction. For such methods it is essential that the probe-reporter bond survives the rather harsh hybridization and washing conditions. Perhaps more important, however, is, thatthe reporter molecule does not interfere with the hybridization reaction. The terminal fluorochrome labeling procedure of RNA probes developed by Bauman *et al.* (1980, 1984), and the direct enzyme labelling procedure of nucleic acids described by Renz and Kurz (1984) meet these criteria.

Boehringer Mannheim has introduced several fluorochrome-labelled nucleotides that can be used for labelling and direct detection of DNA or RNA probes. If antibodies against the reporter molecules are available, direct methods may also be converted to indirect immunochemical amplification methods.

NOTES

Indirect procedures require the probe to contain a reporter molecule, introduced chemically or enzymatically, that can be detected by affinity cytochemistry. Again, the presence of the label should not interfere with the hybridization reaction or the stability of the resulting hybrid. A few years ago, the chemical synthesis of oligonucleotides containing functional groups (e.g., primary aliphatic amines or sulfhydryl groups) was described. These can react with haptens, fluorochromes or enzymes to produce a stable probe which can be used for *in situ* hybridization experiments. Such oligonucleotide probes will undoubtedly be widely used as automated oligonucleotide synthesis makes them available to researchers not familiar with DNA recombinant technology. This manual concentrates on two labelling systems:

- Indirect methods using digoxigenin (detected by specific antibodies) and biotin (detected by streptavidin).
- Direct methods using fluorescein or other fluorochromes directly coupled to the nucleotide.

8.5.1 In situ hybridization to locate transcripts in cell type

Since the first description of in situ hybridization in 1969 the technique has advanced to allow sensitive detection of DNA and mRNA molecules at the cellular and subcellular levels. In particular fluorescence in situ hybridization (FISH) has become a frequently used tool in basic and applied biomedical research since detection is sensitive and allows discrimination of multiple targets in the same sample. By using RNA-FISH we have been able to detect primary transcripts of the human embryonic, fetal, and adult globins in erythroid cells to study the competitive transcription mechanism or variegated expression patterns of the human beta-globin locus. We have correlated such expression patterns with other parameters such as cell type, cell cycle, replication, and stage of differentiation by simultaneous detection of, e.g., incorporated BrdUTPs, proteins (e.g., cyclins A and E, PCNA, histones), and

globin (primary) transcripts and/or locus integration sites. Thus a combination of FISH and immunofluorescence methods allow the visualization of different processes taking place in the nucleus relative to each other in terms of three-dimensional space and structure and time (development, cell cycle).

NOTES

In situ hybridization allows detection of target mRNAs in cells by hybridization with a labeled anti-sense RNA probe obtained by in vitro transcription of the gene of interest. Here we outline a protocol for the in situ localization of gene expression in plants that is highly sensitivity and specific. It is optimized for use with paraformaldehyde fixed, paraffin-embedded sections, which give excellent preservation of histology, and DIG-labelled probes that are visualized by immuno-detection and alkaline-phosphatase colorimetric reaction. This protocol has been successfully applied to a number of tissues from a wide range of plant species, and can be used to analyze expression of mRNAs as well as small RNAs.

8.6 Fluorescence In Situ Hybridization (FISH)

Fluorescence *in situ* hybridization is a cytogenetic technique developed by biomedical researchers in the early 1980s that is used to detect and localize the presence or absence of specific DNA sequences on chromosomes. FISH uses fluorescent probes that bind to only those parts of the chromosome with which they show a high degree of sequence complementarity. Fluorescence microscopy can be used to find out where the fluorescent probe is bound to the chromosomes. FISH is often used for finding specific features in DNA for use in genetic counseling, medicine, and species identification. FISH can also be used to detect and localize specific RNA targets (mRNA,lncRNA and miRNA) in cells, circulating tumor cells, and tissue samples. In this context, it can help define the spatial-temporal patterns of gene expression within cells and tissues.

Probes – RNA and DNA

RNA probes can be designed for any gene or any sequence within a gene for visualization of mRNA, lncRNA and miRNA in tissues and cells. FISH is used by examining the cellular reproduction cycle, specifically interphase of the nuclei for any chromosomal abnormalities. This technique allows the analysis of a large series of archival cases much easier to identify the pinpointed chromosome by creating a probe with an artificial chromosomal foundation

that will attract similar chromosomes. The hybridization signals for each probe when a nucleic abnormality is detected. Each probe for the detection of mRNA and lncRNA is composed of 20 oligonucleotide pairs, each pair covering a space of 40–50 bit/s. For miRNA detection, the probes use proprietary chemistry for specific detection of miRNA and cover the entire miRNA sequence.

NOTES

Probes are often derived from fragments of DNA that were isolated, purified, and amplified for use in the Human Genome Project. The size of the human genome is so large, compared to the length that could be sequenced directly, that it was necessary to divide the genome into fragments. (In the eventual analysis, these fragments were put into order by digesting a copy of each fragment into still smaller fragments using sequence-specific endonucleases, measuring the size of each small fragment using size-exclusion chromatography, and using that information to determine where the large fragments overlapped one another.) To preserve the fragments with their individual DNA sequences, the fragments were added into a system of continually replicating bacteria populations. Clonal populations of bacteria, each population maintaining a single artificial chromosome, are stored in various laboratories around the world. The artificial chromosomes (BAC) can be grown, extracted, and labelled, in any lab. These fragments are on the order of 100 thousand base-pairs, and are the basis for most FISH probes.

Preparation and hybridization process – RNA

Cells, circulating tumor cells (CTCs), or formalin-fixed paraffin-embedded (FFPE) or frozen tissue sections are fixed, and then permeabilized to allow target accessibility. FISH has also been successfully done on unfixed cells. A target-specific probe, composed of 20 oligonucleotide pairs, hybridizes to the target RNA(s). Separate but compatible signal amplification systems enable the multiplex assay (up to two targets per assay). Signal amplification is achieved via a series of sequential hybridization steps. At the end of the assay the tissue samples are visualized under a fluorescence microscope.

Preparation and hybridization process – DNA

First, a probe is constructed. The probe must be large enough to hybridize specifically with its target but not so large as to impede the hybridization process. The probe is tagged directly with fluorophores, with targets

NOTES

for antibodies or with biotin. Tagging can be done in various ways, such as nick translation, or PCR using tagged nucleotides. Then, an interphase or metaphase chromosome preparation is produced. The chromosomes are firmly attached to a substrate, usually glass. Repetitive DNA sequences must be blocked by adding short fragments of DNA to the sample. The probe is then applied to the chromosome DNA and incubated for approximately 12 hours while hybridizing. Several wash steps remove all unhybridized or partially hybridized probes. The results are then visualized and quantified using a microscope that is capable of exciting the dye and recording images.

If the fluorescent signal is weak, amplification of the signal may be necessary in order to exceed the detection threshold of the microscope. Fluorescent signal strength depends on many factors such as probe labelling efficiency, the type of probe, and the type of dye. Fluorescently tagged antibodies or streptavidin are bound to the dye molecule. These secondary components are selected so that they have a strong signal.

FISH experiments designed to detect or localize gene expression within cells and tissues rely on the use of a reporter gene, such as one expressing green fluorescent protein, to provide the fluorescence signal.

Variations on probes and analysis

FISH is a very general technique. The differences between the various FISH techniques are usually due to variations in the sequence and labelling of the probes; and how they are used in combination. Probes are divided into two generic categories: cellular and a cellular. "In situ" in Fluorescence in situ hybridization refers to the placement of the probe placed cellularly. FISH is usually associated with nucleic acid based-type methods that can be further broken down to antimicrobial peptides. These few modifications make possible all FISH techniques.

Probe size is important because longer probes hybridize less specifically than shorter probes, "a short strand of DNA or RNA (often 10-25 nucleotides) which is complementary to a given target sequence, it can be used to identify or locate the target.". The overlap defines the resolution of detectable features. For example, if the goal of an experiment is to detect the breakpoint of a translocation, then the overlap of the probes-the degree to which one DNA

sequence is contained in the adjacent probes-defines the minimum window in which the breakpoint may be detected.

The mixture of probe sequences determines the type of feature the probe can detect. Probes that hybridize along an entire chromosome are used to count the number of a certain chromosome, show translocations, or identify extra-chromosomal fragments of chromatin. This is often called "whole-chromosome painting." If every possible probe is used, every chromosome, (the whole genome) would be marked fluorescently, which would not be particularly useful for determining features of individual sequences. However, it is possible to create a mixture of smaller probes that are specific to a particular region (locus) of DNA; these mixtures are used to detect deletion mutations. When combined with a specific colour, a locus-specific probe mixture is used to detect very specific translocations. Special locus-specific probe mixtures are often used to count chromosomes, by binding to the centromeric regions of chromosomes, which are unique enough to identify each chromosome.

NOTES

A variety of other techniques use mixtures of differently coloured probes. A range of colours in mixtures of fluorescent dyes can be detected, so each human chromosome can be identified by a characteristic colour using whole-chromosome probe mixtures and a variety of ratios of colours. Although there are more chromosomes than easily distinguishable fluorescent dye colours, ratios of probe mixtures can be used to create secondary colours. Similar to comparative genomic hybridization, the probe mixture for the secondary colours is created by mixing the correct ratio of two sets of differently coloured probes for the same chromosome. This technique is sometimes called M-FISH.

The same physics that make a variety of colours possible for M-FISH can be used for the detection of translocations. That is, colours that are adjacent appear to overlap; a secondary colour is observed. Some assays are designed so that the secondary colour will be present or absent in cases of interest. An example is the detection of BCR/ABL translocations, where the secondary colour indicates disease. This variation is often called double-fusion FISH or D-FISH. In the opposite situation—where the absence of the secondary colour is pathological—is illustrated by an assay used to investigate translocations where only one of the breakpoints is known or constant. Locus-specific probes are made for one side of the breakpoint and the other intact chromosome. In

normal cells, the secondary colour is observed, but only the primary colours are observed when the translocation occurs. This technique is sometimes called "break-apart FISH".

Stellaris (R) RNA FISH probes

NOTES

Stellaris RNA FISH, formerly known as Single Molecule RNA FISH, is a method of detecting and quantifying mRNA and other long RNA molecules in a thin layer of tissue sample. Targets can be reliably imaged through the application of multiple short singly labelled oligonucleotide probes. The binding of up to 48 fluorescent labelled oligos to a single molecule of mRNA provides sufficient fluorescence to accurately detect and localize each target mRNA in a wide-field fluorescent microscopy image. Probes not binding to the intended sequence do not achieve sufficient localized fluorescence to be distinguished from background.

Single-molecule RNA FISH assays can be performed in simplex or multiplex, and can be used as a follow-up experiment to quantitative PCR, or imaged simultaneously with a fluorescent antibody assay. The technology has potential applications in cancer diagnosis, neuroscience, gene expression analysis, and companion diagnostics.

8.6.1 Types of FISH

Fiber FISH

In an alternative technique to interphase or metaphase preparations, fiber FISH, interphase chromosomes are attached to a slide in such a way that they are stretched out in a straight line, rather than being tightly coiled, as in conventional FISH, or adopting a random conformation, as in interphase FISH. This is accomplished by applying mechanicalshear along the length of the slide, either to cells that have been fixed to the slide and then lysed, or to a solution of purified DNA. A technique known as chromosome combing is increasingly used for this purpose. The extended conformation of the chromosomes allows dramatically higher resolution - even down to a few kilobases. The preparation of fiber FISH samples, although conceptually simple, is a rather skilled art, and only specialized laboratories use the technique routinely.

Q-FISH

Q-FISH combines FISH with PNAs and computer software to quantify fluorescence intensity. This technique is used routinely in telomere length research.

Flow-FISH

NOTES

Flow-FISH uses flow cytometry to perform FISH automatically using per-cell fluorescence measurements.

8.6.2 Applications of FISH

In Medical area

Often parents of children with a developmental disability want to know more about their child's conditions before choosing to have another child. These concerns can be addressed by analysis of the parents' and child's DNA. In cases where the child's developmental disability is not understood, the cause of it can potentially be determined using FISH and cytogenetic techniques. Examples of diseases that are diagnosed using FISH include Prader-Willi syndrome, Angelman syndrome, 22q13 deletion syndrome, chronic myelogenous leukemia, acute lymphoblastic leukemia, Cri-du-chat, Velocardiofacial syndrome, and Down syndrome. FISH on sperm cells is indicated for men with an abnormal somatic or meiotic karyotype as well as those with oligozoospermia, since approximately 50% of oligozoospermic men have an increased rate of sperm chromosome abnormalities. The analysis of chromosomes 21, X, and Y is enough to identify oligozoospermic individuals at risk.

In medicine, FISH can be used to form a diagnosis, to evaluate prognosis, or to evaluate remission of a disease, such as cancer. Treatment can then be specifically tailored. A traditional exam involving metaphase chromosome analysis is often unable to identify features that distinguish one disease from another, due to subtle chromosomal features; FISH can elucidate these differences. FISH can also be used to detect diseased cells more easily than standard Cytogenetic methods, which require dividing cells and requires labor and time-intensive manual preparation and analysis of the slides by a technologist. FISH, on the other hand, does not require living cells and can be quantified automatically, a computer counts the fluorescent dots present. However, a trained technologist is required to distinguish subtle differences in

banding patterns on bent and twisted metaphase chromosomes. FISH can be incorporated into Lab-on-a-chip microfluidic device. This technology is still in a developmental stage but, like other lab on chip methods, it may lead to more portable diagnostic techniques.

NOTES

Species identification

FISH is often used in clinical studies. If a patient is infected with a suspected pathogen, bacteria, from the patient's tissues or fluids, are typically grown on agar to determine the identity of the pathogen. Many bacteria, however, even well-known species, do not grow well under laboratory conditions. FISH can be used to detect directly the presence of the suspect on small samples of patient's tissue.

FISH can also be used to compare the genomes of two biological species, to deduce evolutionary relationships. A similar hybridization technique is called a zoo blot. Bacterial FISH probes are often primers for the 16s rRNA region.

FISH is widely used in the field of microbial ecology, to identify microorganisms Biofilms, for example, are composed of complex (often) multi-species bacterial organizations. Preparing DNA probes for one species and performing FISH with this probe allows one to visualize the distribution of this specific species within the biofilm. Preparing probes (in two different colors) for two species allows visualize/study co-localization of these two species in the biofilm, and can be useful in determining the fine architecture of the biofilm.

8.7 Genomic In Situ Hybridization (GISH)

Schwarzacher (1989) successfully modified the fluorescent *in situ* hybridization (FISH) protocol as a new cytogenetic technique called genomic *in situ* hybridization (GISH). Total genomic DNA (gDNA) of a parental species is hybridized by simultaneous or subsequent reaction to chromosomes of a hybrid, where it enables to discriminate parental genomes. At the first time, GISH was used in a study of organization of parental genomes in an intergeneric *Hordeum-Secale* hybrid. During two decades, GISH has become a powerful tool for analyzing interspecific and intergeneric hybrids and allopolyploid species as well as introgression, addition and substitution lines. This method's efficacy is largely based on genome-specific dispersed repetitive

sequences. Repetitive sequences are a prominent component of nuclear genomes of plants. Many classes of dispersed repeats are shared among closely related species by a phylogenetic descend. In hybrids and allopolyploids, two or more genomes with a specific repertoire of repeats are merged together within a single nucleus. As the dispersed repeats evolve faster than genes, they enable to differentiate chromosomes from closely related species. Because genomic probes are able to differentiate between parental chromosomes in crossbreeds, they are widely used to identify the parents of hybrids or allopolyploids. Parental chromosomes could be distinguished by using parental gDNA as a probe in cases where the dispersed repeats in each of the parental genomes were highly diverged prior to interspecific hybridisation and when the hybrid or allopolyploid formed is itself relatively recently formed. When the parental species of a hybrid or allopolyploid are closely related, genomic probes less readily label individual parental chromosome sets. In this case a pre-hybridization with a blocking (competitor) DNA is needed. Unlabelled DNA, which could be also enriched for repetitive DNA sequences, effectively prevents *in situ* probe labelling to common dispersed repetitive sequences and thus enables parental chromosomes from more closely related species to be distinguished.

NOTES

Karyotyping by GISH

Karyotypes of many plant species are characterized by chromosomes of rather uniform size, shape and banding patterns. Considering classical banding methods only N banding (in only limited number of species) and C banding are feasible in plants. Besides FISH, whole gDNA can be used in some cases to karyotype chromosomes. Genome-specific repeats have frequently a non-random distribution, forming clusters within heterochromatin blocks. As the hybridization signals often coincide with C-bands, this approach is referred as GISH-banding. Based on GISH-banding, a universal reference karyotype for rye has been created. The GISH banding pattern coincided not only with Giemsa C-banding but also with DAPI patterns, whereby the fluorescence was brighter at the GISH bands. In some species with relatively small genomes, GISH probes label preferentially (peri) centromeric regions, leaving chromosome arms mostly unlabelled. In some species GISH labelling seems to be almost exclusively based on centromeric satellite repeats (centromeric

GISH, cenGISH). On the other hand, hybridization pattern of some repetitive sequences can mimic GISH. One particular repetitive sequence could be dispersed along the whole genome of only one parent; hence it is able to recognize parental genomes in hybrid.

NOTES

8.7.1 Phylogenetic Applications of GISH

A quite new area of GISH experiments is its utilization as a tool in phylogenetic studies, a new application of GISH which we propose to term GISH-phylogenetics. Molecular systematics based on DNA sequences is sometimes limited by the lack of sequence data. Moreover, different algorithms have been used to construct phylogenetic trees that provide for different assumptions and treat data differently. As a result, various phylogenetic approaches applied to the same dataset may generate different phylogenetic trees. In this case, an independent control system could help to decide, which tree represents the true species relationships. In natural polyploid species, of this age and older, distinguishing of parental genomes using GISH was not possible.

Other applications of GISH

- Chromosome mapping
- Genome analysis
- Phylogenetic relationships
- Analysis of somaclonal variations
- Detection of alien chromatin
- Detection of chromosomal aberration

Over the past 20 years GISH has become a very powerful and popular cytogenetic tool. The major overall differences between plant genomes are generated by repetitive DNA sequences, thus, genome-specific repeats within the total genomic DNA can be used as GISH probes to identify chromosome sets, chromosomes and chromosome segments of differential origin. At present, it is impossible to investigate the genome organization of most plant species using genome sequencing methods. GISH provides another tool to study genome organisation and chromosome structure in plants by showing the distribution pattern of characterized and uncharacterized gDNA sequences along chromosomes. The methods also offer new perspectives in phylogenetic

and taxonomic studies by determining and testing hypotheses of genomic relationships between wild and cultivated plant species. As genes are predominantly clustered in discrete blocks between various repetitive DNA elements, GISH can generate chromosome-specific banding patterns. Approaches like cenGISH and self- GISH are useful for elementary characterization of plant genomes, especially in species for which genomic sequence is incomplete or unavailable. Nucleotide sequence-based phylogenetic reconstructions can sometimes result in several alternative phylogenetic trees, depending on a sequence analyzed. As the log DA values are (in the case of *Silene*) well correlated with phylogenetic distances measured by the branch lengths, the GISH phylogenetic approach based on measuring log DA may provide additional information on the relationship of the tested species.

NOTES

8.8 Confocal Microscopy

A confocal microscope creates sharp images of a specimen that would otherwise appear blurred when viewed with a conventional microscope. This is achieved by excluding most of the light from the specimen that is not from the microscope's focal plane. The image has less haze and better contrast than that of a conventional microscope and represents a thin cross-section of the specimen. Thus, apart from allowing better observation of fine details it is possible to build three-dimensional reconstructions of a volume of the specimen by assembling a series of thin slices taken along the vertical axis. Confocal microscopy was pioneered by Marvin Minsky in 1955. Minsky's invention would perform a point-by-point image construction by focusing a point of light sequentially across a specimen and then collecting some of the returning rays. By illuminating a single point at a time Minsky avoided most of the unwanted scattered light that obscures an image when the entire specimen is illuminated at the same time. Additionally, the light returning from the specimen would pass through a second pinhole aperture that would reject rays that were not directly from the focal point. The remaining ''desirable'' light rays would then be collected by a photomultiplier and the image gradually reconstructed using a long-persistence screen.

NOTES

Modern Confocal Microscopy

Modern confocal microscopes have kept the key elements of Minsky's design: the pinhole apertures and point-by-point illumination of the specimen. Advances in optics and electronics have been incorporated into current designs and provide improvements in speed, image quality, and storage of the generated images. Although there are a number of different confocal microscope designs, this entry will discuss one general type-the other designs are not markedly different. The majority of confocal microscopes image either by reflecting light off the specimen or by stimulating fluorescence from dyes (fluorophores) applied to the specimen. The focus of this entry will be on fluorescence confocal microscopy as it is the mode that is most commonly used in biological applications. The difference between the two techniques is small. There are methods that involve transmission of light through the specimen, but these are much less common.

Fluorescence

If light is incident on a molecule, it may absorb the light and then emit light of a different colour, a process known as fluorescence. At ordinary temperatures most molecules are in their lowest energy state, the ground state. However, they may absorb a photon of light (for example, blue light) that increases their energy causing an electron to jump to a discrete singlet excited state. Fluorescein is a common fluorophore that acts this way, emitting green light when stimulated with blue excitation light. The wavelengths of the excitation light and the colour of the emitted light are material dependent. Microscopy in the fluorescence mode has several advantages over the reflected or transmitted modes. It can be more sensitive. Often, it is possible to attach fluorescent molecules to specific parts of the specimen, making them the only visible ones in the microscope and it is also possible to use more than one type of fluorophore.

8.8.1 Principles of Confocal Microscopy

The confocal principle in epi-fluorescence laser scanning microscope is diagrammatically presented. Coherent light emitted by the laser system (excitation source) passes through a pinhole aperture that is situated in a conjugate plane (confocal) with a scanning point on the specimen and a second pinhole aperture positioned in front of the detector (a photomultiplier tube). As

the laser is reflected by a dichromatic mirror and scanned across the specimen in a defined focal plane, secondary fluorescence emitted from points on the specimen (in the same focal plane) pass back through the dichromatic mirror and are focused as a confocal point at the detector pinhole aperture. The significant amount of fluorescence emission that occurs at points above and below the **Objectives** focal plane is not confocal with the pinhole and forms extended Airy disks in the aperture plane. Because only a small fraction of the out-of-focus fluorescence emission is delivered through the pinhole aperture, most of this extraneous light is not detected by the photomultiplier and does not contribute to the resulting image. The dichromatic mirror, barrier filter, and excitation filter perform similar functions to identical components in a widefield epi-fluorescence microscope. Refocusing the **Objectives** in a confocal microscope shifts the excitation and emission points on a specimen to a new plane that becomes confocal with the pinhole apertures of the light source and detector.

NOTES

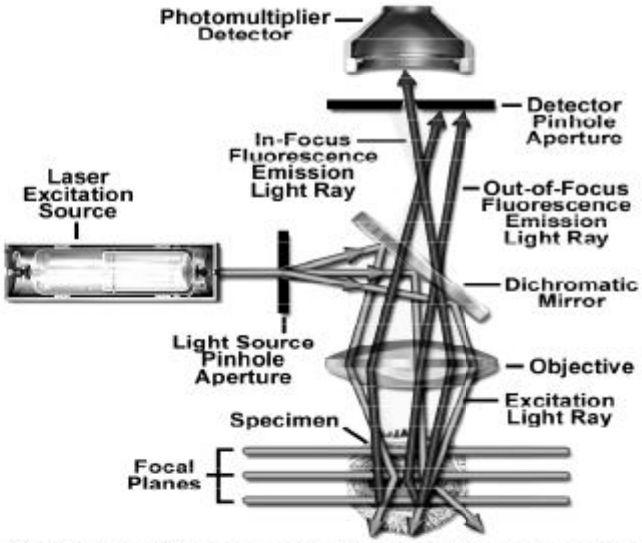


Fig. 8.5 : The Optical pathway and Principal Components in a Confocal Microscope

In traditional widefield epi-fluorescence microscopy, the entire specimen is subjected to intense illumination from an incoherent mercury or xenon arc-discharge lamp, and the resulting image of secondary fluorescence emission can be viewed directly in the eyepieces or projected onto the surface of an electronic array detector or traditional film plane. In contrast to this simple

NOTES

concept, the mechanism of image formation in a confocal microscope is fundamentally different. As discussed above, the confocal fluorescence microscope consists of multiple laser excitation sources, a scan head with optical and electronic components, electronic detectors (usually photomultipliers), and a computer for acquisition, processing, analysis, and display of images. The scan head is at the heart of the confocal system and is responsible for rasterizing the excitation scans, as well as collecting the photon signals from the specimen that are required to assemble the final image. A typical scan head contains inputs from the external laser sources, fluorescence filter sets and dichromatic mirrors, a galvanometer-based raster scanning mirror system, variable pinhole apertures for generating the confocal image, and photomultiplier tube detectors tuned for instruments include diffraction gratings or prisms coupled with slits positioned near the photomultipliers to enable spectral imaging followed by linear unmixing of emission profiles in specimens labeled with combinations of fluorescent proteins or fluorophores having overlapping spectra.

8.8.2 Confocal Microscope Configuration

Basic microscope optical system characteristics have remained fundamentally unchanged for many decades due to engineering restrictions on Objectives design, the static properties of most specimens, and the fact that resolution is governed by the wavelength of light. However, fluorescent probes that are employed to add contrast to biological specimens and, and other technologies associated with optical microscopy techniques, have improved significantly. The explosive growth and development of the confocal approach is a direct result of a renaissance in optical microscopy that has been largely fueled by advances in modern optical and electronics technology. Among these are stable multi-wavelength laser systems that provide better coverage of the ultraviolet, visible, and near infrared spectral regions, improved interference filters (including dichromatic mirrors, barrier, and excitation filters), sensitive low-noise wide band detectors, and far more powerful computers. The latter are now available with relatively low-cost memory arrays, image analysis software packages, high-resolution video displays, and high quality digital image printers.

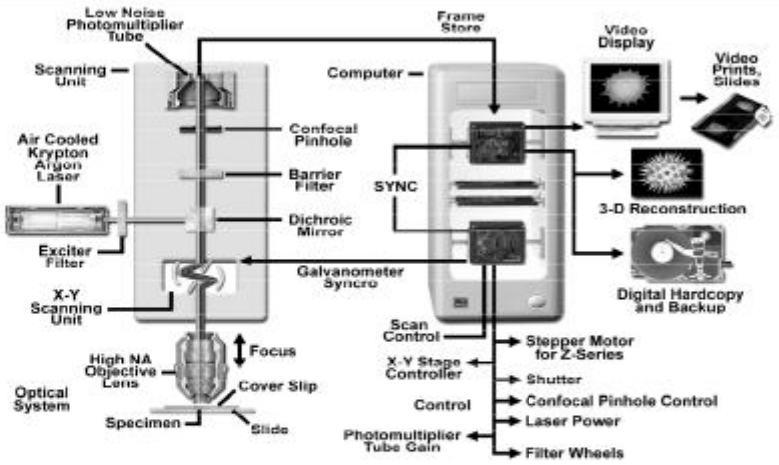


Fig. 8.6 : Confocal Microscope Configuration and Information Flow

8.8.3 Advantages and Disadvantages of Confocal Microscopy

The primary advantage of laser scanning confocal microscopy is the ability to serially produce thin (0.5 to 1.5 micrometer) optical sections through fluorescent specimens that have a thickness ranging up to 50 micrometers or more. The image series is collected by coordinating incremental changes in the microscope fine focus mechanism (using a stepper motor) with sequential image acquisition at each step. Image information is restricted to a well-defined plane, rather than being complicated by signals arising from remote locations in the specimen. Contrast and definition are dramatically improved over widefield techniques due to the reduction in background fluorescence and improved signal-to-noise. Furthermore, optical sectioning eliminates artifacts that occur during physical sectioning and fluorescent staining of tissue specimens for traditional forms of microscopy. The non-invasive confocal optical sectioning technique enables the examination of both living and fixed specimens under a variety of conditions with enhanced clarity. With most confocal microscopy software packages, optical sections are not restricted to the perpendicular lateral (x-y) plane, but can also be collected and displayed in transverse planes. Vertical sections in the x-z and y-z planes (parallel to the microscope optical axis) can be readily generated by most confocal software programs. Thus, the specimen appears as if it had been sectioned in a plane that is perpendicular to the lateral axis. In practice, vertical sections are obtained by combining a series of x-y scans taken along the z axis with the software, and then projecting a

NOTES

view of fluorescence intensity as it would appear should the microscope hardware have been capable of physically performing a vertical section.

8.9 Summary

NOTES

This unit will attempt to summarize current understanding of the immunotechniques and hybridizations. Hybridization has been widely used to transfer genes from wild species into cultivated sunflower. Fluorescent genomic *in situ* hybridization (GISH) has been used to identify alien chromosomes or segments in other crops, but an equivalent technique for sunflower is lacking. The **Objectives** of this study was to develop a GISH procedure for identifying chromosomes or chromosome segments of wild species in the background of cultivated sunflower.

Genetic engineering of commercially important crops has become routine in many laboratories. However, the inability to predict where a transgene will integrate and to efficiently select plants with stable levels of transgenic expression remains a limitation of this technology. Fluorescence *in situ* hybridization (FISH) is a powerful technique that can be used to visualize transgene integration sites and provide a better understanding of transgene behavior. Studies using FISH to characterize transgene integration have focused primarily on metaphase chromosomes, because the number and position of integration sites on the chromosomes are more easily determined at this stage. However gene (and transgene) expression occurs mainly during interphase. In order to accurately predict the activity of a transgene, it is critical to understand its location and dynamics in the threedimensional interphase nucleus.

Interspecific hybridization has been widely used to transfer genes from wild species into cultivated sunflower. Fluorescent genomic *in situ* hybridization (GISH) has been used to identify alien chromosomes or segments in other crops, but an equivalent technique for sunflower is lacking. The **Objectives** of this study was to develop a GISH procedure for identifying chromosomes or chromosome segments of wild species in the background of cultivated sunflower.

Laser scanning confocal microscopy has become an invaluable tool for a wide range of investigations in the biological and medical sciences for imaging thin

optical sections in living and fixed specimens ranging in thickness up to 100 micrometers. Modern instruments are equipped with 3-5 laser systems controlled by high-speed acousto-optic tunable filters (AOTFs), which allow very precise regulation of wavelength and excitation intensity.

NOTES

8.10 Glossary

- **Fluorescence :** If light is incident on a molecule, it may absorb the light and then emit light of a different colour, a process known as fluorescence.
- **Hybridisation :** The process of combining different varieties of organisms to create a hybrid
- **Immunology :** The study of the immune system and its reaction to pathogens, as well as its malfunctions.
- **Antigens :** An antigen(Ag), or antibody generator, is any substance which provokes an adaptive immune response.
- **Antibody :**Used by the immune system to identify and neutralize foreign objects such as bacteria and viruses.
- **ISH :** In Situ Hybridisation
- **FISH :** Fluorescence In Situ Hybridisation
- **GISH :** Genomic In Situ Hybridisation

8.11 Self-Learning Exercise

Section – A (Very Short Answer Type Questions)

1. Which type of cell produce antibody?
2. Who give the name of antigens?
3. Give the name of two adjuvants for antibody production.

Section – B (Short Answer Type Questions)

1. What are antigens?
2. Describe the adjuvants mode of action.
3. Describe direct and indirect methods for In Situ Hybridisaton.
4. Describe In Situ Hybridisation to locate transcript.

NOTES

Section – C (Long Answer Type Questions)

1. Give the detail description of FISH.
2. Describe GISH.
3. Give the illustrative description of confocal microscopy.

Section – A (Answer Key)

1. B-cells.
2. Ladislas Deutsch.
3. Freund's adjuvants and Ribi adjuvants.

8.12 References

- A. R. Hibbs, 2004. *Confocal Microscopy for Biologists*. Kluwer Academic, New York
- Agrawal, S.; Christodoulow, C.; Gait, M. J. (1986). Efficient methods for attaching nonradioactive labels to the 5'-ends of synthetic oligodeoxyribonucleotides. Nucleic Acids Res. 14, 6227–6245.
- B. R. Masters (1996), *Selected Papers on Confocal Microscopy*, SPIE Milestone Series, Volume MS 131, Bellingham, Washington: SPIE Optical Engineering Press.
- Baumann, J. G. J.; Van der Ploeg, M.; Van Duijn, P. (1984) Fluorescent hybridocytochemical procedures: DNA-RNA hybridization in situ. In: Chayen, J.; Bitensky, L. (Eds.) Investigative Microtechniques in Medicine and Microbiology. New York: Marcel Dekker, 41–48.
- Baumann, J. G. J.; Wiegant, J.; Borst, P.; Van Duijn, P. (1980) A new method for fluorescence microscopical localization of specific DNA sequences by in situ hybridization of fluorochrome-labeled RNA. Exp. Cell Res. 138, 485–490.
- Charles A. Janeway Jr. and Paul Travers. Immunobiology. Garland Publishing Inc., New York.
- Chollet, A.; Kawashima, E. M. (1985) Biotin-labeled synthetic oligodeoxyribonucleotides: synthesis and uses as hybridization probes. Nucleic Acids Res. 13, 1529–1541.

- Coghlan, J. P.; Aldred, P.; Haralambridis, J.; Niall, H. D.; Penschow, J. D.; Tregear, G. W. (1985) Hybridization histochemistry. Anal. Biochem. 149, 1–28.
- Current protocols in Immunology. Edited by John C. Coligan, Barbara E. Beier, David H. Marguiles, Ethan M. Shevach and Warren Strober. John Wiley and Sons, Inc.
- Ed Harlow and David Lane. Antibodies – A Laboratory Manual. Cold Spring Harbor Laboratory, USA.
- G.P.Talwar. *A Handbook of Practical Immunology*. Vikas Publishing House Pvt. Ltd., India.
- Kohler G and Milstein C. 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. Nature 256: 495-7.
- Leslie Hudson and Frank C. Hay. *Handbook of Experimental Immunology*. Blackwell Scientific Publications, Oxford, London.
- Leslie Hudson and Frank C. Hay. *Practical Immunology*. Blackwell Scientific Publications, Oxford, London.
- M. Gu. *Principles of Three-Dimensional Imaging in Confocal Microscopes*, New Jersey: World Scientific, 1996.
- M. Minsky, Microscopy Apparatus, US Pat. 3,013,467, 1961.
- M. Muller, *Introduction to Confocal Fluorescence Microscopy*, Maastricht, Netherlands: Shaker, 2002.
- P. M. Conn, *Confocal Microscopy*, in Methods in Enzymology, Volume 307, New York: Academic Press, 1999.
- T. R. Corle and G. S. Kino, *Confocal Scanning Optical Microscopy and Related Imaging Systems*, New York: Academic Press, 1996.
- T. Wilson (ed.), *Confocal Microscopy*, New York: Academic Press, 1990.

Unit-9

Chromatin Organization

NOTES

Structure of the unit

9.0 Objectives

After studying this unit you will be able to understand:

- Organization of Chromatin, Chromosome structure and packaging of DNA, Molecular organization of centromere and telomere,
- Euchrommatin and heterochromatin, Karyotypes of chromosomes,
- Polytene, lampbrush, B- chromosomes and Sex chromosome,
- Molecular basis of Chromosome Pairing

9.1 Introduction

A chromosome is packaged and organized chromatin, a complex of macromolecules found in cells, consisting of DNA and protein. The main

information-carrying macromolecule is a single piece of coiled double-stranded DNA, containing many genes, regulatory elements and other non-coding DNA. The DNA-bound macromolecules are proteins, which serve to package the DNA and control its functions. Chromosomes vary widely between different organisms. Some species also contain plasmids or other extrachromosomal DNA.Compaction of the duplicated chromosomes during cell division (mitosis or meiosis) results in a four-arm structure. Chromosomal recombination during meiosis and subsequent sexual reproduction plays a vital role in genetic diversity. If these structures are manipulated incorrectly, through processes known as chromosomal instability and translocation, the cell may undergo mitotic catastrophe and die, or it may unexpectedly evade apoptosis leading to the progression of cancer.

NOTES

9.2 Chromosomes Structure and Packaging of DNA

The DNA of an organism are packed in their cells to protected, they also regulate the accessibility of the DNA. Packaging of DNA helps conserve space in the cells. Approximately, two meters of the human DNA can fit into a cell that is only a few micrometers wide. Chromosomes are made up of DNA segments. Chromosomes carry all the information that help a cell grows, survive and reproduce. DNA segments with specific patterns are called genes. The chromosomes are found in the nucleus of the cell. In prokaryotic organisms, the DNA is not present in the nucleus; the DNA floats in the cytoplasm in area called the nucleoid. The chromosomes vary widely between different organisms. Eukaryotic cells have large number of linear chromosomes and cells of prokaryotes have smaller and circular DNA. Cells may contain more than one type of chromosome, like in most eukaryotic cells, the mitochondria and the chloroplasts in plant cells possess their own set of chromosomes. In nucleus of eukaryotic organism, the chromosomes are packed by proteins to form a compact structure called chromatin. This condensation allows long molecules of DNA to fit into the cell nucleus.Chromosomes are more condensed then the chromatin and they are essential for cell division. The chromosomes are replicated, divided and passed on to the daughter cells, to ensure genetic diversity and survival of the progeny. Duplicated chromosomes contain tow identical copies known as chromtids or sister chromatids, they are

joined by a centromere. Compaction of the chromosomes during the cell division process results in the four-arm structure. Recombination of chromosome plays a vital role in genetic diversity. Incorrect multiplication of the chromosomes may lead to mitotic failure or death of the cell, it may lead to apoptosis and sometimes may be cancerous.

NOTES

In the nucleus of each cell, the DNA molecule is packaged into thread-like structures called chromosomes. Each chromosome is made up of DNA tightly coiled many times around proteins called histones that support its structure.

9.3 Molecular Organization of Centromere and Telomere

Chromosomes are not visible in the cell's nucleus—not even under a microscope—when the cell is not dividing. However, the DNA that makes up chromosomes becomes more tightly packed during cell division and is then visible under a microscope. Most of what researchers know about chromosomes was learned by observing chromosomes during cell division.

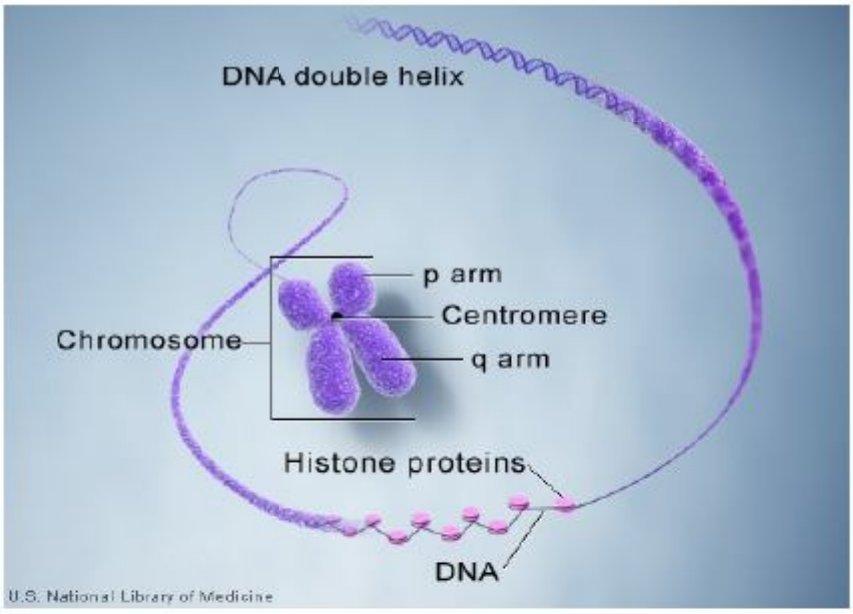


Fig. 9.1 : DNA and histone Proteins are packaged into structures called Chromosomes

Each chromosome has a constriction point called the centromere, which divides the chromosome into two sections, or "arms." The short arm of the chromosome is labeled the "p arm." The long arm of the chromosome is labeled the "q arm." The location of the centromere on each chromosome gives

the chromosome its characteristic shape, and can be used to help describe the location of specific genes.

Chromosomes are organized structure of DNA and proteins found in cells. they are thread-like structures located inside the nucleus of animal and plant cells. Chromosomes are made up of proteins and a molecule of deoxyribonucleic acid (DNA). Chromosomes are passed on from parents to offspring. The term chromosome is derived from a Greek word 'chroma' which means 'color' and 'soma' which means 'body'. The chromosomes are named so because they are cellular structures or cellular bodies and they are strongly stained by some dyes used in research. Chromosomes play an important role that ensures DNA is copied and distruuted accurately in the process of cell division. In most of the organisms chromosomes are arranged in pairs in the nucleus of the cell. We have 23 pairs of chromosomes.

NOTES

In eukarytoic cells, chromosomes are composed of single molecule of DNA with many copies of five types of histones.Histones are proteins molecules and are rich in lysine and arginine residues, they are positively charged. Hence they bind tightly to the negatively-charged phosphates in the DNA sequence. A small number of non-histone proteins are also present, these are mostly transcription factors. Transcription factors regulate which parts of DNA to be transcribed into RNA.

During most of the cell's life cycle, chromosomes are elongated and cannot be observed under the microscope. During the S phase of the mitotic cell cycle the chromosomes are duplicated. At the beginning of mitosis the chromosomes are duplicated and they begin to condense into short structures which can be stained and observed easily under the light microscope.These duplicated condensed chromosomes are known as dyads. The duplicated chromosomes are held together at the region of centromeres. The centromeres in humans are made of about 1-10 million base pairs of DNA. The DNA of the centromere are mostly repetitive short sequences of DNA, the sequences are repeated over and over in tandem arrays. The attached, duplicated chromosomes are commonly called sister chromatids. Kinetochores are the attachment point for spindle fibers which helps to pull apart the sister chromatids as the mitosis process proceeds to anaphase stage. The kinetochores are a complex of about

80 different proteins. The shorter arm of the two arms of the chromosome extending from the centromere is called the p arm and the longer arm is known the q arm.

NOTES

Bacterial Chromosome

Bacterial chromosomes contain circular DNA molecule unlike the linear DNA of vertebrates.Most of chromosomes are circular DNA molecules and there are no free ends to the DNA. The bacterial DNA is packaged into a single chromosome into a continuous loop. The DNA is folded or coiled to fit into the cell. The compaction of the DNA involves the binding of proteins to the DNA that help form initial loops which is then coiled.

Prokaryotic Chromosome

Prokaryotes like the bacteria and archaea typically have a single circular chromosome. The chromosome size of most bacteria is from only 160,000 base pairs to 12,200,00 base pairs.Some bacteria in exceptions contain a single linear chromosome. The base sequences in prokaryotic chromosomes are less than in eukaryotic cells. Bacterial chromosomes have a single origin of replication from which the replication starts. In some archaea there are multiple replication origins. The prokaryotic genes are organized into operons and it usually it does not contain introns. Nucleus is absent in prokaryotes, the DNA is organized into a structure called the nucleoid. The DNA of the archaea are more organized, they are packaged within structures similar to eukaryotic nucleosomes. The chromosomes in the prokaryotes and plasmids are generally supercoiled like that of the eukaryotes. The DNA are released into the relaxed state for the process of transcription, replication and regulation.

Eukaryotic Chromosomes

In eukaryotes the chromosomes are multiple large, linear and are present in the nucleus of the cell. Each chromosome typically has one centromere; one or two arms that project from the centromere, the arms are usually not visible during most of the time. Most of the eukaryotes have a small circular genome in the mitochondria. Some of the eukaryotes have small linear or circular chromosomes in the cytoplasm. To fit into the compartment in which it is contained the DNA has to be condensed and the degree to which it is condensed is expressed as its packaging ratio. Packaging ratio is the length of

the DNA divided by the length into which it is packaged. The chromosome of eukaryotic organisms consists of complexes made of DNA and protein, and it is organized in a condensed manner. This condensation permits the large amount of DNA to be stored in the nucleus of the cell.

NOTES

Human Chromosomes

Human's chromosomes are of two types autosomes and sex chromosomes. Genetic traits that are linked to the sex of the person are passed on through the sex chromosomes. The rest of the genetic information is present in the autosomes. Humans have 23 pairs of chromosomes in their cells, of which 22 pairs are autosomes and one pair of sex chromosomes, making a total of 46 chromosomes in each cell. Many copies of mitochondrial genome are present in human cells.

Sex Chromosomes

Sex chromosomes differ in form of size, behavior from the ordinary chromosome. The sex chromosomes determine the sex of an individual during reproduction. These sex chromosomes differ between the male and the females. Females have two copies of X chromosome; males have one X chromosome and one Y chromosome. In the process of sexual reproduction in humans, two different gametes fuse to form a zygote.

Human Sex Chromosomes

Homologous Chromosomes

Homologous chromosomes are also known as homologs or homologues. The homologous chromosomes are pairs of chromosomes that are approximately of the same length, position of centromere, and pattern of staining, genes for the same characteristic are at a corresponding loci. In an organism one of the homologous chromosomes is inherited from the mother and the other from the father. These chromosomes are usually not identical, but they carry the same type of genes.During the process of mitosis the daughter chromosomes carry the same sequence of nucleotide, assuming there are no errors during the replication process. The genome in diploid organisms is composed of homologous chromosomes One of homologous pair is the maternal chromosome and the other is the paternal chromosome. During the process of

meiosis the homologous chromosomes cross over. Homologous chromosomes are not identical but they are similar. The genes are carried in the same order, but the alleles for the trait may not be similar.

NOTES

Functions of Chromosomes are as follows:

Genetic Code Storage: Chromosome contains the genetic material that is required by the organism to develop and grow. DNA molecules are made of chain of units called genes. Genes are those sections of the DNA which code for specific proteins required by the cell for its proper functioning.

Sex Determination: Humans have 23 pairs of chromosomes out of which one pair is the sex chromosome. Females have two X chromosomes and males have one X and one Y chromosome. The sex of the child is determined by the chromosome passed down by the male. If X chromosome is passed out of XY chromosome, the child will be a female and if a Y chromosome is passed, a male child develops.

Control of Cell Division: Chromosomes check successful division of cells during the process of mitosis. The chromosomes of the parent cells insure that the correct information is passed on to the daughter cells required by the cell to grow and develop correctly.

Formation of Proteins and Storage: Proteins are essential for the activity of a cell. The chromosomes direct the sequences of proteins formed in our body and also maintain the order of DNA. The proteins are also stored in the coiled structure of the chromosomes. These proteins bound to the DNA help in proper packaging of the DNA.

9.4 Euchromatin and Heterochromatin

Chromatin

Chromatin is the collective name for the long strands of DNA, RNA and their associated nucleoproteins. During interphase of the cell cycle .chromatin is dispersed throughout the nucleus, becoming more compact during mitosis or meiosis. Two types of chromatin can be seen with electron microscopy.

Heterochromatin – This is electron dense and distributed around the periphery of the nucleus and in discrete masses within the nucleus. The DNA is in close association with nucleoproteins, and it is not active in RNA synthesis.

Euchromatin – This is electron lucent and represents DNA that is actually or potentially active in RNA synthesis.

NOTES

Nucleosomes

A nucleosome is formed by 146 bp of DNA wound twice around an octamer of histone proteins. The octamer consists of two copies each of the histone proteins H2A, H2B, H3, and H4. Histone proteins are conserved throughout eukaryotic evolution. They contain a high proportion of positively charged amino-acid residues that can form ionic bonds with the negatively charged DNA. This interaction does not depend on DNA sequence and theoretically histones can bind with any piece of DNA. However, *in vivo,* the position of histone binding is influenced by:

1.AT content (bends more easily than GC)

2.the presence of other tightly bound proteins.

Chromatin fibre organization. (A) The nucleosome core particle is composed of pairs of histones. 166 base pairs of DNA wind around each nucleosome. Linker DNA consisting of 8–114 base pairs runs between one nucleosome and the next. (B) Chromatin consists of nucleosomes bound together through their H1 proteins (not shown in this part of the figure). (C) Bound nucleosomes form a solenoid, with six nucleosomes per turn.

Nucleosome bound regions of DNA are separated by a region of linker DNA that varies from 0–80 bp in length. Consequently, on electron micrographs nucleosomes appear as 11-nm 'beads' on a 2-nm DNA 'string'.

DNaseI is an endonuclease that breaks the internal phosphodiester bonds in DNA, irrespective of its base sequence. The regulatory regions of genes are frequently bound by proteins that prevent histones from binding. Since histone binding protects DNA from degradation with DNaseI, the regulatory regions of genes are particularly sensitive to this enzyme, and they are sometimes called 'nuclease-hypersensitive sites'.

The second level of DNA packing is mediated by histone HI, binding together adjacent nucleosomes to condense DNA into the supercoiled 30-nm fibre,

which is also called the solenoid. The solenoid exhibits six to eight nucleosomes per turn of the spiral, corresponding to heterochromatin.

Giant supercoil

NOTES

The third level of organization is thought to involve the formation of transcriptional units of DNA loops radiating from a central scaffold of non-histone proteins.

Chromosomes

At metaphase chromatin is maximally condensed and forms 1400-nm fibres. After cell staining these structures are visible as chromosomes under light microscopy.

Centromeres

Each metaphase chromosome is composed of two identical sister chromatids. Chromatids are connected at a central region called the centromere, above and below which chromatin strands loop across and between chromatids to hold them together.

Centromere position can be used to categorize chromosomes morphologically •

Acrocentric – centromeres located very close to one end, yielding a small short arm, often associated with small pieces of DNA called satellites, encoding rRNA.

Metacentric – centromeres located in the middle, yielding arms of roughly equal length.

Submetacentric – off-centre centromere so that one arm is longer than the other.

The kinetochore is an organelle located at the centromere region. It acts as a microtubule organizing centre and facilitates spindle formation by polymerization of tubulin dimers to form microtubules early in mitosis.

The kinetochore- The kinetochore contains two regions: an inner kinetochore, which is tightly associated with the centromere DNA, and an outer kinetochore, which interacts with microtubules.

Telomeres

The ends of chromosomes are protected by DNA structures called telomeres. Telomeres are tandem repeats of the hexameric sequence 'TTAGGG', ending in a $3'$ single-stranded overhang that ranges in length from about 50–400

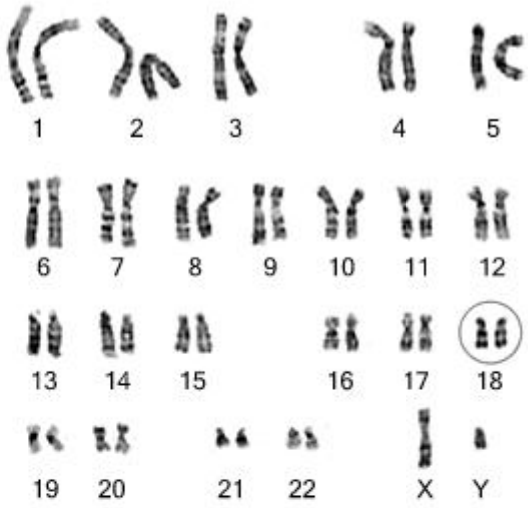
nucleotides and loops back on itself to form the T-loop . Telomeres have several functions in preserving chromosome stability, including:

1. preventing abnormal end-to-end fusion of chromosomes
2. protecting the ends of chromosomes from degradation
3. ensuring complete DNA replication
4. having a role in chromosome pairing during meiosis.

NOTES

9.5 Karyotype of Chromosome

A karyotype is simply a picture of a person's chromosomes. In order to get this picture, the chromosomes are isolated, stained, and examined under the microscope. Most often, this is done using the chromosomes in the white blood cells. A picture of the chromosomes is taken through the microscope. Then, the picture of the chromosomes is cut up and rearranged by the chromosome's size. The chromosomes are lined up from largest to smallest. A trained cytogeneticist can look for missing or extra pieces of chromosome.



Karyotype

There are 22 numbered pairs of chromosomes called autosomes>. The 23rd pair of chromosomes are the sex chromosomes. They determine an individual's gender. Females have two X chromosomes, and males have an X and a Y chromosome.

NOTES

Chromosomes Number

Each chromosome has been assigned a number based on its size. The largest chromosome is chromosome 1. Therefore chromosome 18 is one of the smallest chromosomes in humans.

Polytene chromosomes are giant chromosomes common to many dipteran (two-winged) flies. They begin as normal chromosomes, but through repeated rounds of DNA replication without any cell division (called endoreplication), they become large, banded chromosomes. For unknown reasons, the centromeric regions of the chromosomes do not endoreplicate very well. As a result, the centromeres of all the chromosomes bundle together in a mass called the chromocenter.

Polytene chromosomes are usually found in the larvae, where it is believed these many-replicated chromosomes allow for much faster larval growth than if the cells remained diploid. Simply because each cell now has many copies of each gene, it can transcribe at a much higher rate than with only two copies in diploid cells.

The polytene chromosomes at the right are from the salivary glands of the fruit fly *Drosophila melanogaster*. the bands on each chromosome are like a road map, unique to each chromosome and well defined enough to allow high resolution mapping of each chromosome. The *Drosophila* Genome Project uses polyene chromosomes as a framework for the map.

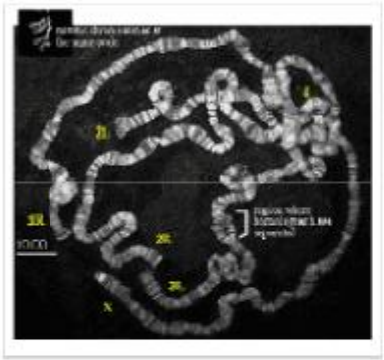


Fig. 9.2 : Polytene Chromosomes alongside normal Mitotic Chromosomes

9.6 Polytene Chromosomes

NOTES

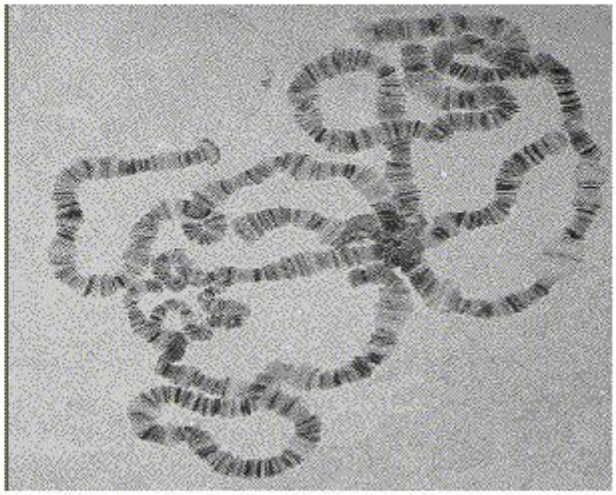


Fig. 9.3 : Drosophila Polytene Chromosomes stained with Orescein and viewed under transmitted light. These chromosomes came from a fly strain called ltx13, which has most of the left arm of chromosome 2 on the right arm of chromosome 3.

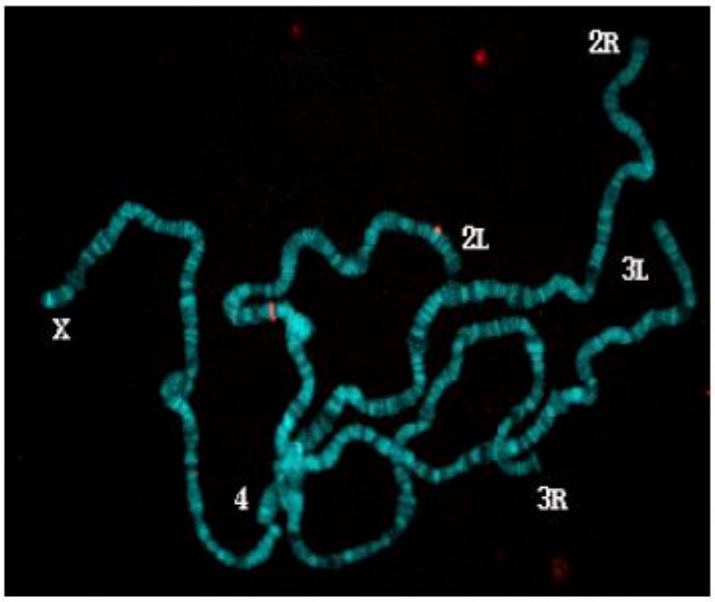


Fig. 9.4 : Low Magnification view of Polytene Chromosomes

9.7 Lampbrush Chromosome

Lampbrush chromosome from the cell nucleus of an ovarial egg from Triton sp., a salamander.

NOTES

Lampbrush chromosomes (first seen by Flemming in 1882) are a special form of chromosomes that are found in the growing oocytes (immature eggs) of most animals, except mammals Lampbrush chromosomes of tailed and tailless amphibians, birds and insects are described best of all.Chromosomes transform into the lampbrush form during the diplotene stage of meiotic prophase I due to an active transcription of many genes. They are highly extended meiotic half-bivalents, each consisting of 2 sister chromatids. Lampbrush chromosomes are clearly visible even in the light microscope, where they are seen to be organized into a series of chromomeres with large chromatin loops extended laterally. Amphibian and avian lampbrush chromosomes can be microsurgically isolated from oocyte nucleus (germinal vesicle) with either forceps or needles.

A given loop always contains the same DNA sequence, and it remains extended in the same manner as the oocytes grows. These chromosomes are producing large amounts of RNA for the oocyte, and most of the genes present in the DNA loops are being actively expressed. Each lateral loop contains one or several transcription units with polarized RNP-matrix coating the DNA axis of the loop. The majority of the DNA, however, is not in loops but remains highly condensed in the chromomeres on the axis, where genes are generally not expressed.

It is thought that the interphase chromosomes of all eukaryotes are similarly arranged in loops. Although these loops are normally too small and fragile to be easily observed in a light microscope, other methods can be used to infer their presence. For example, it has become possible to assess the frequency with which two loci along an interphase chromosome are paired with each other, thus revealing candidates for the sites on chromatin that form the closely apposed bases of loop structures. These experiments and others suggest that the DNA in human chromosomes is organized into loops of different lengths. A typical loop might contain between 50,000 and 200,000 nucleotide pairs of DNA, although loops of a million nucleotide pairs have also been suggested.

Giant chromosomes in the lampbrush form are useful model for studying

chromosome organization, genome function and gene expression during meiotic prophase, since they allow the individual transcription units to be visualized. Moreover lampbrush chromosomes are widely used for high-resolution mapping of DNA sequences and construction of detail cytological maps of individual chromosomes.

NOTES

9.8 B Chromosome

In addition to the normal karyotype, wild populations of many animal, plant, and fungi species contain B chromosomes (also known as supernumerary or accessory chromosomes). These chromosomes are not essential for the life of a species, and are lacking in some (usually most) of the individuals. Thus a population would consist of individuals with 0, 1, 2, 3 (etc.) supernumeraries.

Most B chromosomes are mainly or entirely heterochromatic (and so would be largely non-coding), but some, such as the B chromosomes of maize, contain size-able euchromatic segments. In general it seems unlikely that super numeraries would persist in a species unless there was some positive adaptive advantage, which in a few cases has been identified. For instance, the British grasshopper *Myrmeleotettixmaculatus* has two structural types of B chromosomes: metacentrics and submetacentrics. The supernumeraries, which have a satellite DNA, occur in warm, dry environments, and are scarce or absent in humid, cooler localities. In plants there is a tendency for B chromosomes to be present in the germ-line, but to be lost from other tissues such as root tips and leaves.

General properties of B chromosomes

Form and size-B chromosomes are smaller than A chromosomes except in a few cases, in which they are of equal size. They often have distinct centromere positions and can be readily identified at mitosis. Variants include b chromosomes as isochromosomes or telocentrics, and, in a few species, they appear as microchromosomes.

Structural polymorphism- the normal situation is for only one form of a B chromosome, with variants arising at mutation frequency.

Chromatin- B chromosomes are described as heterochromatic in about half ofplants that carry them (Zea mays). In general, theirheterochromatic content is similar to that of their A chromosomes.

NOTES

Inheritance

The inheritance of B chromosomes is non-mendelian and irregularowing to vagaries in the levels of paring, to degrees of meioticelimination and to various drive processes. Drive is mostly caused bydirected nondisjunction of sister chromatids at the first pollen mitosis,such that the generative nucleus carries the unreduced number, whichthen forms the sperm (e.g. many species of Gramineae). In rye,unusually, this drive happens on both the male and the femaleside, and, in maize, the nondisjunction happens at the second pollenmitosis, followed by preferential fertilization by the B-chromosomecontainingsperm. Meiotic drive and accumulation at earlier developmentalstages in the germ line operate in a few cases. Theseirregularities in transmission generate a numerical polymorphism inpopulations, with a spectrum of B chromosome numbers includingindividuals with none. There is usually a modal number and anequilibrium frequency based on a balance between drive and theharmful effects caused by high numbers. Drive is by no means auniversal process: it is known in,60% of species for whichtransmission data are available. In the others, there is no known driveand no real understanding of how the population equilibriumfrequencies are maintained.

9.9 Sex Chromosomes

The nuclei of human cells contain 22 autosomes and 2 sex chromosomes. In females, the sex chromosomes are the 2 X chromosomes. Males have one X chromosome and one Y chromosome. The presence of the Y chromosome is decisive for unleashing the developmental program that leads to a baby boy.

The Y Chromosome

In making sperm by meiosis, the X and Y chromosomes must separate in anaphase just as homologous autosomes do. This occurs without a problem because, like homologous autosomes, the X and Y chromosome synapse during prophase of meiosis I. There is a small region of homology shared by the X and Y chromosome and synapsis occurs at that region. Crossing over occurs in two

regions of pairing, called the pseudoautosomal regions. These are located at opposite ends of the chromosome.

The Pseudoautosomal Regions

NOTES

The pseudoautosomal regions get their name because any genes located within them (so far only 9 have been found) are inherited just like any autosomal genes. Males have two copies of these genes: one in the pseudoautosomal region of their Y, the other in the corresponding portion of their X chromosome. So males can inherit an allele originally present on the X chromosome of their father and females can inherit an allele originally present on the Y chromosome of their father..

Genes outside the pseudoautosomal regions

Although 95% of the Y chromosome lies between the pseudoautosomal regions, only 24 different functional genes have been found here. Over half of this region is genetically-barren heterochromatin. Of the 24 genes found in the euchromatin, some encode proteins used by all cells. The others encode proteins that appear to function only in the testes. A key player in this latter group is SRY.SRY

SRY (for sex-determining region Y) is a gene located on the short (p) arm just outside the pseudoautosomal region. It is the master switch that triggers the events that converts the embryo into a male. Without this gene, you get a female instead.

On very rare occasions aneuploid humans are born with such karyotypes as XXY, XXXY, and even XXXXY. Despite their extra X chromosomes, all these cases are male.

The X Chromosome

The X chromosome carries nearly 1,000 genes but few, if any, of these have anything to do directly with sex. However, the inheritance of these genes follows special rules. These arise because:

1.males have only a single X chromosome

2.almost all the genes on the X have no counterpart on the Y; thus any gene on the X, even if recessive in females, will be expressed in males.Genes inherited in this fashion are described as sex-linked or, more precisely, X-linked.

X-Linkage: An Example

NOTES

Hemophilia is a blood clotting disorder caused by a mutant gene encoding either clotting factor VIII, causing hemophilia A or clotting factor IX, causing hemophilia B.

Heterozygous females are called "carriers" because although they show no symptoms, they pass the gene on to approximately half their sons, who develop the disease, and half their daughters, who also become carriers.

X	Y	
X	XX	XY
Xh	XhX	XhY

Women rarely suffer from hemophilia because to do so they would have to inherit a defective gene from their father as well as their mother. Until recently, few hemophiliacs ever became fathers.

X-chromosome Inactivation (XCI)

Human females inherit two copies of every gene on the X chromosome, whereas males inherit only one (with some exceptions: the 9 pseudoautosomal genes and the small number of "housekeeping" genes found on the Y). But for the hundreds of other genes on the X, are males at a disadvantage in the amount of gene product their cells produce? The answer is no, because females have only a single active X chromosome in each cell.

During interphase, chromosomes are too tenuous to be stained and seen by light microscopy. However, a dense, stainable structure, called a Barr body (after its discoverer) is seen in the interphase nuclei of female mammals. The Barr body is one of the X chromosomes. Its compact appearance reflects its inactivity. So, the cells of females have only one functioning copy of each X-linked gene — the same as males.X-chromosome inactivation occurs early in embryonic development. In a given cell, which of a female's X chromosomes becomes inactivated and converted into a Barr body is a matter of chance (except in marsupials like the kangaroo, where it is always the father's X chromosome that is inactivated). After inactivation has occurred, all the descendants of that cell will have the same chromosome inactivated. Thus X-chromosome inactivation creates clones with differing effective gene content. An organism

whose cells varies in effective gene content and hence in the expression of a trait, is called a genetic mosaic.

Mechanism of X-chromosome inactivation

NOTES

Inactivation of an X chromosome requires a gene on that chromosome called XIST.

XIST is transcribed into a long noncoding RNA.

XIST RNA accumulates along the X chromosome containing the active XIST gene and proceeds to inactivate all (or almost all) of the hundreds of other genes on that chromosome.

Barr bodies are inactive X chromosomes "painted" with XIST RNA.

During the first cell divisions of the female mouse zygote, the XIST locus on the father's X chromosome is expressed so most of his X-linked genes are silent.By the time the blastocyst has formed, the silencing of the paternal X chromosome still continues in the trophoblast (which will go on to form the placenta) but in the inner cell mass (the ICM, which will go on to form the embryo) transcription of XIST ceases on the paternal X chromosome allowing its hundreds of other genes to be expressed. The shut-down of the XIST locus is done by methylating XIST regulatory sequences. So the pluripotent stem cells of the ICM express both X chromosomes.

However, as embryonic development proceeds, X-chromosome inactivation begins again. But this time it is entirely random. There is no predicting whether it will be the maternal X or the paternal X that is inactivated in a given cell.Some genes on the X chromosome escape inactivation.

X-chromosome Abnormalities

As we saw above, people are sometimes found with abnormal numbers of X chromosomes. Unlike most cases of aneuploidy, which are lethal, the phenotypic effects of aneuploidy of the X chromosome are usually not severe. Examples:

Females with but a single intact X chromosome (usually the one she got from her mother) in some (thus a genetic mosaic) or all of her cells show a variable constellation of phenotypic traits called Turner syndrome. For those girls that survive to birth, the phenotypic effects are generally mild because each cell has

a single functioning X chromosome like those of XX females. Number of Barr bodies = zero.

XXX, XXXX, XXXXX karyotypes: all females with mild phenotypic effects because in each cell all the extra X chromosomes are inactivated. Number of Barr bodies = number of X chromosomes minus one.

NOTES

Klinefelter's syndrome: people with XXY or XXXY karyotypes are males (because of their Y chromosome). But again, the phenotypic effects of the extra X chromosomes are mild because, just as in females, the extra Xs are inactivated and converted into Barr bodies.

Sex Determination in Other Animals

Although the male fruit fly, Drosophila melanogaster, is X-Y, the Y chromosome does not dictate its maleness but rather the absence of a second X. Furthermore, instead of females shutting down one X to balance the single X of the males as we do male flies double the output of their single X relative to that of females.

In birds, moths, schistosomes, and some lizards, the male has two of the same chromosome (designated ZZ), whereas the female has "heterogametic" chromosomes (designated Z and W). In chickens, a single gene on the Z chromosome (designated DMRT1), when present in a double dose (ZZ), produces males while the presence of only one copy of the gene produces females (ZW).

Environmental Sex Determination

In some cold-blooded vertebrates,some fishes,many reptiles (e.g. certain snakes, lizards, turtles, and all crocodiles and alligators)as well as in some invertebrates (e.g. certain crustaceans), sex is determined after fertilization — not by sex chromosomes deposited in the egg.

The choice is usually determined by the temperature at which early embryonic development takes place.In some cases (e.g. many turtles and lizards), a higher temperature during incubation favors the production of females.In other cases (e.g., alligators), a higher temperature favors the production of males.Even in cases (e.g. some lizards) where there are sex chromosomes, a high temperature can convert a genotypic male (ZZ) into a female.

9.10 Molecular basis for Chromosome Pairing

NOTES

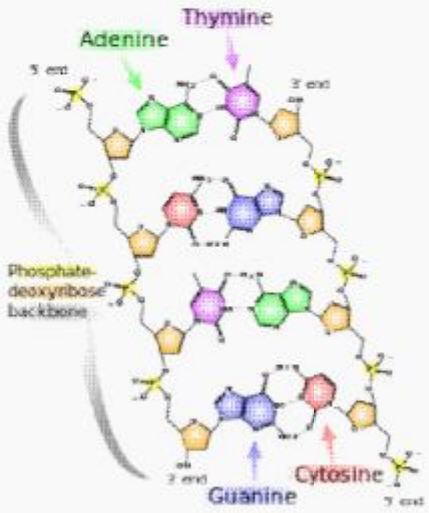


Fig 9.4 : The Molecular Structure of DNA

Bases pair through the arrangement of hydrogen bonding between the strands.

The molecular basis for genes is deoxyribonucleic acid (DNA). DNA is composed of a chain of nucleotides, of which there are four ypes: adenine (A), cytosine (C), guanine (G), and thymine . Genetic information exists in the sequence of these nucleotides, and genes exist as stretches of sequence along the DNA chain. Viruses are the only exception to this rule—sometimes viruses use the very similar molecule RNA instead of DNA as their genetic material. Viruses cannot reproduce without a host and are unaffected by many genetic processes, so tend not to be considered living organisms.

DNA normally exists as a double-stranded molecule, coiled into the shape of a double helix. Each nucleotide in DNA preferentially pairs with its partner nucleotide on the opposite strand: A pairs with T, and C pairs with G. Thus, in its two-stranded form, each strand effectively contains all necessary information, redundant with its partner strand. This structure of DNA is the physical basis for inheritance: DNA replication duplicates the genetic information by splitting the strands and using each strand as a template for synthesis of a new partner strand.

NOTES

Genes are arranged linearly along long chains of DNA base-pair sequences. In bacteria, each cell usually contains a single circular genophore, while eukaryotic organisms (such as plants and animals) have their DNA arranged in multiple linear chromosomes. These DNA strands are often extremely long; the largest human chromosome, for example, is about 247 million base pairs in length. The DNA of a chromosome is associated with structural proteins that organize, compact and control access to the DNA, forming a material called chromatin; in eukaryotes, chromatin is usually composed of nucleosomes, segments of DNA wound around cores of histone proteins.The full set of hereditary material in an organism (usually the combined DNA sequences of all chromosomes) is called the genome.

While haploid organisms have only one copy of each chromosome, most animals and many plants are diploid, containing two of each chromosome and thus two copies of every gene.The two alleles for a gene are located on identical loci of the two homologous chromosomes, each allele inherited from a different parent.

9.11 Summary

Chromosomes are tiny thread-like structures found in the nucleus of a cell. Chromosomes store and transmit the coded information which is responsible for all the life processes of an organism. Hence, chromosomes are commonly described as carriers of heredity.In bacteria, the entire hereditary material is packed into a single, irregularly packed, compact mass called nucleioid or bacterial chromosome.Bacterial chromosomes lack a nuclear membrane and are not associated with histones.Eukaryotic chromosomes are visible only during metaphase stage of mitosis.The chromosome number of a given species is generally described as the diploid number (2n) since chromosomes occur in pairs.Haploid number of chromosomes occur in gametes and spores. Some adult organisms (chlamydomonos, male honey bee) have haploid number of chromosomes.Chromosomes range in size from 0.1 to 30 mm in length and 0.2 to 2.0 mm in thickness. When the number is less, the chromosomes are larger in size.A metaphase chromosome shows two identical components called chromatids, joined by a centromere (primary constriction). Each chromatid has two arms. A secondary constriction may occur sometimes. Such a chromosome

is called sat-chromosome.Chemically the chromosome is composed of nucleoproteins present in the form of a highly coiled chromonema. Chromonema is composed of a chain of bead-like structures called nucleosomes (nu bodies).A nucleosome has a core particle formed by proteins called histones surrounded by a DNA strand.The part of DNA not associated with histones is called Linker DNA.Normally a chromosome has only one centromere (monocentric). It can be sometimes dicentric or polycentric. Very rarely it may lack a centromere (acentric).Monocentric chromosome can be distinguished into metacentric, submetacentric, acrocentric and telocentric, based on the position of centromere.Chromosomes, which are extremely larger than the normal ones are called giant chromosomes. They occur in some animal cells.

NOTES

Lampbrush chromosomes of amphibian oocytes and polytene chromosomes of Drosophile one common examples of giant chromosomes.Genome is the term used to describe the sum total of all the genes present in a haploid set of chromosomes.In unisexual organisms the chromosomes can be distinguished into autosomes (somatic chromosomes) and allosomes (sex chromosomes).Allosomes have a role in sex-determination and in the expression of sex-linked characters.

9.12 Glossary

- **Adenine**: A purine base; one of the four molecules containing nitrogen present in the nucleic acids DNA and RNA; designated by letter A.
- **Allele**: An alternative form of a gene.
- **Allele Frequency**: The proportion of a particular allele among the chromosomes carried by individuals in a population.
- **Amino acid**: Any of a class of 20 molecules that are combined to form proteins in living things. The sequence of amino acids in a protein and hence protein function are determined by the genetic code.
- **Amplification**: An increase in the number of copies of a specific DNA fragment; can be in vivo or in vitro.
- **Autosome**: A chromosome not involved in sex determination. The diploid human genome consists of 46 chromosomes, 22 pairs of

autosomes, and one pair of sex chromosomes (the X and Y chromosomes).

NOTES

- **Basepair**: Two complementary nucleotides joined by hydrogen bonds; basepairing occurs between A and T and between G and C.
- **Base sequence**: The order of nucleotide bases in a DNA molecule.
- **Base sequence analysis**: A method, sometimes automated, for determining the base sequence.
- **Chromosome** :The structure by which hereditary information is physically transmitted from one generation to the next.
- **Complementary sequences**: Nucleic acid base sequences that form a double-stranded structure by matching base pairs; the complementary sequence to G-T-A-C is C-A-T-G.
- **Cytosine**: A pyrimidine base; one of the four molecules containing nitrogen present in the nucleic acids DNA and RNA; designated by letter C.
- **Denaturation**: the process of splitting the complementary double strands of DNA to form single strands
- **Deoxyribonucleic acid (DNA):** The genetic material of organisms, usually double-stranded; a class of nucleic acids identified by the presence of deoxyribose, a sugar, and the four nucleobases.
- **DNA sequence**: The relative order of base pairs, whether in a fragment of DNA, a gene, a chromosome, or an entire genome.
- **Double Helix**: The shape that two linear strands of DNA assume when bonded together
- **Gene**: the basic unit of heredity; a sequence of DNA nucleotides on a chromosome
- **Gene frequency**: the relative occurrence of a particular allele in a population
- **Gene mapping**: Determination of the relative positions of genes on a DNA molecule (chromosome or plasmid) and of the distance, in linkage units or physical units, between them.

- **Genetics**: The study of the patterns of inheritance of specific traits.
- **Genome**: All the genetic material in the chromosomes of a particular organism; its size is generally given as the total number of base pairs.
- **Genome projects**: Research and technology development efforts aimed at mapping and sequencing some or all of the genome of an organism.
- **Genotype**: the genetic makeup of an organism, as characterized by its physical appearance or phenotype
- **Guanine**: a purine base; one of the four molecules containing nitrogen present in the nucleic acids DNA and RNA; designated by letter G
- **Heredity**: the transmission of characteristics from one generation to the next
- **Heterozygosity**: The presence of different alleles at one or more loci on homologous chromosomes.
- **Homologies**: Similarities in DNA or protein sequences between individuals of the same linear sequences each derived from one parent.
- **Homologous chromosomes**: A pair of chromosomes containing the same linear gene sequences each derived from one parent.
- **Kilobase (kb):** Unit of length for DNA fragments equal to 1000 nucleotides.
- **Linkage**: The proximity of two or more markers (genes, etc.) on a chromosome; the closer together the markers are, the lower the probability that they will be separated during DNA repair or replication process, and hence the greater the probability that they will be inherited together.
- **Localize**: Determination of the original position (locus) of a gene or other marker on a chromosome.
- **Locus** (pl. loci): The specific physical location of a gene on a chromosome.
- **Mitosis**: The process of nuclear division in cells that produces daughter cells that are genetically identical to each other and to the parent.
- **Mutation**: Any inheritable change in DNA sequence.

NOTES

NOTES

- **Nucleic acid**: A nucleotide polymer that DNA and RNA are major types.
- **Nucleotide**: A unit of nucleic acid composed of phosphate, ribose or deoxyribose, and a purine or pyrimidine base.
- **Nucleus**: The cellular organelle in eukaryotes that contains the genetic material.
- **Sex chromosomes (x and y chromosomes):** Chromosomes that are different in the two sexes and involved in sex determination.
- **Thymine: a pyrimidine base**; one of the four molecules containing nitrogen present in the nucleic acids DNA and RNA; designated by letter

9.13 Self-Learning Exercise

Section-A (Very Short Answer Type Questions)

1. What is DNA?
2. What is a chromosome?
3. How many chromosomes do people have?
4. Can changes in chromosomes affect health and development?
5. What is mitochondrial DNA?

Section-B (Short Answer Type Questions)

1 What is nucleosome?
2. What is Solenoid?
3. What is heterochromatin?

Section-C (Long Answer Type Auestions)

1. Discuss the molecular basis of chromosome pairing.
2. Describe the characteristic features of lamp-brush chromosomes.
3. What is the difference between B-chromosomes and polytene chromosomes?

9.14 References

- Wilson, E.B. (1925). The Cell in Development and Heredity, Ed. 3. Macmillan, New York. p. 923.

- Mayr, E. (1982). The growth of biological thought. Harvard. p. 749.
- White, M. J. D. (1973). The chromosomes (6th ed.). London: Chapman and Hall, distributed by Halsted Press, New York. p. 28. ISBN 0-412-11930-7

NOTES

Unit - 10

Structural and Numerical Alterations in Chromosomes

Structure of the unit

10.0 Objectives

After studying this unit you will be able to understand:

- Structural and numerical alterations in chromosomes: Origin, meiosis and breeding behavior of duplication, deficiency, inversion and translocation heterozygote.

- Origin, occurrence, production and meiosis of haploids, aneuploids and euploids, origin and production of autopolyploid, chromosome and chromatid segregation, allopolyploids,
- Genome constitution and analysis, evolution of major crop plants.

NOTES

10.1 Introduction

The chromosome set of a species remains relatively stable over long periods of time. However, within populations there can be found abnormalities involving the structure or number of chromosomes. These alterations arise spontaneously from errors in the normal processes of the cell. Their consequences are usually deleterious, giving rise to individuals who are unhealthy or sterile, though in rare cases alterations provide new adaptive opportunities that allow evolutionary change to occur. In fact, the discovery of visible chromosomal differences between species has given rise to the belief that radical restructuring of chromosome architecture has been an important force in evolution.

Two important principles dictate the properties of a large proportion of structural chromosomal changes. The first principle is that any deviation from the normal ratio of genetic material in the genome results in genetic imbalance and abnormal function. In the normal nuclei of both diploid and haploid cells, the ratio of the individual chromosomes to one another is 1:1. Any deviation from this ratio by addition or subtraction of either whole chromosomes or parts of chromosomes results in genomic imbalance. The second principle is that homologous chromosomes go to great lengths to pair at meiosis. The tightly paired homologous regions are joined by a ladder like longitudinal structure called the synaptonemal complex. Homologous regions seem to be able to find each other and form a synaptonemal complex whether or not they are part of normal chromosomes. Therefore, when structural changes occur, not only are the resulting pairing formations highly characteristic of that type of structural change but they also dictate the packaging of normal and abnormal chromosomes into the gametes and subsequently into the progeny.

10.2 Origin, Meiosis and Breeding behavior of Duplication, Deficiency, Inversion and Translocation Heterozygotes

NOTES

Chromosomes are the structures that hold our genes. Genes are the individual instructions that tell our bodies how to develop and function; they govern our physical and medical characteristics, such as hair color, blood type and susceptability to disease. Each chromosome has a p and q arm; p is the shorter arm and q is the longer arm. The arms are separated by a pinched region known as the centromere. Chromosomes are tiny string-like structures in cells of the body that contain the genes. Humans have about 20,000 to 25,000 genes that determine traits like eye and hair color. They also direct the growth and development of every part of the body.

The typical number of chromosomes in a human cell is 46 - two pairs of 23 - holding an estimated 25,000 genes. One set of 23 chromosomes is inherited from the biological mother (from the egg), and the other set is inherited from the biological father (from the sperm).In order for chromosomes to be seen with a microscope, they need to be stained. Once stained, the chromosomes look like strings with light and dark "bands" and their picture can be taken. A picture, or chromosome map, of all 46 chromosomes is called a karyotype. The karyotype can help identify chromosome abnormalities that are evident in either the structure or the number of chromosomes.To help identify chromosomes, the pairs have been numbered from 1 to 22, with the 23rd pair labeled "X" and "Y." In addition, each chromosome arm is defined further by numbering the bands that appear after staining; the higher the number, the further that area is from the centromere.The first 22 pairs of chromosomes are called "autosomes" and the final pair is called the "sex chromosomes." The sex chromosome an individual has determines that person's gender; females have two X chromosomes (XX), and males have an X and a Y chromosome (XY).

A chromosome abnormality reflects an abnormality of chromosome number or structure. There are many types of chromosome abnormalities. However, they can be organized into two basic groups: About 1 in 150 babies is born with a chromosomal abnormality. These are caused by errors in the number or structure of chromosomes. There are many different chromosomal abnormalities. Many children with a chromosomal abnormality have mental

and/or physical birth defects. Understanding what chromosomes are may make it easier to understand the wide range of problems chromosomal abnormalities can cause.

Numerical Abnormalities

NOTES

When an individual is missing either a chromosome from a pair (monosomy) or has more than two chromosomes of a pair (trisomy). An example of a condition caused by numerical abnormalities is Down syndrome, also known as Trisomy 21 (an individual with Down syndrome has three copies of chromosome 21, rather than two). Turner Syndrome is an example of monosomy, where the individual - in this case a female - is born with only one sex chromosome, an X.

Structural Abnormalities

In this the chromosome's structure is altered. This can take several forms:

Deletions: A portion of the chromosome is missing or deleted.

Duplications: A portion of the chromosome is duplicated, resulting in extra genetic material.

Translocations: When a portion of one chromosome is transferred to another chromosome. There are two main types of translocations. In a reciprocal translocation, segments from two different chromosomes have been exchanged. In a Robertsonian translocation, an entire chromosome has attached to another at the centromere.

Inversions: A portion of the chromosome has broken off, turned upside down and reattached, therefore the genetic material is inverted.

Rings: A portion of a chromosome has broken off and formed a circle or ring. This can happen with or without loss of genetic material.

Most chromosome abnormalities occur as an accident in the egg or sperm. Therefore, the abnormality is present in every cell of the body. Some abnormalities, however, can happen after conception, resulting in mosaicism, where some cells have the abnormality and some do not. Chromosome abnormalities can be inherited from a parent (such as a translocation) or be "de novo" (new to the individual). This is why chromosome studies are often performed on parents when a child is found to have an abnormality.

NOTES

Chromosome abnormalities usually occur when there is an error in cell division. There are two kinds of cell division. Mitosis results in two cells that are duplicates of the original cell. In other words, one cell with 46 chromosomes becomes two cells with 46 chromosomes each. This kind of cell division occurs throughout the body, except in the reproductive organs. This is how most of the cells that make up our body are made and replaced. Meiosis results in cells with half the number of chromosomes, 23 instead of the normal 46. These are the eggs and sperm.

In both processes, the correct number of chromosomes is supposed to end up in the resulting cells. However, errors in cell division can result in cells with too few or too many copies of a chromosome. Errors can also occur when the chromosomes are being duplicated.

Other factors that can increase the risk of chromosome abnormalities are:

Maternal Age: Women are born with all the eggs they will ever have. Therefore, when a woman is 30 years old, so are her eggs. Some researchers believe that errors can crop up in the eggs' genetic material as they age over time. Therefore, older women are more at risk of giving birth to babies with chromosome abnormalities than younger women. Since men produce new sperm throughout their life, paternal age does not increase risk of chromosome abnormalities.

Environment: Although there is no conclusive evidence that specific environmental factors cause chromosome abnormalities, it is still a possibility that the environment may play a role in the occurrence of genetic errors.

Sometimes, however, a baby can be born with too many or too few chromosomes, or with one or more chromosomes that are missing a piece or are rearranged. These errors in the number or structure of chromosomes can cause a wide variety of birth defects ranging from mild to severe. Some chromosomal abnormalities result in miscarriage or stillbirth.Chromosomal abnormalities usually result from an error that occurred when an egg or sperm cell was developing. It is not known why these errors occur. As far as we know, nothing that a parent does or doesn't do before or during pregnancy can cause a chromosomal abnormality in his or her child.

Sperm and egg cells are different from other cells in the body. These cells have only 23 unpaired chromosomes. When an egg and sperm cell join together they form a fertilized egg with 46 chromosomes.But sometimes something goes wrong before fertilization. An egg or sperm cell may divide incorrectly, resulting in an egg or sperm cell with too many or too few chromosomes.When this cell with the wrong number of chromosomes joins with a normal egg or sperm cell, the resulting embryo has a chromosomal abnormality. A common type of chromosomal abnormality is called a trisomy. This means that an individual has three copies, instead of two, of a specific chromosome.

NOTES

Down syndrome is an example of a trisomy. Individuals with Down syndrome generally have three copies of chromosome 21. Children with Down syndrome have varying degrees of mental retardation, characteristic facial features and, often, heart defects and other problems. The risk of Down syndrome and other trisomies increases with maternal age.

In most cases, an embryo with the wrong number of chromosomes does not survive. In such cases, the pregnant woman has a miscarriage. This often happens very early in pregnancy, before a woman may realize she's pregnant. Up to 75 percent of first trimester miscarriages are caused by chromosomal abnormalities in the embryo.

Other errors also can occur, usually before fertilization. These errors can alter the structure of one or more chromosomes. Individuals with structural chromosomal abnormalities usually have the normal number of chromosomes. However, small pieces of a chromosome (or chromosomes) may be deleted, duplicated, inverted, misplaced or exchanged with part of another chromosome. These structural rearrangements sometimes have no effect, if the entire chromosome is there but just rearranged, or it may result in pregnancy loss or birth defects.

Errors in cell division also can occur soon after fertilization. This can result in mosaicism, a condition in which an individual has cells with different genetic makeup. For example, individuals with the mosaic form of Turner syndrome are missing an X chromosome in some, but not all, of their cells. Some individuals with chromosomal mosaicism may be mildly affected, but the severity of the condition depends largely on the percentage of abnormal cells.

NOTES

Chromosomal abnormalities can be diagnosed after birth using a blood test, or before birth using prenatal tests (amniocentesis or chorionic villus sampling). Cells obtained from these tests are grown in the laboratory, and then their chromosomes are examined under a microscope. The lab makes a picture (karyotype) of all the person's chromosomes, arranged in order from largest to smallest. The karyotype shows the number, size and shape of the chromosomes and helps experts identify any abnormalities.

Babies also can be born with an extra copy of chromosome 13 or 18. These trisomies are usually more severe than Down syndrome, but fortunately less common. About 1 in 10,000 babies is born with trisomy 13 (also called Patau syndrome), and about 1 in 6,000 with trisomy 18 (also called Edwards syndrome) (3, 6). Babies with trisomies 13 or 18 generally have severe mental retardation and many physical birth defects. Most affected babies die before their first birthday.

The X and Y chromosomes are referred to as sex chromosomes. About 1 in 500 babies has missing or extra sex chromosomes .Generally, females have two X chromosomes, and males have one X chromosome and one Y chromosome. Sex chromosome abnormalities may cause infertility, growth abnormalities, and in some cases, behavioral and learning problems. However, most affected individuals live fairly normal lives.

Turner syndrome is a sex chromosome abnormality that affects about 1 in 2,500 girls .Girls with Turner syndrome are missing all or part of one X chromosome. They usually are infertile and do not undergo normal pubertal changes unless they are treated with sex hormones. Affected girls are short, though treatment with growth hormones can help increase height. Some have other health problems, including heart and kidney defects. Girls with Turner syndrome generally have normal intelligence, though some have difficulties with mathematics and spatial concepts.

About 1 in 1,000 females has an extra X chromosome, referred to as triple X (3). Affected girls tend to be tall. They usually have no physical birth defects, undergo normal puberty and are fertile. Affected girls usually have normal intelligence, though many have learning disabilities. Because these girls are healthy and have a normal appearance, their parents often don't know they have

a chromosomal abnormality. Some parents may learn that their daughter has this abnormality if they have prenatal testing (with amniocentesis or chorionic villus sampling).

NOTES

Klinefelter syndrome is a sex-chromosome abnormality that affects about 1 in 500 to 1,000 boys (Boys with Klinefelter syndrome have two, or occasionally more, X chromosomes along with their Y chromosome (males normally have one X and one Y chromosome). Affected boys usually have normal intelligence, though many have learning disabilities. As adults, they produce lower-than-normal amounts of the male hormone testosterone (and often are treated with this hormone) and are infertile.

About 1 in 1,000 males is born with one or more extra Y chromosomes .(Affected males are sometimes taller than average, have normal sexual development and are fertile. Most have normal intelligence, though some have learning disabilities, and behavioral and speech/language problems. As with triple X females, many affected males and their families don't know they have a chromosomal abnormality unless it is diagnosed with prenatal testing.
New techniques for analyzing chromosomes have made it possible to identify tiny chromosomal abnormalities that may not be visible even under a high-powered microscope. As a result, more parents are learning that their child has a chromosomal abnormality.

Sex Chromosome Abnormalities			
Female Genotype	**Syndrome**	**Male Genotype**	**Syndrome**
XX	normal	XY	Normal
XO	Turner	XXY	Klinefelter
XXX	Triple-X	XYY	XYY

Tabular form of the sex chromosome abnormalities

Some rare disorders can be caused by small chromosomal deletions. Some examples are:

NOTES

- **Prader-Willi syndrome** (deletion on chromosome 15): Affected children usually have mental retardation or learning disabilities, behavioral problems and short stature. They also may develop extreme obesity.
- **Cri-du-chat (cat cry) syndrome** (deletion on chromosome 5): Affected children have a cat-like high-pitched cry during infancy, mental retardation and physical abnormalities.
- **Wolf-Hirschhorn syndrome** (deletion on chromosome 4): This disorder is characterized by severe mental retardation, heart defects, poor muscle tone, seizures, high blood pressure and other problems.
- **22q11 deletion syndrome** (deletion on chromosome 22): Deletions in a specific region of chromosome 22 cause a variety of problems that can include heart defects, cleft lip/palate, immune system abnormalities, characteristic facial features and learning disabilities. Certain combinations of these features are sometimes called DiGeorge or velocardiofacial syndrome. Individuals with this disorder have a 50 percent chance of passing the chromosomal abnormality on to their offspring with each pregnancy.

10.3 Origin, Occurrence, Production and Meiosis of Haploids. Chromosome and Chromatid Segregation

New techniques of analyzing chromosomes sometimes can pinpoint exactly where missing or extra genetic material comes from. If doctors know what genes are contained in that section and their function, they sometimes can give parents a better prediction of a child's future development.

Chromosomal abnormalities, alterations and aberrations are at the root of many inherited diseases and traits. Chromosomal abnormalities often give rise to birth defects and congenital conditions that may develop during an individual's lifetime. Examining the karyotype of chromosomes (karyotyping) in a sample of cells can allow detection of a chromosomal abnormality and counseling can then be offered to parents or families whose offspring are at risk of growing up with a genetic disorder.

Although it is possible to inherit some types of chromosomal abnormalities, most chromosomal disorders (such as Down syndrome and Turner syndrome) are not passed from one generation to the next.

Some chromosomal conditions are caused by changes in the number of chromosomes. These changes are not inherited, but occur as random events during the formation of reproductive cells (eggs and sperm). An error in cell division called non disjunction results in reproductive cells with an abnormal number of chromosomes. For example, a reproductive cell may accidentally gain or lose one copy of a chromosome. If one of these atypical reproductive cells contributes to the genetic makeup of a child, the child will have an extra or missing chromosome in each of the body's cells.

Changes in chromosome structure can also cause chromosomal disorders. Some changes in chromosome structure can be inherited, while others occur as random accidents during the formation of reproductive cells or in early fetal development. Because the inheritance of these changes can be complex, people concerned about this type of chromosomal abnormality may want to talk with a genetics professional.Some cancer cells also have changes in the number or structure of their chromosomes. Because these changes occurs in somatic cells (cells other than eggs and sperm), they cannot be passed from one generation to the next.

Events involved in meiosis: chromosomal crossover

Meiosis is a special type of cell division necessary for sexual reproduction in eukaryotes, such as animals, plants and fungi. The number of sets of chromosomes in the cell undergoing meiosis is reduced to half the original number, typically from two sets (diploid) to one set (haploid). The cells produced by meiosis are either gametes (the usual case in animals) or otherwise usually spores from which gametes are ultimately produced (the case in land plants). In many organisms, including all animals and land plants (but not some other groups such as fungi), gametes are called sperm in males and egg cells or ova in females. Since meiosis has halved the number of sets of chromosomes, when two gametes fuse during fertilisation, the number of sets of chromosomes in the resulting zygote is restored to the original number. Meiotic division occurs in two stages, meiosis I and meiosis II, dividing the cells once at each

NOTES

stage. Before meiosis begins, during S phase of the cell cycle, the DNA of each chromosome is replicated, so that each chromosome has two sister chromatids; a diploid organism now has a tetraploid DNA amount in the cell. The first stage of meiosis begins with a cell that has (if it is from a diploid organism) two copies of each type of chromosome, one from each of the mother and father, called homologous chromosomes, each of which has two sister chromatids. The homologous chromosomes pair up and may exchange genetic material with each other in a process called crossing over. Each pair then separates as two cells are formed, each with one chromosome (two chromatids) from every homologous pair.

In the second stage, each chromosome splits into two; each half, each sister chromatid, is separated into two new cells, which are haploid. This occurs in both of the cells formed in meiosis I. Therefore from each original cell, four genetically distinct haploid cells are produced. These cells can mature into gametes.

While the process of meiosis bears a number of similarities with the 'life-cycle' cell division process of mitosis, it differs in two important respects:

1. Recombination meiosis shuffles the genes between the two chromosomes in each pair (one received from each parent), producing chromosomes with new genetic combinations in every gamete generated. 2. Mitosis does not shuffle the genes, producing chromosomes pairs identical to those in the parent cell.3. Meiosis produces four genetically unique cells, each with half the number of chromosomes as in the parent.4.Mitosis produces the two genetically identical cells, each with the same number of chromosomes as in the parent

It is also noteworthy that during meiosis, specific genes are more highly transcribed, and these are called the meiome, the term used in functional genomics for the meiotic transcriptome. Meiosis is a key feature for all sexually reproducing eukaryotes in which homologous chromosome pairing, synapse and recombination occur. In addition to strong meiotic stage-specific expression of mRNA (the meiome), however, there are also pervasive translational controls (e.g. selective usage of preformed mRNA), regulating the ultimate meiotic stage-specific protein expression of genes during meiosis. Thus, both the meiome and translational controls determine the broad

restructuring of meiotic cells needed to carry out meiosis.Prior to the meiosis process the cell's chromosomes are duplicated by a round of DNA replication, creating from the maternal and paternal versions of each chromosome (homologs) two exact copies, sister chromatids, attached at the centromere region. In the beginning of meiosis, the maternal and paternal homologs pair with each other. Then they typically exchange parts by homologous recombination leading to crossovers of DNA between the maternal and paternal versions of the chromosome. Spindle fibers bind to the centromeres of each pair of homologs and arrange the pairs at the spindle equator. Then the fibers pull the recombined homologs to opposite poles of the cell. As the chromosomes move away from the center the cell divides into two daughter cells, each containing a haploid number of chromosomes composed of two chromatids.

NOTES

After the recombined maternal and paternal homologs have separated into the two daughter cells, a second round of cell division occurs. There meiosis ends as the two sister chromatids making up each homolog are separated and move into one of the four resulting gamete cells. Upon fertilization, for example when a sperm enters an egg cell, two gamete cells produced by meiosis fuse. The gamete from the mother and the gamete from the father each contribute one half of the set of chromosomes that make up the new offspring's genome. Meiosis uses many of the same mechanisms as mitosis, the type of cell division used by eukaryotes like plants and animals to split one cell into two identical daughter cells. In all plants and in many protists meiosis results in the formation of spores: haploid cells that can divide vegetatively without undergoing fertilization. Some eukaryotes, like bdelloid rotifers, do not have the ability to carry out meiosis and have acquired the ability to reproduce by parthenogenesis.

Meiosis does not occur in archaea or bacteria, which generally reproduce via asexual processes such as binary fission. However, a similar "sexual" process, known as bacterial transformation, involves transfer of DNA from one bacterium to another and recombination of these DNA molecules of different parental origin.

NOTES

Cycling meiosis and fertilization events produces a series of transitions back and forth between alternating haploid and diploid states. The organism phase of the life cycle can occur either during the diploid state (gametic or diploid life cycle), during the haploid state (zygotic or haploid life cycle), or both (sporic or haplodiploid life cycle, in which there are two distinct organism phases, one during the haploid state and the other during the diploid state). In this sense there are three types of life cycles that utilize sexual reproduction, differentiated by the location of the organisms phase(s).

In the gametic life cycle or " diplontic life cycle", of which humans are a part, the organism is diploid, grown from a diploid cell called the zygote. The organism's diploid germ-line stem cells undergo meiosis to create haploid gametes (the spermatozoa for males and ova for females), which fertilize to form the zygote. The diploid zygote undergoes repeated cellular division by mitosis to grow into the organism. Mitosis is a related process to meiosis that creates two cells that are genetically identical to the parent cell.

In the zygotic life cycle the organism is haploid instead, spawned by the proliferation and differentiation of a single haploid cell called the gamete. Two organisms of opposing gender contribute their haploid gametes to form a diploid zygote. The zygote undergoes meiosis immediately, creating four haploid cells. These cells undergo mitosis to create the organism. Many fungi and many protozoa are members of the zygotic life cycle.

Finally, in the sporic life cycle, the living organism alternates between haploid and diploid states. Consequently, this cycle is also known as the alternation of generations. The diploid organism's germ-line cells undergo meiosis to produce spores. The spores proliferate by mitosis, growing into a haploid organism. The haploid organism's gamete then combines with another haploid organism's gamete, creating the zygote. The zygote undergoes repeated mitosis and differentiation to become a diploid organism again. The sporic life cycle can be considered a fusion of the gametic and zygotic life cycles.

Because meiosis is a "one-way" process, it cannot be said to engage in a cell cycle as mitosis does. However, the preparatory steps that lead up to meiosis are identical in pattern and name to the interphase of the mitotic cell cycle.

Interphase is divided into three phases:

Growth 1 (G1) phase: This is a very active period, where the cell synthesizes its vast array of proteins, including the enzymes and structural proteins it will need for growth. In G1 stage each of the chromosomes consists of a single (very long) molecule of DNA. In humans, at this point cells are 46 chromosomes, 2N, identical to somatic cells.

NOTES

Synthesis (S) phase: The genetic material is replicated: each of its chromosomes duplicates, so that each of the 46 chromosomes becomes a complex of two identical sister chromatids. The cell is still considered diploid because it still contains the same number of centromeres. The identical sister chromatids have not yet condensed into the densely packaged chromosomes visible with the light microscope. This will take place during prophase I in meiosis.

Growth 2 (G2) phase: G2 phase as seen before mitosis is not present in Meiosis. Actually, the first four stages of prophase I in many respects correspond to the G2 phase of mitotic cell cycle.

Interphase is followed by meiosis I and then meiosis II. Meiosis I consists of separating the pairs of homologous chromosome; each made up of two sister chromatids, into two cells. Entire haploid content of chromosomes is contained in each of the resulting daughter cells; the first meiotic division therefore reduces the ploidy of the original cell by a factor of 2.

Meiosis II consists of decoupling each chromosome's sister strands (chromatids), and segregating the individual chromatids into haploid daughter cells. The two cells resulting from meiosis I divide during meiosis II, creating 4 haploid daughter cells. Meiosis I and II are each divided into prophase, metaphase, anaphase, and telophase stages, similar in purpose to their analogous subphases in the mitotic cell cycle. Therefore, meiosis includes the stages of meiosis I (prophase I, metaphase I, anaphase I, telophase I), and meiosis II (prophase II, metaphase II, anaphase II, telophase II).

Meiosis generates genetic diversity in two ways: (1) independent alignment and subsequent separation of homologous chromosome pairs during the first meiotic division allows a random and independent selection of each chromosome segregates into each gamete; and (2) physical exchange of

homologous chromosomal regions by homologous recombination during prophase I results in new combinations of DNA within chromosomes.

Origin and function

NOTES

Meiosis is ubiquitous among eukaryotes. It occurs in single-celled organisms such as yeast, as well as in multicellular organisms, such as humans. Eukaryotes arose from prokaryotes more than 1.5 billion years ago, and the earliest eukaryotes were likely single-celled organisms. To understand meiosis in eukaryotes, it is necessary to understand (1) how meiosis arose in single celled eukaryotes, and (2) the function of mitosis.

Theory that meiosis evolved from bacterial sex (transformation)

In prokaryotic sex, DNA from one bacterium is released into the surrounding medium, is then taken up by another bacterium and its information integrated into the DNA of the recipient bacterium. This process is called transformation. One theory on how meiosis arose is that it evolved from transformation. By this view, the evolutionary transition from prokaryotic sex to eukaryotic sex was continuous.

Transformation, like meiosis, is a complex process requiring the function of numerous gene products. The ability to undergo natural transformation among bacterial species is widespread. At least 67 prokaryote species (in seven different phyla) are known to be competent for transformation. A key similarity between bacterial sex and eukaryotic sex is that DNA originating from two different individuals (parents) join up so that homologous sequences are aligned with each other, and this is followed by exchange of genetic information (a process called genetic recombination). After the new recombinant chromosome is formed it is passed on to progeny.

When genetic recombination occurs between DNA molecules originating from different parents, the recombination process is catalyzed in prokaryotes and eukaryotes by enzymes that have similar functions and that are evolutionarily related. One of the most important enzymes catalyzing this process in bacteria is referred to as RecA, and this enzyme has two functionally similar counterparts that act in eukaryotic meiosis, Rad51 and Dmc1.

Support for the theory that meiosis arose from bacterial transformation comes from the increasing evidence that early diverging lineages of eukaryotes have the core genes for meiosis. This implies that the precursor to meiosis was already present in the bacterial ancestor of eukaryotes. For instance the common intestinal parasite Giardia intestinalis, a simple eukaryotic protozoan was, until recently, thought to be descended from an early diverging eukaryotic lineage that lacked sex. However, it has since been shown that G. intestinalis contains within its genome a core set of genes that function in meiosis, including five genes that function only in meiosis. In addition, G. intestinalis was recently found to undergo a specialized, sex-like process involving meiosis gene homologs. This evidence, and other similar examples, suggests that a primitive form of meiosis was present in the common ancestor of all eukaryotes, an ancestor that arose from antecedent bacteria.

NOTES

Theory that meiosis evolved from mitosis

Mitosis is the process in eukaryotes for duplicating chromosomes and segregating each of the two copies into each of the two daughter cells upon somatic cell division (that is, during all cell divisions in eukaryotes, except those involving meiosis that give rise to haploid gametes). In mitosis, chromosome number is ordinarily not reduced. The alternate theory on the origin of meiosis is that meiosis evolved from mitosis. On this theory, early eukaryotes evolved mitosis first, but lacked meiosis and thus had not yet evolved the eukaryotic sexual cycle. Only after mitosis became established did meiosis and the eukaryotic sexual cycle evolve. The fundamental features of meiosis, on this theory, were derived from mitosis.

Support for the idea that meiosis arose from mitosis is the observation that some features of meiosis, such as the meiotic spindles that draw chromosome sets into separate daughter cells upon cell division, and processes regulating cell division employ the same, or similar, molecular machinery as employed in mitosis.

However, there is no compelling evidence for a period in the early evolution of eukaryotes during which meiosis and accompanying sexual capability was suspended. Presumably such a suspension would have occurred while the evolution of mitosis proceeded from the more primitive chromosome

replication/segregation processes in ancestral bacteria until mitosis was established.

NOTES

In addition, as noted by Wilkins and Holliday, there are four novel steps needed in meiosis that are not present in mitosis. These are: (1) pairing of homologous chromosomes, (2) extensive recombination between homologs; (3) suppression of sister chromatid separation in the first meiotic division; and (4) avoiding chromosome replication during the second meiotic division. They note that the simultaneous appearance of these steps appears to be impossible, and the selective advantage for separate mutations to cause these steps is problematic, because the entire sequence is required for reliable production of a set of haploid chromosomes.

The two contrasting views on the origin of meiosis are (1) that it evolved from the bacterial sexual process of transformation and (2) that it evolved from mitosis. The two contrasting views on the fundamental adaptive function of meiosis are: (1) that it is primarily an adaptation for repairing damage in the DNA to be transmitted to progeny and (2) that it is primarily an adaptation for generating genetic variation among progeny. At present, these differing views on the origin and benefit of meiosis are not resolved among biologists.

10.4 Aneuploids and Euploids

Meiosis occurs in all animals and plants. The end result, the production of gametes with half the number of chromosomes as the parent cell, is the same, but the detailed process is different. In animals, meiosis produces gametes directly. In land plants and some algae, there is an alternation of generations such that meiosis in the diploid sporophyte generation produces haploid spores. These spores multiply by mitosis, developing into the haploid gametophyte generation, which then gives rise to gametes directly (i.e. without further meiosis). In both animals and plants, the final stage is for the gametes to fuse, restoring the original number of chromosomes.

Chromosome Mutation: Changes in Numbers

I. Aberrant euploidy

A. General Informatiom

1. Euploidy refers to the situation in which an organism has one complete set of chromosomes or an integer multiple of a set.
2. The number of chromosomes in a basic set called monoploid number (x).
3. The "normal" situations are 1x = haploid or 2x=diploid.
4. Polyploids (>2x) are by definition abnormal although they exist in plants in nature.
5. Haploid number (n) refers only to number of chromosomes in gametes. Formost diploid organisms that we are familiar with the haploid and monoploidnumber is the same.However, this is not always the case: wheat is a hexaploid containing six sets ofsimilar but not identical chromosomes with 7 chromosomes per set (x=7 and 6x =42). The gametes however have 21 chromosomes (n=21 and 2n=42).

NOTES

Type 1 Type 2 Type 3 Type 4

Type 5 Type 6 Type 7

Meiosis

Type 1 Type 2 Type 3 Type 4

Type 5 Type 6 Type 7

x = 7 6x = 42

n = 21 2n = 42A. Monoploidy (1x = 1 set of chromosomes instead of normal 2)

1. Monoploidy can arise from spontaneous development of an unfertilized egg (parthenogenesis).
2. Lethal in most animals because unmasking of lethal alleles.
3. Monoploids usually sterile because of a problem with meiosis: No pairing partner for each chromosome. Gametes containing all the chromosomes in the set are formed at a frequency of (1/2)x-1
4. Exception in animals are male bees, wasps, and ants which develop parthenogenetically and in which gametes are formed by mitosis
5. Monoploids are important in plant breeding. For example:Diploids mask advantageous new mutations. Can use monoploidy tocircumvent

NOTES

this: Induced in plants anthers by taking cells destined tobecome pollen and treating with cold will become small mass of dividing cells – embryoid Plate on agar and grows into plantlet look for desirable traits Use colchicine to redouble chromosomes returning to diploidy in sector of plant that can then be transplanted.

C. General information on polyploidy.

1. In plants

   a) Autoploidy (multiple of sets from one species) vs. alloploidy (multiple sets from different species)

   b) Even ploidy vs. odd ploidy: Odd ploidy always have unpaired chromosome for each chromosome type during meiosis; thus, the probability of producing genetically balanced gametes is low, and odd ploidies are effectively sterile.

2. In animals

   a) Polyploidy is rare. Reasons for rareness of polyploidy in animals is not certain. Hypothesis is the sex-determination mechanism depends on delicate balance in chromosomes. Also, since most animals do not undergo self-fertilization, there is no chance for polyploids to successfully produce progeny that will be fertile (i.e. 2n X n = 3x (sterile).

   b) In humans, polyploid zygotes usually die in utero (account for approx. 20% of spontaneous abortions.

   c) There are a few examples of polyploids in animals including flatworms, leeches, brine shrimp, some polyploid frogs and toads, and some fish.

D. Humans have giant polyploid nuclei in such tissues as liver and kidneys. Somatic cell polyploidy called endopolyploidy.D. Triploids (3x = 3 sets of chromosomes instead of normal 2 sets)

1. Usually autopolyploids

2. Can arise from spontaneous aberrant meiosis in nature to give a 2n gamete that fuses with a normal n gamete to give a 3x zygote OR from tetraploid (4x) and diploid (2x) cross.

3. Usually sterile because the frequency of getting exactly 1 set or exactly 2 sets of all the chromosomes to migrate to one pole is (1/2)x-1 Gametes that contain intermediate between 1n and 2n exhibit genome imbalance because there are different ratios of genes depending upon if the gamete has 1 or 2 copies of each chromosome.

NOTES

4. Seedless watermelons and bananas are triploid. They are sterile, but are propagated vegatatively.

E. Autotetraploids in plants (4x where all 4 sets all derived from one species)

   1. Arise from spontaneous aberrant meiosis in nature to give a 2n gamete that fuses with another 2n gamete to give a 4x zygote or by accidental doubling of 2x genome to a 4x genome in the 1 cell stage of development

   2. Important in crop plants because greater chromosomes = larger plant. (alfalfa, coffee, peanuts, Macintosh apples)

   3. Because there are an even of chromosome sets, pairing of homologous chromosomes allows for normal meiosis.

   4. Single locus segregation ratios in are different than simple Mendelian ratios

   a) With a AAaa plant where A/a is close to the centromere:

F. Allotetraploids (4 sets from different species: 2 sets from species 1 and 2 sets from species 2)

   1. Chromosomes are homeologous (similar but not completely homologous)

   2. Important in plant breeding:

   3. In nature, the same events as (2) happen except that chromosome doubling is a rare spontaneous event. Allotetraploidy was

NOTES

important in the evolution of plants (cotton, wheat).II. Aneuploidy

A. General information

1 The situation in which one or several chromosomes are lost from or added to the normal set of chromosomes

2. Caused by nondisjunction during meiosis to generate n+1 or n-1 gametes which when fertilized with n gamete to generate 2n+1 or 2n-1 zygotes

3. If autosome is affected, usually lethal in animals because of either chromosome imbalance or unmasking of lethal alleles; exception is the sex chromosomes).

B. Nullsomics

1. Missing two homologous chromosomes for a certain chromosome #

2. 2n-2

3. Lethal in human diploids because a whole chromosome type has been deleted

4. Can be ok in some polyploid plants

C. Monosomics

1. Missing 1 chromosome of a certain chromosome #

2. for diploids: 2n-1

3. In humans, lethal for autosomes (die in utero)

4. Turner syndrome (only 1 X chromosome) Figure 9-17

5. Can be ok in some plants – in fact can use monosomic banks to map recessive alleles to the chromosomes

D. Trisomy

1. Have 1 extra chromosome of a certain chromosome #

2. For diploids: 2n+1

3. sex chromosome trisomies

a) Klinefelter's syndrome (XXY) Figure 9-20

b XYY syndrome (XYY)

c Triplo-X (XXX)

4. autosomal trisomies

a) Trisomy 21 (Down's syndrome) Figure 9-21

b) Trisomy 13 (Patau syndrome)

c) Trisomy 18 (Edwards syndrome)

5. Can be ok in some plants

3. Single locus segregation ratios in are different than simple Mendelian ratios

a) With an Aaa plant where A/a is close to the centromere: Gamete ratios of 1A: 2Aa: 2a: 1 aa Aberration in haploid organisms where there is an extra chromosome

F. Somatic Aneuploidy

In above mentioned cases, the chromosome mutations all arose in the germ cells during meiosis. However, aneuploidy can arise in the somatic tissues during mitosis resulting in a genetic mosaic of cell types.

NOTES

10.5 Origin and Production of Autopolyploids & Allopolyploids

Polyploidy is pervasive in plants and some estimates suggest that 30–80% of living plant species are polyploid, and many lineages show evidence of ancient polyploidy (paleopolyploidy) in their genome. Huge explosions in angiosperm species diversity appear to have coincided with the timing of ancient genome duplications shared by many species. It has been established that 15% of angiosperm and 31% of fern speciation events are accompanied by ploidy increase. Polyploid plants can arise spontaneously in nature by several mechanisms, including meiotic or mitotic failures, and fusion of unreduced (2n) gametes. Both autopolyploids (e.g. potato) and allopolyploids (e.g. canola, wheat, cotton) can be found among both wild and domesticated plant species. Most polyploids display heterosis relative to their parental species, and

NOTES

may display novel variation or morphologies that may contribute to the processes of speciation and eco-niche exploitation. The mechanisms leading to novel variation in newly formed allopolyploids may include gene dosage effects (resulting from more numerous copies of genome content), the reunion of divergent gene regulatory hierarchies, chromosomal rearrangements, and epigenetic remodeling, all of which affect gene content and/or expression levels. Many of these rapid changes may contribute to reproductive isolation and speciation.

Lomatia tasmanica is an extremely rare Tasmanian shrub that is triploid and sterile; reproduction is entirely vegetative, with all plants having the same genetic constitution.

There are few naturally occurring polyploid conifers. One example is the Coast Redwood *Sequoia sempervirens*, which is a hexaploid (6x) with 66 chromosomes (2n = 6x = 66), although the origin is unclear.Aquatic plants, especially the Monocotyledons, include a large number of polyploids

Polyploid crops

The induction of polyploidy is a common technique to overcome the sterility of a hybrid species during plant breeding. For example, Triticale is the hybrid of wheat (*Triticum turgidum*) and rye (*Secale cereale*). It combines sought-after characteristics of the parents, but the initial hybrids are sterile. After polyploidization, the hybrid becomes fertile and can thus be further propagated to become triticale.In some situations, polyploid crops are preferred because they are sterile. For example, many seedless fruit varieties are seedless as a result of polyploidy. Such crops are propagated using asexual techniques, such as grafting.Polyploidy in crop plants is most commonly induced by treating seeds with the chemical colchicine.

Examples of polyploid crops

Triploid crops: apple, banana, citrus, ginger, watermelon Tetraploid crops: apple, durum or macaroni wheat, cotton, potato, canola/rapeseed, leek, t obacco, peanut, kinnow, Pelargonium

Hexaploid crops: chrysanthemum, bread wheat, triticale, oat, kiwifruit

Octaploid crops: strawberry, dahlia, pansies, sugar cane, oca (*Oxalis tuberosa*) Some crops are found in a variety of ploidies: tulips and lilies are commonly found as both diploid and triploid; daylilies *(Hemerocallis* cultivars) are available as either diploid or tetraploid; apples and kinnows can be diploid, triploid, or tetraploid.Here the relation between genotype and phenotype is illustrated, using a Punnett square, for the character of petal colour in pea.

NOTES

10.6 Genomic Constitution and Analysis

The genotype is the genetic makeup of a cell, an organism, or an individual usually with reference to a specific characteristic under consideration. For instance, the human CFTR gene, which encodes a protein that transports chloride ions across cell membranes, can be dominant (A) as the normal version of the gene, or recessive (a) as a mutated version of the gene. Individuals receiving two recessive alleles will be diagnosed with Cystic fibrosis. It is generally accepted that inherited genotype, transmitted epigenetic factors, and non-hereditary environmental variation contribute to the phenotype of an individual.Non-hereditary DNA mutations are not classically understood as representing the individual's genotype. Hence, scientists and physicians sometimes talk for example about the (geno)type of a particular cancer, that is the genotype of the disease as distinct from the diseased.The genotype of an organism is the inherited instructions it carries within its genetic code. Not all organisms with the same genotype look or act the same way because appearance and behavior are modified by environmental and developmental conditions. Likewise, not all organisms that look alike necessarily have the same genotype. One's genotype differs subtly from one's genomic sequence. A sequence is an absolute measure of base composition of an individual, or a representative of a species or group; a genotype typically implies a measurement of how an individual differs or is specialized within a group of individuals or a species. So typically, one refers to an individual's genotype with regard to a particular gene of interest and, in polyploid individuals, it refers to what combination of alleles the individual carries (see homozygous, heterozygous). The genetic constitution of an organism is referred to as its genotype, such as the letters Bb. (B - dominant genotype and b - recessive genotype)The distinction between genotype and phenotype is commonly

NOTES

experienced when studying family patterns for certain hereditary diseases or conditions, for example, haemophilia. Due to the diploidy of humans (and most animals), there are two alleles for any given gene. These alleles can be the same (homozygous) or different (heterozygous), depending on the individual (see zygote). With a dominant allele, the offspring is guaranteed to inherit the trait in question irrespective of the second allele.In the case of an albino with a recessive allele (aa), the phenotype depends upon the other allele (Aa, aa or AA). An affected person mating with a heterozygous individual (Aa or aA, also carrier) there is a 50-50 chance the offspring will be albino's phenotype. If a heterozygote mates with another heterozygote, there is 75% chance passing the gene on and only a 25% chance that the gene will be displayed. A homozygous dominant (AA) individual has a normal phenotype and no risk of abnormal offspring. A homozygous recessive individual has an abnormal phenotype and is guaranteed to pass the abnormal gene onto offspring.In the case of haemophilia; it is sex-linked thus only carried on the X chromosome. Only females can be a carrier in which the abnormality is not displayed. This woman has a normal phenotype, but runs a 50-50 chance, with an unaffected partner, of passing her abnormal gene on to her offspring. If she mated with a man with haemophilia (another carrier) there would be a 75% chance of passing on the gene.Inspired by the biological concept and usefulness of genotypes, computer science employs simulated phenotypes in genetic programming and evolutionary algorithms. Such techniques can help evolve mathematical in the solution of certain types of otherwise difficult problems.Genotyping is the process of elucidating the genotype an individual with a biological assay. Also known as a genotypic assay, techniques include PCR, DNA fragment analysis, allele specific oligonucleotide (ASO) probes, DNA sequencing, and nucleic acid hybridization to DNA microarrays or beads. Several common genotyping techniques include restriction fragment length polymorphism (RFLP), terminal restriction fragment length polymorphism (t-RFLP)amplified fragment length polymorphism (AFLP) and multiplex ligation-dependent probe amplification (MLPA).DNA fragment analysis can also be used to determine such disease causing genetics aberrations as microsatellite instability (MSI), trisomy or aneuploidy, and loss of heterozygosity (LOH). MSI and LOH in particular have been associated with cancer cell genotypes for colon, breast and cervical

cancer.The most common chromosomal aneuploidy is a trisomy of chromosome 21 which manifests itself as Down syndrome. Current technological limitations typically allow only a fraction of an individual's genotype to be determined efficiently.

NOTES

10.7 Evolution of Major Crop Plants

The origins of agriculture and domesticated crops are intertwined, and the change from a hunter-gatherer mode to tillage, sowing and harvesting was one of the major technologcal innovations of humankind. There is good evidence that this occurred some 10,000 years ago in several different locations, and involved the domestication of wild-relatives of the major crops. Domestication involves changes in the genetic makeup and morphological appearance of plants (and animals). These changes occur because people select the variations of a wild plant that best suit their needs. If the desired features can be passed to offspring, over generations of planting and harvesting, the strains of plants grown become changes. Despite the fact that these domesticated varieties of plants are preferred over their original forebears, the wild-relatives of crop plants continue to be an important resource. Reserves of wild plants offer a pool of genetic diversity. Traits of plants that might have been lost in domestication, sometimes become crucial for protection of domesticated crops from stress and disease . Maintaining wild strains for their gene pool helps ensure food security. Knowledge of crop origins is thus of considerable practical importance, even when these original strains are no longer harvested as crops.Development today of new crops (such as perennial alternatives of currently used annual staples) has potential value in helping meet serious current agricultural challenges such as the need for water use efficiency, better management of land salinization, and soil conservation. A Swiss botanist, Alphonse de Candolle started studies of the origins of crops in 1885, and proposed two approaches to answering these questions. The first is identification of the geographical distribution of wild-relatives of modern crops, based on careful botanical descriptions and tests for cross-pollination between candidate ancestors and the crop in question.Second, archaeological studies often provide clues on how and when a transition from hunter-gatherer existence to agriculture occurred.The earliest origins of major crops based on

carbon-14 dating, date back around 10,000 years, just after the end of the last ice age. This allows the geographical regions in which crop domestication took place to be identified

NOTES

Six independent centers of crop origin

1. Mesoamerica (Southern Mexico and Northern Central America): Maize, Phaseolus beans, Sweet potato, tomato
2. The Andes of South America: Potato, cassava (manioc), pineapple
3. Southwest Asia (including the "Fertile Crescent": Wheat, barley, pea, lentil
4. The Sahel region and Ethiopian highlands of Africa: Sorghum, coffee, melon, watermelon
5. China: Asian rice, soybean, adzuki bean, orange, apricot, peach, tea
6. Southeast Asia: Cucumber, banana, plantain

10.8 Summary

Chromosome abnormalities contribute significantly to genetic disease. This impact is seen in the effect on the fetus or individual directly or in the ability to produce healthy offspring.Autosomal abnormalities are generally more detrimental than sex chromosomes abnormalities. Abnormalities involving entire chromosomes or subtle micro deletion can result in clinically abnormal syndrome.

10.9 Glossary

- **Aneuploidy** - The chromosomal constitution of cells which deviate from the normal by the addition or subtraction of CHROMOSOMES, chromosome pairs, or chromosome fragments. In a normally diploid cell (DIPLOIDY) the loss of a chromosome pair is termed nullisomy (symbol: 2N-2), the loss of a single chromosome is MONOSOMY (symbol: 2N-1), the addition of a chromosome pair is tetrasomy (symbol: 2N+2), the addition of a single chromosome is TRISOMY (symbol: 2N+1).

- **Chromosomal Instability** - An increased tendency to acquire CHROMOSOME ABERRATIONS when various processes involved in chromosome replication, repair, or segregation are dysfunctional.
- Chromosome Deletion - Actual loss of a portion of the chromosome.
- **Inversion (Genetics)** - An aberration in which a chromosomal segment is deleted and reinserted in the same place but turned 180 degrees from its original orientation, so that the gene sequence for the segment is reversed with respect to that of the rest of the chromosome.
- **Nondisjunction, Genetic** - The failure of homologous CHROMOSOMES or CHROMATIDS to segregate during MITOSIS or MEIOSIS with the result that one daughter cell has both of a pair of parental chromosomes or chromatids and the other has none.

NOTES

10.10 Self-Learning Exercise

Section-A (Very Short Answer Type Questions)

1 Which syndrome is also known as Trisomy 18?

2 Which syndrome is also known as Trisomy 21?

3 Which of the following chromosome abnormality is also known as Turner syndrome?

4 Which chromosome abnormality is known as Patau syndrome?

5 XXY is known as which syndrome?

Section-B (Short Answer Type Questions)

1 What is a centromere?

2 What happens if due to an error,centromere is deleted from a chromosome?

3 What are the different types of numerical chromosomal aberrations?

Section-C (Long Answer Type Questions)

1 What is the most common type of triploidy?

2 What are the structural features of a chromosome?

3 What are the outcomes of monosomics and trisomics of sex chromosomes?

Answer Key to Section A

1. Edward syndrome
2. Down syndrome
3. X chromosome abnormality
4. Chromosome 13
5. Klinfelter's syndrome

NOTES

10.11 References

- .NHGRI. 2006. Chromosome Abnormalities
- Rieger, R.; Michaelis, A.; Green, M.M. (1968). "Mutation". A glossary of genetics and cytogenetics: Classical and molecular. New York: Springer-Verlag. ISBN 9780387076683.
- Gardner, R. J. McKinlay, and G. R. Sutherland. *Chromosome Abnormalities and Genetic Counseling.* Oxford: Oxford University Press, 1996.
- John, Bernard. *Meiosis.* New York: Cambridge University Press, 1990.
- Hassold T, Abruzzo M, Adkins K et al. (1996) Human aneuploidy: incidence, origin, and etiology. Environmental and Molecular Mutagenesis 28: 167–175.

Unit - 11

Genetic Recombination and Genetic Mapping

NOTES

Structure of the Unit

11.0 Objectives

The Objectives of this unit is to explain the subject to student in a very simple way. After studied this unit student formal the knowledge about

- Genetic recombination
- Linkage and crossing over
- Mutation
- DNA Damage and Repair

11.1 Introduction

NOTES

Chromosome theory of inheritance based upon genetic recombination and genetic mapping. According to "Chromosome Theory of Inheritance" the genes are carried in the chromosomes. It means each chromosome bears many genes. The genes located on the homologous chromosomes cannot assort independently, while these tend to carried together generation after generation. This phenomenon of inheritance of genes together in offspring as their parental combination is known as linkage. The genes situated in the some chromosome and being inherited together are known as linked genes, and the characters regulated by linked genes are linked characters.

11.2 Linkage

T.H. Morgan in 1911 first time propounded the theory of linkage. Morgan (1910) while working on Drosophila he described 'linkage'. He defined linkage, the tendency of the genes, as present in the same chromosome, to remain in their previous combination and to enter together in the same gamete.

11.2.1 Chromosome Theory of Linkage

Morgan and Castle observed some characteristic and formulated. 'The Chromosome Theory of Linkage'.

1. Linkage genes are situated in the some chromosome.
2. Linkage genes are arranged in a linear sequence in the chromosome.
3. Closely located genes show strong linkage whereas distance between the linked gene make it weak linkage.
4. Linked genes remain together during sexual reproduction.

Linkage gene and independent assortment

Each chromosome must contain many gene and these genes would not be expected to assort independently since the basic of independent assortment is the independent segregation of the different pairs of homologous chromosomes during the reduction division of meiosis.

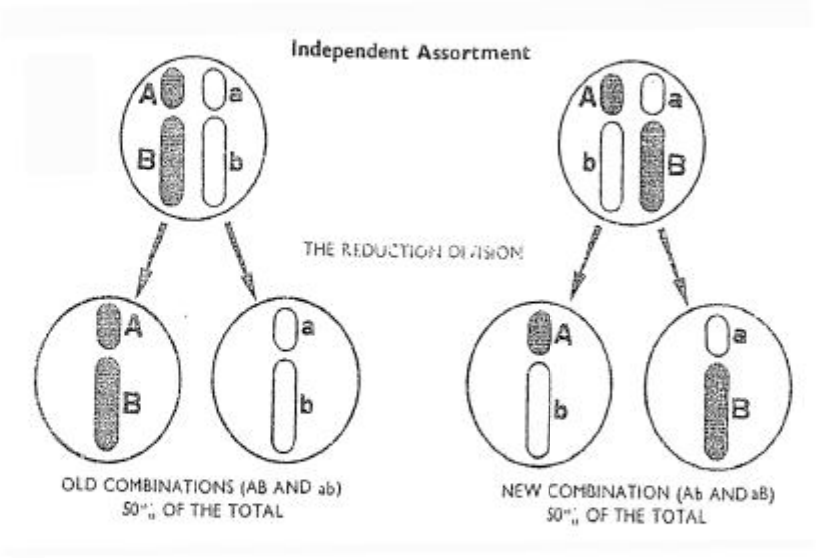


Fig. 11.1 : Independent Assortment

The difference between linkage and independent assortment can be explained by the following figure.

In this figure linkage group of gene located on the single chromosome.

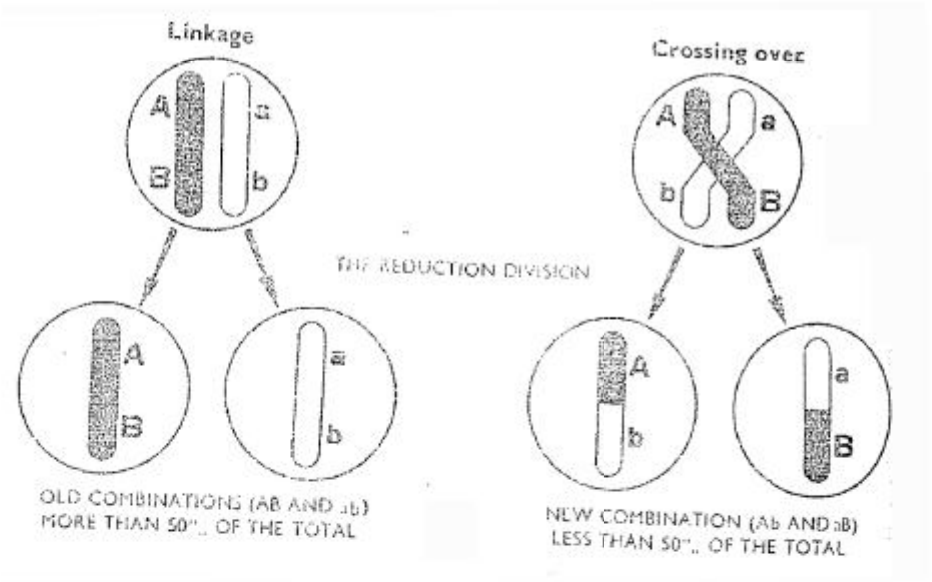


Fig. 11.2 : Linkage and Crossing Over

11.2.2 Complete and Incomplete Linkage

Linkage can be understand with the help of hypothetical example, where gene A and B are involved, a and b being their recessive alleles respectively. What types and frequencies of offspring would be expect from such a test cross. A

NOTES

cross AB/AB × ab/ab would give rise to a F_1 dihybrid AB/ab. This dihybrid will then be crossed with double recessive parent ab/ab to get the test cross progeny. The result of this dihybrid (AB/ab) is presence and absence of crossing over.

What types and frequencies of offspring would we expect from such a testcross? We know that, all the gametes produced by the doubly homozygous recessive test cross parent will be ab in genotype. The F_1 parents would then be AbBb and F_1 parents would produce gametes of four type AaBb, aabb, Abab, aBab in equal proportion (1:1:1:1).

Thus, the expected frequencies of testcross progeny will be 50% with parental phenotypes and 50% of recombinant phenotypes.

On the basis of absence or presence of non-parental or new combinations of linked genes linkage has been divided complete and incomplete linkage.

1. Complete linkage: In this type linked genes are transmitted together to the offspring only in their original or parental combination for several germinations.

eg. complete linkage in male Drosophila

A cross between wild type Drosophila with grey body and vestigial wings (BBvv) and Drosophila with black body and long wings (bbvv) produce F_1 offspring, all of which have grey body and long wings (BbVv). These F_1 male hybrids when back crossed with a double recessive female (test cross) having black body and vestigial wing (bbvv) produce off-springs of two types in equal proportion.

Complete linkage in Drosophila

In the above example, the off springs exhibit only the parental combinations on characters here new combinations are not formed. This shows complete linkage (Fig. 11.3).

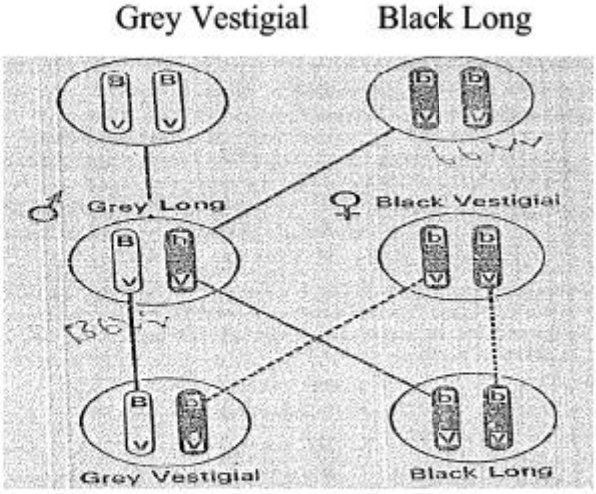


Fig.11.3 : Complete linkage in Drosophila

2. Incomplete linkage: During meiosis in gametogenesis the homologous chromosomes pairing and make new combination with the help of interchange of chromosome segments. This process called crossing over.

We can understand the incomplete linkage with a example in maize.

When one variety of maize having coloured and normally filled seeds, is crossed with another variety having colourless and shrunken seeds all the F_1 plants produce coloured and full seeds. But in a test cross, when these F_2 female hybrids are cross pollinated with the pollens from a plant having colourless and shrunken seeds (double recessive) four types of seeds are produced.

(i) Coloured full (CS)
(ii) Colourless shrunken (cs)
(iii) Coloured shrunken (Cs)
(iv) Colourless full (cS)

In this example the parental combinations are more numerous rather than new combination. This clearly indicates that these characters are linked together. Their genes are locked on the same chromosome and only in 3.6% individuals these have become separated by crossing over (Fig. 11.4).

NOTES

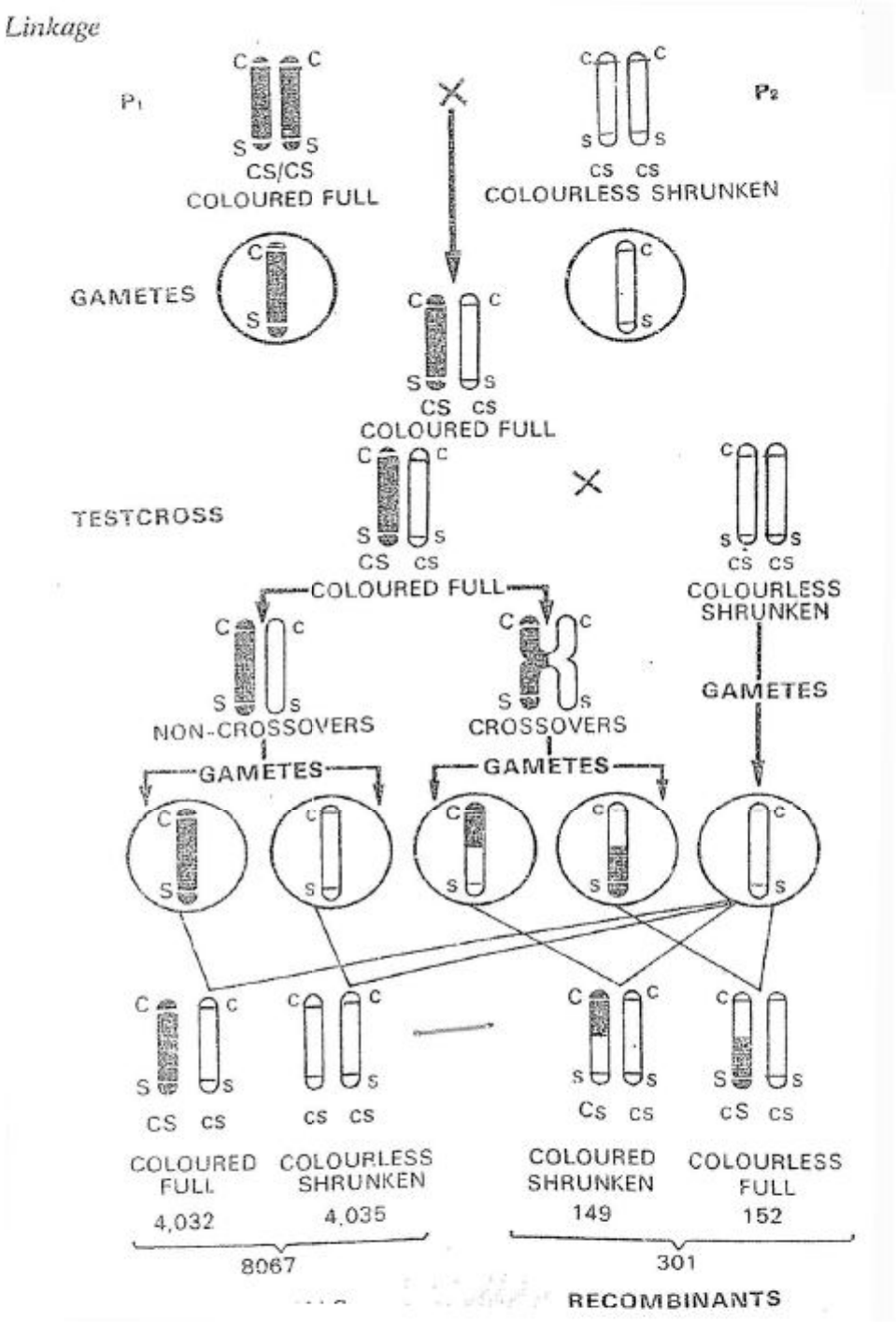


Fig. 11.4 : The results of a cross in maize between coloured full seeds and colourless shrunken seeds showing linkage when characters are in coupling phase (Experiment conducted by Hutchinson)

11.3 Crossing over

Crossing-over, is a process where homologous chromosomes are interchange and as a result new combination of genes brought out. Crossing over and independent assortment are mechanisms that produce new combinations of gene.

NOTES

Mechanism of recombination or crossing over

We can understand the process of crossing over as follows:

1. Synapsis: During prophase 1 of meiosis the homologues chromosomes from parents come close together and pair at zygotene substage (synapsis). The homologous chromosomes come close together at one or more points by mutual attraction between the allelic genes. In phase of zygotene they lie side by side in close approximation all along their lengths. The pairing is exact and point to point. In this process chromosome are known as bivalents.

Ultrastructure of synapsis

During synapsis a synaptinemal complex was observed by Montrose J. Moses (1955). It occurs as a highly organized.

Structure of filaments between the paired chromosomes in zygotene and pachytene stage of meiosis shown in Fig. 11.5.

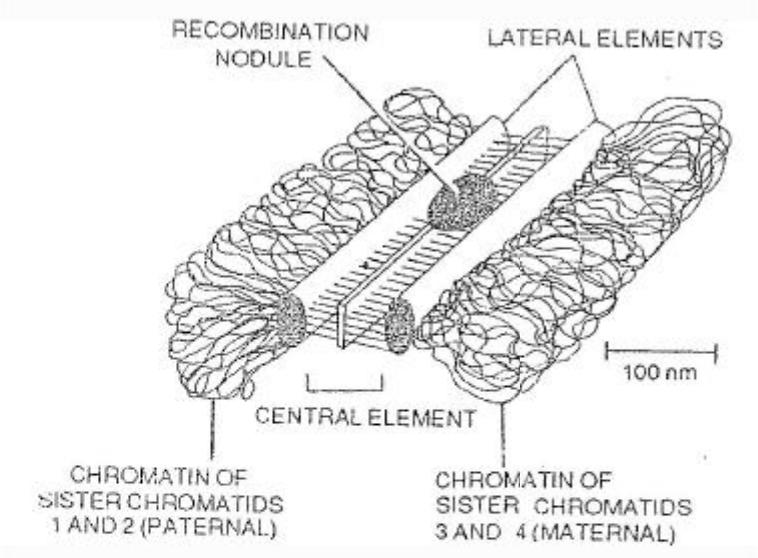


Fig. 11.5 : Structure of Synaptonemal Complex Formed During Meiotic Pairing

NOTES

2. Duplication of chromosomes: Homologous chromosomes in a bivalent split longitudinally into two sister chromatids. Duplication take place in diplotene stage. Thus the bivalent now consists of four chromatids and is known as tetrad. The longitudinal splitting of chromosomes is achieved by the separation of already duplicated DNA molecules.

3. Crossing over: During diplotene stage, when the paired chromosomes start separating, the chromatids remain in contact at one or more points and thus establish one or more exchanges per bivalent. Chromosomes's contact point called chiasmata.

All each chaiasma two non-sister chromatids of the bivalent break at the corresponding points and then rejoin with the exchange of segments. The new chromatids formed as a result of exchange of segments are formed of segments derived from two non-sister chromatids of the bivalent.

Breakage and fusion of broken segments take place due to action of enzyme respectively by endonuclease and ligase (Fig. 11.6).

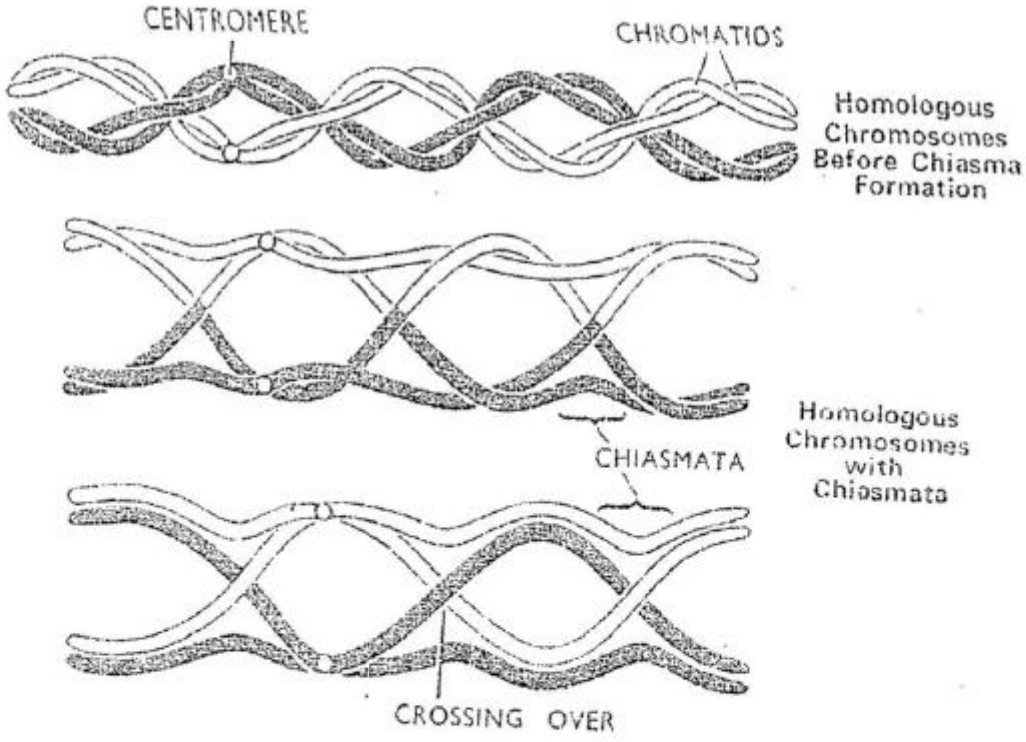


Fig. 11.6 : Mechanism of Crossing Over

4. Terminalization: After crossing over, the nonsister chromatids start repelling each other, because the forces of attraction keeping them together lapse. The chromatids separate from the middle centromere to the tip of chromosome and chiasmata also start moving in zipper like fashion towards the ends. Terminalization is the process where chiasmata movement occurs.

NOTES

Essential features of crossing over as follows:

1. The location of a gene on a chromosome is called a locus.
2. The two alleles of a gene in a heterozygote occupy corresponding positions in the homologous chromosomes, like allele A occupies the same position in homology 1 that allele a occupies in homolog 2.
3. The crossing over occurs between the non-sister chromatids of the homologous pair of chromosomes.
4. In meiosis prophase I after the synopsis of the homogenous chromosomes, crossing over occurs at pachytene stage.
5. Chiasmata depend upon the length of chromosomes. The chiasmata are greater when longer the chromosomes.
6. The probability that crossing over will occur between two loci increase with increasing distance between the two loci on the chromosome.

Theories Explaining Mechanism of Crossing Over

1. Darlington's strain theory: Darlington explained crossing over to be the result of strain or torsion produced due to coiling of homologous chromosomes and sister chromatids.

When chromosomes start separating, their relational coiling unravels in one direction and the sister chromatids unravel in the opposite direction (Fig. 11.7).

NOTES

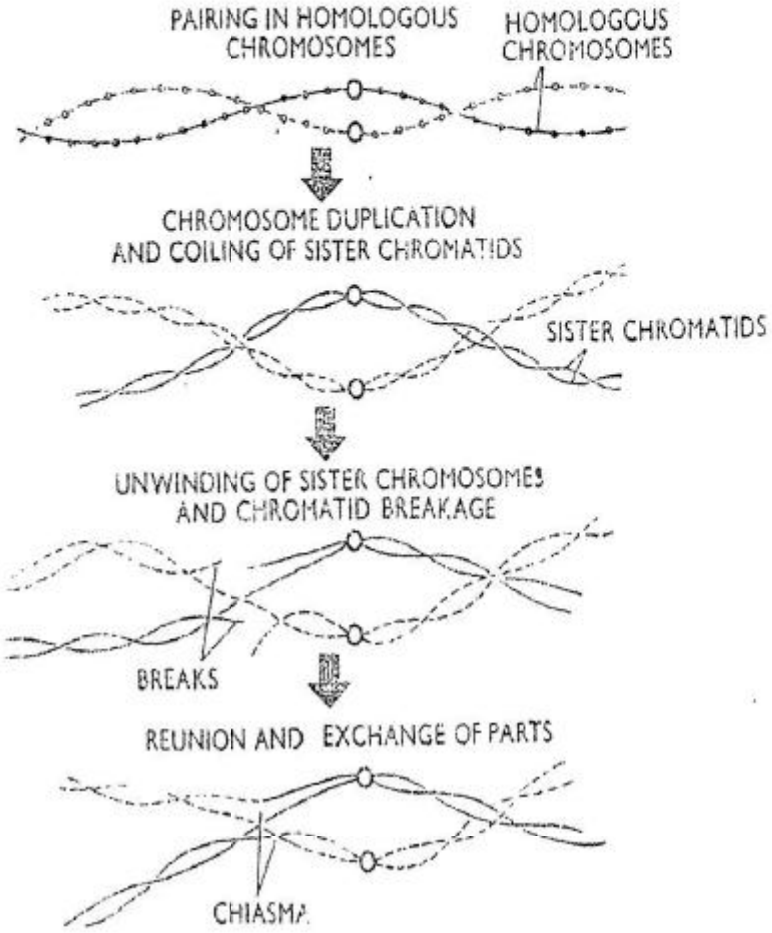


Fig. 11.7 : Darlington's Strain Theory of Crossing Over

2. Belling's hypothesis and copy choice theory: According to Belling crossing over is related with the duplication of chromosomes. Crossing over and recombination occur during the synthesis of new chromatids. The parent chromatids act as templates upon which new chromatids are synthesized. In other words, copy choice model meant that a newly synthesized daughter chromatid is derived due to copying of one chromosomes upto a certain distance and then switching on to the other homologous chromosomes for copying the remaining distance or region of the chromosomes (Fig. 11.8).

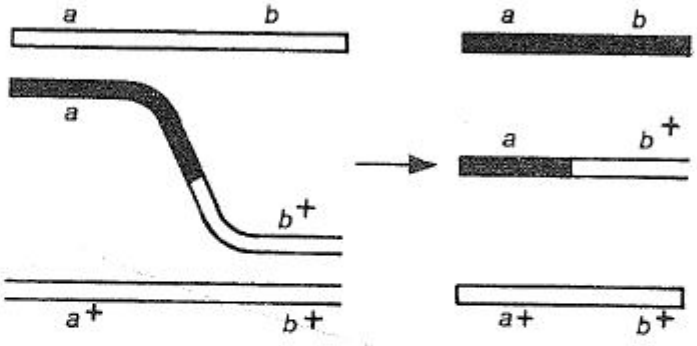


Fig. 11.8 : Mechanism of Genetic Recombination Based on Copy Choice Mechanism

This copy choice model, as a mechanism of recombination, was inadequate for the following reasons: (i) In this model DNA replicates in a semi-conservative manner (ii) The experimental evidence suggests that three strands or four strands may actually be involved in multiple crossovers.

3. Holliday Model for General Recombination: During 1964, R. Holliday proposed a model which had become very popular in name of hybrid DNA models.

In these models, only one strand in each of two DNA duplexes belonging to non-sister chromatids breaks.

Segments of the single strands on one side of each cut are displaced from their complementary strands, probably with the aid of DNA binding proteins. The displaced strands then exchange pairing partners, base pairing with the intact complementary strands of the homologous chromosomes.

Each broken DNA single strand invades the opposite helix and establishes base pairs with its complementary nucleotides. Ligase enzymes seal the broken ends and make the X-shaped points.

The chromosomes with chix (X) are called 'Holliday intermediate' or chi-form or half-chromatid chiasma.

The Holliday intermediate is cut in the bridge region to produce two independent duplexes. The cut might occur along east-west axis and along north-south axis. In first case two noncross over strands having parental configuration and other two hybrids or recombinant strands are produced. In second case all the strands of both duplexes are hybrids (Fig. 11.9).

NOTES

NOTES

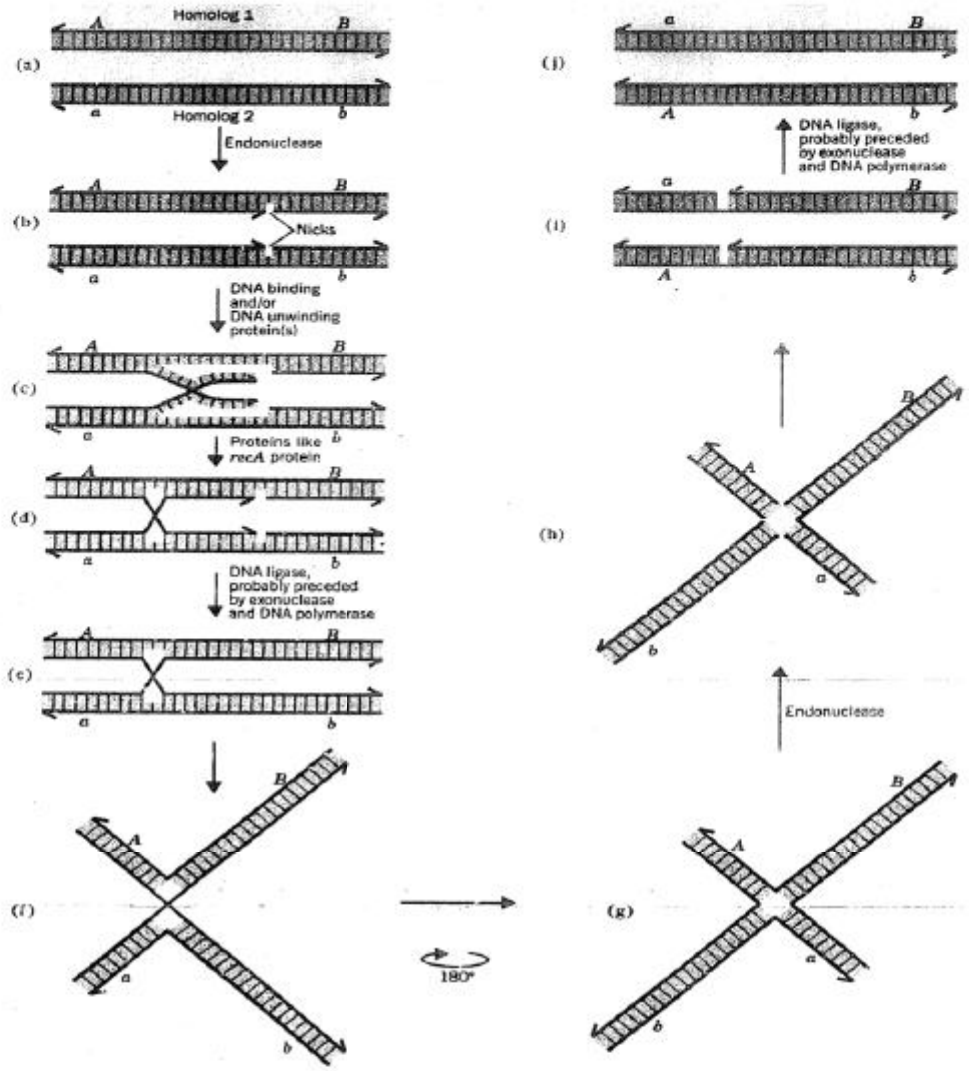
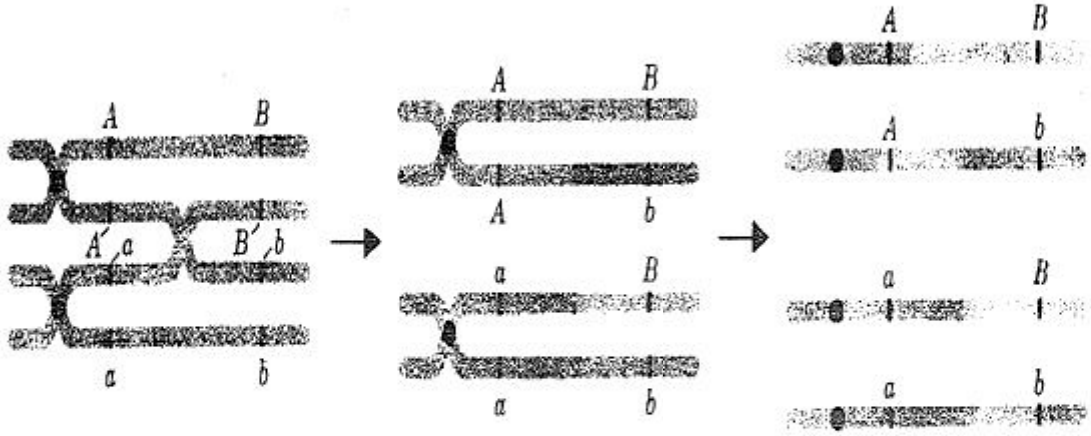


Fig. 11.9 : The Holiday Model for Genernal Recombination

Crossing over occurs in the Tetrad stage of Meiosis

Crossing over occurs after chromosomes duplication, when each homologous chromosome is represented by two chromatids. Each pair of synapsed homologs is called a tetrad, because it consists of four chromatids. Note that crossing over refers to the exchange of segments of chromosomes whereas recombination implies the formation of new combinations of gene. Thus crossing over occur in completely homogzygous organisms, but new combinations can be formed only in organisms that are heterozygous at two or more loci (Fig. 11.10).



NOTES

Fig. 11.10 : Crossing Over

11.4 Genetic Mapping

Strutevant (1913) developed the idea from linkage and crossing over observations that frequencies or percentage of crossing over can be used as a tool to determine the relative distance between the genes in a linkage group and their arrangement at the chromosomes.

Strutevant and Morgan marked the five position of genes on the X-chromosome of Drosophila. This graphic representation of genes is now known as chromosome map.

We can explained linkage maps by using a hypothetical example.

Let us presume that there are three genes A, B and C present on the same chromosome. These are A-B-C, A-C-B or B-A-C three possible linear orders could be present on chromosome. Therefore, in finding out the linear order, one has to really find out the gene which is present in the centre. For this purpose a three point test cross is made, which involves crossing of a trihybrid ABC/abc with triple homozygous recessive abc/abc. The progeny obtained will represent the gametes formed by the hybrid.

Presuming A-B-C as the order of genes, results expected can be express as in Fig. 11.11.

NOTES

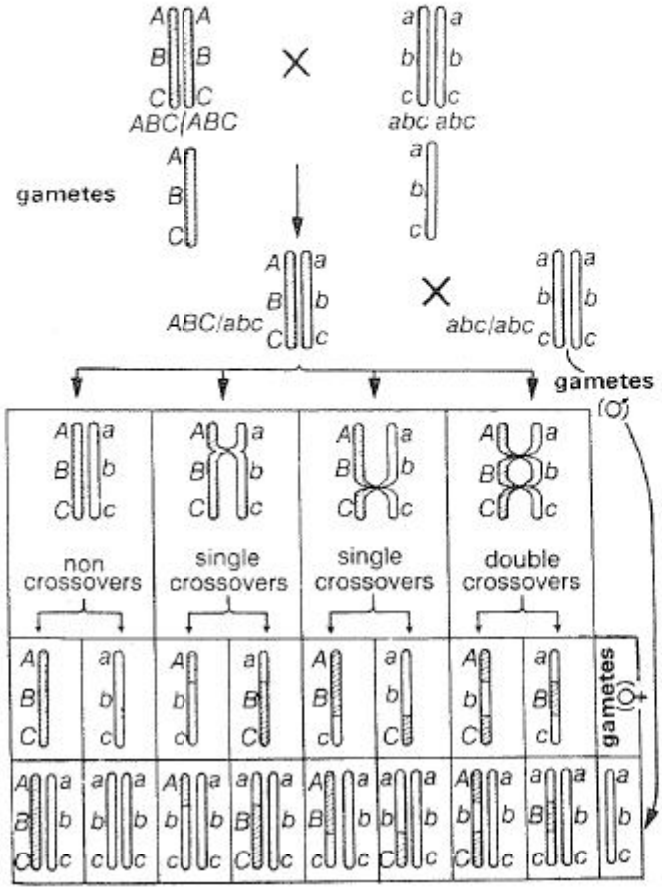


Fig. 11.11 : A Three Point Testcross Involving Three Hypothetical Genes A,B and C

Construction of chromosome maps

The percentage of recombination between two genes represents the distance between these two genes. We can study under the following steps of chromosome mapping process.

1. Methods of ascertaining the linkage groups of genes

Cross breading experiments are conducted between the animals of a species considering three or more characteristics simultaneously and it is ascertained these crosses produce results which shows expected ratio on the basis of independent assortment, i.e. such crosses produce individuals in which parental combinations of genes are found to be more numerous than the new combinations. This indicates that the genes in question form a linkage group.

2. Locating the relative position of genes in a chromosome

In fact genes are plotted on the chromosome on the basis of crossing over results between different pairs of linked genes. The actual distance between two genes is said to be equivalent to the percentage of crossing over between these genes i.e. 10%.

NOTES

Crossing over chances between genes a and b suggest that these are to be placed on the chromosome at a distance of 10 units (Fig. 11.12).

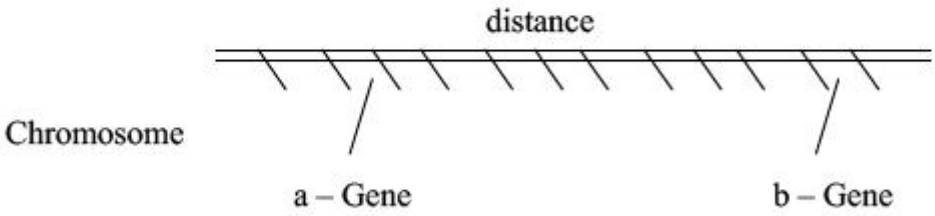


Fig. 11.12 : Crossing over chances 10%

Interference and coincidence

The crossing over is greatly reduced by the phenomenon of interference and coincidence.

Interference. In the cases of double crossing over the frequency of double cross over is found to be much below the expected value. This is because the crossing over and chiasma formation at one point in the homologous non-sister chromatids interferes with the crossing over and chiasma formation at other points nearly. This is known as interference.

Coincidence. Coincidence or coefficient of coincidence is inverse measure of interference and is expressed as the ratio between the actual number of double cross overs and the expected number of such double cross overs. That is

$$\text{Coincidence} = \frac{\text{Actual number of double cross over}}{\text{Expected number of double cross overs}}$$

Linkage maps of Drosophila

The chromosome maps of *Drosophila melanogaster* include four linkage groups corresponding to four chromosome pairs.

The genetic length for the four chromosomes, measured in terms of percentage frequencies of crossing over between genes are estimated to be 66 units for the

X-chromosome, 107.3 for the second, 106.2 for the third and only 0.2 units for the fourth chromosome (Fig. 11.13).

NOTES

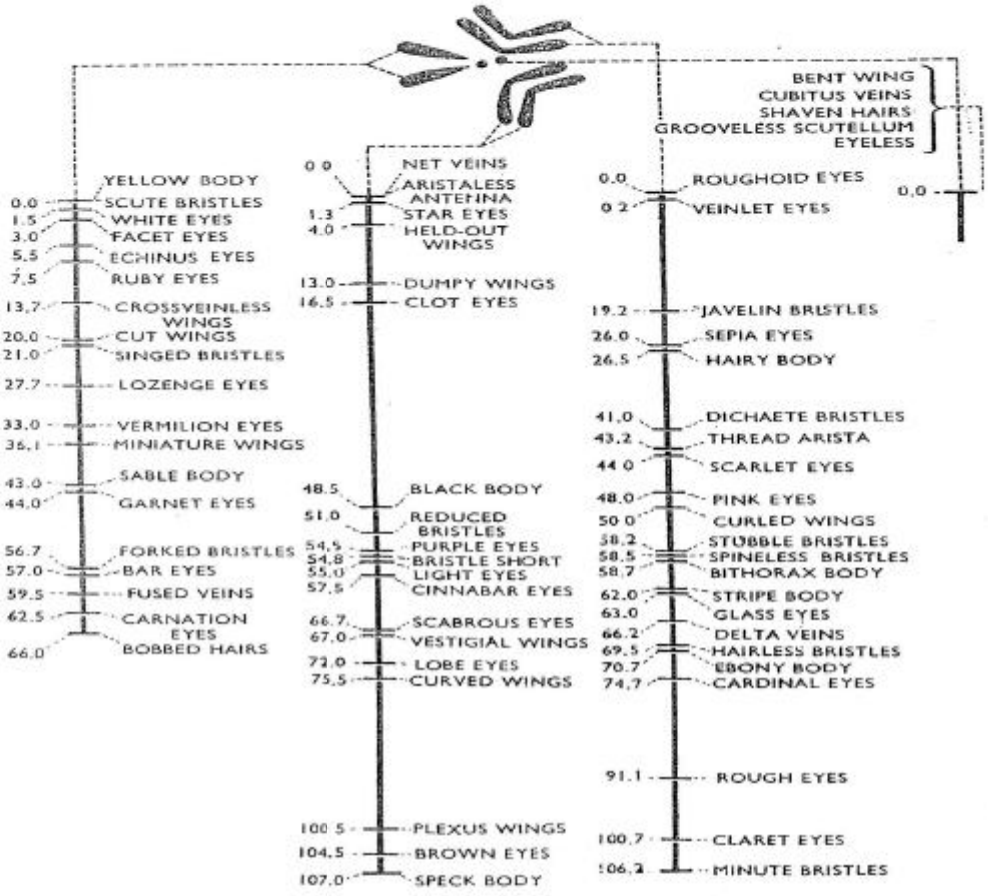


Fig. 11.13 : Linkage Map of Chromosomes of *Drosophila* Showing Alignment of Different Genes on the Four Chromosomes

11.5 Mutation

The term of mutation was introduced by Hugo Devries. Mutation can be define any changes in the gene structure and composition which occur naturally reproducing population. "Mutation is hereditary change in the genetic make up of an individual".

Spontaneous versus Induced Mutations

Mutations are rare events in nature and are then described as spontaneous mutations. Due to their rare occurrence, sometimes, it is difficult to identify and score them.

Spontaneous mutation also called Reverse Mutations and forward mutations because most mutant events consist of a change from normal or wild type to a new genotype same mutant genotype changes back to the wild type.

Spontaneous mutations are easy to be identified in microorganisms, especially the nutritional reverse mutations. The occurrence of reverse mutation usually helps in distinguishing the point mutations from large mutational effects.

NOTES

Point mutations caused by base substitution without a significant loss or gains of genetic material are more easily reversible than the mutations caused by deletion or duplications.

11.5.1 Induced Mutations

Mutations can be artificially induced with the help of mutagenic agents which can be broadly classified into two groups:

(a) Physical mutagens

(b) Chemical mutagens

11.5.2. Mutagens

Mutagens are any physical and chemical agents, which can be produce mutations.

There are two main classes of mutagens:

(i) Radiation (ii) Chemical agents

I. Radiation: Radiations which radiate all forms of energy by naturally in the environment can be produce mutation in almost all the organism.

Biological Effects of Radiations

On the basis of penetration and ionizing power of the rays radiations can be of two types:

(i) Effect of ionizing radiations: The ionizing radiations include X-rays, gamma rays, alpha and beta rays, neutrons, protons and other fast moving particles. These radiations produce breaks in the chromosome and chromatids and abnormal mitosis in the irradiate cells.

The chromosome breaks then lead to the losses of chromosomes, chromosome segments, deficiencies, duplications, translocations or inversions.

These changes result in the formation of abnormal chromosomes and abnormal chromosome number in the daughter cells causing abnormal functioning of the cell or the cell death.

NOTES

(ii) Effects of non-inonizing radiations: The nonionizing radiations, like ultraviolet light rays have longer wave lengths and carry much lower energy. Therefore, their penetration power is much less than X-rays. The ultraviolet rays are absorbed by nucleic acids and cause alterations in the bond characteristics of purines and pyrimidines.

2. Chemical mutagens

Many geneticists feel that chemical mutagens are more hazardous to man than radiations. The list of chemical mutagens becomes longer day by day.

Most powerful chemical mutagens are:

(i) Mustard gas
(ii) Ethyl urethane
(iii) Formaldehyde
(iv) Organic peroxides
(v) Ferrous and manganous salts
(vi) L.S.D. (Lysergic acid dimethylamide)

11.5.3 Molecular Basis of Gene Mutations

Mutations at molecular level, should mean permanent alternation in sequence of nucleotides (base) in the nucleic acid. The smallest change may involve the addition, deletion or substitution of a single nucleotide pair in the DNA molecule, which may change the reading of genetic code and ultimately may be manifested into an altered phenotype.

In molecular level mutations include very limited segment of DNA, these are called point mutations.

Point mutation can be divided into two types.

1. Substitution mutations
2. Frame shift mutations

1. In a substitution mutation a nitrogenous based a triplet codon of DNA is replaced by another nitrogen base or some derivative of the nitrogen

base, changing the codon. The altered codon may code for a different amino acid and may result in the formation of a protein molecule with a single amino acid substitution.

The substitution mutations may be the following two types.

NOTES

1. Transitions 2. Transversions

1. Transitions are those base pair replacements, where a purine is replaced by another purine and a pyrimidine is replaced by another pyrimidine.

A ⟷ G

T ⟷ C

It means that AT is replaced by GC and vice versa (Fig. 11.14).

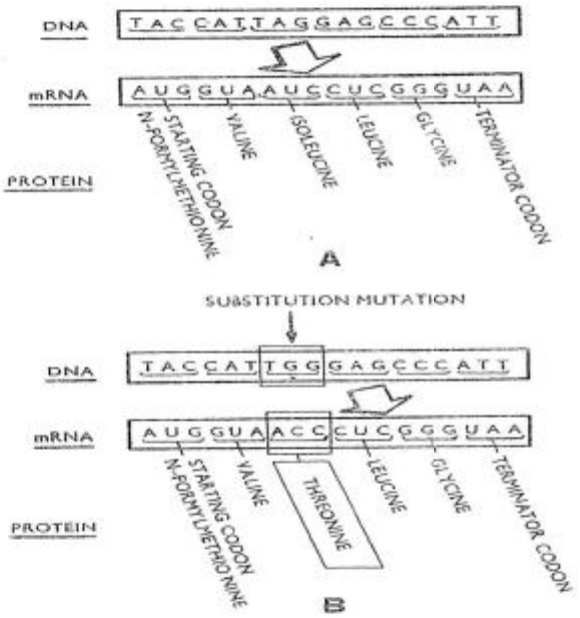


Fig. 11.14 : Effect of Substitution : A Normal Nitrogen Base Sequence B – Base Sequence Changed due to Substitution

The transitional substitutions can be possible by many ways either during DNA replication or other mechanism.

11.6 Tautomerisation

In a molecule of DNA, the purine is linked to the pyrimidine. A is linked to T by two hydrogen bond while G is linked to C by three hydrogen bond. These

arrangement are called tautomers and are formed by the rearrangement in the distribution by hydrogen atoms – tautomeric shifts.

In this tautomeric shift, the amino ($-NH_2$) group of cytosine and adenine is converted in imino (-NH) group and linkwise ketogroup (C=O) or thymine and guanine is converted to enol group (-OH) (Fig. 11.15).

NOTES

Due to tautomeric state unusual base pairs are found, and cause mutation.

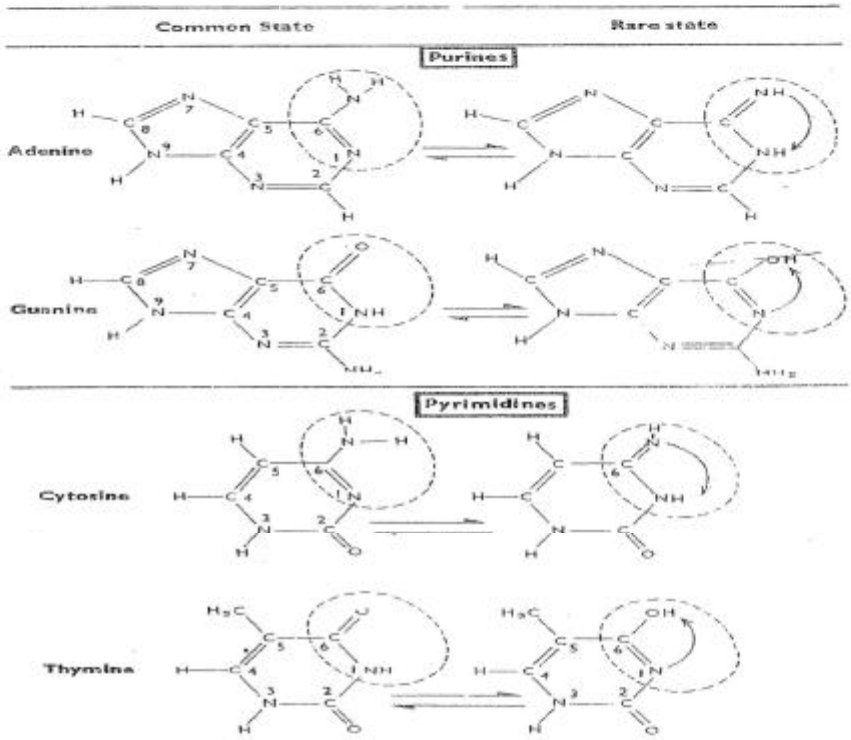


Fig. 11.15 : Common Base of DNA and Their Tautomers

2. Ionization. At the time of DNA replication, ionization of a base may create transitions; Ionization involves the loss of the hydrogen from number-1 nitrogen of a N-base.

For example, in its ionized state, T pairs with normal G and ionized G links with normal thymine (Fig. 11.16).

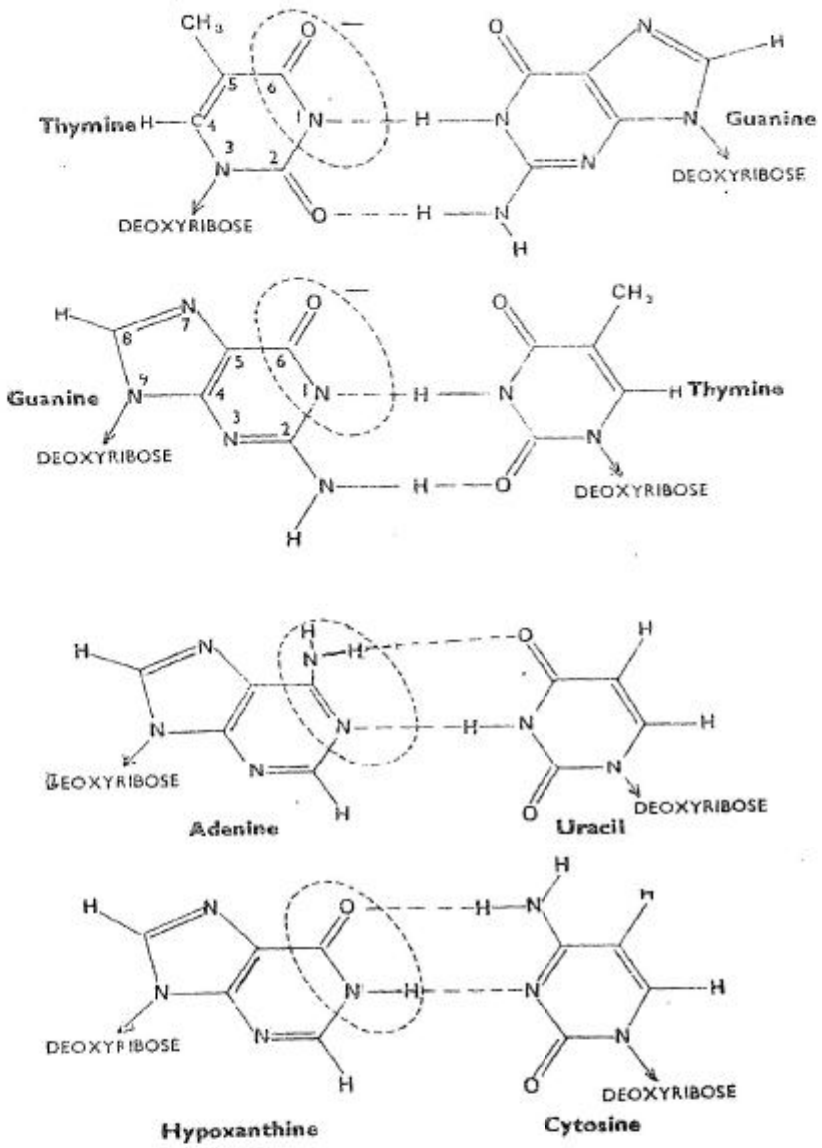


Fig.11.16 : Forbidden Base Pairs of Thymine and Guanine Resulting from the Ionization of no. 1 Nitrogen

(iii) Base analogs. In a normal molecular DNA certain N-base modified through methylation and some time other chemical compounds have molecular structure similar to the nitrogenous base. These are called base analogs. These are derivatives of N-bases of DNA. Some of the natural base analogs are 5-methyl cytosine, 5-hydroxymethyl cytosine and 5-glucosyl hydroxymethyl cytosine, 5-hydroxymethyl uracil and 6-methyl purine. The artificial base analogs are 5-bromo uracil (5-Bu), 5-iodo uracil, 2-bromo and 5-methyl-cytosine (Fig. 11.17).

NOTES

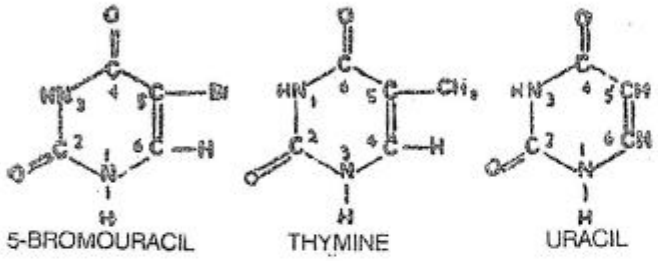


Fig. 11.17 : Bromouracil, Uracil and Thymine Molecules to Show Similarities

(iv) Deamination- Certain chemical substances like nitrous acid and hydroxylamine cause deamination of N-base by replacing amino group ($-NH_2$) by hydroxyl group (-OH). The deamination of cytosine leads to the formation of uracil, deamination of adenine forms hypoxanthine (H) and of guanine forms Xanthine (Fig. 11.18).

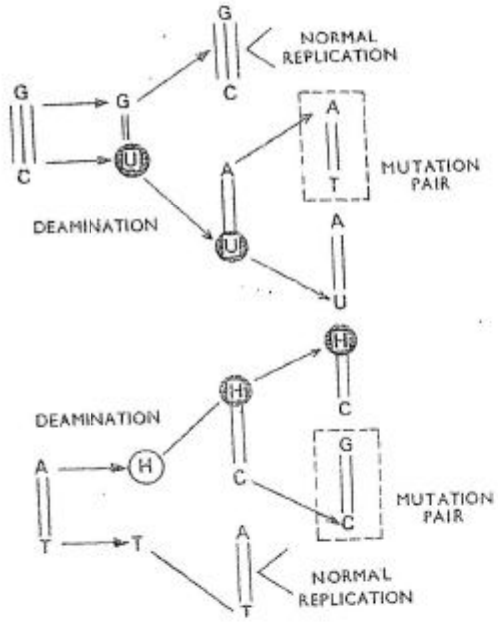


Fig.11.18 : Mechanism of Mutation Caused by Nitrous Acid

At the time of DNA replication uracil pairs with adenine and hypoxanthine pairs with cytosine. This leads to the substitution of A=T for G≡C and G=C for A=T.

2. Transversions. The process of transversions was first seen by E. Freese in 1959. Certain alkylating agents, like ethyl methane sulphate (EMS) and methyl methane sulphate (MMS) induce substitutions. In

transversion purine substituted by pyrimidine or a pyrimidine substituted by purine, i.e. change A=T to C≡G.

These chemicals alkylate the purin N-base in the nitrogen at the seventh position in the guanine and adenine and finally lead to its separation from the DNA sterol. This is known as depurination.

NOTES

Depurination leaves a gap at that point. At the time of replication, any other four bases can possibly get inserted at this place in the complementary strand.

If the nucleotide inserted contains a pyrimidines, it is transition, and if purine then it is transversion (Fig. 11.19).

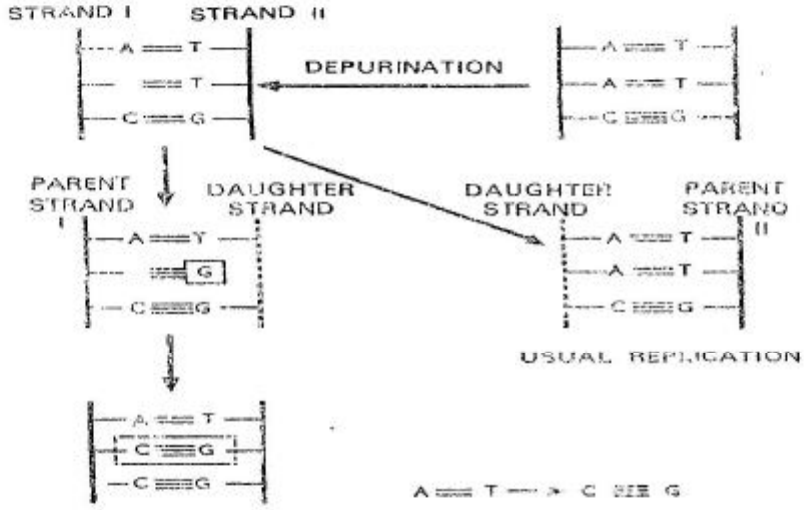


Fig.11.19 : Substitution of A=T by C≡G as a Result of Transversion Frame-Shift Mutations

The insertion or deletion of a nucleotide into a molecule of DNA would shift the reading of genetic message and would alter the codons beyond the point of insertion or deletion.

(i) a. If the normal sequence of nitrogenous bases in DNA is ……T7ACCATTAG……

(b) It will be read (TAC) (CAT) (TAG)……

(ii) a. If a nitrogenous base C is added between CC nitrogenous bases. the sequence becomes ……TACCCATTAG……

(b) It will be read as (TAC) (CCA) (TTA)G……

NOTES

(iii) (a) if a nitrogenous base T is deleted from third codon the sequence becomes TACCATAG……..

(b) It will be read as (TAC) (CAT) AG…….. such mutations in which the entire reading of genetic message is altered beyond the point of change, are known as frame-shift mutations (Fig. 11.20).

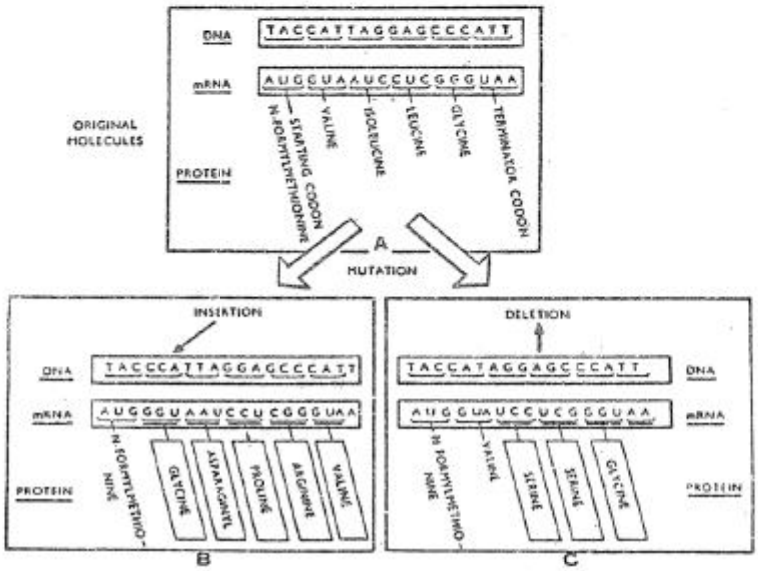


Fig. 11.20 : Effect of Insertion or Deletion on mRNA

Transposons in Prokaryotes and Eukaryotes

Transposons describe additional extra chromosomal elements, which are transposable and can occupy different sites on the main DNA molecules of plasmid or bacterial chromosome.

Transposons in Prokaryotes

Transposon is a term used in 1974 by R.W. Hedges and A.E. Jacob in London, for a DNA segment or genetic element which could more from one molecule to another and carried resistance for antibiotic ampicillin.

They observed that transfer of such antibiotics resistance from one plasmid to another is accompanied by an increase in size of the recipient DNA molecule. This recipient plasmid could donate this resistance to another plasmid, which also showed a similar increase in size thus proving that transfer of DNA segment was involved. This DNA segment called a transposon, can occupy different sites in the genome, can be transposed between two lateral

chromosomes or two plasmid or between a plasmid and a lateral chromosome. It was also shown that transfer of transposon carrying ampicilin resistance could take place not only in plasmid as shown in fig. 11.21, but even in those bacteria which were mutant for the gene *sec A* responsible for recondition, thus suggesting that these transfer could not involve normal recombination process .

NOTES

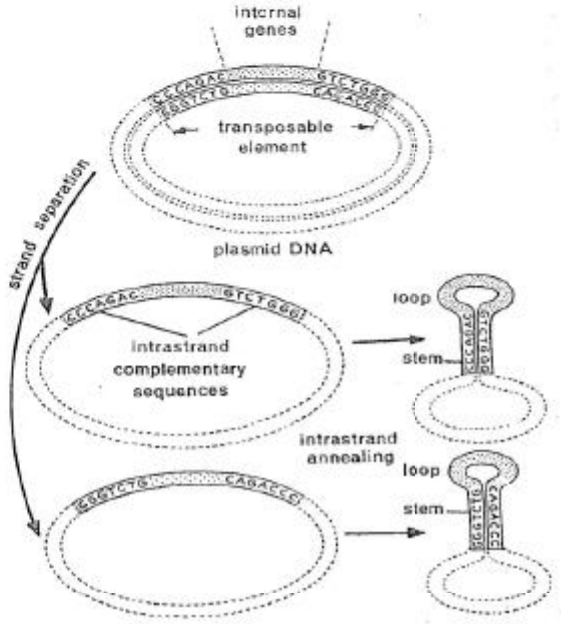


Fig.11.21 : Presence of inverted repeats at the two ends of each strand of a transposon in a plasmid and the formation of characteristic stem and loop in each strand on strand separation due to denaturation. Stem and loop formation is an evidence for inverted repeats.

It was also shown that two ends of each of the two strands in transposon consisted of nucleotide sequences that were complementary to each other, but in reverse order. For instance if one and had CCCAGAC, the other end will have GTCTGGG. These inverted reports will from stem and loop structure characteristic of such repeats when a plasmid with transposon was denatured and each single strand was allowed to base pair among its ownself, such stem and loop structure were actually observed.

Transposons in Eukaryotes

Eukaryotes transposons can be classified into four types. The basis s of their classification include the size and nature of the terminal repeats, which could be long or short and could be direct reports or inverted repeats. Following are the four types of transposons (i) with long terminal direct repeats (eg. copia in

NOTES

Drosophila, Ty in yeast and IAP in nice), (ii) with long terminal inverted repeats (eg. FB, TE in Srosophilia); (iii) with short terminal inverted repeats (eg. P and I in Drosophilia; Ac/Ds in maize; Tam I in snapdragon and Tcl in coenorhabditis) and (iv) without terminal repeats (Alu in mammals). In another classification, on the basis of mechanism of transposition, transposable elements are classified into two classes:

(i) those which transpose by reverse transcription of an RNA intermediate (DNA-RNA-DNA)- class. I. (eg. copia like, Ty, IAP, L1, F) and

(ii) those which transpose directly, DNA to DNA class II (eg. P, Ac, spm/En,/En, Tam, Tcl)

Transposable elements (TE) in drosophila as much as 10% of the genome in drosophila may consist of transposons; most important of these are copia like elements, the fold back (FB) and P (and I) elements. The salient features of those three types are shown in fig. 11.22.

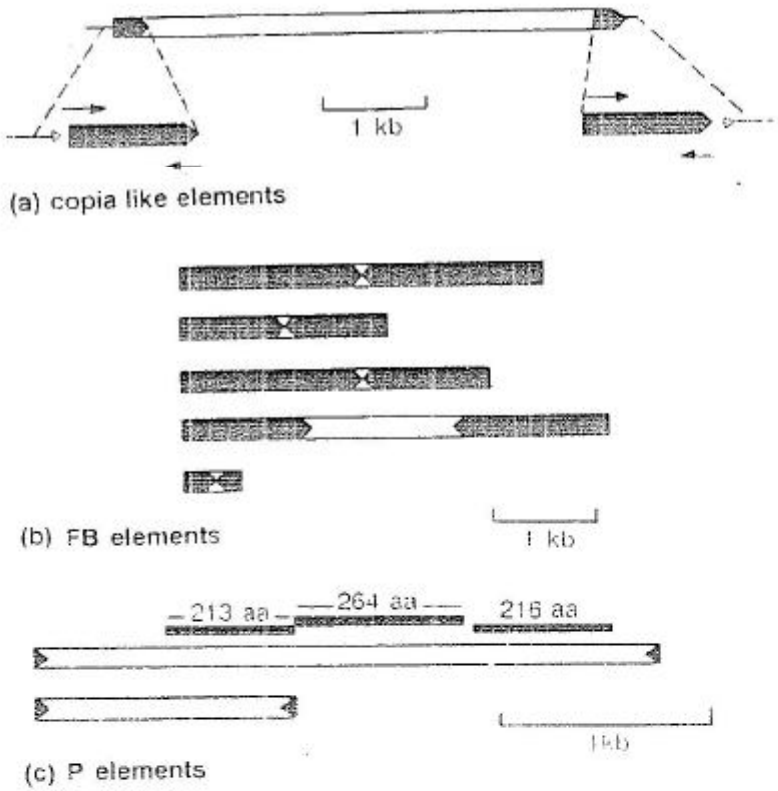


Fig.11.22 : Three classes of Transposable Elements in *Drosophila*

(a) Copia like element- Insertion of those elements cause mutations, a classical example being apricot (W^a) mutation for eye colour caused by insertion of a copia like element.

(b) **FB elements** – These elements range from a few hundred to a few thousand base pairs and carry long inverted repeats at the terminal. They many cause mutations by insertion or by their effect on gene expression.

(c) **P and elements** – The P and I elements are responsible for hybrid digenesis in Drosophila. Hybrid digenesis is a phenomenon responsible of unusual characteristics like sterility, high mutation rate, high frequency of chromosomal aberrations and non disjunction which appears in the progeny, when either.

NOTES

(1) M strain is crossed as female with P strain as male or (II) I (Inducer) strain is crossed as male with R (reactive) strain as female. These hybrid progeny are called dysgenic, hence the name hybrid digenesis (Fig. 11.23).

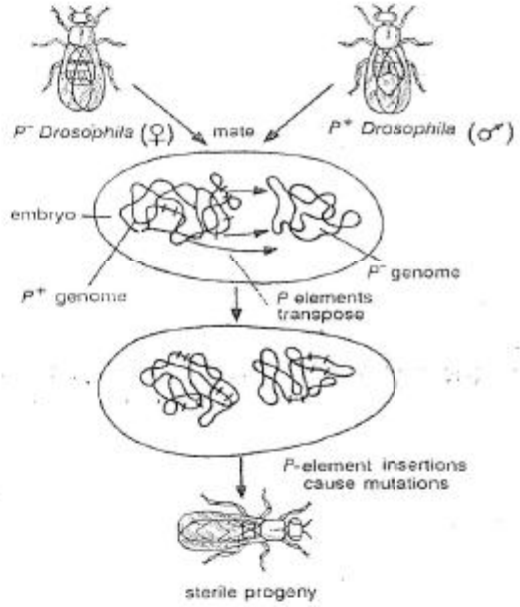


Fig.11.23 : The Phenomenon of Hybrid Dysgenesis (Due to Insertion of P Element) in *Drosophila*

Ty elements in yeast- An example of Ty elemental in yeast is Ty I, which is present in 35 copies in yeast genome. At the ends are present direct repeats called delta (□); of which 100 copies per genome are found. Ty elements generate a repeated sequence (5 by long of the target DNA) during transposition and cause mutations due to insertion.

Controlling elements in corn- The transposable elements although studied initially only in bacterial cells, but were originally discovered in maize plant by

Marcus Rhoades and Barbares Mc Clintock. These genetic elements in maize were found to be responsible turning the expression of genes on or off hence the name 'controlling elements'. It was also shown that these controlling elements when located on maize chromosomes represented specific sites for breakage and reunion of chromosomes leading to gross changes in chromosome structure.

NOTES

Mechanism of Transposition- It was been shown that transposons have genes coding for an enzyme called transposas which observes off DNA segment of transposon by recognizing the reverse repeats (Fig. 11.24).

A suggested pathway involving five steps to explain transposition .

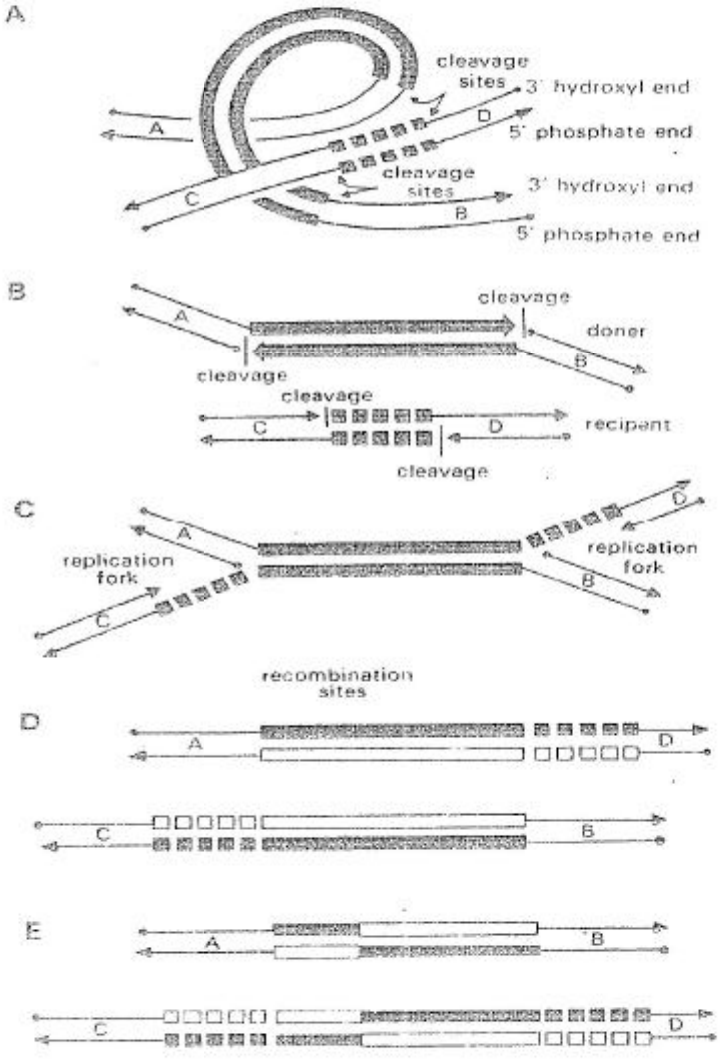


Fig.11.24 : A suggested pathway involving five steps (A, B, C, D, E) to explain transposition.

The donor DNA includes a transposon, and cleavage takes place at the ends of transposon sequence in two strands of donor DNA and also at the two ends of target nucleotide sequence in two strands of recipient DNA. Both the transposon and target sequence are duplicated which is followed by reciprocal recombination.

The insertion of a transposable genetic element in one orientation will switch off a specific gene and switch on another gene. A reverse effect may observed due to insertion in a different orientation. It is in this respect that transposons and other transposable genetic elements in bacteria and higher organism control gene regulation.

Uses of Transposons – Transposons may be used as genetic markers because they change the pattern of restriction fragment analysis, for instance with the help of probe it was possible to characterize different strains of *Plasmodium falciparum* the causative agent of human malaria.

Transposons, which change the pattern of restriction fragment analysis, may also be used as genetic markers to construct linkage maps.

These have also been used in humans for distinguishing carriers from non-carriers of disease like sickle cell trait.

Transposons as mutagens- Insertion of transposons can be used as a method for inducing mutations as has been shown in a number of spontaneous mutations like Ac-Ds system in maize, and P-M and I-R systems of hybrid dysgenesis in *Drosophila melanogaster*. Transposons usually cause mutations due to insertion in structural or regulatory region, rather than due to addition, deletion or substitution of bases. Therefore, these mutations can be used for a study of structural and regulatory regions of a gene. Further, a probe carrying a transposon can be used to screen the restriction fragments containing the mutant gene.

11.7 DNA Damage and Repair

Both in Prokaryotes and Eukaryotes these are repair enzymes system to deal with DNA damage. The changes in DNA damage are broadly divided into two classes.

NOTES

(i) Single base changes- May be caused by conversion of one base to another. These are corrected through DNA replication leading to changes in DNA sequences.

(ii) Structural distortions – May result from a single strand nick (thymine dimer).

A cell may have several system to deal with DNA damage.

(i) Direct repair

(ii) Excision repair

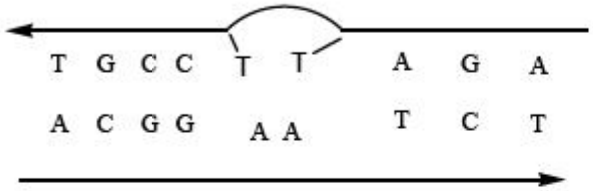
(iii) Mismatch repair

(i) Direct repair or photo reactivation- It involves an enzymes that splits thymine dimers directly without the removal of any nucleotides.

- Enzymes bind to T.D. in dark.
- The enzyme splits the cynobutane ring only on activation by light.
- It can illustrate by this diagram.

(1) Segment of DNA contains two adjacent thymine both.

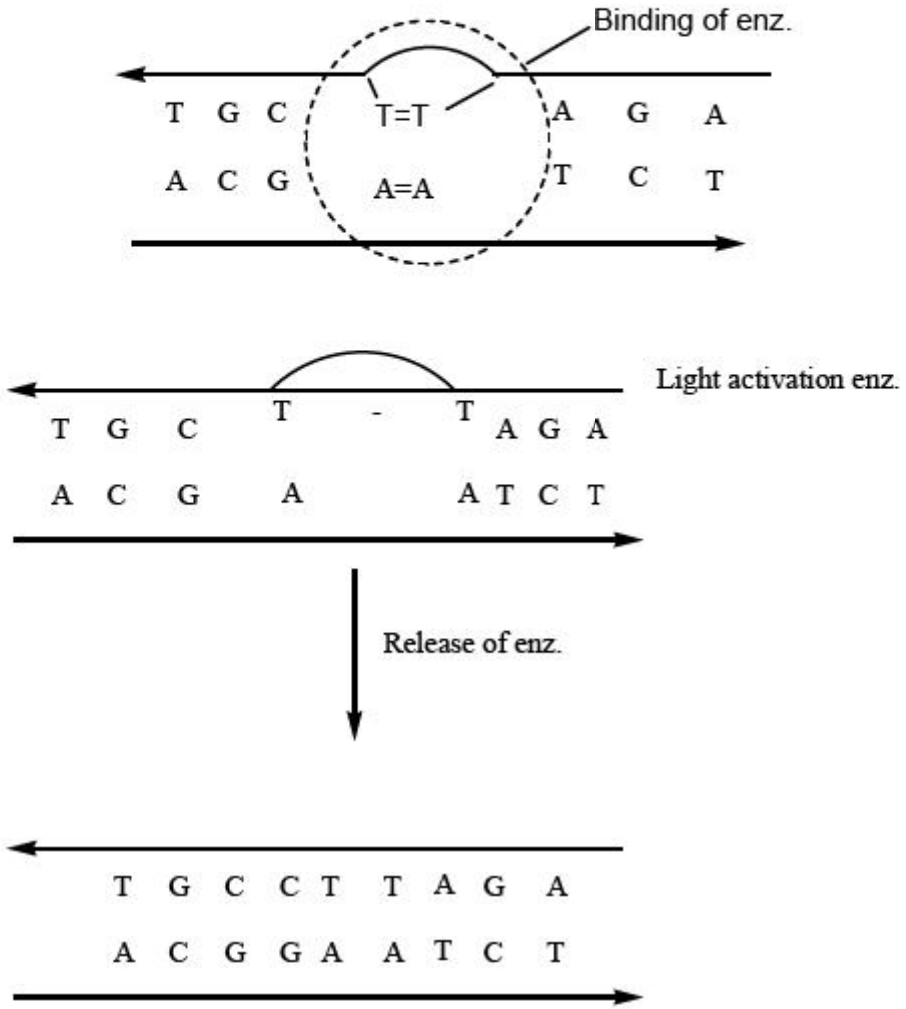
(2) Thymine dimer formation as a result of UV irradiation.



(3) Binding of photo activating enzyme to thymine dimer containing segment of the DNA molecule.

(4) Cleavage of the dimer cross links by the photo reactivating enzymes causing energy from the absorption of blue light.

(5) Release of the photo activating enzymes leaving a repaired normal DNA.



NOTES

(2) Excision repair- Classified into two

(1) BER (Base excision repair)- Initiated by DNA glycolyas.

(2) NER (Nucleotide excision repair)- initiated by endonucleolytic excision, either only one side of the lesion endo, exo.

BER in three steps-

(1) Inversion- Endonuclease cleavage the DNA on one or both sides of damage.

(2) Excision- a 5'-3' exonuclease DNA poly I removes a structure of the damaged strand.

(3) **Synthesis step-** The single strand of region of DNA is used as a template for synthesis of DNA by DNA poly–I.

In BER- Causing agent.

NOTES

$r \longrightarrow n_3$- meAde

$\longrightarrow$ 8 oxo guanine

Uracil, thymine glycols

Thymine hydracts.

- **Excision repair as above involves different length of DNA is described as-**
  (1) Very short patch pair
  (2) Short patch pair
  (3) Long patch pair
  (4) Very long patch pair.
- Excision repair involves a sequence of enzymes Catalyzed step in which the thymine dimer are removed from the DNA molecules and a new segment of DNA is synthesize.
- It occurs in the dart.

(1) Segment of DNA containing two adjacent thymine bases.

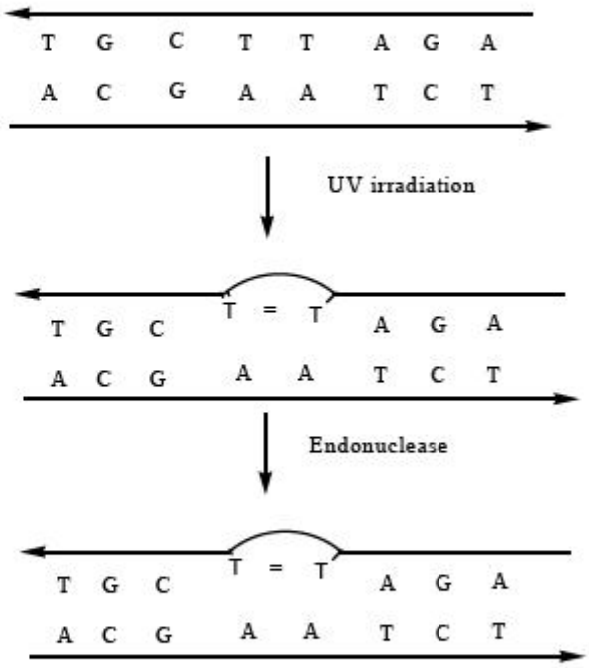
(2) T.D. as a result of irradiation.

(3) Endonuclease recognize the T.D. cleave the phosphodiester backbone of the DNA T.D.

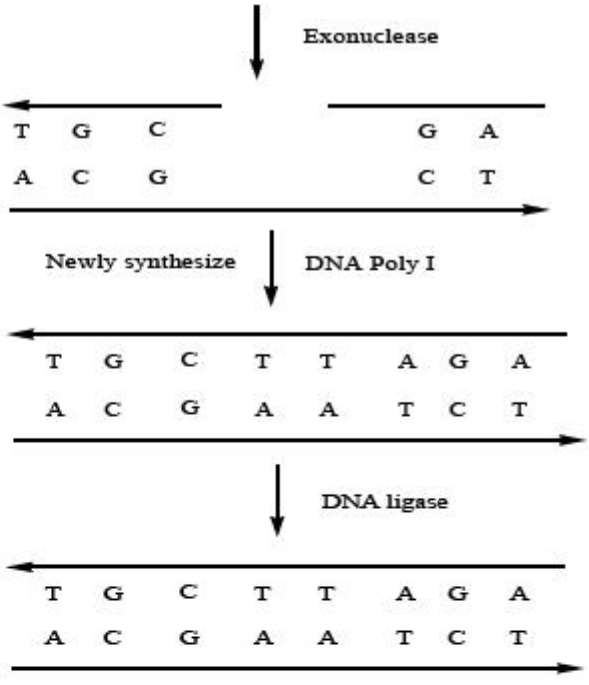
(4) Exonuclease removed of a nucleotide including the thymine dimer.

(5) DNA pol. I fill the gap using the complementary stand as or we can say Resynthesis of excised segment by DNA poly – I.

(6) Formation of final phosphodiester bond by DNA ligase.



(Thymine Dimer) T.D.

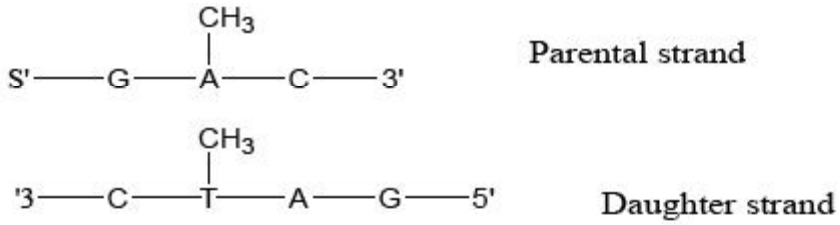


NOTES

NOTES

(3) Mismatch repair- Many spontaneous mutations are point mutations, which involve a change in a single base pair in the DNA sequence. Bacterial end Eukaryotic cells have a mismatch repair system that recognizes and repairs all single base mispairs except C°C, as well as small insertion and deletions.

In *E. coli* DNA adenine residues in a GATC sequence are methylated at the 6 position. Since DNA polymerases incorporate adenine, not methyl – adenine, into DNA, adenine residues in newly replicated DNA are methylated only on the parental strand. The adenines in GATC sequence on the daughter strands are methylated by a specific enzyme, called **Dam methyl transferase** only after a lag of several minutes. During this lag period, the newly replicated DNA contains hemimethylated GATC sequence.



An *E. coli* protein designated Muth, which binds specifically to hemimethylated sequence, is able to distinguish the methylated parental strand from the unmethylated daughter strand. If an error occurs during DNA replication, resulting in a mismatched base pair near a GATC sequence, another protein, Muts binds to this abnormally paired segment (Fig. 11.25).

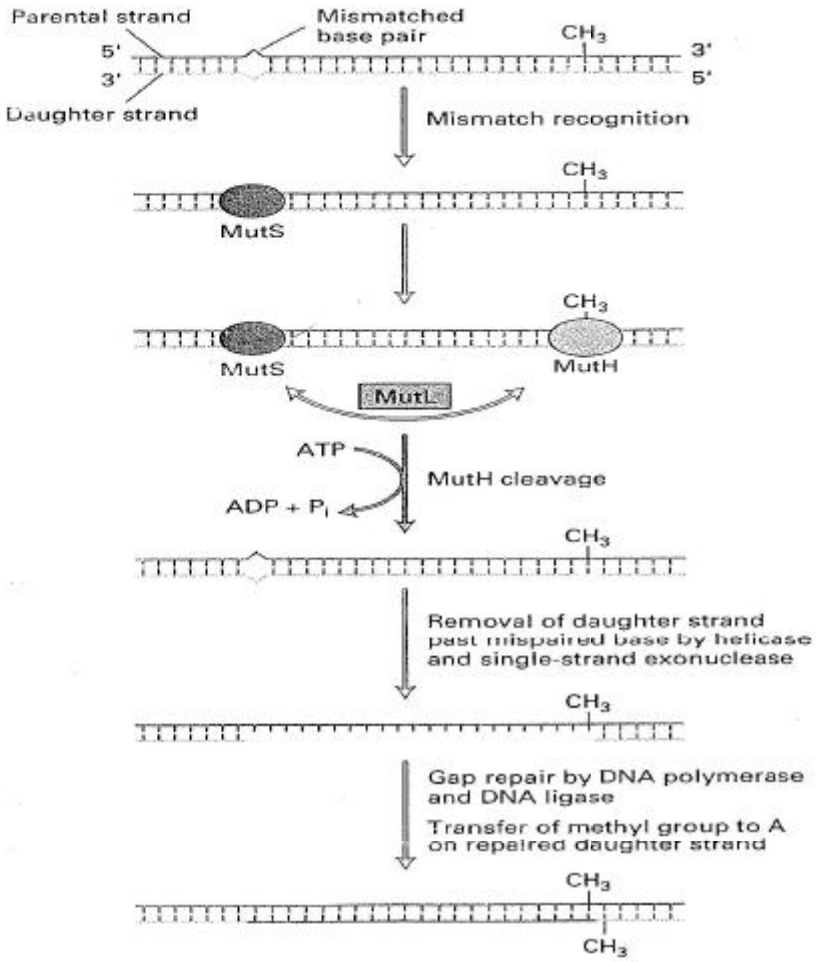


Fig.11.25 : Model of Mismatch Repair by the *E. Coli* MutHLS system.

Binding of MutS triggers binding of Muth, a linking protein's that connects MutS with a nearly MutH. This cross linking activates a lateral endonuclease activity of MutH, which then cleaves specifically the ummethylated daughter strand. Following this initial incirion, the segment of the daughter strand containing the misincorporated base is excised and replaced with the correct DNA sequence.

E. coli strains that back the MutS, MutH, or MutL protein have a higher rate of spontaneous mutations than wild type cells.

11.8 Summary

In this unit we emphasize the role of linkage and crossing over in genetic recombination and genetic mapping, recombination covers a much wider

NOTES

spectrum of events like reciprocal exchange of segments, gene conversion exchange between sisters chromatids repair of DNA damage. Mutation in a broad sense includes all those heritable changes, which after the phenotype of an individual. Mutations at molecular level, should mean permanent alterations in sequence of nucleotides in the nucleic acid, which forms the genetic material.

The changes in DNA strand leading to damage of DNA. A cell may have several system to deal with DNA damage.

11.9 Glossary

- **Linkage group:** All of the genes linked together on the particular chromosome.
- **Recombination:** Exchange of homologous chromosomes formed a new gene combination.
- **Crossing over:** Process by which chromatids of homologous exchange positions of genetic material.
- **Synapsis:** Pairing of homologous chromosomes at Zygotene substage.
- **Genetic mapping:** Allocation of gene on the chromosome.
- **Mutation:** A permanent, heritable change involving a gene or chromosome.
- **Mutagens:** Factors capable of causing mutation.
- **Deletion:** A chromosomal abnormality in which a piece of the chromosome is missing.

11.10 Self - Learning Exercise

Section –A (Very Short Answer Type Questions)

1. General recombination was studied by …….
2. Crossing over occurs at stage ……
3. Linkage is discovered by ……..
4. Define mutation.
5. Define linkage.
6. What is autosome?

7. What is transposable element?
8. Name the plant which made Hugo-devries famous.

Section –B (Short Answer Type Questions)

1. Write any two chemical mutagens.
2. What is mutagenesis?
3. Define induced mutation
4. What is spontaneous mutation?
5. Write any two differences in crossing over and linkage.
6. Write any two physical mutagens.

Section –C (Long Answer Type Questions)

1. What is crossing over and explain the significance of crossing.
2. Give an account of molecular basis of mutation.
3. Write a note on transpoble elements.
4. Give an account of different known transposons in eukaryotes. How do they differ from tansposon in prokaryotes?

Answer of section A

1. Holiday
2. Four strand stage
3. Bateson

11.11 References

- Rastogi, V.B. (1991-92) : A Text Book of Genetics
- Gupta, P.K. (2000) : A Reference Book of Genetics
- Lodish and Berk (2004): Molecular Cell Biology
- Gardner/Simmons/Snusted - Principle of genetics sixth edition.

NOTES

Unit - 12

Gene Structure and Expression

NOTES

Structure of the Unit

12.0 Objectives

The Objectives of this unit is to explain the subject to student in very simple way. After studied this unit student found the knowledge about:

- Fine structure of gene,
- Regulation of gene expression in prokaryotes and eukaryotes Catabolites repression.

12.1 Introduction

In gene expression process information from a gene is used in the synthesis of functional gene product. These products are generally proteins but in non protein coding gene such as ribosomal RNA (rRNA) gene or tRNA gene the product is functional RNA.

NOTES

The process of gene expression is used by all eukaryote, prokaryotes and virus to generate the macromolecular machinery for life. There are several steps in the gene expression including the transcription, RNA splicing translation and post-translation modification of protein .By the gene regulation the cell control over structure and function .It is the basis for cellular differentiation, morphogenesis and the versatility and adaptability for any organism.

12.2 Fine Structure of Gene

In molecular terms, a gene commonly is defined as the entire nucleic acid sequence that is necessary for the synthesis of a functional gene product (polypeptide or RNA). According to this definition, a gene includes more than the nucleotides encoding the amino acid sequence of a protein, referred to as the coding region. A gene also includes all the DNA sequences required for synthesis of a particular RNA transcript. In eukaryotic genes, transcription-control regions known as **enhancers** can lie 50 kb or more from the coding region. Other critical noncoding regions in eukaryotic genes are the sequences that specify 3' cleavage and polyadenylation, known as poly (A) sites, and splicing of primary RNA transcripts, known as splice sites . Mutations in these RNA-processing signals prevent expression of a functional mRNA and thus of the encoded polypeptide. Although most genes are transcribed into mRNAs, which encode proteins, clearly some DNA sequences are transcribed into RNAs that do not encode proteins (e.g., tRNA sand rRNAs). These DNA regions generally are referred to a stRNA and rRNA genes, even though the final products of these genes are RNA molecules and not proteins.

Eukaryotic gene includes additional sequences that lie within the coding region, interrupting the sequence that represents the protein. The sequences of DNA comprising an interrupted gene are divided into the two categories:

NOTES

1. The **exons** are the sequences represented in the mature RNA. By definition, a gene starts and ends with exons, corresponding to the 5' and 3' ends of the RNA.
2. The **introns** are the intervening sequences that are removed when the primary transcript is processed to give the mature RNA (Fig. 12.1).

Prokaryotic genes lack introns. Bacterial mRNAs are *polycistronic;* that is, a single mRNA molecule includes the coding region for several proteins that function together in a biological process. In contrast, most eukaryotic mRNAs are *monocistronic;* that is, each mRNA molecule encodes a single protein. This difference between polycistronic and monocistronic mRNAs correlates with a fundamental difference in their translation. Within a bacterial polycistronic mRNA a ribosome binding site is located near the start site for each of the protein coding regions, or cistrons, in the mRNA. Translation initiation can begin at any of these multiple internal sites, producing multiple proteins. In most eukaryotic mRNAs, however, the 5'-cap structure directs ribosome binding, and translation begins at the closest AUG start codon. As a result, translation begins only at this site. In many cases, the primary transcripts of eukaryotic protein-coding genes are processed into a single type of mRNA, which is translated to give a single type of polypeptide. Unlike bacterial and yeast genes, which generally lack introns, most genes in multi cellular animals and plants contain introns, which are removed during RNA processing. In many cases, the introns in a gene are considerably longer than the exons.

The cluster of genes that form a bacterial operon comprises a single **transcription unit,** which is transcribed from a particular promoter into a single primary transcript. In other words, genes and transcription units often are distinguishable in prokaryotes. In contrast, most eukaryotic genes and transcription units generally are identical, and the two terms commonly are used interchangeably. Eukaryotic transcription units, however, are classified into two types, depending on the fate of the primary transcript. The primary transcript produced from a *simple* transcription unit is processed to yield a single type of mRNA, encoding a single protein. Mutations in exons, introns, and transcription-control regions all may influence expression of the protein encoded by a simple transcription unit. Quite common in multicellular

organisms, the primary RNA transcript can be processed in more than one way, leading to formation of mRNAs containing different exons. Each mRNA, however, is monocistronic, being translated into a single polypeptide, with translation usually initiating at the first AUG in the mRNA. Multiple mRNAs can arise from a primary transcript by using of different splice sites or alternative poly (A) sites or alternative promoters.

NOTES

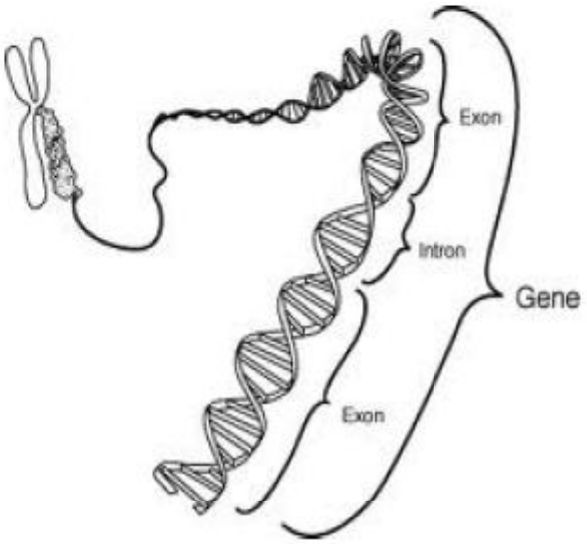


Fig. 12.1 : Fine Structure of Gene

12.3 *Cis*-trans test

Definition

A method of genetic analysis that determines whether recessive mutation occur in a single gene or in different genes.

- The test was devised by the American geneticist E. Lewis in 1951.
- Cis trans test also known as complementation test with two or more interacting gene placed in cis and in trans relationship to each other.
- A double mutant genome is used in the cis test made from the two single mutant genome used in the trans test by recombination.
- This test to determine whether two mutant sites are in the same functional unit or gene.

- This test can be done with bacteria yeast and other organism that are homogygous for the two mutations.
- If the combination of two genomes containing different recessive mutation yields a mutant phenotype, then there are three possibilities.
- (1)Mutation occur in the same gene .(ii) One mutation affects the expression of the others.(iii) One mutation may result in an inhibitory product.

NOTES

An explanation for cis-trans test was given by E.B. Lewis. According to him if both mutant are on one chromosome and other chromosome will be normal and will be normal and will be able to produce the end produce this type of situation occurs of steps involved in synthesis will be interrupted due to mutation on either of the two homologous chromosomes thus leading to a mutant phenotype (Fig. 12.2).

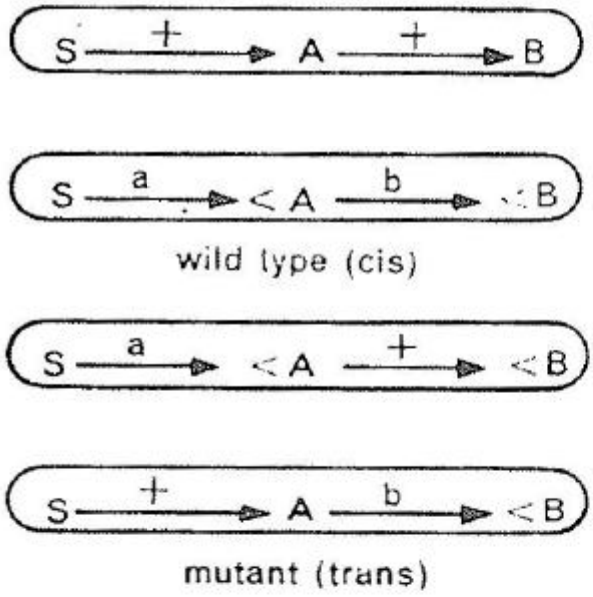


Fig.12.2 : Synthesis of Polypeptides in *cis* and *trans* Configurations

12.4 Intron and their Significance

The term intron refers to both the DNA sequence within a gene and the corresponding sequence in RNA transcripts. An **intron** is any nucleotide sequence within a gene . Introns are common in the genes of most organisms and many viruses, and can be located in including those gene that generate

proteins, ribosomal RNA (rRNA), and transfer RNA (tRNA). When proteins are generated from intron-containing genes, RNA splicing takes place as part of the RNA processing pathway that follows transcription and precedes translation.

NOTES

The word *intron* is derived from the term *intragenic region*, i.e. a region inside a gene. Although introns are sometimes called *intervening sequences*, the term "intervening sequence" can refer to any of several families of internal nucleic acid sequences that are not present in the final gene product, including inteins, untranslated sequences (UTR), and nucleotides removed by RNA editing, in addition to introns.

Discovery: Introns were first discovered in protein-coding genes of adenovirusand were subsequently identified in genes encoding transfer RNA and ribosomal RNA genes. Introns are now known to occur within a wide variety of genes throughout organisms and viruses within all of the biological kingdoms (Fig. 12.3).

The fact that genes were split or interrupted by introns was discovered independently in 1977 by Phillip Allen Sharp andRichard J. Roberts, for which they shared the Nobel Prize in Physiology or Medicine in 1993. The term *intron* was introduced by American biochemist Walter Gilbert.Introns are extremely common within the nuclear genome of higher vertebrates (e.g. humans and mice), where protein-coding genes contain multiple introns.

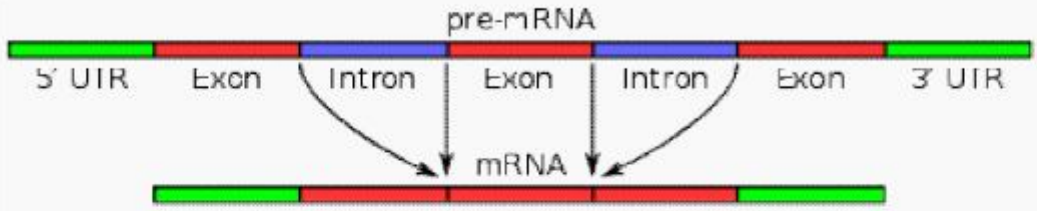


Fig. 12.3 : Simple illustration of an unspliced mRNA precursor, with two introns and three exons (top)

12.4.1 Classification

Splicing of all intron-containing RNA molecules. Following four distinct classes of introns have been identified

NOTES

1. Introns in nuclear protein-coding genes which is removed by spliceosomes.
2. Introns in nuclear and archaeal transfer RNA genes that are removed by tRNA splicing enzymes.
3. Self-splicing group I introns that are removed by RNA catalysis.
4. Self-splicing group II introns that are also removed by RNA catalysis

Group I and group II introns are found in genes encoding proteins (messenger RNA), transfer RNA and ribosomal RNA in a very wide range of living organisms. Group I and group II introns are isolated by different sets of internal conserved sequences and folded structures, and by the fact that splicing of RNA molecules containing group II introns generates branched introns (like those of spliceosomal RNAs), while group I introns use a non-encoded guanosine nucleotide (typically GTP) to initiate splicing, adding it on to the 5'-end of the excised intron.

Transfer RNA introns that depend upon proteins for removal occur at a specific location within the anticodon loop of unspliced tRNA precursors, and are removed by a tRNA splicing endonuclease. The exons are then linked together by a second protein, the tRNA splicing ligase. Note that self-splicing introns are also sometimes found within tRNA genes.

12.4.2 Significance

Introns might have explored new genetic space and acquired functions which provided a positive pressure for their expansions.

Introns contain several short sequences that are important for efficient splicing, such as acceptor and donor sites at either end of the intron as well as a branch point site, which are required for proper splicing by the spliceosome. Some introns are known to enhance the expression of the gene that they are contained in by a process known as intron-mediated enhancement (IME).

12.5 RNA Splicing

Splicing is a modification of the nascent pre-messenger RNA (pre-mRNA) transcript in which introns are removed and exons are joined. Splicing is needed for the typical eukaryotic messenger RNA (mRNA) before it can be used to

produce a correct protein through translation. For many eukaryotic introns, splicing is done in a series of reactions which are catalyzed by the spliceosome, a complex of small nuclear ribonucleoproteins (snRNPs) (Fig. 12.4).

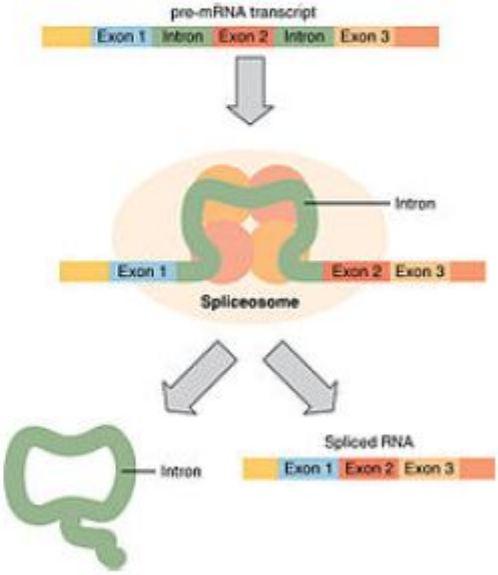


Fig. 12.4 : Exons and Introns in pre-mRNA and the formation of mature mRNA by splicing Introns

The *intron* is *intragenic region*, that is, a region inside a gene. It refers to both the DNA sequence within a gene and the corresponding sequence in the unprocessed RNA transcript. Introns are common in genes of most organisms and many viruses. Spliceosomal introns often reside within the sequence of eukaryotic protein-coding genes. Within the intron, a donor site (5' end of the intron), a branch site (near the 3' end of the intron) and an acceptor site (3' end of the intron) are required for splicing. The splice donor site includes an almost invariant sequence GU at the 5' end of the intron, within a larger, less highly conserved region. The splice acceptor site at the 3' end of the intron terminates the intron with an almost invariant AG sequence. Upstream (5'-ward) from the AG there is a region high in pyrimidines (C and U), or polypyrimidine tract (Fig. 12.5).

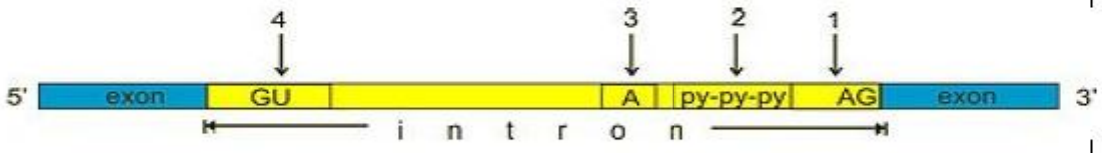


Fig. 12.5 : Intron

NOTES

Formation and activity

Splicing is catalyzed by the spliceosome, which is a large RNA-protein complex composed of five small nuclear ribonucleo proteins (snRNPs, pronounced 'snurps'). The RNA components of snRNPs interact with the intron and may be involved in catalysis. Two types of spliceosomes have been identified (the major and minor) which contain different snRNPs.

- **Major**

The major spliceosome splices introns containing GU at the 5' splice site and AG at the 3' splice site. It is composed of the U1, U2, U4, U5, and U6 snRNPs and is active in the nucleus. In addition, a number of proteins including U2AF and SF1 are required for the assembly of the spliceosome. E Complex-U1 binds to the GU sequence at the 5' splice site, along with accessory proteins/enzymes ASF/SF2, U2AF (binds at the Py-AG site), SF1/BBP (BBP=Branch Binding Protein);

- **Minor**

The minor spliceosome is very similar to the major spliceosome, however it splices out rare introns with different splice site sequences. While the minor and major spliceosomes contain the same U5 snRNP, the minor spliceosome has different, but functionally analogous snRNPs for U1, U2, U4, and U6, which are respectively called U11, U12, U4atac, and U6atac.

Trans-splicing is a form of splicing that joins two exons that are not within the same RNA transcript.

(a) Self-splicing

Self-splicing occurs for rare introns that form a ribozyme, performing the functions of the spliceosome by RNA alone. There are three kinds of self-splicing introns, *Group I*, *Group II* and *Group III*. Group I and II introns perform splicing similar to the spliceosome without requiring any protein. This similarity suggests that Group I and II introns may be evolutionarily related to the spliceosome. Self-splicing may also be very ancient, and may have existed in an RNA world present before protein.

(b) tRNA splicing

tRNA (also tRNA-like) splicing is another rare form of splicing that usually occurs in tRNA. The splicing reaction involves a different biochemistry than the spliceosomal and self-splicing pathways. Ribonucleases cleave the RNA and ligases join the exons together.

NOTES

Biochemical mechanism

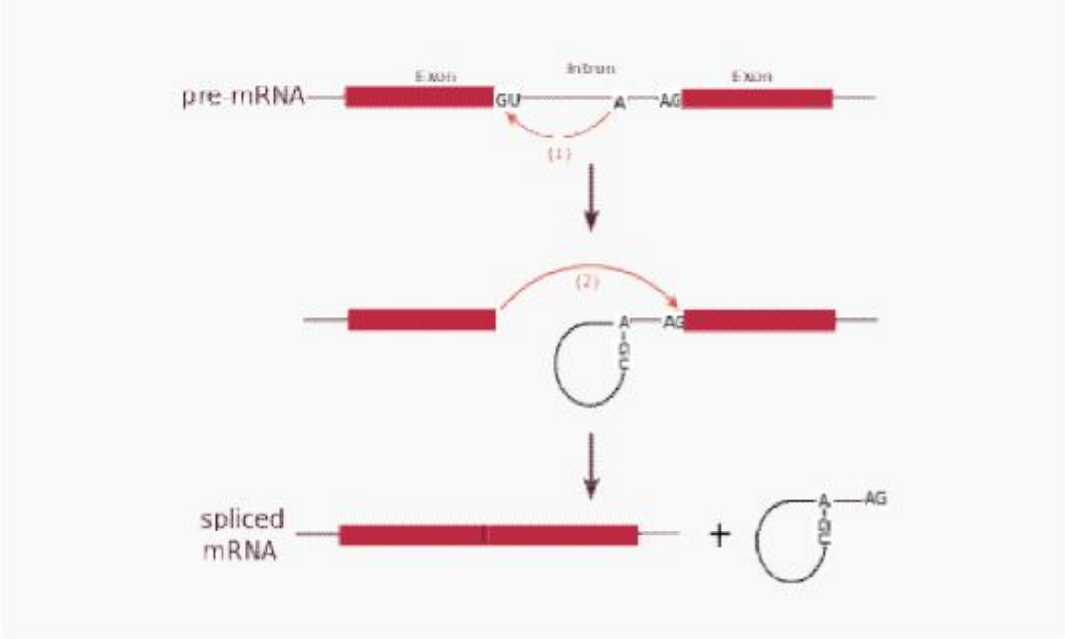


Fig. 12.6 : The Two-Step Biochemistry of Splicing

Spliceosomal splicing and self-splicing involves a two-step biochemical process. Both steps involve transester-ification reactions that occur between RNA nucleotides. tRNA splicing, however, is an exception and does not occur by transesterification.

Spliceosomal and self-splicing transesterification reactions occur in two sequential transesterification reactions. In first, the 2'OH of a specific branch point nucleotide within the intron that is defined during spliceosome assembly performs a nucleophilic attackon the first nucleotide of the intron at the 5' splice site forming the *lariat intermediate*. Second, the 3'OH of the released 5' exon then performs a nucleophilic attack at the last nucleotide of the intron at the 3' splice site thus joining the exons and releasing the intron lariat (Fig. 12.6).

12.6 Regulation of Gene Expression

12.6.1 Gene Expression in Prokaryotes (Operon Concept – lac and Trp operon)

NOTES

The first step in gene expression is **transcription**, the synthesis of an mRNA copy of the DNA template that encodes a protein. Transcription is followed by **translation**, the synthesis of the protein on the ribosome. Transcription is controlled by proteins that bind DNA, and these DNA binding proteins are themselves subject to various types of regulation. Gene regulation in eukaryotes is far more complex than in prokaryotes. The added complexity of gene expression in eukaryotes is what allows cells and tissues to differentiate and makes possible the diverse life cycles of plants and animals. DNA-Binding Proteins Regulate Transcription in Prokaryotes. In prokaryotes, genes are arranged in **operons**, sets of contiguous genes that include **structural genes** and **regulatory sequences**. A famous example is the *E. coli* lactose (*lac*) operon, which was first described in 1961 by François Jacob and Jacques Monod of the Pasteur Institute in Paris. The *lac* operon is an example of an **inducible** operon-that is, one in which a key metabolic intermediate induces the transcription of the gene. The *lac* operon is responsible for the production of three proteins involved in utilization of the disaccharide lactose. This operon consists of three structural genes and three regulatory sequences. The structural genes (*z*, *y*, and *a*) code for the sequence of amino acids in three proteins: β-galactosidase, the enzyme that catalyzes the hydrolysis of lactose to glucose and galactose ; permease, a carrier protein for the membrane transport of lactose into the cell; and transacetylase, the significance of which is unknown. The three regulatory sequences (*i*, *p*, and *o*) control the transcription of mRNA for the synthesis of these proteins .Gene *i* is responsible for the synthesis of a **repressor protein** that recognizes and binds to a specific nucleotide sequence, the **operator**. The operator, *o*, is located downstream of the **promoter** sequence, *p*, where RNA polymerase attaches to the operon to initiate transcription, and immediately upstream of the transcription start site, where

transcription begins. In the absence of lactose, the lactose repressor forms a tight complex with the operator sequence and blocks the interaction of RNA polymerase with the transcription start site, effectively preventing transcription. When present, lactose binds to the repressor, causing it to undergo a conformational change. The *lac* repressor is thus an allosteric protein whose conformation is determined by the presence or absence of an **effector** molecule, in this case lactose. As a result of the conformational change due to binding lactose, the *lac* repressor detaches from the operator (Fig. 12.7). When the operator sequence is unobstructed, the RNA polymerase can move along the DNA, synthesizing a continuous mRNA. The translation of this mRNA yields the three proteins, and lactose is said to induce their synthesis. The *lac* repressor is an example of **negative control**, since the repressor blocks transcription upon binding to the operator region of the operon. The *lac* operon is also regulated by **positive control**, which was discovered in connection with a phenomenon called the *glucose effect*. If glucose is added to a nutrient medium that includes lactose, the *E. coli* cells metabolize the glucose and ignore the lactose. Glucose suppresses expression of the *lac* operon and prevents synthesis of the enzymes needed to degrade lactose. Glucose exerts this effect by lowering the cellular concentration of cyclic AMP (cAMP). When glucose levels are low, cAMP levels are high. Cyclic AMP binds to an **activator protein**, the *catabolite activator protein* (CAP), which recognizes and binds to a specific nucleotide sequence immediately upstream of the *lac* operator and promoter sites (Fig. 12.7). In contrast to the behavior of the lactose repressor protein, when the CAP is complexed with its effector, cAMP, its affinity for its DNA-binding site is dramatically *increased* (hence the reference to positive control). The ternary complex formed by CAP, cAMP, and the lactose operon DNA sequences induces bending of the DNA, which activates transcription of the lactose operon structural genes by increasing the affinity of RNA polymerase for the neighboring promoter site. Bacteria synthesize cyclic AMP when they exhaust the glucose in their growth medium. The lactose operon genes are thus under opposing regulation by the absence of

NOTES

glucose (high levels of cyclic AMP) and the presence of lactose, since glucose is a catabolite of lactose. In bacteria, metabolites can also serve as *corepressors*, activating a repressor protein that blocks transcription. Repression of enzyme synthesis is often involved in the regulation of biosynthetic pathways in which one or more enzymes are synthesized only if the end product of the pathway—an amino acid, for example—is not available. In such a case the amino acid acts as a corepressor: It complexes with the repressor protein, and this complex attaches to the operator DNA, preventing transcription. The tryptophan (*trp*) operon in *E. coli* is an example of an operon that works by corepression (Fig. 12.8).

NOTES

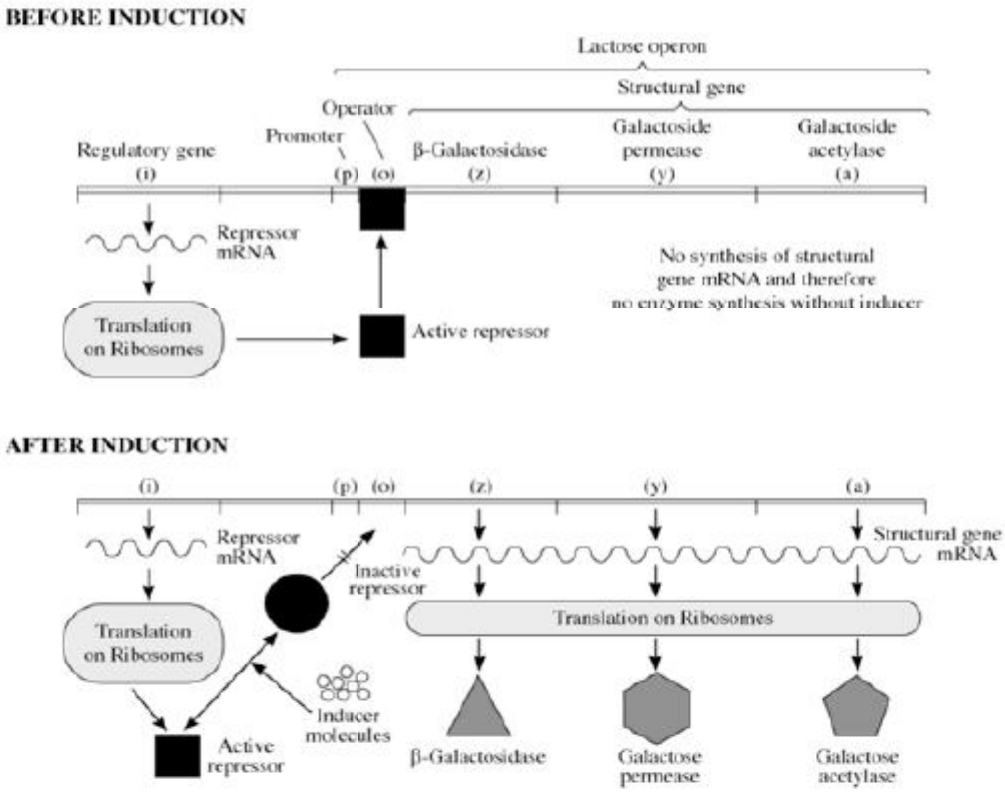


Fig.12.7 : lac operon

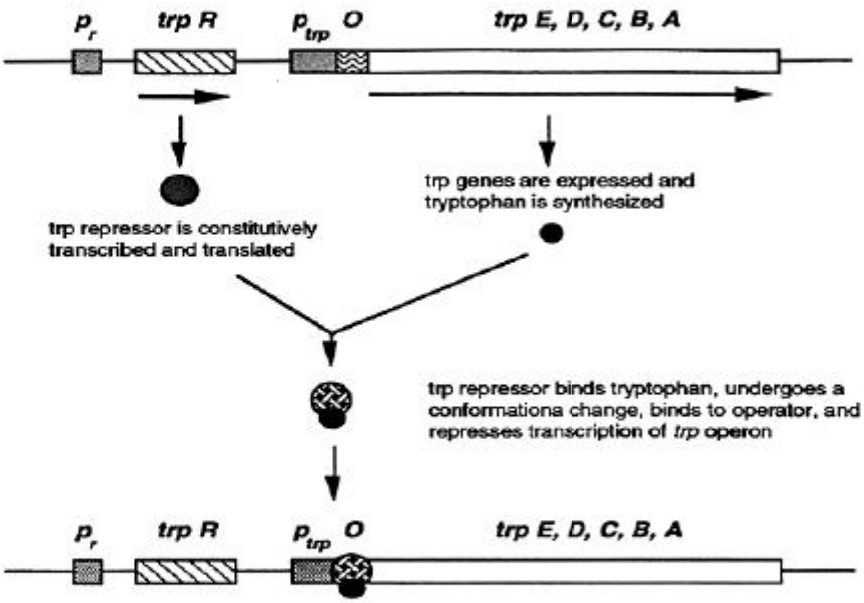


Fig. 12.8 : Trp Operon

12.6.2 Regulation of Gene Expression in Eukaryotes

The study of bacterial gene expression has provided models that can be tested in eukaryotes. However, the details of the process are quite different and more complex in eukaryotes. In prokaryotes, translation is coupled to transcription: As the mRNA transcripts elongate, they bind to ribosomes and begin synthesizing proteins (translation). In eukaryotes, however, the nuclear envelope separates the genome from the translational machinery. The transcripts must first be transported to the cytoplasm, adding another level of control. Eukaryotic nuclear transcripts require extensive processing. Eukaryotes differ from prokaryotes also in the organization of their genomes. In most eukaryotic organisms, each gene encodes a single polypeptide. The eukaryotic nuclear genome contains no operons. Furthermore, eukaryotic genes are divided into coding regions called **exons** and noncoding regions called **introns**. Since the primary transcript,or pre-mRNA, contains both exon and intron sequences, the pre-mRNA must be processed to remove the introns. RNA processing involves multiple steps. The newly synthesized pre-mRNA is immediately packaged into a string of small protein-containing particles, called **heteronuclear ribonucleoprotein particles**, or **hnRNP particles**. Some of these particles are composed of proteins and small nuclear RNAs, and are called **small nuclear ribonucleoproteins**, or **snRNPs**.Vario2us snRNPs

NOTES

NOTES

assemble into **spliceosome complexes** at exon–intron boundaries of the pre-mRNA and carryout the splicing reaction. In some cases, the primary transcript can be spliced in different ways, a process called **alternative RNA splicing**. For example, an exon that is present in one version of a processed transcript may be spliced out of another version. In this way, the same gene can give rise to different polypeptide chains. Before splicing, the pre-mRNA is modified in two important ways. First it is **capped** by the addition of 7-methylguanylate to the 5′ end of the transcript via a 5′-to-5′ linkage. The pre-mRNA is capped almost immediately after the initiation of mRNA synthesis. One of the functions of the 5′ cap is to protect the growing RNA transcript from degradation by RNases. At a later stage in the synthesis of the primary transcript, the 3′ end is consisting of about 100 to 200 adenylic acid residues, is added by the enzyme poly-A polymerase (figure4).The poly-A tail has several functions: (1) It protects against RNases and therefore increases the stability of mRNA molecules in the cytoplasm, (2) both it and the 5′cap are required for transit through the nuclear pore, and (3) it increases the efficiency of translation on the ribosomes. The requirement of eukaryotic mRNAs to have both a 5′ cap and a poly-A tail ensures that only properly processed transcripts will reach the ribosome and be translated. Each step in eukaryotic gene expression can potentially regulate the amount of gene product in the cell at any given time. Like transcription initiation, splicing may be regulated. Export from the nucleus is also regulated. For example, to exit the nucleus an mRNA must possess a 5′ cap and a poly-A tail, and itmust be properly spliced. Incompletely processed transcripts remain in the nucleus and are degraded. Various posttranscriptional regulatory mechanisms have been identified in eukaryotes. The stabilities or **turnover rates** of mRNA molecules differ from one another, and may vary from tissue to tissue, depending on the physiological conditions. In addition to RNA turnover, the **translatability** of mRNA molecules is variable. For example, RNAs fold into molecules with varying secondary and tertiary structures that can influence the accessibility of the translation initiation codon (the first AUG sequence) to the ribosome. Another factor that can influence translatability of an mRNA is codon usage. There is redundancy in the triplet codons that specify a given amino acid during translation, and each

cell has a characteristic ratio of the different amino acylated tRNAs available, known as **codon bias**. If a message contains a large number of triplet codons that are rare for that cell, the small number of charged tRNAs available for those codons will slow translation. Finally, the **cellular location** at which translation occurs seems to affect the rate of gene expression. Free polysomes may translate mRNAs at very different rates from those at which polysomes bound to the endoplasmic reticulum do; even within the endoplasmic reticulum, there may be differential translation rates (Fig. 12.9).

NOTES

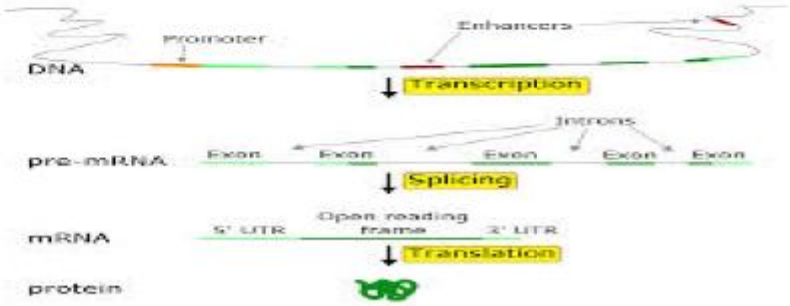


Fig.12.9 : Gene Expression in Eukaryotes

12.7 Catabolite Repression

Catabolite repression is an important part of global control system of various bacteria and other micro-organisms. It is an intracellular regulatory mechanism whereby glucose, or any other carbon source is an intermediate in catabolism, prevents formation of inducible enzymes. This is usually achieved by inhibition of synthesis of enzymes involved in catabolism of carbon sources.

The catabolite repression was first shown to be initiated by glucose sometimes referred to as the **glucose effect**. However, the term "glucose effect" is actually a misnomer since other carbon sources are known to induce catabolite repression.

Following examples are explaining the catabolic repression mechanism in different micro-organisms.

- 1 Escherichia coli
- 2 Bacillus subtilis

NOTES

1. Escherichia coli

Catabolite repression mostly studied in *Escherichia coli*. *E. coli* grows faster on glucose than on any other carbon source. For example, if *E. coli* is placed on an agar plate containing only glucose and lactose, the bacteria will use glucose first and lactose second. When glucose is available in the environment, the synthesis of β-galactosidase is under repression due to the effect of catabolite repression caused by glucose. The catabolite repression in this case is achieved through the utilization of phosphotransferase system.

An important enzyme from the phosphotranferase system called Enzyme II A (**EIIA**) plays an important role in this mechanism. There are different catabolite-specific **EIIA** in a single cell, even though different bacterial groups have specificities to different sets of catabolites. In specefic bacterial group one of the **EIIA** enzymes in their set is specific for glucose transport only. When glucose levels are high inside the bacteria, **EIIA** mostly exists in its un-phosphorylated form, resulting inhibition of adenylyl cyclase and lactose permease, therefore cAMP levels are low and lactose can not be transported inside the bacteria. After some time, the glucose is all used up and the secondary source lactose has to be used by bacteria. Lack of glucose will "turn off" catabolite repression.

Further, when glucose levels are low the phosphorylated form of **EIIA** accumulates and later activates the enzyme adenylyl cyclase, which will produce high levels of cAMP. It binds to catabolite activator protein (CAP) and together they will bind to a promoter sequence on the lac operon (Fig. 12.10A). However, this is not enough for the lactose genes to be transcribed. Lactose must be present inside the cell to remove the lactose repressor from the operator sequence i.e. transcriptional regulation. When these two conditions are satisfied, it means for the bacteria that glucose is absent and lactose is available. Next, bacteria start to transcribe lactose gene and produce β-galactosidase enzymes for lactose metabolism. Given example is a simplification of a complex process. Catabolite repression is a part of global control system and further it affects more genes rather than just lactose gene transcription.

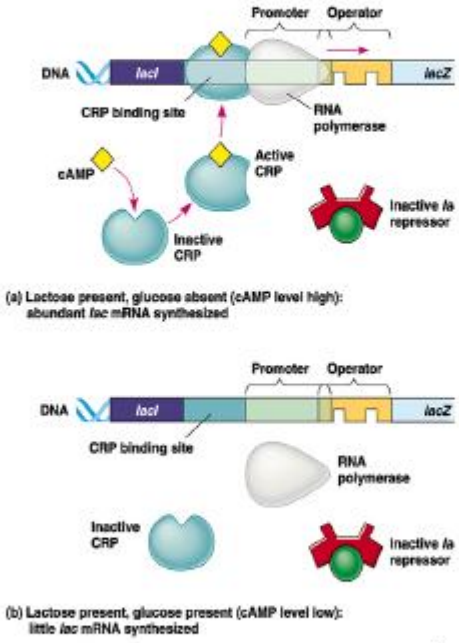


Fig.12.10A: Catabolites Repression in *E.coli.*

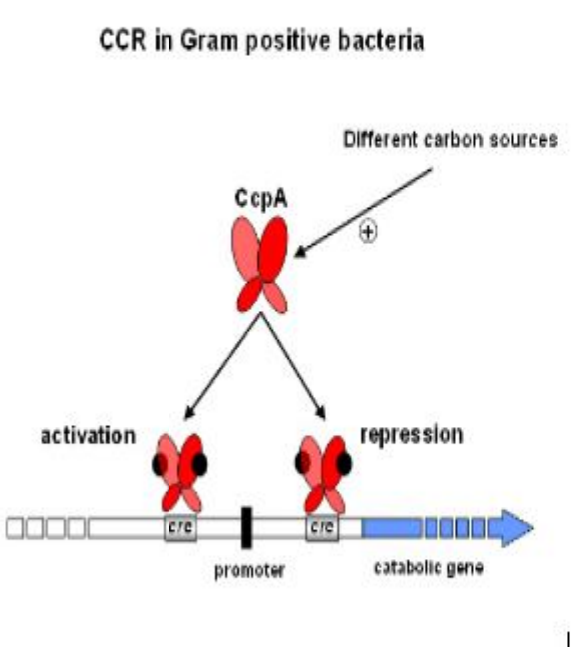


Fig.12.10. B: Catabolites Repression in *Bacillus subtilis* (Gram+ve Bacteria).

2. *Bacillus subtilis*

Bacillus subtilis (Gram+ve) is consist cAMP-independent catabolite repression mechanism regulated by catabolite control protein A (CcpA). In this secondary pathway CcpA negatively suppresses other sugar operons so they are off in the presence of glucose (Fig.12.10B). It works by the fact that Hpr is phosphorylated by a specific mechanism, when glucose enters through the cell membrane protein EIIC, and when Hpr is phosphoralated it will then allow CcpA to block transcription of the alternative sugar pathway operons at their respective cre sequence binding sites. *E. coli* has a similar cAMP-independent catabolite repression mechanism that utilizes a protein called catabolite repressor activator (Cra).

12.8 Regulation of Gene Exprerssion by Trans-cription Attenuation and Antitermination

The 20 common amino acids are required in large amounts for protein synthesis, and *E. coli* can synthesize all of them. The genes for the enzymes needed to synthesize a given amino acid are generally clustered in an operon

NOTES

and are expressed whenever existing supplies of that amino acid are inadequate for cellular requirements. When the amino acid is abundant, the biosynthetic enzymes are not needed and the operon is repressed. The *E. coli* tryptophan (*trp*) operon (Figure 3) includes five genes for the enzymes required to convert chorismate to tryptophan. When tryptophan is abundant it binds to the Trp repressor, causing a conformational change that permits the repressor to bind to the *trp* operator and inhibit expression of the *trp* operon. Different cellular concentrations of tryptophan can vary the rate of synthesis of the biosynthetic enzymes over a 700-foldrange. Once repression is lifted and transcription begins, the rate of transcription is fine-tuned by a second regulatory process, called **transcription attenuation,** in which transcription is initiated normally but is abruptly halted *before* the operon genes are transcribed. The frequency with which transcription is attenuated is regulated by the availability of tryptophan and relies on the very close coupling of transcription and translation in bacteria. The *trp* operon attenuation mechanism uses signals encoded in four sequences within a 162 nucleotide **leader** region at the 5' end of the mRNA, preceding the initiation codon of the first gene. Within the leader lies a region known as the **attenuator,** made up of sequences 3 and 4. These sequences base-pair to form a GqC-rich stem-and-loop structure closely followed by a series of U residues. The attenuator structure acts as a transcription terminator. Sequence 2 is an alternative complement for sequence 3. If sequences 2 and 3 base-pair, the attenuator structure cannot form and transcription continues into the *trp* biosynthetic genes; the loop formed by the pairing of sequences 2 and 3 does not obstruct transcription. Regulatory sequence 1 is crucial for a tryptophan sensitive mechanism that determines whether sequence 3 pairs with sequence 2 (allowing transcription to continue) or with sequence 4 (attenuating transcription). Formation of the attenuator stem-and-loop structure depends on events that occur during *translation* of regulatory sequence 1, which encodes a leader peptide of 14 amino acids, two of which are Trp residues. The leader peptide has no other known cellular function; its synthesis is simply an operon regulatory device. This peptide is translated immediately after it is transcribed, by a ribosome that follows closely behind RNA polymerase as transcription proceeds. When tryptophan concentrations are high, concentrations of charged tryptophan tRNA (Trp-tRNATrp) are also high. This allows translation to

proceed rapidly past the two Trp codons of sequence 1 and into sequence 2, before sequence 3 is synthesized by RNA polymerase. In this situation, sequence 2 is covered by the ribosome and unavailable for pairing to sequence 3 when sequence 3 is synthesized; the attenuator structure (sequences3 and 4) forms and transcription halts. When tryptophan concentrations are low, however, the ribosome stalls at the two Trp codons in sequence 1, because charged tRNA-Trp is less available. Sequence 2 remains free while sequence 3 is synthesized, allowing these two sequences to base-pair and permitting transcription to proceed. In this way, the proportion of transcripts that are attenuated declines as tryptophan concentration declines. Many other amino acid biosynthetic. The *leu* operon leader peptide has four contiguous Leu residues. The leader peptide for the operon contains seven contiguous His residues. In fact, in the operon and a number of others, attenuation is sufficiently sensitive to be the *only* regulatory mechanism.

NOTES

Antitermination : In antitermination the prokaryotic cell's help to fix premature termination of RNA synthesis during the transcription of RNA. It occurs when the RNA polymerase by pass the termination signal, and it provides a mechanism whereby one or more genes at the end of an operon can be switched either on or off, depending on the RNA polymerase either recognizing or not recognizing the termination signal. Antitermination is used by some phages to regulate progression from one stage of gene expression to the next.

12.9 Summary

Gene is a unit of genetic material controlling the inheritance of one phenotyphic characteristic or one trait. The gene is operationally defined by the cis trans or complementation test.

The utility of the cis-trans test is united in some cases by (1.) intragenic complementation (2) polar effect (3) Cis-acting genes or gene products.Gene vexpression is frequently under the control of regulator gene.

Operon is unit of transcription and repressors are allosteric proteins. Operons can not be induced in the presence of high concentration of glucose this phenomena is known as catabolites repression. Operon controlling enzymes

involved in amino acid biosynthetic pathways are frequently controlled by attenuation.

The regulation of gene expression in eukaryotes occurs at the level of transcription.

NOTES

12.10 Glossary

- **Catabolite repression :** Glucose mediated reduction in transcription.
- **Citron :** A unit of functional DNA
- **Cytoplasmic-** inheritance: heredity transmission dependent on the cytoplasm.
- **Operon :** A group of gene making up a regulatory or control unit includes an operator, a promoter and structural gene.

12.11 Self-Learning Exercise

Section –A (Very Short Answer Type Questions)

1. One gene one enzyme theory was proposed by ……..
2. Operon model was proposed by ……..
3. The term gene was proposed by ……..
4. …….a unit of function, unit of mutation and unit of recombination.
5. What is intron.
6. Define muton.
7. What recon.
8. What is citron.

Section–B (Short Answer Type Questions)

1. What is attenuation?
2. What RNA splicing.
3. What is the cis Trans test?
4. Write two significance of introns.

5. What is antitermination.

Section–C (Long Answer Type Questions)

1. Write a note on catabolite repression.
2. What is intron .write the ir significance?
3. Explain the gene regulation in prokaryotes.
4. Write the note on following.

   (a) RNA splicing (b) cis-trans test.

Answer of section A

1. Bedle and tatum
2. Jacob and monad
3. Johanson
4. Gene

NOTES

12.12 References

- Deutscher, Josef (April 2008). "The mechanisms of carbon catabolite repression in bacteria". *Current Opinion in Microbiology***11** (2): 87–93. doi:10.1016/ j.mib.2008. 02.007. ISSN 1369-5274. PMID 18359269.
- Madigan, M.T., J.M. Martinko, P.V. Dunlap, and D.P. Clark. *Brock biology of microorganisms*. 12th ed. San Francisco, CA: Pearson/Benjamin Cummings, 2009.
- Rastogi, V.B. (1991-92) : A Text Book of Genetics.
- Gupta, P.K. (2000) : A Reference Book of Genetics.
- Lodish and Berk (2004): Molecular Cell Biology.
- Gardner/Simmons/Snusted - Principle of genetics sixth edition.

NOTES

Unit - 13

Genetics of Prokaryotes and Eukaryotic Organelles

Structure of the Unit

13.0 Objectives

The structure and context of this unit is guided by our basic belief that can easily explain to the genetic study.

- Genetics of Prokaryotes and Eukaryote
- Recombination in Bacteriophage
- Genectics of Mitochondria and Chloroplast.

13.1 Introduction

NOTES

Genetics of prokaryotes and eukaryotic organelles have interest in all scientific world. Bacteriophage genome analysis, genetics transformation, conjugation and transduction in bacteria, genetics of chloroplast and mitochondria are explained in this unit.

Bacteriophages are viruses that parasitize and kill bacteria, can be used in several difofferent types ofgenetic analysis . Genetic recombination is defined as the rearrangement of genes or parts of genes. Bacterial transformation may be referred to as a stable genetic change brought about by the uptake of naked DNA (DNA without associated cells or proteins) to increase DNA quantity and competence refers to the state of being able to take up exogenous DNA from the environment. There are two forms of transformation and competence: natural and artificial. **Chloroplasts** and **Mitochondria** are organelles, specialized subunits, in plant and algal cells. The main role chloroplast is to conduct photosynthesis, where the photosynthetic pigment chlorophyll captures the energy from sunlight, and stores it in the energy storage molecules ATP and NADPH while freeing oxygen from water.

Chloroplasts have their own **DNA** often abbreviated as **ctDNA**, or is also known as the plastome. Its existence was first proved in 1962, and first sequenced in 1986-when two Japanese research teams sequenced the chloroplast DNA of liverwort and tobacco. Since then, hundreds of chloroplast DNAs from various species have been sequenced, but they are mostly those of land plants and green algae-glaucophytes, red algae, and other algal groups are extremely under represented, potentially introducing some bias in views of "typical" chloroplast DNA structure and content.

13.2 Mapping the Bacteriophage Genome and Phage Phenotypes

Bacteriophages are viruses that parasitize and kill bacteria, can be used in several different types of genetic analysis. T2 and its close relative T4 are viruses that infect the bacterium *E. coli*. The infection ends with destruction

(lysis) of the bacterial cell so these viruses are examples of bacteriophages ("bacteria eaters").

Each virus particle (**virion**) consists of:

NOTES

- a protein head (~0.1 μm) inside of which is a single, circular molecule of double-stranded DNA containing 166,000 base pairs.
- a protein tail from which extend
- thin protein fibers

Life Cycle:

- The virus attaches to the *E. coli* cell
- (**a**). This requires a precise molecular interaction between the fibers and the cell wall of the host.
- The DNA molecule is injected into the cell (**b**) (Fig. 13.1).

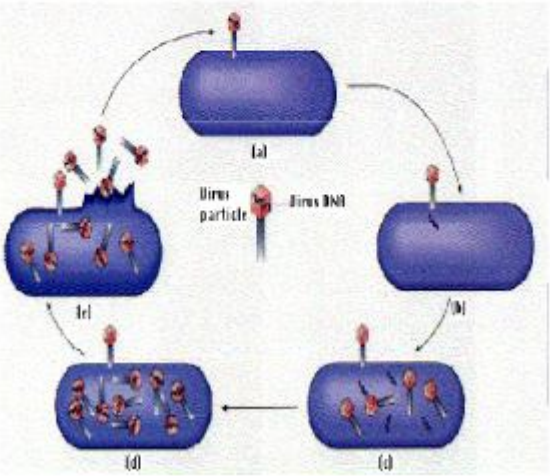


Fig. 13.1 : Life-Cycle of Bateriophage

- Within 1 minute, the viral DNA begins to be transcribed and translated into some of the viral proteins, and synthesis of host proteins is stopped.
- At 5 minutes, viral enzymes needed for synthesis of new viral DNA molecules are produced (**c**).
- At 8 minutes, some 40 different structural proteins for the viral head and tail are synthesized.
- At 13 minutes, assembly of new viral particles begins (**d**).
- At 25 minutes, the viral lysozyme destroys the bacterial cell wall and the viruses burst out — ready to infect new hosts (**e**).
    - If the bacterial cells are growing in liquid culture, it turns clear.

- If the bacterial cells are growing in a "lawn" on the surface of an agar plate, then holes, called **plaques**, appear in the lawn.

Occasionally, new phenotypes appear such as a change in the appearance of the plaques or even a loss in the ability to infect the host.

NOTES

As with so many organisms, the occurrence of mutations provides the tools to learn about such things as

- the function of the gene;
- its location in the DNA molecule (mapping).

Mapping by Recombination Frequencies

As we have seen, *E. coli* strain B can be infected by both h^+ and h strains of T2. In fact, a single bacterial cell can be infected simultaneously by both.

Let us infect a liquid culture of *E. coli* **B** with two different mutant T2 viruses

- $h\ r^+$ and
- h^+ r

When this is done in liquid culture, and then plated on a mixed lawn of *E. coli* **B** and **B/2**, four different kinds of plaques appear (Fig. 13.2).

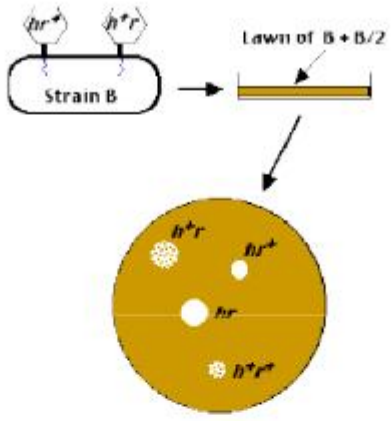


Fig. 13.2 : Genotype

Phenotype	Number of Plaques
hr^+	clear, small	460
h^+r	turbid, large	460

NOTES

h^+r^+	turbid, small	40
hr	clear, large	40
	Total =	1000

The most abundant (460 each) are those representing the parental types; that is, the phenotypes are those expected from the two infecting strains. However, small numbers (40 each) of two new phenotypes appear. These can be explained by **genetic recombination** having occasionally occurred between the DNA of each parental type within the bacterial cell.

Just as in higher organism, one assumes that the frequency of recombinants is proportional to the distance between the gene loci. In this case, 80 out of 1000 plaques were recombinant, so the distance between the *h* and *r* loci is assigned a value of 8 **map units** or **centimorgans** (cM).

13.3 Genetic Recombination in Phage

Genetic recombination is defined as the rearrangement of genes or parts of genes. Throughout evolution, the acquisition of 'new' genes and the rearrangement of 'old' ones has been driven by genetic recombination between bacteriphage and bacterial genomes. In *Escherichia coli*, genetic recombination is carried out 'cautiously' through carefully regulated mechanisms. Although genetic recombination may be highly active in repairing DNA damage during periods of stress or in promoting genetic diversity, uncontrolled recombination causes genomic instability. This relatively 'static' recombination mechanism shifts dramatically when *E. coli* is infected by the temperate bacteriphage λ (lambda). Each and every λ progeny undergoes at least one recombination event, despite possessing a tiny chromosome of only 50 kb in length. Thus, the bacterium is rendered a hotbed of genetic exchange in its last few minutes of existence during the natural history of the cycle of infection. The highly efficient rate of recombination induced by λ is called 'hyper rec' and requires just 40 bp of homology.

Genetic recombination in bacteriophage λ

NOTES

During rolling circle replication, each λ progeny acquires a double-strand break at *cos* (the cohesive site). If a cut at *cos* is the only double-strand-break available, λ recombination is generally RecA-dependent. RecA-dependent genetic recombination with a linear λ chromosome is a 'break-join' mechanism. This mechanism involves the λ exonuclease Exo, which degrades the 5′-ending strand of duplex DNA, while the 3′-ending strand is preserved, generating a 3′-ssDNA (single-stranded DNA) overhang. Subsequently, the *E. coli* RecA protein binds to the processed 5′-ssDNA-overhang generated by Exo. This step allows the assembly of multiple RecA monomers that form a presynaptic filament on ssDNA. Invasion of the filament into a homologous duplex results in synapse formation via a replication-dependent double-strand invasion recombination pathway .In the presence of a homologous partner ssDNA, however, λ may mediate recombination with Exo-processed substrates via a RecA-independent single-strand annealing pathway. This requires the λ-encoded ssDNA-binding protein Bet, which anneals partner ssDNA. The λ-encoded proteins involved in these reactions are analogous to the *E. coli* RecBCD enzyme (Red system), the RecFOR enzyme (Orf) and the RuvC resolvase (Rap). In the absence of RecFOR, Orf is required for early events of recombinational exchange. The Rap resolvase targets recombination intermediates such as Holliday junctions, which arise as a consequence of host restriction, terminase cleavage, DNA damage, rolling circle replication or replication fork collapse (Fig. 13.3).

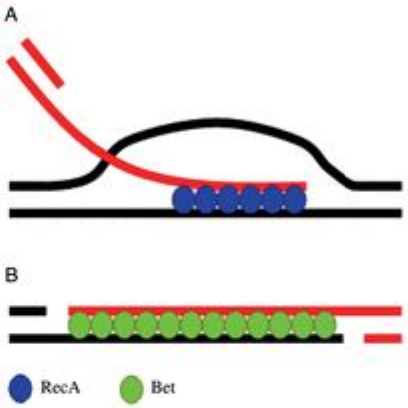


Fig. 13.3 : Genetic Recombination in Bacterio-Phage λ

Pathways for λ-mediated double-strand invasion vs. single-strand annealing. (**A**) Double-strand invasion is RecA-dependent. (**B**) Bet-mediated single-strand annealing is RecA-independent.

NOTES

Bet is required for Rec A-independent single-strand annealing

The ssDNA-binding protein Bet has Mg^{2+}-dependent and adenosine triphosphate (ATP)-independent ssDNA annealing properties. Bet loads onto the 3′-end of Exo-processed ssDNA due to 3′–5′ polar cooperativity. This activity allows Bet to spontaneously form ring structures or helical filaments on >35 nucleotide long ssDNA *in vitro*. The half-time for Bet-mediated renaturation, like RecA protein, is independent of ssDNA concentration. This feature contrasts with the renaturation reactions mediated by the *E. coli* single-stranded binding protein and is a second-order reaction that allows Bet to form filaments on dsDNA analogous to the eukaryotic Rad52 protein). Thus, Bet promotes renaturation of homologous ssDNA partner sequences that have been processed by Exo preferentially at sites rich in AT base pairs, while the RecA protein prefers GC-rich sequences. The strand exchange activities of Bet may extend to a homologous duplex with a single-strand gap, allowing displacement of the complementary strand with the free energy of the annealing reaction. Whether these reactions require RecA is unclear.. Bet remains bound to recombinant dsDNA after ssDNA assimilation, although Bet does not bind to dsDNA directly. This has been suggested to confer resistance against nucleases attack. Because Bet forms a complex with Exo *in vitro*. Bet modulates the nucleolytic and recombinational activities of Exo through a direct mechanism. This interaction may allow Bet to protect recombination intermediates from degradation by inhibiting Exo. Thus, the main function of Bet is to promote 3′-end assimilation, while preferentially presenting 5′-ends to Exo for degradation.

Gam protects recombination intermediates from nuclease attack

Gam is a 276-amino acid homodimer that binds to and inhibits the nuclease activity of *E. coli*-encoded nucleases. Gam inhibits all the known activities of the *E. coli* RecBCD enzyme, including its helicase, nuclease and ATPase activity. Expression of Gam confers the types of defects associated with

RecBCD mutants. The inactive RecBCD-Gam complex is thought to exist in an equilibrium between free and complex forms, allowing some portion of the enzyme to exist in active state. Nevertheless, Gam prevents RecBCD-mediated degradation of linear duplexes *in vitro*. Indeed, *RecBCD* mutants and RecBCD-Gam recombination both produce DNA concatemeres during λ replication *in vivo*. Inhibition of RecBCD likely occurs by Gam preventing RecBCD from binding to dsDNA ends. The mode of Gam binding to RecBCD is thought to involve disruption of amino acid residues in RecB that are important for DNA binding. Addition of Gam to RecBCD–dsDNA complexes has been shown to lead to complex dissociation, suggesting that Gam interferes with the RecBCD DNA-binding site even when that site is occupied. Interestingly, Gam has been reported to confer radioresistance to X-rays, but sensitivity to ultraviolet (UV). The reason for this paradox is likely to centre on the types of DNA lesions produced by these modes of ionizing radiation. UV-induced lesions tend to be ssDNA daughter-strand gaps caused by blockage of the replisome at pyrimidine dimers or abasic sites. Only occasionally do these lesions generate dsDNA ends that are processed by RecBCD. In contrast, the proportion of dsDNA breaks generated by X-rays is thought to be higher, because X-rays produce a higher proportion of clustered single-strand breaks that become 'frank' double-strand breaks that are processed by RecBCD. Therefore, Gam prevents the RecBCD-mediated digestion of dsDNA that may be repaired by genetic recombination. RecBCD-mediated reactions are usually destructive, until the activity of the enzyme is modified by the χ recombination hotspot.

NOTES

13.4 Genetic Transformation

Bacterial transformation may be referred to as a stable genetic change brought about by the uptake of naked DNA (DNA without associated cells or proteins) to increase DNA quantity and competence refers to the state of being able to take up exogenous DNA from the environment. There are two forms of transformation and competence: natural and artificial.

Transformation is the process by which foreign DNA is introduced into a cell. Transformation of bacteria with plasmids is important not only for studies in bacteria but also because bacteria are used as the means for both storing and replicating plasmids. Because of this, nearly all plasmids, even those designed

for use in mammalian cells; carry both a bacterial origin of replication and an antibiotic resistance gene for use as a selectable marker in bacteria.

NOTES

"Transformation" may also be used to describe the insertion of new genetic material into nonbacterial cells, including animal and plant cells; however, because "transformation" has a special meaning in relation to animal cells, indicating progression to a cancerous state, the term should be avoided for animal cells when describing introduction of exogenous genetic material. Introduction of foreign DNA into eukaryotic cells is often called "transfection". Trans-formation was first demonstrated in 1928 by British bacteriologist Frederick Griffith.

Griffith Experiment

Griffith discovered that a harmless strain of Streptococcus pneumoniae could be made virulent after being exposed to heat-killed virulent strains. Griffith hypothesized that some "transforming principle" from the heat-killed strain was responsible for making the harmless strain virulent. In 1944 this "transforming principle" was identified as being genetic by Oswald Avery, Colin MacLeod, and Maclyn McCarty. They isolated DNA from a virulent strain of *S. pneumoniae* and using just this DNA was able to make a harmless strain virulent. They called this uptake and incorporation of DNA by bacteria "transformation" (Fig. 13.4).

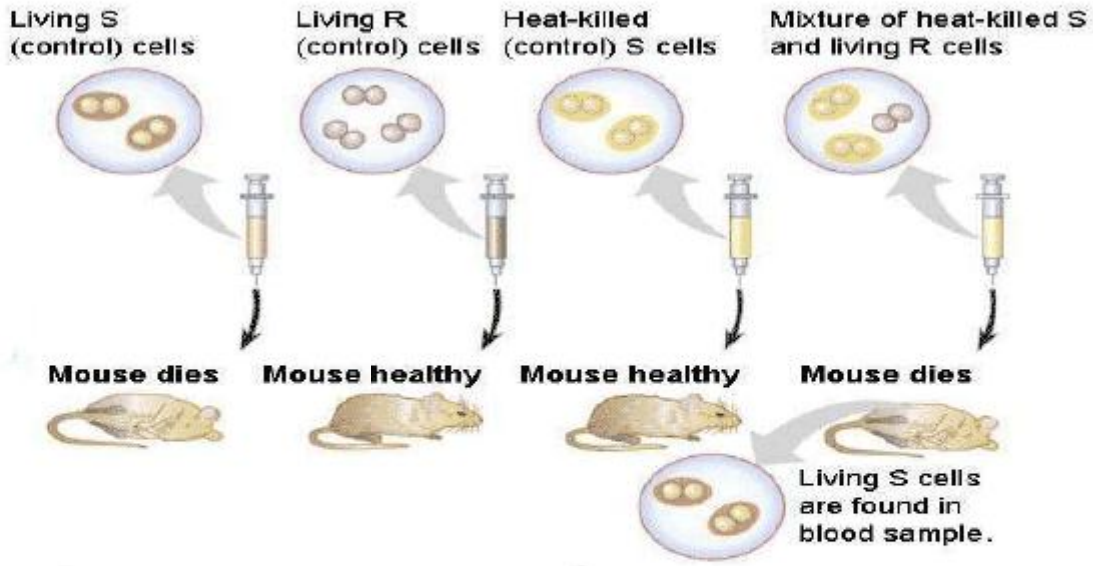


Fig. 13.4 : Griffith Transformation Experiment

Natural transformation

Natural transformation is a bacterial adaptation for DNA transfer that depends on the expression of numerous bacterial genes whose products appear to be designed to carry out this process. In general, transformation is a complex, energy requiring developmental process. In order for a bacterium to bind, take up and recombine exogenous DNA into its chromosome it must become competent, that is, enter a special physiological state. Competence development in Bacillus subtilis requires expression of about 40 genes. The DNA integrated into the host chromosome is usually derived from another bacterium of the same species, and is thus homologous to the resident chromosome.

NOTES

Natural competence

About 1% of bacterial species are capable of naturally taking up DNA under laboratory conditions; more may be able to take it up in their natural environments. DNA material can be transferred between different strains of bacteria, in a process that is called horizontal gene transfer. Some species upon cell death release their DNA to be taken up by other cells; however transformation works best with DNA from closely related species. These naturally competent bacteria carry sets of genes that provide the protein machinery to bring DNA across the cell membrane(s).

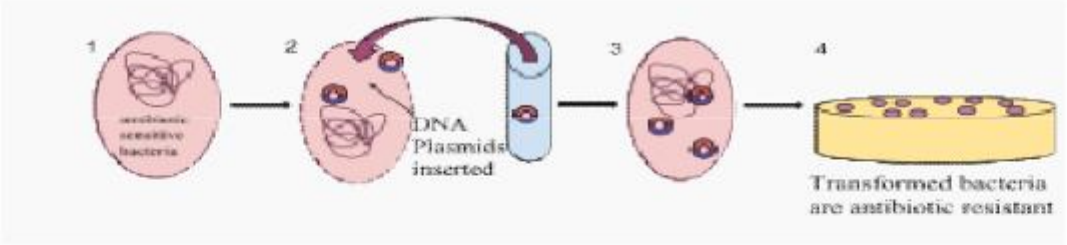


Fig. 13.5 : Artificial Competence

Schematic of bacterial transformation — for which artificial competence must first be induced (Fig. 13.5).

Artificial competence can be induced in laboratory procedures that involve making the cell passively permeable to DNA by exposing it to conditions that do not normally occur in nature. Typically the cells are incubated in a solution containing divalent cations (often calcium chloride) under cold conditions, before being exposed to a heat pulse (heat shock).

13.5 Conjugation in Bacteria

NOTES

Bacterial conjugation was discovered in 1946 by Joshua Lederberg and Edward Tatum, It is the transfer of genetic material (plasmid) between bacterial cells by direct cell-to-cell contact or by a bridge-like connection between two cells.

Bacterial conjugation is often regarded as the bacterial equivalent of sexual reproduction or mating since it involves the exchange of genetic material. During conjugation the donor cell provides a conjugative or mobilizable genetic element that is most often a plasmid or transposon. Most conjugative plasmids have systems ensuring that the recipient cell does not already contain a similar element.

The genetic information transferred is often beneficial to the recipient. Benefits may include antibiotic resistance, xenobiotic tolerance or the ability to use new metabolites, Such beneficial plasmids may be considered bacterial endosymbionts.

Mechanism

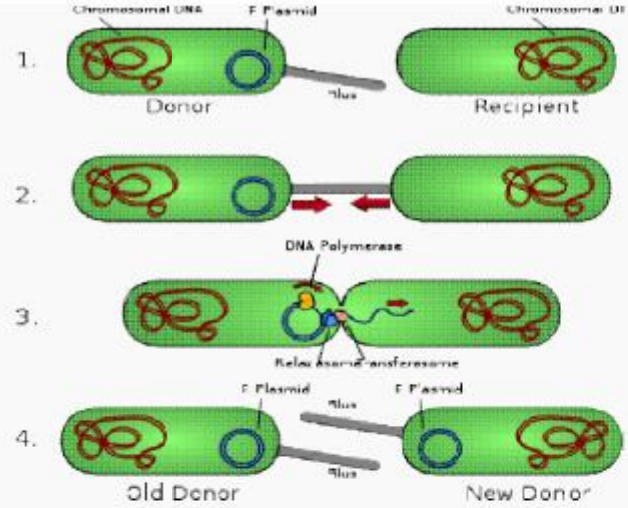


Fig. 13.6 : Bacterial Conjugation.

Steps of Conjugation

1. Donor cell produces pilus.
2. Pilus attaches to recipient cell and brings the two cells together.

3. The mobile plasmid is nicked and a single strand of DNA is then transferred to the recipient cell.

4. Both cells synthesize a complementary strand to produce a double stranded circular plasmid and also reproduce pili; both cells become viable donors now.

NOTES

The conjugative plasmid is the **F-plasmid**, or F-factor. The F-plasmid is an episome (a plasmid that can integrate itself into the bacterial chromosome by homologous recombination) with a length of about 100 kb. It carries its own origin of replication, the oriV, and an origin of transfer, oriT, There can only be one copy of the F-plasmid in a given bacterium, either free or integrated, and bacteria that possess a copy are called *F*-positive or *F*-plus (denoted F^+). Cells that lack F plasmids are called F-negative or F-minus (F^-) and as such can function as recipient cells (Fig. 13.6).

Among other genetic information the F-plasmid carries a *tra* and *trb* locus, which together are about 33 kb long and consist of about 40 genes. The *tra* locus includes the *pilin* gene and regulatory genes, which together form pili on the cell surface. The locus also includes the genes for the proteins that attach themselves to the surface of F^- bacteria and initiate conjugation.

When conjugation is initiated by a signal the **relaxase** enzyme creates a nick in one of the strands of the conjugative plasmid at the *oriT*. Relaxase may work alone or in a complex of over a dozen proteins known collectively as a **relaxosome**. The remaining strand is replicated either independent of conjugative action (vegetative replication beginning at the *oriV*) or in concert with conjugation (conjugative replication similar to the rolling circle replication of lambda phage). Conjugative replication may require a second nick before successful transfer can occur.

If the F-plasmid that is transferred has previously been integrated into the donor's genome (producing an Hfr strain ["High Frequency of Recombination"]) some of the donor's chromosomal DNA may also be transferred with the plasmid DNA. The amount of chromosomal DNA that is transferred depends on how long the two conjugating bacteria remain in contact. In common laboratory strains of *E. coli* the transfer of the entire

bacterial chromosome takes about 100 minutes (Fig. 13.7). The transferred DNA can then be integrated into the recipient genome via homologous recombination.

NOTES

Since integration of the F-plasmid into the *E. coli* chromosome is a rare spontaneous occurrence, and since the numerous genes promoting DNA transfer are in the plasmid genome rather than in the bacterial genome.

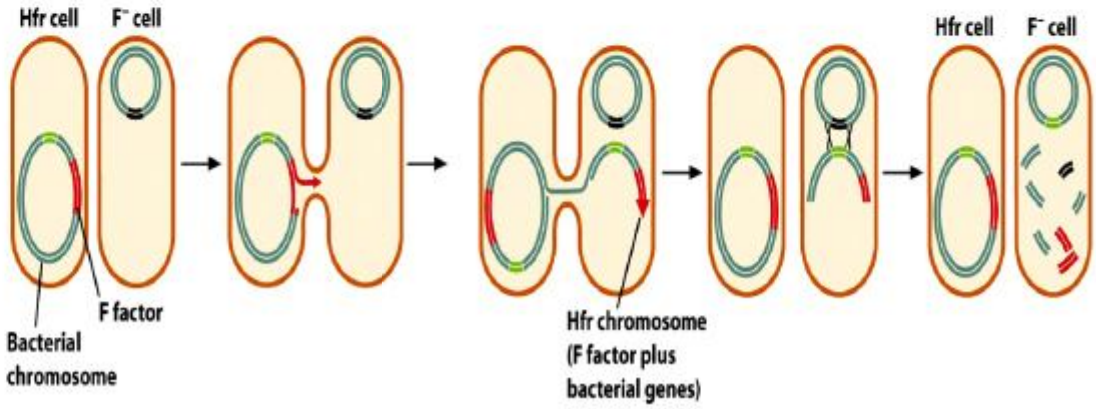


Fig. 13.7 : Bacterial Hfr Conjugation

13.6 Transduction in Bacteria

Transduction is the process in which DNA is transferred from one bacterium to another by a virus, it also refers to the process where foreign DNA is introduced into bacterium via a viral vector. Transduction does not require physical contact between the cell donating the DNA and the cell receiving the DNA, and it is DNase resistant .Transduction is a common tool used by molecular biologists to stably introduce a foreign gene into a host cell's genome.

When bacteriophages (viruses that infect bacteria) infect a bacterial cell, their normal mode of reproduction is to use the replicational, transcriptional, and translational machinery of the host bacterial cell to make numerous virions, or complete viral particles, including the viral DNA or RNA and the protein coat (Fig. 13.8).

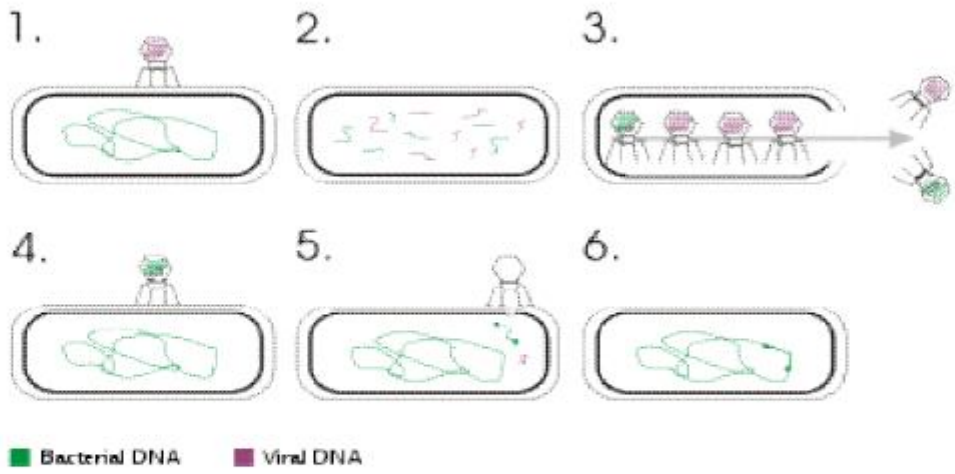


Fig. 13.8 : Transduction in Bacteria

Discovery : In 1951, Joshua Lederberg and Norton Zinder discovered transduction in the bacterium *Salmonella typhimurium.*

Lytic and lysogenic (temperate) cycles

Transduction happens through either the lytic cycle or the lysogenic cycle. If the lysogenic cycle is adopted, the phage chromosome is integrated (by covalent bonds) into the bacterial chromosome, where it can remain dormant for thousands of generations. If the lysogen is induced (by UV light for example), the phage genome is excised from the bacterial chromosome and initiates the lytic cycle, which culminates in lysis of the cell and the release of phage particles. The lytic cycle leads to the production of new phage particles which are released by lysis of the host.

Transduction as a method for transferring genetic material

The packaging of bacteriophage DNA has low fidelity and small pieces of bacterial DNA, together with the bacteriophage genome, may become packaged into the bacteriophage genome. At the same time, some phage genes are left behind in the bacterial chromosome.

There are generally three types of recombination events that can lead to this incorporation of bacterial DNA into the viral DNA, leading to two modes of recombination.

13.6.1 Types of Transduction as a method for transferring genetic material. It is mainly of two types as follows:-

NOTES

1. Generalized transduction
2. Specialized transduction

Generalized Transduction

NOTES

Generalized transduction is the process by which any bacterial gene may be transferred to another bacterium via a bacteriophage, and typically carries only bacterial DNA and no viral DNA. In essence, this is the packaging of bacterial DNA into a viral envelope. This may occur in two main ways, recombination and headful packaging.

If bacteriophages undertake the lytic cycle of infection upon entering a bacterium, the virus will take control of the cell's machinery for replicating its own viral DNA. If by chance bacterial chromosomal DNA is inserted into the viral capsid which is usually used to encapsulate the viral DNA, the mistake will lead to *generalized transduction*.

If the virus replicates using 'headful packaging', it attempts to fill the nucleocapsid with genetic material. If the viral genome results in spare capacity, viral packaging mechanisms may incorporate bacterial genetic material into the new virion.

The new virus capsule now loaded with part bacterial DNA continues to infect another bacterial cell. This bacterial material may become recombined into another bacterium upon infection.

When the new DNA is inserted into this recipient cell it can fall to one of three fates:-

1. The DNA will be absorbed by the cell and be recycled for spare parts.
2. If the DNA was originally a plasmid, it will re-circularize inside the new cell and become a plasmid again.
3. If the new DNA matches with a homologous region of the recipient cell's chromosome, it will exchange the DNA material (Fig. 13.9).

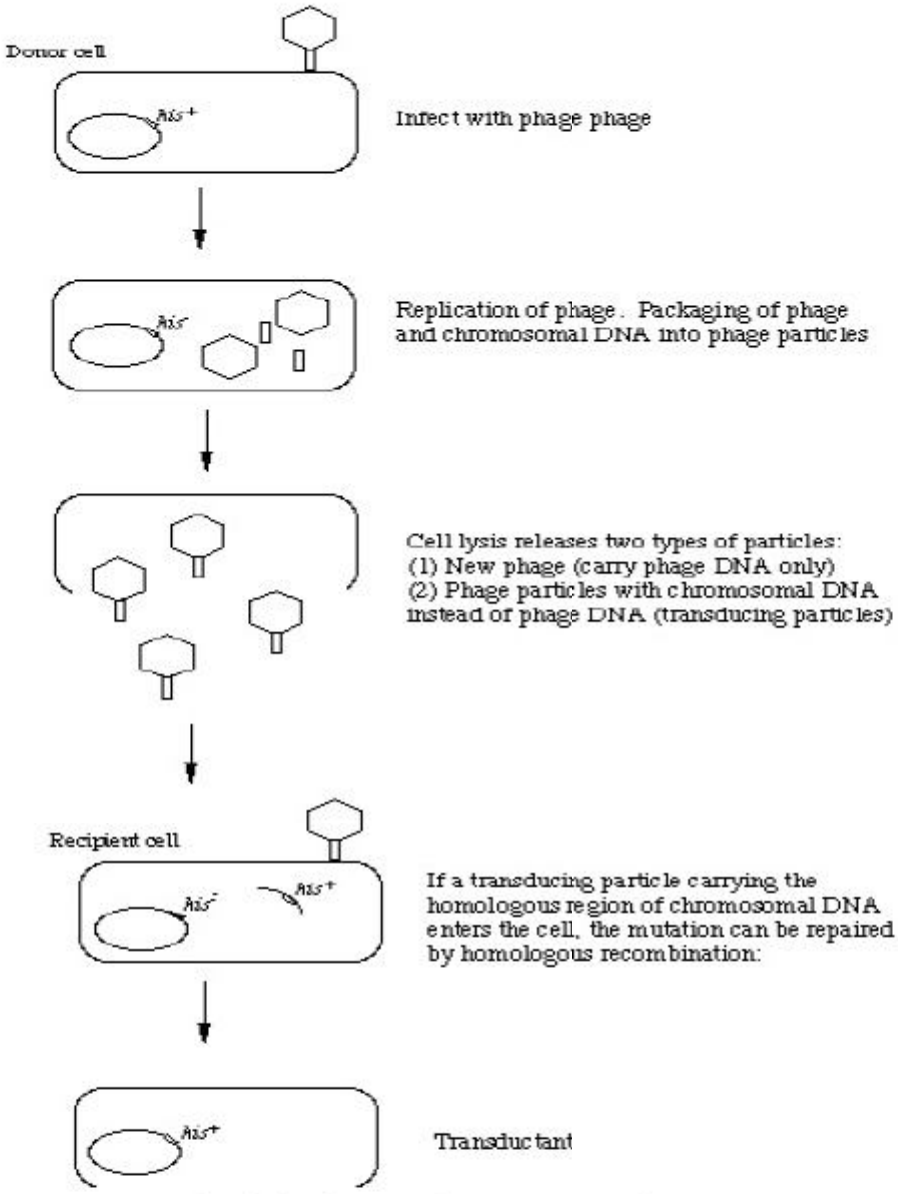


Fig. 13.9 : Generalized Transduction

NOTES

Specialized Transduction

Specialized transduction is the process by which a restricted set of bacterial genes is transferred to another bacterium. The genes that get transferred (donor genes) depend on where the phage genome is located on the chromosome. Specialized transduction occurs when the prophage excises imprecisely from the chromosome so that bacterial genes lying adjacent to the prophage are

NOTES

included in the excised DNA. The excised DNA is then packaged into a new virus particle, which then delivers the DNA to a new bacterium, where the donor genes can be inserted into the recipient chromosome or remain in the cytoplasm, depending on the nature of the bacteriophage.

When the partially encapsulated phage material infects another cell and becomes a "prophage" (is covalently bonded into the infected cell's chromosome), the partially coded prophage DNA is called a "heterogenote".

Example of specialized transduction is λ phages in *Escherichia coli* discovered by Esther Lederberg (Fig. 13.10).

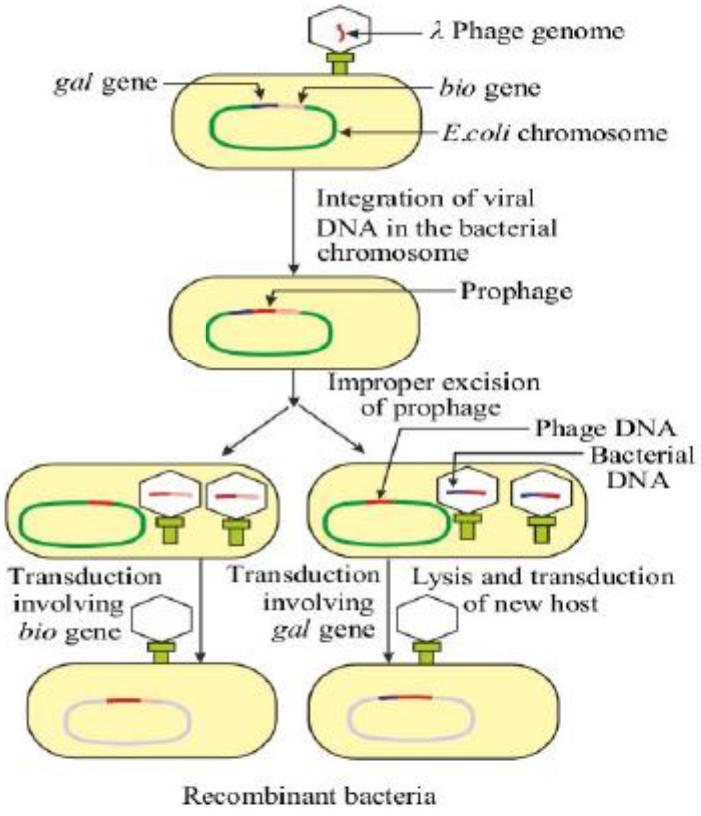


Fig. 13.10 : Specialized Transduction

13.7 Genetics of Chloroplast

Chloroplasts are organelles, specialized subunits, in plant and algal cells. Their main role is to conduct photosynthesis, where the photosynthetic pigment chlorophyll captures the energy from sunlight, and stores it in the energy storage molecules ATP and NADPH while freeing oxygen from water.

They then use the ATP and NADPH to make organic molecules from carbon dioxide in a process known as the Calvin cycle. Chloroplasts carry out a number of other functions, including fatty acid synthesis, much amino acid synthesis, and the immune response in plants.

NOTES

A chloroplast is one of three types of plastids, characterized by its high concentration of chlorophyll. (The other two types, the leucoplast and the chromoplast, contain little chlorophyll and do not carry out photosynthesis). Chloroplasts are highly dynamic—they circulate and are moved around within plant cells, and occasionally pinch in two to reproduce. Their behavior is strongly influenced by environmental factors like light color and intensity. Chloroplasts, like mitochondria, contain their own DNA, which is thought to be inherited from their ancestor—a photosynthetic cyanobacterium that was engulfed by an early eukaryotic cell. Chloroplasts cannot be made by the plant cell, and must be inherited by each daughter cell during cell division.

Chloroplasts have their own **DNA** often abbreviated as **ctDNA**, or is also known as the plastome. Its existence was first proved in 1962, and first sequenced in 1986—when two Japanese research teams sequenced the chloroplast DNA of liverwort and tobacco. Since then, hundreds of chloroplast DNAs from various species have been sequenced, but they are mostly those of land plants and green algae—glaucophytes, red algae, and other algal groups are extremely under represented, potentially introducing some bias in views of "typical" chloroplast DNA structure and content. Figure 13.11 showing the structure of chloroplast.

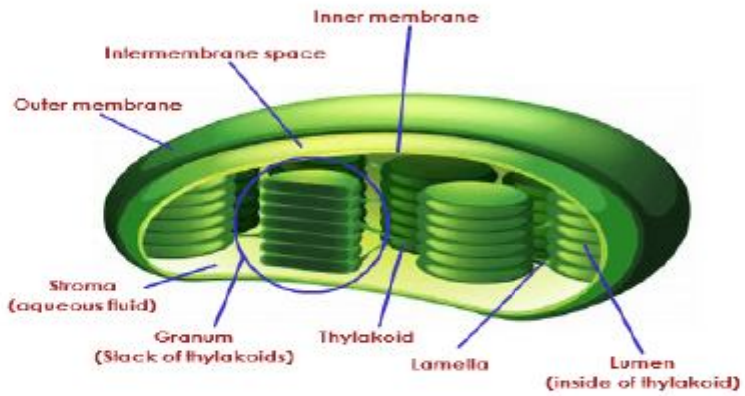


Fig. 13.11 : Structure of Chloroplast

NOTES

Molecular structure

Chloroplast DNAs are circular, and are typically 120,000–170,000 base pairs long. They can have a contour length of around 30–60 micrometers, and have a mass of about 80–130 million daltons. Most chloroplasts have their entire chloroplast genome combined into a single large ring, though those of dinophyte algae are a notable exception—their genome is broken up into about forty small plasmids, each 2,000–10,000 base pairs long. Each mini circle contains one to three genes but blank plasmids, with no coding DNA, have also been found.

Inverted repeats

Chloroplast contains two inverted repeats, which separate a long single copy section (LSC) from a short single copy section (SSC).The inverted repeats vary wildly in length, ranging from 4,000 to 25,000 base pairs long each. Inverted repeats in plants tend to be at the upper end of this range, each being 20,000–25,000 base pairs long. The inverted repeat regions usually contain three ribosomal RNA and two tRNA genes, but they can be expanded or reduced to contain as few as four or as many as over 150 genes. While given pair of inverted repeats are rarely completely identical, they are always very similar to each other, apparently resulting from concerted evolution.

Linear structure

Chloroplast DNAs have long been thought to have a circular structure, but some evidence suggests that chloroplast DNA more commonly takes a linear shape. Over 95% of the chloroplast DNA in corn chloroplasts has been observed to be in branched linear form rather than individual circles.

Nucleoids

New chloroplasts may contain up to 100 copies of their DNA, though the number of chloroplast DNA copies decreases to about 15–20 as the chloroplasts age. They are usually packed into nucleoids which can contain several identical chloroplast DNA rings. Many nucleoids can be found in each chloroplast. Though chloroplast DNA is not associated with true histones, in red algae, a histone-like chloroplast protein (HC) coded by the chloroplast DNA that tightly packs each chloroplast DNA ring into a nucleoid has been

found. In primitive red algae, the chloroplast DNA nucleoids are clustered in the center of the chloroplast, while in green plants andgreen algae, the nucleoids are dispersed throughout the stroma.

Gene content and protein synthesis

The chloroplast genome most commonly includes around 100 genes which code for a variety of things, mostly to do with the protein pipeline and photosynthesis. As in prokaryotes, genes in chloroplast DNA are organized into operons. Interestingly, introns are common in chloroplast DNA molecules, while they are rare in prokaryotic DNA molecules (plant mitochondrial DNAs commonly have introns, but not human mt-DNA). Among land plants, the contents of the chloroplast genome are fairly similar—they code for four ribosomal RNAs, 30–31tRNAs, 21 ribosomal proteins, and four RNA polymerase subunits, involved in protein synthesis.

Chloroplast genome reduction and gene transfer

Over time, many parts of the chloroplast genome were transferred to the nuclear genome of the host, a process called *endosymbiotic gene transfer*. As a result, the chloroplast genome is heavily reduced compared to that of free-living cyanobacteria. Chloroplasts may contain 60–100 genes whereas cyanobacteria often have more than 1500 genes in their genome. Endosymbiotic gene transfer is how we know about the lost chloroplasts in many chromalveolate lineages. Even if a chloroplast is eventually lost, the genes it donated to the former host's nucleus persist, providing evidence for the lost chloroplast's existence. For example, here have been a few recent transfers of genes from the chloroplast DNA to the nuclear genome in land plants. Of the approximately three-thousand proteins found in chloroplasts, some 95% of them are encoded by nuclear genes. Many of the chloroplast's protein complexes consist of subunits from both the chloroplast genome and the host's nuclear genome.

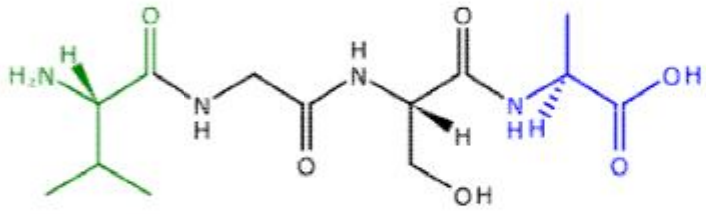
Protein synthesis

Protein synthesis within chloroplasts relies on an RNA polymerase coded by the chloroplast's own genome, which is related to RNA polymerases found in bacteria. Chloroplasts also contain a mysterious second RNA polymerase that is encoded by the plant's nuclear genome. The two RNA polymerases may

recognize and bind to different kinds of promoters within the chloroplast genome. The ribosomes in chloroplasts are similar to bacterial ribosomes.

Cytoplasmic translation and N-terminal transit sequences

NOTES



A polypeptide with four amino acids linked together. At the left is the N-terminus, with its amino (H_2N) group in green. The blue C-terminus, with its carboxyl group(CO_2H) is at the right.

Polypeptides, the precursors of proteins, are chains of amino acids. The two ends of a polypeptide are called the N-terminus, or *amino end*, and the C-terminus, or *carboxyl end*. For many (but not all) chloroplast proteins encoded by nuclear genes, *cleavable transit peptides* are added to the N-termini of the polypeptides, which are used to help direct the polypeptide to the chloroplast for import (N-terminal transit peptides are also used to direct polypeptides to plant mitochondria). N-terminal transit sequences are also called *presequences* because they are located at the "front" end of a polypeptide—ribosomes synthesize polypeptides from the N-terminus to the C-terminus. Chloroplast transit peptides exhibit huge variation in length and amino acid sequence. They can be from 20–150 amino acids long—unusually long lengths, suggesting that transit peptides are actually collections of domains with different functions. Transit peptides tend to be positively charged, ich in hydroxylated amino acids such as serine, threonine, and proline, and poor in acidic amino acids like aspartic acid and glutamic acid. In an aqueous solution, the transit sequence forms a random coil. Not all chloroplast proteins include a N-terminal cleavable transit peptide though. Some include the transit sequence within the functional part of the protein itself. A few have their transit sequence appended to their C-terminus instead. Most of the polypeptides that lack N-terminal targeting sequences are the ones that are sent to the outer chloroplast membrane, plus at least one sent to the inner chloroplast membrane.

Phosphorylation, chaperones, and transport

NOTES

After a chloroplast polypeptide is synthesized on a ribosome in the cytosol, ATP energy can be used to phosphorylate, or add a phosphate group to many (but not all) of them in their transit sequences. Serine and threonine (both very common in chloroplast transit sequences—making up 20–30% of the sequence) are often the amino acids that accept the phosphate group. The enzyme that carries out the phosphorylation is specific for chloroplast polypeptides, and ignores ones meant for mitochondria or peroxisomes. Phosphorylation changes the polypeptide's shape, making it easier for 14-3-3 proteins to attach to the polypeptide. In plants, 14-3-3 proteins only bind to chloroplast preproteins.

The translocon on the outer chloroplast membrane (TOC)

The TOC complex, or *translocon on the outer chloroplast membrane*, is a collection of proteins that imports preproteins across the outer chloroplast envelope. Five subunits of the TOC complex have been identified—two GTP-binding proteins Toc34 and Toc159, the protein import tunnel Toc75, plus the proteins Toc64 and Toc12. The first three proteins form a core complex that consists of one Toc159, four to five Toc34s, and four Toc75s that form four holes in a disk 13 nanometers across. The whole core complex weighs about 500 kilodaltons. The other two proteins, Toc64 and Toc12, are associated with the core complex but are not part of it.

13.8 Genetics of Mitochondria

The **mitochondrion** (plural **mitochondria**) is a membrane bound organelle found in most eukaryotic cells (the cells that make up plants, animals, fungi, and many other forms of life). The word mitochondrion comes from the Greek μίτος, *mitos*, i.e. "thread", and χονδρίον, *chondrion*, i.e. "granule. Mitochondria range from 0.5 to 1.0 micrometer (μm) in diameter. These structures are sometimes described as "cellular power plants" because they generate most of the cell's supply of adenosine triphosphate (ATP).

NOTES

Mitochondrial proteins vary depending on the tissue and the species. In humans, 615 distinct types of proteins have been identified from cardiac mitochondria, whereas in rats, 940 proteins have been reported. The mitochondrial proteome is thought to be dynamically regulated. Although most of a cell's DNA is contained in the cell nucleus, the mitochondrion has its own independent genome. Further, its DNA shows substantial similarity to bacterialgenomes.

A mitochondrion contains outer and inner membranes composed of phospholipid bilayers and proteins. The two membranes have different properties. Because of this double-membraned organization, there are five distinct parts to a mitochondrion. They are:

1. the outer mitochondrial membrane,
2. the intermembrane space (the space between the outer and inner membranes),
3. the inner mitochondrial membrane,
4. the cristae space (formed by infoldings of the inner membrane), and
5. the matrix (space within the inner membrane).

The human mitochondrial genome is a circular DNA molecule of about 16 kilobases. It encodes 37 genes: 13 for subunits of respiratory complexes I, III, IV and V, 22 for mitochondrial tRNA (for the 20 standard amino acids, plus an extra gene for leucine and serine), and 2 for rRNA. One mitochondrion can contain two to ten copies of its DNA. As in prokaryotes, there is a very high proportion of coding DNA and an absence of repeats. Mitochondrial genes are transcribed as multigenic transcripts, which are cleaved and polyadenylated to yield mature mRNAs.

Mitochondrial genomes have far fewer genes than the bacteria from which they are thought to be descended. Although some have been lost altogether, many have been transferred to the nucleus, such as the respiratory complex II protein subunits. This is thought to be relatively common over evolutionary time. A few organisms, such as the *Cryptosporidium*, actually have mitochondria that lack any DNA, presumably because all their genes have been lost or transferred. In *Cryptosporidium*, the mitochondria have an altered ATP

generation system that renders the parasite resistant to many classical mitochondrial inhibitors such as cyanide, azide, and atovaquone.

Replication and inheritance

Mitochondria divide by binary fission, similar to bacterial cell division. The regulation of this division differs between eukaryotes. In many single-celled eukaryotes, their growth and division is linked to the cell cycle. For example, a single mitochondrion may divide synchronously with the nucleus. This division and segregation process must be tightly controlled so that each daughter cell receives at least one mitochondrion. In other eukaryotes (in mammals for example), mitochondria may replicate their DNA and divide mainly in response to the energy needs of the cell, rather than in phase with the cell cycle. When the energy needs of a cell are high, mitochondria grow and divide. When the energy use is low, mitochondria are destroyed or become inactive. In such examples, and in contrast to the situation in many single celled eukaryotes, mitochondria are apparently randomly distributed to the daughter cells during the division of the cytoplasm. Understanding of mitochondrial dynamics, which is described as the balance between mitochondrial fusion and fission, has revealed that functional and structural alterations in mitochondrial morphology are important factors in pathologies associated with several disease conditions.

An individual's mitochondrial genes are not inherited by the same mechanism as nuclear genes. Typically, the mitochondria are inherited from one parent only. The mitochondria, and therefore the mitochondrial DNA, usually come from the egg only. The sperm's mitochondria enter the egg but do not contribute genetic information to the embryo. Instead, paternal mitochondria are marked with ubiquitin to select them for later destruction inside the embryo. The egg cell contains relatively few mitochondria, but it is these mitochondria that survive and divide to populate the cells of the adult organism. Mitochondria are, therefore, in most cases inherited only from mothers, a pattern known as maternal inheritance. This mode is seen in most organisms including the majority of animals. However, mitochondria in some species can sometimes be inherited paternally. This is the norm among certain coniferous plants, although not in pine trees and yew trees. For Mytilidae mussels paternal inheritance only occurs within males of

the species. It has been suggested that it occurs at a very low level in humans. There is a recent suggestion mitochondria that shorten male lifespan stay in the system because mitochondria are inherited only through the mother. By contrast natural selection weeds out mitochondria that reduce female survival as such mitochondria are less likely to be passed on to the next generation. Therefore it is suggested human females and female animals tend to live longer than males. The authors claim this is a partial explanation.

NOTES

13.9 Cytoplasmic Male Sterility

Cytoplasmic male sterility is total or partial male sterility in plants as a result of specific nuclear and mitochondria interactions. Male sterility is defined as the failure of plants to produce functional anthers, pollen, or male gametes. Cytoplasmic male sterility was first observed by Joseph Gottlieb Kölreuter. He observed anther abortion within species and specific hybrids. Cytoplasmic male sterility has now been identified in over 150 plant species. It is more prevalent than female sterility. Male sterility is easy to detect because a large number of pollen grains are produced and are easily studied. Male sterility is detected through staining techniques (carmine, lactophenol or iodine), while detection of female sterility is by the absence of seeds. Male-sterile plants may be propagated, since they can still set seed, while female-sterile plants cannot be propagated. Male sterility can be either cytoplasmic or cytoplasmic–genetic. Cytoplasmic male sterility (CMS) is caused by the extranuclear genome (mitochondria or chloroplast) and shows maternal inheritance. Manifestation of male sterility in CMS may be controlled either entirely by cytoplasmic factors or by interaction between cytoplasmic and nuclear factors.

Cytoplasmic male sterility is under extranuclear genetic control (under control of the mitochondrial or plastid genomes). It shows non-Mendelian inheritance, with male sterility inherited maternally. In general there are two types of cytoplasm: N (normal) and aberrant S (sterile) cytoplasms. These types exhibit reciprocal differences.

Cytoplasmic-genetic male sterility

While CMS is controlled by an extranuclear genome, nuclear genes may have the capability to restore fertility. When nuclear restoration of fertility genes

("Rf") is available for a CMS system in any crop, it is cytoplasmic–genetic male sterility; the sterility is manifested by the influence of both nuclear (with Mendelian inheritance) and cytoplasmic (maternally inherited) genes. There are also restorers of fertility i.e. (*Rf*) genes that are distinct from genetic male sterility genes. The *Rf* genes have no expression of their own unless the sterile cytoplasm is present. *Rf* genes are required to restore fertility in S cytoplasm that causes sterility. Thus plants with N cytoplasm are fertile and S cytoplasm with genotype *Rf*- leads to fertiles while S cytoplasm with *rfrf* produces only male steriles. Another feature of these systems is that *Rf* mutations (*i.e.*, mutations to *rf* or no fertility restoration) are frequent, so that N cytoplasm with *Rfrf* is best for stable fertility (Fig. 13.12).

NOTES

Cytoplasmic–genetic male sterility systems are widely exploited in crop plants for hybrid breeding due to the convenience of controlling sterility expression by manipulating the gene–cytoplasm combinations in any selected genotype. Incorporation of these systems for male sterility evades the need for emasculation in cross-pollinated species, thus encouraging cross breeding producing only hybrid seeds under natural conditions.

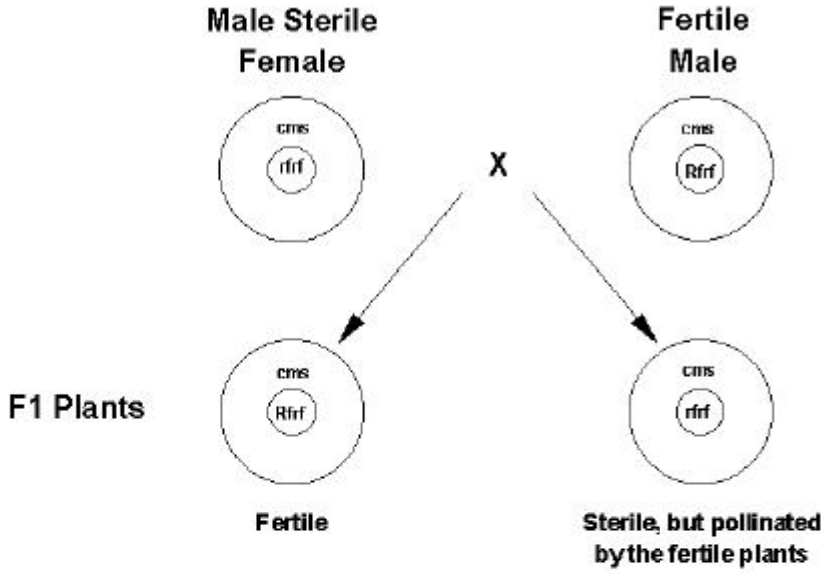


Fig.13.12 : Pattern of Cytoplasmic Genetic Male Sterility

NOTES

Cytoplasmic male sterility in hybrid breeding

Hybrid production requires a female plant in which no viable male gametes are borne. Emasculation is done to make a plant devoid of pollen so that it is made female. Another simple way to establish a female line for hybrid seed production is by identifying or creating a line that is unable to produce viable pollen. Since this male-sterile line cannot self-pollinate, seed formation is dependent upon pollen from the male line. Cytoplasmic male sterility is used in hybrid seed production. In this case, the sterility is transmitted only through the female and all progeny will be sterile. This is not a problem for crops such as onions or carrots where the commodity harvested from the F1 generation is produced during vegetative growth. These CMS lines must be maintained by repeated crossing to a sister line (known as the maintainer line) that is genetically identical except that it possesses normal cytoplasm and is therefore male-fertile. In cytoplasmic–genetic male sterility restoration of fertility is done using restorer lines carrying nuclear genes. The male-sterile line is maintained by crossing with a maintainer line carrying the same nuclear genome as the MS line but with normal fertile cytoplasm.

Cytoplasmic male sterility in hybrid maize breeding

Cytoplasmic male sterility is an important part of hybrid maize production. The first commercial cytoplasmic male sterile, discovered in Texas, is known as CMS-T. The use of CMS-T, starting in the 1950s, eliminated the need for detasseling. In the early 1970s, plants containing CMS-T genetics were susceptible to southern corn leaf blight and suffered from widespread loss of yield. Since then CMS types C and S are used instead. Unfortunately these types are prone to environmentally induced fertility restoration and must be carefully monitored in the field. Environmentally induced, in contrast to genetic, restoration occurs when certain environmental stimuli signal the plant to bypass sterility restrictions and produce pollen anyway.

The systematic sequencing of new plant species in recent years has uncovered the existence of several novel RF genes and their encoded proteins. A unified nomenclature for the RF defines protein families across all plant species and facilitates comparative functional genomics. This nomenclature accomm-odates

functional RF genes and pseudogenes, and offers the flexibility needed to incorporate additional RFs as they become available in future.

13.10 Summary

NOTES

Bacteriophage is a virus that attack bacteria and they destroy their bacterial hosts.

Genetic recombination is defined as the rearrangement of genes or parts of genes. Throughout evolution, the acquisition of 'new' genes and the rearrangement of 'old' ones have been driven by genetic recombination between bacteriphage and bacterial genome. In bacteria three different mechanism for transfer of genetic material from one cell to another cell during (1) Transformation naked DNA molecules are taken up by competent recipient cells. (2) In transduction a fragment of the donor chromosome is carried to and injected in to the recipient cell by bacteriophage and in (3) conjugation the transfer of donor DNA to recipient Cells through a conjugation tube that forms between the two cells.

Transformation .transduction and conjugation almost always produce cells that are partial zygotes or partial diploids called merozygotes. Mitochondria and chloroplast carry small amount of unique DNA that is independently with respect to nuclear gene. Male sterility in maize and some other plants is controlled by cytoplsmic factors. Maternal effects are controlled by nuclear gene.

13.11 Glossary

- **Bacteriophage:** Virus that attack bacteria such virus is called bacteriophage because they destroy their bacterial hosts.
- **Conjugation:** A way of transfer of genetic material fro donor (male) to a recipient (female).
- **Transduction:** Genetic recombination in bacteria by bacteriophage.
- **Transformation:** Genetic alternation of bacterial brought about by the incorporation of foreign DNA in the bacterial cells.

- **Maternal effect:** Trait controlled by gene of the mother but expressed in the progeny.

13.12 Self -Learning Exercise

NOTES

Section –A (Very Short Answer Type Questions)

1. The Area of lysis on a bacterial lawn culture produced by a phage is known as he process by……..
2. Which phage reproduction is initiated in lysogenized culture is called……...?
3. One of the first enzymes synthesized by many bacteriophages is _____, an RNA-dependent RNA polymerase.
4. Who discovered the transduction in bacteria?
5. Define hfr recombination.
6. Define bacteriophage.
7. Write the type of transduction.
8. Which cell organelle known as power house?

Section –B (Short Answer Type Questions)

1. What are bacteria?
2. What is bacterial conjugation When it was discovered
3. Define transformation.
4. What is transduction
5. Define cytoplasm male sterility
6. Write any two function of mitochondria

Section-C (Long Answer Type Questions)

1. Write a note on genetic of chloroplast.
2. What is transduction .discuss different type of transduction?
3. Explain the conjugation method of gene transfer in bacteria
4. Write a note on cytoplasmic male sterility.

Answer key of Section -A

1. Induction
2. Plaque
3. RNA replicase

NOTES

13.13 References

- Gómez-Campo, C. (1999). *Biology of Brassica Coenospecies*. Elsevier. pp. 186–189. ISBN 978-0-444-50278-0.
- Schnable PS, RP Wise (1998) The molecular basis of cytoplasmic male sterility. Trends in Plant Science, 3(5): 175-180.
- Weider, Christophe Stamp, Peter Christov, Nikolai Husken, Alexandra Foueillassar, Xavier Camp, Karl-Heinz Munsch, Magali (2009) Stability of Cytoplasmic Male Sterility ins Maize under Different Environmental Conditions. Crop Science 49: 77-84.
- Rastogi, V.B. (1991-92) : A Text Book of Genetics
- Gupta, P.K. (2000) : A Reference Book of Genetics
- Lodish and Berk (2004): Molecular Cell Biology
- Gardner/Simmons/Snusted - Principle of genetics sixth edition.

NOTES

Unit-14

Biostatistics

Structure of the Unit

NOTES

14.0 Objectives

After going through this unit you will be able to:

- Define biostatistics, important terms &symbols.
- Define and describe quantitative data.
- Prepare various type of graphs.
- Understand and compute Mean, Mode& Median
- Understand and explain the various methods of data analysis.
- Compute measures of central tendency ,variance, standard
- deviation
- Apply t-test and Chi-squre test.
- Understand & apply correlation and regression.
- Define and describe probability distributions

14.1 Introduction

When we want to explain questions like “how much”, “how many”, “how fast”, “what time”, etc. all these questions need numbers for answers. The numbers have become so important that a separate branch of knowledge has developed,

NOTES

which is known as statistics. In know days, the subject became more and more popular and its application became wider and wider. In early years it was considered a branch of economics, but now the statistics has applications in virtually all areas of human endeavour. No matter where we look, we find statistics being used there. The naturalists, doctors, businessmen, meteorologists, economists, retailers, insurance companies, administrators, bankers, etc., all employ statistics in their fields of activity. In biological sciences, statistical approach is obligatory in ecology, plant breeding, genetics, medicines, microbiology, etc.

By the use of statistics we can able to collect not only numerical data but also to provide a methodology for handling, analyzing and drawing valid inferences from data. Statistics, if approached with a proper frame of mind, can be the most exciting field of study.

When we use statistics in biology it called as biostatistics Biostatistics is the application of statistical techniques to scientific research in life science including plant and animal and the development of new tools to study these areas.

In this unit we shall focus on various type of data and describe the methods use in biostatistics .The aim of this unit is make you understand the important terms and symbols used in biostatistics. In this unit we dealt with various topic like Mean, Mode, Median, ANOVA,t-test.Chi-squre test& probability distributions.

14.2 Important Terms & Symbols

14.2.1 Biostatistics

Biostatistics *is the application of statistical methods to the problems of biology, including human biology, medicine and public health. Statistical methods include the collection, organization, summary-classification and analyzing and measured evaluation of facts to reach some inference.*

Biometry (literally meaning 'biological measurement') Biostatistics is also called. The word biometry has Greek origin ('bios' means life and 'metron' means measured).

14.2.2 Statistical Terms

NOTES

1. **Population**

   In a statistical term, *population refers to any well defined group of individuals who are being studied or of observations of a particular type*. More simply, a group of study elements is called **population.** For example, all university students in Jaipur could be a population. All patients of a hospital suffering from Malaria or patients suffering from cancer and treated with a new drug may be considered as population.

2. **Sample**

   Sample *is a small group or portion of a population selected using a suitable method so that it can be regarded as representative of the entire population and can be used for investigating its properties.* **.** For example, we want to study the average height of students studying in class X of different Senior Secondary Schools in Kota. It is not necessary to observe the height measurements of all the students. In fact, we can take a small representative sample of a few students from different schools for measurements of all the students and can give results.

Sampling may be Random or Biased

1. **Random sample** gives every member of the population an equal chance of being selected. No one in the population is favoured over other in the selection process.
2. **Biased sample or Nonrandom sample** does not provide equal opportunity for all members of the population of being selected. Sample is drawn with a purpose.
3. **Unit**

   Unit *is the smallest object or individual that can be investigated as the source of basic information, e.g.,* small subareas of land, individual patients, etc. Units are expressed in two ways:

   (a) Sampling units during surveys.

   (b) Experimental units during experiments.

4. **Variable**

NOTES

Variable is a quantity which can vary from one individual to another. For example, animals of same species may differ in their length, weight, age, sex, etc. These characteristics are variables. Thus, variable can be defined as *'the characteristics by which individuals differ among themselves'*. The particular values of a variable are termed as **variate** or **variate values**. Variables may be of two types :

1. **Quantitative variable** is a characteristic which can be measured on a scale in some appropriate units, e.g., measurement of age, weight, length, etc. Quantitative variables can be further subdivided into two types :

(i) **Discontinuous or discrete variable** is one which is incapable of taking all possible values, e.g., the number of rooms in a house or the number of persons in the family can take only the integral values such as 2,3,4, etc. Here a count of 2½ is not possible.

(ii) **Continuous variable** is one which can take any numerical value within a certain range, e.g., the height of a child at various ages when he grows from 120 cm. to 150 cm., assumes all possible values within the limit even in fractions.

2. **Qualitative variable** is one which is unmeasurable and unexpressible in magnitude. It can be expressed in qualities which are called **attributes**, e.g., texture of flowers, colour of leaves, etc.

5. **Constant**

Constant *is a quantity which does not vary from one member of a group to another or within a particular set of defined conditions, e.g.,* the number of enrolled students in a particular course or a particular college at the time when data were collected.

6. **Parameter**

Parameter *is any numerical property, characteristic or fact that is descriptive of a population.* Population mean, median and mode are the **'parameter of location'.** These relate the numerical values of a variable with the general location of a population. Usually all the characteristics of a population can be specified in terms of a few parameters, e.g., if the characteristic is length and the measurements of length are variable then the mean length can be regarded as a parameter.

7. Data

Data *is a set of facts expressed in quantitative form*. It can be primary or secondary. The data collected by investigator from personal experimental studies is called **primary data** while the data obtained from some secondary source such as journals, magazines, newspapers, etc., is known as **secondary data.** A data can be primary for one person and secondary for the other.

NOTES

8. Accuracy

Accuracy is the closeness of a measured or computed value to its true value.

9. Precision

Precision is the closeness of repeated measurements of the same quantity.

14.3 Graphical Presentation of Quantitative Data

14.3.1 Bar Graphs

The bar graph consists of parallel, usually vertical bars or rectangles with lengths proportional to the frequency with which specified quantities occur in a set of data. Refer to Figure 14.1 which illustrates a bar graph, which depicts data related to distribution of the haemoglobin concentration in children (9-36 months of age). Bar graph differs from a histogram in the sense that the bars are separated by spaces. Bar charts help in making the categories stand out from one another, therefore making it easier to compare each category.

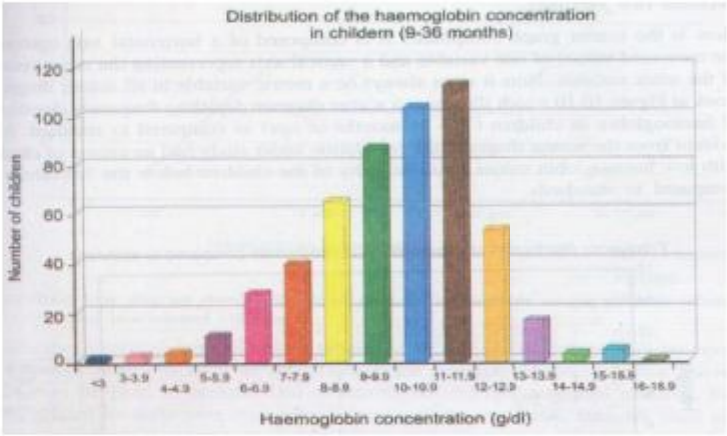


Fig. 14.1: Bar Graph

14.3.2 Histograms

NOTES

A histogram or column diagram is a graph in which class-intervals are represented along the horizontal axis and their corresponding frequencies are represented by areas in the form of rectangular vertical bars down on the intervals.

Let us consider the data presented in table for drawing and explain a histogram.

Table 1: Frequency distribution of the scores of 40 students

Class Intervals	Exact Units of Class Intervals	Mid Point	Freq uency (f)	Cumulative Frequency (F)	Cumulative Percentage Frequency
35-39	34.5-39.5	37	4	40	100.00
30-34	29.5-34.5	32	8	36	90.00
25-29	24.5-29.5	27	11	28	70.00
20-24	19.5-24.5	22	8	17	42.50
15-19	14.5-19.5	17	6	9	22.50
10-14	9.5-14.5	12	3	3	7.50
N = 40					

The histogram drawn for the above data is shown in Figure-14.2.

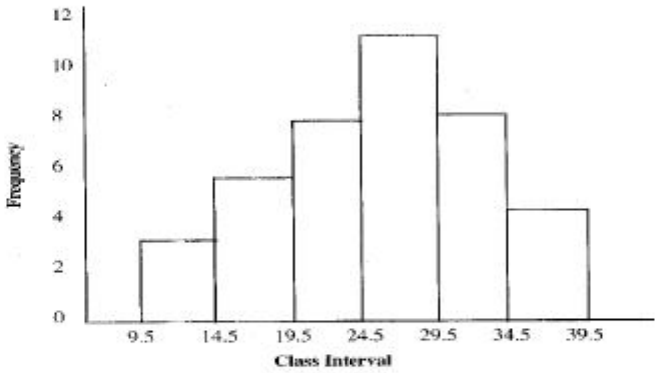


Fig. 14.2: Histogram plotted graph

Thus histogram you would have noticed in Figure 14.2 is a set of contiguously drawn bars. The bars are drawn for each group/interval of values such that the area is proportional to the frequency in that group.

14.3.3 Frequency Polygon

Frequency polygon is drawn by plotting the mid-point of each class-interval (i.e. bars in the histogram) at a height proportional to its respective frequency and then joining the points by straight lines including those with zero frequency at the two ends. The first two steps are identical to those used in the construction of a histogram. The next step to be followed is given as under :

NOTES

Step 3 **:** Directly above the mid-point of each

…… frequency polygon

frequency curve class-interval along the horizontal axis plot the points at a

height proportional to the respective frequencies. Join these points by straight lines. A frequency curve is a smooth curve as drawn in Figure 14.3.

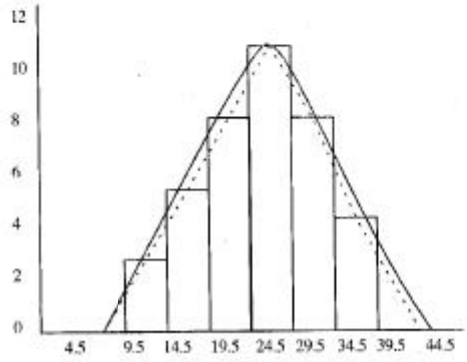
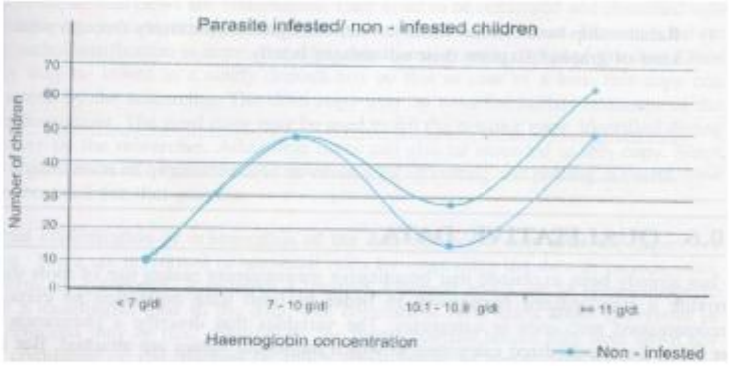


Fig. 14.3: Frequency Polygon plotted graph

14.3.4 Line Diagram

A line graph now than actually plots points and then connects them with a line. It is used to show trend of one variable over the other. Refer to figure, which shows the trend of mean haemoglobin among parasitic infested and non-infested children. As you must have noticed two lines are drawn one for infested and the other for non-infested. The plotted points are joined by a line (Fig-14.4).



.Fig. 14.4: Line Diagram graph

14.4 Frequency Distribution & Centering Constants

Measures of Central Tendency

(Mean, Mode, Median)

NOTES

A population can be expressed by means of a location parameter, around which most of the values of the population should be clustered. Since this central value locates the population, the location parameters are also called *measures of central tendency*. There are at least three measures of central tendency which are commonly used (i) arithmetic mean, (ii) median and (iii) mode.

14.4.1 Arithmetic mean

The arithmetic average of a distribution is known as its mean. The mean of a set of observations or measures is obtained by dividing the sum of all values by the total number of values.

Suppose a finite population consists of values y_1, y_2……y_N, population size being N (*n* is used for sample size), then the arithmetic mean will be denoted as follows:

$$\mu = \frac{y_1 + y_2 + y_3 \ldots . y_N}{N} = \frac{\sum_{i=1}^{N} y_i}{N} \quad (\mu = \text{arithmetic mean of population})$$

Similarly for a sample for which data has been recorded, arithmetic mean will be

$$\bar{y} = \frac{y_1 + y_2 + y_3 \ldots . y_n}{n} = \frac{\sum_{i=1}^{n} y_i}{n} \text{ or } \frac{\sum y}{n} \quad (\bar{y} = \text{arithmetic mean of sample})$$

If each of the *k* values occurs more than once with a frequency f_1, then

$$\bar{y} = \frac{f_1 y_1 + f_2 y_2 + f_3 y_3 \ldots . f_k y_k}{f_1 + f_2 + f_3 \ldots . . f_k} = \frac{\sum f_i y_i}{\sum f_i}$$

Arithmetic mean is considered to be the best measure of central tendency and one can calculate it from data given as follows :$\Sigma fy = 342$, $n = \Sigma f = 100$

$$SS = \sum yi^2 - \frac{(\sum yi)^2}{n} \qquad \bar{y} = \frac{342}{100} = 3.42 \text{ children/family}$$

$$S^2 = \frac{\sum yi^2 - (\sum yi)^2 / n}{n-1}$$

NOTES

In a table of frequency distribution for a continuous variable with class intervals, mid class values are considered to be $x_1, x_2, x_3 \ldots\ldots x_n$ or $y_1, y_2, y_3 \ldots\ldots y_k$ and the arithmetic means then can be calculated for height of 50 students as follows :

$$\bar{y} = \frac{\sum f_i x_i}{\sum f_i} \quad \text{or} \quad \frac{\sum f_i y_i}{\sum f_i}$$

14.4.2 Median

If observations in a population or a sample are arranged in order of magnitude from smallest to largest (represented as $y_{(1)}, y_{(2)} \ldots\ldots y_{(n)}$; the bracket on either side of number shows that they are arranged in order), then the middle value (when n= odd number) or the average of two middle values (when n = odd number) or the average of two middle values (when n = even number) gives the median, which another measure of central tendency. For example, the median of set of data 2, 5, 7, 10, 12 is 7 and that of another set 2, 5, 7, 9, 10, 12 is (7+9)/2 = 8. Sometimes median is defined as a value from which half the remaining values are greater or equal and the other half are less than or equal to the median. The number of the median item will be calculated as

$$Sm = \frac{n+1}{2} \text{ (if n is odd number)}$$

If n = 87 then $Sm = \frac{87+i}{2} = 44$; and if it is 86 $Sm = \frac{86+i}{2} = 43\frac{1}{2}$, so that we use 43 and 44 for calculating median as an average of two values.

14.4.3 Mode

The mode is that observation, which occurs most frequently. In a frequency distribution, modal class is the class with highest frequency. It is not always necessary that a mode may really be found. In other cases, it is also possible that more than one mode may occur in a distribution e.g. bimodal distribution with two modes. This is illustrated in the following three examples.

Original data	Modes
2, 3, 4, 6, 6, 6, 7, 7	one mode i.e. 6
2, 4, 5, 5, 7, 8, 8, 9	two modes i.e. 5, 8
2, 2, 3, 3, 5, 5, 6, 6	no mode

NOTES

When a class interval represents a modal class, it is not always possible to specify a modal value, but in some such cases a particular modal value can be calculated even for such grouped data.

14.5 Measures of Variation

The measures of central tendency are very useful in describing the nature of a distribution of measures, but they do not give the researcher a complete picture of the data. These measures will not tell the researcher how the scores tend to be distributed. For this, we use a different set of measures which are called measures of *'variability'* or measures of *'spread'* or *'dispersion'*. The most commonly used measures of variability include the variance and standard deviation. Let us study about these measures.

14.5.1 The Variance and Standard Deviation

The average of the squared deviations of the measures or values from their means is known as the *variance*.

The standard *deviation* is the positive square root of variance. Let us understand this concept by calculating the variance and standard deviation for grouped/ungrouped data.

a) The *Variance and Standard Deviation for the ungrouped data*

The variance for the ungrouped data is found by using the formula:

$$\sigma^2 = \frac{\sum x^2}{N} \qquad (1)$$

where

σ^2 = variance of the sample

x = deviation of raw measures or values from the mean.

N = number of values or measures

Let us consider the following data of scores for the application of formula (5):

10, 10, 9, 9, 8, 8, 7, 7, 6, 6.

At the deviation of each score from the mean is required, the first thing to do is to calculate the mean.

NOTES

Using formula (1)

$$M = \frac{\sum x}{N} = \frac{80}{10} = 8$$

Now, from each raw score, the mean is substracted to get the value of x as shown in Table2.

Table 2 : Distribution of the test scores of ten learners

Score (X)	**Deviation (X-M) (*x*)**	**Deviation Squared (x^2)**
10	2	4
10	2	4
9	1	1
9	1	1
8	0	0
8	0	0
7	-1	1
7	-1	1
6	-2	4
6	-2	4
		$x^2 = 20$

Using formula (1)

$$\sigma^2 = \frac{\sum x^2}{N}$$

$$= \frac{20}{10} = 2$$

Now to get the standard deviation, we need the positive square root of the variance, σ^2

NOTES

Standard Deviation = $\sigma = \sqrt{\frac{\sum x^2}{N}}$

$= \sqrt{2} = 1.41$

The raw scores instead of deviation scores may also be used. The raw score formulae for variance and standard deviation are given as follows:

$$\textit{Variance} = \sigma^2 = \frac{N\sum X^2 - \left(\sum X\right)^2}{N^2} \quad \textbf{(2)}$$

$$\textit{Standard Deviation} = \sigma = \frac{\sqrt{N\sum X^2 - \left(\sum X\right)^2}}{N} \quad (3)$$

in which

X = raw score

N = the number of scores in the distribution

Using the same set of data, we can calculate variance and standard deviation with the help of raw-score formulae. Refer to Table3.

Table 3 : The calculation of variance and standard deviation from original (raw) score when the assumed mean is taken at zero and the data is ungrouped

Score (X)	X^2
10	100
10	100
9	81
9	81
8	64
8	64
7	49
7	49

6	36
6	36
$\Sigma X = 80$	$\Sigma X^2 = 660$

NOTES

Using Formulae (2)

$$\text{Variance} = \frac{N\sum X^2 - \left(\sum X\right)^2}{N^2}$$

$$= \frac{10 \times 660 - (80)^2}{100}$$

$$= \frac{6600 - 6400}{100} = \frac{200}{100} \qquad = 2$$

Using formula (3)

$$\text{Standard Deviation} = \sigma = \frac{\sqrt{N\sum X^2 - \left(\sum X\right)^2}}{N}$$

$$= \frac{\sqrt{100 \times 600 - (80)^2}}{N}$$

$$= \frac{\sqrt{6600 - 6400}}{10}$$

$$= 1.414$$

Next, let us learn how to calculate variance and standard deviation for grouped data.

b) *Variance and Standard Deviation for grouped data*

In the case of grouped data in a frequency distribution, the variance and standard deviation are calculated by using the formulae :

$$\textbf{\textit{Variance}} = \sigma^2 = \frac{i^2}{N^2}\left[N\sum fx'^2 - \left(\sum fx'\right)^2\right] \qquad \textbf{(4)}$$

$$\textbf{\textit{Standard Deviation :}}\ \frac{i}{N}\left[\sqrt{N\sum fx'^2 - \left(\sum fx'\right)^2}\right] \qquad (5)$$

where

i = width of the class interval

N = total number of measures

F = frequency of class-interval

x' = deviation of the raw measure from the assumed mean divided by the length of class-interval

NOTES

To illustrate the use of these formulae let us consider the distribution given in Table4.

Table 4 : The calculation of variance and standard deviation from data grouped in a frequency distribution

Class Interval	x	f	X'*	fx'	fx'^2
71 – 75	73	3	3*	9	27
66 – 70	68	4	2	8	16
61 – 65	63	9	1	9	9
56 – 60	58AM	15	0	0	0
51 – 55	53	8	-1	-8	8
46 – 50	48	6	-2	-12	24
41 – 45	43	5	-3	-15	45
		N = 50		Σfx' = - 9	$\Sigma fx'^2$ = 129

$*3 = \frac{73-58}{5}$; AM = Assumed mean

Using formula (3)

$$\text{Variance} = \sigma^2 = \frac{i^2}{N^2}\left[N\sum fx'^2 - \left(\sum fx'\right)^2\right]$$

$$= \frac{(5)^2}{(50)^2}[50 + 129 - (-9)]$$

$$= 6.69$$

Using formula (4)

$$\text{Standard Deviation} = \sigma = \frac{i}{N}\left[N\sum fx'^2 - \left(\sum fx'\right)^2\right]$$

$$= \frac{5}{50}\sqrt{\left(50\times 129 - (-9)^2\right)}$$

$$= \frac{1}{10}\sqrt{6369}$$

$$= \frac{1}{10}\times 79.81$$

$$= 7.98$$

NOTES

The standard deviation is a very useful device for comparingcharacteristics that may be different or expressed in different units of measurement. It is also used in describing the status or position of an individual in a group. But before this concept is developed further, it is essential to understand the nature of the 'normal probability distribution'.

14.5.2 Standard error or standard deviation of means

If we calculate mean for a large number of samples of same size from a population and prepare frequency polygons, we will notice that the standard deviation of the distribution of these means (σ_y') will be different from the standard deviation of the original population (σ_y). Standard deviation of means, also called standard error can be calculated as follows from data in Table.

$S'_y = \frac{S_y}{n}$ (n = sample size)

$S'_y = 1.58/\sqrt{5} = 0.7066$

This relationship indicates that the standard error will decrease as the sample size increase.

14.6 Analysis of Variance

14.6.1 ANOVA

The statistical technique used to compare means of variations of more than two populations is called **'analysis of variance (ANOVA).** The concept of

NOTES

ANOVA or analysis of variance was introduced by **R.A. Fisher.** ANOVA is based on two types of variations :

1. Variations existing within the sample
2. Variations existing between the sample

The ratio of these two variations is an indication of sample differences. Therefore, ANOVA helps in estimating whether more variations exist among the group means or within the groups. The ratio of these two variations is denoted by **'F'**. The **F- value** is the indication of differences between the samples.

14.6.2 Assumptions of Anova

ANOVA is based on following assumptions :

1. All ANOVA require random sampling.
2. The items for analysis of variance should be independent. But this may not be true in case there is correlation in time or space.
3. Equality of variances in a group of samples is an important precondition for several statistical tests.
4. The residual components should be normally distributed.
5. The variable of interest for each population or group has a normal probability distribution.
6. The variance associated with the variable of interest must be the same for each population or for each group of data.

14.6.3 Test of Anova

For performing the test of analysis of variance following statistical steps are carried out :

1. Sum of Squares (S.S.)

Sum of squares is computed as the sum of squares deviation of the values for mean of the sample. It means :

$$S.S._{\cdot} = \sum \left(X - \overline{X}\right)^2$$

$$S.S._{\cdot} = \sum x^2 - \frac{\sum x^2}{n}$$

Here, X represents sum of squares

$\bar{X}$ represents mean of sample

x represents deviation

n represents total number of observations.

2. Mean Square (M.S.)

NOTES

Mean square is the mean of all the sum of squares. It is obtained by dividing S.S. by appropriate degree of freedom (df) :

$$M.S. = \frac{\text{Sum of squares}}{\text{Degree of freedom}} = \frac{S.S.}{df}$$

Therefore, following three values or quantities are required for computing ANOVA :

1. Total sum of squares
2. Between sample sum of squares
3. Residual sum of squares

One Way Analysis of Variance

Simplest type of analysis of variance is known as **one way analysis of variance.**

In one way ANOVA only one source of variation (or factor) is investigated. But investigations are carried out in three or more samples simultaneously.

One way analysis of variance is used to test the null hypothesis that three or more treatments are equality effective.

Computation of Analysis of Variance (Anova)

For the computation of ANOVA the variance is estimated at two different levels :

1. Population variance between the groups
2. Population variance within the groups.

1. Between Groups Estimate of Population Variance

In this ANOVA procedure, the estimate of population variance σ^2 is calculated based on the variability between the k sample mean. Let us presume that in a given data :

Number of populations is represented by = k

Mean of each population is represented by = μ

Common population variance of the variable of interest is represented by = σ^2

NOTES

If we assume that a random sample of size n is selected from each of the populations (of numbers *k*), their **sample means** will be $\overline{X}_1 + \overline{X}_2 + \overline{X}_3 ..... \overline{X}_k$.

The sample mean corresponding to *j*th population will be $\overline{X}_j$ and the sampling distribution of $\overline{X}_j$ will have mean μ and variance $\sigma_{\bar{x}^2} = \frac{\sigma^2}{n}$

The mean μ of this sampling distribution is calculated by taking average of *k* sample means. This value of μ is represented as $\boxed{S_B^2 = \frac{\sum_j n_j \left(\overline{X} - \overline{\overline{X}}\right)^2}{k-1}}$ (double bar). It is calculated by the following formula :

$$\boxed{\overline{\overline{X}} = \frac{\overline{X}_1 + \overline{X}_2 + \overline{X}_3 ..... \overline{X}_k}{k} = \sum^{k} \overline{X}_j \quad \text{Here j} = 1}$$

The variance $\sigma_{\bar{x}^2}$ of the sampling distribution can be calculated by using the variance of individual sample means about their overall mean $\overline{\overline{X}}$ as follows :

$$\boxed{S^2_{\bar{x}} = \frac{\sum \left(\overline{X}_j - \overline{\overline{X}}\right)^2}{k-1}}$$

$$\text{or} \qquad \sigma^2_{\bar{x}} = \frac{\sigma^2}{n}$$

$$\text{or} \qquad \sigma^2 = n.\sigma^2_{\bar{x}} \quad \text{but } \sigma^2_{\bar{x}} \text{ can be represented as } S^2_{\bar{x}}$$

$$\text{It means} \qquad \sigma^2 = n.S^2_{\bar{x}}$$

$$= \frac{n.\sum \left(\overline{X} - \overline{\overline{X}}\right)}{k-1}$$

In ANOVA procedure σ^2 is also represented by the symbol S_B^2 , where *B* stands for the word **'between the group'.** It means :

$$S_B^2 = n.S_{\bar{x}}^2$$

When sample sizes are different, the overall sample mean $\overline{\overline{X}}$ is computed as weighted average of the individual sample mean as follows :

$$\overline{\overline{X}} = \frac{n_1\overline{X}_a + n_2\overline{X}_b + n_3\overline{X}_c \ldots\ldots n_k}{n_1 + n_2 + n_3 \ldots\ldots n_k}$$

(Overall sample mean)

NOTES

Between groups estimate of population variance is given by the following formula:

$$S_B^2 = \frac{\Sigma_j\, n_j \left(\overline{X} - \overline{\overline{X}}\right)^2}{k-1}$$

Here, S_B^2 = Population variance between the groups

Σ_j = Summation of $\overline{X}_a + \overline{X}_b \ldots\ldots\ldots \overline{X}_j$

n_j = Number of population from $n_a \ldots\ldots n_j$

$\overline{X}$ = Sample mean

$\overline{\overline{X}}$ = Overall sample mean

k = Total number of populations

2. Within Groups Estimate of Population Variance

In this ANOVA procedure, the estimate of population variance σ^2 is based upon the variation within the sample observations. It means variance is estimated from observations within their respective data sets. The estimate is denoted as S_w^2.

The subscript *w* refers to 'within'. Each of the *k* sample variances provides an estimate of variance within the population. For each sample, the sample variance is calculated in the usual fashion. It means variance of sample $j = S_j^2$ and $j = S_j^2$ calculated by the following formula :

$$S_j^2 = \frac{\Sigma\left(X_{ij} - \overline{X}_j\right)^2}{nj-1}$$

Here, $\overline{\overline{X}}$ S_j^2 is variance of sample *j*

X_{ij} is it item in sample j, and

n_j is the number of items in sample j.

The weighted average used to provide the within group estimates of variance is denoted by S_w^2. It is computed by the following formula :

NOTES

$$S_w^2 = \frac{\Sigma_j (n-1).S_j^2}{\Sigma_j n_j - k}$$

In case the sample size for each k populations is the same, i.e. $n_1 = n_2 = n_3 \ldots\ldots n_j = n_k$, then :

$$S_w^2 = \frac{\Sigma_j (n-1).S_j^2}{kn-k}$$

$$= \frac{(n-1)\Sigma S_j^2}{k(n-1)} = \frac{\Sigma S_j^2}{k}$$

Summary of Steps for Calculating ANOVA

Step 1. Calculate mean of each sample (*a to k*) and denote these means by

$$\overline{X}_a, \overline{X}_b, \overline{X}_c \ldots\ldots \overline{X}_k$$

Step 2. Take the average off all the means. It is denoted as $\overline{\overline{X}}$. It is calculated

by the formula :

$$\overline{\overline{X}} = \frac{\overline{X}_a + \overline{X}_b + \overline{X}_c \ldots\ldots \overline{X}_k}{k}$$

When sample sizes are different, the overall sample mean $\overline{\overline{X}}$ is computed as weighted average of individual means by the formula :

$$\overline{\overline{X}} = \frac{n_1\overline{X}_a + n_2\overline{X}_b + \ldots\ldots n_k\overline{X}_k}{n_1 + n_2 \ldots\ldots n_k}$$

Step 3. Calculate variance between the groups S_B^2 by using the following formula :

$$S_B^2 = \frac{\Sigma_j n_j \left(\overline{X} - \overline{\overline{X}}\right)^2}{k-1}$$

Here, k = the number of populations

n_j = the sample size of *j*th sample

$\overline{X}$ = Mean of a sample

$\overline{\overline{X}}$ = Average of all means

NOTES

When sample size is same for all the samples, the formula used for calculating variance between the group is :

$$S_B^2 = \frac{n\sum_j \left(\overline{X}_j - \overline{\overline{X}}\right)^2}{k-1}$$

Step 4. Calculate variance within the group S_w^2 by adding all sample variances and dividing the sum by total number of population, i.e.

$$S_w^2 = \frac{\sum S^2{}_j}{k}$$

If sample sizes are different, then use the following expression to calculate S_w^2

$$S_w^2 = \frac{\sum_j \left(n_j - 1\right) S^2{}_j}{\sum_j n_j - k}$$

Step 5. Calculate $\frac{S_B^2}{S_w^2}$

Step 6. Find the value of *fa* from the table with *k*-1 degrees of freedom in the numerator and *Sn-k* in the denominator.

Step 7. Compare the calculated value of *f* with table value of *fa*.

If $S_B^2 / S_w^2 > fa$ reject H_0

If $\frac{S_B^2}{S_w^2} \leq fa$ accept H_0

Example : The haemoglobin levels of three groups of children fed on three different diets are given in the following table 5:

NOTES

Table 5 : Haemoglobin levels (%) of children fed with three different diets

S. No.	Group I	Group II	Group III
1	11.6	11.2	9.8
2	10.3	8.9	9.7
3	10.0	9.2	11.5
4	11.5	8.8	11.6
5	11.8	8.4	10.8
6	11.8	9.1	9.1
7	12.1	6.3	10.5
8	10.8	9.3	10.0
9	11.9	7.8	12.4
10	10.7	8.8	10.7
11	11.5	10.0	-
12		9.7	-
	$t_1 = 124.0$	**$t_2 = 107.5$**	**$t_3 = 106.1$**

Test whether the means of these three groups differ significantly

Solution :

No. of observations $n_1 = 11$ $n_2 = 12$ $n_3 = 10$

Total of all samples $t_1 = 124.0$ $t_2 = 107.5$ $t_3 = 106.1$

Sample Mean $\overline{X}_1 = \frac{124}{11} = 11.27$ $\overline{X}_2 = \frac{107.5}{12} = 8.96$

$$\overline{X}_3 = \frac{106.1}{10} = 10.61$$

$S_1^2 = 124 \times 11.27$ $S_2 = 107.5 \times 8.96$ $S_3 = 106.1 \times 10.61$

$= 1397.48 \qquad = 963.2 \qquad = 1125.72$

$$N = n_1 + n_2 + n_3$$

$$= 11 + 12 + 10 = 33$$

$$\Sigma X \text{ or } T = t_1 + t_2 + t_3$$

$$= 124.0 + 107.5 + 106.1 = 337.6$$

Common Mean = $\overline{\overline{X}} = \dfrac{124.0+107.5+106.1}{33} = \dfrac{337.6}{33} = 10.23$

The common mean $\overline{\overline{X}}$ of all the sample is 10.23 whereas mean of sample $T_1\left(i.e.\ \overline{X}_1\right)$ is 11.27, of Sample $T_2\left(i.e.\ \overline{X}_2\right)$ is 8.86 and mean of sample $T_3\left(i.e.\ \overline{X}_3\right)$ is 10.61. We have to see whether the mean levels of 3 different samples in comparison to common mean $\overline{\overline{X}}$ show significant variance.

14.7 t-Test (Two Sample Means)

If two samples are available and their means calculated, one is always interested in finding whether these samples came from same population of iron two populations having different means. In such a case, the standard error needs to be calculated as follows :

$$S_{\bar{y}_1-\bar{y}_2} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \quad \left(\begin{array}{l}\text{where } S_1^2 = \text{ variance of sample 1} \\ \text{and } S_2^2 = \text{ variance of sample 2}\end{array}\right)$$

or $$S_{\bar{y}_1-\bar{y}_2} = \sqrt{S^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \qquad (\text{where } S_1^2 = S_2^2 = S^2)$$

or $$S_{\bar{y}_1-\bar{y}_2} = \sqrt{S_1^2 + S_2^2/n} \qquad (\text{if } n_1 = n_2)$$

or $$S_{\bar{y}-\bar{y}_2} = \sqrt{\frac{2S^2}{n}} \qquad \left(\begin{array}{l}\text{where, } S_1^2 = S_2^2 = S^2 \\ \text{and} \qquad n_1 = n_2 = n\end{array}\right)$$

Therefore,

$$t_{(n_1+n_2-2)} = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \qquad (df = n_1 + n_2 - 2)$$

NOTES

NOTES

$$\text{or} \qquad t_{(n_1+n_2-2)} = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{S^2(n_1+n_2)/n_1 n_2}}$$

$$\left(\text{where } S_1^{\,2} = S_2^{\,2} = S_2; n_1 \neq n_2\right)$$

$$\text{or} \qquad t_{(n_1+n_2-2)} = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{S_1^{\,2} + S_2^{\,2}/n}} \qquad (\text{where } n_1 = n_2 = n)$$

$$\text{or} \qquad t_{(n_1+n_2-2)} = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{2S^2}{n}}} \qquad (\text{where } S_1^{\,2} = S_2^{\,2} = S^2;\ n_1 = n_2 = n)$$

Example --Following are the counts of plants of a species at 15 different places randomly taken in a specified area (like college campus) in May and June. Ascertain if the number of plants remained constant in two months.

May	37	27	60	44	22	26	10	38	15	24	10	33	23	142	34
June	6	9	2	6	6	4	18	9	13	14	10	8	1	4	7

Note : 15 places were taken at random in both the months

$$H_0 : \mu_1 = \mu_2; H : \mu_1 \neq \mu_2$$

$$t_{(n_1+n_2-2)} = \frac{\bar{y}_1 - \bar{y}_2}{\sqrt{\frac{S_1^{\,2} + S_2^{\,2}}{n}}}$$

$$\sum y_1 = 545; \bar{y}_1 = 36.333; \sum y_1^{\,2} = 34177; \left(\sum y_1\right)^2 / n = 19801.7$$

$$\sum y_2 = 117; \bar{y}_2 = 7.800; \sum y_2^{\,2} = 1209 \left(\sum y_1\right)^2 / n = 912.6$$

$$S_1^{\,2} = \frac{34177 - 19801.7}{14} = 1026.8$$

$$S_2^{\,2} = \frac{1209 - 912.6}{14} = 1.4$$

$$t_{2n-2} = \frac{36.333 - 7.800}{\sqrt{\frac{1026.8 + 1.4}{15}}} = \frac{28.533}{8.27} = 3.45 \qquad \begin{pmatrix} t_{28} = 2.048 \text{ at } 5\% \\ t_{28} = 2.763 \text{ at } 1\% \end{pmatrix}$$

The estimated t value (3.45) is more than the tabular value of t for 28 *df* even at 1% level (t-2.763); therefore it is highly significant so that H_0 is rejected and we conclude that the number of plants did not remain constant in the two months.

14.8 Chi-Square (χ^2) Tests

Chi-square test is most frequently used for testing whether the observed results are in agreement with the hypothetical results expected on the basis of sound principles, e.g. Mandel's law in genetics. If theoretical ratios are given, we try to find the goodness of fit of the data to the theoretical ratio. The chi-square (χ^2) statistics is calculated as follows :

NOTES

$$\chi^2 = \Sigma\left(\frac{(O-E)^2}{E}\right),$$

where O = observed frequency in a class

E = expected frequency in a class

Σ = summation over all classes

From this equation, the value of χ^2 will be zero, if O = E in each class, but due to chance error, this never happens and the observed frequencies differ from the expected frequencies, even if the observed results are based on the same principle from which expected frequencies are calculated. The χ^2 calculated from different samples of the same population has a distribution which is given in Table 4 given in Appendix 4. The values of χ^2 depend on number of classes or in other words on the number of degrees of freedom (n-1) and on the level of probability, so that if we know the degrees of freedom (*df)* and the critical level of probability (α= 5%or 1%), we can find the expected value of χ^2 from the table. This expected value can be compared with the value calculated from data. If the tabular value is lower calculated from data. If the tabular value is lower, the results are significant (observed results are not in agreement with the expected results), otherwise these are insignificant (observed results agree with expected results).

Chi-square (χ^2) test for independence

(2x2 contingency table)

Another very important use of χ^2 test is to test independence of two separate events or occurrences that are classified in the form of what we call *contingency tables.* A contingency table is the two way classification of data as shown in the following example.

NOTES

A plant heterozygous for two genes (AaBb) gives progeny which can be classified into four phenotypic ckasses (AB, Ab, aB, ab) which are expected to segregate in 9 AB : 3 Ab : 3aB : 1ab on the basis of independent segregation and assortment of the two genes, which are not linked, Following are the observed results :

AB : A*b* : aB : a*b*

102 : 16 : 35 : 7

Determine the goodness of fit to 9 : 3 : 3 : 1 ratio.

Observed frequencies can be set out in a contingency table as follows, where expected values are given in brackets. The χ^2 can be calculated either by using an equation available for contingency table, or by using the formula $\chi^2 = \Sigma\left\{(O-E)^2/E\right\}$.

	Segregation for A			
		A	a	Total
Segregation for B	B	102 a (90)	35 c (30)	137 r_1
	b	16 b(30)	7 d(10)	23 r_2
	Total	**118** c_1	**42** c_2	**160**

c_1 and c_2 are columns ; r_1 and r_2 are rows

(i) $$\chi^2_{(1)} = \frac{\left\{\left|ad-bc\right|-n/2\right\}^2 n}{r_1 r_2 c_1 c_2}$$

note : straight lines on either side of *ad* – *bc* mean that it will always be used as +ve value.

(degree of freedom = (r-1) (c-1) where R = number of rows ; c = number of columns)

$$= \frac{(154-80)^2 \times 160}{137 \times 23 \times 118 \times 42} = \frac{(74)^2 \times 160}{137 \times 23 \times 118 \times 42} = 0.056(for\ 1\ df)$$

The value of chi-square for one *df* at 5% level is 3.84; therefore the results are not significant and suggest that the 2 genes A and B, function independently.

(ii) $$\chi^2_{(3)} = \frac{(102-90)^2}{90} + \frac{(35-30)^2}{30} + \frac{(16-30)^2}{30} + \frac{(7-10)^2}{10}$$

$$= 1.60 + 0.83 + 6.53 + 0.90 = 9.86 \text{ (for 3 } df\text{)}$$

NOTES

The value of chi-square for 3 *df* at 5% level is 7.81 ; therefore the results are significant and the data do not give a good fit with the ratio 9 : 3 : 3 : 1.

In order to know the reason for this discrepancy, one may like to partition the chi-square into three separate chi-square values (for A-a segregation; for B-b segregation and A-B linkage) as follows :

(1) $\chi^2_{(1)}$ for A *vs* $a = \chi^2_{(1)} = (a+b-3c-3d)^2 / 3n$

$$= (102 + 16 - 105 - 21)^2 / 480 = 0.1333$$

This chi square value is insignificant suggesting normal segregation of A and a.

(2) $\chi^2_{(1)}$ for B *vs* $b = \chi^2_{(1)} = (a-3b+c-3d)^2 / 3n$

$$= (102 + 48 + 35 - 21)^2 / 480 = 9.63$$

This chi square value is significant, suggesting that B and b are not segregating as expected.

(3) $\chi^2_{(1)}$ for linkage between A and B

$$= (a - 3b - 3c + 9d)^2 / 9n$$

$$= (102 - 48 - 105 + 63)^2 / 1440 = 0.10$$

This value of χ^2 is not significant suggesting no linkage of A and B.

14.9 Correlation and Regression

14.9.1 Defination

According to **Connor** 'if two or more quantities vary in sympathy and the movements in one tend to be accompanied by corresponding movements in the other, then these two quantities are said to be correlated

Correlation analysis- we are concerned whether two variables are independent or they vary together in positive or negative direction. In correlation the two variables are not related as independent and dependent variables. It means in

correlation both the variables are affected by a common cause and the degree to which these variables vary together is estimated. The concept of correlation analysis and term correlation originated with **Galton** in 1888

NOTES

Regression Analysis,-- the dependence of one variable on another variable is determined. Therefore, the two variables are related as independent and dependent variables.

Regression analysis is employed to predict or estimate the value of one variable corresponding to a given value of another variable. Regression equations are applied to determine changes in Y due to changes in X variable.

14.9.2 Significance of Correlation

The study of correlation is of great significance in practical life, because of the following reasons :

1. The study of correlation enables us to know the nature, direction and degree of relationship between two or more variables.
2. Correlation studies help us to estimate the changes in the value of one variable as a result of change in the value of related variables. This is called **regression analysis.**
3. Correlation analysis helps us in understanding the behaviourof certain events under specific circumstances. For example, we can identify the factors for rainfall in a given area and how these factors influence paddy production.
4. Correlation facilities the decision making in the business world. It reduces element of uncertainty in decision-making.
5. It helps in making predictions.

14.9.3 Types of Correlation

Depending on its extent and direction the correlation between two variables may be of following types :

1. Positive and Negative Correlations

The positive and negative correlations are based on the direction of change in the value of two variables :

1. **Perfect Positive Correlation :** When two variables move proportionately in the same direction, i.e., the increase in the value of

one variable leads to corresponding increase in the values of other variable, the correlation between them is called **perfect positive**. For example, increase in body weight with the increase in height presents positive correlation. It is also called **direct correlation**.

NOTES

2. **Moderately Positive Correlation :** When two variables are partially positively correlated the correlation is termed **moderately positive correlation**, e.g., tallness of plants and the quantity of manure used etc.

3. **Perfect Negative Correlation :** The two variables show negative correlation when one variable increases with a constant interval and another decreases with constant interval. Thus, variables deviate in opposite directions. This is also called **inverse correlation.** Examples of perfect negative correlation are very rare in nature but some approaching to that extent are temperature and lipid content of the body , etc.

4. **Moderately Negative Correlation :** When two variables are partially negatively correlated, the correlation is termed as **moderately negative correlation.** e.g., economic condition of state and cases of tuberculosis, income and infant mortality rate, etc.

5. **Absolutely no Correlation :** When two variables are completely independent of each other, the correlation is termed as **absolutely no correlation,** e.g., body, weight and I.Q. In this case no imaginary mean line is formed which could indicate the trend of correlation.

2. Linear and Non-linear Correlations

The correlation can also be classified as linear or nonlinear on the basis of ratio of variations in the related variables :

1. **Linear Correlation :** Correlation between two variables is said to be linear if there is some constant relationship between the two variables. When the values of two variables are plotted as points in the XY plane, a straight line is formed
2. **Non-Linear Correlation :** The relationship between two variables is said to be **nonlinear** or **curvilinear** if corresponding to a unit change in one variable, the other variable does not change at the same constant rate but fluctuates.

3. Simple, Partial and Multiple Correlation

Based on the number of variables involved, the correlation may be of following three types :

NOTES

1. **Simple Correlation :** In simple correlation only two variables are involved. Therefore, in simple correlation the relationship is between two variables such as intelligence of students and their performance (marks) in the examination.
2. **Multiple Correlation :** In multiple correlation relationship between three or more variables is studied. Simultaneous study of relationship between yield of wheat per acre, the amount of rainfall and the amount of fertilizer applied are the examples of multiple correlation.
3. **Partial Correlation :** In partial correlation, relation between more than two variables is considered but correlation is studied only between two variables. Other variables are assumed to be constant. For example, the correlation between the amount of fertilizers and the yield of wheat per acre is partial correlation in case rainfall is assumed to be normal.

14.9.4 Measures Of Correlation

Correlation analysis measures the degree of association of two variables. Following methods are used to measure the correlation between two variables :

1. Scatter diagram method
2. Karl Person's coefficient of correlation
3. Spearman's rank correlation coefficient

1. Scatter Diagram Method Or Scatter Plot Method

Scatter diagram is the simplest method of studying relationship between two variables. It is in the form of graphic representation of degree and direction of correlation between two variables.

Say we take two variables X and Y for n number of samples (X_1, X_2, X_3, …..X_n) and plot X_1 against Y_1 as a dot (.) in the XY-plane, the diagram of dots so obtained is known as **scatter diagram** or **dot diagram**. It is customary to take dependent variable along Y-axis, i.e., along vertical axis and independent variable along X-axis or horizontal axis. Placement of dots on the graph reveals whether the changes in the variable are in the same direction or in opposite direction.

2. Karl Pearson's Correlation Coefficient

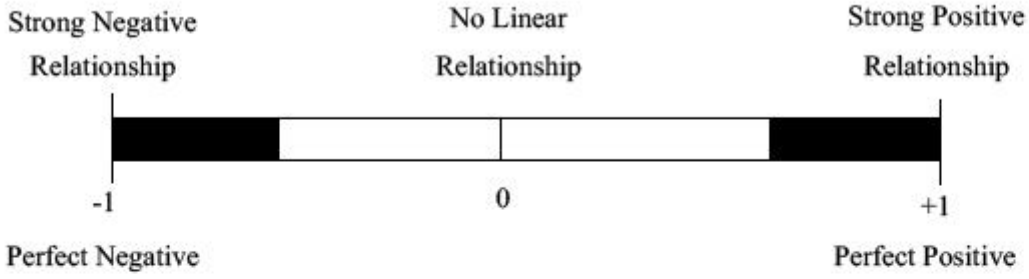
A scattered diagram, like a histogram, is a convenient way of displaying existence of correlation and direction of correlation, but is does not give any correlation value. To measure correlation, the coefficient of correlation is worked out by Karl Pearson's Coefficient of Correlation.

NOTES

Coefficient of correlation is the degree to which two variables are inter-related. T is mathematical method of measuring the tendency of linear relationship between two variables. This was introduced by **Karl Pearson** (1867-1936). This measure of correlation is also known as **Pearsonian Correlation Coefficient**. If two variables are denoted by X and Y, the coefficient of correlation between them is represented by rxy or r. Pearson's coefficient correlation method can be used to measure correlation for individual series as well as for grouped data.

14.9.5 Properties of Coefficient of Correlation

1. Coefficient of correlation is a measure of closeness between two variables.
2. The correlation may be positive or negative.
3. The range of correlation coefficient is from -1 to +1.
4. If r = +1, the correlation between two variables is **perfect** and **positive.**
5. If r = -1 the correlation is perfect and negative.



The Range of Correlation Coefficient

If there is a strong positive linear relationship between two variables, the value of r will be close to +1.

6. If there is strong negative linear relationship between the variables, the value of r will be close to -1.
7. If r = 0, there is no correlation between two variables. It means variables are independent.

The value of r is calculated by using following formula :

$$r = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma(x-\bar{x})^2\,\Sigma(y-\bar{y})^2}} \quad \text{or} \quad \frac{\Sigma x - \bar{x}\Sigma y}{\left(\sqrt{\Sigma x^2 - \bar{x}\Sigma x}\right)\left(\sqrt{\Sigma y^2 - \bar{y}\Sigma y}\right)}$$

NOTES

Computation of Coefficient of Correlation from Ungrouped Data

Coefficient of correlation for ungrouped data is calculated by the following formula :

$$r = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma(x-\bar{x})^2\,\Sigma(y-\bar{y})^2}} \quad \text{or}$$

$$r = \frac{\Sigma\left(\frac{x-\bar{x}}{\sigma x}\right)\left(\frac{y-\bar{y}}{\sigma y}\right)}{n}$$

Here, x and y are measurements for two variables,

$\bar{x}$ and $\bar{y}$ are means of two distributions of measurements and

σx and σy are standard deviations of measurements.

Example : For five patients, temperature (x) and pulse (y) are given. The correlation of coefficient for these two measurements can be calculated as follows

Patient	**Temperature X**	**Pulse y**	$\left(\frac{x-\bar{x}}{\sigma x}\right)$	$\left(\frac{y-\bar{y}}{\sigma y}\right)$	$\left[\frac{x-\bar{x}}{\sigma x}\right]\left[\frac{y-\bar{y}}{\sigma y}\right]$
A	102	100	$\frac{102-100}{\sqrt{2}} = \frac{2}{\sqrt{2}}$	$\frac{100-80}{\sqrt{200}} = \frac{20}{\sqrt{200}}$	$\frac{40}{\sqrt{400}}$
B	101	90	$\frac{101-100}{\sqrt{2}} = \frac{1}{\sqrt{2}}$	$\frac{90-80}{\sqrt{200}} = \frac{10}{\sqrt{200}}$	$\frac{10}{\sqrt{400}}$
C	100	80	$\frac{100-100}{\sqrt{2}} = \frac{0}{\sqrt{2}}$	$\frac{80-80}{\sqrt{200}} = \frac{0}{\sqrt{200}}$	$\frac{0}{\sqrt{400}}$
D	99	70	$\frac{99-100}{\sqrt{2}} = \frac{-1}{\sqrt{2}}$	$\frac{70-80}{\sqrt{200}} = \frac{-10}{\sqrt{200}}$	$\frac{10}{\sqrt{400}}$

E	98	60	$\frac{98-100}{\sqrt{2}}=\frac{-2}{\sqrt{2}}$	$\frac{60-80}{\sqrt{200}}=\frac{-20}{\sqrt{200}}$	$\frac{40}{\sqrt{400}}$
Total	$\Sigma x = 500$	$\Sigma x = 400$	0	0	$\frac{100}{\sqrt{400}}$

NOTES

$\bar{x} = 500/5 = 100$

$\bar{y} = 400/5 = 8$

$S_x = \sqrt{2} \qquad S_y = \sqrt{200}$

$$r = \frac{\left(\frac{x-\bar{x}}{\sigma x}\right)\left(\frac{y-\bar{y}}{\sigma y}\right)}{n}$$

$$= = \frac{100/\sqrt{400}}{5} = \frac{100/20}{5} = \frac{5}{5} \qquad = 1$$

When coefficient of correlation's value is +1, it means there perfect and positive correlation, i.e. the two variables, i.e., temperature and pulse rate change in the same direction.

14.9.6 Regression

In analyzing data, we find that it is frequently desirable to learn something about the relationship between two variables. For example, we may be interested in studying the relationship between blood pressure and age, height and weight, concentration of an injected drug and heart rate, the consumption level of some nutrient and weight gain, the intensity of a stimulus and reaction time or total family income and medical care expenditures. The nature of relationship between variables such as these may be examined by **regression analysis..3**

The term regression was coined by **F**. Galton in 1885 to explain the data obtained during the study oh inheritance. **Galton** observed the height of off springs during a few generations of a family and came to the conclusions that the height of off springs tend to occupy median position. He expressed the regression as **'the tendency to remain towards central position.**

NOTES

Types of Regression Analysis

The regression analysis can be of two types : **Simple** and **multiple**.

1. **Simple Regression :** The regression analysis confined to the study of only two variables at a time is termed as **simple regression.**
2. **Multiple Regression :** The regression analysis for studying more than two variables at a time is known as **multiple regression.**

Regression Lines and Linear Regression

When observations from two variables are plotted as a graph, and if the points so obtained fall in a straight line then the relationship is linear and it is said that there is **linear regression** between the variables under study. However, if the line is not a straight line, the regression is termed as **non-linear regression.**

When the points are obtained on a scattered diagram, the process of deciding the line of the best fit to summarize a particular set of points on a graph is called regression analysis. This is worked out by deriving an equation called **regression equation.**

Regression Equation

The equation that describes position of any line on a graph is called **regression equation.** For a linear regression, the equation for a dependent variable XY against independent variable x can be given as follows :

$$y = a + bX$$

Here, values of 'a' and 'b' are constant and are fixed for a particular line. If the values of 'a' and 'b' are known, y can be obtained for any corresponding value of x. The values of 'a' and 'b' are calculated by the following equation :

$$b = \frac{\Sigma(X-\bar{X})(y-\bar{y})}{\Sigma(X-\bar{X})^2} = r\frac{(\text{SD of y})}{\text{SD of x}} \text{ or } = r\frac{\sigma_x}{\sigma_y}$$

or $$b = \frac{\Sigma xy - \bar{X}\Sigma y}{\Sigma X^2 - \bar{X}\Sigma X}$$

After obtaining the value of 'b' the value of 'a' can be calculated.

The constant 'a' is known as intercept, and denotes the value of y when the value of x is zero.

The constant 'b' measures the slope of the line and is called "regression coefficient". The constant 'b' gives an idea of that how change occurs in variable y when the variable X varies by 1 unit. For instance, if the value of 'b' is 5.8, then a change in X by one unit will bring out a change in y by 5.8 units. The positive value of 'b' indicates the increase in the value of y. It is associated with the increase of X while a negative value will tell the decrease in y with an increase in X.

NOTES

$$X = a' + b' y$$

Procedure

1. Plot a graph between two variables taking independent variable of X-axis and dependent variable on Y-axis.
   Find out the values of a and b using the equations given earlier. For drawing the line of best fit (regression line), find out any two values of y associated with corresponding value of x by using the equation :
   $$y = a + bx$$
2. Plot these two values on the graph on which all the points of original values have been put.
3. Make a straight line intersecting through these two points to get the fittest regression line.

Example : Find out regression equation from the following data for 7 fishes of a species :

X	13.4	15.1	15.3	16.8	17.5	19.2	21.2
Y	2.1	2.3	2.3	2.6	2.7	3.0	3.3

Solution :

The following table is prepared first :

X	Y	X.Y	X^2
13.4	2.1	28.14	179.56
15.1	2.3	34.73	228.01
15.3	2.3	35.19	234.09

NOTES

16.8	2.6	43.68	282.24
17.5	2.7	47.25	306.25
19.2	3.0	57.6	368.64
21.2	3.3	69.96	449.44
$\Sigma X = 118.5$	$\Sigma Y = 18.3$	$\Sigma X.Y = 316.55$	$\Sigma X^2 = 2048.23$

Equation (i) $\Sigma y = \Sigma x.a + \Sigma x.b$ or $y = a + bx$

Equation (ii) $\Sigma xy = \Sigma x.a + \Sigma x^2.b$ or $xy = x.a + b.x^2$

Putting the values in the above formula :

$18.3 = a + 118.5b$ .....(i)

$\Sigma y = 18.3$

$\Sigma x = 118.5$

$316.55 = 118.5a + 2048.23b$ ....(ii)

$\Sigma XY = 316.55$

$\Sigma X = 118.5$

$\Sigma X^2 = 2048.23$

Multiply equation (i) with 118.5

$= 18.3 \times 118.5 = 118.5a + 118.5 \times 118.5b$

$2168.55 = 118.5a + 14042.25b$ ..... (iii)

Subtracting equation (ii) from (iii)

$(2168.55 - 316.55) = (118.5 - 118.5)\, a + (1042.25 - 2048.23)b$

$(2168.55 - 316.55) = 0 + (14042.25 - 2048.23)b$

$\therefore \quad b = \dfrac{2168.55 - 316.55}{14042.25 - 2048.23} = \dfrac{1852}{11994.02} \quad = b = 0.16$

Now put the value of b in equation (i)

$18.3 = 7a + 0.16 \times 118.5$

$$18.3 = 7a + 18.96$$

$$7a = 18.3 - 18.96 = -0.66$$

$$a = \frac{-0.66}{7} = -0.0943$$

Therefore, the Regression equation $y = a+bx$

$$Y = - 0.0943 + 0.16 x$$ **Ans**

NOTES

14.10 Probability

Defination-It is defined as "the ratio of number of times a particular event occurs to the total number of trials during which the event could have happened".

We come across the concept of probability in our daily life when we use expressions 'probable' or 'unlikely' for future events. In these cases, by knowing the situation, we know the possible future outcomes. One popular example is the tossing of coin. If the coin is unbiased (newly minted coin), the probability that on tossing, it will fall head i.e., P(H) is 0.5, and the probability that it will fall tail i.e. P(T) is also 0.5. In other words, if we toss the coin 100 times, approximately 50 heads and 50 tails will be obtained. Similarly, in biology if we count the number of children born in a hospital in a year, approximately 50% male and 50% females will be available, because the probability of a male or a female child being born to a mother are equal i.e. P(M) = P(F) = 0.5, although in families of 2 children, 2 females, 2 males or 1 male + 1 female children will have different probabilities (see later). Similarly, in genetics, in a heterozygote Aa, the gametes with A and a occur with equal probability P(A) = 0.5 and P(a) = 0.5.

Mutually Exclusive Events and Exhaustive Set

There may also be cases where, if we toss a dice, six possible octcomes (1, 2, 3, 4, 5, 6), are there, each with a probability P =1/6. In all these cases, when one event occurs, the other cannot occur. Such a set of events is called *mutually exclusive events*, so that if one event occurs, no one other event can occur. This set is also called exhaustive, since no other possibility exists. Therefore, a set of head and tall in one coin is an exhaustive set of mutually exclusive events. Similar is the case with a set of six events of a dice.

NOTES

In all above cases, if a trial has 'n' possible, equally likely and mutually exclusive outcomes and if 'n' is the number of favorable outcomes, then the probability of the favourable outcomes will be P=n'/n. When probabilities of a certain set of events are available prior to conducting the experiment, these are conveniently described as a priori probabilities. The probabilities are not equally likely or when the number of possible events is not certain. This will happen when the coin is damaged resulting in a baised coin with $P(H) \neq P(T) \neq 0.5$. Similarly in genetics, if the expression of a gene is influenced by another gene, the probability of its phenotypic expression will not be known and will not be an a priori probability. In all such cases we need to estimate the probability through an experiment. Therefore, we should recognize two types of situations : (i) where a priori probability is available and (ii) where probability is not available a priori and the same will have to be estimated through an experiment.

Random Samples Or Randomization

If any experiment, where statistics is to be used, drawing a random sample or randomization of treatments (e.g. tossing of coin, throwing of dice; or taking a sample of plants from a population; arrangement of crop varieties in an experimental plot) is an essential requirement, to get an unbiased estimate of the different parameters or statistics. Estimation of probability, when probability is not known a priori also requires randomization. Randomization is therefore, sometimes defined as a requirement to get an unbiased estimate of experimental error, so that it is analogous to insurance, providing precaution against disturbances that may or may not occur, and which may or may not be serious, if they do occur.

We will discuss more about probability and its estimation under deferment situations i.e. normal and binomial distributions, since different methods of computation of probability are needed in different kinds of distributions.

14.11 Normal and Binomial Distributions

Normal distributions and binomial distribution are both extensively used in biology and in both cases probability is available a priori from the nature of the events. While normal distribution usually applies to continuous distribution or

to variable like height or body weight of individuals (sometimes it also applies to the attributes like number of seeds per plant etc.), the binomial distribution is used for discrete variable like tossing of coins, number of male and female children in a family of a definite size, inheritance of individual traits etc. These two distributions will be discussed here with particular reference to their use in biological sciences.

NOTES

14.11.1 Normal Distribution

The use of normal distribution to biological data was pioneered by Sir Frances Galton (1822-1911), who studied the quantitative characters and their inheritance. The word normal in the term *"normal distribution"* is not intended to distinguish it from abnormal, but has a specific interpretation that will be discussed in this section.

The normal distribution is often characterized by the bell shaped appearance of the curve that is described as *normal curve* or *Gaussian curve*, which means it is a symmetrical distribution with equal number of observations on either side of mean show in figure. The observations on either side of the mean are so distributed that in the centre, the curve is concave downwards, but at a distance of one standard deviation on either side of mean i.e. at the points $\mu \pm \sigma$ (called points of inflection), the curve becomes concave upwards is show in figure. The normal distribution is actually characterized by the following equation, where f(y) represents the frequency of a particular value of the variable y.

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e - \left\{ (y-u)^2 / 2\sigma^2 \right\}$$ where π and e are constants

(π = 22/7 = 3.14159 and e = 2.71828)

The shape of the normal curve implies that observations occur most frequently in the vicinity of mean value and gradually decrease as we go away from the mean value, so that the area covered under $\mu \pm \sigma$ represents 68% of the total area, the area under $\mu \pm 2\sigma$ represents 95% and that under $\mu \pm 3\sigma$ represents complete area under the curve, which is regarded to be unity for convenience. The probabilities of the variable Y falling in different regions can then be expressed as follows and can be obtained from normal distribution tables

prepared using the equation given above assuming mean, $\mu_i = 0$ and $\sigma=1$ (see table in Appendix 1).

NOTES

$P\{(\mu - \sigma) \leq Y \leq (\mu + \sigma)\} = 0.6827$

$P\{(\mu - 2\sigma) \leq Y \leq (\mu + 2\sigma)\} = 0.9545$

$P\{(\mu - 3\sigma) \leq Y \leq (\mu + 3\sigma)\} = 0.9973$

In the equation given above, it can be shown that f(y) is maximum, when $y = \mu$, so that $e^{-\{(y-\mu)^2/2\sigma^2\}} = e^0 = 1$ and $(fy) = 1/\sigma\sqrt{2\pi}$. Further, if σ increases, f(y) decreases. This would mean that the normal curve flattens as σ increases. For a distribution with $\mu=0$ and $\sigma = 1$ or in other words for a distribution of $Z = (y-\mu)/\sigma$ (Z is called the normal deviate), the probabilities are given in Table in Appendix 1, which is popularly called single tail distribution, since only area for specified value of Z on one side of μ has been considered. If the situation demands that both tails be considered, then the probability values of the table need to be doubled. You will notice that left marginal column has values to one decimal place, and second decimal place is given at the top of the table. The body of the table contains areas under the curve which represents probabilities associated with the tail of the distribution. Sometimes the normal distribution table may be given with probabilities in the marginal column and top row with the Z values in the body of the table (Table in Appendix 2). The use of the table will be illustrated with the help of the following example.

Example : The mean length of the ear in a wheat variety is 9.978 cm and the standard deviation is 1.441 cm. Calculate the probability of the occurrence of –

(a) An ear having a length of 12.128 cm. or more,

(b) An ear of 6.536 cm. or less,

(c) An ear deviating from the mean by $\pm$ 2.581 cm.

Solution : (a) We first calculate Z (normal deviate) as follows

$$Z = \frac{x-\mu}{\sigma} = \frac{12.128 - 9.978}{1.441} = \frac{2.150}{1.441} = 1.492$$

$\because P(Z \geq 1.49) = 0.0681$ $\quad\quad$ $\therefore P(Z \geq 1.492) = 0.0681 - (.0013) \times \frac{2}{10}$

$P(Z \geq 1.50) = 0.0668$ $= 0.06784$

Difference .0013 $= 0.06784 \times 100 = 6.784\%$

Therefore, the probability of the occurrence of an ear of the length 12.128 cm or more will be 6.784%.

NOTES

(b) The normal deviate is calculated again

$$Z = \frac{6.356 - 9.978}{1.441} = \frac{3.442}{1.441} = 2.389$$

$\because P(Z \leq 2.38) = .0087$ $\therefore P(Z \leq 2.389) = .0084 + \frac{.0003}{10}$

$P(Z \leq 2.39) = .0084$ $= .00843 = 0.843\%$

Difference : .0003

Therefore, the probability of the ear length being 6.536 or less is 0.843% only (much lower than that for the ear length to be 12.128 cm. or more).

(c) The value given is $x = \mu \pm 2.581$, so that

$$Z = \frac{x - \mu}{\sigma} = \frac{2.581}{1.441} = 1.791$$

$\because P(Z \geq 1.79) = 0.0367$

$P(Z \geq 1.80) = 0.0359$

.0008

$$\therefore P(Z \geq 1.791) = 0.0359 + \frac{.0008}{10} = 0.03598$$

Since in this case both the tails will have to be included, the probability will be $2 \times 0.03598 = .07196 = 7.196\%$ or 7.2%

14.11.2 Binomial Distribution

A distribution having only two possible outcomes each with a known probability (whether a priori or not) is called binomial distribution. The most popular example of binomial distribution is tossing of a coin with head and tail with equal probabilities i.e. P(H) = P(T) = 0.5. Sometimes a distribution with more than two outcomes may also be converted into a binomial distribution having two outcomes. For instance a dice has six faces outcomes (1,2,3,4,5,6), but we may be interested in finding out the probability of success P(S) = 1/6

and probability of failure is P(F) = 5/6. Also in binomial distribution, if the probability of one outcome (success) is known to be p, the probability of the second outcome (failure) will always be (1-p).

NOTES

Binomial distribution and probabilities

If we toss a coin 5 times, then we may have 0,1,2,3,4 or 5 heads and remaining tales i.e. 0H, 5T ; 1H, 4T ; 2H, 3T; 3H, 2T; 4H, 1T or 5H, 0T. If the experiment is repeated 100 times, frequencies of six possible outcomes can be recorded and compared with the theoretical distribution. In other words from a theoretical distribution of $(½+½)^5$ we can calculate the theoretical probabilities of all six outcomes listed above as follows :

$$(p+q)^n = \binom{n}{0}p^0q^n + \binom{n}{1}p^1q^{(n-1)} + \binom{n}{2}p^2q^{(n-2)} \ldots \binom{n}{n}p^nq^0$$

As an example of 5 tosses as above

$$\left(\tfrac{1}{2}+\tfrac{1}{2}\right)^5 = \binom{5}{0}\left(\tfrac{1}{2}\right)^0\left(\tfrac{1}{2}\right)^5 + \binom{5}{1}\left(\tfrac{1}{2}\right)^1\left(\tfrac{1}{2}\right)^4 \ldots\ldots \binom{5}{5}\left(\tfrac{1}{2}\right)^5\left(\tfrac{1}{2}\right)^0$$

$$= (1)\,(½)^5 + 5(½)^5 \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)\,(½)^5$$

$$= 1/32 + 5/32 + 10/32 + 10/32 + 5/32 + 1/32 = 1$$

Thus the theoretical probabilities of binomial distribution are as shown in Table 4.1.

Table : Probability distribution in 5 tosses of an unbiased coin

	Number of heads					
	0 (0H, 5T)	1 (1H, 4T)	2 (2H, 3T)	3 (3H, 2T)	4 (4H, 1T)	5 (5H, 1T)
Probability (expected)	1/32 (0.03125)	5/32 (0.15025)	10/32 (0.31250)	10/32 (0.31250)	5/32 (0.15025)	1/32 (0.03125)

Mean and standard deviation in binomial distribution

If in an experiment, involving binomial distribution, if we know values of n, p and q, then we can calculate the expected mean $\begin{matrix}\overline{y}_i = np; \\ \overline{y}_j = nq;\end{matrix}$ $(\overline{y})$ of success or failure and the standard deviation (S)

NOTES

$\overline{y}_i = np;$	$\overline{y}_i = np;$
(success)	(failure)

$$S = \sqrt{npq}; S^2 = npq$$

14.12 Summary

In this unit we studied the various terms and symbol which are important in biostatics, In this unit we discussed the nature of quantitative and qualitative data, the various methods of representing the quantified data graphically. The unit further elaborated the measures of central tendency, mean, median and mode .We also discussed the variance, standard deviation which are commonly used measures of variability. We also describe the application of t-test and Chi-square test and ANOVA. We further elaborated the correlation and regression, probability distribution.

14.13 Glossary

- **Quantitative Data** : data which are expressed in nominal, ordinal, interval or ratio scales of measurement.
- **Qualitative Data :** data which are available in the form of detailed Descriptions of situations, events, people, interactions, and observed behavior, direct quotations from people about there experiences, attitudes, beliefs, and thoughts, and excerpts from documents, correspondence, records, and case histories.
- **Central Tendency :** a measure of central tendency provides a single most typical value as representative of a group of values; the 'trend' of a

group of measures as indicated by some type of averages, usually the mean, median or mode.

NOTES

- **Mean :** a kind of average obtained by dividing the sum of a set of measures by their number.
- **Median :** the middle value in a distribution or set of ranked values; the point that divides the group into two equal parts.
- **Mode :** the value that occurs most frequently in a distribution.
- **Variability :** the spread or dispersion of measures or values.
- **Variance :** a measure of variability of a distribution. It is the average of the squared deviations of the measures or values from the mean.
- **Standard Deviation :** the positive square root of variance.
- **Standard Score :** a general term referring to any of the variety of 'transformed' scores, in terms of which raw scores may be expressed for reasons of convenience, comparability, ease of interpretation, etc. Sigma Scores, T-Scores etc. are the examples of standard score.
- **Normal Distribution :** a distribution of measures that in graphic form has a distinctive bell-shaped appearance. It is symmetrical and asymptotic. The mean, mode and median for this type for distribution have equal values.

14.14 Self - Learning Exercise

Section-A (Very Short Answer Type Questions)

1 Define biostatistics.

2 What is inferential biostatistics ?

3 Who is called father of Biostatistics

4 When are two variables positively related ?

5 What is correlation ?

6 Who proposed coefficient of correlation ?

7 Under what conditions P.E. can be used ?

Section-B (Short Answer Type Questions)

1 Give brief note on biostatistics.

2 Explain probable error of correlation coefficient.

3 What are the uses of probable error ?

4 Explain coefficient of determination.

5 What is a scatter diagram ? How does it help in studying the correlation between two variables ?

6 What is meant by correlation and coefficient correlation

Section-A (Long Answer Type Questions)

1 What is Biostatistics? Mention its aims and applications in Biology. Write an account of the statistical methods in Biological Sciences.

2. What are parameters and statistics? Explain their meaning with reference to population and sample.

3 How does standard error differ from standard deviation? How will standard error change with a change in sample size?

4 Write short notes on : (i) standard error, (ii) mean deviation, (iii) median, (iv) mode, (v) arithmetic mean, (vi) standard deviation, (vii) sum of squares,

(viii) correction term.

5 What are tests of significance? Discuss the different tests of significance and describe situations under which each of these tests can be utilized.

6 How does *t-tests* for comparing a single mean with a given mean differ from a *t-test* meant to compare two means? Give formulate of standard errors that will be used in each case for calculating *t-statistics*

7 Define correlation and describe methods of measuring correlation.

8 Explain the term regression analysis. How the regression equation is derived

NOTES

9 What is probability? Discuss the addition and multiplication laws of probability?

10 What are mutually exclusive events and exhaustive set of events? How can these be used for calculating theoretical probability of an event.

NOTES

14.15 References

- Bollen, K. (1987). Structural Equations with Latent Variables. New York: Wiley
- Everitt, B. S. (1998). The Cambridge Dictionary of Statistics. New York: Cambridge University Press
- Kotz, S. (1988). Encyclopedia of Statistical Sciences, vol. 1-9 plus supplements. New York: Wiley
- Tukey, J. W. (1977). Exploratory Data Analysis. Reading, MA: Addison-Wesley
- Marcello Pagano &, Kimberlee Gauvreau (2000) Principles of Biostatistics
- V.B Rastogi. Fundamental of biostatistics

Unit - 15

Fundamentals of Computer and Bioinformatics

NOTES

Structure of the Unit

15.0 Objectives

After the study of this unit learner came to know about-

- Basic know how of the computers
- Basics of operating systems and applications of MS office
- Software used in biomedical sciences
- Basics of Bioinformatics and its application

15.1 Introduction

NOTES

To know use of computers and Information and communication technology (ICT) in every field of academics is become more and more essential day by day. For this one has to know basics of computers and common used software. By study present unit you came to know about basics of computers and operating system. Use of ICT in biological sciences emerged as a new subject known as bioinformatics. In this unit we also learn about basic of bioinformatics and its applications. Let we start with fundamentals of computers.

15.2 Fundamentals of Computers

Computer is the most adaptable electronic machine man has ever created. A computer is basically a programmable computing device, which does accept and store an input data then process the data input and finally output the processed data in required format. Computers have made a great impact in our daily work. Their presence is felt in our life be it homes, schools, colleges, offices, industries, hospitals, banks, retail stores, railways, airways, research and design organizations etc.

The word "computer" comes from the word *compute* which means to calculate. So a computer is normally considered to be a calculating device that can perform arithmetic operations at an enormous speed. A computer can not only store data and process but also retrieve data i.e. take out from its memory or storage as and when desired.

Fig. 15.1: Computer Fundamental Process

TYPES OF COMPUTERS:

(A) Digital Computers
(B) Analog Computers

(A) **Digital Computers:** Digital computers process information which is essentially in a binary or two-state form, i.e. 0 and 1. These are based on the measuring of analogues or equivalent physical value.

Digital computers fall into ranges called microcomputers, minicomputers, mainframes and supercomputers, classified on ascending order of size – small, medium, large and very large.

(B) **Analog Computers:** Analog computers process information which is of physical quantities, such as temperature, voltage, pressure etc.

NOTES

Generation of Computers:

The evolution of computers has passed through a number of stages before it reached the present state. Form the early 1950s, growth of the computers was very rapid but this development look distinct phases known as generation of computers. Different generations are characterized by their technology of basic computing elements:

First Generation: The first generation computers were the voluminous computers. These computers used electronic valves (vacuum tubes) and had the main limitations of very large electric power consumption and very little reliability.

Second Generation: The invention of transistor in 1948 led to the development of second generation computers. Transistors replaced the valves (vacuum tubes) completely as they were far more superior in performance on account of their miniature size, smaller power consumption and less expensive.

Third Generation: These computers use integrated circuits (ICs) in place of transistors and they were having higher speed, large storage capacities and less expensive. An ICs arranges thousands of switches on circuit boards known as a chip. These machines used more versatile programmes like Multi Programming Techniques and Database management.

Fourth Generation: Due to development of large scale integrated chips known as micro-processor chips in 1971 by Intel Corporation, another breed popular as fourth generation computer.

The fourth generation of computers is marked with the increased speed retrieval capabilities and facilities for communication. They can be connected with the satellite communication lines to transform information from one part of the world to the other part of the world at a very high speed. These machines

NOTES

used advanced software like controllers, computer aided design, electronic spreadsheet etc.

Fifth Generation: These machines will use parallel processing techniques and artificial intelligence techniques. Therefore these computers will be able to think as human beings can think. These machines will be able to process non-numerical data such as pictures, graphs, etc. They will use new types of integrated circuits for faster speed of operation. These machines are under development.

Characteristics of Computers

1. **Fast Speed calculation and Storing of Information:** Computers works at a fast speed without losing their accuracy. Their speed working measured in Mips, that called mega instructions per second. Another measure of computer speed is in terms of the speed of the microprocessor. Microprocessor is the heart of a computer and the execution speed of the computer will depend on the speed of operation of the processor. A microprocessor based computer which has higher clock frequency will operate faster than the one which has lesser clock frequency.
2. **Small Variety of Instructions:** The development of computers has progressed to the extent that the programs can be written by using a small set of instructions which can do very complex manipulations. Such functions when calculated manually take long time based on the number of variables and their arguments. But when using computer functions, you can do all such calculations in a split of second.
3. **Accuracy:** Computers give consistently accurate results. Their accuracy does not go down even when they are used continuously for days together. The accuracy also depends on the type of machine. For example microcomputer may give accuracy of results up to 8 decimal places only where's a minicomputer may give accuracy up to 32 decimal places.

Functional Units

The five major functional units of a digital computer are:

(a) CPU
(b) Input units
(c) Output Units

(d) Storage Units

(e) Communication Interface

CPU: Central Processing Unit is the brain of a computer. The CPU is responsible for activating and controlling the operations performed by all other units of the computer system. The major parts of a CPU are:

NOTES

(a) **Arithmetic and Logic unit (ALU):** All Calculations are performed and all comparisons are made in ALU.

(b) **Control Unit:** Control Unit obtains instructions from the program stored in the main memory, interprets the instructions, and issues signals that cause other units of the system to perform their functions.

(c) Primary Memory: It is also called main memory. Primary memory is a small and relatively fast storage unit that stores data and instructions which are being used by the CPU.

Inputting: Data input refers to the process of entering data into the computer, by the using of input devices.

Storing: Storing refers to the holding of data and provides instructions to the computer's main memory, for manipulation.

Processing: Processing refers to performing operations or manipulations of data entered into the computer so that useful information may be taken out of the entered data.

Outputting: Outputting refers to the process of showing the information or result to the user by output devices.

Controlling: Controlling refers to directing all the above processes, in coordination. This controlling is done by the Control unit in a CPU.

15.3 Computer Peripherals & Archietecture

Computer system has mainly two components, i.e. Hardware and Software. Hardware refers to the physical units of computer, which includes all electronic and electric circuit components and devices. Anything, which you see or touch, is considered as hardware. Software is a set of instructions that makes the hardware work to get desired result.

NOTES

Hardware

Hardware refers to any all parts on, connected to, or about the computer that you can actually physically touch. Parts like the computer case, the mouse, monitor, keyboard and speakers are all a good definition of computer hardware.

- Input devices
- Process unit
- Output devices
- Storage devices

Input Devices: Input devices accept data and instructions from the user. An input device first converts desired input data and instructions into a suitable binary form, i.e. 0 and 1. And then feed it into the CPU. The input devices are:

(a) Keyboard
(b) Mouse
(c) Joystick
(d) Light Pen
(e) Optical/Magnetic Scanner
(f) Graphic Tablet
(g) Touch Screen
(h) Microphone for voice as input

Fig. 15.2 : Input and Output Peripherals of Computer

Process Unit: The complex procedure that transforms raw input data into useful information for output is called processing. To perform this transform, the computer uses two components: The Processor and Memory.

The processing unit of a small computer normally has a single processor. The processing unit of a large computer may contain a number of processor. Each processor, in a large processing unit, performs a specific task assigned to it. Types of processors are: ADM. Celeron, Intel.

NOTES

Output Devices: After the computer has processed the data, the final result has to be brought out from the machine in some intelligent and readable form. The result that comes out of the machine is called output and the equipment which enables the results to be brought out are termed output devices.

There are several types of output devices manufactured for this purpose, but the most common used devices are-

Monitor:	uses a large vacuum tube
Printer:	writes out results on paper
Magnetic tape drive unit:	writes out results on magnetic tape
CD writers:	writes out results on compact disc

Storage Devices: Storage devices are a source for data storing in different sources. Storage devices are Floppy diskettes, hard disks, external hard disks, Pen drives etc.

Softwares

Software is the set of instructions to make computer work in the way as we like. Set of instructions is called Programmes. Software is a general term to describe all the forms of programs associated with computer.

Software is of two types;

(A) System Software

(B) Operating Software

(A) System Software: Set of programs supplied by manufacturer to make the computer work. The system software is an indispensable part of a total computer system. It's function is to compensate for the differences that exist between the user needs and the capabilities of the hardware. A computer

without some kind of system software would be ineffective and impossible to operate. This software coordinates the functioning of different parts of the computer.

NOTES

- Assemblers
- Compilers
- Loaders
- Linkers
- Operation System

(B) Application Software: Application programmes are user-written programmes to perform certain specific jobs. They are unique in their construction and can be used only for identical jobs. Many such application programmes have been made commercially available as PACKAGES. This type software coordinates two parts-

- User written software
- Ready made software

15.4 Elementary Idea about Operating Systems

An Operating System (OS) is system software that controls the internal activities of the computer hardware and user interface. These programmes are in-built into the computer resources such as processors, memory and input/output devices.

All operating systems are classified into four categories:

(a) **Single User Operating System:** One user allows to work at a time on a computer called single user operating system. Examples – DOS, Windows 9x etc.

(b) **Multiuser Operating System:** A multiuser operating system allows a number of users to work together on a single computer. Examples- Linux, Unix, Windows NT etc.

(c) **Single Tasking Operating System:** Operating system which can execute only a single task at a time is known as single tasking operating system. Example – DOS.

(d) **Multitasking Operating System:** Multitasking operating system supports execution of more than one job at a time on a computer. Example- Windows 2000, OS/2, UNIX, LINUX etc.

15.4.1 Disk Operating System (Dos)

NOTES

MS-DOS (Microsoft Disk Operating System) is a single user single taking operating system that can support only one user and only one task at a time. DOS commands classified into three categories-

1. **Simple Dos Commands:** When the DOS's files executed only corresponding program files are copied into the system called simple DOS command. These are two types-
   - Internal Commands – Internal commands are such programs in DOS which get loaded in the memory of a PC automatically at the time of booting or starting of a computer. For example – dir, del, rename, copy etc.
   - External Commands – External commands are such short programs which are available on hard disk/ output unit. These get loaded in the memory of the PC when specially asked. Example – format, chkdsk, print, diskcopy.
2. **File Management Commands:** These commands are specially designed to manage and operate different kinds of files stored in disks and directories. For Example- Copy, Del, Rename, Attrib, Find, Backup, Rename etc.
3. **Directory Commands-** Directory commands are specific to view the contents of the hard disk. These commands are also used for creating directory and sub-directory on the storage device. Example - Dir, MD, RD, Tree etc.

15.4.2 Windows

Windows is more than an operating system. It provides functions that let you manage many programs simultaneously, customize PC and do many of the maintenance tasks associated with a PC. Windows does the following jobs:

- Provides ways to start programs.

- Runs more than one programme at a time
- Provides a way for you to manage files.
- Helps to set up internet access.

NOTES

- It provides a way to customize the desktop and other screen areas to monitor.
- And many programs like games, music, outlook, etc was provided by winows.

15.5 Applications of MS-Office

Microsoft Office is an office suite of desktop applications, servers and services for the Microsoft Windows and OS X operating systems. It was first announced by Bill Gates of Microsoft on August 1, 1988 at COMDEX in Las Vegas. Initially a marketing term for a bundled set of applications, the first version of Office contained Microsoft Word, Microsoft Excel and Microsoft PowerPoint. Over the years, Office applications have grown substantially closer with shared features such as a common spell checker, OLE data integration and Visual Basic for Applications scripting language.

Major applications of MS-Office are as follows:

Microsoft Word: Microsoft Word is a word processor developed by Microsoft. It was first released in 1983 under the name *Multi-Tool Word* for Xenix systems. Microsoft Word's native file formats are denoted either by a .doc or .docx file extension. It is available for the Windows and OS X platforms. Word for Mac was the first graphical version of Microsoft Word.

Microsoft Excel: Microsoft Excel is a spreadsheet program that originally competed with the dominant Lotus 1-2-3, and eventually outsold it. It is available for the Windows and OS X platforms. Microsoft released the first version of Excel for the Mac OS in 1985, and the first Windows version in November 1987. It provided more functionality than the previous version.

Microsoft PowerPoint: Microsoft Access is a database management system for Windows that combines the relational Microsoft Jet Database Engine with a graphical user interface and software-development tools. Microsoft Access

stores data in its own format based on the Access Jet Database Engine. It can also import or link directly to data stored in other applications and databases.

Microsoft Outlook: Microsoft Outlook is a personal information manager. The replacement for Windows Messaging, Microsoft Mail, and Schedule+ starting in Office 97, it includes an e-mail client, calendar, task manager and address book.

NOTES

On the Mac OS, Microsoft offered several versions of Outlook in the late 1990s, but only for use with Microsoft Exchange Server. In Office 2001, it introduced an alternative application with a slightly different feature set called Microsoft Entourage. It reintroduced Outlook in Office 2011, replacing Entourage.

Microsoft OneNote: Microsoft OneNote is a freeware notetaking program. It gathers notes (handwritten or typed), drawings, screen clippings and audio commentaries. Notes can be shared with other OneNote users over the Internet or a network. OneNote was initially introduced as a standalone app that was not included in any of Microsoft Office 2003 editions. However, OneNote eventually became a core component of Microsoft Office; with the release of Microsoft Office 2013, OneNote was included in all Microsoft Office offerings before eventually becoming completely free of charge.

Fig. 15.3: Microsoft- Office Applications

15.6 Sound Spectrum Analysis

The frequency spectrum of a time-domain signal is a representation of that signal in the frequency domain. Similarly, a source of sound can have many different frequencies mixed. A musical tone's timbre is characterized by its harmonic spectrum. Sound in our environment that we refer to as *noise*

includes many different frequencies. When a sound signal contains a mixture of all audible frequencies, distributed equally over the audio spectrum, it is called white noise.

NOTES

Sound spectrum is one of the determinants of the timbre or quality of a sound or note. It is the relative strength of pitches called harmonics and partials (collectively overtones) at various frequencies usually above the fundamental frequency, which the actual note is named. The spectrum analyzer is an instrument which can be used to convert the sound wave of the musical note into a visual display of the constituent frequencies. This visual display is referred to as an acoustic spectrogram. Software based audio spectrum analyzers are available at low cost, providing easy access not only to industry professionals, but also to academics, students and the hobbyist. The acoustic spectrogram generated by the spectrum analyzer provides an acoustic signature of the musical note.

15.7 Computer Simulation

A **computer simulation** is a simulation, run on a single computer, or a network of computers, to reproduce behaviour of a system. The simulation uses an abstract model (a **computer model**, or a **computational model**) to simulate the system. Computer simulations have become a useful part of mathematical modeling of many natural systems in physics (computational physics), astrophysics, chemistry and biology, human systems in economics, psychology, social science, and engineering.

Computer simulations vary from computer programs that run a few minutes to network-based groups of computers running for hours to ongoing simulations that run for days. The scale of events being simulated by computer simulations has far exceeded anything possible (or perhaps even imaginable) using traditional paper-and-pencil mathematical modeling.

Computer models can be classified according to several independent pairs of attributes, including:

- Stochastic or deterministic
- Steady-state or dynamic
- Continuous or discrete

- Dynamic system simulation
- Local or distributed

For time-stepped simulations, there are two main classes:

NOTES

- Simulations which store their data in regular grids and require only next-neighbor access are called stencil codes. Many CFD applications belong to this category.
- If the underlying graph is not a regular grid, the model may belong to the meshfree method class.

15.8 Elementary Idea of Bioinformatics

Bioinformatics is a combination of molecular biology and computer sciences. It is that technology in which computers are used to gather, store, analyze and integrate biological and genetic information. The need for Bioinformatics arose when a project to determine the sequence of the entire human genome was initiated. This project was called the Human Genome Project. Bioinformatics is very important for the use of genomic information to understand human diseases and to identify new ways for gene-based drug discovery and development. Therefore, many universities, government institutions and pharmaceutical companies have come forward to form bioinformatics groups to do research related to computational biology so that better ways are used to make processes more efficient and less time consuming.

15.8.1 Application of Bioinformatics

Bioinformatics is the use of IT in biotechnology for the data storage, data warehousing and analyzing the DNA sequences. In Bioinformatics knowledge of many branches are required like biology, mathematics, computer science, laws of physics & chemistry, and of course sound knowledge of IT to analyze biotech data. Bioinformatics is not limited to the computing data, but in reality it can be used to solve many biological problems and find out how living things works.

With adequate data and right tools, it is possible to explore a number of new areas in biology. Chief among these are given below:

NOTES

- **Molecular medicine:** Biology and medicine are living in a time of rapidly increasing involvement of knowledge about processes at the molecular level, and bioinformatics can help cope with this rapid increase of required skills. We can find gene, those are directly associated with different disease and begin to understand the molecular level of these more clearly. This information of the molecular mechanisms of disease will enable better treatments, cures and even preventative tests to be developed and in the future, doctors will be able to analyse a patient's genetic profile and prescribe the best available drug therapy and dosage from the beginning.

- **Analysis of Gene Expression:** A physical or chemical change in a living system is not caused by a single gene but by the combined effect of many genes. Understanding the action of many genes on a single condition will, provide a genetic basis for disease and change control. The expression of many genes can be determined by measuring mRNA levels with multiple techniques including microarrays, expressed cDNA sequence tag (EST) sequencing, serial analysis of gene expression (SAGE) tag sequencing, massively parallel signature sequencing (MPSS), RNA-Seq, also known as "Whole Transcriptome Shotgun Sequencing" (WTSS), or various applications of multiplexed in-situ hybridization.

- **Analysis of Gene Regulation:** Regulation is the chain of events, starting with an extracellular event (such as temperature change) and leading to a change in the activity of proteins. The analysis of what promotes and regulates the activity of genes and proteins forms a part of bioinformatic study.

- **Microbial genomics:** Microorganisms are ubiquitous, that is they are found everywhere. They have been found surviving and thriving in extremes of heat, cold, radiation, salt, acidity and pressure. They are present in the environment, our bodies, the air, food and water. The genome of bacteria can help throw light on energy sources, environmental monitoring to detect

pollutants, find disease-producing properties of genes, and improve industrial efficiency.

- **Genomics:** Genomics is the study of complex sets of genes, their expression and the most vital role they play in biology. The most important application of bioinformatics in genomics is the Human Genome Project through which more than 30,000 genes have been identified and secured through the sequencing of chemical base pairs which make up the DNA. It has thus enabled us to obtain necessary knowledge as to how these genes inter-relate and what functions they perform. Cures for many diseases are being discovered through this inter - relation where bioinformatics, no doubt, plays a pivotal role.

- **Proteomics:** Proteomics is a branch of biotechnology that deals with the techniques of molecular biology, biochemistry, and genetics to analyze the structure, function, and interactions of the proteins produced by the genes of a particular cell, tissue, or organism. This technology is being improved continuously and new tactics are being introduced. In the current day and age it is possible to acquire the proteome data. Bioinformatics makes it easier to come up with new algorithms to handle large and heterogeneous data sets to improve the processes. To date, algorithms for image analysis of 2D gels have been developed. In case of mass spectroscopy, data analysis algorithms for peptide mass fingerprinting and peptide fragmentation fingerprinting have been developed.

- **Bioarchaeology, anthropology, evolution and human migration:** Understanding human and other genomes will help to understand evolution, inheritance, traits, mutations and disease carriers. The study of the genome comparison across organisms can help to understand similar genes matching organ donors.

- **DNA forensics (identification):** DNA profile of individual, called DNA fingerprints can help in identify criminal, establishing family relationship, protecting rare wildlife species, and matching organ donors.

NOTES

NOTES

- **Agriculture, livestock breeding and bio-processing:** Genome research on plants can provide nutritious, disease-resistance, pesticide-free crop. Even alternate use of crops can be found, e.g. tobacco has been found to produce bacterial enzymes that break down explosives such as TNT and dinitroglycerin.

15.8.2 Tools of Bioinformatics

Bioinformatics tools provide state-of-the-art descriptions of the various areas of applied bioinformatics, from the analysis of sequence, literature, and functional data to the function and evolution of organisms. The ability to development and interpret large volumes of data is essential with the application of new high throughput DNA sequencers providing an overload of sequence data. These can be classified as homology and similarity tools, protein functional analysis tools, sequence analysis tools and miscellaneous tools.

Homology Tools:

Homologous sequences are sequences that are related by divergence from a common ancestor. Thus the degree of similarity between two sequences can be measured while their homology is a case of being either true of false. This set of tools can be used to identify similarities between novel query sequences of unknown structure and function and database sequences whose structure and function have been elucidated.

Protein Function Analysis:

This group of programs allows comparing protein sequence to the secondary (or derived) protein databases that contain information on motifs, signatures and protein domains. Highly significant hits against these different pattern databases allow to approximate the biochemical function of query protein.

Structural Analysis:

This set of tools allows comparing structures with the known structure databases. The function of a protein is more directly a consequence of its structure rather than its sequence with structural homolog's tending to share

functions. The determination of a protein's 2D/3D structure is crucial in the study of its function.

Sequence Analysis:

This set of tools allows carrying out further, more detailed analysis on query sequence including evolutionary analysis, identification of mutations, hydropathy regions, CpG islands and compositional biases. The identification of these and other biological properties are all clues that aid the search to elucidate the specific function of sequence.

NOTES

15.9 Proteomics and Genomics

Proteomics is a branch of biotechnology that deals with the techniques of molecular biology, biochemistry, and genetics to analyze the structure, function, and interactions of the proteins produced by the genes of a particular cell, tissue, or organism. This technology is being improved continuously and new tactics are being introduced. In the current day and age it is possible to acquire the proteome data. Bioinformatics makes it easier to come up with new algorithms to handle large and heterogeneous data sets to improve the processes. To date, algorithms for image analysis of 2D gels have been developed. In case of mass spectroscopy, data analysis algorithms for peptide mass fingerprinting and peptide fragmentation fingerprinting have been developed.

Genomics is the study of complex sets of genes, their expression and the most vital role they play in biology. The most important application of bioinformatics in genomics is the Human Genome Project through which more than 30,000 genes have been identified and secured through the sequencing of chemical base pairs which make up the DNA. It has thus enabled us to obtain necessary knowledge as to how these genes inter-relate and what functions they perform. Cures for many diseases are being discovered through this inter-relation where bioinformatics, no doubt, plays a pivotal role.

15.10 Softwares for Visualization of Secondary Structures of Bio Molecules

NOTES

Secondary structure prediction is a set of techniques in bioinformatics that aim to predict the local secondary structures of proteins and RNA sequences based only on knowledge of their primary structure — amino acid or nucleotide sequence, respectively. For proteins, a prediction consists of assigning regions of the amino acid sequence as likely alpha helices, beta strands.

The best modern methods of secondary structure prediction in proteins reach about 80% accuracy; this high accuracy allows the use of the predictions in fold recognition and ab initio protein structure prediction, classification of structural motifs, and refinement of sequence alignments. The accuracy of current protein secondary structure prediction methods is assessed in weekly benchmarks such as LiveBench and EVA.

The software for visualization of secondary structures are as follows:

- Agadir • Prediction of the helical content of peptides
- APSSP • Advanced Protein Secondary Structure Prediction
- Ascalaph • Molecular modeling software
- CFSSP • Protein secondary structure prediction
- Click2Drug • Directory of computational drug design tools
- COILS • Prediction of Coiled Coil Regions in Proteins
- CPHmodels • Protein homology modeling
- DisEMBL • Prediction of disordered protein regions
- DLP-SVM • Domain linker predictor
- FUGUE • Sequence-structure homology recognition
- Geno3D • Protein molecular modelling
- GlobPlot • Protein disorder/globularity/domain predictor
- GOR • Protein secondary structure prediction

- HHpred • Homology detection & structure prediction
- HTMSRAP • Helical TransMembrane Prediction
- Jmol • 3D molecule visualization tool
- Jpred • Secondary Structure Prediction Server
- LiveBench • Continuous Benchmarking of Structure Prediction Se
- MakeMultimer.py • Reconstruction of multimeric molecules in crystals
- MARCOIL • coiled-coils prediction
- MarvinSpace • Visualization, modeling and interactive rendering
- MeDor • Metaserver for protein disorder prediction
- MolTalk • Environment for structural bioinformatics
- MovieMaker • Generates movies of protein dynamics
- NetSurfP • Protein Surface Accessibility Prediction
- NetTurnP • Prediction of Beta-turn regions in proteins
- NQ-Flipper • Correction of unfavorable rotamers of Asn and Gln
- OpenStructure • molecular modelling and visualization
- PDBePISA • Protein Interfaces, Surfaces and Assemblies
- Phyre2 • 3D structure prediction with HMM
- Poodle • Prediction of disordered protein regions
- PoPMuSiC • Thermodynamic stability prediction upon mutations
- PROCHECK • Stereochemical quality of a protein structure
- PROF • Secondary Structure Prediction System
- ProSA-web • Program of error recognition in 3D structures
- ProtBud • Comparison of asymmetric units and biological unit
- Protein Disorder Predictors • Protein Disorder Predictors

NOTES

NOTES

- Protein Model Portal • structural information for a protein
- PSIPRED • Various protein structure prediction methods
- PyMOL • Molecular graphics visualization
- QMEAN • estimate quality of protein models
- RasMol • Molecular graphics visualization
- SAM-T08 • HMM-based Protein Structure Prediction
- Scratch Protein Predictor • protein structure & structural features prediction
- SOPMA • Secondary structure prediction.
- Strap • Structural alignments of proteins
- Swiss-PdbViewer • analyse protein 3D structures
- SwissParam • topology, parameters for small organic molecules
- TLS Motion Determination • Translation/Libration/Screw Motion Determination
- TMPred • membrane-spanning region prediction
- TopMatch • Protein structure comparison
- TopPred • Topology prediction of membrane proteins
- TopSearch • Gateway to protein structures
- VMD - Visual Molecular Dynamics • Molecular visualization program
- WHAT IF • Molecular modelling package
- YASARA • Molecular graphics, modeling, simulations

15.11 Summary

Computer is our daily need now days. Present day computer serves our requirements at every field of science with adaptive and emerging application softwares.

Present day education especially science education is highly dependent on tools of computers. The use of computer in the field of biological science emerged as a new field of academics i.e. Bioinformatics. Bioinformatics enables the scientists to analyse each and every aspect of biomolecules using their sequence. Various tools of Bioinformatics are very much helpful in biomedical sciences and open new vistas in the field of diagnostics and therapeutics.

NOTES

15.12 Glossary

- **System Software :** Set of programs supplied by manufacturer to make the computer work.
- **Application Software :** Application programmes are user-written programmes to perform certain specific jobs.
- **Bioinformatics :** It is that technology in which computers are used to gather, store, analyze and integrate biological and genetic information.
- **Sound Spectrum :** Sound spectrum is one of the determinants of the timbre or quality of a sound or note.
- **Proteomics :** Proteomics is a branch of biotechnology that deals with the techniques of molecular biology, biochemistry, and genetics to analyze the structure, function, and interactions of the proteins produced by the genes of a particular cell, tissue, or organism.
- **Genomics :** Genomics is the study of complex sets of genes, their expression and the most vital role they play in biology.

15.13 Self-Learning Excercise

1. Write short notes on

   i. DOS

   ii. Computer simulation

2. Write a detail note on Proteomics and Genomics.

3. What do you understand by computer simulation? Explain.

4. Write an essay on applications of Bioinformatics.

15.14 References

NOTES

- David W. Mount (2013), Bioinformatics: Sequence and Genome Analysis, Pub. Laboratory Press
- Arthur M. Lesk (2002), Introduction to Bioinformatics, Pub. Oxford University Press, USA
- Neil C. Jones, Pavel A. Pevzner (2004), An Introduction to Bioinformatics Algorithms, Pub. Mit Press
- Glyn Moody (2004), Digital Code of Life: How Bioinformatics Is Revolutionizing Science, Medicine, and Business, Pub. John Wiley & Sons.