PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
MINISTRY OF HIGHER EDUCATION AND SCIENTIFIC RESEARCH

20th August 1955 University of Skikda

Faculty of Sciences

Department of Nature and Life Sciences

GENETICS

Course for 2nd year Biological Sciences, Marine and Continental Hydrobiology students

By Dr. Safia Boulechfar

2024-2025

Preface

The course of Genetics is specifically designed for 2nd year Biological Sciences and Marine and Continental Hydrobiology students.

The overall aim of this course, is to provide a clear, comprehensive, rigorous, and balanced introduction to fundamental genetics. It is a guide to help students to:

- understand the basic concepts of genetics
- be able to make connections between the main concepts of genetics
- apply the concepts and the principles to solve problems of several types during tutorial sessions, such as single-concept exercises that require application of definitions or the basic principles of genetics, and quantitative problems that call for some numerical calculation.

The course comprises 11 chapters. Chapter 1 presents information about the genetic material including DNA and RNA structures, DNA replication, organization of DNA in chromosomes. Chapter 2 deals with cell cycle and cell divison. Chapter 3 covers the principles of inheritance including Mendel's laws, independent and linked genes, X-linked inheritance. Chapter 4 deals with ordered tetrad analysis. Chapter 5 describes the mechanisms of genetic transfer between bacteria. Chapters 6 and 7 deal with topics in molecular genetics including transcription, translation and mutations. It is worth-mentioning that chapter 7 combined two chapters (genic mutations and chromosomal mutations) into just one chapter, which is mutations. Chapters 8 and 9 focus mainly on the structure of genes and control of gene expression in both prokaryotic and eukaryotic oganisms. Chapter 10 presents extrachromosomal genetics. Chapter 11 deals with some basic concepts of population genetics.

Table of contents

Chapter 11 : Population genetics

CHAPTER 1

Genetic Material

ABSTRACT

This chapter provides information about the structure of nucleic acids, the mode of DNA replication and the organisation of DNA in the cell.

1.1. The structure of DNA and RNA

1.1.1. The structure of DNA

DNA is composed of repeating subunits called nucleotides. Each nucleotide is composed of a five-carbon sugar, or pentose, a nitrogenous base and a phosphate group. In DNA, the sugar is β-D-2'-deoxyribose, thus the name deoxyribonucleic acid (Figure 1.1).

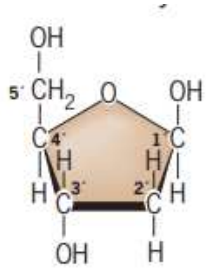


Figure 1.1. β-D-2'-deoxyribose structure

Four different bases commonly are found in DNA: adenine (A), guanine (G), thymine (T), and cytosine (C). Adenine and guanine are double-ring bases called *purines*; cytosine and thymine are single-ring bases called *pyrimidines* (Figure 1.2).

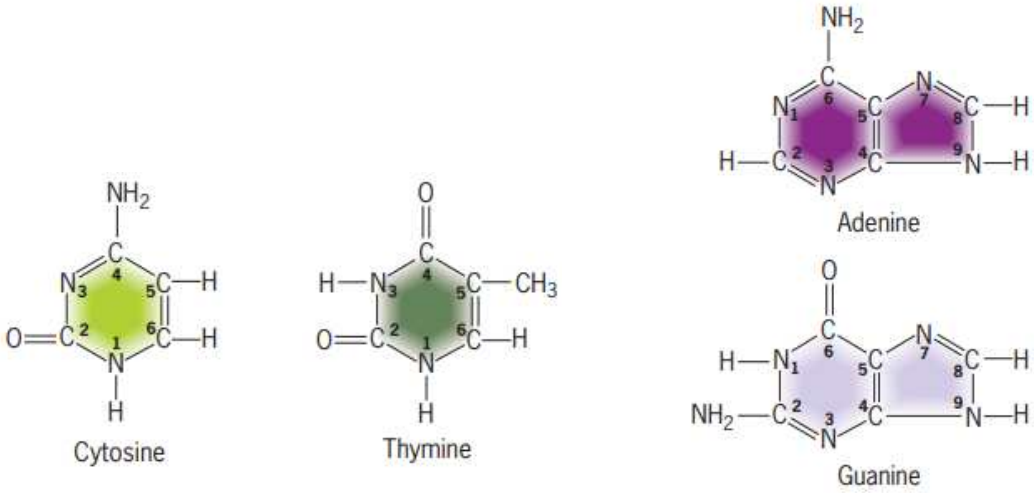


Figure 1.2. Structures of nitrogenous bases present in DNA

In DNA, each base is chemically linked to one molecule of the sugar deoxyribose, forming a compound called a *nucleoside.* When a phosphate group is also attached to the sugar, the nucleoside becomes a nucleotide (Figure 1.3). Thus a nucleotide is a nucleoside plus a phosphate.

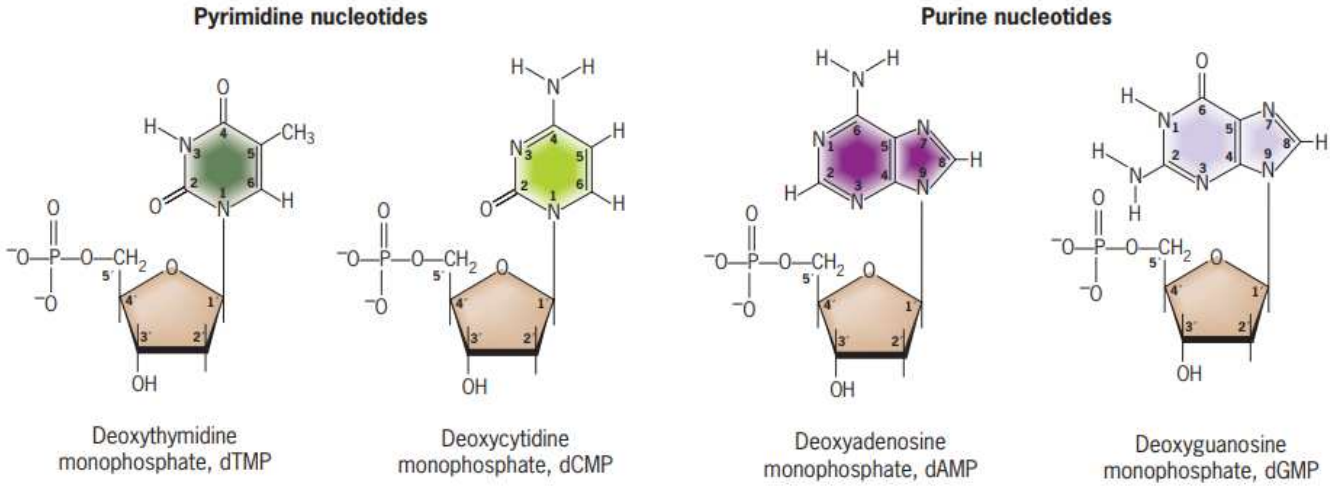


Figure 1.3. Structures of the four common deoxyribonucleotides present in DNA.

The nomenclature of the nucleoside and nucleotide derivatives of the DNA bases is summarized in table 1.1.

Table 1.1. Nomenclature of nucleoside and nucleotide present in DNA

Base	Nucleoside	Nucleotide
Adenine (A)	Deoxyadenosine	Deoxyadenosine-5' monophosphate (dAMP) Deoxyadenosine-5' diphosphate (dADP) Deoxyadenosine-5' triphosphate (dATP)
Guanine (G)	Deoxyguanosine	Deoxyguanosine-5' monophosphate (dGMP) Deoxyguanosine-5' diphosphate (dGDP) Deoxyguanosine-5' triphosphate (dGTP)
Thymine (T)	Deoxythymidine	Deoxythymidine-5' monophosphate (dTMP) Deoxythymidine-5' diphosphate (dTDP) Deoxythymidine-5' triphosphate (dTTP)
Cytosine (C)	Deoxycytidine	Deoxycytidine-5' monophosphate (dCMP) Deoxycytidine-5' diphosphate (dCDP) Deoxycytidine-5' triphosphate (dCTP)

In DNA, the nucleotides are joined to form a polynucleotide chain, in which the phosphate attached to the 5' carbon of one sugar is linked to the hydroxyl group attached to the 3' carbon of the next sugar in line (Figure 1.4). The chemical bonds by which the sugar components of adjacent nucleotides are linked through the phosphate groups are called *phosphodiester bonds*. The 5'-3'-5'-3' orientation of these linkages continues throughout the

chain. The terminal groups of each polynucleotide chain are a 5'-phosphate (5'-P) group at one end and a 3'-hydroxyl (3'-OH) group at the other.

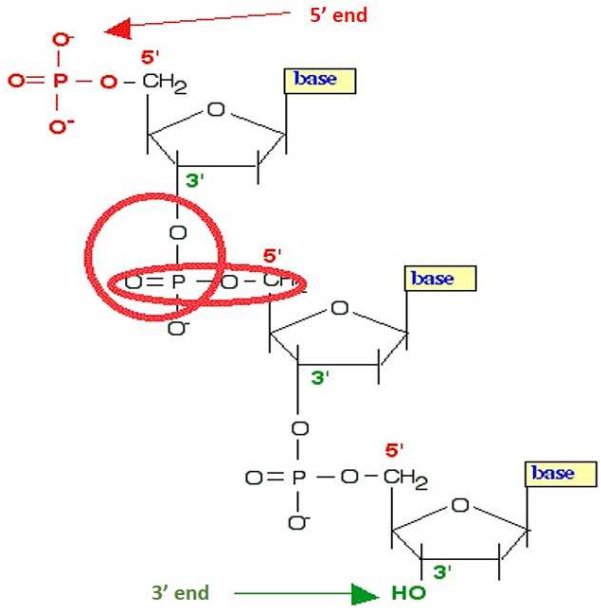


Figure 1.4. Phosphodiester bond

DNA is a normally double stranded macromolecule composed of two polynucleotide strands wound around each other (Figure 1.5). The sugar-phosphate backbones of the two DNA strands wind around the helix axis like the railing of a sprial staircase, whereas the base pairs are stacked between the two chains perpendicular to the axis of the molecule like the steps of a spiral staircase. The two strands are held together in their helical configuration by hydrogen bonding between bases in opposing strands. The base-pairing is specific : adenine is always paired with thymine, and guanine is always paired with cytosine. Adenine and thymine form two hydrogen bonds, and guanine and cytosine form three hydrogen bonds. The two strands of a DNA double helix are thus said to be complementary. According to Chargaff's rules, in a double stranded DNA :

- The amount of adenine equals that of thymine: [A] = [T].
- The amount of guanine equals that of cytosine: [G] = [C].
- The amount of the purine bases equals that of the pyrimidine bases: [A] + [G] = [T] + [C].

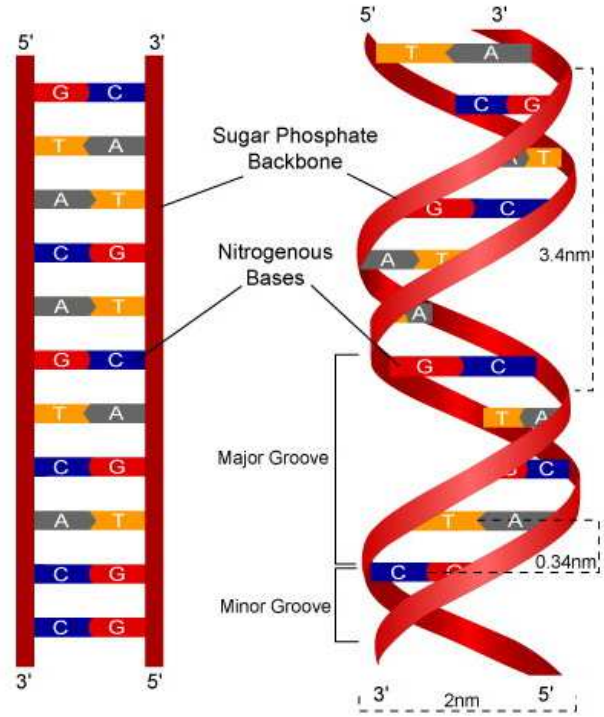


Figure 1.5. DNA structure

The two strands are "antiparallel"; that is, one strand runs 5′ to 3′ while the other runs 3′ to 5′. However, since the DNA strands are assembled in the 5′ to 3′ direction, by convention, we read them the same way.

1.1.1.1. Forms of DNA double helix

Under physiological conditions, DNA exists in the B conformation (B-DNA) (Figure 1.6). The vast majority of the DNA molecules present in the aqueous protoplasms of living cells exist in such conformation. However, the structures of DNA molecules change as a function of their environment. In high concentrations of salts or in a partially dehydrated state, DNA exists as A-DNA, which is a right-handed helix like B-DNA, but is a shorter and thicker double helix. DNA molecules almost certainly never exist as A-DNA *in vivo*. Certain DNA sequences have been shown to exist in a left-handed, double-helical form called Z-DNA (Z for the zigzagged path of the sugar-phosphate backbones of the structure).

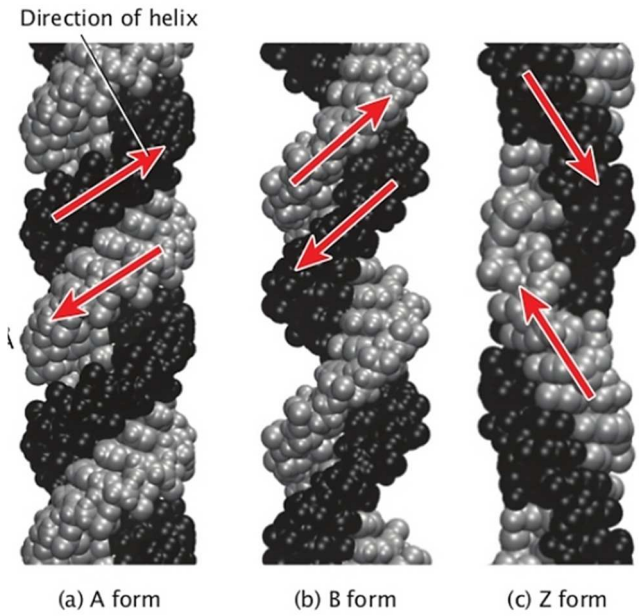


Figure 1.6. DNA Double helix forms

1.1.2. The structure of RNA

The structure of RNA is similar to that of DNA; the main chemical differences are the presence of ribose instead of 2′-deoxyribose and uracil instead of thymine (Figure 1.7).

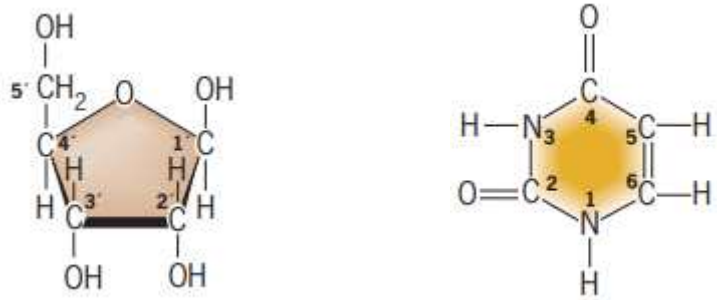


Figure 1.7. Structures of ribose and uracil

RNA is also most commonly single-stranded, although short stretches of double-stranded RNA may be found in self-complementary regions. The two most common structures of double-stranded RNA are hairpins and pseudoknots (Figure 1.8).

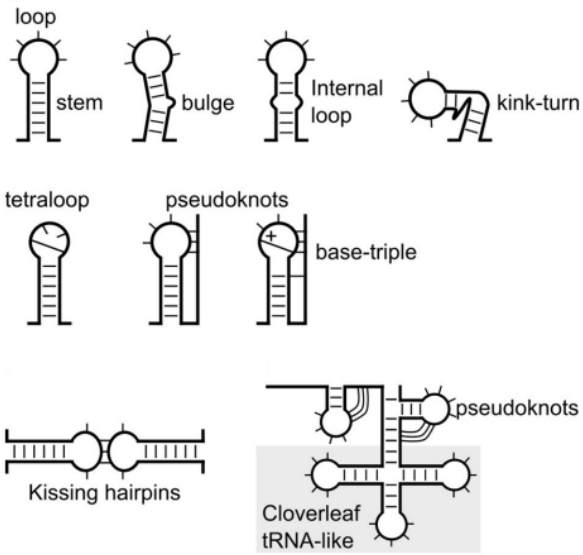


Figure 1.8. Hairpins and pseudoknots

N.B : The names of nucleosides and nucleotides in RNA are similar to that in DNA but without the word « deoxy » before the names of nucleosides and nucleotides and without « d » before the abbreviation. For example : Adenosine, Adenosine-5'triphosphate (ATP), Uridine, Uridine-5' triphospate (UTP)…etc.

1.1.2.1. RNA types

There are three main types of RNA molecule found in cells: messenger RNA (mRNA), ribosomal RNA (rRNA), and transfer RNA (tRNA). Ribosomal RNA is the most abundant class of RNA molecule, making up some 85% of total cellular RNA. It is associated with ribosomes, which are an essential part of the translational machinery. Transfer RNAs make up about 10% of total RNA and provide the essential specificity that enables the insertion of the correct amino acid into the protein that is being synthesised. Messenger RNA, as the name suggests, acts as the carrier of genetic information from the DNA to the translational machinery and usually makes up less than 5% of total cellular RNA.

1.2. Physico-chemical properties of DNA

- *Solubility* : DNA is polar in nature and thus soluble in water. Its highly charged phosphate-sugar backbone gives it its polarity. However, in the presence of salt and alcohol, it is insoluble.
- *Absorption* : Nucleic acids can absorb UV light at the wavelength of 260 nm.
- *Denaturation and renaturation* : Denaturation is the process of separation of double-stranded DNA (dsDNA) to two single strands (ssDNA) by breaking of hydrogen bonds

between bases due to an increase in temperature. When the two separated strands of DNA bind together again, it is called as renaturation of DNA. DNA is renatured by cooling.

Melting temperature (Tm) : T° at which 50% of DNA has denatured from double-stranded DNA (dsDNA) to single-stranded DNA (ssDNA) (Figure 1.8). Tm is influenced by many factors :

- ✓ Salt concentration : Tm increases as [NaCl] decreased
- ✓ DNA strand length : Tm increases as DNA strand length increases.
- ✓ G+C% content : Tm increases as the proportion of G+C bases increases.

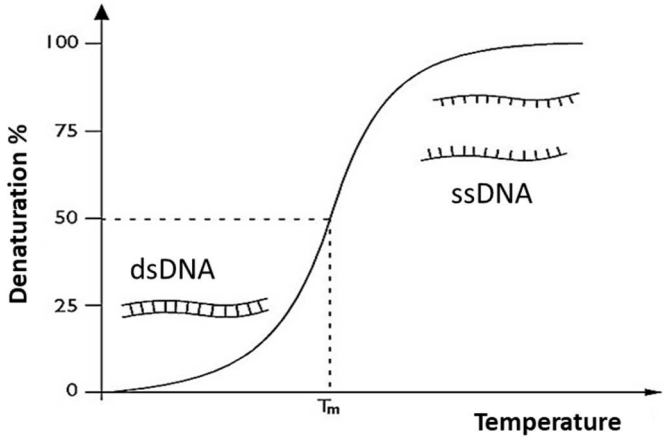


Figure 1.8. DNA melting curve

1.3. DNA replication

A replication is a process by which DNA is duplicated prior to cell division. DNA replication is semi concervative in which each double helix contains an old and a newly synthesized strand. The strands of the original (parent) duplex separate, and each individual strand serves as a pattern, or template, for the synthesis of a new strand (replica).

1.3.1. Steps of DNA replication

Here are the key steps involved in DNA replication :

1.3.1.1. Initiation

DNA replication requires extreme precision. Consequently, replication does not start randomly at any point in the DNA. The replication process in both prokaryotes and eukaryotes begins at a specific region known as the origin of replication. This is where the replication starts.

In prokarytes, the replication of circular DNA molecule stars at one origin of replication called replicon, from which the replication procceds in the two directions of replication forks. The mode of replication in prokaryotes is usually called θ (theta) replication.

DNA duplex in a eukaryotic chromosome is linear and also replicates bidirectionally. However, since it is a long molecule, the replication is initiated at many sites in the DNA (Figure 1.9).

The process of replication initiates with the detection of this origin, followed by the unwinding of the two DNA strands. Due to the high energy input, it is not feasible to unzip the entire length of the DNA strands. Thus, a replication fork is first created. Helicase enzyme, unzips the DNA strands, giving rise to two replication forks. Replication complexes assemble at both of these and proceed in opposite directions. Then, single stranded binding proteines (SSB proteines) bind to the separated strands to prevent their reassociation.

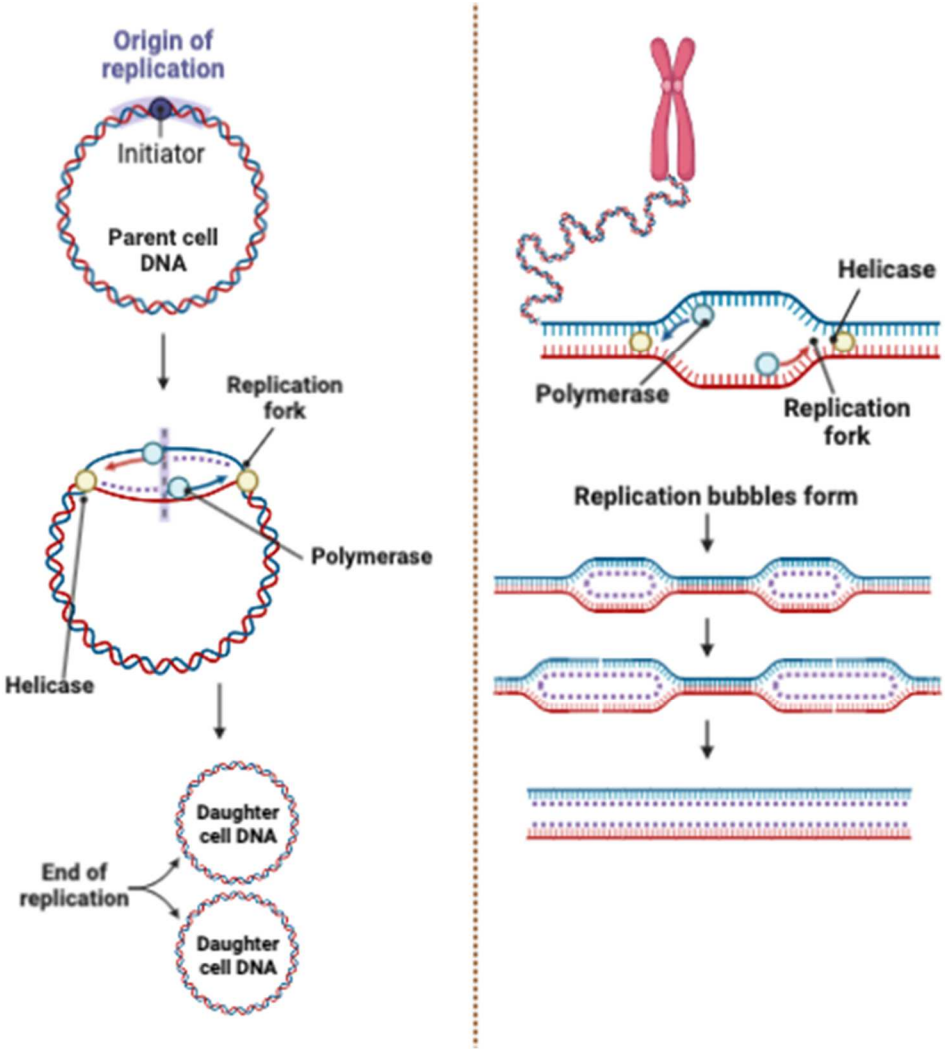


Figure 1.9. Replication in prokarytes and eukaryotes

1.3.1.2. Elongation

Once the strands are separated, the polymerase enzymes start synthesizing the complementary sequence in each of the strands, using the parental strands as templates for the newly

synthesizing daughter strands (Figure 1.10). However, DNA polymerase is unable to initiate new chains but rather requires a preexisting primer. The primer is a short RNA sequence synthesized near the origin with the help of a special type of RNA polymerase called primase. Then DNA polymerase adds deoxyribonucleotides to the 3' end position of RNA primer.

Notably, DNA is always polymerized in the 5′ to 3′ direction. Thus, in one strand (the template 3'→5'), it is continuous replication, the strand is called the leading strand. However, on the other strand (the template 5'→3'), it is discontinuous replication, and the strand called lagging strand. Its replication occurs as fragments known as Okazaki fragments, which are later joined by the enzyme DNA ligase.

1.3.1.4. Termination

The termination of replication varies among different organisms. For instance, in organisms like *E.coli*, which have circular chromosomes, replication terminates when the two replication forks meet each other. Then, DNA polymerase degrades the RNA primer and replaces it with short DNA segment. The newly synthesized segment is joined to the main DNA strand with the help of DNA ligase enzyme.

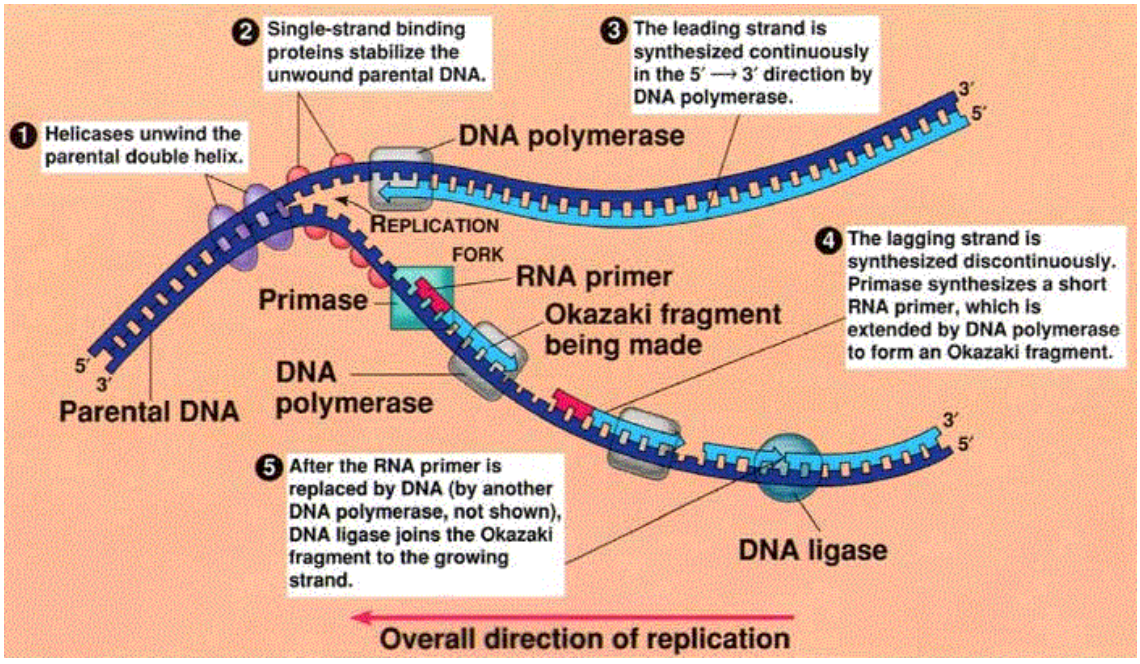


Figure 1. 10. DNA replication process

1.4. Organisation of DNA in cell

DNA is packaged into chromosomes in a highly organized way. DNA associate with 8 proteins called histones to form nucleosomes, the structural unit of chromosome. The second level of

DNA packaging is mediated by histone HI, binding together adjacent nucleosomes to condense DNA into the supercoiled 30-nm fiber, which is also called the solenoid. The third level of organization is thought to involve the formation of transcriptional units of DNA loop. At metaphase, chromatin is maximally condensed and forms 1400-nm fibers.

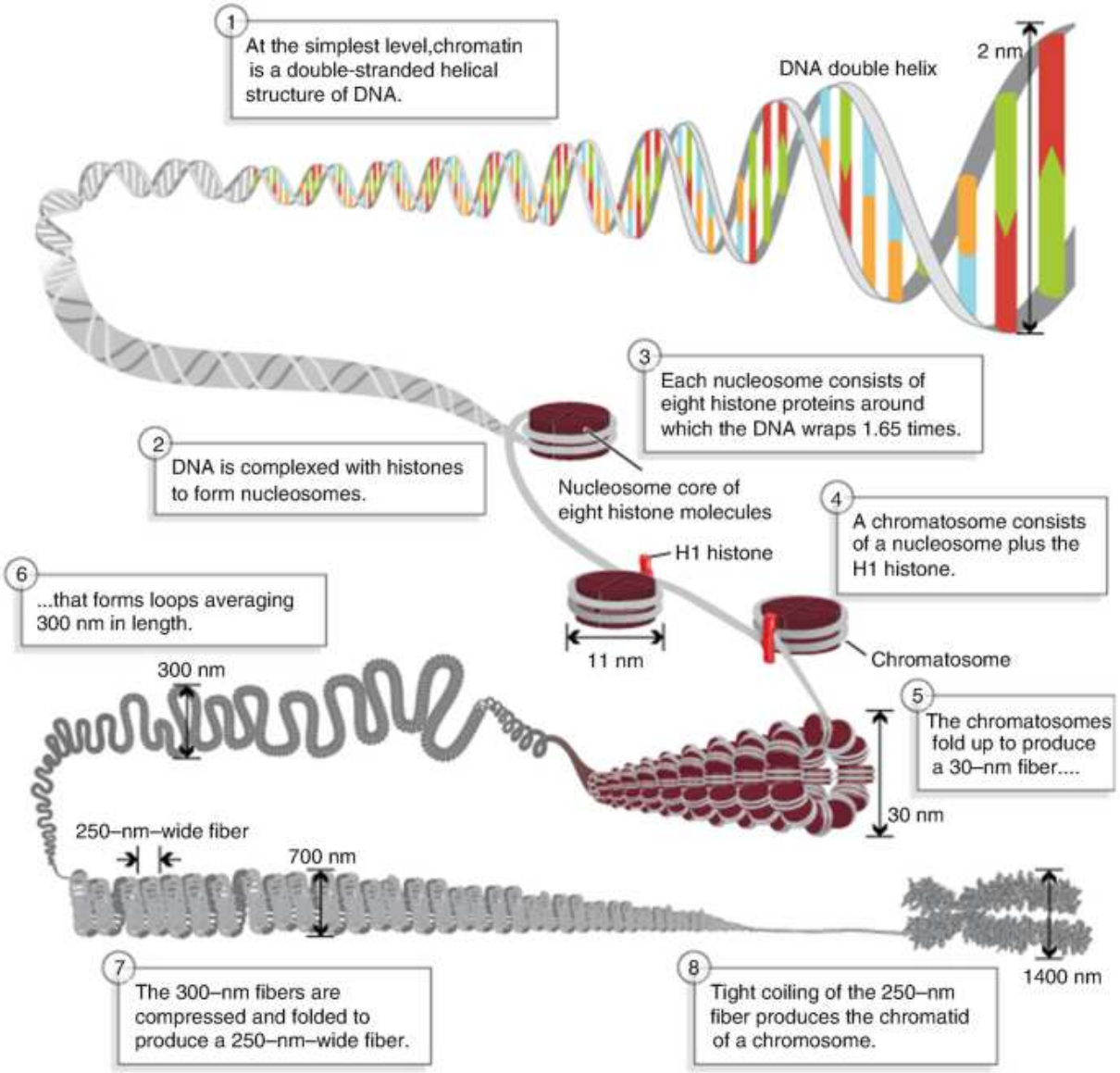


Figure 1.10. DNA packaging into chromosome

1.5. Number and morphology of chromosomes

1.5.1. Chromosome number

The basic number of chromosomes varies among species. Chromosome number is unrelated to the size or biological complexity of an organism. *Drosophila* has 8 chromosomes, corn has

20 chromosomes, 200 chromosomes in crustaceans have 200 hundreds and humans have 46 chromosomes in their genome.

1.5.2. Morphological features of chromosomes

Chromosomal morphology is clearly observable under optical microscope during metaphase. A chromosome consists of a centromere, telomere, and one or two sister chromatids (the non-duplicated chromosome has only one chromatid, and the duplicated chromosome has two chromatids).

On the basis of the position of the centromere, chromosomes are classified into four types (Figure 1.11) :

- *Metacentric chromosomes* : the centromere is located in the middle of a chromosome forming two equal arms.
- *Sub-metacentric chromosomes* : the centromere is located near the centre of the chromosome forming two unequal arms.
- *Acrocentric chromosomes* : the centromere is closer to one end of the chromosome.
- *Telocentric chromosomes* : the centromere is located on the proximal end of the chromosome.

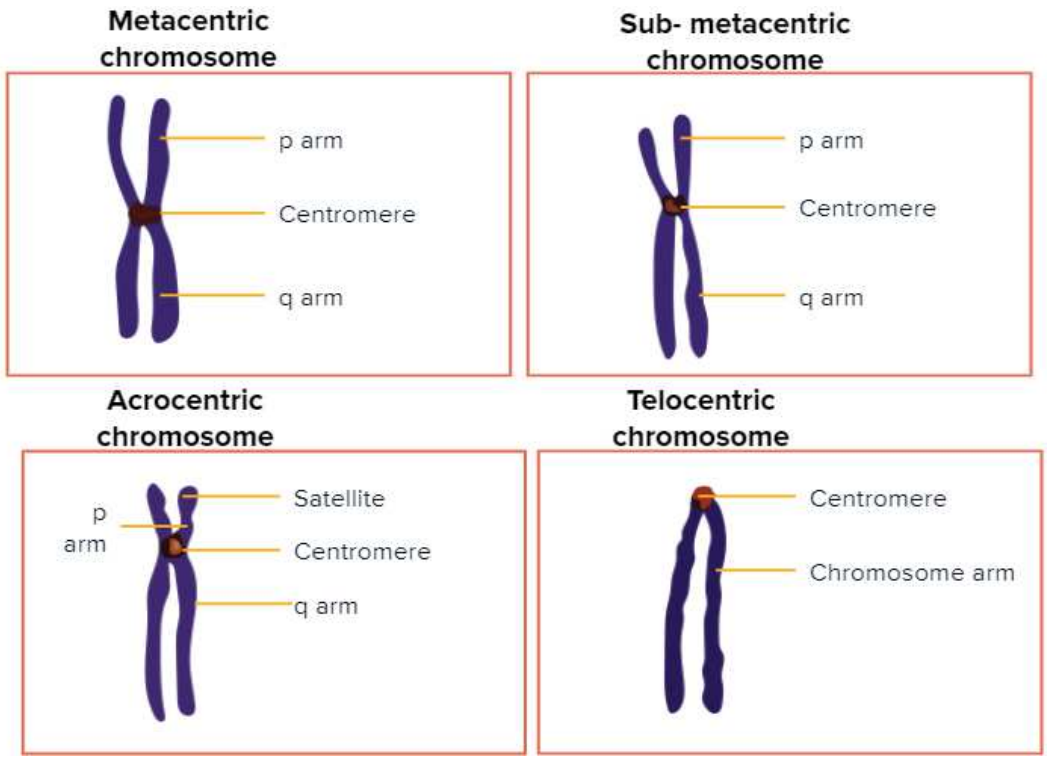


Figure 1.11. Chomosome types

CHAPTER 2

Transmission Of Genetic Characteristics In Eukaryotic Organisms

ABSTRACT

This chapter provides information about cell cycle, mitosis and meiosis as well as changes of DNA quantity throughout the cell cycle.

2.1. Cell cycle

The cell cycle is the repeating pattern of cell growth. It is the life story of a cell, the stages through which it passes from one division to the next.

Cell cycle consists of two major stages : **interphase** and **M phase** or cell division (mitosis or meiosis).

2.1.1. Interphase

Interphase is the period between cell divisions, in which the cell grows, develops, and prepares for cell division. Interphase is divided into three phases : **G1**, **S**, and **G2** (Figure 2.).

- **G1 (for gap 1) :** Interphase begins with G1. In this phase, the cell grows, and proteins necessary for cell division are synthesized. This phase typically lasts several hours. There is a critical point in the cell cycle, termed the G1/S checkpoint, in G1; after this checkpoint has been passed, the cell is committed to divide, if not, the cell pass into a nondividing phase (quiescent state) called G0, which is a stable state during which cells usually maintain a constant size.
- **S phase (for DNA synthesis) :** After G1, the cell enters the S phase, in which each chromosome duplicates. Before S phase, each chromosome is composed of one chromatid; following S phase, each chromosome is composed of two chromatids.
- **G2 (gap 2)** : After the S phase, the cell enters G2. In this phase, several additional biochemical events necessary for cell division take place. The important G2/M checkpoint is reached in G2; after this checkpoint has been passed, the cell is ready to divide and enters M phase.

2.1.2. M phase or cell division

The cell division phase resulting in the production of two daughter cells It is a rapid phase (lasting about 1 hour).

2.2. The cell cycle checkpoints

A checkpoint is a mechanism that halts progression through the cycle (arrest cell cycle) until a critical process such as DNA synthesis is completed, or until damaged DNA is repaired. When a checkpoint is satisfied, the cell cycle can progress.

Two types of proteins play important roles in this progression : the cyclins and the cyclin-dependent kinases, often abbreviated CDKs. Complexes formed between the cyclins and the CDKs cause the cell cycle to advance.

To minimize the possibility of errors, checkpoints exist at the four different phases of the cell cycle : G1 (to prevent enry into S phase), intra S, G2 (to prevent mitosis), and M (to avoid mitotic exist with abnormal segregation of chromosomes. In general these checkpoints monitor the status and structure of DNA during cell cycle progression. In Addition checkpoints monitor te proper structure of chromosomes ans their biooriontation to ensure equal distribution between the two daughter cells.

2.2.1. G1/S Checkpoint

G1/S Checkpoint is a DNA integrety checkpoint, which ensures that any damage is repaired before DNA synthesis.

2.2.2. Intra-S phase checkpoint

The goal of this checkpoint is to prevent the deterious consequences of replicating damaged DNA. Therefore, S-phase cells must respond instantaneously to DNA damage to halt initiation of new replication forks throughout the S phase.

2.2.3. G2 checkpoint

The G2 checkpoint is required to prevent the passage of DNA lesions to the two daughter cells during M phase.

2.2.4. Spindle assembly checkpoint (SAC)

SAC delays the exit from mitosis in the presence of defective chromosome alignement. This checkpoint prevents chromosome segregation until all condensed sister chromatid pairs are aligned at the metaphase plate with the appropriate biorientation.

2.3. Changes in quantity of DNA during the cell cycle

The quantity of DNA also changes throughout the cell cycle. A cell that enters the G1 phase has 1 copy of the full genome in its nucleus. But then during the S phase, DNA gets replicated to create 2 copies of all the genetic material within the cell. So in the G2 phase, the DNA quantity has doubled. But when the cell divides, in the M phase, the quantity of DNA halves again (goes back to what it was in G1) as each daughter cell gets 1 copy of DNA each.

During G1 phase, DNA content is, for example, C (Figure 2.1). S phase is marked by replication of DNA and the amount of DNA per cell is doubled (2C). During G2 phase, synthesis of RNA and proteins takes place and still the amount of DNA is 2C. During mitosis (M phase), DNA content gets equally distributed among two daughter cells and DNA again equal to the parent cell (C).

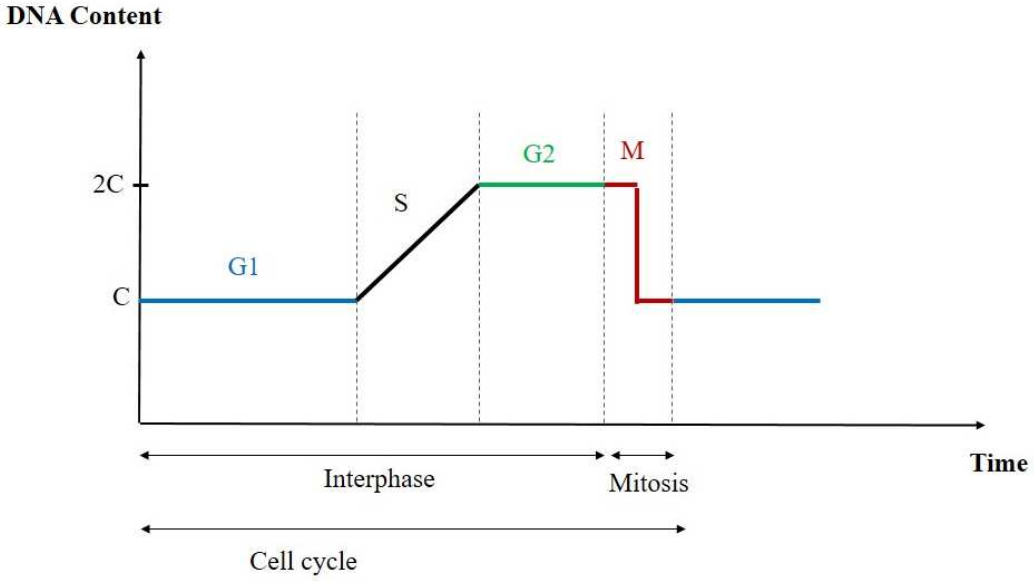


Figure 2.1. Changes of DNA content during cell cycle

N.B :

The number of chromosomes = the number of centromeres
The number of DNA molecules = The number of chromatids

2.4. Mitosis

Mitosis is the process by which the somatic cells of all multicellular organisms multiply. The significance of mitosis is its ability to produce daughter cells which are exactly the same as the parent cell. It is important for growth and for repairing damaged cells.

The process of mitosis is divided into four stages : *prophase, metaphase, anaphase* and *telophase* (Figure 2.2).

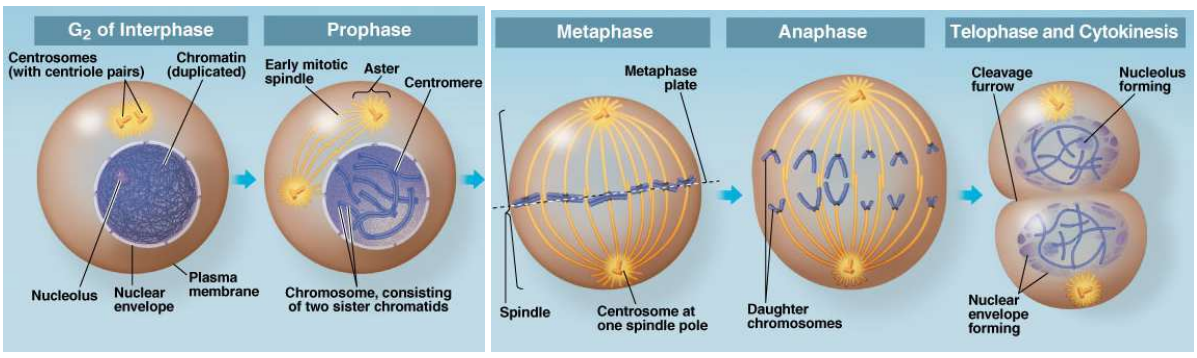


Figure 2.2. Mitosis stages

2.4.1. Prophase

During this stage, the replicated chromosomes condense and become visible. The centrioles duplicate and migrate towards opposite poles of the cell. The spindle of microtubules is formed. Nucleolus disappears and the nuclear envelope breaks.

2.4.2. Metaphase

The chromosomes become attached to the spindle at the kinetochore and align at the equatorial plate (metaphase plate). At this point the chromosomes are maximally condensed and are most visible.

2.4.3. Anaphase

Spindle fibers contract and separate the sister chromatids from each other. Then the sister chromatides move toward opposite poles of the cell.

2.4.4. Telophase

During this phase, the spindle disappears, chromosomes become less compact and gather at opposite ends of the cell. The nuclear envelope reforms and the nucleoli reappear. Then, cytoplasm divide by cytokinesis allowing creation of 2 daughter cells with identical genetic information.

2.5. Meiosis

Meiosis is the process by which the gametes are produced. Meiosis produces a total of four haploid cells from each original diploid cell. Meiosis provides an opportunity for the exchange of the genes (by crossing over) and, thus, causes the genetical variations among the species.

Like mitosis, meiosis is preceded by an interphase stage that includes G1, S, and G2 phases. Meiosis consists of two distinct divisions : meiosis I and meiosis II.

2.5.1. Meiosis I (reductional division)

The first meiotic division (meiosis I) is termed reductional division because the number of chromosomes per cell is reduced by half. Meiosis I includes the following four stages (Figure 2.3).

2.5.1.1. Prophase I

Prophase I is a stage of meiosis I in which chromosomes condense and pair, crossing over takes place, the nuclear membrane breaks down, and the mitotic spindle forms. It is a lengthy stage, divided into five substages :

- *Leptotene (threadlike) :* The chromosomes condense and become visible as long, thread-like structures
- *Zygotene (pairing)* : The chromosomes continue to condense, homologous chromosomes (Homologos) begin to pair up and begin synapsis. Each homologous pair of synapsed chromosomes consists of four chromatids called a bivalent or tetrad.
- *Pachytene (thickening)* : In pachytene, the chromosomes become shorter and thicker; Crossing over (genetic exchange) takes place.
- *Diplotene (appearing double) :* The synapsed homologous chromosomes begin to separate at centromere but remain attached at each chiasma.
- *Diakinesis (further condensation)* : In this stage, the chromosomes coil tightly, and so become shorter and thicker (Maximal condensation of chromosomes).

2.5.1.2. Metaphase I

In this phase, the bivalent aligns at the equatorial plate.

2.5.1.3. Anaphase I

In anaphase I, homologs separe from each other and move to opposite poles.

2.5.1.4. Telophase I

During this phase, a haploid set of chromosomes gather near each pole of the spindle. Cytoplasm divides by cytokinesis forming 2 haploid daughter cells with different genetic information.

2.5.2. Meiosis II (equational division)

The second meiotic division (meiosis II) is sometimes called the equational division because the chromosome number remains the same in each cell.

The second meiotic division follows the first without an interphase. It resembles mitosis. Meiosis II includes the following four stages (Figure 2.3):

2.5.2.1. Prophase II

In prophase II, the second-division spindles form, The nuclear membrane and the nucleous disappear.

2.5.2.2. Metaphase II

During metaphse II, the chromosomes aligh at metaphase plate

2.5.2.3. Anaphase II

In this stage, the centromeres divide and the chromatids of each chromosome move to opposite poles of the spindle.

2.5.2.4. Telophase II

In this phase, a haploid set of chromosomes gather near each pole of the spindle. The cytokinesis occurs, creating 4 haploid daughter cells that are genetically variable.

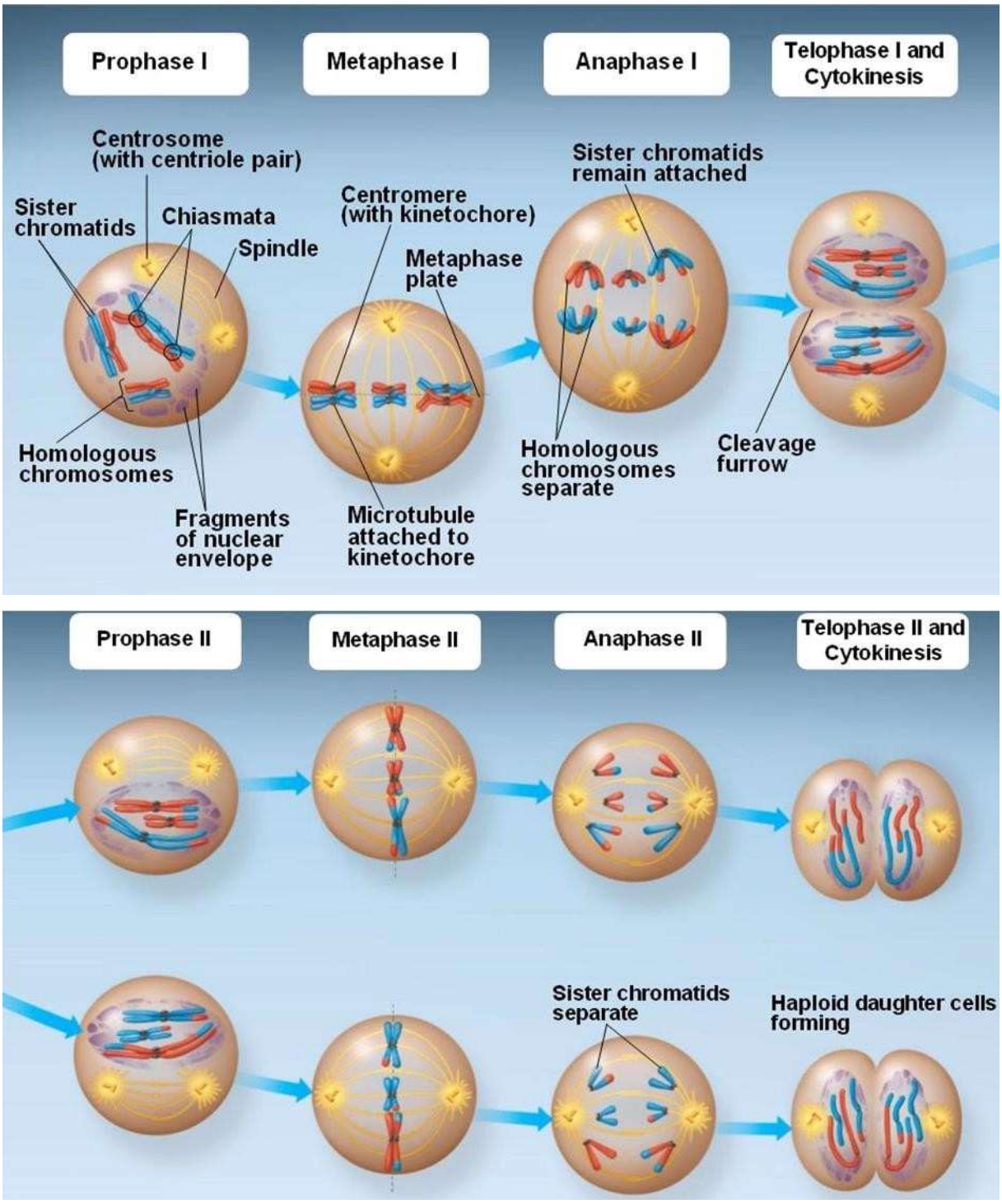


Figure 2.3. Meiosis stages

CHAPTER 3

Diploid Genetics

ABSTRACT

This chapter describes the basic principles of inheritance

3.1. Basic concepts

Diploid refers to the presence of two complete set of chromosomes in an organism's cells.

A *gene* is an inherited factor that determines a characteristic. Genes frequently come in different versions called alleles. All alleles for any particular gene will be found at a specific place on a chromosome called the locus (The plural of locus is loci).

The *genotype* is the set of alleles that an individual organism possesses. A diploid organism that possesses two identical alleles is homozygous for that locus. One that possesses two different alleles is heterozygous for the locus.

The *phenotype* is the manifestation or appearance of a characteristic. A phenotype can refer to any type of characteristic : physical, physiological, biochemical, or behavioral. Thus, the condition of having round seeds is a phenotype, a black hair is a phenotype, and having sickle-cell anemia is a phenotype.

A *true-breeding line (pure breeding line)* : is A group of genetically identical homozygous species that have been intercrossed for many generation and have retained a common phenotype.

3.2. Monohybrid Cross

A monohybrid cross is a cross between two parents that differ in a single charachteristic. Each parent exhibits one of the two contrasting forms (traits) of the character under study.

3.2.1. Mendel's experiment

Mendel studied the transmission of a single character of pea plants which is the color of the flower petals. He crossed a pea plant homozygous for purple petals with one that was homozygous for white petals. This first generation of a cross is the P (parental) generation. The offspring from the parents in the P generation are the F1 (first filial) generation. When Mendel examined the F1 of this cross, he found that they expressed only one of the phenotypes present in the parental generation: all the F1 were purple flowers (F1 100% purple).

Mendel carried out many times such crosses and always obtained this result. Then, he planted the F1 seeds, cultivated the plants that germinated from them, and allowed the plants to self-fertilize, producing a second generation (the F2 generation). Both of the traits from the P generation emerged in the F2. Mendel noticed that the number of the purple (705) and white (224) flowers constituted approximately a 3 to 1 ratio (3 :1) ; that is, about ¾ of the F2 plants had purple flower petals and ¼ had white flower petals. He concluded that each plant must

therefore possess two genetic factors coding for a character. The genetic factors that Mendel discovered (alleles) are, by convention, designated with letters ; the allele for purple is represented by P, and the allele for white by p. The plants in the P generation of Mendel's cross possessed two identical alleles : PP in the purple-flowered parent and pp in the white-flowered parent (Figure 3.1).

A second conclusion that Mendel drew from his monohybrid crosses was that the two alleles in each plant separate when gametes are formed, and one allele goes into each gamete. When two gametes (one from each parent) fuse to produce a zygote, the allele from the male parent unites with the allele from the female parent to produce the genotype of the offspring.

Thus, Mendel's F1 plants inherited an P allele from the purple-flowered plant and an p allele from the white-flowered plant.

However, only the trait encoded by purple allele (P) was observed in the F1. All the F1 progeny had purple flowers. Those traits that appeared unchanged in the F1 heterozygous offspring Mendel called dominant, and those traits that disappeared in the F1 heterozygous offspring he called recessive. When dominant and recessive alleles are present together, the recessive allele is masked, or suppressed. **The concept of dominance** was a third important conclusion that Mendel derived from his monohybrid crosses.

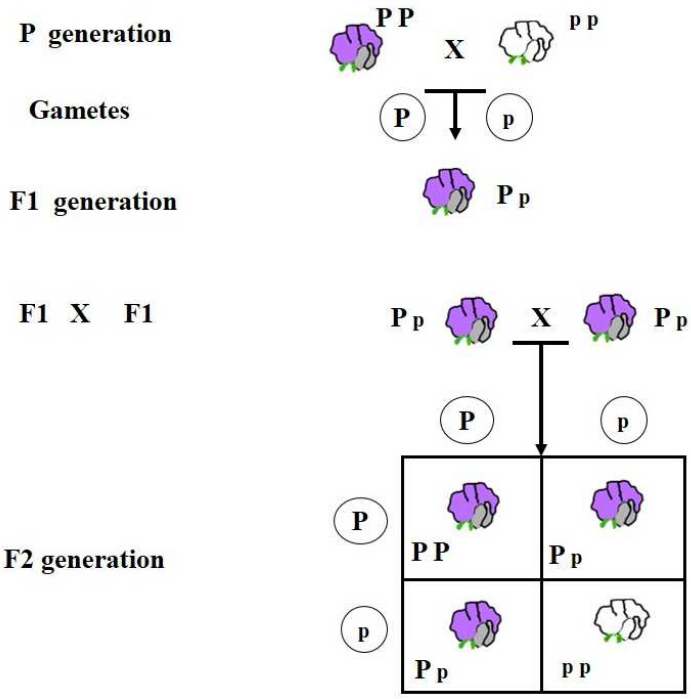


Figure 3.1. Monohybrid cross

Mendel's fourth conclusion was that the two alleles of an individual plant separate with equal probability into the gametes. When plants of the F1 (with genotype Pp) produced gametes, half of the gametes received the P allele for purple flowers and half received the p allele for white flowers. From this conclusion he derived the **the principle of segregation** (**Mendel's first law**) that states that each individual diploid organism possesses two alleles for any particular characteristic. These two alleles segregate (separate) when gametes are formed, and

one allele goes into each gamete. Furthermore, the two alleles segregate into gametes in equal proportions.

3.2.2. The allelic interaction

3.2.2.1. Complete dominance

In complete dominance one allele is dominant over the other. When dominant and recessive alleles are present together, the recessive allele is masked, or suppressed.

3.2.2.2. Incomplete dominance

In incomplete dominance, neither allele is dominant. The heterozygous individuals display an intermediate trait.

For example, in snapdragon plant there are two types of flower color : red and white. A cross between red and white flowered snapdragon will yield 100% pink offspring (intermediate flower color). This pattern of inheritance is described as incomplete dominance (Figure 3.2).

In F1, the genotypic and phenotypic ratios are identical. In the second generation, F2, the phenotypic and genotypic ratios are 1:2:1 (1/4, 2/4, 1/4).

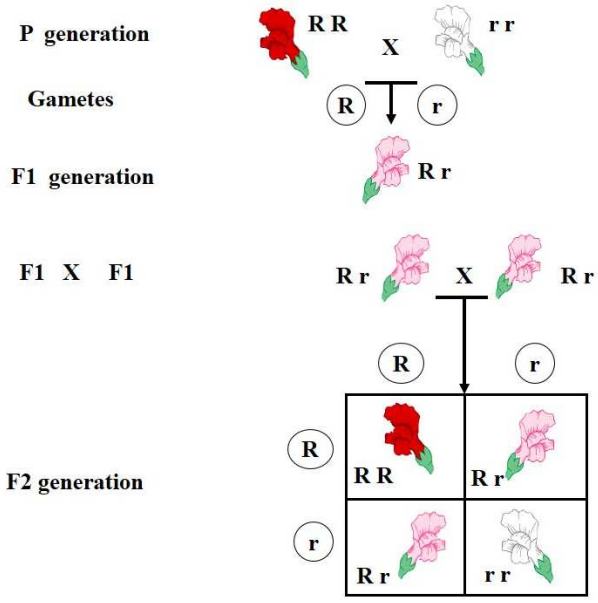


Figure 3.2. Incomplete dominance

3.2.2.3. Codominance

In case of codominance both alleles express their phenotypes in heterozygote. Both alleles are equally dominant.

The example is AB blood group in human (Figure 3.3). The people who have blood type AB are heterozygous exhibiting phenotypes for both the I^A and I^B alleles. The AB phenotype is due to the presence of an I^A allele and an I^B allele, which results in the production of A and B antigens on red blood cells.

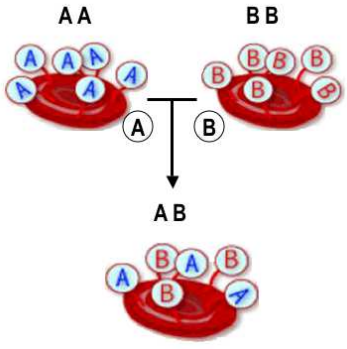


Figure 3.3. Codominance

3.2.2.4. Multiple Alleles

In living organisms there are sometimes more than two alleles governing the inheritance of a trait. This condition is called multiple alleles. The inheritance of characteristics encoded by multiple alleles is not different from the inheritance of characteristics encoded by two alleles, except that a greater variety of genotypes and phenotypes are possible. No more than two of the alleles are found in any one individual.

Examples of multiple alleles in plants, animals and humans can be seen, such as fur color of rabbits and human blood types.

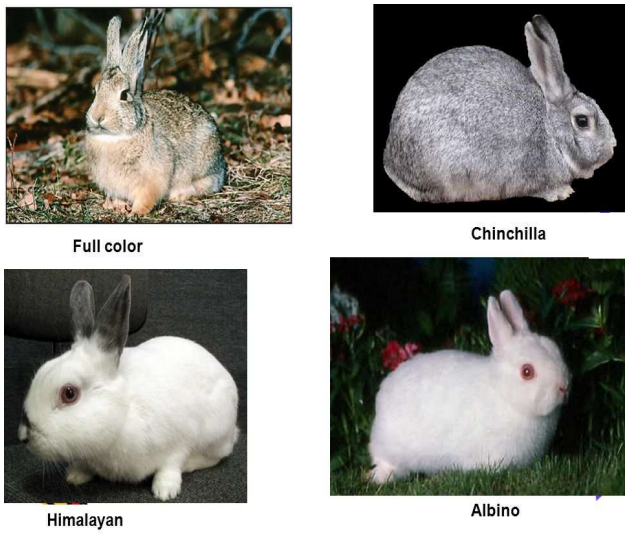


Figure 3.4. Example of Multiple alleles

- The fur color of rabbits

In rabbits, the four common alleles of C gene are C, cch, ch and c. A rabbit with genotype CC has Agouti fur (full color), cchcch have chinchilla or greyish coat color, chch have himalayan pattern and cc have white coat color (albino). Due to multiple alleles, there are many combinations of alleles possible.

- The ABO Blood Group system

The three common alleles for the ABO blood group locus are : I^A codes for the A antigen; I^B codes for the B antigen; and i codes for no antigen (O). The dominance relations among the ABO alleles is as follows : $I^A > i$, $I^B > i$ and $I^A = I^B$. The I^A and I^B alleles are both dominant over i and are codominant with each other.

3.2.2.5. Lethal alleles

A lethal allele is one that causes death at an early stage of development, often before birth, and so some genetypes may not appear among the progeny. In homozygous condition, the gene causes the death of its carrier.

In 1905, Lucien Cuenot reported a peculiar (strange) pattern of inherance in mice. When he mated two yellow mice, approximately 2/3 of their offspring were yellow and 1/3 were

nonyellow. When he test-crossed the yellow mice, he found that all were heterozygous; he was never able to obtain a homozygous yellow mouse.

Cuenot originally crossed two mice heterozygous for yellow: Yy x Yy. Normally, this cross would be expected to produce YY, Yy, and yy. The homozygous YY mice are conceived but never complete development, which leaves a 2 : 1 ratio of Yy (yellow) to yy (nonyellow) in the observed offspring; all yellow mice are heterozygous (Yy).

So the normal segregation ratio of 3:1 (3/4, 1/4) is modified into 2:1 (2/3, 1/3) ratio.

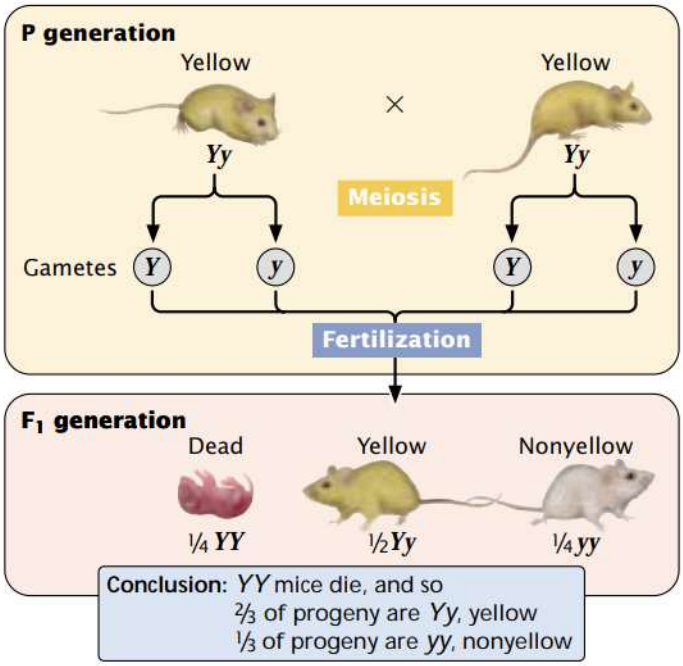


Figure 3.5. Lethal allele

3.2.3. Test cross

To determine the genotype of a specific organism, a test cross can be performed, in which the organism with an uncertain (unknown) genotype is crossed with an organism that is homozygous recessive. Thus, test cross is used to determine whether an organism with a dominant phenotype is homozygous dominant or heterozygous.

For example, if you were given a pea plant with purple flowers it might be a homozygote (PP) or a heterozygote (Pp). You could cross this purple-flowered plant to a white-flowered plant as a tester, since we know the genotype of the tester is pp. Depending on the genotype of the purple-flowered parent (Figure 3.6), we will observe different phenotypic ratios in the F generation. If the purple-flowered parent was a homozgyote, all of the F progeny will be purple. If the purple-flowered parent was a heterozygote, the F progeny should segregate purple-flowered and white-flowered plants in a 1:1 ratio (1/2, 1/2).

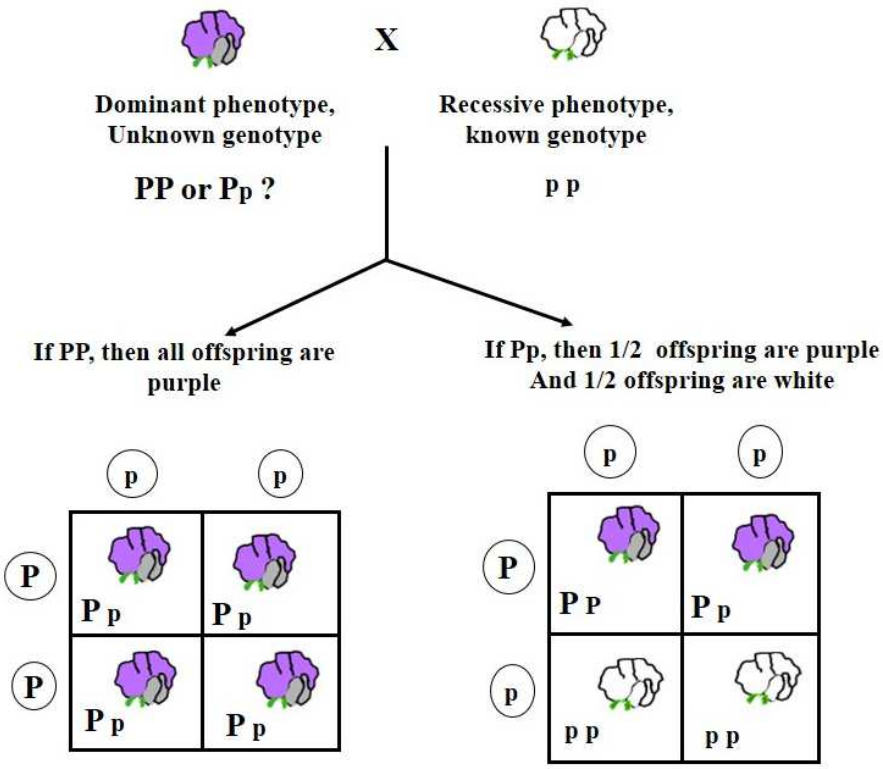


Figure 3.6. Test cross

3.2.4. Back-cross

Back-crosss or Back-crossing is the process of crossing or mating a hybrid offspring (F1 hybrids) with one of its parents or an organism genetically similar to its parent.

In such back cross, when F1 is back crossed to the parent with dominant phenotype, no recessive individuals are obtained in the progeny. However, when it is crossed with recessive parent, both phenotypes appear in the progeny.

3.3. Dihybrid cross

A dihybrid cross is a cross between two parents that differ in two characteristics.

3.3.1. Independent assortment

The principle of independent assortment states that alleles at different loci separate independently of one another.

In one of his hybridization experiment Mendel crossed a homozygous pea plant having yellow round seeds with the homozygous pea plant having green wrinkled seeds (Figure 3.7). The F1 hybrids were found to have yellow round seeds. When the F1 hybrids were allowed to cross among themselves they produced four types of seeds in the ratio of 9 : 3 : 3 : 1 given as follows:

Yellow Round 9

Yellow Wrinkled 3

Green Round 3

Green Wrinkled 1.

Beside getting the ratio of 3 : 1 (3/4, 1/4) of the monohybrid crosses, Mendel got the ratio of 9 : 3 : 3 : 1. This irregularity in the ratio of F2 offspring was explained by Mendel stating that "when the parents differ from each other in two or more pairs of contrasting characters or factors then the inheritance of one pair of factors is independent to that of the other pair of factors".

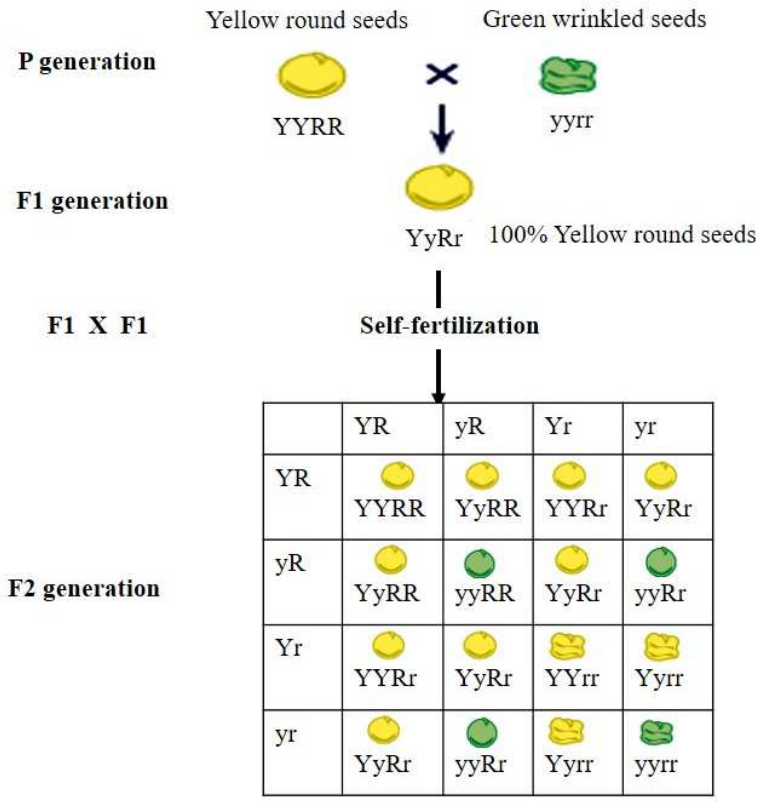


Figure 3.7. Dihybrid cross

This is the Mendel's law of independent assortment. Which means that genes segregate independently during gamete formation.

The phenotypic ratio of 9 : 3 : 3 : 1 (9/16, 3/16, 3/16, 1/16) in F2 implies the idependent assortment of two genes => genes are located on nonhomologous chromosomes.

3.3.3. Phenotypic ratio of back-cross :

The back-cross between F1 offspring (yellow rounded seeds) and a homozygous recessive parent (green Wrinkled seeds) produce four types of seeds in the ratio of 1 : 1 : 1 : 1 (1/4, 1/4, 1/4, 1/4) given as follows :

Yellow Round 1/4
Green Wrinkled 1/4
} **½ parental types (PT)**

Yellow Wrinkled 1/4
Green Round 1/4
} **½ recombinant types (RT)**

When PT% = RT% = 50%, in this case, the genes are unlinked (independent). This means they are located on nonhomologous chromosomes or they are so far apart in the chromosome.

3.3.2. Linked genes

3.3.2.1. Back-cross linkage

In the cross between homozygous tomato plants having purple fruits and hairy stems with the homozygous tomato plant having red fruits and smooth stems, the F1 hybrids were found to have purple fruits and hairy stems. This means that the phenotype purple fruits and hairy stems is dominant over the phenotype red fruits and smooth stems.

Thus : Purple fruits and hairy stems (PPHH)
Red fruits and smooth stems (pphh)

To determine whether the genes are linked, a back cross is performed between the hybrid (F1 hybrids (PpHh) with the double recessive parent (pphh). We obtained four types of tomato plants :

220 Purple fruits, hairy stems
210 Red fruits, smooth stems
} **parental types (PT)**

32 Purple fruits, smooth stems
38 Red fruits, hairy stems
} **Recombinant types (RT)**

Out of a total of 500 plants, 430 plants are parental types and 70 plants are recombinants.
Parental type % = 430/500 x 100 = 86%
Recombinants % = 70/500 x 100 = 14%
Thus, the frequency of recombination (R%) = 14%
RT % < PT => the genes are linked, which means they are located on the same chromosome.

The recombination between genes takes place by an exchange of segments between homologous chromosomes in the process of crossing-over during prophase I.

3.3.2.2. Frequency of recombination

The minimum recombination frequency between two genes is 0. The recombination frequency also has a maximum : no matter how far apart two genes may be, the maximum frequency of recombination between any two genes is 50 percent. Fifty percent recombination is the same value that would be observed if the genes were on nonhomologous chromosomes and assorted independently.

This implies the following principle :

Genes with recombination frequencies smaller than 50 percent are present in the same chromosome (linked). Two genes that undergo independent assortment, indicated by a recombination frequency equal to 50 percent, either are in nonhomologous chromosomes or are located far apart in a single chromosome.

3.3.2.3. Linkage map

The linkage of the genes in a chromosome can be represented in the form of a genetic map, which shows the linear order of the genes along the chromosome with the distances between adjacent genes proportional to the frequency of recombination between them. A genetic map is also called a linkage map or a chromosome map.

The unit of distance in a genetic map is called a map unit; 1 map unit is equal to 1 percent recombination. One map unit is also called a centimorgan, abbreviated cM, in honor of Morgan. A distance of 3.5 map units therefore equals 3.5 centimorgans and indicates 3.5 percent recombination between the genes (between two loci).

3.4. Trihybrid Cross

A trihybrid cross is a cross between organisms that differ in three characteristics. It gives us information regarding relative distances between these genes, and also shows us the linear order in which these genes should be present on chromosome.

Suppose that we back-cross a heterozygous organism having the genotype AaNnRr and a homozygous recessive organism aannrr and find in the progeny the following 8 phenotypes :

347 ANR
357 anr } **Parental type**

52 ANr
49 anR } Single Crossing-over between N and R

90 Anr
92 aNR } Single crossing-over between A and N

6 AnR
7 aNr } Double C.O (1 C.O between A and N, and 1C.O between N and R)

} **Recombinants types**

1) **<u>Map distance between A and N [D (A, N)] :</u>**

To find the distance between A and N, we must count all crossing-over (both singles and

doubles) that occured in the region between A and N.

1 C.O between A and N ⇒ RT% = [(90+ 92)/ 1000] x 100 = 18.2%

double C.O between **A and N**, and between N and R ⇒ RT% = [(6+ 7)/ 1000] x 100 = 1.3%

D (A, N)= 18.2% + 1.3% = 19.5% of recombination (19.5 cM between the loci A and N).

2) <u>Map distance between N and R [D (N, R)]</u>

To find the distance between N and R, we must count all crossing-over (both singles and doubles) that occured in the region between N and R.

1 C.O between N and R ⇒ RT% = [(52+ 49)/ 1000] x 100 = 10.1%

double C.O between A and N, and between **N and R** ⇒ TR% = [(6+ 7)/ 1000] x 100 = 1.3%

D (N, R) = 10.1% + 1.3%= 11.4% of recombination or 11.4 cM.

3) <u>Map distance between A and R [D (A, R)]</u>

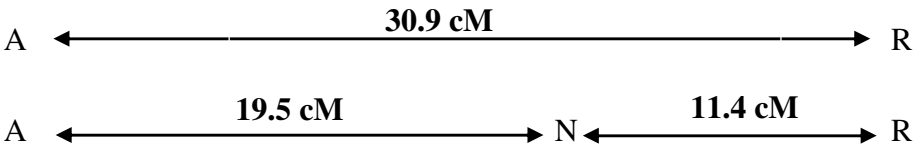
To find the distance between A and R, we must count all the recombinants.

We also must count twice the doubles recombinants.

D (A, R) = {[90+92+52+49+ 2x (6+7)]/1000} x100

D (A, R) = 30.9% of recombination or 30.9 cM.

All the three RT are below 50% ⇒ genes are linked and are located on the same chromosomes.



3.5. Sex Determination and Sex-Linked inherance

3.5.1. Sex determination systems

Organisms employ a variety of mechanisms to determine sex, whether an individual is male or female.

In XX-XO system, females have two X chromosomes (XX), and males possess a single X chromosome (XO). There is no O chromosome; the letter O signifies the absence of a sex chromosome. Because males produce two different types of gametes with respect to the sex chromosomes, they are said to be the *heterogametic sex*. Females, which produce gametes that

are all the same with respect to the sex chromosomes, are the *homogametic sex*. The XX-XO system is found in grasshoppers.

In XX-XY system, the cells of females have two X chromosomes (XX) and the cells of males have a single X chromosome and a smaller sex chromosome called the Y chromosome (XY). In humans and many other organisms, the Y chromosome is acrocentric. In this type of sex-determining system, the male is the heterogametic sex (half of his gametes have an X chromosome and half have a Y chromosome), however, the female is the homogametic sex since all her egg cells contain a single X chromosome. Many organisms, including some plants, insects, and reptiles, and all mammals (including humans), have the XX-XY sex-determining system.

In ZZ-ZW sex determination system, the female is heterogametic and the male is homogametic. The sex chromosomes in this system are labeled Z and W, but the chromosomes do not resemble Zs and Ws. Females in this system are ZW; after meiosis, half of the eggs have a Z chromosome and the other half have a W. Males are ZZ; all sperm contain a single Z chromosome. The ZZ-ZW system is found in birds, butterflies moths, some amphibians, and some fishes.

3.5.2. Sex-linked inheritance

Sex-linked characteristics are determined by genes located on the sex chromosomes. Genes on the X chromosome determine X-linked characteristics; whereas those on the Y chromosome determine Y-linked characteristics. Because little genetic information exists on the Y chromosome in many organisms, most sex-linked characteristics are X linked. Males and females differ in their sex chromosomes; so the pattern of inheritance for sex-linked characteristics differs from that exhibited by genes located on autosomal chromosomes.

3.5.2.1. The X-linked recessives genes

The X-linked recessives genes can be detected in human pedigrees through the following clues:

- The X-linked recessive phenotype is usually found more frequently in the male than in the female. This is because an affected female can result only when both mother and father bear the X-linked recessive allele (e.g., $X^A X^a \times X^a Y$), whereas an affected male can result when only the mother carries the gene.
- Further, if the recessive X-linked gene is very rare, almost all observed cases will occur in males.

- Usually none of the offspring of an affected male will be affected, but all his daughters will carry the gene in masked heterozygous condition, so one half of their sons will be affected.
- None of the sons of an affected male will inherit the X-linked recessive gene, so not only will they be free of the defective phenotype; but they will not pass the gene along to their offspring.

Example : X-Linked Color Blindness in Humans

The most common types of human color blindness are caused by defects of the red and green pigment. Mutations that produce defective color vision are generally recessive and, because the genes (C) coding for the red and green pigments are located on the X chromosome, red–green color blindness is inherited as an X-linked recessive characteristic.

The possible genotypes are as follows :

X^CX^C : female with normal vision.

X^CX^c : female with normal vision but carries the color blind allele.

X^cX^c : color-blind female

X^CY : male with normal vision

X^cY : color-blind male

3.5.2.2. Dominant X-linked genes

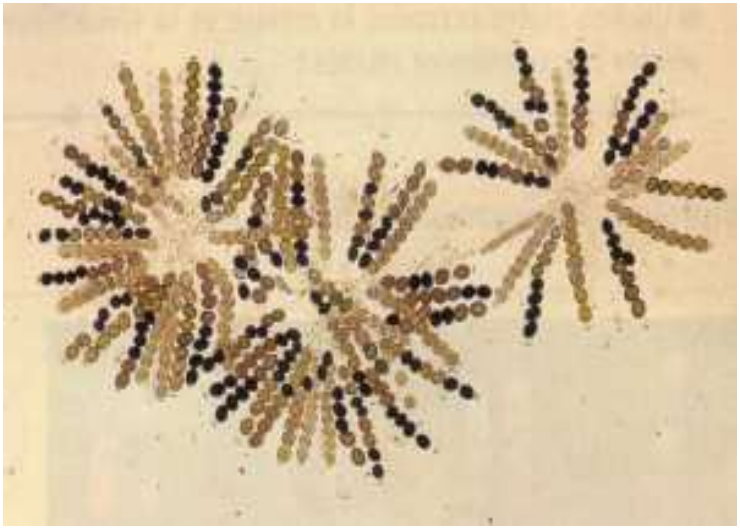
Dominant X-linked genes can be detected in human pedigrees through the following clues :

- It is more frequently found in the female than in the male of the species.
- The affected males pass the condition on to all of their daughters but to none of their sons.
- Females usually pass the condition (defective phenotype) on to one-half of their sons and daughters.

3.5.3. Pedigree

A pedigrees is a diagram that shows the relationships among the members of a family, which can be used to determine the pattern of certain genes or disease inheritance within a family. It is customary to represent males as squares and females as circles. A horizontal line connecting a circle and a square represents a mating. The offspring of the mating are shown beneath the mates, starting with the first born at the left and proceeding through the birth order to the right.

Individuals that have a genetic condition are indicated by coloring or shading. The generations in a pedigree are usually denoted by Roman numerals.



CHAPTER 4

Haploid Genetics

ABSTRACT

This chapter provides information about the linkage of gene-centromere and between two genes in ordered tetrads.

4.1. Introduction

Haploid refers to the presence of a single set of chromosomes in an organism's cells. In humans, only the gametes (egg and sperm cells) are haploid (n=23).

4.2. The analysis of ordered tetrads

Tetrad analysis is the technique of using each of the individual spores of a tetrad for analysis to acquire information on linkage and recombination .

Neurospora crassa is a species of fungi used extensively in genetic investigations. It is a haploid organism.

4.2.1. Life cycle of *Neurospora crassa*

Neurosopra crassa reproduces asexually and also sexually (Figure 4.1).

There are two mating types : A and a. The 2 haploid nuclei fuse to give a diploid nucleus (zygote nucleus), the only diploid phase to this life cycle. The zygote nucleus undergoes meiosis and produces 4 haploid spores within an ascus. Subsequent mitotic division produces 8 haploid, sexual ascospores in a linear arrangement (ordered array). All the ascospore are found together in the ascus, and, the order of the 4 spore pairs is the same as the order of the 4 chromatids of each tetrad during Metaphase I.

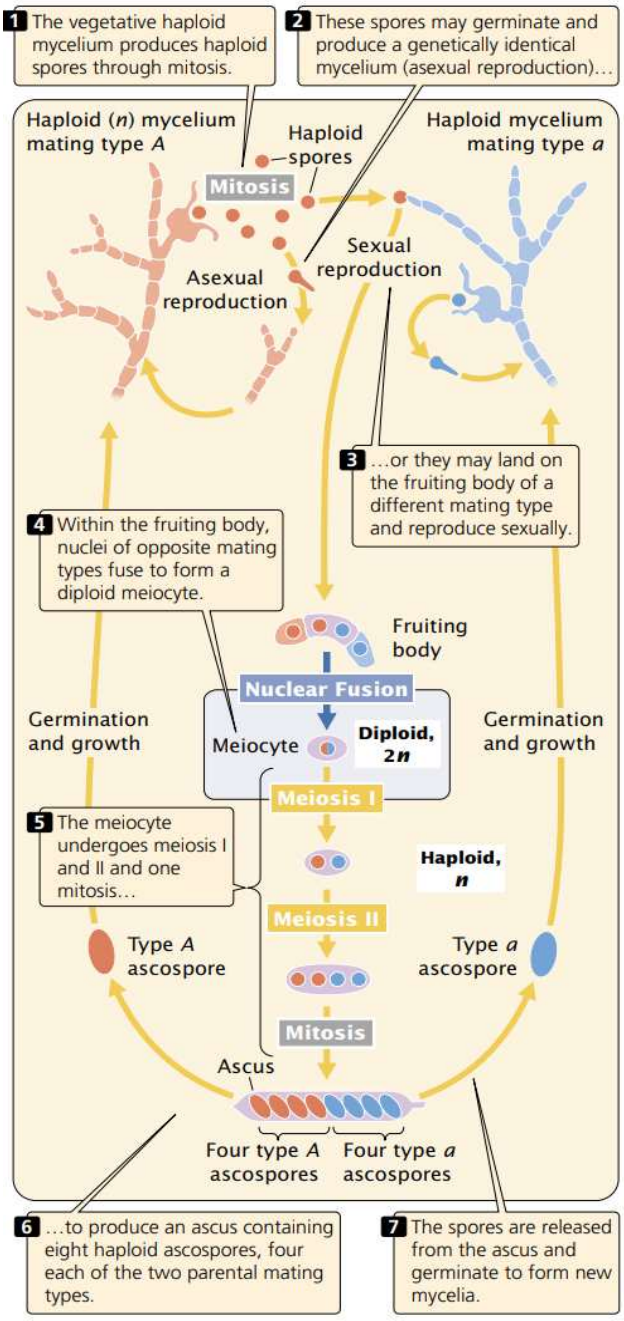


Figure 4.1. Life cycle of *Neurospora*

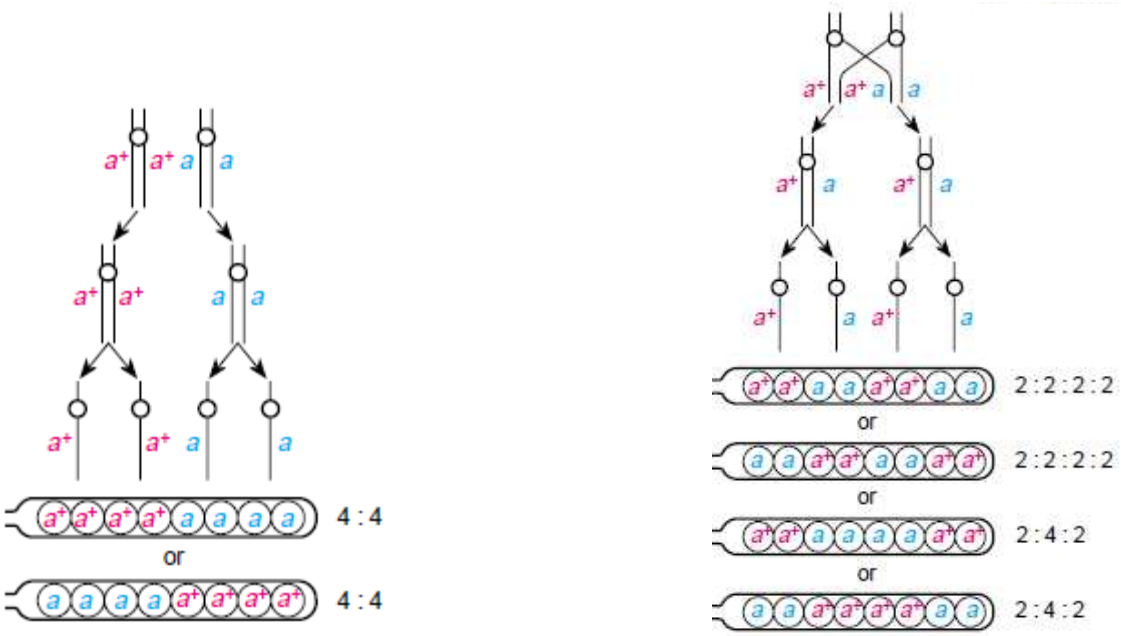
4.2.2. Linkage of gene-centromere in ordered tetrad

In *Neurospora*, a cross between normal (a+) and mutant (a) strain, will give rise to different type of asci (Figure 4.2).

In the absence of crossing-over between a gene and its centromere, the alleles of the gene (for example, a+ and a) must separate in the first meiotic division; this separation is called first-division segregation. It leads to 4 : 4 arrangement of spores (only two possible arrangements of the products of meiosis can yield first-division segregation).

If a crossing over is formed between the gene and its centromere, the a+ and a alleles do not become separated until the second meiotic division; this separation is called second-division segregation. Four patterns of second-division segregation are possible.

N.B : Centromeres of homologous parental chromosomes separate at the first meiotic division, while the centromeres of sister chromatids separate at the second meiotic division.



First division segregation with no crossing over between gene and its centromere

Second division segregation with crossing over between gene and its centromere

Figure 4.2. First-division and second-division segregation

4.2.2.1. Gene-centromere distance

The ordered arrangement of meiotic products makes it possible to determine the recombination frequency between any particular gene and its centromere.

The percentage of asci with second-division segregation patterns for a gene can be used to map the gene with respect to its centromere.

Since, in each cell in which crossing-over takes place, two of the chromatids are recombinant and two are nonrecombinant, the map distance can be calculated using the following formula :

$$\text{Map distance} = \frac{\text{½ Asci with second division segregation}}{\text{Total number of Asci}} \times 100$$

4.2.3. Linkage of two genes in ordered tetrad

In the tetrads, when two pairs of alleles are segregating, three patterns of segregation are possible.

For example, in the cross AB × ab, the three types of tetrads are :

- AB AB ab ab referred to as parental ditype, or PD. Only two genotypes are represented, and their alleles have the same combinations found in the parents (two parental genotypes).
- Ab Ab aB aB referred to as nonparental ditype, or NPD. Only two genotypes are represented, but their alleles have nonparental combinations (only recombinant combinations).
- AB Ab aB ab referred to as tetratype, or TT. All four of the possible genotypes are present.

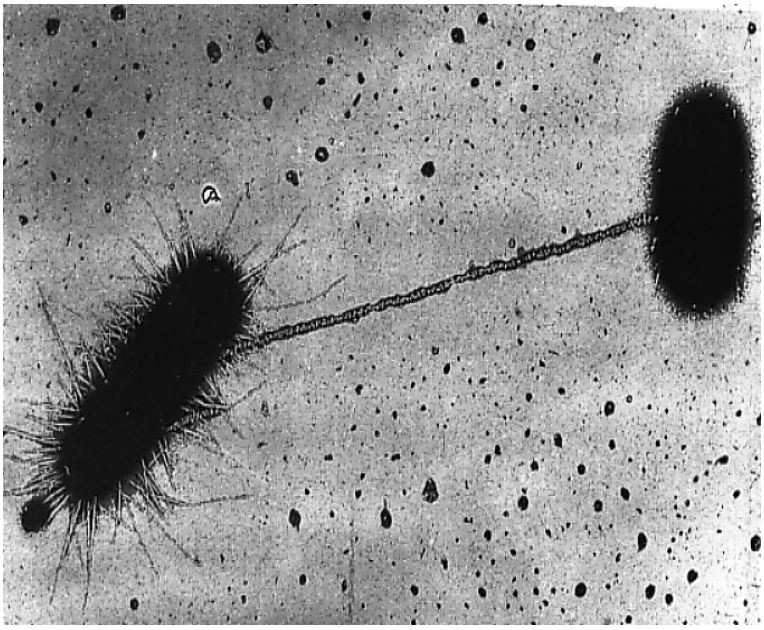
PD	NPD	TT
AB	Ab	AB
AB	Ab	Ab
ab	aB	aB
ab	aB	ab

Linkage is indicated when nonparental ditype tetrads appear with a much lower frequency than parental ditype tetrads (NPD << PD). In this case, the genes are located on the same chromosome.

When genes are unlinked, the parental ditype tetrads and nonparental ditype tetrads are expected in equal frequencies (PD = NPD). In this case, the genes either they are located on different chromosomes, or they are far apart on the same chromosome.

The map distance is calculated using the following equation :

$$\text{Map distance} = \frac{\sum \text{NPD} + \sum \text{TT}/2}{\text{Total number of tetrads}} \times 100$$



CHAPTER 5

Bacterial Genetics

ABSTRACT

This chapter provides information about the mechanisms of transfer of genetic material between bacteria.

5.1. Introduction

Bacteria are single-celled (unicelular) prokaryotic organisms. They have a simple structure with a plasma membrane ; most contain a single double-stranded circular DNA molecule (a few have more) and no nucleus. Bacteria contain additional small DNA molecules that can be either independent structures, plasmids (capable of self-replicating independently of the bacterial chromosome). Plasmids carry genes that could confer useful properties to the bacteria. There are many different plasmids : resistance (R) plasmids, fertility (F) plasmids, degradative plasmids, virulence plasmids, etc. A single bacterium can contain several types of plasmids.

Wild-type, or prototrophic, bacteria can use simple ingredients to synthesize all the compounds that they need for growth and reproduction. For instance, strain of genotype (Leu+) are able to grow on a medium lacks such amino acid because they are able to synthesize it. However, mutant strains called auxotrophs lack one or more enzymes necessary for synthesizing essential compounds and thus, will grow only on medium supplemented with those essential molecules. For example, strain of genotype (Leu^-) are auxotrophic strains, because they are unable to synthesize the amino acid leucine and will grow only on a medium supplemented with leucine.

A medium that contains all the substances required by bacteria for growth and reproduction is called a complete medium. A medium that contains only the nutrients required by prototrophic bacteria is termed a minimal medium. Hence, prototrophic strains are able to grow on a minimal medium, whereas auxotrophic bacteria cannot grow on minimal medium but instead they grow on a complete medium.

5.2. Mechanisms of gene transfer in bacteria

Bacteria can exchange genetic material by three mechanisms : transformation, conjugation and transduction.

Transfer of genes between bacteria in all mechanisms is an unidirectional DNA transfer, from donor bacterium to recipient bacterium.

5.2.1. Transformation

In transformation, small pieces of naked extracellular DNA are taken up by a living bacterial cell, then incorporated into the bacterial chromosome trough recombination. The recipient bacterium has to be competent (able to incorporate exogenous DNA fragment). Most bacterial cells are only competent during a restricted part of their life cycle. During the competent state, the cell produces one or more proteins called *competence factors* that modify the cell wall so

it can bind exogenous DNA fragments. Thus, recepror sites are present only during competent state.

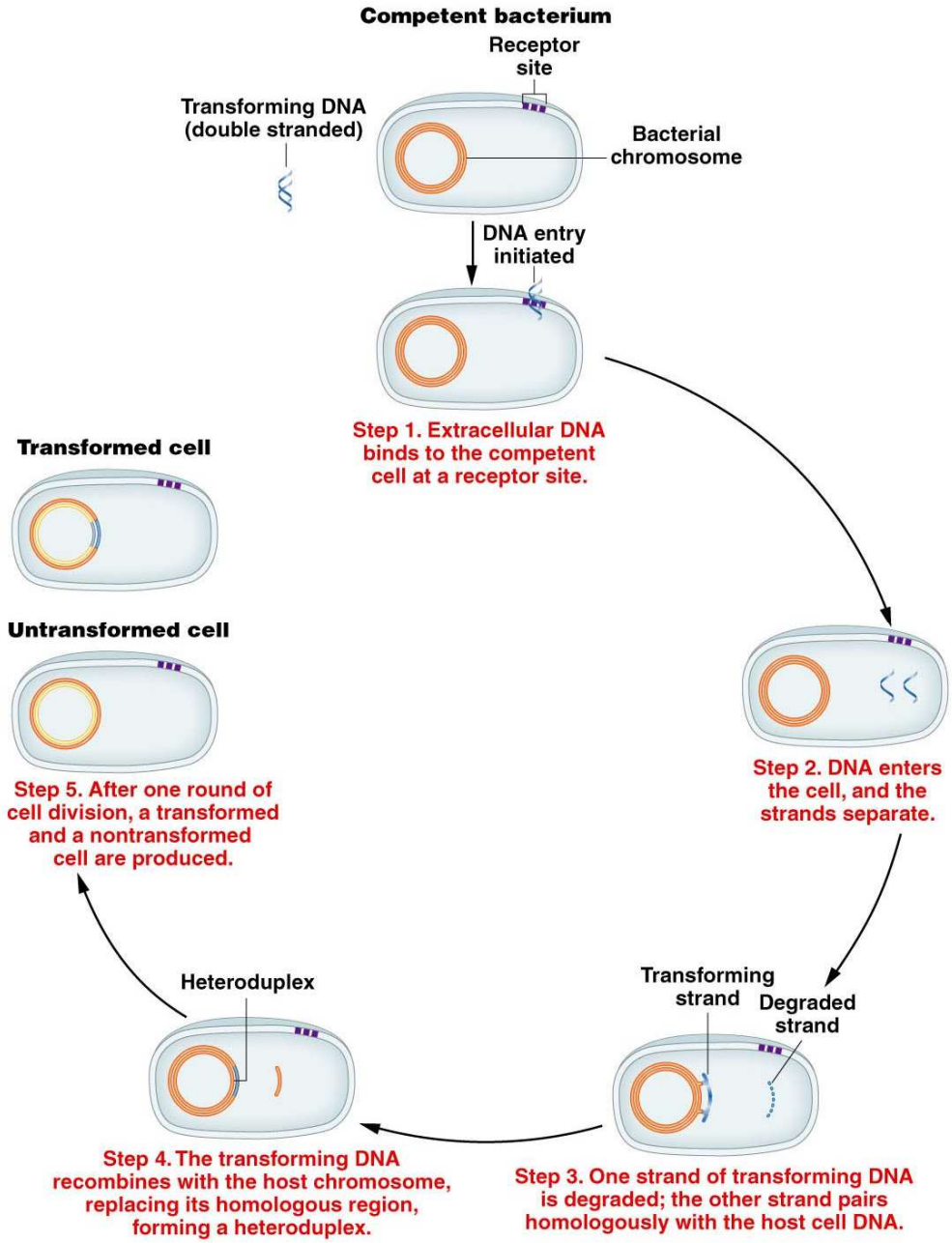


Figure 5.1. Transformation mechanism

As a DNA fragment enters the cell in the process of transformation, one of the strands is hydrolyzed, whereas the other strand associates with proteins as it moves across the membrane. Once inside the cell, this single strand may pair with a homologous region and become integrated into the bacterial chromosome (Figure 5.1). This integration requires two crossover events, after which the remaining single-stranded DNA is degraded by bacterial enzymes.

Transformation in bacteria was first demonstrated in 1928 by Frederick Griffith, in his famous 'transforming principle' experiment. Microbiologists identified two strains of the bacterium *Streptococcus pneumoniae*. The R-strain produced rough colonies on a bacterial plate, while the other S-strain was smooth. More importantly, the S-strain bacteria caused fatal infections when injected into mice, while the R-strain did not (Figure 5.2). Griffith noticed that upon mixing "heat-treated" S-strain cells together with some R-type bacteria, the mice died and there were S-strain, pathogenic cells recoverable. Thus, some non-living component from the S-type strains contained genetic information that could be transferred to and transform the living R-type strain cells into S-type cells.

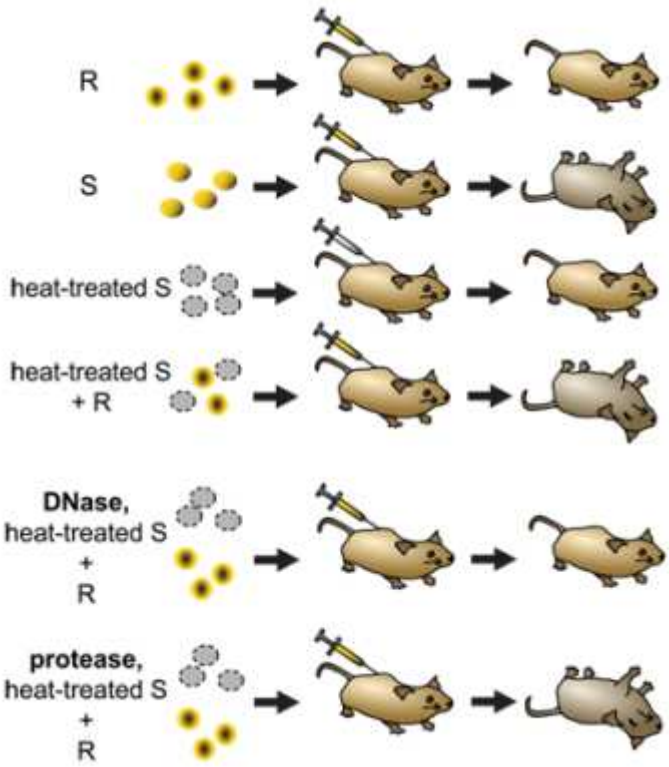


Figure 5.2. Griffith's transformation experiment

After, Avery, MacLeod and McCarty's in 1944 determined the molecule responsible for transformation. The researchers separated the S-type cells into various components, such as proteins, polysaccharides, lipids, and nucleic acids. Only the nucleic acids from S-type cells were able to make the R-strains smooth and fatal. Furthermore, when cellular extracts of S-type cells were treated with DNase (an enzyme that digests DNA), the transformation ability was lost. The researchers therefore concluded that DNA was the genetic material, which in this case controlled the appearance (smooth or rough) and pathogenicity of the bacteria.

5.2.2. Conjugation

Conjugation is a process in which DNA is transferred from a bacterial donor cell to a recipient cell by direct physical contact between cells. It has been observed in many bacterial species and is best understood in *E. coli*, in which it was discovered.

In most bacteria, conjugation depends on a fertility (F) factor that is present in the donor cell and absent in the recipient cell. The cells that carry the sex factor F are known as F+ or male, whereas females do not have the sex factor and are symbolized F^-.

5.2.2.1. Conjugation F+ X F^-

The F factor contains an origin of replication (oriT) and a number of genes required for conjugation. For example, some of these genes encode sex pili (singular, pilus), slender extensions of the cell membrane. A cell containing F produces the sex pili, which makes contact with a receptor on an F^- cell and pulls the two cells together.

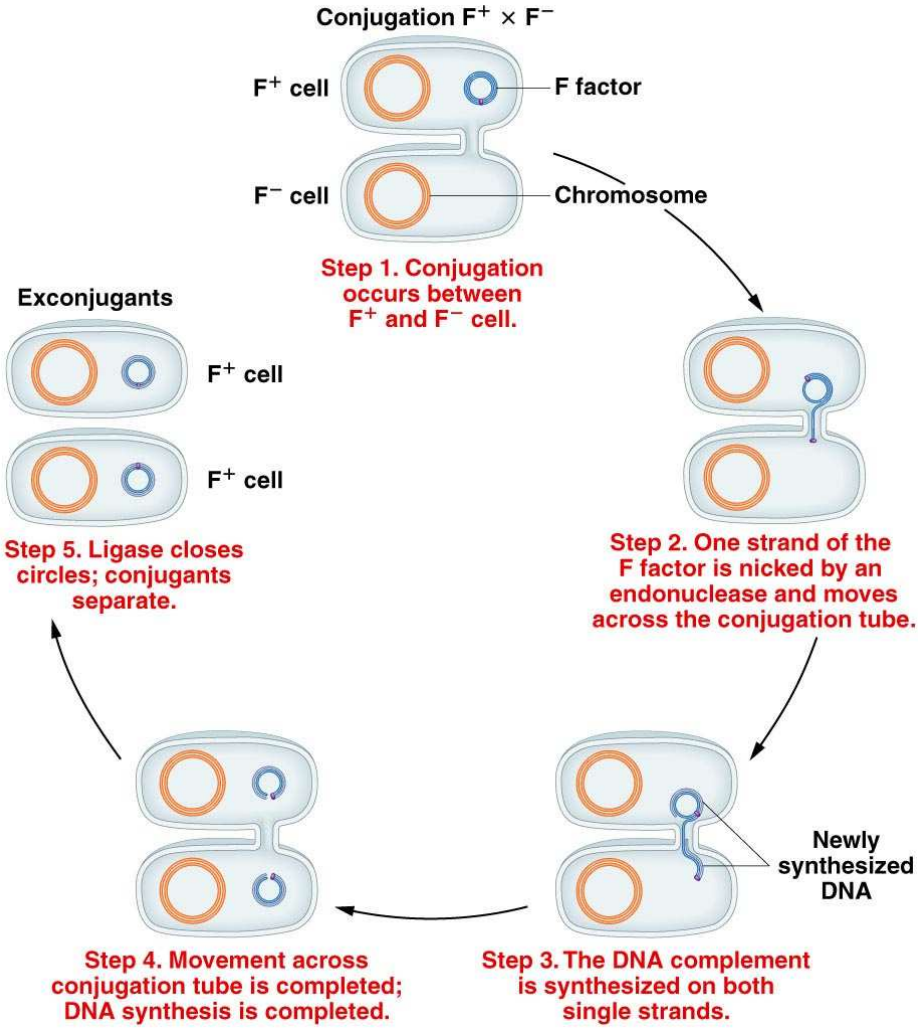


Figure 5.3. Conjugation F+ X F^-

DNA is then transferred from the F+ cell to the F^- cell. Once inside the recipient cell, the single strand is replicated, producing a circular, double-stranded copy of the F plasmid. If the entire F factor is transferred to the recipient F^- cell, that cell becomes an F+ cell (Figure 5.3).

Conjugation can take place only between a cell that possesses F and a cell that lacks F. Transfer of the entire *E. coli* chromosome requires about 100 minutes; if conjugation is interrupted before 100 minutes have elapsed, only part of the chromosome will pass into the F^- cell and have an opportunity to recombine with the recipient chromosome.

5.2.2.2. Conjugation Hfr X F^-

Hfr stands for high frequency of recombination, which refers to the relatively high frequency with which donor genes are transferred to the recipient.

In Hfr (high-frequency) strains, the F factor is integrated into the bacterial chromosome. Hfr cells behave as F+ cells, forming sex pili and undergoing conjugation with F^- cells. The stages of conjugation between Hfr and F^- strains are shown in figure 5.4.

In Hfr X F^- conjugation, chromosomal genes of Hfr are transferred and the remaining part of F is the last DNA to enter the F^- cell. Because the conjugating cells usually break apart long before the entire bacterial chromosome is transferred, the final segment of F is almost never transferred into the recipient (F^- cell). Therefore, F^- cells remain F^-.

In Hfr transfer, some regions in the transferred DNA fragment become incorporated into the recipient chromosome. The incorporated regions replace homologous regions in the recipient chromosome. The result is that some F- cells become recombinants containing one or more genes from the Hfr donor cell. For example, in a mating between Hfr leu+ and F^- leu-, some F^- leu+ cells arise. However, the genotype of the donor Hfr cell remains unchanged.

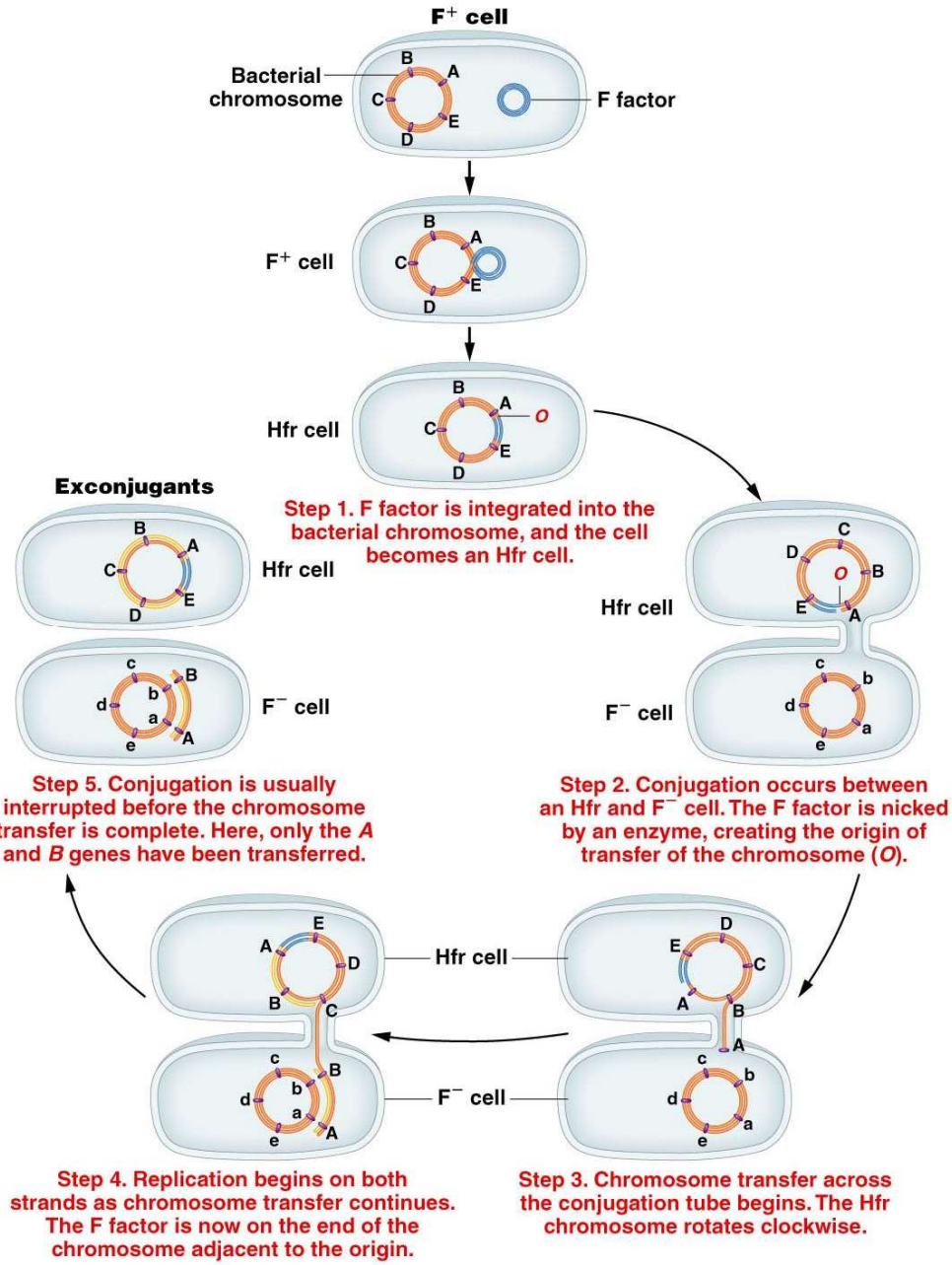


Figure 5.4. Conjugation Hfr X F–

5.2.2.5. Interrupted conjugation

Gene mapping can be done by interrupting the conjugation at multiple times and taking account of which genes the recipient cell has. Genes that take longer to enter the recipient cell are further from the origin. The percentage of recipient cells with a given gene also provides information on the order of genes on a chromosome. The sex pilus can randomly break at any time. When a gene takes a longer amount of time to enter a recipient cell (because it is further from the

origin), there will be more chances that the sex pilus could break before it is donated. Because there are more chances for conjugation to stop, less recipient cells have genes far from the origin. The fertility factor Hfr is the last gene to enter a recipient cell so very few recipient cells receive it.

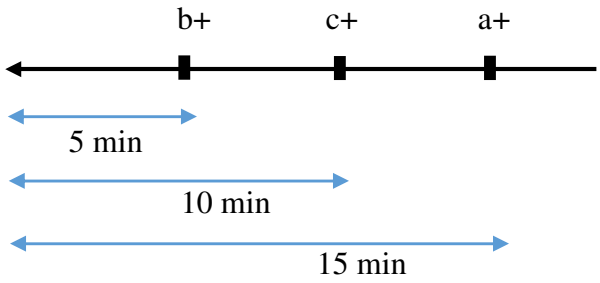
In this technique, Hfr and F^- strains are mixed and allowed to conjugate for a short period of time. The conjugation is stopped at any desired time by agitation of the mixture which breaks the conjugation tube. The mixture is diluted immediately and plated at selective media, incubated and then scored for recombinants. The lengh of donor chromosome transmitted could then be determined and mapped in terms of time units required for the transfer.

Example

An Hfr bacterial strains carrying the protophic markers such as a+, b+ and c+ is mixed with a F^- strain carrying the auxotrophic alleles a, b and c. The conjugation was interrupted at 5 minutes intervals and plated on media which revealed the presence of recombinants :

Time (minutes)	Recombinants
5	a b+ c
10	a b+ c+
15	a+ b+ c+

The order of genes in the donor Hfr strain is :



5.2.3. Transduction

Transduction is defined as the transfer of genetic information between cells through the mediation of a virus (bacteriophage) particle. It therefore does not require cell to cell contact.

Viruses that infect bacterial cells are called bacteriophages or simply phages. Most viruses consist of a single molecule of genetic material (DNA or RNA) enclosed in a protective coat composed of one or more kinds of protein molecules; however, their size, molecular constituents, and structural complexity vary greatly.

Bacteriophages have two alternative life cycles : the lytic and the lysogenic cycle. In addition, there are two types of bacteriophage :

- Virulent phages reproduce strictly through the lytic cycle in which the phages infects a susceptible cell and multiplies within it to form a large number of phages and then the new phages are released by lysing the host cell. This type always kill their host cells.
- Temperate phage can utilize either the lytic or the lysogenic cycle. The lysogenic cycle begins like the lytic cycle but, inside the cell, the phage DNA integrates into the bacterial chromosome, where it remains as an inactive prophage. The prophage is replicated along with the bacterial DNA and is passed on when the bacterium divides.

During infection by bacteriophage, the phage makes a nuclease that cuts the bacterial DNA into fragments (Figure 5.5). Single fragments of bacterial DNA comparable in size to phage DNA get packaged into a phage coat instead of phage DNA. These phage particles are called transducing phages. The tranducing phages are released and infect new cell. Then, the bacterial DNA contained in the phage head is injected into the cell and may then become integrated into the bacterial chromosome by a double crossover. The recombinant bacteria that are formed called transductants.

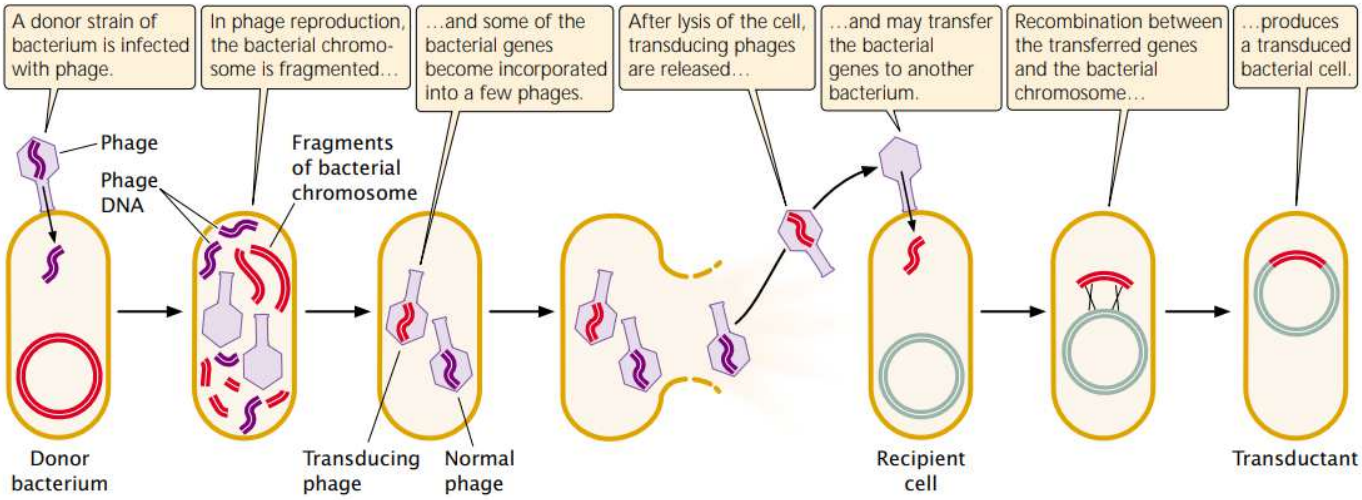
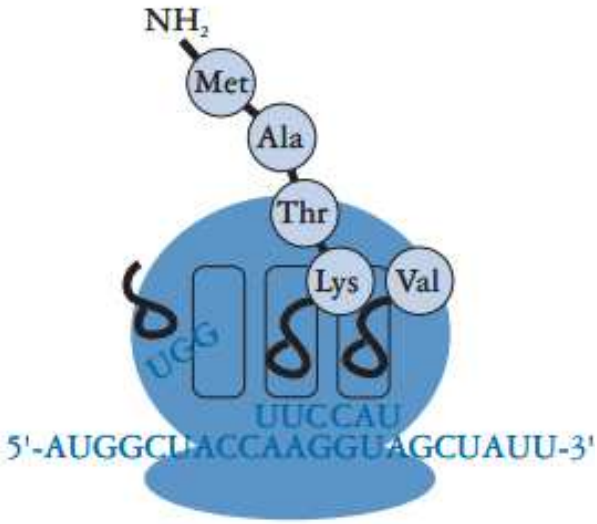


Figure 5.5. Transduction mechanism



CHAPTER 6

Protein Synthesis

ABSTRACT

This chapter provides information about the steps of protein synthesis

6.1. Introduction

Protein synthesis is a vital process consisting of two steps : transcription and translation.

6.2. Transcription

Transcription is the process in which a DNA sequence is transcribed into an RNA molecule with the help of enzyme RNA polymerase.

6.2.1. Elements required for transcription

This process takes place in the nucleus and requires some elements including :

- RNA polymerases: While prokaryotes have a single RNA polymerase, eukaryotic cells possess three RNA polymerases. RNA synthesis requires RNA polymerase I for ribosomal RNA, RNA polymerase II for mRNA and microRNA, and RNA polymerase III for tRNA and other small RNAs.
- DNA template strand
- Ribonucleotide triphosphates (ATP, UTP, CTP and GTP)
- Divalent metal ions Mg++ or Mn++ as a co-factor.

6.2.2. General characteristics of transcription

DNA is a double helical molecule. It contains two strands. One runs in 5' to 3' and the other is presented in a 3' to 5' direction. The template strand (transcribed strand) is considered as a non-coding strand which gets transcribed to RNA. It is also called antisense strand because its sequence is antiparallel to that of the newly synthetized mRNA. Non-template strand (not transcribed strand) is called a coding or sense strand because its sequence is the same as transcribed RNA, except for uracil.

- In transcription, only one of the DNA strands is transcribed, progressing in the 5′→3′ direction.
- The RNA polymerase synthesizes the RNA in the 5' to 3' direction.
- Unlike DNA polymerase, RNA polymerase can initiate RNA synthesis without a primer.
- Transcription is not as accurate as replication and produces more errors. The replication process is highly accurate, due to the proofreading ability of DNA polymerase.

6.2.3. Transcription steps

6.2.3.1. Initiation

RNA polymerase recognizes the initiation site called promoter, and binds to it directly, unwinding (unziping) of the DNA double helix.

The promoter region is known as a TATA box due to the presence of a high frequency of adenine and thymine. It signals the beginning of RNA synthesis. The promoter also determines the DNA strand that is to be transcribed.

6.2.3.2. Elongation

In this phase, RNA polymerase continues to separate the double helix and synthesize RNA in a 5' to 3' direction. The RNA polymerase places complementary bases to the template strand, except instead of placing thymine with every adenine, the enzyme places a uracil (U). This process is the elongation phase as the RNA polymerase continues down the template creating a new complementary single-stranded RNA.

6.2.3.3. Termination

When RNA polymerase arrives at a sequence called « terminator site », transcription stops, RNA polymerase dissociates and the DNA duplex reforms.

Transcription products are : mRNA (messenger RNA), tRNA (transfer RNA), rRNA (ribosomal RNA) and other small RNAs.

6.2.4. Post-transcriptional modification

RNA is transcribed, but must be processed into a mature form before translation can begin. This processing after an RNA molecule has been transcribed, but before it is translated into a protein, is called post-transcriptional modification.

In prokaryotes, pre-mRNA does not undergo modification, however, in eukaryotes pre-mRNA must undergo a number of modifications. These include :

- Addition of a G-cap (7-methylguanosine) to the 5′ end. The cap may protect the mRNA from degradation by nucleases and helps in recognition of ribosomes and thereby facilitates translation of mRNA.
- Addition of poly-A tail (100-200 adenine) to the 3′ end by poly-(A) polymerase. The poly(A) tail confers stability on many mRNAs, increasing the time during which the mRNA remains intact and available for translation before it is degraded by cellular enzymes.

- Splicing : The process of removing introns (non-coding regions) and reconnecting (rejoining) exons (coding regions) is called splicing.

6.3. Translation

Translation is the process of translating a genetic message, encoded in a messenger RNA, from one language (nucleic acid language- sequence of nucleotides) to another language (protein language- sequence of amino acids). The polypeptide in its turn determines the phenotype of the organism.

6.3.1. Translation steps

6.3.1. 1. Initiation

During initiation, the small ribosomal subunit recognizes the 5' G-cap and binds to the start of the mRNA sequence. Then a transfer RNA (tRNA) molecule carrying the amino acid methionine binds to what is called the start codon of the mRNA sequence. The start codon in all mRNA molecules has the sequence AUG and codes for methionine. Next, the large ribosomal subunit binds to form the complete initiation complex.

6.3.1.2. Elongation

The ribosome reads the mRNA codon by codon (from the 5' end of the mRNA to the 3' end). The codons in mRNA are recognized by tRNA molecules, which contain a 3-base sequence complementary to a codon and called an anticodon. Amino acids get attached to each tRNA by a specific enzyme called aminoacyl-tRNA synthetase. Then, molecules of tRNA bring amino acids to the ribosome in the correct sequence and each corresponding amino acid is added to the growing chain and linked via a bond called a peptide bond. Elongation continues until all of the codons are read.

The peptide chain always grows in a sequence from the free terminal amino ($–NH_2$) group towards the carboxyl (–COOH) end.

6.3.1.3. Termination

Termination occurs when the ribosome reaches a stop codon (UAA, UAG, and UGA). Since there are no tRNA molecules that can recognize these codons, the ribosome recognizes that translation is complete. The new protein is then released, and the translation complex (the ribosome dissociates from mRNA) comes apart (dissociates).

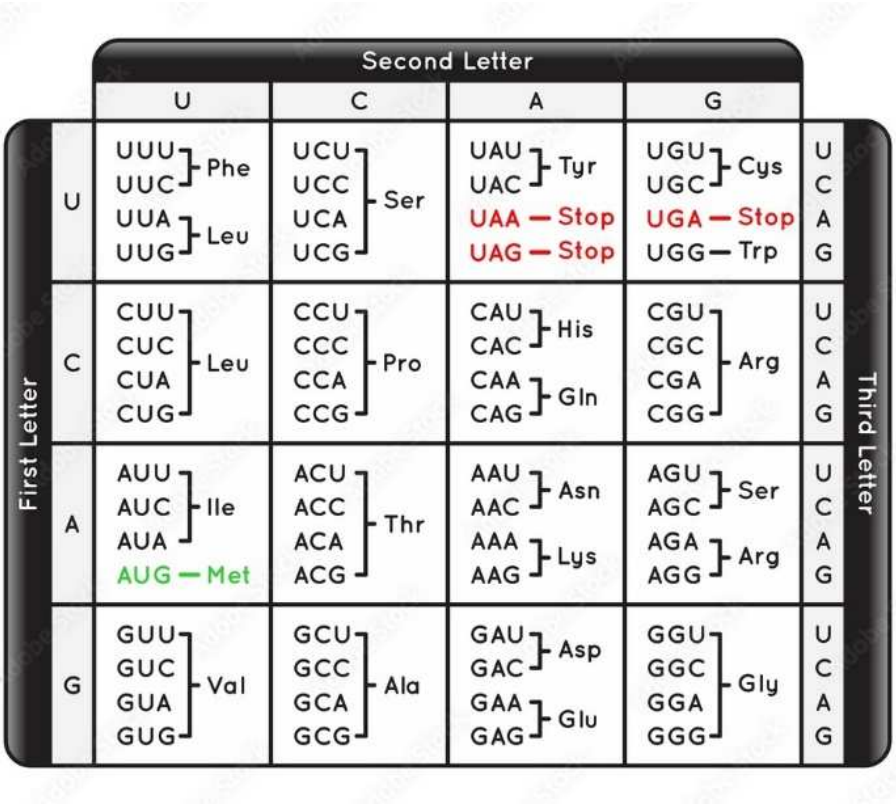
6.3.2. Genetic code

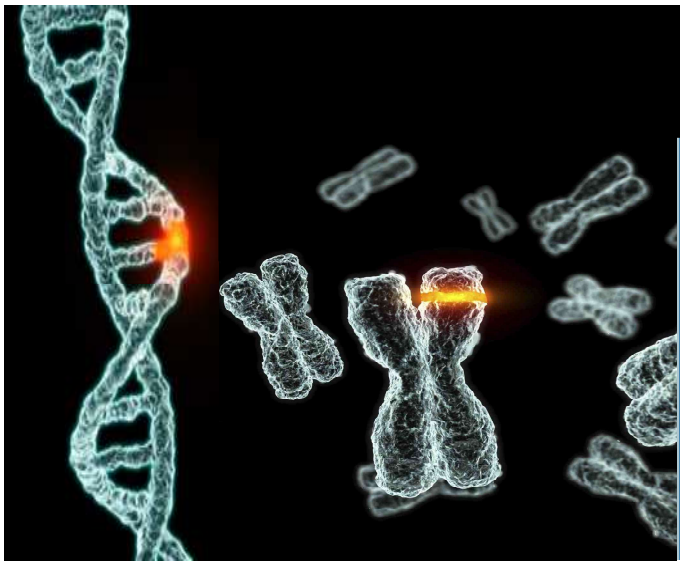
During translation, an mRNA sequence is read using the genetic code, which is a set of three-letter combinations of nucleotides called triplet codons, each of which corresponds with a specific amino acid or stop signal.

There are only 4 nucleotides in DNA to specify 20 amino acids. Thus : 4^3 =64 possible codes enough to code all amino acids. Among the 64 codons, 61 correspond to the 20 amino acids, of which 1 (AUG) is a start codon. The remaining 3 codons (UAA, UAG, and UGA) are stop codons (table 6.1).

The genetic code is highly redundant : many amino acids have several codons.

Table 6.1. Genetic code





CHAPTER 7

Mutations

ABSTRACT

This chapter provides fundamental information about gene mutations and chromosomal mutations.

7.1. Introduction

A mutation is a change in the genome of an organism.

There are two main types of mutations : gene mutations and chromosomal mutations.

7.2. Gene mutations

Gene mutations result from alterations in the nucleotide sequence in a gene.

7.2.1. Molecular basis of gene mutations

Gene mutations can result from:

a- Base substitutions: one pair of DNA nucleotides is replaced by another pair. There are two types:

- *Transition:* a purine is substituted for a purine or a pyrimidine is substituted for a pyrimidine.

  Example : **ACT TAT GCG**

  → **ACC TAT GCG**

- *Transversion:* a purine replaces a pyrimidine or vice versa.

  Example: **ACT TAT GCG**

  → **ACA TAT GCG**

b- Addition (insertion): one pair of DNA nucleotides is inserted.

Example: **ACT TAT GCG**

→ **ACG TTA TGC G**

c- Deletion : one pair of DNA nucleotides is removed.

Example : **ACT TAT GCG**

→ **ATT ATG CG**

7 .2.2. Types of gene mutations

There are two main types of gene mutations : point mutation and frameshift mutation.

7.2.2.1. Point mutation

Point mutation refers to a single-nucleotide change in a gene as a consequence of base substitution, insertion, or deletion.

Point mutations can be **silent, missense (conservative or non-conservative)** or **nonsense mutations**.

7.2.2.1.1. Silent mutation

Silent mutation is a point mutation that results in no change in the amino acid sequence of the protein encoded. The mutation has no effect.

Example :

5’ – AUG CAG UCG CAG – 3’
Met Gln Ala Glu

5’ – AUG CAA UCG CAG – 3’
Met Gln Ala Glu

7.2.2.1.2. Nonsense mutation

Nonsense mutation is a point mutation in which a sense codon (encoding for an amino acid) is changed to a nonsense codon (stop codon), resulting in premature chain termination. Usually it results in loss of the protein’s function.

Example :

5’ – AUG CAG UCG CAG – 3’
Met Gln Ala Glu

5’ – AUG CAG UCG UAG – 3’
Met Gln Ala Stop

7.2.2.1.3. Missense mutation

Missense mutation is a point mutation in which a single nucleotide change results in a codon that codes for a different amino acid.

Example :

5’ – AUG CAG UCG CAG – 3’
Met Gln Ala Glu

5’ – AUG CAG UGG CAG – 3’
Met Gln Trp Glu

Missense mutation can be **conservative** (new AA chemically similar to old AA) or **non-conservative** (new AA chemically different from old AA).

7.2.2.2. Frameshift mutation

Frameshift mutation is a genetic mutation caused by a deletion or insertion of one or two nucleotides in a DNA sequence that shifts the way the sequence is read.

Note : Deletion or insertion alters the triplet reading frame of a DNA sequence, resulting in premature termination of the protein synthesis.

Example 1 :

GUU – CAU – UUG – ACU – CCC – GAA – GAA
Val His Leu Thr Pro Glu Glu

Inserion of G

GUU – CAU – GUU – GAC – UCC – CGA – AGA A
Val His Val Ala Ser Arg Arg

Example 2 :

GUU – CAU – UUG – ACU – CCC – GAA – GAA
Val His Leu Thr Pro Glu Glu

Deletion of A

GUU – CAU – UUG – CUC – CCG – AAG – AA
Val His Leu Leu Pro Lys

7.3. Chromosomal mutations

Chromosomal mutations involve changes in the number or structure of chromosomes.

7.3.1. Changes in chromosome number

7.3.1.1. Aneuploidy

Aneuploidy is the presence of an abnormal number of chromosomes in a cell either in the autosomes or the sex chromosomes. Chromosome segregation errors (non-disjunction) during meiosis lead to aneuploid cells containing a number of chromosomes that differs from the exact haploid number (Figure 7.1).

The most common aneuploid conditions are : monosomy and trisomy.

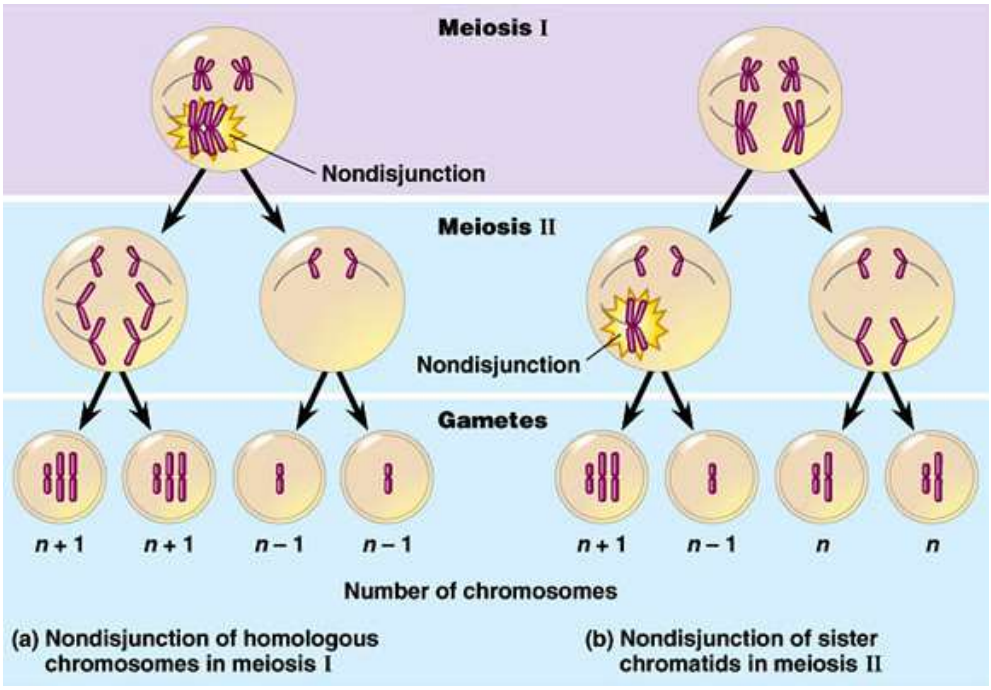


Figure 7.1. Nondisjunction phenomene

7.3.1.1.1. Monosomy

Monosomy is a form of aneuploidy with the presence of only one chromosome from a pair. There is one less chromosome than the normal (2n-1).

In humans, monosomy X causes Turner syndrome (45, X).

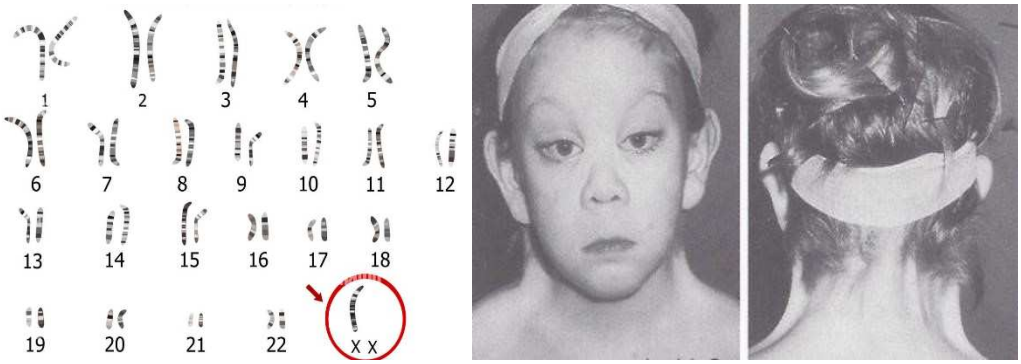


Figure 7.2. Turner syndrome

7.3.1.1.2. Trisomy

Trisomy is a form of aneuploidy with the presence of one extra chromosome (2n+1) than the normal.

The best known is trisomy 21 (47, XX or XY, + 21), also called Down syndrome (Figure 7.3). Individuals with Down syndrome are typically mentally retarded and have a characteristic

broad face with a flat nasal bridge, wide-set eyes, and prominent epicanthic folds. The hands are also broad. Individuals with Down syndrome are prone to the early appearance of Alzheimer's disease and typically have a shortened life span.

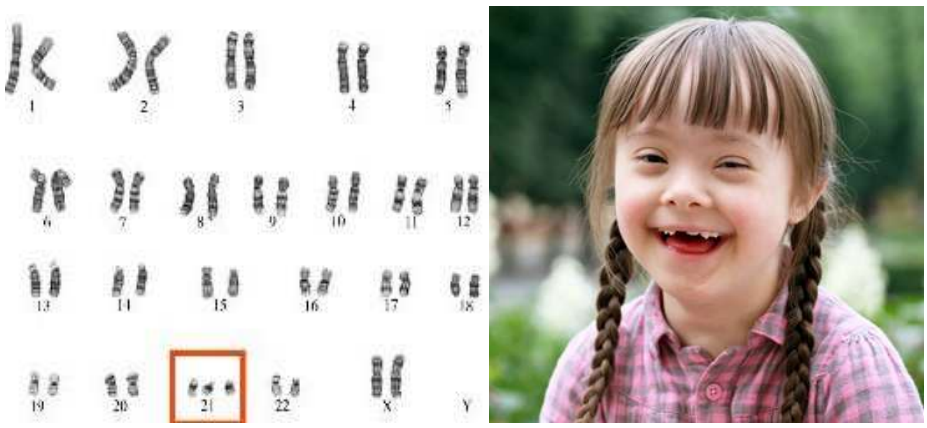


Figure 7.3. Down syndrome

Trisomies of chromosomes 13 and 18 known as Patau syndrome and Edwards syndrome, respectively, result in severely malformed fetuses, many of which do not survive to birth (Figures 7.4 and 7.5). Infants with trisomy 13 (47, XX or XY, + 13) and trisomy 18 (47, XX or XY, + 18) show severe mental retardation and other defects of the central nervous system. Cleft lip and cleft palate are common. Most infants born with trisomy 13 or trisomy 18 die within the first 1 or 2 months after birth.

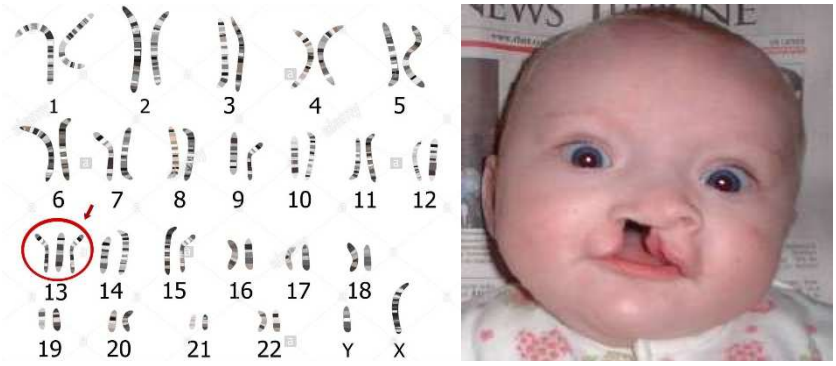


Figure 7.4. Patau syndrome

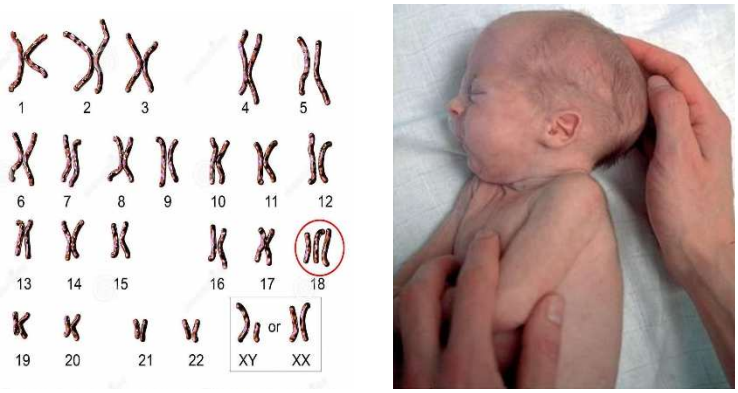


Figure 7.5. Edwards syndrome

7.3.1.2. Polyploidy

Polyploidy is the condition of a cell or organism possessing extra haploid sets of chromosomes. Although species with up to eight sets of chromosomes (octoploidy=8n) are known in plants, the most typical polyploidies observed in humans are triploidy (3n) and tetraploidy (4n) (Figure 7.6).

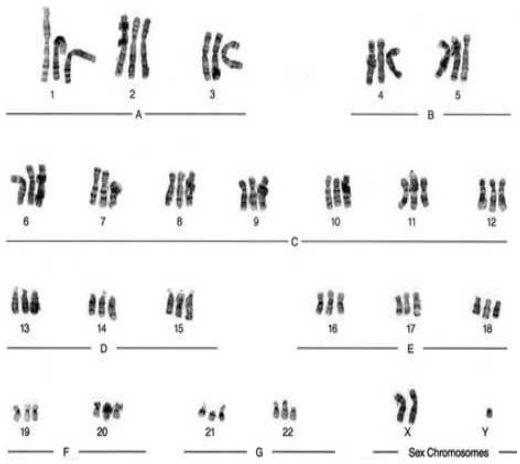


Figure 7.6. Human triploid karyotype

Polyploidy can arise in several different ways. If gametes are produced which are diploid and these self-fertilise, a tetraploid is produced. If the diploid gamete fuses with a normal haploid gamete, a triploid results. Polyploidy can also occur when whole sets of chromosomes double after fertilisation.

7.3.2. Changes in chromosome structure

Structural chromosomal aberrations refer to changes in chromosome structure. These mutations can change the location of genes on chromosomes and can even change the number of copies of some genes.

There are four types : Deletion, Duplication, Inversion and Translocation.

7.3.2.1. Insertion/Duplication

When part of a chromatid breaks off and attaches to its sister chromatid, an insertion occurs. Duplication is when an extra copy of a gene is made.

7.3.2.2. Deletion

When a part of a chromosome is left out, a deletion occurs.

7.3.2.3. Inversion

When part of a chromosome breaks off and reattaches backwards, an inversion occurs.

7.3.2.4. Translocation

When part of one chromosome breaks off and is added to a non-homologous chromosome, a translocation occurs.

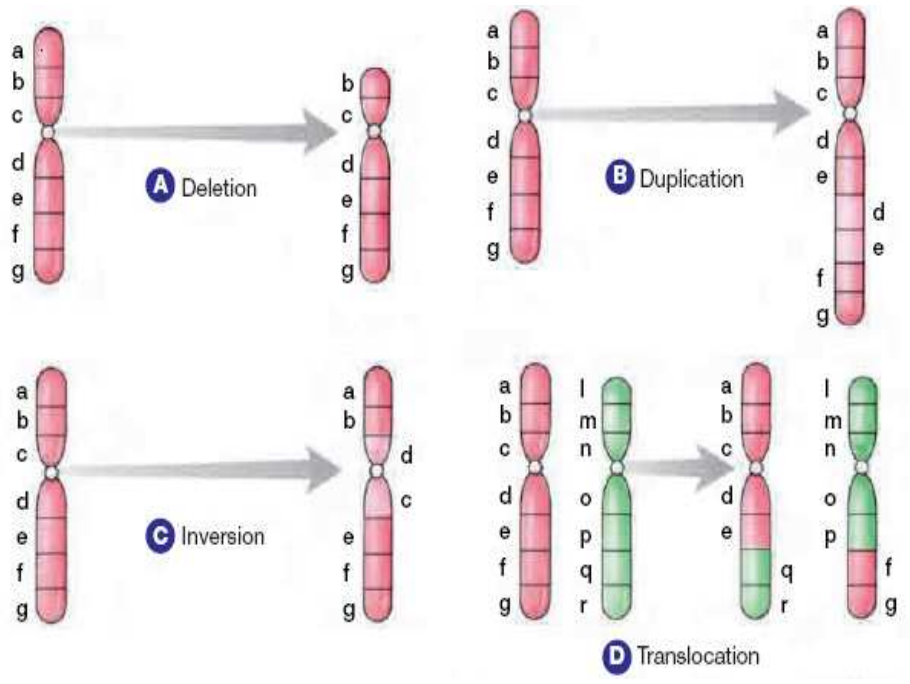


Figure 7. Structural chromosomal aberrations

7.4. Causes of mutations

Mutations can be caused by errors in transcription, cell division, or by external agents. Any agent that can cause a change in DNA is called a **mutagen**.

Mutagens inlcude chemical or physical agents in the environment.

- *Chemical mutagens* include certain pesticides, a few natural plant alkaloids, tobacco smoke, and environmental pollutants.
- *Physical mutagens* include some forms of electromagnetic radiation (such as X-rays, ultraviolet light) and even high temperature.

7.5. Effects of mutations

7.5.1. Harmful effects

Most mutations, if expressed, are harmful. However, in diploid organisms such as us, mutations usually result in recessive alleles. These are expressed only in the homozygous condition unless the mutation is on the X chromosome.

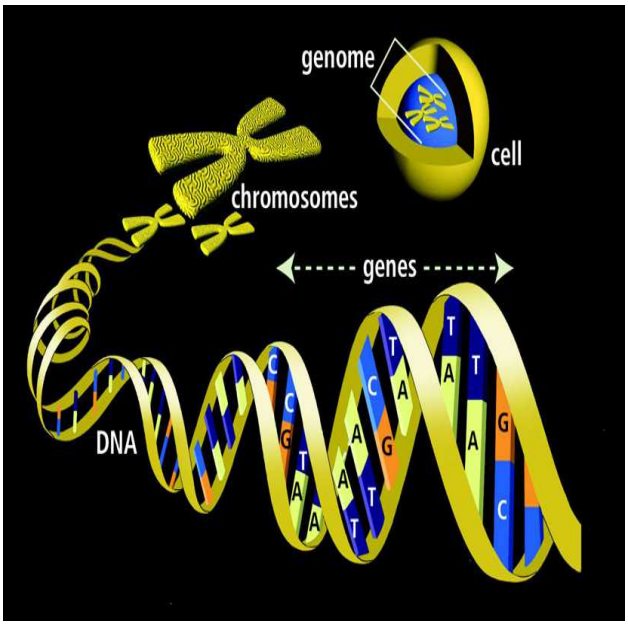
The defective proteins produced by these mutations can disrupt normal biological activities, and result in genetic disorders.

Some cancers, for example, are the product of mutations that cause the uncontrolled growth of cells.

Sickle cell disease is a disorder associated with changes in the shape of red blood cells. Normal red blood cells are round. Sickle cells appear long and pointed. This disease is caused by a point mutation in one of the polypeptides found in hemoglobin, the blood's principal oxygen-carrying protein.

7.5.2. Beneficial effects

Mutations can be beneficial. For example, mutations have helped many insects resist chemical pesticides. Some mutations have enabled microorganisms to adapt to new chemicals in the environment.



CHAPTER 8

Gene Structure And Function

ABSTRACT

This chapter provides information about the structure and funtion of genes in prokaryotes and eukaryotes.

8.1. Gene structure

The gene is the fundamental unit of hereditary. A gene determines a particular trait. At the molecular level, a gene is a particular segment of DNA molecule that codes for a cellular product usually polypeptide. In the cells, DNA contain the information needed to make all of the proteins that the cells use. Eye color, hair color, such traits are visible because of protein synthesis. DNA does not control protein synthesis directly. It uses RNA. The flow of genetic information from DNA to RNA to protein is called *central dogma*.

8.1.1. Gene structure in prokaryotes

In prokaryotic cells such as bacteria, genes are usually found grouped together in operons. The operon is a cluster of genes that are related and that are under the control of a single promoter/regulatory region.

The fact that structural genes in prokaryotes are often grouped together means that the transcribed mRNA may contain information for more than one protein. Such a molecule is known as a polycistronic mRNA, with the term cistron equating to the 'gene' as we have defined it (i.e. encoding one protein). Thus, much of the genetic information in bacteria is expressed via polycistronic mRNAs whose synthesis is regulated in accordance with the needs of the cell at any given time. This system is flexible and efficient, and it enables the cell to adapt quickly to changing environmental conditions.

8.1.1.1. Operon structure

Operon consists of 3-parts : promoter, operator and structural genes (Figure 9.1).

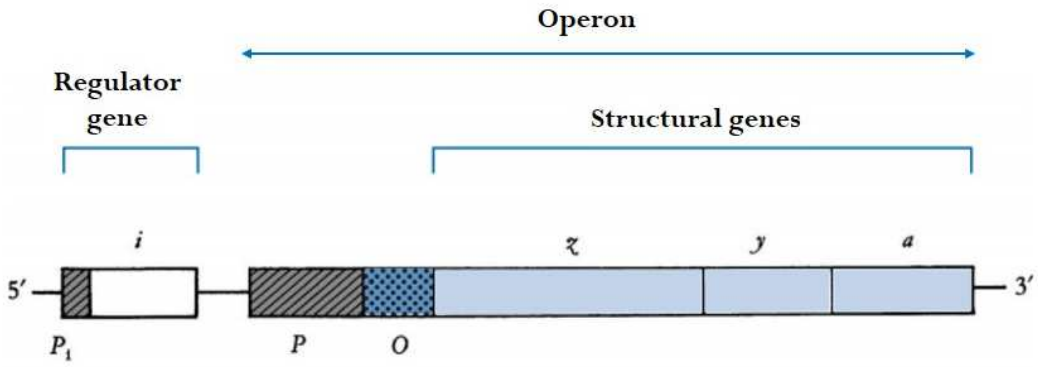


Figure 8.1. Operon structure

- **Regulator gene :** Regulator gene is out side of operon and is not considered part of the operon, although it affects operon function. It is a gene that codes for a regulator protein, (repressor) which controls the expression of another gene. This gene codes for the production of a protein that can bind to the operator. When this repressor binds to an operator, it prevents RNA polymerase from binding to the promoter region of the Operon. So repressor prevents the expression of structural genes.
- **Repressor :** The regulator protein coded by regulatory gene. Repressor can bind to the operator.
- **Co-repressor** : A small molecule that cooperates with a repressor to switch an operon off.
- **Promoter** : The part of the operon where RNA polymerase binds to begin transcription of the structural genes. The structural gene and operon all share a single promoter site when RNA polymerase binds to this site, it transcripts all of the structural genes onto one mRNA strand, which may go on to be translated into the individual proteins (enzymes).
- **Operator :** The part of DNA that controls expression of structural genes. It acts as a control switch on or off by allowing or preventing of RNA polymerase enzyme on the promoter.
- **Structural genes :** The part of DNA that codes synthesis of a group of enzyme that are involved in the same function.(for example the digestion of lactose).

8.1.1.2. Types of operons

There are two main types of operons : **inducible and repressible**.

8.1.1.2.1. Inducible operon

The operon is said to be inducible, because transcription is normally off (inhibited) and must be turned on (induced). Transcription is turned off because the repressor is bound to the operator. However, the transcription is turned on when a small molecule, an inducer, binds to the repressor, so, the repressor can no longer bind to the operator and therefore, RNA polymerase readily binds to the promoter and transcription of the structural genes takes place.

This system is often found in the control of the synthesis of enzymes that participate in degradative (catabolic) pathways. Example: Lactose operon.

8.1.1.2.2. Repressible operon

The operon is said to be repressible because the transcription normally takes place and must be turned off, or repressed.

The regulator protein in this type of operon also is a repressor but is synthesized in an inactive form that cannot by itself bind to the operator. Because there is no repressor bound to the operator, RNA polymerase readily binds to the promoter and transcription of the structural genes takes place.

To turn transcription off, a small molecule called a co-repressor binds to the repressor forms the repressor-co repressor complex. This complex binds to operator and blocks the protein synthesis.

This type is often found in the control of the synthesis of enzymes that participate in biosynthetic (anabolic) pathways. Example : Tryptophan operon

8.1.2. Gene structure in eukaryotes

Eukaryotic genes are not organized into operons and are rarely transcribed together into a single mRNA molecule; instead, each structural gene typically has its own promoter and is transcribed separately.

Many eukaryotic genes contain coding regions called *exons* and noncoding regions called intervening sequences or *introns*. For example, the ovalbumin gene has eight exons and seven introns; the gene for cytochrome *b* has five exons and four introns. All the introns and the exons are initially transcribed into RNA but, after transcription, the introns are removed by splicing and the exons are joined to yield the mature RNA (Figure 8.2).

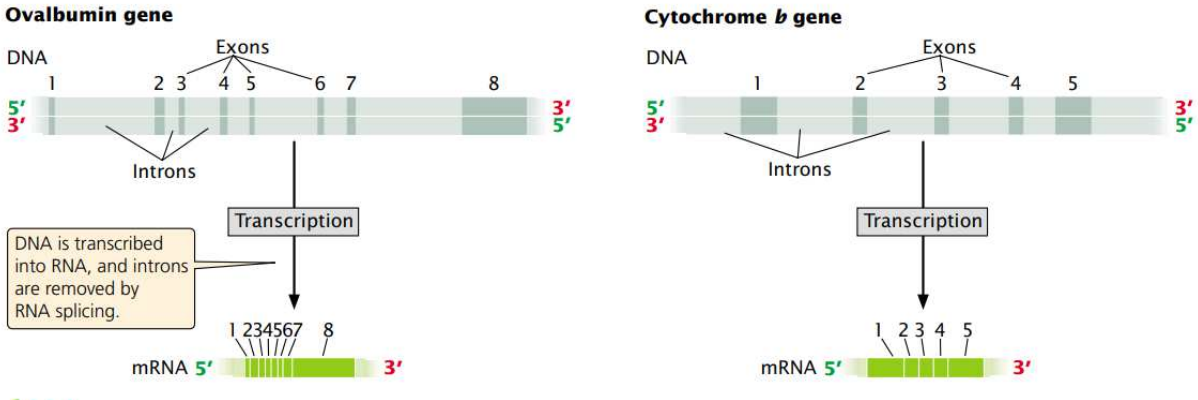


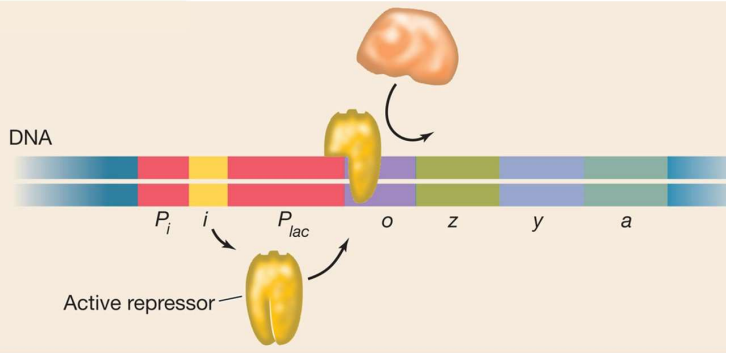
Figure 8.2. Ovalbumin and cytochrome *b* genes

8.2. Gene function

A gene is a contiguous region of DNA that encodes instructions for how the cell can make a macromolecule (or potentially multiple different macromolecules). A macromolecule is called a gene product, and can be of two types, a protein (the most common type) or a non-coding RNA. A gene product can act as a molecular machine, which can perform a chemical action that we call an activity. Gene products from different genes can combine into a larger molecular machine, called a macromolecular complex.

Most genes and gene products have multiple functions. Furthermore, even if they interact with only one other molecule or catalyse just one reaction, they often act in a variety of different circumstances, at a variety of different times and places, and under a variety of different conditions.

Certain aspects of gene function can be determined by the analysis of gene and protein expression. For example, if a gene is expressed specifically in response to growth factor signaling, this means that the gene's function is linked to the control of cell proliferation. Genes that are expressed in a disease state but not in healthy tissue are likely in some way reflecting the pathogenesis of that disease or the body's response to it. For example, over expression (high serum alpha-fetoprotein levels) of alpha-fetoprotein can be indicative of certain cancers and liver diseases.



CHAPTER 9

Regulation of Gene Expression

ABSTRACT

This chapter provides information about the mechanisms that are used to control gene expression in prokaryotes and eukaryotes.

9.1. Introduction

Gene expression is the process by which the information encoded in a gene is turned into a function. Transcription and translation provide the mechanisms by which genes are expressed. However, it is vital that gene expression is controlled so that the correct gene products are produced in the cell the right time.

Bacteria carry the genetic information for many proteins, but only a subset of this genetic information is expressed at any time. When the environment changes, new genes are expressed, and proteins appropriate for the new environment are synthesized.

In addition, all of human's cells carry the same genetic information, but only a subset of genes are expressed in each cell type. Genes needed for other cell types are not expressed. For instance, a nerve cell and a blood cell produce different proteins although they carry the same genetic instructions.

Thus, gene expression is carefully regulated in both prokaryotic and eukaryotic cells. Several strategies are used to regulate gene expression.

9.2. Gene regulation in prokaryotes

Prokaryotic cells such as *Escherichia coli* (*E. coli*) can regulate their gene expression to adapt to the changes in available nutrients in the surrounding environment. Their regulation of gene expression occurs mainly through regulation of the rate of mRNA synthesis, since their transcription and translation is coupled.

9.2.1. Regulation of inducible operon (Lac operon)

9.2.1.1. Lac operon structure

The lactose operon (also known as the lac operon) is a set of genes that are specific for uptake and metabolism of lactose and is found in *E. coli* and other bacteria.

Lactose (a disaccharide) is one of the major carbohydrates found in milk; it can be metabolized by *E. coli* bacteria that reside in the gut of mammals. Lactose does not easily diffuse across the E. coli cell membrane and must be actively transported into the cell by the enzyme permease. To utilize lactose as an energy source, *E. coli* must first break it into glucose and galactose, a reaction catalyzed by the enzyme β-galactosidase. This enzyme can also convert lactose into allolactose, a compound that plays an important role in regulating lactose metabolism. A third enzyme, thiogalactoside transacetylase, also is produced by the lac operon,

which transfers an acetyl group from acetyl-CoA to β- galactoside (i.e., this enzyme is indirectly involved in lactose utilization).

A lac operon contains promotor, operator and three structural genes or cistrons, namely z, y and a, whose products (=enzymes) are involved in the breakdown (catabolism) of the sugar lactose (Figure 9.1). Gene z codes for an enzyme, β- galactosidase, which converts lactose into glucose and galactose, while gene y determines the structure of an enzyme, galactoside permease which is a plasma-membrane bound protein and facilitates the entrance of lactose into the cell. Gene a specifies an enzyme, thiogalactoside transacetylase, which transfers an acetyl group from acetyl-CoA to β-galactoside (i.e., this enzyme is indirectly involved in lactose utilization).

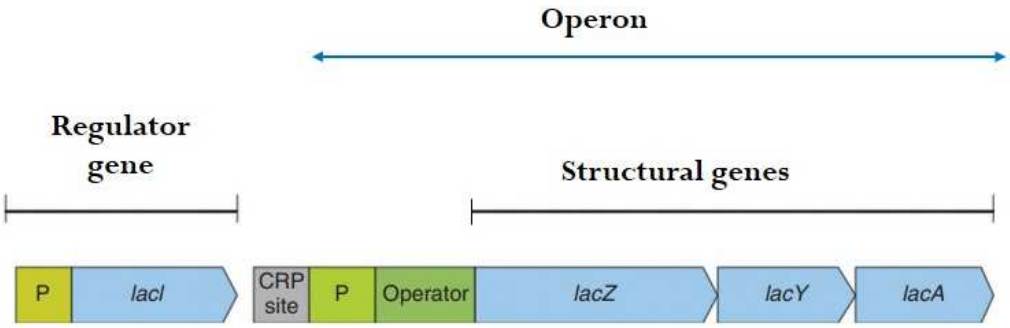


Figure 9.1. Lac operon structure

In addition, in front of the lac operon there are binding sites for proteins (CBS : CAP-binding site) involved in transcriptional regulation of the operon.

9.2.1.2. Negative control of lac operon

Regulation of the lac operon by repressor is called negative control. Expression of this operon is activated only when lactose levels outside the cell are high and glucose levels are low. E. coli utilizes preferentially glucose and, as a result, will not activate the genes to metabolize lactose until there is a sufficiently high level of external lactose.

When the culture medium contains low glucose and high lactose concentrations, the permease protein transports external lactose into the cell. The internal lactose is converted to allolactose, which then binds to the repressor causing a conformational change. As a result, the inactive repressor cannot bind to the operator, allowing RNA polymerase to bind at the promoter and transcription to begin.

9.2.1.3. Positive control of lac operon

The lac operon is also under positive control. This mode of gene regulation is conferred by a cAMP-CRP complex that enhances (stimulates) the attachment of RNA polymerase to the promoters and initiation of mRNA synthesis.

Glucose levels help regulate the cAMP (cyclic AMP) concentration in the cell, and cAMP regulates the activity of the lac operon. However, both CRP (cyclic AMP receptor protein) function and cAMP are required for lac mRNA synthesis. CRP and cAMP bind to one another, forming a complex denoted cAMP-CRP, which is an active regulatory element in the lac system.

The concentration of cAMP is inversely proportional to the abundance of glucose : when glucose concentrations are low, an enzyme called adenylate cyclase is able to produce cAMP from ATP. Evidently, E. coli prefers glucose over lactose, and so expresses the lac operon at high levels only when glucose is absent and lactose is present. Only in the presence of lactose, and in the absence of glucose is the operon expressed at its highest levels.

When glucose [Glc] and lactose [Lac] are both high, the lac operon is transcribed at a moderate level, because CAP (in the absence of cAMP) is unable to bind to its corresponding cis-element (yellow) and therefore cannot help to stabilize binding of RNA pol at the promoter (Figure 9.3). Alternatively, when [Glc] is low, and [Lac] is high, CAP and cAMP can bind near the promoter and increase further the transcription of the lac operon (Figure 9.2).

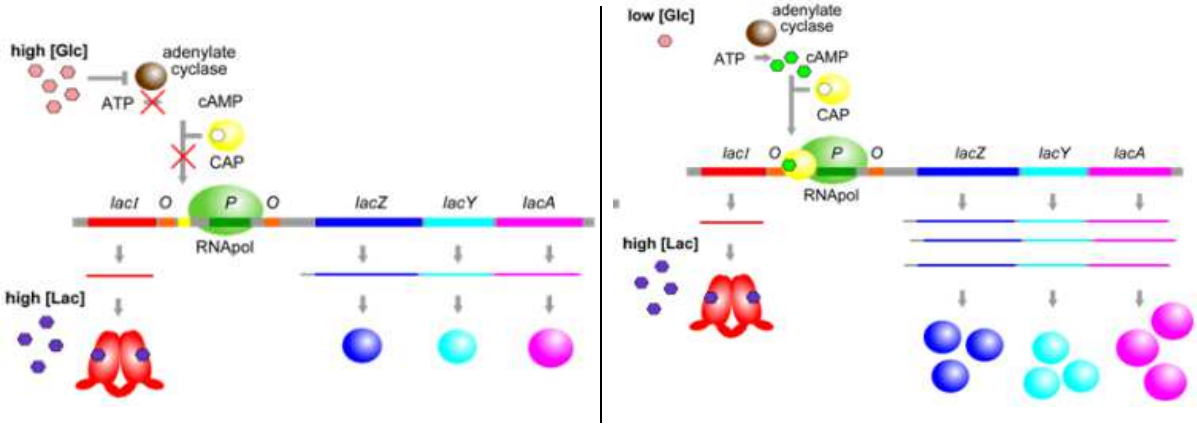


Figure 9.2. Positive control of lac operon

9.2.2. Regulation of repressible operon (Trp operon)

9.2.2.1. Structure of Trp operon

The tryptophan operon of *E. coli* is a negative repressible operon that controls the biosynthesis of tryptophan.

The trp operon includes 5-structural genes (trpE, trpD, trpC, trpB, and trpA), which produce components of three enzymes required to convert chorismate into tryptophan. Anthranilate synthetase with component I is encoded by trpE, component II is encoded by trpD; N-(5′-phosphoribosyl)- anthranilate isomerase/indole-3-glycerol phosphate synthase is encoded by trpC); tryptophan synthase with the α-subunit is encoded by trpA) and the β-subunit is encoded by trpB (Figure 9.3).

Adjacent to the TrpE gene are the promoter, the operator and two regions called the leader and the attenuator, which are designated TrpL and Trpa. Some distance from the trp operon is a regulator gene, trpR, which encodes a repressor helps regulate expression of these genes.

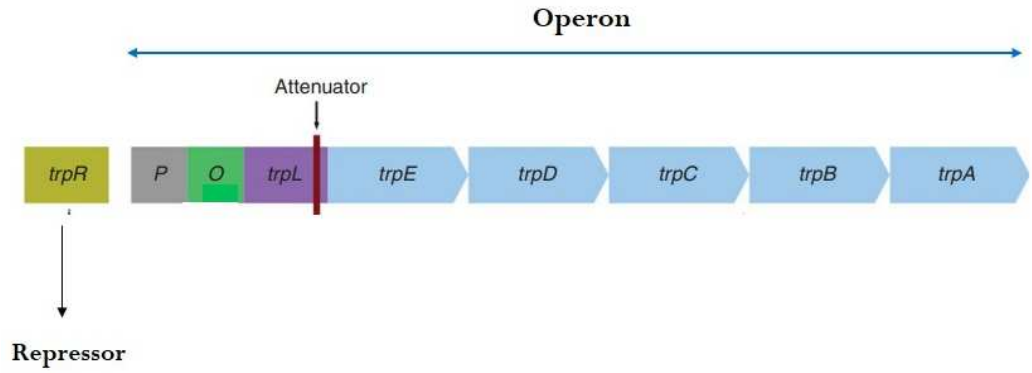


Figure 9.3. Trp operon structure

Regulation of the trp operon is determined by the concentration of tryptophan: when adequate tryptophan is present in the growth medium, transcription of the operon is repressed (turned off) and when tryptophan is absent, transcription is turned on.

Expression of trp operon is regulated by a repressor and by attenuation.

9.2.2.2. Regulation Trp operon by repressor

Regulation of the trp operon is determined by the concentration of tryptophan: when adequate tryptophan is present in the growth medium, transcription of the operon is repressed (turned off) and when tryptophan is absent, transcription is turned on.

In a repressible operon, transcription is normally turned on and must be repressed.

The repressor protein is synthesized in an inactive form that can not bind to the operator. It has two binding sites, one that binds to DNA at the operator site and another that binds to tryptophan. Binding with tryptophan causes a conformational change in the repressor that (activates it) makes it capable of binding to DNA at the operator site, which overlaps the promoter. When the operator is occupied by the tryptophan repressor, RNA polymerase cannot bind to the promoter and thus the structural genes cannot be transcribed.

9.2.2.3. Regulation of Trp operon by attenuation

Attenuation is a second mechanism of negative feedback (control) in the trp operon. While the TrpR repressor decreases transcription by a factor of 70, attenuation can further decrease it by a factor of 10, then the combined action of both mechanisms result in 700-fold reduction in expression. These mechanisms ensure that expression of the trp operon is turned off when tryptophan levels are high. Attenuation is possible because in prokaryotes the ribosomes begin translating the mRNA while RNA polymerase is still transcribing the DNA sequence. This allows the process of translation to directly affect transcription of the operon.

In attenuation, transcription is initiated but terminates prematurely (termination of transcription before the first structural gene of the operon is transcribed.

9.2.4. Comparison between Lac and Trp operon

Regulation in the trp operon is similar to that of the lac operon because mRNA synthesis is regulated negatively by a repressor. However, it differs from regulation of lac in that tryptophan acts as a co-repressor, which stimulates binding of the repressor to the trp operator to shut off synthesis.

The trp operon is a repressible rather than an inducible operon, although both the lac and the trp operons are negatively regulated.

Furthermore, because the trp operon codes for a set of biosynthetic enzymes rather than degradative enzymes, neither glucose nor cAMP-CRP functions in regulation of the trp operon.

9.3. Gene Regulation in Eukaryotes

Eukaryotic regulation of gene expression is much more complex than in prokaryotic gene expression and involves a variety of factors and elements.

9.3.1. Mechanisms of gene regulation in eukaryotes

Eukaryotic gene expression can be regulated at five major levels: chromatin, transcription, post-transcription, translation and post-translation.

9.3.1.1. Chromatin level

9.3.1.1.1. Histone acetylation

One factor affecting chromatin structure is acetylation, the addition of acetyl groups (CH_3CO) to histone proteins. Acetyl groups are added to histone proteins by acteyltransferase enzymes; the acetyl groups destabilize the nucleosome structure, perhaps by neutralizing the positive charges on the histone tails and allowing the DNA to separate from the histone (Figure 9.4).

9.3.1.1.2. DNA methylation

Another change in chromatin structure is the methylation of cytosine bases, which yields 5-methylcytosine. Heavily methylated DNA is associated with the repression of transcription in vertebrates and plants (Figure 9.4).

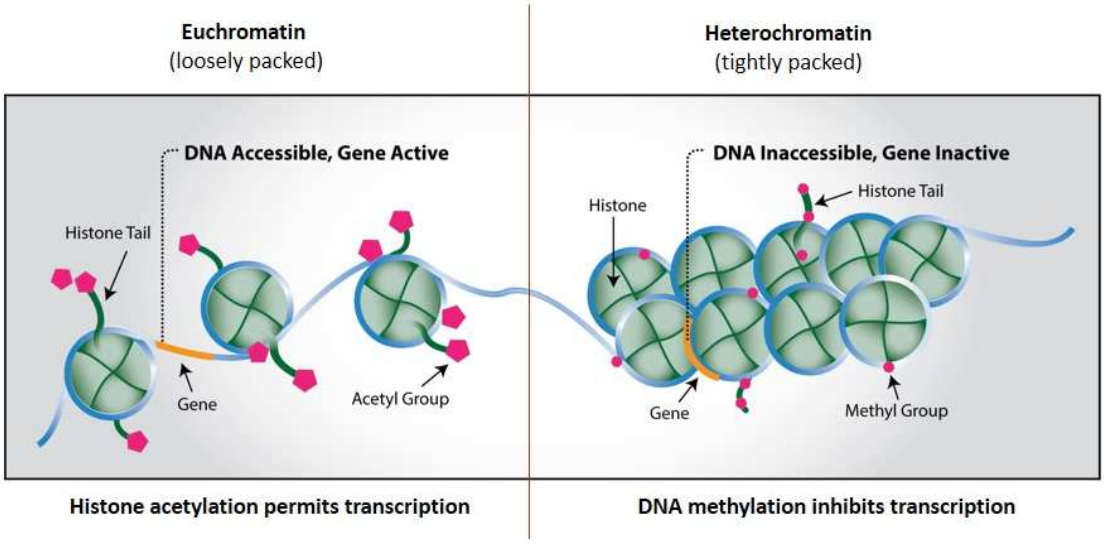


Figure 9.4. Histone acetylation and DNA methylation

9.3.1.2. Transcriptional level

Eukaryotic transcription is regulated by a large number of proteins. This process requires the interaction of numerous proteins with DNA in the region of the promoter. However, before the transcription can be initiated, the transcription factor IID (TFIID) must bind to the promoter. This is followed by the binding of other proteins called general transcription factors (GTFs) to DNA in the region of the promoter. The various GTFs facilitate binding of RNA polymerase II

to the promoter at the correct nucleotide for initiation, destabilize the DNA at the promoter, and initiate transcription.

9.3.1.3. Post-transcriptional

9.3.1.3. 1. The stability of the mRNA in the cytoplasm

Another type of post-transcriptional control involves the stability of the mRNA in the cytoplasm. Many factors contribute to mRNA stability, including the length of its poly-A tail.

9.3.1.3. 2. RNA editing

RNA editing involves processing of RNA in the nucleus by enzymes that change a single nucleotide (insertion, deletion, or substitution).

9.3.1.3. 3. Alternative RNA splicing

Alternative RNA splicing allows a pre-mRNA to be spliced in multiple ways, generating different proteins in different tissues or at different times in development (Figure 9.5).

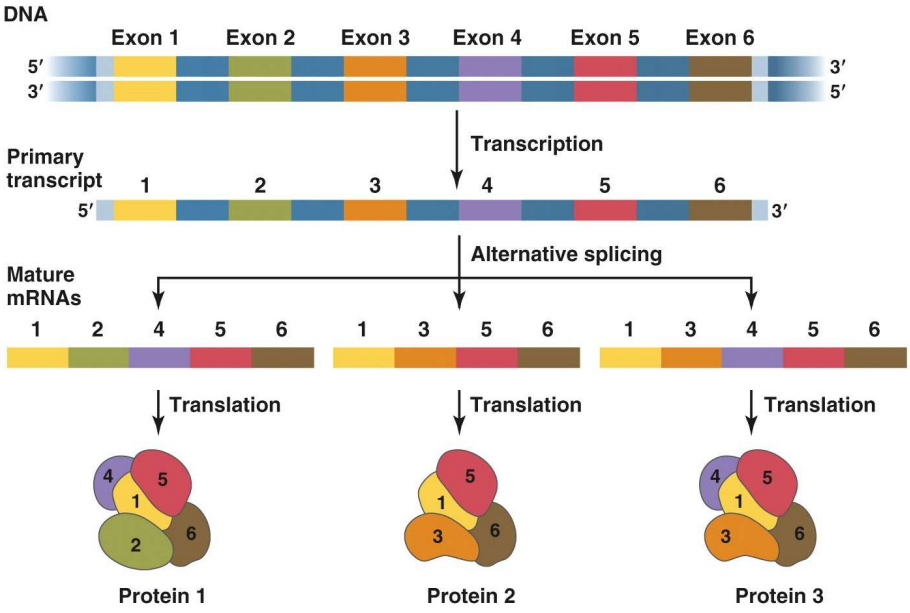


Figure 9.5. Alternative RNA splicing

9.3.1.3.4. RNA Transport from the Nucleus to the Cytoplasm

The mature mRNA molecules have to be translocated into the cytoplasm, where protein synthesis occurs.

During the pre-mRNA processing phase, many nuclear proteins are recruited to the mRNA being processed to form RNA-protein complexes, and only a successfully processed mRNA will form a proper mRNA-protein complex, which is required to pass through the

nuclear pore complexes. Therefore, defective mRNA and other byproduct RNA molecules, such as spliced-out introns, will be subjected to degradation by a multi-subunit exosome complex specialized for RNA degradation.

9.3.1.3.5. RNA Silencing

The expression of some genes may be suppressed through RNA silencing, also known as RNA interference and post-transcriptional gene silencing.

Two classes of small noncoding RNA molecules, microRNA (miRNA) and small interfering RNA (siRNA) are involved in RNA silencing.

- **miRNAs** are short single-stranded RNA molecules of up to 25 nucleotides that regulate the expression of other genes by binding to mRNAs (particularly the 3′ untranslated region, or 3′ UTR, in humans) and causing the degradation of the latter or blocking their translation into proteins.
- **siRNAs** are double-stranded RNA (dsRNA) molecules, up to 21 nucleotides in length. siRNAs combine with proteins to form an RNA-induced silencing complex (RISC). In a process called RNA interference, or RNAi, the RISC pairs with complementary sequences on mRNA and either cleaves the mRNA or prevents it from being translated.

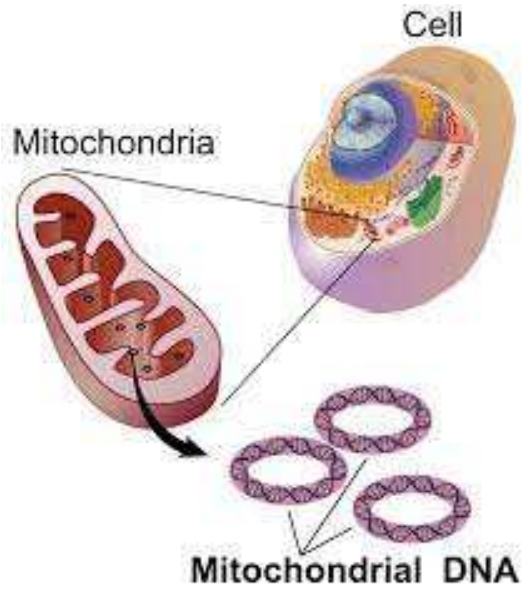
9.3.1.4. Translational level

The regulation of the translation of mRNA in eukaryotic cells is critical for gene expression, and occurs principally at the initiation phase which is mainly regulated by eukaryotic initiation factors (eIFs).

9.3.1.5. Post-translational control

The final level of control of gene expression in eukaryotes is post-translational regulation. This type of control involves modifying the protein after it is made, in such as way as to affect its activity.

The activity and/or stability of proteins can also be regulated by adding functional groups, such as methyl, phosphate, or acetyl groups. Sometimes these modifications can regulate where a protein is found in the cell, for example, in the nucleus, the cytoplasm, or attached to the plasma membrane.



CHAPTER 10

Extrachromosomal Genetics

ABSTRACT

This chapter provides information about mitochondrial and chloroplast DNAs, their replications, transctiprtion, translation and mechanism of inherence.

10.1. Introduction

The genetic material in eukaryotic cells is enclosed in a nucleus. However, additional genetic material is found in some of the organelles (mitochondria, chloroplast) of eukaryotic cells and in the cytoplasm of prokaryotic cells (plasmides).

10.2. Mitochondrion and Chloroplast Structure

Mitochondria are from 0.5 to 1.0 µm in diameter (Figure 10.1.). They are surrounded by two membranes that enclose a region called the matrix that contains enzymes, ribosomes, RNA, and DNA. The inner membrane is highly folded, forming cristae (folds). Mitochondria are present in almost all eukaryotic cells. However, the number of mitochondria present in a cell depends upon the metabolic requirements of that cell, and may range from a single large mitochondrion to thousands of the organelles. Mitochondria are responsible for cell energy production (synthesis of ATP), oxidative phosphorylation cell signaling (sending cell messages), cell death or apoptosis and free radicals production.

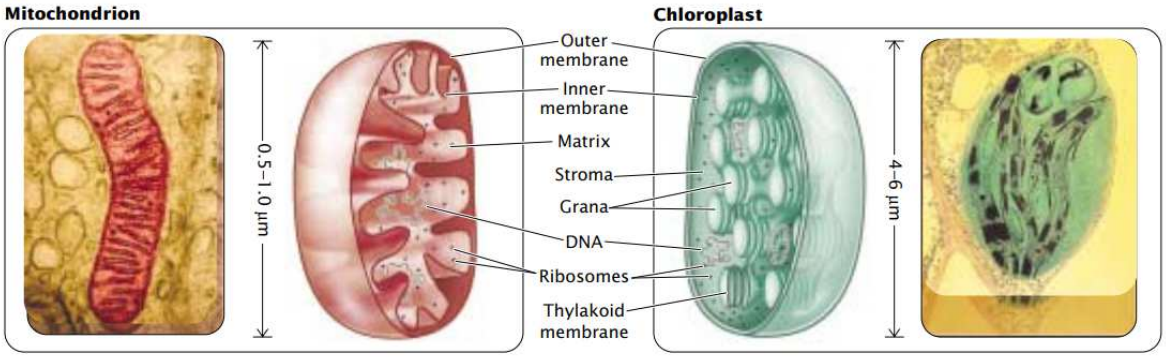


Figure 10.1. Structures of mitochondria and chloroplasts

Chloroplasts are specialized organelles found in photosynthetic organisms. They are typically from about 4 to 6 µm in diameter (Figure 10.1). They are surrounded by two membranes that enclose a region called the stroma that contains enzymes, ribosomes, RNA, and DNA. Chloroplasts have a third membrane, called the thylakoid membrane, which is highly folded and stacked to form aggregates called grana. This membrane bears the pigments and enzymes required for photophosphorylation.

10.3. Origin of mitochondrion and chloroplast

Both mitochondrial and chloroplast DNA are believed to have originated from bacteria, through endosymbiosis. Mitochondria are thought to have evolved from a proteobacterium, while

chloroplasts originated from a cyanobacterium. This theory is supported by their structural similarities to bacteria, and the presence of genes that are similar to those found in bacteria.

10.4. Multiplication of organelles

New mitochondria and chloroplasts arise by the division of existing organelles. They propagate through binary fission from pre-existing mitochondria and chloroplasts (Figure 10.2).

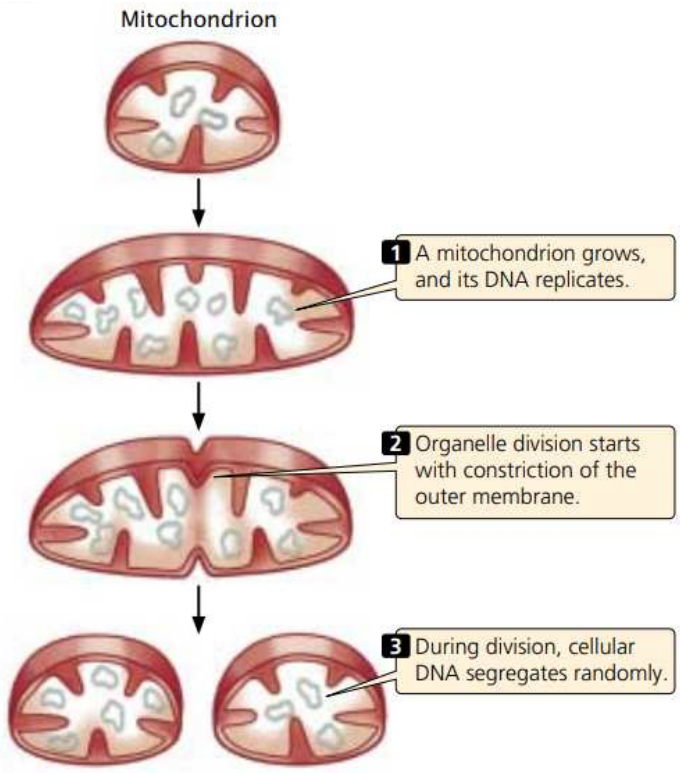


Figure 10.2. Multiplication of mitochondria

10.5. Organellar Genomes

10.5.1. Mitochondrial DNA

Most mtDNA molecules are circular, but in some species, such as the alga *Chlamydomonas reinhardtii* and the ciliate *Paramecium aurelia*, they are linear.

10.5.1.1. Human mtDNA

Human mtDNA is a circular double-stranded molecule encompassing 16,569 bp that encode two rRNAs, 22 tRNAs, and 13 proteins. mtDNA lacks histone proteins (Figure 10.3). The two nucleotide strands of the molecule differ in their base composition : the heavy (H) strand has more guanine nucleotides, whereas the light (L) strand has more cytosine nucleotides. The H strand is the template for both rRNAs, 14 of the 22 tRNAs, and 12 of the 13 proteins, whereas

the L strand serves as template for only 8 of the tRNAs and one protein. The origin of replication for the H strand is within a region known as the D loop, which also contains promoters for both the H and L strands.

Human mtDNA is highly economical in its organization: there are few noncoding regions (NCR) between the genes; almost all the mRNA is translated (there are no 5' and 3' untranslated regions); and there are no introns. Each strand has only a single promoter; so transcription produces two very large RNA precursors that are later cleaved into individual RNA molecules. Many of the genes that encode polypeptides even lack a complete termination codon, ending in either U or UA; the addition of a poly(A) tail to the 3' end of the mRNA provides a UAA termination codon that halts translation.

Human mtDNA also contains very little repetitive DNA (repetitive regions). The one region of the human mtDNA that does contain some noncoding regions (NCR) is the D loop.

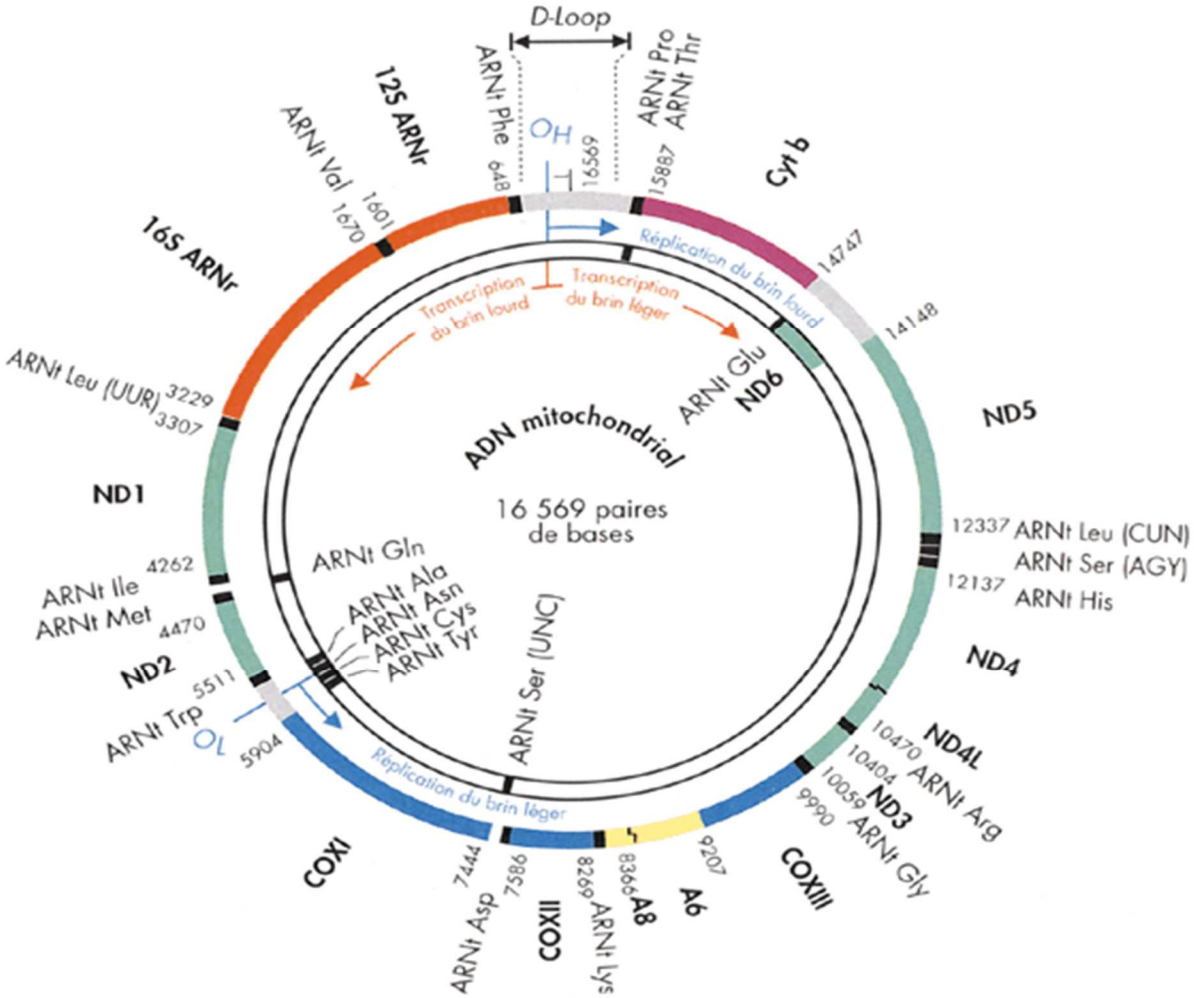


Figure 10.3. Human mtDNA

10.5.2. Chloroplast DNA

The chloroplast genome (cpDNA) is a double-stranded DNA molecule, highly coiled and lacks of associated histone proteins. cpDNAs typically range from 120 to 160 kb in size, but it could be much larger, according to species (chloroplast genome size varies between species).

Chloroplast genome is circular but can also be branched and linear. The number of cpDNA present in a cell depends the number of chloroplasts and the number of cpDNA molecules within each chloroplast. However, the cpDNA number is highly variable during the plant development.

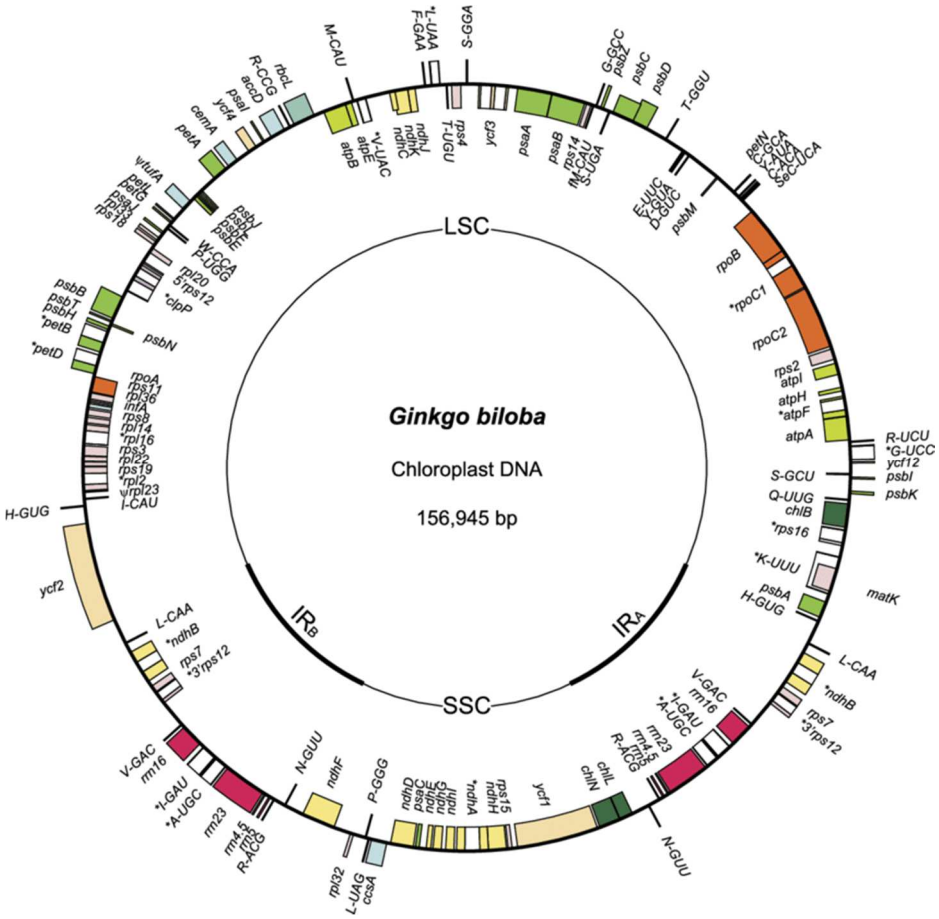


Figure 10. cpDNA map of *Ginkgo biloba*

The chloroplast genome includes 120–130 genes, includes genes for ribosomal RNAs, transfer RNAs, some ribosomal proteins, various polypeptide components of the photosystems involved in capturing solar energy, the catalytically active subunit of the enzyme ribulose 1,5-

bisphosphate carboxylase (RuBisCO = rbcl), and four subunits of a chloroplast-specific RNA polymerase.

This genome is very AT-rich, especially in non-coding regions. The typical structure of the chloroplast genome consists of a large single-copy (LSC) and a small single-copy (SSC) region separated by two identical inverted repeat (IR) regions. Much of cpDNA consists of non-coding sequences.

10.6. Replication of mitochondrial and chloroplast DNA

The replication of mitochondrial and chloroplast DNA occurs independently of nuclear DNA replication, with both organelles using a bacteria-like replication process. However, variations exist between their replication mechanisms.

10.7. Transcription and protein synthesis of organellar DNA

Similar to replication, the processes of organelle DNA transcription and protein synthesis take place where the genome is located : in the matrix of mitochondria and the stroma of chloroplasts. Although the proteins that mediate these genetic processes are unique to the organelle, most of them are encoded in the nuclear genome because the protein-synthesis machinery of the organelles resembles that of bacteria rather than that of eucaryotes. For example, chloroplast ribosomes are very similar to *E. coli* ribosomes, In addition, protein synthesis in mitochondria and chloroplasts starts with N-formyl methionine, as in bacteria, and not with the methionine as in eucaryotic cells.

10.8. The genetic codes of organelles

Translation in organelles follow the standard decoding process in which the organelle tRNA molecules translate codon by codon along the mRNA. However, the genetic code in mitochondria may differ somewhat from that used to encode proteins in prokaryotes, in archaea, and in nuclear genes of eukaryotes. Human mitochondria depart from the standard code in three principal ways : (1) The UGA codon is not a stop codon but encodes tryptophan, (2) the AGA and AGG codons are stop codons rather than arginine codons, and (3) AUA encodes methionine rather than isoleucine. The genetic code in yeast mitochondria resembles that in humans in that UGA is a tryptophan codon; however, in yeast, AUA is an isoleucine codon, AGA

10.9. Cytoplasmic inheritance

Cytoplasmic inheritance differs from the inheritance of characteristics encoded by nuclear genes. A zygote inherits nuclear genes from both parents, but typically, all its cytoplasmic organelles, and thus all its cytoplasmic genes, come from only one of the gametes, usually the

egg. A sperm from the male parent generally contributes only a set of nuclear genes. Thus, most cytoplasmically inherited traits are present in both males and females and are passed from mother to offspring, never from father to offspring.

CHAPTER 11

Population Genetics

ABSTRACT

This chapter covers the basic concepts in population genetics including the frequency of genes and genotypes in a population and the Hardy-Weinberg law of a stable population.

11.1. Introduction

Popultation is a group of organisms of the same species that live in the same geographical area (Ecosystem) at the same time.

Population genetics is the quantitative study of the distribution of genetic variation in populations and of how the frequencies of genes and genotypes are mainttained or change over time both within and between populations.

Population genetics is concerned both with genetic factors, such as mutation and reproduction, and with environmental and societal factors, such as selection and migration, which together determine the frequency and distribution of alleles and genotypes in families and communities.

11.2. The genetic composition of population

The genetic composition of an individual is called its genotype, while the genetic composition of a population is called a gene pool (Figure 11.1).

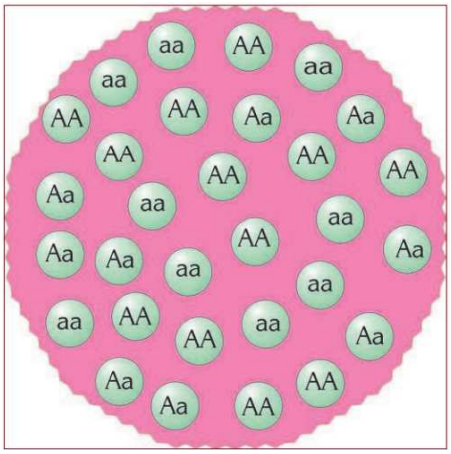


Figure 11.1. Gene pool

Every population has a unique genetic structure. This genetic structure consists of the genes found in the population and the frequency with which the genes are found.
To understand a population's genetic structure, we need to know the number of individuals and their genotypes.

11.2.1. Genotype frequency

Genotype frequency in a population is the number of individuals with a given genotype divided by the total number of individuals in the population.

We consider a population **n** carrying a gene with two alleles B and b.

The possible genotypes are : **BB**, **Bb** and **bb**.

If **n1** individuals have the genotype BB, **n2** individuals have the genotype Bb and **n3** individuals have the genotype bb, so: n1 + n2 + n3 = n

Therefore:

F (BB) = n1/n

F (Bb) = n2/n

F (bb) = n3/n

N.B. In a population, the sum of the genotypic frequencies is equal to 1.

F(BB) + F(Bb) + F(bb) = 1

11.2.2. Allele Frequency

The percentage of a specific allele out of the total of all alleles of a given locus in the population. The allele frequency is obtained by dividing the observed number of each allele by the total number of alleles in a population.

Allele frequency of B allele is :

f(B) = n1/n + 1/2 n2/n

= f (BB) + 1/2 f (Bb)

Allele frequency of b allele is :

f(b) = n3/n + 1/2 n2/n

= f (bb) + 1/2 f (Bb)

N.B. In a population, the sum of the allelic frequencies is equal to 1.

f(B) + f(b) = 1

11.2.3. Application exercices

Exercice n° 1 : A and a are two alleles. Complete the following table :

Genotype				
	AA	Aa	aa	Total
Number of individuals	30	60	10	100
Number of A allele				
Number of a allele				
Total number of alleles (A+a)				

Solution

Genotype				
	AA	Aa	aa	Total
Number of individuals	30	60	10	100
Number of A allele	60	60	0	120
Number of a allele	0	60	20	80
Total number of alleles (A+a)	60	120	20	200

Exercise n°2 : A study of human blood groups ABO in a population of 5000 individuals showed the results presented in the table bellow:

Genotype	AO	BB	AB	OO
Number	1600	700	200	2500

Solution

F(A)= ½(1600/5000) + ½ (200/5000) = 0.18

F(B)= 700/5000 + ½ (200/5000) = 0.16

F(O)= 2500/5000 + ½ (1600/5000) = 0.66

11.2.4. Summary

In a population :

- The sum of the allelic frequencies is equal to 1
- The sum of the phenotypic frequencies is equal to 1
- The sum of the genotypic frequencies is equal to 1

11.3. Hardy-Weinberg equilibrium

11.3.1. Hardy-Weinberg principle

Genetic equilibrium is a condition in which the gene frequencies remain the same from one generation to the next.

The Hardy-Weinberg principle is valid only if the population is stable (in equilibrium). The principle is that: if, in a population, the frequency of genes and individuals remains the same during the passage from one generation to the next, then the population remains stable.

The conditions required to satisfy the law are :

- the population is large
- mating is random
- no mutation
- alleles are not selected for (i.e. confer no survival or reproductive advantage)
- there is no migration into or out of the population.

The Hardy–Weinberg equation is derived by considering a population carrying an gene with two alleles A and a. The frequency of the dominant allele A in gametes is represented by p, and the frequency of the recessive allele a is represented by q.

Since there are only two alleles, **$p+q= 1$**.

If two heterozygotes mate, Aa x Aa, the frequencies of the genotypes AA, Aa, and aa in a population correspond to p^2, 2pq, and q^2, respectively.

$F(AA) = p^2$

$F(Aa) = 2pq$

$F(aa) = q^2$

Therefore :

$$\mathbf{p^2 + 2\,pq + q^2 = 1}$$

Where :

2pq : Frequency of heterozygous individuals

p^2 : Frequency of homozygous dominant individuals

q^2 : Frequency of homozygous recessive individuals

N.B. Hardy-Weinberg equation can be used to predict allele frequencies given genotype frequencies or predict genotype frequencies given allele frequencies.p^2+2pq : Frequency of individuals with the dominant phenotype.

11.3.2. Application exercise

Exercise : Within a population of butterflies, the color brown (B) is dominant over the color white (b).The population is in equilibrium. If the frequency of dominant phenotype is 0.98, calculate the allelic and genotypic frequencies for this population.

Solution

The frequency of dominant phenotype + the frequency of recessive phenotype =1

We have : the frequency of dominant phenotype = 0.98.

So, the frequency of recessive phenotype = 1- 0.98 = 0.02

The frequency of recessive phenotype = the frequency of recessive genotype .

So : F(bb)= 0.02.

$F(bb)= q^2 \Rightarrow q=\sqrt{0.02} \Rightarrow q= 0.14 \Rightarrow F(b)= 0.14$

We know that : p+q = 1. Thus : p= 1-q = 1- 0.14 => p= 0.86 => F(B) = 0.86

$F(BB)= p^2 = (0.86)^2 \Rightarrow F(BB)= 0.74$

F(Bb) = 2pq = 2x 0.86 x 0.14 => F(Bb) = 0.24

References

1. Abdelhakim Aouf A. (2015). Cours de Biologie Moléculaire et Génie Génétique. Editions Université Ferhat AbbasSétif.
2. Ameur Ameur A (2015). Génétique générale. Editions Al-Djazair
3. Bhagavan NV & Ha CE. (2015). Regulation of Gene Expression in Essentials of Medical Biochemistry, 2nd edition. Elsevier. http://dx.doi.org/10.1016/B978-0-12-416687-5.00024-5. P. 447-464.
4. Bousseboua H. (2003). Cours de Microbiologie générale, Imprimerie de l'université Mentouri Constantine, Algérie.
5. Daniell H., Lin CS, Yu M & Chang WJ. (2016). Chloroplast genomes: diversity, evolution, and applications in genetic engineering. Genome Biology, 17, 134 DOI 10.1186/s13059-016-1004-2.
6. Dobrogojski J, Adamiec M & Luciński R. (2020). The chloroplast genome: a review. Acta Physiologiae Plantarum, 42, 98. https://doi.org/10.1007/s11738-020-03089-x
7. Falus A, László V, Oberfrank F, Pap E, Szalai C &Tóth S. (2016). Medical Genetics And Genomics, 2nd eddition. Typotex Kiadó publisher, Budapest.
8. Gomes TMB, Vincent AE, Menger KE, Stewart JB & Nicholls TJ. (2024). Mechanisms and pathologies of human mitochondrial DNA replication and deletion formation. Biochemical Journal, 481, 683–715. https://doi.org/10.1042/BCJ20230262
9. Hartl DL & Jones EW. (1998). Genetics : Principles and Analysis, 4th edition. Jones & Bartlett Publishers.
10. Hao P, Yu J, Ward R, Liu Y, Hao Q, An S & Xu T. (2020). Eukaryotic translation initiation factors as promising targets in cancer therapy. Cell Communication and Signaling, 18:175. https://doi.org/10.1186/s12964-020-00607-9
11. Lim CS and & Brown CM. (2018). Know your enemy: successful bioinformatic approaches to predict functional RNA structures in viral RNAs. Frontiers in Microbiology. https://doi.org/10.3389/fmicb.2017.02582.
12. Lin CP, Wu CS, Huang YY & Chaw SM. (2012). The complete chloroplast genome of *Ginkgo biloba* reveals the mechanism of inverted repeat contraction. Genome Biology, 4(3):374–381. doi:10.1093/gbe/evs021.
13. Malumbres M (2020). Control of the cell cycle in Abeloff's Clinical Oncology, 6th edition. Elsevier. PP. 56-73.e5, https://doi.org/10.1016/B978-0-323-47674-4.00004-9.
14. Nicholl DST. (2008). An introduction to genetic engineering, 3rd edition. Cambridge university press.
15. Nussbaum RL, McInnes RR & Willard HF. (2016) Thompson and thompson genetics in medecine, 8th eddition. Elsevier.
16. Passarge E (2008). Atlas de poche de génétique. 3ème édition. Médecine-science, Flammarion.
17. Pierce BA. (2017). Genetics : A Conceptual Approach. 6th edition. W.H. Freeman edition. New york.

18. Ramos JL, García-Salamanca A, Molina-Santiago C & Udao Z. (2013). Operon in Brenner's Encyclopedia of Genetics, 2nd edition, Elsevier. Vol 5. P. 176-180. doi:10.1016/B978-0-12-374984-0.01096-2
19. Schleif R. (1993). Genetics and Molecular Biology, 2nd eddition. Johns Hopkins University Press.
20. Serre JL. (2001). Génétique : Rappels de cours, exercices et problèmes corrigés. DUNOD, Paris.
21. Singh NR (2019). Open Genetics, Thompson Rivers University press.
22. Snustad DP & Simmons MJ. (2012). Principles of genetics, 6th eddition, John Wiley & Sons.
23. Stansfield W (2003). Génétique. EdiScience.
24. Stubbs M & Suleyman N. (2013). Cell biology and genetics, 4th edition. Elsevier.
25. Tobias ES, Connor M & Ferguson-Smith M. (2011). Essential medical genetics, 6th edition, Willey Blackwell.
26. Turudic A, Liber Z, Grdiša M, Jakše J, Varga F & Šatovic Z. (2021). Towards the well-tempered chloroplast DNA sequences. Plants, 10, 1360. https://doi.org/10.3390/plants10071360.
27. Verma PS & Agarwa VK (2005). Cell biology, genetics, molecular biology, evolution and ecology. S. Chand & Company Ltd, India.
28. William S. Klug, Michael R. Cummings, Charlotte A. Spencer, Michael A. Palladino, Darrell J. Killian. (2020). Concepts of genetics, 12th eddition. Pearson Education Limited.
29. Winter PC, Hickey GI & Fletcher HL. (2006). L'essentiel en génétique. BERTI éditions, Paris.