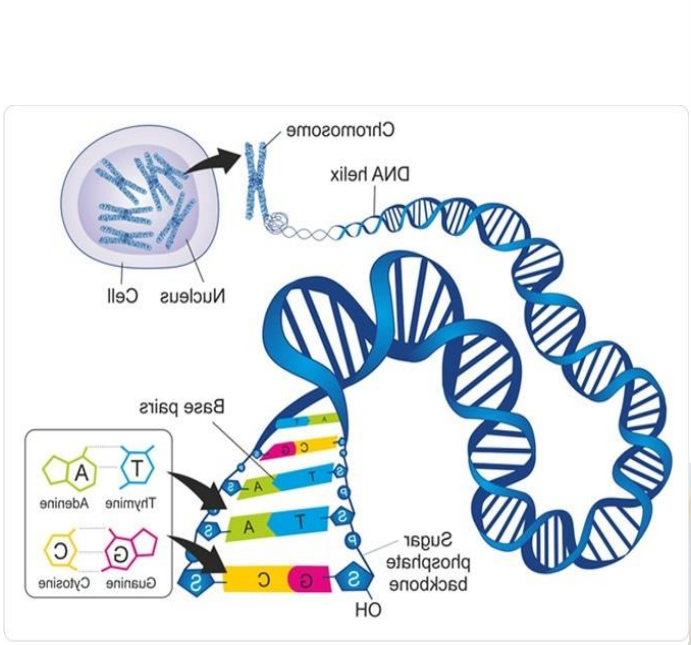
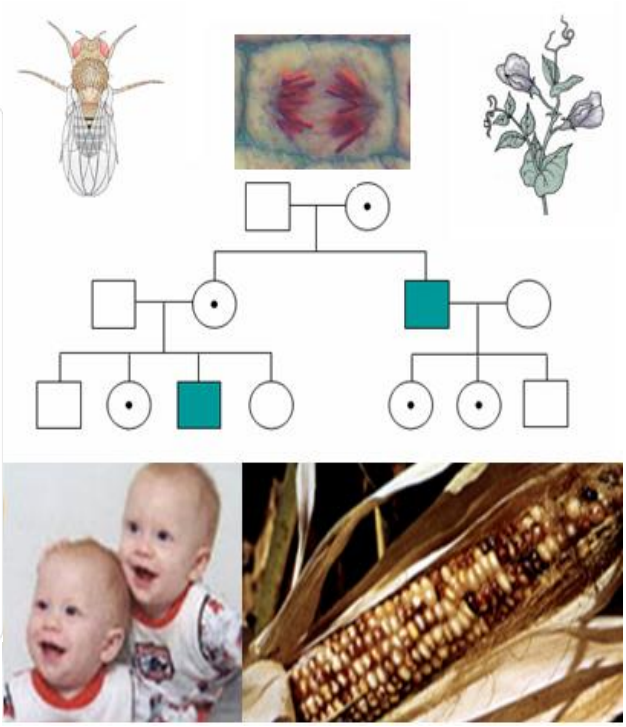
# Lectures on Genetics



**Prepared by**

**Dr\ Hoida Zaki**

***Botany and Microbiology Department***

**2022-2023**

*Chapter 1*

# MENDELAN GENETICS

In genetic experiments, it is important to obtain the correct experimental organism for the experiments you which to perform. For classical genetics, genetic stocks that contain the alternate alleles that you will be studying at the morphological level are required. When performing an experiment in molecular genetics, you will need a genetic stock containing the specific gene that you want to study at the molecular level, and a probe to identify and analyze the specific alleles. For many studies, the allele to be studied is the wild type allele and its alternatives may be mutants. In population and evolutionary genetic studies, a population with varying gene frequencies is necessary. To determine what type of genetic stock or population is best for your specific experiments, it is important to understand the advantages and disadvantages of the different biological species that are used as test organisms. In principle, a good genetic experimental organism should have the following criteria: -

**1.** Have short life cycle
**2.** Easy to grow at low costs
**3.** Produces a large number of progeny
**4.** Have many observable contrasting characters

Gregor Mendel (1822–1884) was the first to carry out controlled experiments on genetics. The results of his work on the garden peas (*Pisum sativum* L.) prepared the ground for our modern understanding of genetics. He did so by tackling the problem in a logical fashion applying the following ideas: -

1. Concentrating on one or few observable contrasting characters
2. Making controlled crosses in a carefully designed experiments
3. Keeping careful records and analysis of the results

4. Suggesting factors as the particular causes of genetic patterns

Mendel was fortunate to choose the garden peas because.

**1.** Garden pea has a short life cycle.

**2.** Garden pea has several observable contrasting characters.

**3.** Garden pea is a self pollinating plant.

**The standard monohybrid cross**

A cross involving contrasting expression of the same character is referred to as a monohybrid cross. For example, if we want to learn something of the inheritance of vestigial wing in the fruit fly *Drosophila melanogaster*, we would cross a fly having normal, full sized wings with one vestigial wing ignoring all other characters. However, here we shall study the monohybrid cross along the general lines of Mendel's work. Mendel made controlled crosses on seven traits of garden peas; the traits he used are listed in table 1-1 and illustrated in figure1.

Mendel made a cross between tall plants and dwarf plants of garden peas. The resulting seeds were collected and planted the following season. All resulting plants were tall. The flowers of these plants were self-pollinated, and the resulting seeds were collected and planted the following season. Mendel found that from the growing plants 787 were tall and 277 were dwarf a ratio of **3** tall : **1** dwarf. The two originally used plants are called parents and designated as "P" The first generation comprising only tall plants is known as the first filial generation or the "$F_1$". The second generation comprising 75% tall plants and 25% dwarf plants is the second filial generation or the "$F_2$".

The results for the seven monohybrid crosses in the garden peas were similar to those described for the plant height and their results are given in table 1-1.

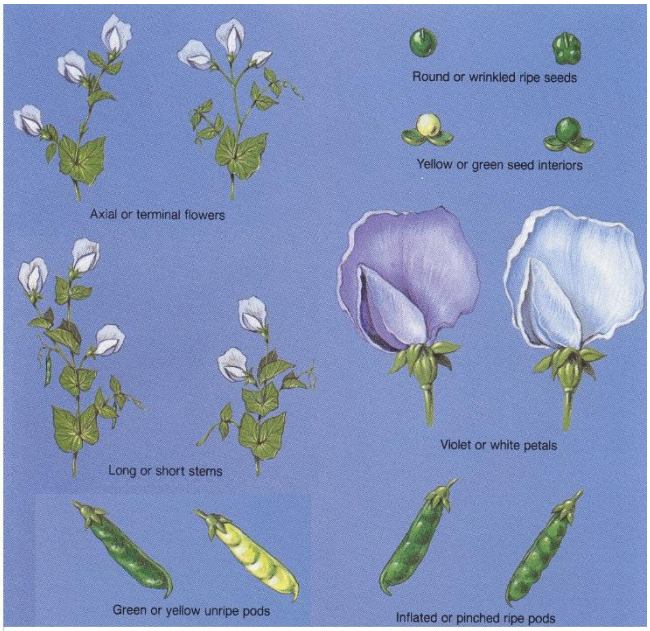



Figure 1-1: The seven traits studied by Gregor Mendel in his experiments on the inheritance of traits in garden peas.

It is clear from table 1-1 that $F_1$ always resembled one of the parents. Such characters or trait which expresses itself in all the individuals of the $F_1$ are termed dominant characters, and that which failed to be expressed in the $F_1$ are referred to as recessive. The $F_2$ plants obtained by selfing the garden peas plants of the $F_1$ or by interbreeding members of the $F_1$ showed 3 times as dominant plants as recessive plants.

**Table 1-1:** List of the seven monohybrid crosses studied by Mendel in the garden peas.

| Parents cross | $F_1$ | $F_2$ | Ratio |
|---|---|---|---|
| Round x Wrinkled seeds | Round | 5474 R., 1850 W | 2.96:1 |
| Yellow x Green cotyledons | Yellow | 6022 Y, 2001 G | 3.01:1 |
| Violet x White petals | Violet | 705 B, 224 W | 3.15:1 |
| Long x Short stems | Long | 787 L, 277 S | 2.84:1 |
| Axial x Terminal flowers | Axial | 651 A, 207 T | 3.14:1 |
| Inflated x Constricted pods | Inflated | 882 I, 299 C | 2.95:1 |
| Green x Yellow pods | Green | 428 G, 152 Y | 2.81:1 |

Because the $F_1$ plants segregate in the $F_2$ to produce dominant and recessive plants, Mendel recognized that the $F_1$ plants must contain factors responsible for the dominant and recessive characters. If the factor for tallness is represented by **T** and for dwarfism by **t,** the factors contained in the $F_1$ plants may be designated as **Tt**. Since the $F_1$ plants

contain the genes **T** and **t** for plant height, the tall parent plant must contain two identical genes **TT** responsible for tallness and the dwarf parent plant must contain two genes **tt** responsible for dwarfism. Thus, the parent plants whether they are tall or dwarf are homozygous because they contain two similar genes. While the $F_1$ plants which contain two different genes are called heterozygous.

The appearance of an organism with regard to a character under consideration, constitute the phenotype which is generally designated by a descriptive word or phrase. On the other hand, and individual's genetic make up is termed genotype, and is generally represented by letters of alphabet or other convenient symbol. In the case of Mendel experiments tall and dwarf represent the phenotypes, while **TT**, **Tt** and **tt** represent symbols to genotypes. According to the general genetic terminology the results are often summarized in the following simple diagram and are usually illustrated as in Fig. 3.

| | | | |
|---|---|---|---|
| **Parents** | **Tall** | **x** | **Dwarf** |
| **$F_1$** | **Tall** | | |
| **F2** | **75% Tall** | | **25% Dwarf** |

Because the sex cell or gametes constitute the only link in sexually reproducing organisms between parents and offspring, it is obvious that genes must be transmitted from generation to generation via the gametes. The sperm and egg in a particular type of cross do make equal genetic contribution to the next generation. This is strongly indicated by the fact that whether dwarf plants in Mendel's experiments served as the female (egg contributor) or male (sperm contributor) the result was the same, all the $F_1$ plants were tall. Mendel assumed that tall homozygous plants **TT** would produce **T** pollen and **T** eggs and the dwarf homozygous plants **tt** would give **t** pollen and **t** eggs. On selfing plants of the $F_1$ or crossing two heterozygous plants of the F1, there would be two types of gametes from

each plant; **T** and **t** and the possibility of fertilization is one of the following possibilities that would occur in equal probabilities (Figure 1-3)

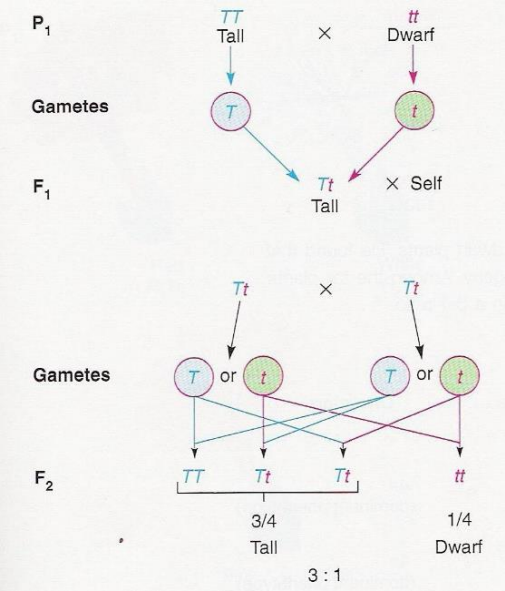



**Figure 1-3:** Diagram illustrating the transmission of plant height through the segregation of the pair of alleles governing this trait in garden peas.

The results of the Monohybrid cross are often illustrated by the so called Punnett square, presented as a checker board table, based on the following possibilities (Table 1-2).

**1-** **T** pollen may fertilize **T** egg resulting in a **TT** genotype and tall homozygous plants

**2-** **T** pollen may fertilize **t** egg resulting in a **Tt** genotype and tall heterozygous plants

**3-** **t** pollen may fertilize **T** egg resulting in a **Tt** genotype and tall heterozygous plants

**4-** **t** pollen may fertilize **t** egg resulting in a **tt** genotype and dwarf homozygous plants

**Table 1-2:** A Punnett's square illustrating the genotypes and phenotypes of the $F_1$ garden pea plants based on the possibilities that male gametes and female gametes of the $F_1$ plants are mated at equal chances.

| ♀ \ ♂ | **T** | **t** |
|---|---|---|
| **T** | **Tall, TT** | **Tall, Tt** |
| **T** | **Tall, Tt** | **Dwarf, tt** |

In the monoybrid cross, the $F_2$ genotype ratio = **1 TT** : **2 Tt** : **1 tt** and the $F_2$ genotype ratio = **3** dominant (tall): **1** recessive (dwarf).

Mendel concluded that each character or trait is controlled by a pair of factors known as allelomorphic genes or alleles. Segregation of the members of allelic genes during meiosis and their reunion in the hybrid is in fact the bases of Mendel's law of segregation which may be stated as follows: "Each character is controlled by a pair of factors (genes), the members of this pair of genes segregate from each other during gametes formation and reunite in the offspring". In brief, the principle of segregation means that heterozygotes produce equal numbers of gametes that have the two different alleles.

## The test cross

As we have seen **TT** and **Tt** individuals in the garden pea cannot be distinguished visually from each other. This raises the question that, can homozygous dominant and heterozygous individuals be distinguished in any manner? The answer is yes, and the test cross can supply the answer. In the test cross, the individual having the dominant phenotype (which can be represented by a genotype **TT** or **Tt** is crossed with one having a recessive phenotype (and of course a homozygous recessive genotype **tt**. In our example the test cross of homozygous dominant is as follows:

| | | | |
|---|---|---|---|
| **Parents** | **Tall, TT** | **tt** | **Dwarf** |
| **P Gametes** | **T** | **t** | |
| **F1** | **All are tall heterozygous Tt** | | |

If the parent dominant phenotype had been heterozygous, the results would be a classic **1** : **1** monohybrid test cross ratio.

| | | | |
|---|---|---|---|
| **Parents** | **Tall, Tt** | **tt** | **Dwarf** |
| **P Gametes** | **T or t** | **t** | |
| **F1** | **Tall Tt** | **Dwarf** | **tt** |

It is important to emphasize that the term $F_1$ is used to designate the first generation resulting from any given mating regardless of the parental genotype. Therefore, the term $F_1$ does not necessarily imply heterozygosity, it depends on the parent gene types and may be either homozygous or heterozygous.

## Conclusions on monohybrid cross

Gene for stem height in pea plants has two different forms or alleles. Tall allele is dominant (only parental phenotype seen in $F_1$). Dwarf allele is recessive (parental phenotype missing in $F_1$ but reappears in $F_2$). $F_2$ generation segregates into two phenotypic classes (tall and dwarf) in **3:1** ratio. Similarly, each of the other six traits examined by Mendel has a dominant allele and a recessive allele that segregate at meiosis.

## Dihybrid inheritance

A cross involving a single pair of genes is equally applied to crosses involving two pairs or more of genes. In tracking segregation for two traits; Mendelian dihybrid crosses assume that:

1. Two clearly distinguishable phenotypic traits; each trait controlled by a single gene locus.
2. Only two alleles at each locus (gene only exists in two forms).
3. No interaction between the two loci.

In addition to plant height we shall consider seed shape. Seeds of the garden peas may be round or wrinkled. The round shape is dominant while wrinkled shape is recessive. This was evident from the results of a simple monohybrid cross involving these two characters, carried out by Mendel himself. A cross of a double homozygous tall plants having round seeds with dwarf plants having wrinkled seed using T and t again to represent the tall and dwarf genotypes and **R** and **r** to denote round and wrinkled seeds.$F_1$ plants are usually referred to as dihybrid individuals because they are heterozygous for each of the two pairs of genes. These

heterozygous plants produce four types of gametes in equal numbers of four possible gamete genotypes: **TR, Tr, tR** and **tr** (Table 1-3)

Selfing a dihybrid individual of the first filial generation or crossing two of the $F_1$ plants to produce the $F_2$ plants produced one of the classic phenotypic ratios, based on a random combination of phenotypic classes. The ratio is **9/16** tall plants with round seeds, **3/16** tall plants with wrinkled seeds, **3/16** dwarf plants with rounded seeds and **1/16** dwarf plants with wrinkled seeds.

The results outlined above are based on expectation that, in gametes, association of **T** will be with **R** or **r** and of **t** with **R** or **r** is random based on the law of independent assortment illustrated by the inheritance of plant height and seed shape in garden peas. A Punnett square expressed as checkerboard diagram illustrating the expected genotypes in the $F_2$ plants are illustrated in table 1-4.

It is clear that the nine genotypes **T-R-** give tall plants with rounded seeds, three genotypes **T-rr** give tall plants with wrinkled seeds, three genotypes **ttR-** give dwarf plants with round seeds and only one, the **ttrr** genotype gives dwarf plants with wrinkled seeds.

**Table 1-3:** Punnett's Square illustrating a cross of a double homozygous tall plants having round seeds (**TTRR**) with dwarf plants having wrinkled seeds (**ttrr**) to produce double heterozygous plants that produce four types of gametes in equal numbers of four possible gamete genotypes: **TR, Tr, tR** and **tr.**

| Parents | Tall plants, Round seeds<br>TTRR | Dwarf plants, wrinkled seeds<br>ttrr |
|---|---|---|
| P. Gametes | TR | tr |
| F1 | Phenotype: Tall plants Round seeds<br>Genotype: TtRr | |
| F1 Gametes | T → R → TR; T → r → Tr<br>t → R → tR; t → r → tr | |

The dihybrid cross is two monohybrid crosses acting independently. When we consider the $F_2$ progeny in term of each character alone, 3 : 1 ratio will be obtained. On counting the tall plants of the $F_2$ we get 12/16 and on counting the dwarf plants we get 4/16 and on doing the same on seed shape a ratio of 12 round and 4 wrinkled is obtained (Table 1-4). That is, each pair of genes behaved exactly as it would in a one pair cross, and this forms the bases for Mendel's law of independent assortment which states that: "Each pair of genes segregate independently of all other such pairs during the formation of gametes".

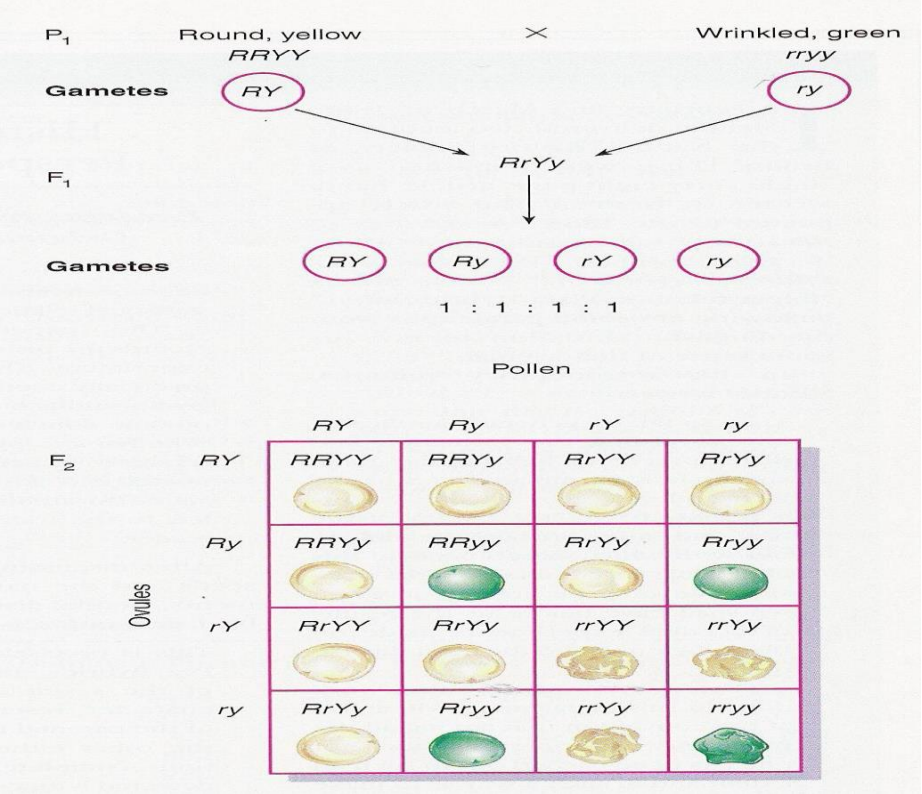
**Table 1-4:** A Punnett's square illustrating production of four phenotypes in the $F_2$ of through mating of the four gamete: **TR, Tr**, **tR** and **tr** produced by the F1 plants.

| ♂ ♀ | **TR** | **Tr** | **tR** | **tr** |
|---|---|---|---|---|
| **TR** | **TTRR** Tall, Round | **TTRr** Tall, Round | **TtRR** Tall, Round | **TtRr** Tall, Round |
| **Tr** | **TTRr** Tall, Round | **TTrr** Tall, wrinkled | **TtRr** Tall, Round | **TtRr** Tall, wrinkled |
| **tR** | **TtRR** Tall, Round | **TtRr** Tall, Round | **ttRR** Dwarf, Round | **ttRr** Dwarf, wrinkled |
| **tr** | **TtRr** Tall, Round | **Ttrr** Tall, wrinkled | **ttRr** Dwarf, wrinkled | **ttrr** Dwarf, wrinkled |

Another example illustrating the tracking of the segregation of Mendelian dihybrid crosses assume is the inheritance of seed shape and seed color in garden peas. Like plant height and seed shape, these two traits are clearly distinguishable phenotypic traits each of which is controlled at single locus by two alleles at each locus (gene only exists in two forms) and no

interaction exists between the two loci. The inheritance of these two traits is best illustrated in by the diagram given in figure 1-4.

**Figure 4-4:** Punnett's square illustrating the inheritance of seed shape and color as two independent traits in garden peas.




## Polyhbrid inheritance

As seen earlier, a monohybrid such as **Tt** produces two kinds of gametes **T** and **r**, in equal numbers, that can combine by fertilization to form three different zygote genotypes **TT**, **Tt** and **tt**, from which two phenotypes (tall and dwarf plants) were produced in progeny. Likewise a doubly heterozygous individual **TtRr** produced four kinds of gametes **TR, Tr, tR** and **tr**, again in equal numbers. With random mating, nine different zygote genotypes are produced these are **TTRR, TTRr, TtRR, TtRr, TTrr**, **Ttrr, ttRR, ttRr and ttrr**. From these nine genotypes four phenotypes are produced, these are tall plants with rounded seeds, tall plants with wrinkled seeds, dwarf plants with rounded and dwarf plants with wrinkled seeds.

A third pair of genes can be symbolized as **Y** for yellow cotyledons and **y** for green cotyledons. When studying the inheritance of three characters in the same organism it is called a trihybrid cross. A heterozygous trihybrid of garden peas will have the genotype **TtRrYy** this is a tall plant with rounded seeds and yellow cotyledons. What gamete combination can be formed by such a trihybrid? If each of the three pairs of genes is on a

different chromosome pair, then either allele of any pair can combine with either allele of any other pair. A simple application of the probability method indicates eight possible gamete genotypes as in figure 1-5.

| | | |
|---|---|---|
| **P** | **Tall plants, round yellow seeds TTRRYY** | **Dwarf plants, wrinkled, green seeds ttrryy** |
| **P Gametes** | **TRY** | **try** |
| **F1** | **Tall plants, round yellow seeds TtRrYy** | |

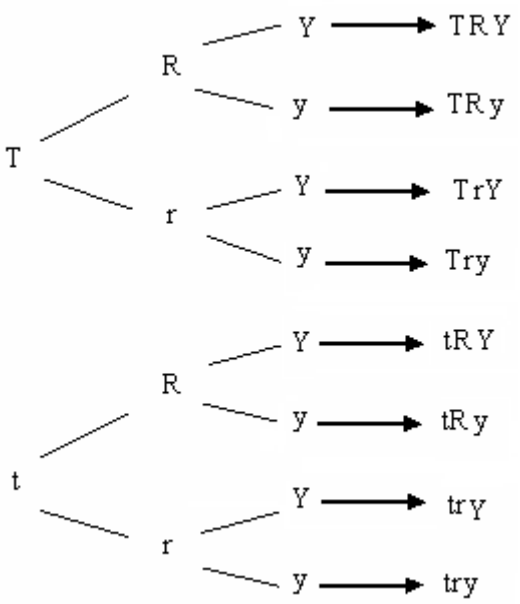



**Figure 1-5:** The probability method indicates eight possible gamete genotypes ware produced by the trihybrid $F_1$.

If we apply again the probabilities used for the monohybrids and dihybrids these eight gamete genotypes will randomly combine into 64 combinations forming 27 different zygote genotype combinations as in Table 1-5.

**Table 1-5:** Details of the 27 different zygote combinations as a result of crossing two heterozygous trihybrids **TtRrYy**.

| ? \ ? | TRY | TRy | TrY | Try | tRY | tRy | trY | try |
|---|---|---|---|---|---|---|---|---|
| TRY | TTRRYY | TTRRYy | TTRrYY | TTRrYy | TtRRYY | TtRRYy | TtRrYY | TtRrYy |
| TRy | TTRRYy | TTRRyy | TTRrYy | TTRrYy | TtRRYy | TtRRyy | TtRrYy | TtRryy |
| TrY | TTRrYY | TTRrYy | TTrrYY | TTrrYy | TtRrYY | TtRrYy | TtrrYY | TtrrYy |
| Try | TTRrYy | TTRryy | TTrrYy | TTrryy | TtRrYy | TtRryy | TtrrYy | Ttrryy |
| tRY | TtRRYY | TtRRYy | TtRrYY | TtRrYy | ttRRYY | ttRRYy | ttRrYy | ttRrYy |
| tRy | TtRRYy | TtRRyy | TtRrYy | TtRryy | ttRRYy | ttRRyy | ttRrYy | ttRryy |
| trY | TtRrYY | TtRrYy | TtrrYY | TtrrYy | ttRrYY | ttRrYY | ttrrYY | ttrrYy |
| try | TtRrYy | TtRryy | TtrrYy | TtrrYy | ttRrYy | ttTRryy | ttrrYy | ttrryy |

Considering that dominance is complete, eight phenotypic classes will be produced in the ratios given in Table 1-6.

Figure 1-6: Number of the phenotypic classes and possible genotypes in the F2 of a trihybrid cross of garden pea.

| Number | Phenotypes | Genotypes |
|---|---|---|
| 27 | Tall plants, round seeds, yellow cotyledons | T-R-Y- |
| 9 | Tall plants, round seeds, green cotyledons ons | T-R-yy |
| 9 | Tall plants, wrinkled seeds, yellow cotyledons | T-rrY- |
| 9 | Dwarf plants, round seeds, yellow cotyledons | ttR-Y- |
| 3 | Tall plants, wrinkled seeds, green cotyledons | T-rryy |
| 3 | Dwarf plants, round seeds, green cotyledons | ttW-yy |
| 3 | Dwarf plants, wrinkled seeds, yellow cotyledons | ttrrY- |
| 1 | Dwarf plants, wrinkled seeds, yellow cotyledons | ttrryy |

Again when examine the inheritance of each character independently we shall observe a 3:1 ratio dominant to recessive phenotypes for each of the three traits.

It is clear from trihybrid example that the number of gamete genotypes produced by parents is **8**, i.e, $\mathbf{2^3}$, the number of progeny phenotypic classes is **8**, i.e. $\mathbf{2^3}$, and the number of progeny genotypic classes is **27**, i.e. $\mathbf{3^3}$. When the number of gene pairs is increased to 4, 5 or n the number of gamete genotypes, the number of progeny phenotypes and the number of progeny genotypes can be deduced from the Table 1-7.

**Table 1-7:** Number of gene pairs, number of gamete genotypes, number of progeny phenotypes and the number of progeny genotypes of traits inherited in a Mendelian fashion**.**

| Number of pairs of heterozygous genes | Number of gamete genotypes | Number of progeny phenotypes | Number of progeny genotypes | Number of possible combination of gametes |
|---|---|---|---|---|
| **1** | **$2^1 = 2$** | **$2^1 = 2$** | **$3^1 = 3$** | **$4^1 = 4$** |
| **2** | **$2^2 = 4$** | **$2^2 = 4$** | **$3^2 = 9$** | **$4^2 = 16$** |
| **3** | **$2^3 = 8$** | **$2^3 = 8$** | **$3^3 = 27$** | **$4^3 = 64$** |
| **4** | **$2^4 = 16$** | **$2^4 = 16$** | **$3^4 = 81$** | **$4^4 = 256$** |
| **n** | **$2^n$** | **$2^n$** | **$3^n$** | **$4^n$** |

## Test cross of dihybrids and polyhybrids

The test cross is very valuable in determining genotypes in situation with two or more pairs of genes. In a two pair cross, where each gene is on a different chromosome pair (i.e. genes are not linked), the resulting progeny genotype ratio will be product of two one-pair ratios. Thus a dihybrid testcross produces a 1:1:1:1 ratio when one parent is heterozygous for both gene pairs as follows:

| **Parents** | Tall plants, round seeds **TtRr** | Dwarf plants, wrinkled seeds **ttrr** |
|---|---|---|
| **P Gametes** | **TR, Tr, tR, or tr** | **tr** |
| **F1** | Tall plants, round seeds **TtRr**<br>Tall plants, wrinkled seeds **Ttrr**<br>Dwarf plants, round seeds **ttRr**<br>Dwarf plants, wrinkled seeds **ttrr** | |

**Chapter 1 Post-test**

**A- Mark the correct answer with √ and the wrong answer with x**

1. A monohybrid cross involves contrasting expressions of 4 different characters
2. The test cross involves crossing two heterozygous contrasting phenotypes
3. In a tri-hybrid heterozygous hybrid (**AaBbCc**) the number of possible gametes is eight
4. The genotype of a dominant expression of a monohybrid trait is always homozygous

**B- Circle the correct answer of the following:-**

1**-** The external appearance of traits is called:-

a. Phenotype b. Genotype

c. Cytotype d. Ecotype

2- The seed coat color in garden pea is controlled by:-

a. Two allele b. Four alleles

c. Pseudo-alleles d. Six alleles

3- The first law of Mendel is called law of:-

a. Segregation b. Separation

c. Delimitation d. Identification

*Chapter 2*

# GENE INTERACTION

## 1- Modification of the 3:1 ratio:

Variations in the classic 3:1 ratio of the Mendelian monohybrid cross have been found in many cases. This variation is due to an interaction of the two alleles governing the character under consideration by a number of ways which include:

### a. Incomplete dominance

When cells of the cotyledons of the three genotypes (**RR**, **Rr**, round and **rr**, wrinkled) were examined microscopically an abundance of well-formed starch grains was seen in the **RR** plants, and few in the **rr** plants. Heterozygotes showed an intermediate number of grains, many of which are imperfect or eroded. Therefore, the gene **R** must be considered incompletely dominant to its allele **r**. Whether round versus wrinkled seeds is considered as complete or incomplete dominance depends entirely on the level of analysis. The number of starch grains reflects the relationship of **R** and w genes at the physiological level, while round versus wrinkled seed shape illustrate their relationship at the morphological level. Because the physiological level more accurately reflects the actual operation of genes, most geneticists would prefer to regard this case as incomplete dominance.

It is important to note that in the incomplete dominance the heterozygotes are phenotypically intermediate between the two homozygous types. Example of incomplete dominance at the morphological level in both plants and animals are numerous.

A classical case and one of the earliest to be described is that of the common ornamental flowering plant snapdragon (*Antirrhinum majus*). Homozygotes are either red or white flowered. When an individual with red

flowers is crossed to another with white flowers the resulting $F_1$ plants are pink flowered.

When a pink flowered plant of the $F_1$ is crossed to another $F_1$ pink flowered plant, the $F_2$ plants show 1 red : 2 pink : 1 white. This cross is summarized as in table 3-1: Another example of incomplete dominance in plants is found in radish (*Raphanus sativus*). Radishes may be long, oval or round. Crosses of long with round shaped radishes produce an $F_1$ of wholly oval phenotype.

**Table 3-1**: Punnett square illustrating a cross of a red homozygous plants having red flower (**RR**) with plants having white seeds (**rr**) to produce heterozygous plants with pink flowers

| **Parents** | **Red**<br>**RR** | **White**<br>**rr** |
|---|---|---|
| **P gametes** | **R** | **r** |
| **F1** | **All are Pink - Heterozygous Rr** | |
| ♂<br>♀ | **R** | **r** |
| **R** | **Red**<br>**RR** | **Pink**<br>**Rr** |
| **r** | **Pink**<br>**Rr** | **White**<br>**rr** |

## b. Co-dominance

In the shorthorn cattle genes for red and white coat color control the coat color. Crosses between red and white produce offspring whose coat color may be reddish gray or roan. Superficially this would seem to be a case of incomplete dominance, but close examination of roan animals disclosed the coat to be composed of mixture of red hairs and white hairs, rather than hairs all of a color intermediate between red and white. Instances such as this, where the heterozgotes exhibit a mixture of the phenotypic characters

of both homozygotes, instead of a single intermediate expression, illustrate co-dominance. The above example is summarized as in table 3-2.

**Table 3-2**: Punnett square illustrating a cross of a red homozygous shorthorn cattle having red flower ($r_1r_1$) with cattle having white hairs ($r_2r_2$) to produce heterozygous animals with roan hairs ($r_1r_2$)

| **Parents** | **Red** $r_1r_1$ | **White** $r_2r_2$ |
|---|---|---|
| **P gametes** | $r_1$ | $r_1$ |
| **F1** | **All are Roan - Heterozygous** $r_1r_2$ | |
| ♀ ♂ | $r_1$ | $r_2$ |
| **R** | **Red** $r_1r_1$ | **Roan** $r_1r_2$ |
| **r** | **Roan** $r_1r_2$ | **White** $r_2r_2$ |

Several cases of co-dominance occur in human genetics. An important case is found in sickle-cell phenotype, which is particularly prevalent in peoples of the Negro race. The hemoglobin of most persons is of a type known as hemoglobin A, but some persons have a hemoglobin different form A and referred to as hemoglobin S, the presence of which causes the sickle-cell disorder. Genes responsible for the two hemoglobin types are $\mathbf{Hb^A}$ and $\mathbf{Hb^S}$. Most persons belong to a genotype $\mathbf{Hb^AHb^A}$, and their erythrocytes (red blood cells) contain only $\mathbf{Hb^A}$ and $\mathbf{Hb^A}$ genes and are disc-shaped. Persons with sickle-cell anaemia are of the genotype $\mathbf{Hb^SHb^S}$ and are characterized by some symptoms, chiefly a chronic hemolytic

anaemia. In their blood the erythrocytes are distorted being sickle- shaped and can not perform their functions.

In heterozygotes **$Hb^AHb^S$**, some red cells contain hemoglobin **A** and others hemoglobin **S.** Microscopic examination of heterozygotes discloses normal and sickle-shaped cells (Fig. 3-3). In normal conditions heterozygotes show none of the severe symptoms of **$Hb^SHb^S$** persons, though they may suffer periodic discomfort and even develop anaemia after a time at high altitude. Because both types of hemoglobin are produced, this is considered a case of co-dominance.

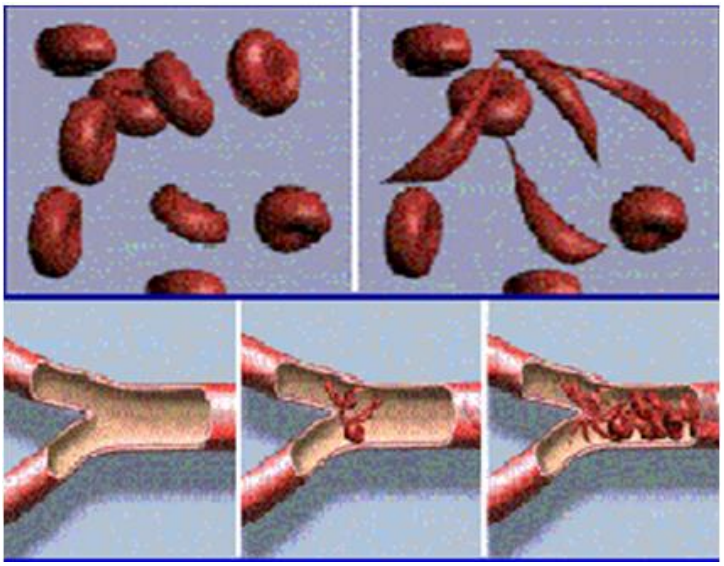


**Figure 3-2**: Co-dominance in humans; Sickle cell anemia, Symptoms: Sickle shaped red blood cells in **$Hb^AHb^S$** persons, circulatory problems, loss of blood cells (anemia), organ damage, death

## General information on sickle cell anemia

**Cause:** Autosomal recessive mutation on chromosome 11

**A** changed to **T** (**glu to val**) see figure 3-3.

**Function:** Gene that codes for part of hemoglobin protein

**Frequency:** More common in Africans; One in 500 = is carrier for gene **$Hb^AHb^S$**

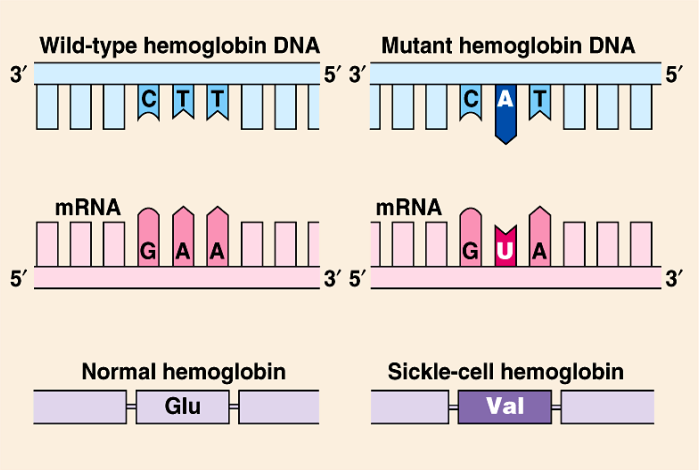



**Figure 3-3**: Diagrammatic representation of the mutation of hemoglobin gene **A** to hemoglobin gene **S** that leads to sickle cell anemia

## c. Lethal genes

A major modification of the classic monohybrid ratio (3:1) is produced by some genes whose effect is lethal to some bearers of certain genotypes. Lethal genes, however, mostly exhibit their effect in a homozygous genotype. They are common in plants, animals and man. The following are common examples.

### Albinism, a recessive lethal gene in plants

An example of lethal genes is the albinism in humans, animals and plants. In man, albinism is known as the lack of melanin pigment in skin and eyes. Albinism in plants is expressed as lack of chlorophyll which is a pigment required for photosynthesis. In maize several pairs of genes affecting chlorophyll production have been described. One such gene which often designated as **G** for normal chlorophyll production is completely dominant over its allele g. It was found that plants wit a genotype **GG** or **Gg** contain chlorophyll and are photosynthetic. On the other hand, **gg** plants produce no chlorophyll and are yellow-white. On the average one fourth of progeny of two heterozygous parents are without chlorophyll as illustrated in table 3-3.

**Table 3-2:** Punnett's square illustrating the inheritance of chlorophyll production a completely dominant trait in maize, a recessive lethal gene in plants

| Parents | Green | Gg × Gg | Green |
|---|---|---|---|
| ♀ ♂ | G | | g |
| G | GG<br>Green | | Gg<br>Green |
| g | Gg<br>Green | | gg<br>White, Dies |
| | 3- Green | | 1- white, Dies |

## Coat color in mice, a dominant lethal gene

The coat color in mouse is controlled by a gene designated as **Y** in its dominate form and **y** for its recessive form. Crosses of heterozygous yellow with heterozyhous yellow always produced yellow and black in a 2 : 1 ratio. In such crosses the numbers of F1 individuals were one fourth smaller than those of other crosses. Since the ratio of such mating should be 3 yellow to 1 black, it was concluded that homozygous individuals **YY** die in an early stage of development:

**Table 3-3**: Punnett's square illustrating the inheritance of coat color in mice as a dominant lethal gene

| Parents | Yellow | Yy × Yy | Yellow |
|---|---|---|---|
| ♀ ♂ | Y | | Y |
| Y | YY<br>Yellow Dies | | Yy<br>Yellow |
| y | Yy<br>Yellow | | yy<br>Black |
| | 2- Yellow | | 1- Black |

This conclusion was supported by supported the results of a test cross between heterozygous yellow **Yy** and homozygous black **yy**. The results were a 1 : 1 ratio yellow to black is follows:

| | | |
|---|---|---|
| **Parents** | **Yellow Yy** | **Black yy** |
| **P Gametes** | **Y or y** | **y** |
| **F1** | **Yellow Yy** | **Black yy** |

The gene **Y** appears to be dominant with respect to coat color, but recessive as to lethality. For some time the action of **YY** was unknown Roberson (1942) and Eaton and Green (1962) were able to demonstrate that about one fourth of the embryos of pregnant yellow **Yy** females that were mated to yellow **Yy** males did die shortly after conception. In this case death usually occurs at gastrulation.

**2- Modification of the 9:3:3:1 ratio:**

Interaction between two pairs of genes occurs in several cases and causes a modification of the classical dihybrid ratio 9:3:3:1. This results in a gene masking or modifying the effect of one or both alleles of another gene. Such interaction operates in several ways and various modified ratios have been reported in both plants and animals. The first case of two genes controlling one trait was reported by Bateson and Punnett and demonstrated by the inheritance of crown shape in domestic fowl. The crown varies in size, shape and color, both roosters and hens have combs, but the male's crown is usually larger and brighter than the female's. A diagram illustrating the two genes controlling the crown shape in domestic fowl is shown in Fig. 3-5.

The phenotypic ratio of four shapes is 9:3:3:1 indicating the interaction of two genes (**R & P**) in the production of crown shape. The walnut shape is produced in 9/16 when the two genes are dominant (**R-P-**). A rose shape is produced in 3/16 when the gene R is dominant and the gene **P**

is recessive (**R-pp**) and a pea shape is 3/16 when produced in the gene **P** is dominant and the gene **R** is recessive (**rrP-**) and a singe shape in produced in 1/16 when the two genes are recessive (**rrpp**).

For convenience the interacting pair of gene has been given certain terms to facilitate their explanation. Form these the following are the most common: epistatic genes, modifying (Supplementary) genes, complementary genes, and duplicate genes.

**a. Epistatic genes**

A gene which masks the effect of another gene is said to be epistatic; the masked gene or genes are called hypostatic. When pure white leghorn and white silkie domestic fowls were crossed, the $F_1$ hybrids were all with white feather. Mating of $F_1$ hybrids together resulted in an $F_2$ with a phenotype ratio of 13 white to 3 colored. These results were explained by postulating that the leghorn individual have a genotype **CCII** in which the gene **I** masks the effect of the gene **C**, while the genotype of the silkei individuals is **ccii**, and the genotype of the $F_1$ members is **CcIi** and they are white colored because the gene **I** masks the effect of the gene **C.** The result of mating two individuals is as illustrated in table 3-4, 13 white and 3 colored.

| | | |
|---|---|---|
| **P** | **White Leghorn** | **White Silkie** |
| | **CCII** | **ccii** |
| **P Gametes** | **CI** | **ci** |
| **F1** | **White** | |
| | **CcIi** | |

The three colored feather are produced because of the presence of the gene **C** and the absence of gene **I**. The other 13 members are white either because of the presence of the gene **I** which masks the effect of **C** or because of the presence of the two genes in a homozygous recessive form (**ccii**).

**Table 3-4:** Punnett's square illustrating the production of 13 white and three colored feather in the $F_2$.

| ♀ \ ♂ | CI | Ci | cI | ci |
|---|---|---|---|---|
| **CI** | **CCII** | **CCIi** | **CcII** | **CcIi** |
| | **White** | **White** | **White** | **White** |
| **Ci** | **CCIi** | **CCii** | **CcIi** | **Ccii** |
| | **White** | **White** | **White** | **White** |
| **cI** | **CcII** | **CcIi** | **ccII** | **ccIi** |
| | **White** | **White** | **Colored** | **Colored** |
| **ci** | **CcIi** | **Ccii** | **ccIi** | **Ccii** |
| | **White** | **White** | **Colored** | **White** |

## b. Modifying (Supplementary) genes

Modifying or supplementary genes are examples of recessive epistasis and is demonstrated by the color phenotypes in the flowers of plants and hairs in animals. In this case a gene controls the color production and the other interacts to modify the color. The cross black × albino mice produced a uniform $F_1$ of agouti, which in certain instances, when inbred results in an $F_2$ of 9 agouti, 3 black and 4 albinos. This result indicates that a 1/16 and 3/16 classes are indistinguishable phenotypically and further indicates that $F_1$ individuals are heterozygous for both genes. It is assumed that one of the two pairs of genes include one allele for color production and anther allele for no color. The other gene includes one allele for agouti and one for black as follows:

| | | | |
|---|---|---|---|
| **A** | **Agouti (full color)** | **C** | **Color production** |
| **a** | **Black color** | **c** | **No color** |

The genotype of a black individual would be **aaCC** and that of an albino would be **AAcc**. A cross between these two individuals will give an agouti individual with a genotype **AaCc** as follows:

| | | |
|---|---|---|
| **P** | **Albino**<br>**AAcc** | **Black**<br>**aaCC** |
| **P Gametes** | **Ac** | **aC** |
| **F1** | **Agouti**<br>**AaCc** | |

Breeding members with this genotype produced $F_2$ progeny in a ratio of 9:3:4 as in table 3-5. The four individuals marked with * are albino because the gene for color inhibition (**cc**) is present in a homozygous form. The three black genotypes are black because the gene for black (**aa**) is a homozygous form. The other nine members of the 16 are agouti because both genes for agouti and color production are present (**A-C-**).

**Table 3-4:** Punnett's square is illustrating the production of nine agouti, four albino and three black mice the $F_2$.

| ♀ \ ♂ | **AC** | **Ac** | **aC** | **ac** |
|---|---|---|---|---|
| **AC** | **AACC**<br>**Agouti** | **AACc**<br>**Agouti** | **AaCC**<br>**Agouti** | **AaCc**<br>**Agouti** |
| **Ac** | **AACc**<br>**Agouti** | **AAcc***<br>**Albino** | **AaCc**<br>**Agouti** | **Aacc***<br>**Albino** |
| **aC** | **AaCC**<br>**Agouti** | **AaCc**<br>**Agouti** | **aaCC**<br>**Black** | **aaCc**<br>**Black** |
| **ac** | **AaCc**<br>**Agouti** | **Aacc***<br>**Albino** | **aaCc**<br>**Black** | **Aacc***<br>**Albino** |

## Molecular basis for modifying genes

A molecular mechanism for recessive epistasis in flower color production may be simplified by postulating that two genes encode enzymes catalyzing successive steps in the synthesis of a blue petal pigment. In this explanation, a wild gene (**$w^+$**) is assumed to produce the colored pigment and a mutant gene ($m^+$) interacts to change the final color of the petal. In a dihybrid heterozygous plant, the blue pigment in the petals is developed when the wild gene and a mutant gene are found together (**$w^+$- $m^+$-**). The presence of the wild gene alone (**$w^+$- /$m^+$-**) results in a pink color and the presence of both genes as a double recessive results in white petals. This epistasis is revealed because the pathway for the production of the colored petals is blocked by a mutant allele of another gene. Hence the mutation in the earlier gene precludes expression of any alleles of a gene acting at the later step (Fig. 3-5).

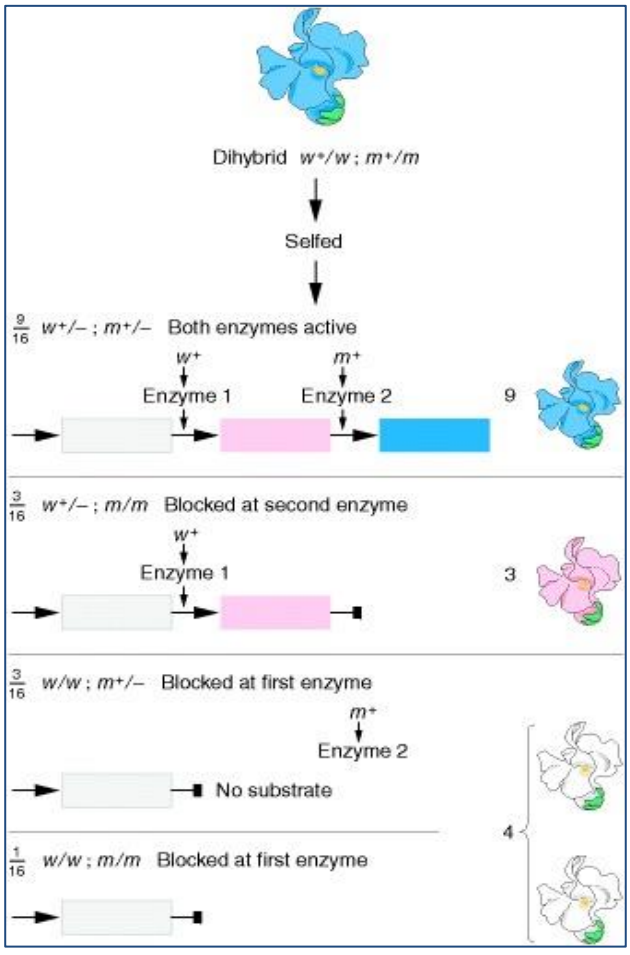



**Figure 3-5:** A diagram illustrating the interaction of a wild and a mutant gene (**w** & **m**) that control the color development in flower petals

## c. Complementary genes

Complementary genes are two genes that interact to produce a dominant phenotype when both are found as dominant alleles, the presence of one or the two genes are found in a homozygous recessive form a recessive phenotype is produced. The common example for this case is again found in the flower color of some plants. When Bateson crossed two white flowered plants he obtained $F_1$ plants which had purple flowers. When the purple-flowered plants were crossed together, plants of the $F_2$ were produced in a phenotypic ration of 9 purple to 7 white-flowered. This 9:7 ratio indicates that there are two pairs of alleles for flower color. Considering that the genotype of one parent is **CCww** and that of the other parent is **ccWW**, the F1 plants had a genotype **CcWw** and they were colored because of the presence of **C-** and **W-** together in a complementary form.

| **P** | **White flowers**<br>**CCww** | **White flowers**<br>**wwCC** |
|---|---|---|
| **P Gametes** | **Cw** | **wC** |
| **F1** | **Colored Flowers**<br>**CcWw** | |

Crossing of this heterozygous $F_1$ produced $F_2$ progeny in a ratio of 9 purple flowered plants to7 white-flowered pants as illustrated in table 3-6. The 7 members marked with * are white-flowered because of the presence of one or both genes in a homogenous recessive form, while the other 9 member are purple-flowered because both genes are present in a dominant form either homozygous (**CCWW**) or heterozygous (**C-W-**).

**Table 3-5:** Punnett's square illustrating the production of nine color flowered plants and seven white flowered ones in the $F_2$.

| ♀ \ ♂ | CW | Cw | cW | cw |
|---|---|---|---|---|
| CW | CCWW<br>Colored | CCWw<br>Colored | CcWW<br>Colored | CcWw<br>Colored |
| Cw | CCWw<br>Colored | CCww*<br>White | CcWw<br>Colored | Ccww*<br>White |
| cW | CcWW<br>Colored | CcWw<br>Colored | ccWW*<br>White | ccWw*<br>White |
| cw | CcWw<br>Colored | Ccww*<br>White | ccWw*<br>White | Ccww*<br>White |

## Molecular basis for complementary genes

An important explanation of complementary genes at the molecular level is the interaction between a regulatory gene (**r**) and the gene that it regulates (**a**). The actions of the two genes show a type of complementation as demonstrated in figure. 3-6. A common situation is that the regulatory gene produces a regulatory protein that binds to the upstream regulatory site of the target gene, possibly facilitating the action of RNA polymerase. In the absence of the regulatory protein, the target gene would be transcribed at very low levels, inadequate for cellular needs. This type of gene interaction can be followed in a situation in which a dihybrid is heterozygous for a mutation of the regulatory gene (***r*$^+$*/r***) and heterozygous for a translation termination mutation (***a*$^+$*/a***). Normally translation termination codons are at the 3′ end of every mRNA, but mutation can introduce a termination codon within the coding sequence. In this position, it results in a short polypeptide. Assume that the mutation in the present example is close to the 5′ end of the protein-coding sequence. The possible genotypes are shown in Fig. 3-6. The stop codon will lead to premature transcriptional termination. In this case, a protein of such small

size that it is negligible. The ***r⁺/r*; *a⁺/a*** dihybrid will give the following progeny.

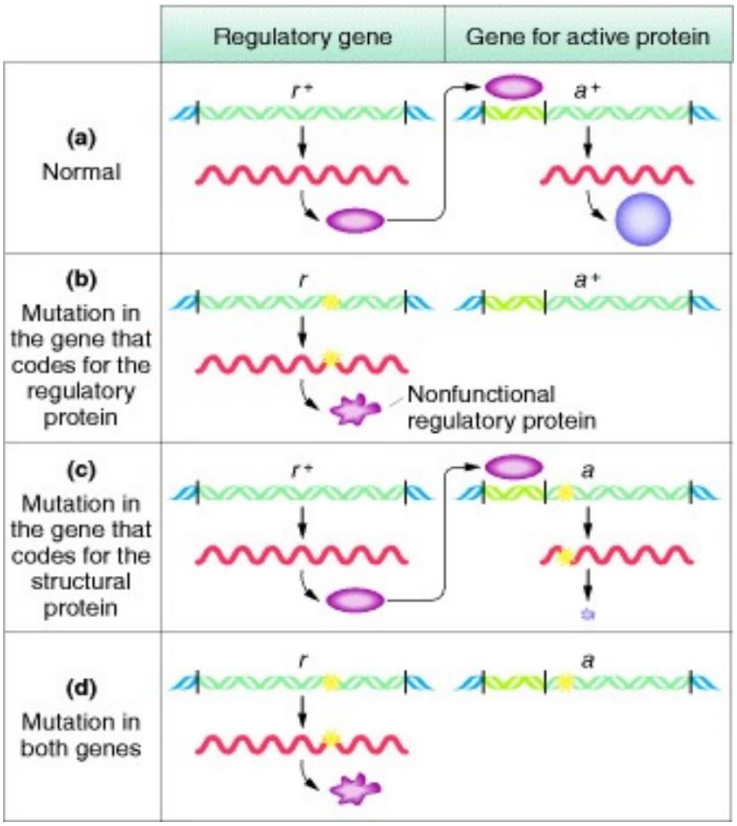



**Figure 3-6:** Diagram illustrating the interaction of two genes, a regulatory gene (**r**) and the gene that it regulates (**a**). The actions of the two genes show a type of complementation as demonstrated in the example of flower color.

## d. Duplicate Genes

Duplicate genes are two genes that interact to produces a dominant phenotype when both or either of them is found as dominant allele in a homozygous or heterozygous genotype, a recessive phenotype is only produced when the two genes are found in a homozygous genotype. The common example of duplicate genes is best presented by fruit shape in the Shepher's purse. There are two species of Shepher's purse (*Capslla bursa pastoris*) that produces heart shaped fruits (triangular capsules) and the other species (*Capslla heegeri*) produces spindle-shaped capsules.

| P | Hear-shaped fruit<br>$A_1A_1A_2A_2$ | Spindle-shaped fruit<br>$a_1a_1a_2a_2$ |
|---|---|---|
| P Gametes | $A_1A_2$ | $a_1a_2$ |
| F1 | Hear-shaped fruit<br>$A_1\ a_1A_2\ a_2$ | |

Crossing the two species produced $F_1$ plants with triangular capsules indicating that triangular shape of capsule is dominant over the spindle shape. Crossing two members of the $F_1$ produced $F_2$ plants with triangular and spindle shaped capsules in a ratio of 15:1 respectively. This indicates that two genes control this character. The genotype of the dominant triangular capsule is **$A_1A_1A_2A_2$** while that of the spindle capsules is **$a_1a_1a_2a_2$,** the genotype of the $F_1$ plants will be **$A_1a_1A_2a_2$**. When plants of the $F_2$ were crossed or self pollinated $F_2$ plants were produced in a ratio of 15 heart shaped fruits to one spindle shaped fruits as illustrated in figure 3-7. This indicates that two genes control this character. The only case of the possible 16 combinations that produced spindle-shaped capsules is produced because both genes are homozygous recessive **$a_1a_1a_2a_2$**. The other 15 plants have triangular capsules because both or one gene are present in a dominant form (Fig. 3-7)

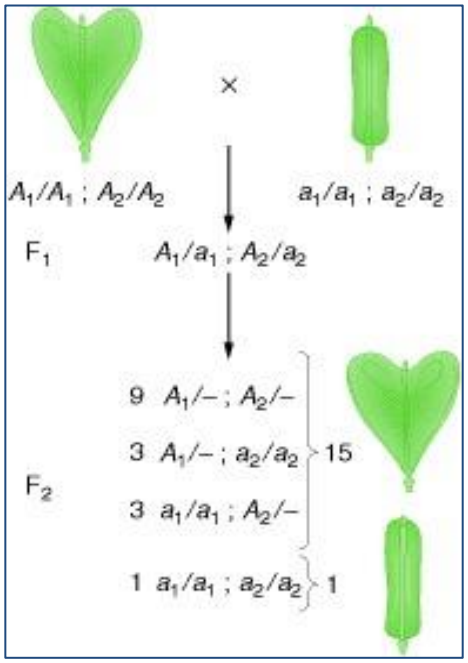



**Figure 3-7:** A diagram illustrating the interaction of two genes duplicate genes that control the fruit shape in the Shepher's purse.

**Chapter 2 Post-test**

**A. Mark the correct statements with the sign (√) and the wrong answers with the sign (x)**

1. Sickle cell anemia is an example of co-dominance and lethal genes in man.
2. The failure to synthesize chlorophyll in maize is caused by lethal genes
3. Yellow coat color in mice is due to dominant lethal genes

**B. Circle the correct answer of the following**

1. Complementary genes modify the classical dihybrid 9:3:3:1 ratio to:-

a. 5:1 - b. 9:7 - c. 13:3 - d. 9:4:3

2. Co-dominance characterizes:

a. Heterozygous genotype

b. Homozygous genotype

c. Dominant phenotype

d. Recessive phenotype

**C. Write a brief essay on each of the following topics:**

**1.** Incomplete dominance

**2.** Epistatic genes

**3.** Molecular basis for complementary genes

**4.** Molecular basis for modifying genes

*Chapter 3*

# NONMENDELIAN INHERITANCE

Since Mendel's time, our knowledge of the mechanisms of genetic inheritance has grown immensely. For instance, it is now understood that inheriting one allele can, at times, increase the chance of inheriting another or can affect how and when a trait is expressed in an individual's phenotype. The simple rules of Mendelian inheritance do not apply in these and other exceptions. The traits that do not fit the Mendelian rules are said to have non-Mendelian inheritance patterns. Examples of traits resulting in the interaction of two alleles of the same gene or of two genes were explained in chapter 3. Patterns associated with multiple alleles, polygenes and related mechanisms of transmission are addressed in this chapter.

## 1-Multiple Alleles

Most traits in living organisms are governed by one gene but many, with these genes often have several allelic variants. Some alleles of the same gene may be dominant over others but neither of the alleles for the same gene can dominant over the others. Multiple alleles pattern of inheritance is exemplified by many traits in animals, humans and plants. Here we give three examples, the coat color of rabbits, The ABO blood type system and the incompatibility factors in plants.

### a. Blood groups in man:

Blood consists of two components, cells (red, white and platelets) and liquid (plasma). Plasma minus the clotting proteins fibrinogen is called serum. In 1910 Karl Landsteiner observed that red blood cells of certain individuals clump together into microscopically visible groups when mixed with the serum of some other persons. This clumping is due to an antigen antibody reaction. Further studies discovered the occurrence of two natural antibodies in the blood serum and two antigens on the

surface of red blood cells. With regard to antigens, an individual may produce either, both or neither, he may produce either, neither or both antibodies. The antigens are designated as **A** and **B** and the corresponding antibodies as an anti A or **α** and anti B or **β**. An individual's blood group is denoted by the type of antigen he produces as in the following Table and illustrated in figure 4-2

| Blood group | Antigen in RBCs | Antibody in serum |
|---|---|---|
| **A** | **A** | **anti (β) $_{\beta}$** |
| **B** | **B** | **anti (α$_{)\ \alpha}$** |
| **AB** | **A and B** | **neither** |
| **O** | **neither** | **α and β** |

The inheritance of blood groups have shown that children produce the **A** antigen only if at least one parent produced it. Similarly the **B** antigen is found in individuals where at least one parent has it. However, individuals having a blood group **O** may be produce by parents of **A** and recessiveness of the gene for group **O**. Marriages of **A** and **B** parents produce, in some cases, children having both **A** and **B** blood groups. Antigens are indicative of co-dominance of the genes for **A** and **B** antigens.

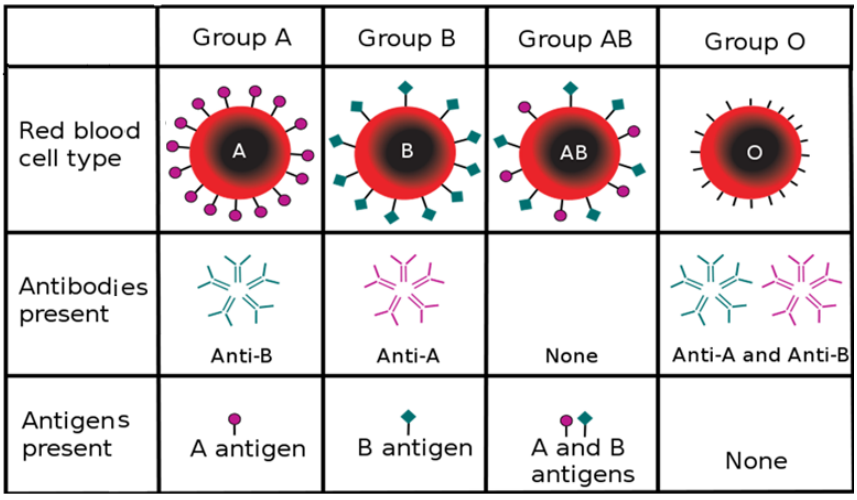



**Figure 4-2:** Diagrammatic representation of the types of antigens and anti-bodies in the red blood cells in humans

The studies on the inheritance of the **ABO** blood type is controlled by a single gene with three alleles (the ABO gene): The three alleles are designated as $I^A$, $I^B$ and i for antigen **A**, antigen **B** and neither antigens respectively. The gene encodes a glycosyl-transferase, an enzyme that

modifies the carbohydrate content of the red blood cell antigens. The gene is located on the long arm (q) of chromosome 9. The $I^A$ allele gives type **A** blood group, $I^B$ gives type **B**, and *i* gives type **O**. As both $I^A$ and $I^B$ are dominant over *i*, only *ii* people have type **O** blood group. Individuals with $I^AI^A$ or $I^Ai$ have type A blood, and individuals with $I^BI^B$ or $I^Bi$ have type **B**. $I^AI^B$ people have both phenotypes because $I^{A1}$ and $I^B$ express a special dominance relationship.

Dominance relationships of these three genes may be as follows ($I^A = I^{B)} > i$. Further atudies have showed that the gene $I^A$ may occur in at least four allelic forms, these are symobolized $I^{A1}$, $I^{A2}$, $I^{A3}$, and $I^{A4}$. $I^{A1}$ is dominant to all other $I^A$ alleles; $I^{A2}$ is recessive to $I^{A1}$ but dominant to the other three and so on. Considering four forms of $I^A$, one of $I^B$ and one of i dominance within the series may be shown as follows:

**$\{(I^{A1} > I^{A2} > I^{A3} > I^{A4}) = I^B\} > i$**

$I^{A4}$ is a very rare allele, and when omitted from the series, the 15 genotypes and 8 phenotypes are produced as follows:

| Genotype | Phenotype | Genotype | Phenotype |
|---|---|---|---|
| $I^{A1}I^{A1}$ | $A_1$ | $I^{A1}I^B$ | $A^1B$ |
| $I^{A1}I^{A2}$ | $A_1$ | $I^{A2}I^B$ | $A^2B$ |
| $I^{A1}I^{A3}$ | $A_1$ | $I^{A3}I^B$ | $A^3B$ |
| $I^{A1}i$ | $A_1$ | $I^BI^B$ | B |
| $I^{A2}I^{A2}$ | $A_1$ | $I^B{}_i$ | B |
| $I^{A2}I^{A3}$ | $A_2$ | ii | O |
| $I^{A2}i$ | $A_2$ | | |
| $I^{A3}I^{A3}$ | $A_3$ | | |
| $I^{A3}i$ | $A_3$ | | |

Study of the blood groups in man is very important in blood transfusion in hospitals. If a blood from a person of group A is transferred into persons of group **A** or **AB** no agglutination occurs because such individuals do not possess the α antibody, whereas if blood of group A was injected to blood of the groups **B** or **O** agglutination will occur because each of these blood contain **α** antibody. Similarity blood of group **B** is agglutinated by blood of the

groups **A** and **O** because the latter bloods contain the antibody **β**. Blood of the group **O** are not agglutinated by any other blood, so can be introduced into persons of **A**, **B**, or **AB** types of blood. Finally bloods of group **AB** are agglutinated by any of the other group and can only be introduced into persons of the same group. For these reasons, group **AB** is known as universal receptor while group **O** is known as universal donor. The possible transfusions of blood groups are summarized in Table 4-1. The sign (+) indicates instances were agglutination takes place and indicates agglutination which means danger to the recipient and the sign (-) indicates possible transfusion with no danger. However, Tests of compatibility with regards to other blood groups are necessary in hospitals before transfusion.

**Table 4-1:** A summary of the possible transfusions in man

| Recipient's blood group | Donor's blood group | | | |
|---|---|---|---|---|
| | A | B | AB | O |
| A | - | + | + | - |
| B | + | - | + | - |
| AB | - | - | - | - |
| O | + | + | + | - |

ABO Blood tests can also be used in cases of illegitimacy and disputed parentage. Tests, however, cannot prove a man to be the father, but they can in some cases show that he cannot be. Using this system the probability of exclusion for a man wrongly accused is about 50%.

The routine method for the determination of the blood group of a person is test his blood cells against two sera, one containing α and the other containing β antibodies at the same time on a glass slide to be examined under the microscope. If no agglutination occurs after mixing the person's blood with both sera, the blood is group O. If agglutination occurs with both sera, the person's blood group is AB. If agglutination occurs with serum α, the blood group is A, while agglutination of the blood with serum β is indication that the blood group is B.

## Pseudoalleles

Research has uncovered evidence that many traits that were though to be governed by multiple alleles are inherited by separate but extremely closely linked genes. These are genes that (i) govern expression of the same trait (ii) found at different, but closely linked chromosome loci, (iii) show a low frequency of recombination (crossing over) and (iv) exhibit the cis-trans position effect. To demonstrate the pseudo-alleles two examples are given, the inheritance of eye color in *Drosophila melanogaster* and the RH blood factor in man.

### a. The Rh blood factor in man

The well-known Rh factor in man was discovered in 1940 by Landsteiner and Wiener. When rabbits were injected with a blood of the rhesus monkey antibodies were formed by the rabbits. When these antibodies were mixed with the blood of the monkey its red blood cells agglutinated. Thus the red cells have on their surface an antigen designated as Rh Tests of human beings showed that most persons produce this antigen. Such persons are designated as $Rh^-$positive or (**$Rh^+$**), the much smaller proportion of persons who do not produce the Rh antigen are termed Rh negative or (**$Rh^-$**).

Earlier evidence indicated that $Rh^+$ is a dominant phenotype and a single pair of alleles were postulated with $Rh^+$ persons having genotype RR or Rr and $Rh^-$ persons having a genotype rr. However, further experiments revealed more Rh antigens; the number is now over 40 and the genetics of its inheritance is much more complex than was originally believed. Two major theories have been developed to explain the complexity or Rh blood groups. The wiener system postulates a series of at least multiple alleles, some are common and others are rare, at a single locus. These alleles are designated as **$R_1$, $R_2$, $R_0$, R', R'', $R_z$, $R_y$**, and **r**. The most common of these alleles is the **$R_1$** for the production of the **Rh** antigen and **r** for no antigen production. Each of the dominant alleles is responsible for production of one or more of the **Rh** antigens.

The other system is postulated by Fisher and is elaborated by Race and Sanger (1968). Fisher proposed a group of at least three very closely linked pseudo-alleles **C-D-E-** situated very close together on the same chromosome. At the first locus we have the gene **C** or gene c; at the second **D** or **d**; at the third **E** or **e.** Here, we also have 8 combinations of the three genes as follows: **CDE, CDe, CdE, Cde, cDE, cDe, cdE** and **cde,** the latter combination (**cde**) produce no Rh antigen, while each of the others is responsible for the production of one or more Rh antigens. The most common combinations of the **C-D-E-** genes are the **cDe** (0.42) and the **CdE** (0.14) for the antigen production and the **cde** (0.39) for no antigen production. A list of these combinations and other common genotypes are given in table 4-2. These genotypes account for 98% of the genetic constitution found in the human genome.Each individual possesses two of these complexes and therefore these are 36 possible combinations.

**Table 4-2:** The most commonly encountered Rh genotypes and genotypes in the human genome as proposed the Fisher-Race pseudoalleles notation and the Wiener multiple alleles notation

| **Rh phenotypes and genotypes in man** | | | |
|---|---|---|---|
| **Phenotype** | **Genotype** | | **Prevalence** |
| | **Fisher-Race notation** | **Wiener notation** | |
| **($Rh^+$)** | **CDe/CDE** | **$R_1R_Z$** | **11.2** |
| | **CDe/cdE** | **$R_1r''$** | **0.1** |
| | **CDe/cDE** | **$R_1R_2$** | **7.8** |
| | **CDe/cDce** | **$R_1R_0$** | **2.1** |
| **($Rh^+$)** | **CDe/CDe** | **$R_1R_1$** | **17.6** |
| | **cDE/cde** | **$R_2r$** | **10.2** |
| **($Rh^+$)** | **cDE/cDE** | **$R_2R_2$** | **1.9** |
| | **CDe/cde** | **$R_0r$** | **32.8** |
| **($Rh^-$)** | **cde/cdE** | **Ryr** | **0.3** |
| | **Cde/cdE** | **r'r''** | **0.3** |
| **($Rh^-$)** | **cde/cde** | **rr** | **15.2** |

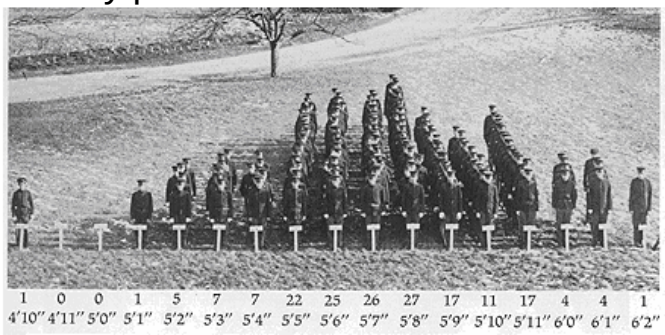
The Rhesus system is of great importance in medicine. The Rh antigens are of far greater importance than any other system in the causation of haemolytic disease of the foetus and newborn. This disorder occurs only

when a number of coincident conditions are met. The mother must be $Rh^-$, the foetus $Rh^+$, therefore marriage of $Rh^-$, women and $Rh^+$ men are involved. There must also be a placental defect whereby foetal blood, whose red cells carry the Rh antigen passes from the embryo into the maternal circulation. This occurs just before or during birth. As a consequence Rh antibody concentration is gradually increased in the mother; she will have been sensitized. In the second pregnancy involving an $Rh^+$ child these antibodies may return to foatus where they destroy the antigen-carrying red cells.

It is now common practice to test pregnant women, particularly in second or later pregnancies to see whether they have developed Rhesus antibodies. Unless the child dies in uterus, an exchange blood transfusion is usually a successful treatment. However, sensitization of Rh- women can now be prevented. In blood transfusion Rh incompatibility is second in importance to the ABO incompatibility as a potential cause of transfusion reactions.

**3- Polygeneic Inheritance:**

Not all inherited traits are expressed in a discontinuous fashion as we have seen in the Mendelian genetics. For example, in man height was recognized as a genetically controlled character as early as 1914 but classifying a random sample of conscripts recruited to join the army according to height, would quickly find that height in man is a trait of continuous phenotypic variation (Figure 4-5). Many other traits are expressed in a similar fashion, including intelligence, skin and eye color in man and crop yield in plants and size in many plants and animals



*Differences in height* of conscripts (From Blakeslee J., *Hered*. 5, 1914.)

**Figure 4-5:** Photograph illustrating the continuous variation in human height in a photograph of conscripts indicating the control of this trait by multiple

In order to explain the continuous variation, our concepts of simple Mendelian inheritance with the reappearance of parental phenotypes as distinct and separate classes in the $F_1$ and $F_2$ must be modified. We are dealing not with just tall versus dwarf but with how tall and how dwarf, that is with continuous characters of degree rather than discontinuous characters of kind. When compared Mendelian inheritance is often termed qualitative while polygenic inheritance is termed quantitative. Quantitative inheritance more often deals with a population in which possible matings occur, and less often with individual matings. Polygenic traits are controlled by multiple loci and called Quantitative trait loci (QTLs). To understand the concept of polygenic inheritance the following examples are presented.

**1- Human eye color:**

Eye color in man depends on the amount of melanin pigment, except for albinos who have eyes without eye pigmentation. Those people with the least pigment have eyes that appear blue; those with the greatest appear brown. However, as you can see in figure 4-6, there is a gradation in eye colors, ranging from the lightest blue to the darkest brown or "black". Although the inheritance of eye color is complex and only incompletely understood, at least nine phenotypic classes may be recognized, and four pairs of genes are operating in a codominant fashion. According to this hypothesis a simplified basis for human eye color would be as detailed in figure 4-6.

| Number of genes for melanin production | Eye color |
|---|---|
| 8 B | Dark brown |
| 7 B | Medium Brown |
| 6 B | Light Brown |
| 5 B | Hazel |
| 4 B | Green |
| 3 B | Gray |
| 2 B | Dark Blue |
| 1 B | Medium Blue |
| 0 B | Light Blue |

**Figure 4-6:** The possible genotypes and corresponding eye color in man based on the hypothesis that four pairs of genes are operating.

## 3- Skin color in man

Skin color in man also depends upon the relative amount of melanin in the skin and is also controlled by a series of polygenes. Attempts to relate frequencies of various degrees of pigmentation to models based on different numbers of pairs of polygenes, show best agreement between observation and theoretical expectation for at least four genes rather than fewer or more. The inheritance of skin color in man was studied by observing children of marriages between members of the Negro race (black) and of the Caucasian (white) race. Such marriages produced children with an intermediate shade of skin color termed mulatto. When two mulattoes resulting from interracial marriages were married, they produced children with a gradation of skin color from the extreme white Caucasian to the black Negro color.

A model based on the contribution of four genes may be similar to that described for eye color. If four contributing genes for pigment production were symbolized by $\mathbf{B_1, B_2, B_3}$ and $\mathbf{B_4}$ each of these genes contributes an additive small effect to color production. The non-contributing alleles are symbolized by $\mathbf{b_1, b_2, b_3}$ and $\mathbf{b_4}$ have no effect on black color production. The Negros have the homozygous genotype for all the four pairs $\mathbf{B_1B_1B_2B_2B_3B_3B_4B_4}$ and the Caucasians have the homozygous genotype $\mathbf{b_1b_1b_2b_2b_3b_3b_4b_4}$ respectively. The mulatto's offspring have the heterozygous genotype $\mathbf{B_1b_1B_2b_2B_3b_3B_4b_4}$**.** In the $F_2$ at least nine phenotypic classes will result with a graduation of color depending on the number of contributing alleles.

The Melanin in skin color is found in two types: pheomelanin (red) and eumelanin (very dark brown). Davenport, in the second decade of the 20$^{th}$ century, was the first to study the genetics of variation in human skin color. He calculated the amount of melanin in mixed marriages using skin reflectance of light and confirmed the genetic bases for its inheritance although the evolutionary causes are not completely certain. According to

more recent scientific studies, natural human skin color diversity within populations is highest in Sub-Saharan African populations, with skin reflectance values ranging from 19 to 46 compared with European and East Asian populations which have skin reflectance values of 62 to 69 and 50 to 59 respectively.

In brief, quantitative traits are controlled by multiple genes, each segregating according to Mendel's laws. The genes controlling the same trait have additive, equal and small effect on the phenotype of the trait. In brief, a quantitative trait is an aggregate of effects from different genes, each of which contributes specific effect that differ in magnitude and biological effect that can also be affected by the environment to varying degrees. In order to explain the continuous variation of quantitative traits, our concept of simple Mendelian inheritance controlled by one or two genes is extended to three genes with the reappearance of parental phenotypes as distinct and separate classes in the $F_1$ and $F_2$. When the number of genes operating on the trait is increased to N a continuous variation in the degree rather than discontinuous variation of kind will be observed as demonstrated in Fig. 4-8.

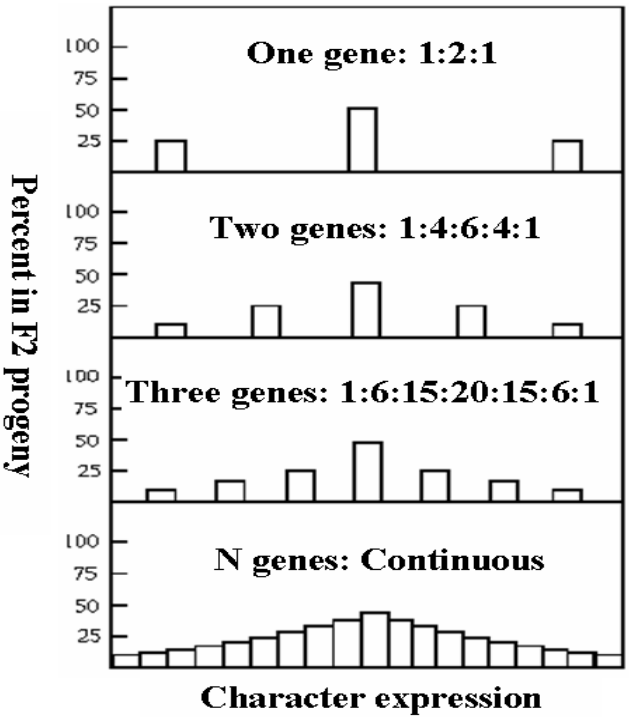



**Figure 4-8:** Diagrammatic representation of the concept of quantitative traits as compared to Mendelian traits.

As the number of genes operating on the trait is increased the continuous variation in the degree rather than discontinuous variation is said to follow the natural distribution curve or a bell shape variation.

Instances of polygenic inheritance are known for other characters in man and for many characters in other organisms. Many human phenotypes such as height, IQ, learning ability, blood pressure and diabitis are quantitative traits. Many important agricultural traits such as crop yield, weight gain in animals, fat content of meat are quantitative traits. In fact, much of the pioneering research into the modes of inheritance of these traits was performed by agricultural geneticists. For example, in tomato crosses between the large-fruited (Golden-Beuty) and a smaller fruited (Red Cherry) produced $F_1$ having fruits intermediate in size between the parental types. The height, yield, seed weight, size of fruits are typical examples of quantitative traits in plants.

**Pleiotropy**

Although genes are identified by main effects, it is possible for some genes to have effects on many separate characters of the individual. This phenomenon is called pleitropy. Pleiotropic gene is best defined as the single gene controlling or influencing a number of possibly unrelated phenotypic traits. This was though to be a rare exception but it is now realized that manifold of genes are more the rule than the exception, and it is believed that most genes are pleiotropic. The manifold effects of genes may be caused by a number of separate actions, or by a single action complicated by developing into a number of apparently separate effects. When a gene has a number of different actions, each causing a distinct effect it is referred to as true pleiotropism, while genes with a single action developing into different effects is termed spurious pleiotropism.

One of the examples illustrating the concept of pleiotropy is the recessive dwarf gene (dw) in the house mouse. In a homozygous form it causes a marked reduction in the body size; the occurrence of this gene is diagnosed

from this effect. Other effects associated with the presence of this gene are that the dwarf mice are sterile, with little or no development of the secondary sex characters and underdeveloped thyroid glands. There are other effects, but these sufficiently illustrate how the manifold effects of the (dw) genes are caused by a simple initial action. An interesting extension of the effects of the dw gene is that the development of the secondary sex structures depends on production of sex hormones by glands (gonads). Secondary sex structures do not develop in dwarf mice because the pituitary gland is defective in producing the gonadotrophic hormone.

In man albinism is pleiotropic trait. The gene for this trait not only results in the typical albino deficiency of skin, hair, and eye pigmentation but also causes defects in vision. However, the sickle cell anaemia is a more common case of pleiotropism. Persons with this disorder suffer from other systems which may lead to their death. The pleiotropic effects of the sickle cell hemoglobin gene in homozygous recessive persons are illustrated in the diagram shown in figure 4-11.

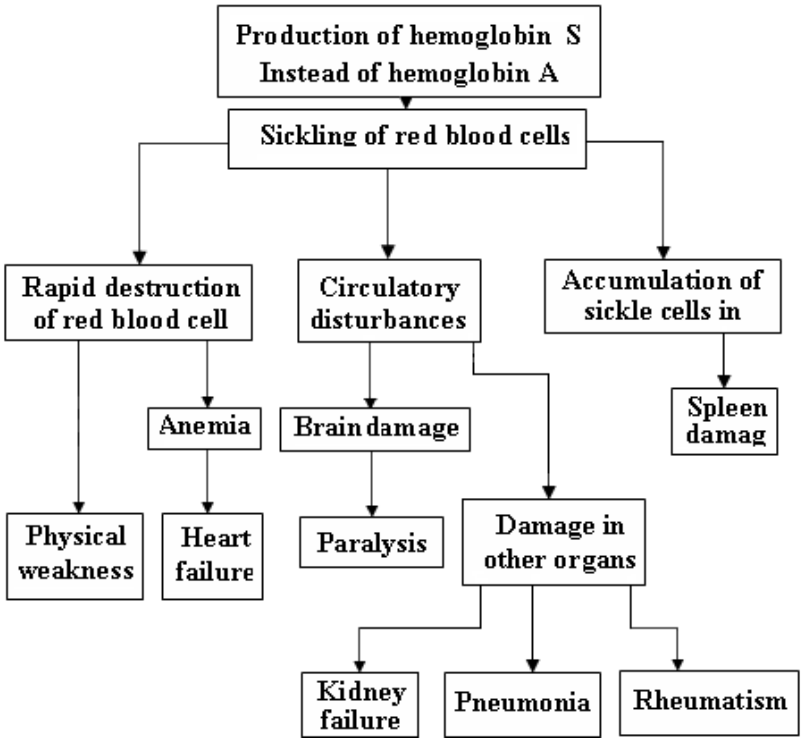



**Figure 4-11**: Diagram illustrating the pleiotropic effects of production of hemoglobin S instead of hemoglobin A and sickling of red blood cells in man.

**Chapter 3 test**

**A- Circle the correct answer of the following:-**

1. The coat color in rabbits is controlled by:

a. Two allele b. Four alleles c. Pseudo-alleles d. Six alleles

2. In human eye color, four pairs of genes are operating in a co-dominant manner resulting in:-

a. Five phenotypic classes b. Seven phenotypic classes

c. Nine phenotypic classes d. 11 phenotypic classes

3. Nilsson-Ehle discevered the qntitative traits by his work on

a. pea b. wheat c. cabbage d. tobacco

on what trait was the experiments of Ehle?

**B- Mark the correct answers with √ and wrong answers with x**

4. The ABO blood type is controlled by a single gene (the ABO gene) with five alleles
5. Females of humans with a blood group O has a homozygous recessive genotype
6. Polygeneic traits are expressed in a continuous fashion not as a discontinuous fashion as in the Mendelian genetics.
7. Dwarfism is a typical example of quantitative characters.
8. Self sterility in plants is the phenomenon in which the pollen grains from a plant fail to bring, about fertilization in the ovules of the same plant
9. Pleiotropic gene is single gene controlling or influencing one single phenotypic traits
10. Quantitative traits are mostly affected by the environmental conditions
11. In plants the failure to produce seeds by the same plant is called self-sterility and is controlled by psedo-alleles

*Chapter 4*

# Cytogenetics

## Mitotic Cell cycle and the DNA C-value

When somatic cells are grown, they establish a repeated pattern of growth and duplication. The mitotic cycle consists of interphase, in which the chromosomes are not visible and the division stage (mitosis). A eukaryotic cell cannot divide into two, the two into four, etc. unless two processes alternate:

1. Doubling of its genome (DNA) in S phase (synthesis phase) of the cell cycle.
2. Halving of that genome during mitosis (M phase).

The period between M and S is called G1; that between S and M is G2. Interphase is divided into three stages; the G1 stage, S- stage and G2-stage (Figure 6-1).

Following mitosis, the daughter cells enter the G1 period and have a DNA content equivalent to 2C (Figure 2-12). All diploid organisms at this stage contain 2C DNA content because the two homologous chromosomes are present as single chromatids. During the S-stage the DNA is doubled and during the G2, cells contain 2 times the amount of DNA present in the original G1 cell (4C). Gametes are haploid and therefore have half the DNA content (1C). Some tissue, like liver, and many plants contains occasional cells that are polyploid and their nuclei have a correspondingly higher DNA. Each species has a characteristic content of DNA in the chromatids which is constant and has thus been called the C-value.

Eukaryotes vary greatly in DNA content but always contain much more DNA than prokaryotes.

## The G1 stage

After mitosis interphase commences with a period referred to as the first gap (G1 phase). During this stage chromosomes are fully extended and the genes are active sending messages for new enzymes to carry out the activities of the cell. G1 phase usually lasts from 10 to 24 hours, but can vary from virtual non-existence to several days. After this period DNA replication begins and the S-period starts.

## The S-stage

This is a period during which DNA replication takes place and the chromosome number is doubled; it lasts 5–10 hours. During S-period not all growth stops because not all genes replicate at the same time. The protein component of the chromosomes is also duplicated so that at the end of the S-stage each chromosome is double i.e., made of two chromatids.

## The G2 stage

This is a second gap period after replication. This stage lasts of about 4 hours and the genes are again fully functional. G2- period is followed by the four mitotic stages; prophase, metaphase, anaphase and telophase; during which duplicated chromosomes condense and the identical halves (sister chromatids) separate equally into two daughter nuclei. Mitosis usually lasts from 1–5 hours.

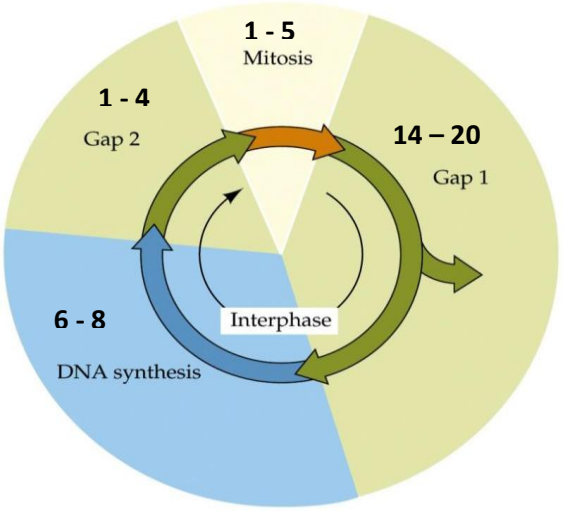



**Figure 6-1**: Diagrammatic representations of cell cycle stages.

## Cell cycle and the DNA C-value

Following mitosis, the daughter cells enter the G1 period and have a DNA content equivalent to 2C (Figure 2-12). All diploid organisms at this stage contain 2C DNA content because the two homologous chromosomes are present as single chromatids. During the S-stage the DNA is doubled and during the G2, cells contain 2 times the amount of DNA present in the original G1 cell (4C). Gametes are haploid and therefore have half the DNA content (1C). Some tissue, like liver, and many plants contains occasional cells that are polyploid and their nuclei have a correspondingly higher DNA. Each species has a characteristic content of DNA in the chromatids, which is constant and has thus been called the C-value. Eukaryotes vary greatly in DNA content but always contain much more DNA than prokaryotes.

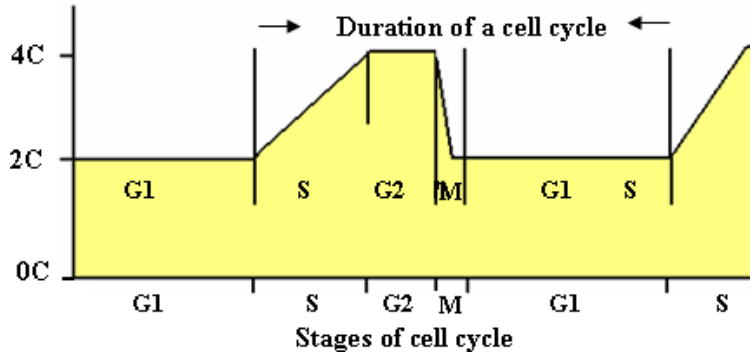



**Figure 6-2:** Diagrammatic representations of the changes in the C value during the stages of the mitotic cell cycle.

## Control of the Cell Cycle

The passage of a cell through the cell cycle is controlled by proteins of three types in the cytoplasm called cyclins, cyclin- dependent kinases (Cdks) and anaphase-promoting complex (APC). Three types of cyclins act as three mitotic cycle check-points; these are:-

- G1 cyclins (D cyclins)
- S-phase cyclins (cyclins E and A)
- mitotic cyclins (B cyclins)

The levels of these cyclins, in the cell, rise and fall with the stages of the cell cycle. Three types of cyclin-dependent kinases (Cdks); act in the control of cell cycle; these are:-

- a G1 Cdk (Cdk4)
- an S-phase Cdk (Cdk2)
- an M-phase Cdk (Cdk1)

The levels of these enzymes in the cell remain fairly stable, but each must bind the appropriate cyclin (whose levels fluctuate) in order to be activated (Figure 6-3). They add phosphate groups to a variety of protein substrates that control processes in the cell cycle. The anaphase-promoting complex (APC) is also called the cyclosome is often designated as the APC/C.) The APC/C has two major functions:-

- Triggers the events leading to destruction of cohesin allowing the sister chromatids to separate.
- Degrades the mitotic (B) cyclins.

**Steps in the cycle**

1. A rising level of G1-cyclins binds to their Cdks and signal the cell to prepare the chromosomes for replication.
2. A rising level of S-phase promoting factor (SPF) - which includes A cyclins bound to Cdk2 - enters the nucleus and prepares the cell to duplicate its DNA (and its centrosomes).
3. As DNA replication continues, cyclin E is destroyed, and the level of mitotic cyclins begins to rise (in G2).
4. Translocation of M-phase promoting factor (the complex of mitotic B cyclins with the M-phase Cdk [Cdk1]) into the nucleus initiates
   a. Assembly of the mitotic spindle
   b. Breakdown of the nuclear envelope cessation of all gene transcription.
   c. Condensation of the chromosomes

The above events take the cell to metaphase of mitosis. At this point, the M-phase promoting factor activates the anaphase-promoting complex

(APC/C) which allows the sister chromatids at the metaphase plate to separate and move to the poles (= anaphase), completing mitosis. Separation of the sister chromatids depends on the breakdown of the cohesin that has been holding them together. It works like this. Cohesin breakdown is caused by a protease called separase (also known as separin) which is kept inactive until late metaphase by an inhibitory chaperone called securin.

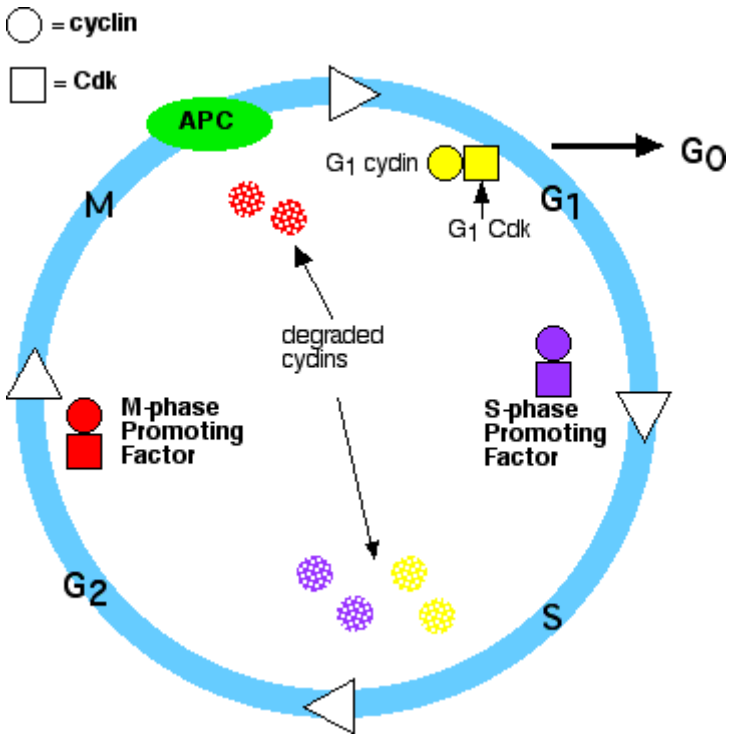



**Figure 6-3:** Diagram illustrating the passage of a cell through the cell cycle is controlled by cyclins and Cdks (http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/C/CellCycle.html

Anaphase begins when the anaphase promoting complex (APC/C) destroys securin (by tagging it with ubiquitin for deposit in a proteasome thus ending its inhibition of separase and allowing separase to break down cohesion and destroys B cyclins. This is also done by attaching them to ubiquitin which targets them for destruction by proteasomes. Separase also turns on synthesis of G1 cyclins (D) for the next turn of the cycle and degrades geminin, a protein that has kept the freshly-synthesized DNA in S phase from being re-replicated before mitosis.

**The G0 point**

A cell may leave the cell cycle, temporarily or permanently often at G1 and enters a stage designated G0 (G zero). A G0 cell is often called "quiescent", but that is probably more a reflection of the interests of the scientists studying the cell cycle than the cell itself. Many G0 cells are anything but quiescent. They are busy carrying out their functions in the organism. e.g. secretion, attacking pathogens. Often G0 cells are terminally differentiated: they will never reenter the cell cycle but instead will carry out their function in the organism until they die. For other cells, G0 can be followed by reentry into the cell cycle. Most of the lymphocytes in human blood are in G0. However, with proper stimulation, such as encountering the appropriate antigen, they can be stimulated to reenter the cell cycle (at G1) and proceeds on to new rounds of alternating S phases and mitosis. The G0 represents not simply the absence of signals for mitosis but an active repression of the genes needed for mitosis. Cancer cells cannot enter G0 and are destined to repeat the cell cycle indefinitely.

## Checkpoints controlling cell cycle

The cell has several systems for interrupting the cell cycle if something goes wrong; these are:-

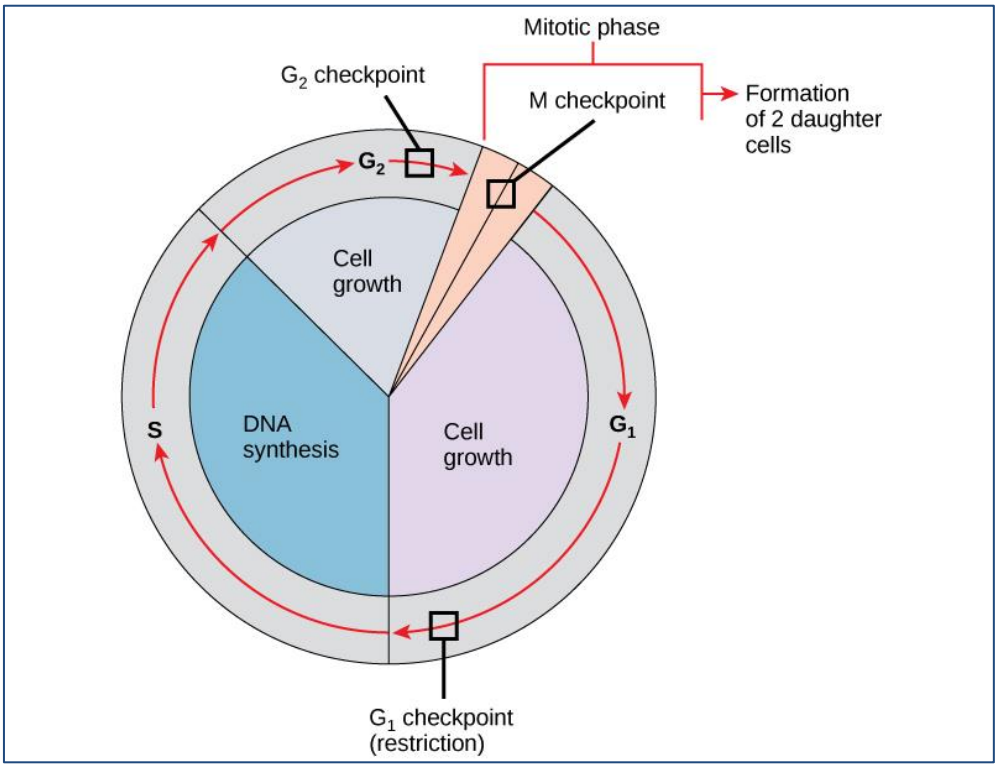
- **DNA damage checkpoints.** These sense DNA damage both before the cell enters S phase (a G1 checkpoint) as well as after S phase (a G2 checkpoint).
    - Damage to DNA before the cell enters S phase inhibits the action of Cdk2 thus stopping the progression of the cell cycle until the damage can be repaired. If the damage is so severe that it cannot be repaired, the cell self-destructs by apoptosis.
    - Damage to DNA after S phase (the G2 checkpoint), inhibits the action of Cdk1 thus preventing the cell from proceeding from G2 to mitosis.

If replication stops at any point on the DNA, progress through the cell cycle is halted until the problem is solved.

- **Spindle checkpoints.** Some of these have been discovered, their functions are: -

  - Detect any failure of spindle fibers to attach to kinetochores and arrest the cell in metaphase until all the kinetochores are attached correctly (M checkpoint)
  - Detect improper alignment of the spindle itself and block cytokinesis.
  - Trigger apoptosis if the damage is irreparable.

All the checkpoints examined require the services of a complex of proteins. Mutations in the genes encoding some of these have been associated with cancer; that is, they are oncogenes. This should not be surprising since checkpoint failures allow the cell to continue dividing despite damage to its integrity.



# The chromosomes

In a diploid cell chromosome occur as pairs of homologous chromosome or homologs. One member of each homologous pair is inherited from either the male or female parent; these homologs pair in bivalents in the meiotic prophase. During cell division, the chromatin of the nucleus condenses into individual chromosomes. The dividing chromosomes are best studied at metaphase as each chromosome appears as two chromatids, two arms, centromere and two telomeres (Fig. 7-1). The chromatids appear to be made of coiled loops of 20-30 nm. However, it is evident from the picture of chromosomes that it is difficult to appreciate details of chromosome structure with the electron microscope.

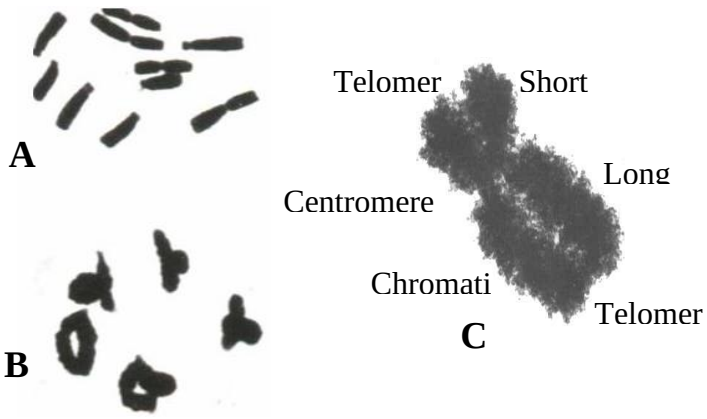



**Figure 7-1:** Photographs illustrating that chromosomes are found in homologous pairs in somatic cells and in bivalent at meiosis I and that chromosomes are formed of two chromatids, two arms, centromere and two telomeres.

The number of chromosomes is the most stable feature of the species; all individuals of the same species usually have the same number of chromosomes

## The Chromosome's organization

The word chromosome is derived from the two Greek words “chromos” meaning color and “soma” which means body, i.e. the chromosome is a colored body. The chromosomes are only seen during cell division as a result of coiling up the chromatic material in the nucleus. They are

colorless in natural living cells. They can however, be seen after fixing the cells followed by staining with dyes specific to the chromatin. The chromatin can be isolated biochemically by purifying nuclei and then lysing them in a hypotonic solution. When prepared in this way, chromatin appears as a viscous, gelatinous substance that contains DNA, RNA, basic proteins called histones, and non-histone (more acidic) proteins. Histones are small proteins that are basic because they have a high content (10-20%) of the basic amino acids arginine and lysine. Being basic, histones bind tightly to DNA, which is acid.

The chromosomes play an important role in heredity, since it carries the genes. It is established that genes are arranged in a linear form in the chromosomes, each chromosome carry thousands of genes. For example, a typical human body cell contains about 35,000 genes in the 46 chromosomes, if extended, the DNA in the human cell is about 2.8 cm long whereas the chromosomes are about 150 μm long and the question is how this long stands of DNA are packaged in the chromosomes.

**DNA Packaging in the Chromosome**

The DNA strands and the histone proteins exist as a fibril of repeating units called nucleosomes. Duplicates of four main histone proteinss, H2A, H2B, H3 and H4, which are very similar in all living eukaryotes and are among the most conserved proteins, are associated in octamers that contain two of each of them to form the nucleosome core. A piece of DNA that is about 140 bp is coiled on the outside of the ocatamere histones. Another histone called H1 is not conserved between species and has tissue-specific forms, links adjacent nucleosomes to each other and is associated with about 80 bp of DNA. Thus, every nucleosome involves about 200 bp of DNA.

The nucleosome fibril is folded in a specific manner to form a solenoid that is about 30 nm wide, this solenoid is folded further to form a 300 nm filament that is coiled again to form the 700 nm, which is about the width of

the chromatid (Fig. 7-2). Geneticists consider the chromatid as the fundamental unit of the chromosome during cell division. In addition to the basic histones, the chromosomes include acidic proteins, which are found in a rather loose association with the DNA and are involved in the maintenance of a higher-order folding of the chromatin and in the forming of the scaffold of the chromosome that hold sister chromatids together.

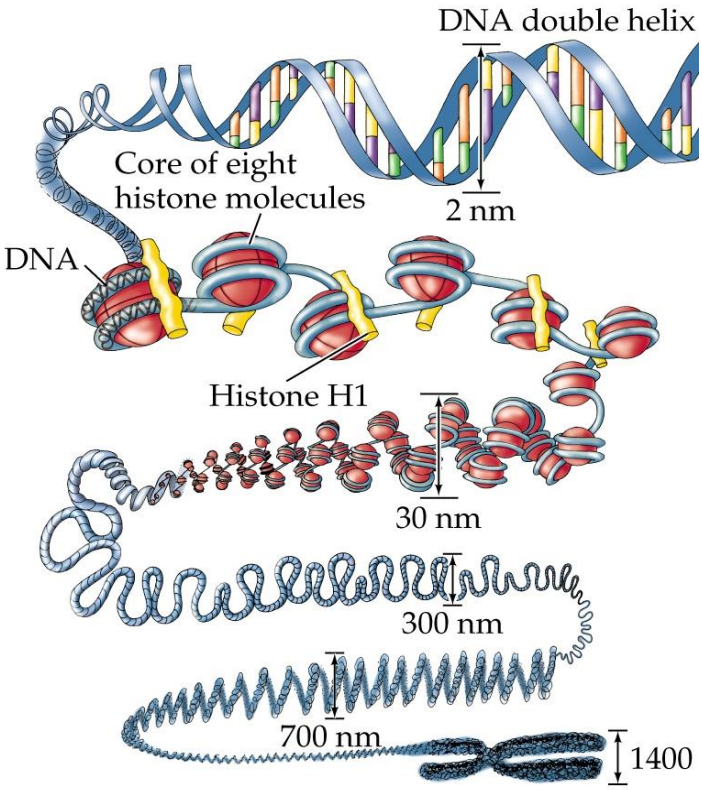



**Fig. 7-2:** The nucleosome model for the packaging

## The Karyotype

The chromosome constitution of an individual is called the karyotype and can be analyzed in dividing cells arrested during mitotic metaphase, by some mitotic inhibitors such as colchicine. The chromosomes are arranged in homologous pairs in order of their length to construct the karyotype. A metaphase chromosome spread, and the karyotype of a human female are presented in Fig. 3-16. The pairs of autosomal chromosomes are arranged in the karyotype from the longest, #1, to the shortest, #22. The sex chromosomes are often placed to the right of the smallest autosomal chromosomes. Chromosomes vary in shape depending upon the position of the centromere. The short arm of the chromosomes is labeled p (for

petite) and a long arm labeled q (the next letter after p). If the centromere is close to the end, the chromosome is acrocentric, and the very short arm consists of a stalk and a knob (satellite).

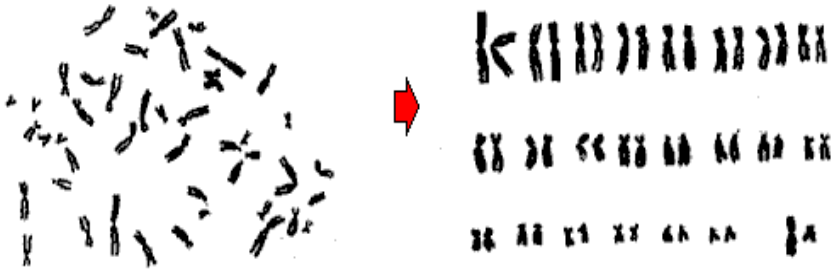


**Fig. 7-3:** The chromosomes of a human female stained with Feulgen reagent as seen under light microscope to the left and arranged in a karyotype to the right

Levan *et al*. (1965) proposed a nomenclature system for chromosome shape based on the position of the centromere (Table 3-1). According to this system, different major types of chromosomes are recognized (Fig. 7-4).

**a- Metacentric or median centromere**: The centromere occurs in the center of the chromosome dividing it into two equal arms. The chromosome is U-shaped at anaphase.

**b- Submetcentric or submedian centromere**: The centromere occurs near the center and thus forming two unequal arms. With the percentage between arms not exceeding 1:7. Such chromosome appears as either J or L shaped at anaphase.

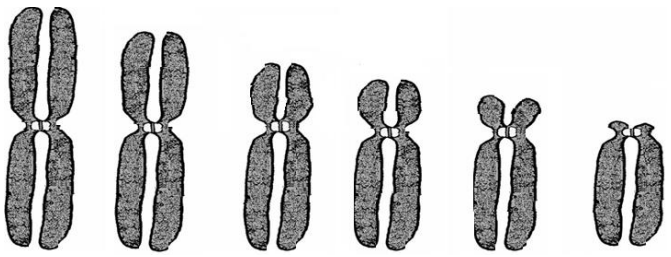
**c- Acrocentric or subterminal centromere**: The centromere is present near the one end of the chromosome so that one arm is very small and the other is much longer.

**d- Telocentric centromeret**: The centromere is almost present in the terminal end of the chromosome. The chromosome appears rod shaped at anaphase.

**Table 7-1:** The terms proposed by Levan et al. (1964) for chromosome types based on the arm ratio and centromere position

| Symbol | Centromere position | Arm ratio | Chromosome type |
|---|---|---|---|
| **M** | **Median point** | **1.0** | **Metacentric** |
| **m** | **Median region** | **1.0 – 1.7** | **Metacentric** |
| **sm** | **Submedian region** | **1.7 – 3.0** | **Submetacentric** |
| **st** | **Subterminal region** | **3.0 – 7.0** | **Subtelocentric** |
| **t** | **Terminal region** | **> 7.0** | **Acrocentric** |
| **T** | **Terminal point** | | **Telocentric** |

The position of the centromere is always constant for the same chromosome and for all cells of the same organism. Consequently, the centromere is considered as one of the features, which distinguishes the chromosomes of a species.



**M m sm st t T**

**Figure 7-4:** Diagrammatic illustration of the major chromosome types, based on arm ratio and centromere as defined by Levan et al., 1964

## Chromosome Banding

Chromosome banding, which produce characteristic staining pattern for each chromosome pair make it possible to identify each chromosome in the karyotype and to identify structural rearrangements with much greater accuracy. A band is defined as a part of a chromosome that is clearly distinguished from adjacent parts by virtue of a lighter or darker staining intensity. Bands that stain light by one method may stain dark by another. The chromosomes are visualized as consisting of a continuous series of light and dark bands so that by definition, there are no interbands.

The banding techniques fall into two fundamental groups, those resulting

in bands distribution along the length of whole chromosome, such as Q-, G and R bands, which demonstrate patterns of DNA synthesis, and those that stain a restricted number of specific bands or structures, and include methods, which reveal centromeric (constitutive hetercromatin) bands (C-bands), telomeric bands (T- bands), and nucleolus organizing regions (NOR-bands). Various causes have been reported for occurrence of chromosome bands; these include the presence of repetitive DNA, differences in the composition of DNA, the protein components, and the degree of DNA packing in the chromosomes.

**Genetic Importance of Chromosomal Banding**

**1**. Identifying different chromosome in the same species, which facilitates matching of homologous chromosomes and are useful in karyotype analysis.

**2.** Chromosomes banding allows cytogeneticists to discover chromosomal abnormalities (aberrations) very accurately, by comparing the bands in the normal chromosomes with those appearing in the aberrant chromosomes.

**3.** The analysis of extended prometaphase or prophase chromosomes was improved to obtain more bands. This was achieved by using synchronization of the cell culture by thymidine or a methopherin and using G- or R-banding methods. The use of such techniques makes the localization of breakpoint and analysis of abnormal chromosomes much easier.

**4.** The study of R-banded chromosomes are particularly important for the analysis of some chromosomal abnormalities in man, where phenotype malformations are often associated with alteration of R- banded chromosomal regions.

**5.** The similarities and differences in banding patterns of chromosomes of different species make it possible to follow up the changes in species related together on the evolutionary scale of taxonomic groups.

# Chromosomal changes

Natural chromosomal changes are important for the induction of variation in living organisms and essential for the evolution to proceed. Chromosomal changes can also be induced artificially by physical, chemical or biological agents. Induced changes in chromosomes are regarded as chromosomal mutations and are divided into two types; changes in number and changes in structure. Very few mutations which cause visible changes in the autosomes are compatible with life. Conor and Ferguson-Smith made the analogy that if the length of the human haploid genome was drawn stretching from London to New York, the smallest visible deletion (about 4Mb) would represent about 8 km gap and that on this scale the average gene would be about 30 m long. So, even the smallest gap will usually contain many genes.

In humans, about 20% of conceptions have some sort of chromosomal disorder but because of the lethal effects of such disorders, the number actually born is only about 0.6%. However, 60 % of spontaneous abortions are due to chromosomal changes in the foetus, about 15% of them are due to polyploidy (triploidy, tetraploidy) and 40% to aneuploidy (Trisomy, 2*n*=47 and monosomy, 2*n*=45 X; Table 3-2).

**Table 3-2:** The frequency of chromosome abnormalities in early spontaneous abortions.

| **Abnormality type** | **Frequency** |
|---|---|
| Triploidy, 2*n*=69 | 10% |
| Tetraploidy, 2*n*=92 | 5% |
| Trisomy, 2*n*=47; e.g. Down's syndrome | 30% |
| Trisomy, 2*n*=45-X; e.g. Turner syndrome | 10% |
| Others | 5% |
| Total | 60% |

## Numerical changes in chromosomes

Chromosome numbers, which deviate from the normal number of chromosomes for a species, are said to be heteroploid. They result from disturbance in the distribution of chromosomes or chromatids at mitotic or meiotic cell division and are of two types: euploidy.and aneuploidy.

## Euploidy

Euploidy is the category of chromosome changes, which involve the addition or loss of complete sets of chromosomes, it is divided into the following classes. Multiples of the haploid number are referred to as polyploids or may be specifically designated triploid (3n), tetraploid (4n), pentaploid (5n), hexaploid (6n) and so on. Polyploidy had taken place in over 40% of plants but is a very rare event in animals and in man. Numerical deviation in animals is regarded as mutations that may lead to damaging occurrences.

Two kinds of polyploids are generally recognized, differing in their origin, frequently and behavior. These are autopolyploids having three or more homologous genomes and allopolyploids, which are derived from two or more non- homologous genomes. However, there are many conditions where plants are autopolyploid for some chromosomes and allopolyploid for other chromosomes thus constituting a range of autoallopolyploids or segmental polyploids.

## Autopolyploidy

Autopolyploidy usually results from the occurrence of karyokinesis (nuclear division) and failure of cytokinesis (cell division) simultaneously. This may be influenced by a number of mechanisms, each of, which enables the number of chromosomes in a cell to be doubled. One mechanism is the C-metaphase in which the spindle apparatus fails to function and thus the duplicated chromosomes remain in the same cell. In an alternative mechanism known as endomitosis, chromosome replication, condensation and chromatid separation occur within the nuclear envelope

with no involvement of any spindle apparatus. This absence of spindle formation is also characteristic of a process called endo-reduplication in which chromosomes go through a replication cycle without contracting or undergoing chromatid separation.

If a diploid species has its two similar genomes designated AA, and then an autotriploid will be AAA and an autotetraploid AAAA. The autotetraploids are derived directly from diploids by doubling of the chromosome number either by somatic doubling or by the union of two diploid (unreduced gametes). Whereas, autotriploids could arise as an offspring of a tetraploid and diploid parents or from two diploid parents by the union of reduced and unreduced gametes.

Cytologically, autopolyploids are characterized and identified by the presence of multivalents formed at metaphase I of meiosis. Autotriploids form trivalents and autotetraploids form quadrivalents. Autopolyploids are, in general, larger than their related diploids as a result of an increase in cell size. Consequently there is, in general, an increase in size of various plant parts like stomata and pollen (Fig. 3-23), a delay in growth rate and flowering. Beyond the tetraploid level however, increase in chromosome number often result in abnormalities such as dwarfing, wrinkled foliage and weak plants.

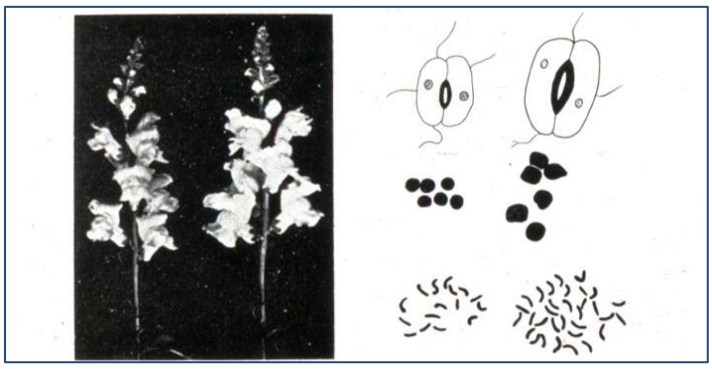


Fig. 3-23: Inflorescence, stomata, pollen and chromosomes of diploid (left) and tetraploid (right) snapdragon (*Antirrhinum majus*)

## Allopolyploidy

Allopolyploids arise by the doubling of the chromosome number of a hybrid between two different species. If the genome of one species is designed as A and the genome of the other species is designated B and if

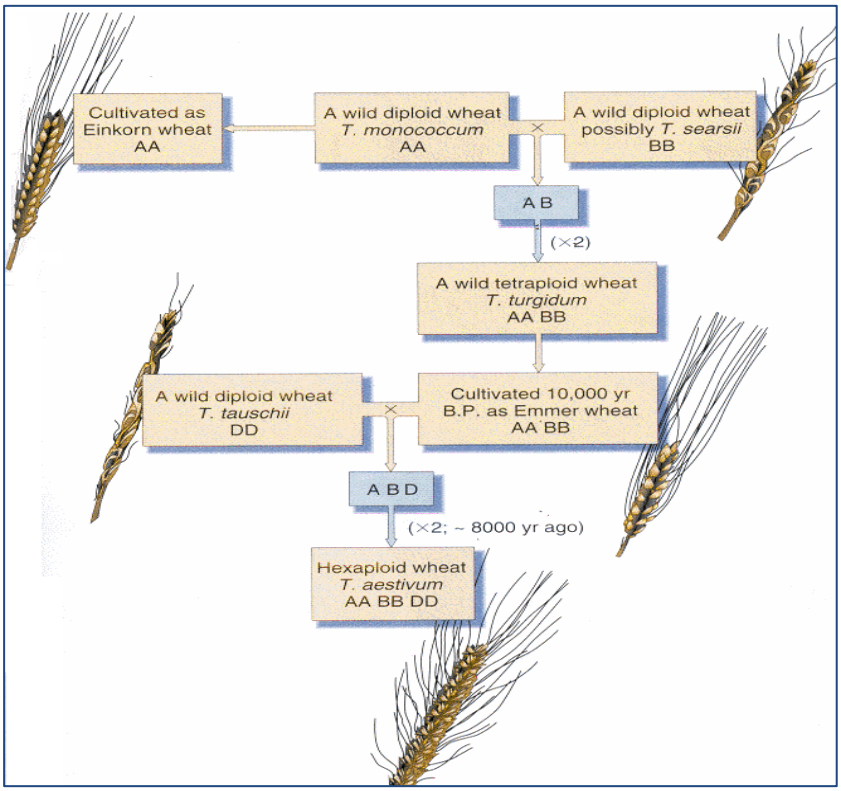
the two genomes are sufficiently dissimilar structurally, no synapsis (pairing of homologous chromosomes at meiosis) will occur in the diploid hybrid. Consequently the chromosomes will be segregated randomly leading to a high degree of sterility. Doubling of chromosome number in such a hybrid will give rise to an allotetraploid with the genomic structure AABB. In this allotertraploid regular synapsis and segregation takes place. Genome A will pair with genome A and genome B with the other B genome. The gametes will then have a genome constitution AB and if the sterility in the F1 hybrid is due only to irregular chromosome distribution, then the allotetraploid can be expected to have a high degree of fertility.

One famous example of allopolyploids is the intergeneric hybrid of the cross between *Raphanas* (radish) and *Brassica* (cabbage) made by the Russian cytogeneticist Karpechenko in the 1930s of the twentieth century. Both plants have a diploid chromosome number of 2n=18. In the F1 the nine haploid *Raphanas* chromosomes were distinctly different from the nine *Brasica* chromosomes and no pairing had occurred leading to a high degree of sterility. However, a number of tetraploid plants were recovered in the F2. Meiosis in these plants was normal and the plants were similar to a normal diploid in many respects and had the origin been unknown it might have been considered a distinct species.

Allopolyploids of more complex nature can arise and probably have been arisen many times in the evolution of plant kingdom. For example the common wheat (*Triticum aestivum*) is a hexaploid (2*n*=42) of three different genomes A, B and D. The evolutionary history of the wheat is demonstrated in Fig. 3-24. In wheat, only bivalents are formed at metaphase I giving gametes with 21 chromosomes suggesting that the three genomes are dissimilar. However, it has been shown that loss of a particular gene cause multivalent formation and consequently leads to irregular distribution of chromosomes at meiosis.

The autopolyploids are extremes of a spectrum in which there is a range from homology between the sets of chromosomes (auto-) to a complete

lack of homology (allo-) between two kinds of genomes; these are known as segmental allopolyploids. The segmentals fall between the autopolyploids and allopolyploids, and it is believed that they are more common in nature than either of the two extreme types.




**Aneuploidy**

Aneuploidy describes changes in chromosome numbers, which are not multiples of the haploid number. Changes resulting from a decrease of one chromosome (i.e. 2*n*-1) are known as monosomic while increase of one chromosome (i.e. 2*n*+1) is called trisomic. In monosomics one member of a pair of chromosomes is missing from diploid set. In trisomics one extra chromosome, which is usually homologous with one pair of chromosomes in the diploid set, is present in the chromosome set. Cells lacking one pair of chromosomes are termed nullisomics while cells having one extra pair are termed tetrasomics. Trisomics and monosomics are of common occurrence in man and cause serious genetic disorders.

In general aneuploid cells result from non-disjunction that is a term used to describe the failure of paired chromosomes or sister chromatids to pass

to opposite poles of the spindle of meiosis or mitosis. Occasionally aneuploidy results from a lagging chromosome that moves noticeably slower to the spindle poles and subsequently fails to be included in the nucleus of either of the two daughter cells. In mitosis this will result in one daughter cell being a normal diploid and the other monosomic. An increase or a decrease of one or a few chromosomes is responsible for the changes in the basic chromosome number, a process that has played a major role in karyotype evolution.

Aneuploidy may only involve gain or loss of centromeres without significant change in the total amount of genetic material. Gain of centromere results by misdivision of the centromere (centric fission) of a metacentric chromosome that gives two telocentric chromosomes. Such telocentrics may be abnormal, commonly showing non-disjunction and misdivision, when they are univalent. Meanwhile, two telocentric chromosomes can give rise to a metacentric isochromosome both telocentrics and the derived isochromosomes are often irregularly distributed at mitosis and meiosis.

However, stable telocentric chromosomes do occur in some plants, for example in *Vicia faba*, *Campanula persicifolia, Nigella doerfleri, Tradescantia micrantha* and a number of species in the genus *Zebrina.* Cases demonstrating this view are found in subgenus Molium of the genus *Allium*. Plants in this group grow basically in the Mediterranean and North America. Plants from North America e.g. *A. cernuum* and *A. acuminatum* and the majority of the species in Europe and the e.g. *A. moly, A. subhirsutum* (Fig. 3-25) and *A. trifoliatum* are characterized by a basic number of $x$=7 with a symmetric karyotype. Other Mediterranean species e.g. *A. roseum* and *A. erdelii* have $2n$=16 and a basic number of $x$=8 that includes two telocentric chromosomes (Fig. 3-25). It is assumed that the karyotype with $x$=8 is derived from $x$=7 through centromere misdivision.

**Fig. 3-25:** The karyotype of *Allium subhirsutum* with *x*=7 (above) and of *Allium erdelii* with *x*=8 including two telocentric chromosomes.

Gain of a centromere is achieved by a reverse process of centromere misdivision known as centric fusion. In this process the centromeres of two telocentric chromosomes are fused to give one metacentric chromosome. It is regarded as a type of translocation known as Robertsonian translocation that is one of the major chromosome structural changes.

## Changes in Chromosome Structure

Structural rearrangements in chromosomes are the results of spontaneous or artificial chromosome breakage with the subsequent reunion of broken ends in such away that the linear order of genes is disturbed. Under normal conditions such breakage is a rare event but a wide range of physical, chemical or biological agents may increase its frequency. Certain chromosomes and chromosome segments are more likely to break than others, for example centromeric and heterochromatic regions are particularly susceptible to breakage. Breaks can occur at any stage of chromosome cycle, those occurring during the G1-phase give rise to chromosome breaks, while those occurring during the S phase or G2 phase give rise to chromatid breaks. Structural changes in the chromosomes depend on behavior of the broken ends. Such changes are often classified into four principal types, deficiencies (deletions), duplications (additions), inversions and translocations (Fig.3-29).

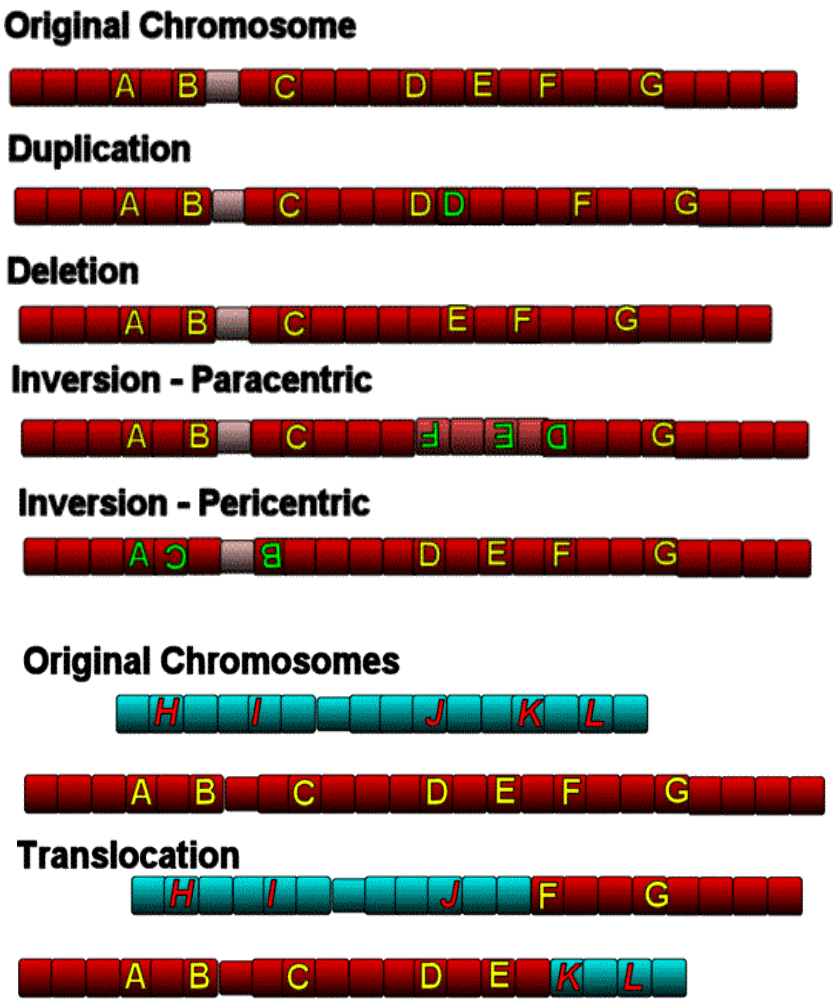



**Fig. 3-29:** The four basic types of chromosomal structural changes: deficiencies (deletions), duplications (additions), inversions (paracentric and pericentric) and translocations

## Deficiencies (deletions)

Breakage of a chromosome segment without reunion leads to a deficient chromosome. The deleted segment may be terminal or interstitial, although most deletions are likely to be the result of the loss of interstitial segment. The cytological detection of deficiency in possible if the deleted segment is long enough to be observed by conventional or banding methods (Fig. 3-30). The close pairing of chromosomes during meiotic prophase can also be useful in heterozygous deficiencies to show the location of the missing segment.

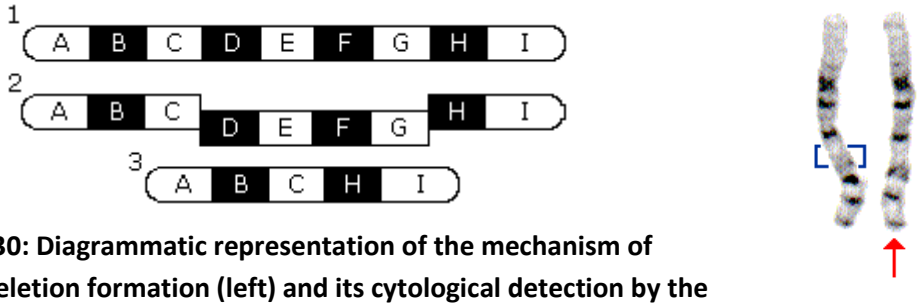



**Fig. 3-30: Diagrammatic representation of the mechanism of deletion formation (left) and its cytological detection by the C-banding method (right)**

## Ring chromosomes and Isochromosomes

A deletion resulting from two breaks in both arms of a chromosome followed by loss of the terminal segments and reunion of the two broken ends give rise to a ring chromosome (Fig. 3-31). Another type of deletion occurs when the centromere divides transversely during mitosis or meiosis followed by the reunion of the centromeric elements of the sister chromatids. This split the chromatids so that two chromosomes are formed each includes duplicates of one end and deletions of the other end these are known as isochromosomes (Fig. 3-31) which are metacentric chromosomes in which the two arms are structurally and genetically identical.

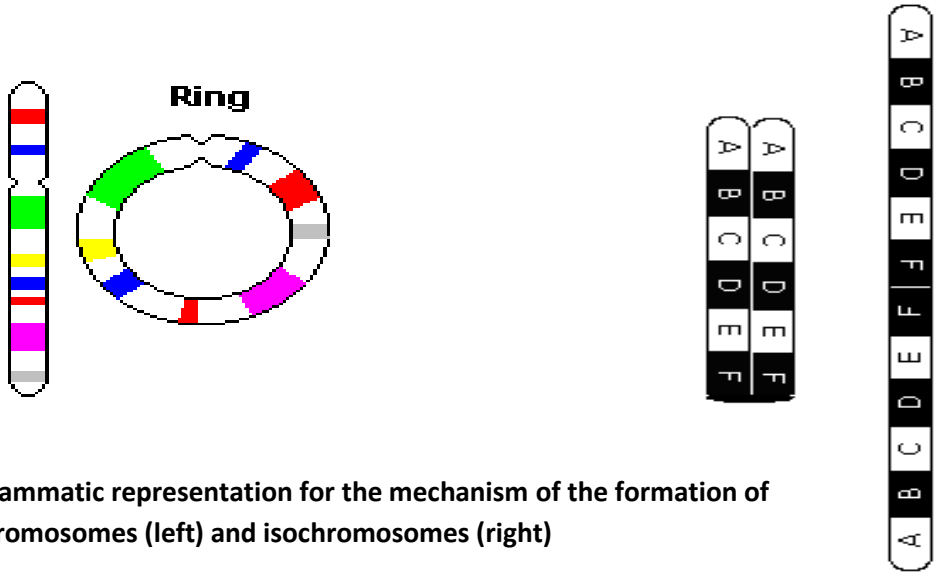



**Fig. 3-31: Diagrammatic representation for the mechanism of the formation of ring chromosomes (left) and isochromosomes (right)**

The absence of the genes lost with the deleted segment is observed depending upon the importance of the deleted region. A large deletion will disturb the balance of genes and cause very harmful phenotypic effects. A very small deletion may not cause serious upset in gene balance, but it can allow recessive genes in the homologous chromosome to be expressed,

because the part of the chromosome with the dominant allele has been deleted. This phenomenon is known as pseudodominance.

In humans, interstitial deletion is the $15q^-$ deletions remove many genes including the two responsible for the two syndrome Wilms tumour, Aniridia, ambiguous Genitalia and mental Retardation. Terminal deletions have only one breakpoint, they extend to the telomere. In humans, the deletion of a small part of the distal region of the short arm of chromosome 5 gives birth to a baby with the Cri du chat (Cry of the cat) syndrome. Babies with this syndrome have a combination of symptoms, which include pinched facial features, mental retardation and developmental delay. The characteristic cry may be separated genetically from the facial dysmorphology and developmental delay, since a small terminal deletion may have the cry only, whereas a larger deletion extending further towards the centromere will include the genes whose hemizygosity is responsible for the other symptoms.

**Duplications (additions)**

The insertion of a broken chromosome segment into a chromosome result in some genes being represented twice (duplicated). Duplications, like deletions can only be observed if the chromosome segment involved is large enough to be detected in Feulgen-stained karyotypes but C-bands can easily demonstrate the presence of duplications (Fig. 3-32). The origin of duplication was most clearly demonstrated by studies on the bar-eye phenotype in *Drosophila*. Individuals with normal eyes had a chromosome segment(s) with 6 transverse bands visible on the X chromosome in the salivary glands, while flies with narrow bar eyes showed the segment S to be duplicated.

Duplications are less deleterious than deletions and may be expected to occur widely in natural populations**.** Pseudoalleles are considered to result from duplication. Pseudoalleles are genes that are found at different, but closely linked chromosome loci, govern different expressions of the same

trait, show low frequency of crossing over and exhibit the cis trans position effect. Duplication thus involves a mechanism by which new genes can arise, and following subsequent mutation enlarge the range of possible functions available to the organism. This has the advantage that the normally functional original segment protects new variability in the duplicated segment.

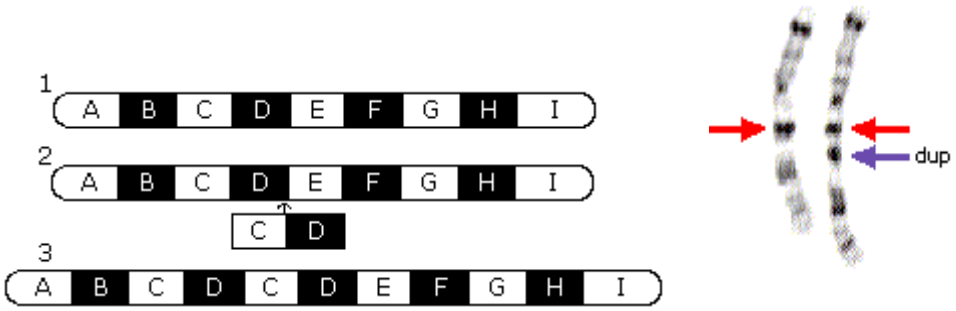



Fig. 3-32: Diagrammatic representation of the mechanism of duplication formation (left) and its cytological detection by the C-banding method (right)

## Inversions

In an inversion, a piece of chromosome is lifted out, turned arround and reinserted (Fig. 3-33). If this includes the centromere then the inversion is termed pericentric. If it excludes the centromere then it is a paracentric inversion. The two have slightly different genetic consequences. 1% of the UK population is heterozygous for a pericentric inversion of chromosome 9. This is absolutely without genetic consequences.

At meiosis, homozygous inversions show normal pairing but inversions heterozygous from a loop at pachytene as homologous chromosomes pair to bring together the homologous segments. The subsequent configuration of the chromosomes depends upon whether or not crossing over takes place within the inverted segment. If no crossing over occurs the chromosomes segregate normally later in meiosis, but when crossing over occurs the resultant chromosomes may have some genes duplicated and other deficient and consequently some gametes will be inviable.

Although the two types of inversion have the disadvantage of reducing fertility they have a role as crossing over suppressors, which serve to

preserve the linkage between gene loci within the inverted segment. However, inversion in one part of a chromosome may increase the level of crossing over in other parts of the chromosome or elsewhere in the karyotype.

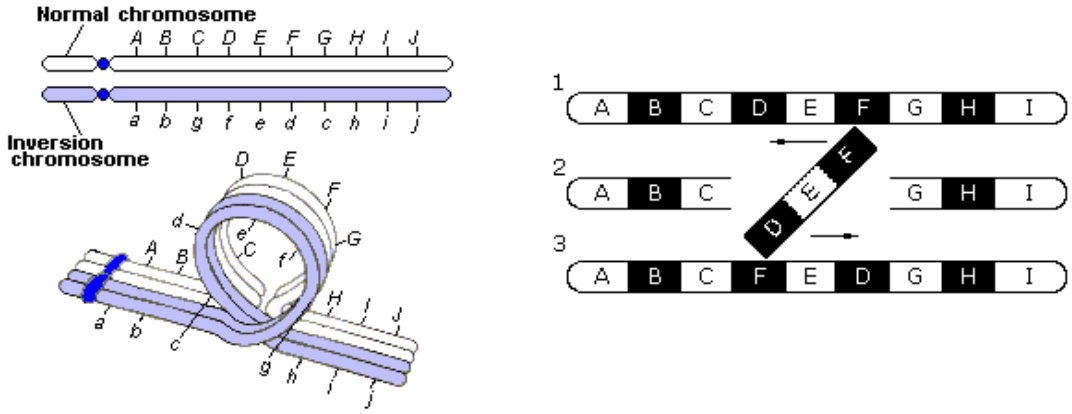



**Fig. 3-33:** Diagrammatic representation of the mechanism for the induction of inversion n (left) and the pairing of a normal and inversion chromosome at the pachytene of meiosis (right)

## Translocations

Translocations occur when a chromosome segment is transferred from its usual position to a new position in the same or a different chromosome. When there is a mutual exchange of segments between two chromosomes the resulting rearrangement is described as a reciprocal translocation or interchange. Non-reciprocal translocation is sometimes referred to as insertion or transposition. A shift is another type of translocation, in which a segment of chromosome is inserted into different part of the same chromosome (Fig.3-34). Translocations may be detected in somatic cells by simple karyotype analysis of Feulgen's stained chromosome spreads and by chromosome banding as revealed by G-banding in animal and human chromosomes (Fig. 3-34) and Q-banding and C-banding in plant chromosomes.

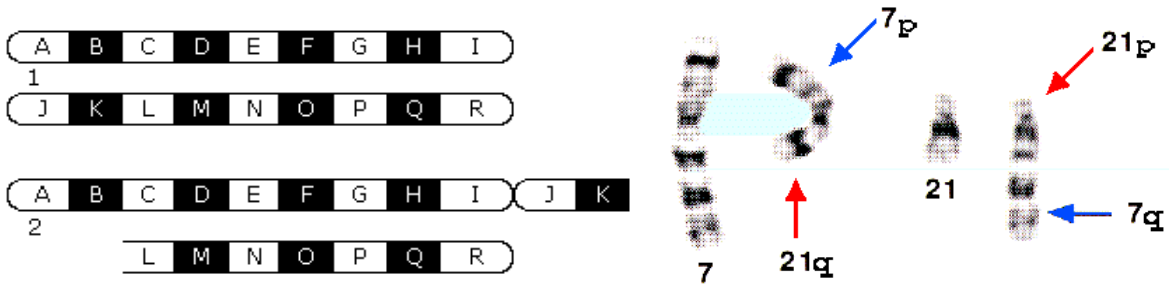



**Fig. 3-34: Diagrammatic representation of a non reciprocal translocation (right) and detection of a reciprocal translocation between chromosome 22 and chromosome 7 in man**

When no genes are lost the translocation is said to be balanced translocation, and usually have no phenotypic effects. Translocations that involve loss of genes on the other hand are called unbalanced tanslocations. In a balanced translocation, two chromosomes have been broken and rejoined in the wrong combination (Fig. 3-35). The figure shows a translocation between the imaginary chromosomes "M" and "N". Balanced reciprocal translocation is unlikely to have any severe consequence for the cell because, even if one of the breakpoints lies within a gene, most mutations are recessive. A complete pairing in the meiotic prophase can only be achieved by the formation of a cross-shaped quadrivalent. More complicated situations occur when crossing over occurs between adjacent points of interchange resulting in reduced fertility and low level of recombination.

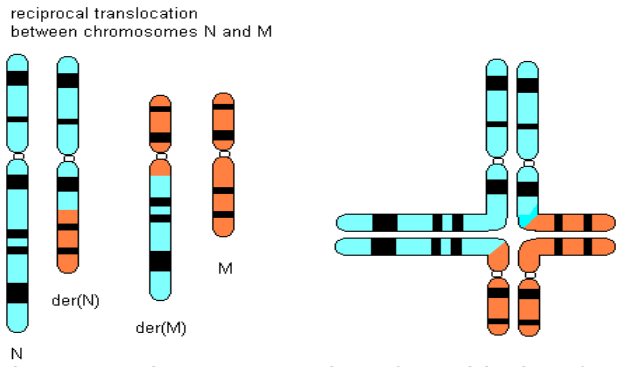


Fig. 3-35: Diagrammatic representation of quadrivalent formation

A translocation may arise spontaneously and is also likely to arise as an offspring of a balanced carrier. There are likely to be symptoms which may be severe. Their exact nature will be unpredictable. Such translocation chromosomes are extremely useful to science in helping to pinpoint genes responsible for the conditions expressed by their bearers. Observations of translocations, in the field of human cytogenetics, revealed that a deleted chromosome 22, which is associated with the chronic myelogenous leukemia, is part of an interchange between chromosome 22 and chromosome 9. The deleted 22 chromosome is called Philadelphian chromosome because it was first discovered in a person from Philadelphia in the USA. The way this interchange occurs is illustrated in Fig. 3-36.

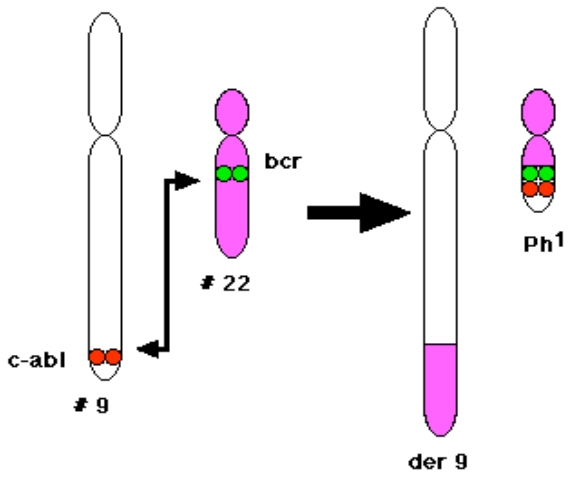



Fig. 3-36: Diagrammatic representation of interchange between chromosome 22 and chromosome 9 leading to the formation of ph1chromosome associated with chronic myelogenous leukemia

## Robertsonian translocations

Robertsonian translocation or centric fusion occurs between two acrocentricor telocentric chromosomes. The breaks occur at or very close to the centromere and the broken ends rejoin to form one metacentric chromosome. The acentric fragments are usually lost resulting in a reduction in chromosome number (Fig. 3-37). The loss of the short arms does not matter because the chromosomes mostly have copies of the rRNA gene blocks that are also often present in the telomeres of the chromosomes. Robertsonian translocations are of significance in the evolution of karyotypes, as they may lead to reduction in basic chromosome number.

In the same way as balanced reciprocal translocation carriers have difficulties at meiosis in ensuring the correct segregation of the chromosomes to make a balanced set, so too do Robertsonian translocation carriers. In their case the chromosomes pair in meiotic prophase to form a trivalent and balanced gametes will only be formed when the translocation chromosome goes to the opposite pole to both of the normal chromosomes. Robersonian carriers therefore suffer a similar reduction to their fertility as do carriers of reciprocal translocations and couples in which one of the partners has a translocation may have a number of early spontaneous abortions.

Robertsonian translocations involving chromosome 21 in man possess a special problem of their own, one of the possible unbalanced gametes will contain effectively two copies of chromosome 21. They are thus at risk of producing a baby with Down syndrome.

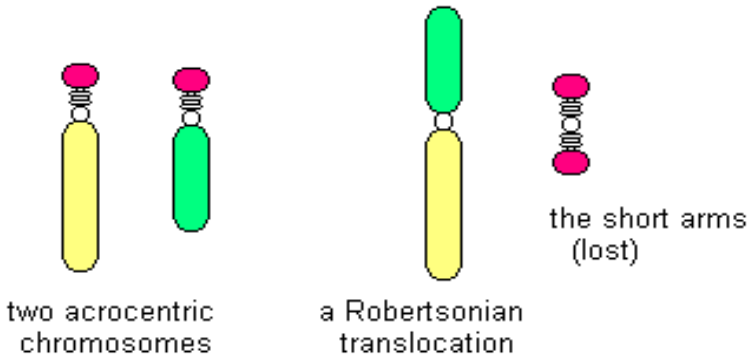



**Fig. 3-37: Diagrammatic representation of Robertsonian translocation**

*Chapter 5*

# Principles of Molecular Genetics

## Identification of DNA as the genetic material

The genes, which constitute the link between generations, must contain some form of information which, when passed to a new generation, influences the form and characteristics of the offspring. This is called the genetic information. This information also, in some way, directs the many complex processes leading to adult form and function. Early in the 20$^{th}$ Morgan and his coworkers demonstrated that genes are inherited in the chromosomes but until 1944 it was not clear what chemical component of the cell constitutes the genetic material, which in turn contains the genetic information.

In the 1940s and 1950s it was clearly demonstrated that DNA is the genetic material. It is now believed that DNA and genes are one and the same. The genetic material has several functions, which are only found in the DNA molecule; these are: DNA must carry information from parent cell to daughter cell. It must contain information for replicating itself. It must be chemically stable, relatively unchanging. However, it must be capable of mutational change. Without mutations there would be no process of evolution.

## Evidence that DNA is the genetic material

DNA was first isolated from fish sperm and the pus of open wounds by the German scientist Friedrich Meischer in 1869. Since it came from nuclei, Meischer named it nuclein. Subsequently the name was changed to nucleic acid and lastly to deoxyribonucleic acid (DNA). Robert Feulgen, in 1914, discovered that the Fuchsin dye, when prepared in a colorless liquid dye, stains DNA in the nucleus of eukaryotic cells. This finding demonstrated that like genes DNA is located in the chromosomes. The evidences that proved that DNA is the genetic material came from four different sources;

bacterial transformation, phage infection of bacterial cells, transduction of bacterial genes from one strain to another and DNA measurements in eukaryotic cells.

### A - Evidence from bacterial transformation

The first evidence that DNA is the genetic material came from the work of Frederick Griffith on the pneumonia causing bacterium *Streptococcus peumoniae*, He distinguished two strains of the pneumococcus: one caused pneumonia (virulent) and had smooth colony outlines on the agar and its cells are surrounded by a polysaccharide capsule, which he designated the S strain and a strain that did not cause pneumonia, had rough colony outlines and lack the capsule, which he designated as the R strain. Griffith (1928) injected the different strains of bacteria into mice; the S strain killed the mice; the R strain did not. He further noted that if heat killed S strain was injected into a mouse, it did not cause pneumonia. When he combined heat-killed S with Live R and injected the mixture into a mouse that the mouse developed pneumonia and died. Bacteria recovered from the mouse had a capsule and killed other mice when injected into them. Griffith concluded that a transforming principle in the killed pathogenic S strain transforms the R strain to the pathogenic S strain. In 1944, Oswald Avery, Colin MacLeod, and Maclyn McCarty revisited Griffith's experiment and concluded the transforming factor was DNA. They fractionated the killed S strain into DNA, RNA, proteins, carbohydrates, lipids mixed each component with the R strain and injected the mixture into a mouse. Only live R strains mixed with DNA were able to transform nonpathogenic into pathogenic strain; DNA treated DNase failed to tansform the R strain into S strain. Avery and his coworkers concluded that DNA should be considered the favorite candidate as the genetic material. A Diagrammatic representation summarizing the experiments of transformation is shown in Figure 7-1.

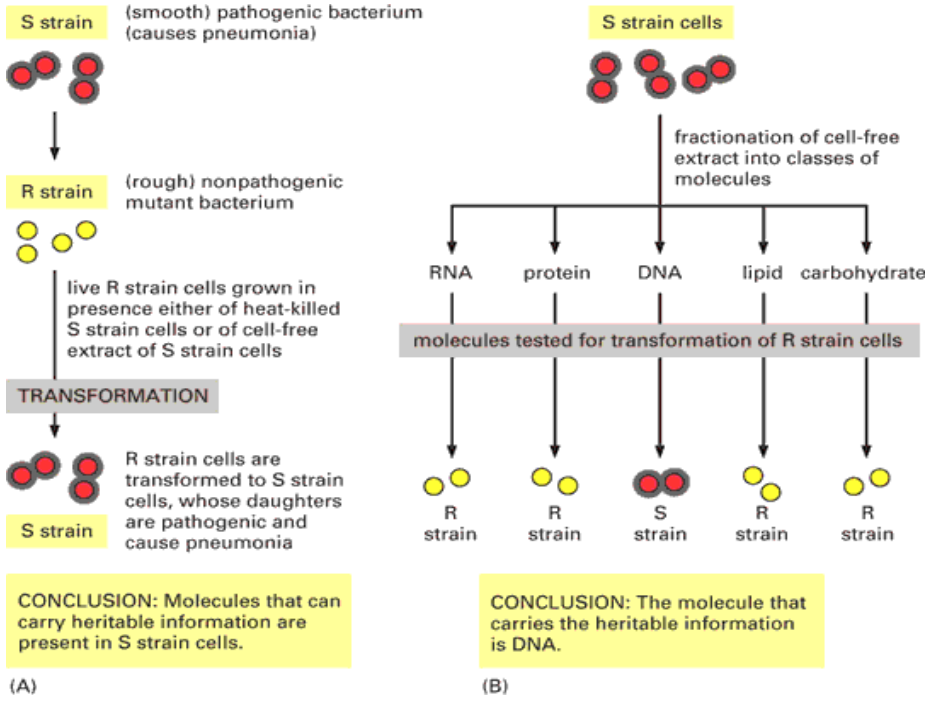



**Figure 7-1:** Summary of the Griffith's transformation experiments (A) and of the Avery, McLeod and McCarty's Experiments (B)

## B- Evidence from phage infection of bacterial cells

The breakthrough in the quest to determine the genetic material came from the work of Max Delbruck and Salvador Luria in the 1940s on the mode of virus infection of bacterial cells. The viruses that they worked with were types of the human intestines bacterium *Escherichia coli* known as the $T_2$ phages. Delbruck and Luria found that these phages have a head and tail shape and consists of protein coats covering DNA in the head; the tail is linked to the head via a neck and a collar and ended a base plate and tail fibers (Figure 7-2).

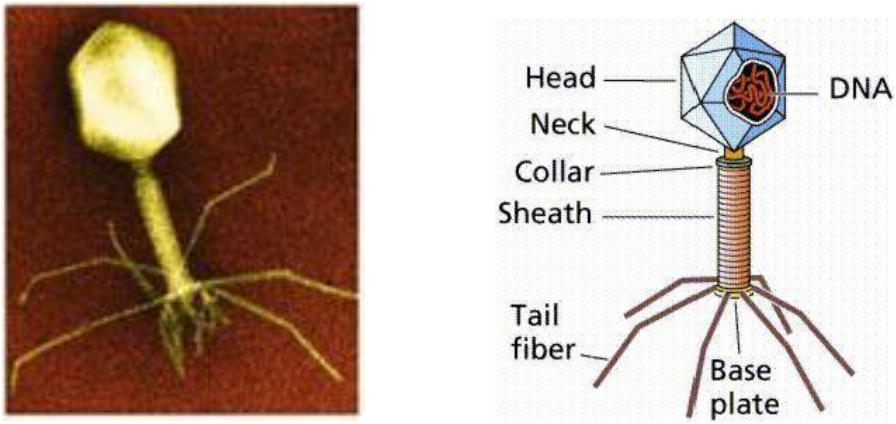



**Figure 7-2:** Photomicrograph (left) and diagram (right) of the $T_2$ bacteriophage.

The phages infect the bacterial cells by injecting DNA into the host cell. This phage DNA then "disappears" while taking over the bacterial machinery and beginning to make new virus instead of new bacteria. After 25 minutes the host cell bursts, releasing hundreds of new bacteriophage. Phages have DNA and protein, making them ideal to resolve the nature of the hereditary material (Figure 7-3).

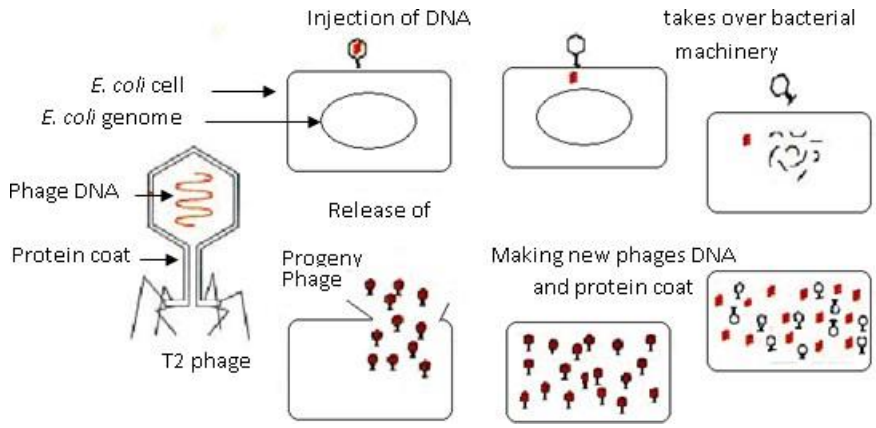



**Figure 7-3:** Diagrammatic representation of the $T_2$ bacteriophages life cycle

In 1952, Alfred Hershey and Martha Chase conducted a series of experiments to determine whether protein or DNA was the genetic material. They proposed that by labeling the DNA and protein with different (and mutually exclusive) radioisotopes, they would be able to determine which chemical (DNA or protein) was getting into the bacterial cells. Such material must be the hereditary material (Griffith's transforming agent). Since DNA

contains Phosphorous (P) but no Sulfur (S), they tagged the DNA with radioactive Phosphorous-32. ($^{32}P$). conversely, protein lacks P but does have S, thus it could be tagged with radioactive Sulfur-35 ($^{35}S$). Hershey and Chase found that the radioactive $^{35}S$ remained outside the cell while the radioactive $^{32}P$ was found inside the cell, indicating that DNA was the physical material of genes. A summary of Hershey and Martha experiments is given in Figure 7-4.

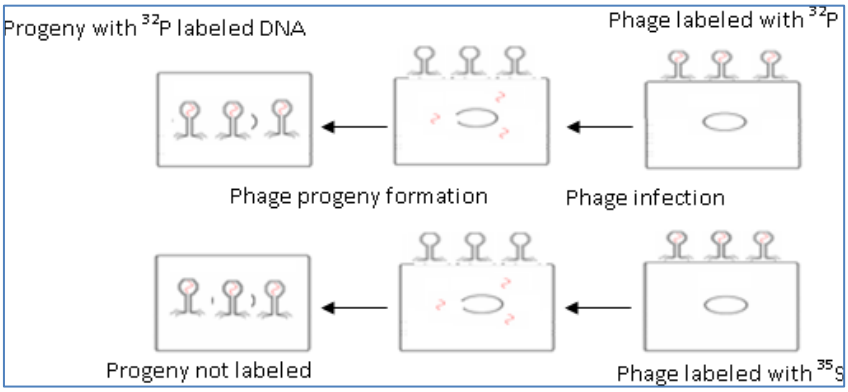



**Figure 7-4:** Diagrammatic representation of the Hershey and Chase experiment

## C - Evidence from Transduction

In 1952, Zinder and Lederberg performed a series of experiments on the typhoid bacterium *Salmonella typhimurium*; the results of which elucidated the phenomenon of transduction. They used two strains of the *Salmonella*; one is prtotrophic for the amino acid methionine and auxotrophic for the amino acid thrionine that may be labeled $meth^{+}thr^{-}$ and the other is prototrophic for the amino acid thrionine and auxotrophic for the amino acid methionine and may be labeled $met^{-}thr^{+}$. Zender and Lederberg found that when cultures of the two strains are mixed and re-cultured on medium lacking the two amino acids, an prtotrophic strain was produced. In another experiment they used $met^{+}thr^{-}$ strain that was infected by the lysogenic virus $P_{22}$, centrifuged the culture to throw down the living cells and the supernated fluid was heated. When this fluid was mixed with $met^{-}$ $thr^{+}$ cells

prtotrophic cells for methionine and thrionine were produced. From these results Zinder and Lederberg suspected that a gene for prototrophy is transferred in the supernatant with the viral DNA. The $P_{22}$ phage is a lysogenic phage that integrates its DNA, as a prophage, in the bacterial genome and remains in this temperate state (Figure 7-5).

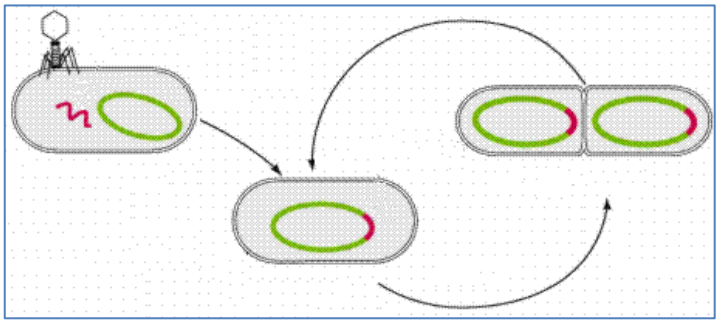


**Figure 7-5:** Diagrammatic representation of the lysogenic cycle of the $P_{22}$ phage

To prove this hypothesis, Zinder and Lederberg grew $meth^{+}thr^{-}$ strain that was infected with $P_{22}$ prophage in one arm of a U-shape tube called Davis tube and a non-infected $meth^{-}thr^{+}$ strain in the other arm of the Davis tube (Figure 7-6). The two arms of the tube are separated by a bacterial filter that allows the passage of phage DNA but does not allow bacterial cells to pass. After hours of bacterial growth prototrophic bacteria was found in the two arms of the tube. The results of Zinder and Lederberg experiments demonstrated that DNA is transferred from one strain of *Salmonella* to another strain with the transfer of the $P_{22}$ phage DNA. The phenomenon of transduction is not restricted to the phage $P_{22}$ and the *Salmonella* bacterium that was used in the used in the experiments of Zinder and Lederberg, but is a general phenomenon that plays important roles in the field of applied microbiology.

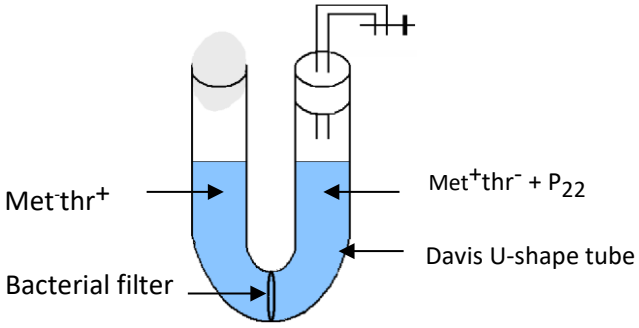



**Figure 7-6:** Diagram illustrating Zinder and Lederberg Davis U tube experiment demonstrating that the gene of methionine prototrophy in *Salmonella* is transferred with the phage $P_{22}$ DNA to an auxotrophic strain

## D - Evidence from DNA measurements

Measurements of DNA amount, in eukaryotic nuclei, using microspectrophotometric methods, which determine the amount of light absorbed by the nucleus, when stained with Feulgen's reagent gives an estimate of the amount of DNA in the nucleus. Such measurements have indicated that:

1. Cells of the same organism contain the same amount of DNA.
2. Somatic cells contain twice as much DNA as gametic cells; this is compatible with the segregation of genes and chromosomes during meiosis.
3. The amount of DNA is congruent with the ploidy level i.e. number of sets of chromosomes in polyploid species.
4. DNA does not undergo turnover during cell metabolism; all other molecules are made and degraded in the cells.

## RNA is the genetic material in some viruses

A group of viruses contains no DNA; among the best known examples are the tobacco mosaic virus and a number of animal viruses known as retroviruses, which include polio. These viruses contain only protein and RNA. Frankel Conrat and his coworkers separated the TMV RNA from TMV proteins. By mixing the two components, they were able to

reconstitute fully infectious TMV. They also mixed RNA from a strain of TMV with proteins of another strain. The reconstituted virus reflected the composition of the parental strain that had contributed the RNA not the one that had contributed the protein. The experimental work of Conrat and his co-workers further demonstrated that the RNA of the TMV determined the amino acid composition of the protein coat. Their experiments demonstrated that where DNA is absent, RNA serves as the genetic material. Further studies on the structure of TMV showed that RNA is found as a spiral shape in the center of the virus surrounded by layers of protein coat (Fig. 7-7).

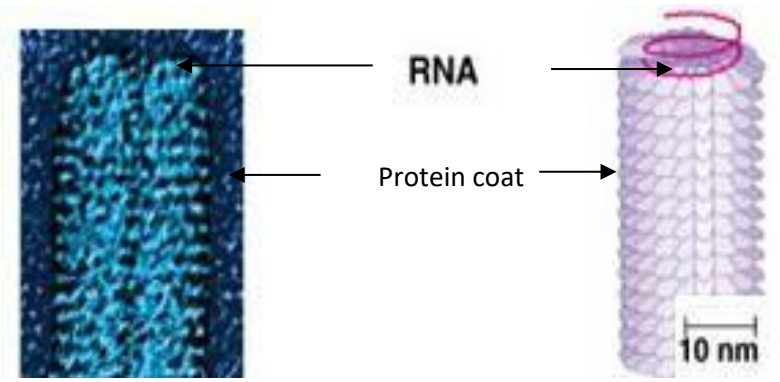



**Figure 7-7:** Photomicrograph of the TMV (left) and a diagram of its structure (right)

## DNA Structure

DNA is known to contain purine and pyrimidine nitrogenous bases, deoxyribose, and phosphate group linked together in nucleotides, which represent the building blocks of DNA. The purines are comprised of Guanine (G) and Adenine (A) and the pyrimidines comprise Cytosine (C) and Thymine (T). The chemical structure of these bases and of a guanine nucleotide is shown in Figure 7-8.

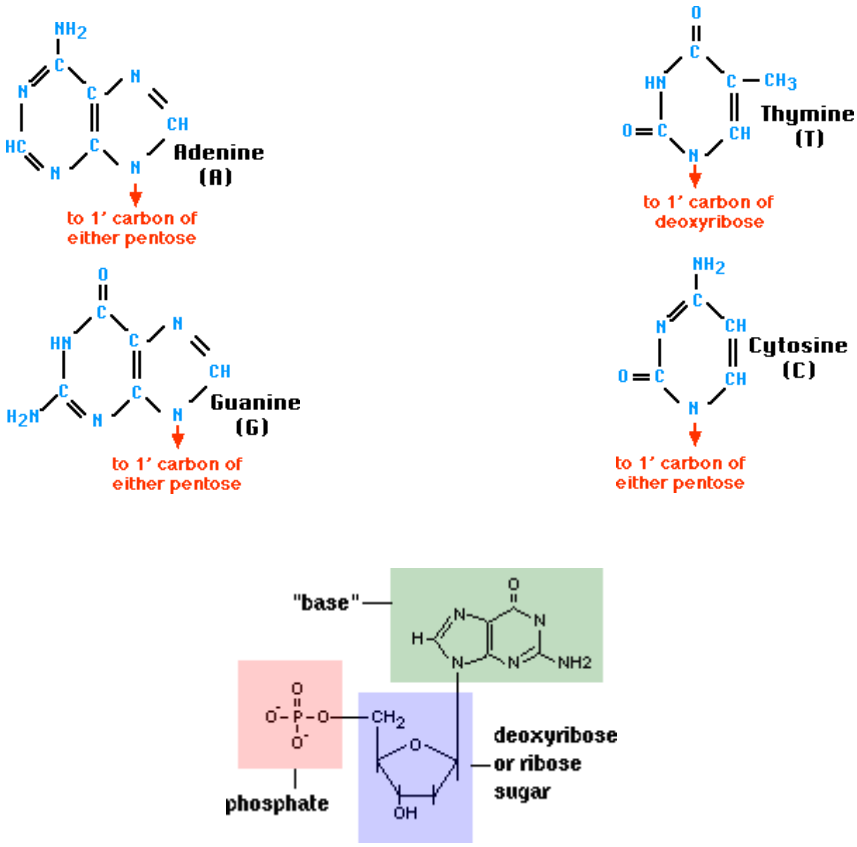


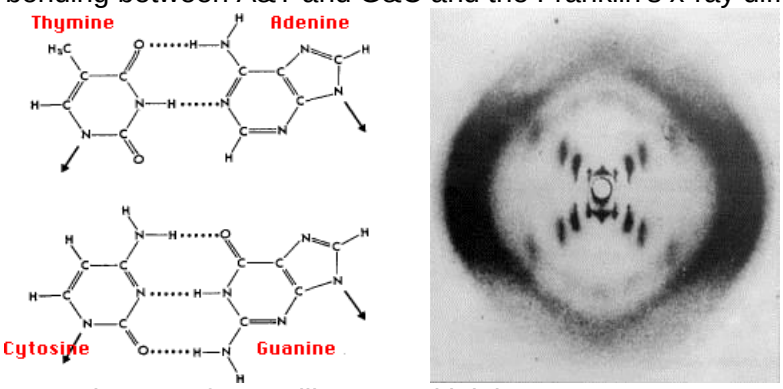

**Figure 7-8:** The chemical structure of the DNA purine nitrogenous bases (left), the

## Chargaff rules

Erwin Chargaff and his coworkers, in the 1940s, estimated the proportion of nitrogenous bases in many different forms of life, and published data showing that the percentage of nitrogenous bases varies from organism to organism. However, in 1947, Chargaff concluded that the amount of purines and pyrimidines show certain regularities, which is now known as the Chargaff rules where the amount of purines equals the amount of

pyrimidines and the amount of adenine equals the amount of thymine and the amount of guanine equals that of cytosine (Figure 7-9). In the attempts to develop a model of DNA structure Franklin (1952) took x-ray diffraction photomicrographs of crystalline DNA extract. In 1952 they published the x-ray pictures of DNA taken by Franklin in Wilkins laboratory in London showing some kind of helix. This picture indicated that the DNA strands are twisted (Figures 7-9 & 7-11).

**Figure 7-9:** Hydrogen bonding between A&T and G&C and the Franklin's x-ray diffraction pictures of crystalline DNA (right)




## The Double Helix

During the late 1940s and early 1950s, many scientists were interested in deciphering the structure of DNA; among them were Francis Crick, James Watson, Rosalind Franklin, and Maurice Wilkins. The data known at the time was that DNA was a long molecule of nucleotides. From the Chargaff rules, Watson and Crick deduced that DNA must be comprised of two strands of nucleotides with bases to the center (like rungs on a ladder) and sugar-phosphate units along the sides of the ladder (Fig.7-10). The two strands are complementary where A pairs with T and C pairs with G by hydrogen bonds (Fig. 7-10).

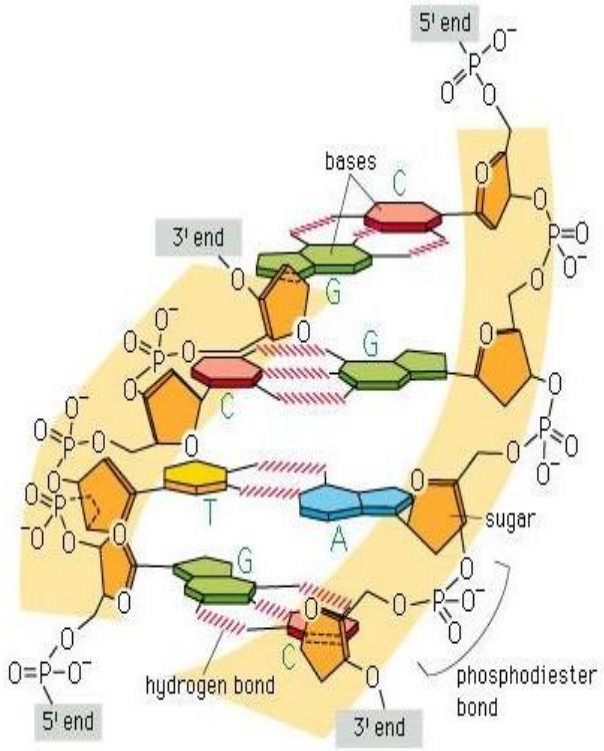



**Figure 7-10:** Illustration of the hydrogen bonding between the A & T and G & C in the DNA strands and of phosphodiester bonds between the sugar and phosphate in opposite directions.

In 1953 Watson and Crick published the DNA double helix model built from Chargaff rules, Franklin data and the chemical properties of the bases, deoxyribose, and the phosphate group. In the DNA double helix, the sugar and the nitrogenous bases are linked by covalent bonds that attach carbon 1 of the sugar and the N in position 9 of the purines or position 1 of thepyrimidines. The nucleotides are linked together by phosphodiester bond where a phosphate group is attached to the carbon 5 of the sugar and the next group to carbon 3 of the same sugar (Fig. 7-10). The DNA strands are pf opposite polarity: that is; the sugars in one strand is linked 5 → 3 direction and in the other strand in 3 → 5 direction. Finally the width of the DNA molecule is 0.2 nm (20Å) the distance between nucleotides is 0.34 nm and the length of the twists is 10 nucleotides i.e. 3.4 nm (Figure 7-11)

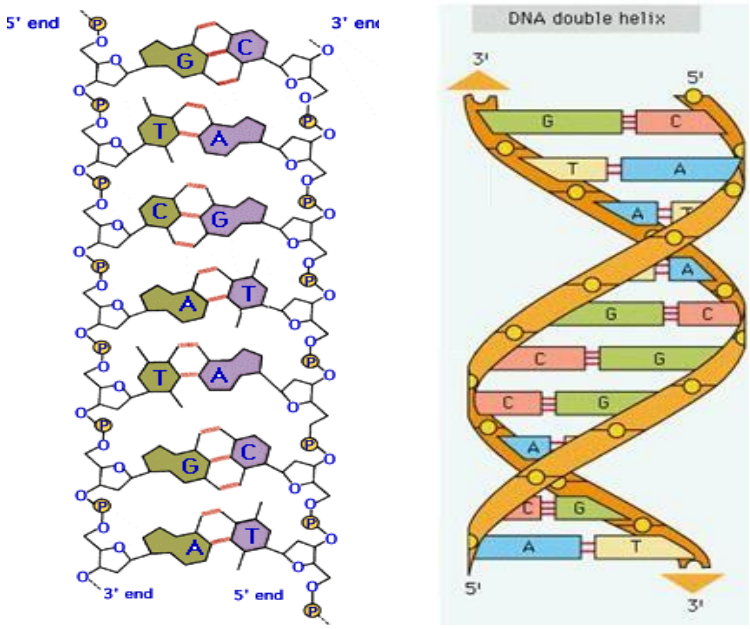



**Figure 7-11:** Illustration of DNA structure details (left) and the double helix of the DNA molecule (right)

## The Basic Functions of DNA

DNA has three basic functions in the life and reproduction of living organisms**:**

1. DNA must carry information from parent cell to daughter cell and from one generation to the next generation. Therefore, it must contain information for replicating itself and repairing its replication mistakes.
2. Genes in the DNA must be expressed as phenotypic traits; the process of gene expression is demonstrated by the control of DNA in protein synthesis.
3. The DNA is chemically stable, relatively unchanging. However, it must be capable of mutational change that could not be repaired. Without mutations there would be no process of evolution.

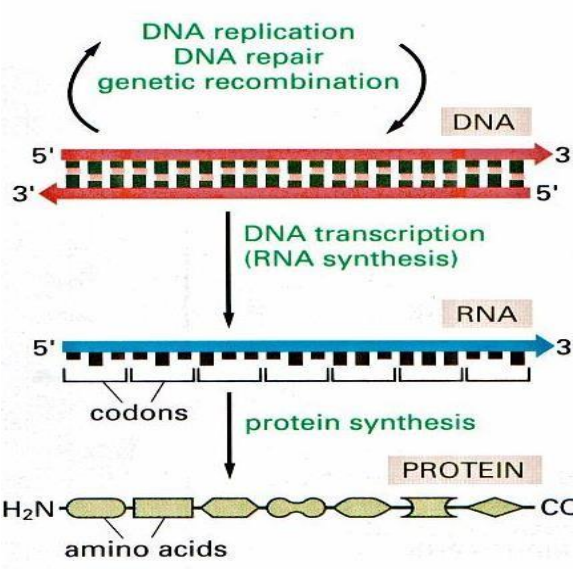



**Figure 7-12:** Diagrammatic representation of the three major functions of DNA in the life and reproduction of living organisms

## DNA replication

The process of DNA replication depends on the ability of both strands of DNA to act as template for new strands. During DNA replication, each DNA strand is used as a template to synthesize the second DNA strand. The new DNA strand is always synthesized in the 5' to 3' direction**.** As a result, the whole DNA molecule is duplicated. DNA replication is therefore termed semi-conservative i.e. in the "next generation" molecule one strand is "old" and another is "new"**.** Semiconservative replication would produce two DNA molecules, each of which is composed of one-half of the parental DNA along with an entirely new complementary strand. In other words the new DNA would consist of one new and one old strand of DNA (Figure 7-13).

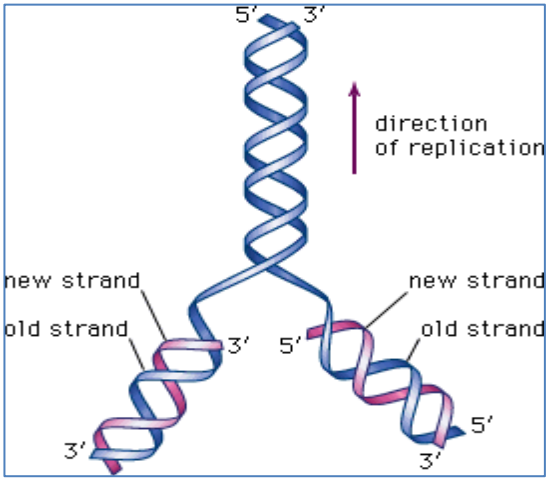



**Figure 7-13:** A simplified representation of the semiconservative nature of DNA replication

The semi-conservative nature of DNA replication in prokaryotes was confirmed by the Messelson and Stahl (1958) using the then new methods of isotope labeling of chemical compounds and ultracentrifugation. Using ultracentrifugation, it was possible to distinguish the DNA containing heavy $^{15}N$ isotope from DNA containing the normal $^{14}N$ isotope. Messelson and Stahl grew the bacterium *Escherichia coli* for several generations on a medium to which $^{15}N$ was added as a nitrogen source instead of the $^{14}N$ isotope. In this way the DNA in the cell came to contain the $^{15}N$ isotope. They then transferred the bacterial cells to a new liquid medium containing $^{14}N$ and withdrew sample from the culture every 30 minutes. They observed that after one generation, the bacterial cells contained DNA with a density that is intermediate between the $^{15}N$ containing DNA and the $^{14}N$ containing DNA. After another generation the DNA was of two densities one characteristic of the hybrid DNA and the other had intermediate density. These findings (Figure 7-14) proved the semiconservative nature of DNA replication in prokaryotes.

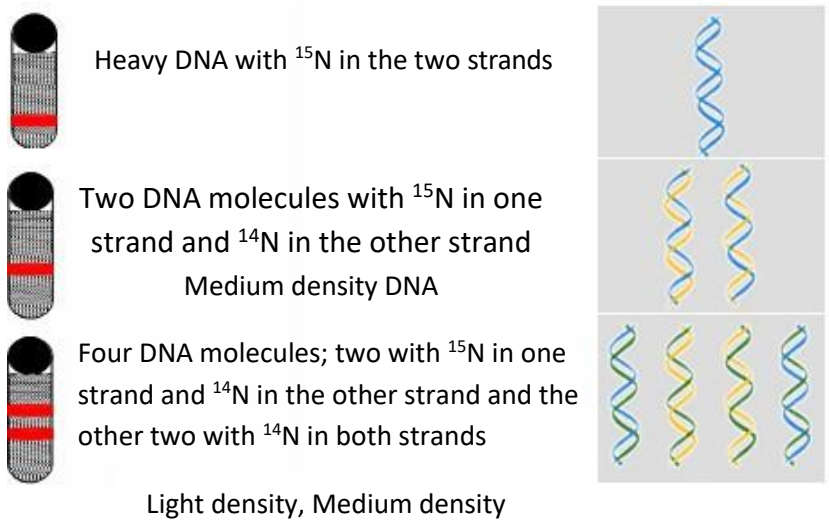



**Figure 7-14:** Diagrammatic representation of the Messelson and Stahl experiment that confirmed the semiconservative mechanism of DNA replication

## Molecular Mechanisms of DNA Replication

DNA replication starts within a special region of DNA called replication origin which is defined by a specific nucleotide sequence. Replication of DNA is bidirectional; two Y-shaped replication forks are moving in the opposite directions. Since two strands in the DNA are antiparallel, the synthesis of one strand is continuous and of the other is discontinuous; the former is called leading strand and the latter called lagging strand. In the lagging strand, DNA synthesis is discontinuous: new DNA is synthesized in small pieces called Okazaki fragments, which then are connected together by DNA ligase (Figure 7-15).

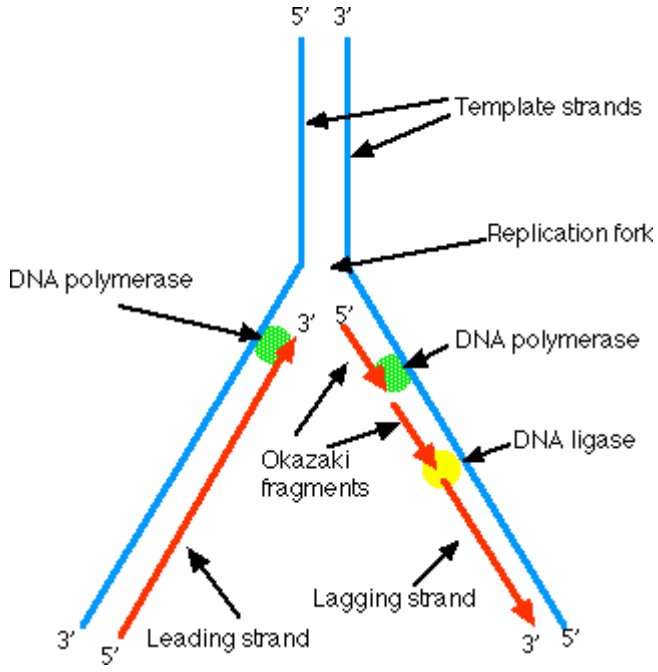



**Figure 7-15:** Diagrammatic representation showing continuous synthesis of the leading strand and the discontinuous synthesis of the lagging strand in Okazaki fragments

DNA replication involves many building blocks, enzymes and a great deal of ATP energy. Nucleotides have to be assembled and made available in the nucleus, along with energy to make bonds between nucleotides. The enzyme DNA helicase unwinds DNA, in front of the replication origin, opening replication fork. Meanwhile, single-stranded DNA binding proteins bind to separated DNA strands to prevent base-pairing back together. The enzyme that assembles the nucleotides in the new DNA strands is called

DNA polymerase, this enzyme cannot start a growing strand from scratch; it needs a short primer of a few RNA nucleotides to add DNA nucleotides to it. This is carried out by DNA-dependent RNA primase that makes very short piece of RNA by base-pairing RNA nucleotides with template DNA. The DNA polymerase then adds DNA nucleotides to the primer. A diagram illustrating the molecular reactions and enzymes involved in DNA replication is shown if Figures 7-16. & 7-17.

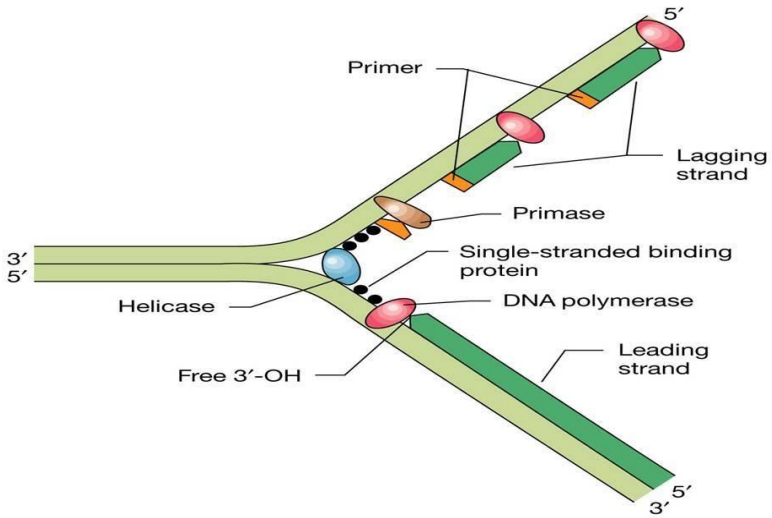



**Figure 7-16:** Diagrammatic representation illustrating the action of DNA replication enzymes

The DNA polymerase was first isolated by Kornberg and requires $Mg^{++}$ and template DNA. Energy for forming new sugar- phosphate phosphodiester bonds, in the new strands, comes from splitting a high-energy phosphate bond in the ATP with the release of Pi. This reaction always occurs at free 3'-OH group on deoxyribose. The new strands thus grow by addition of the phosphate groups at the 3'-end, not at 5'-end of the sugar molecule deoxyribose.

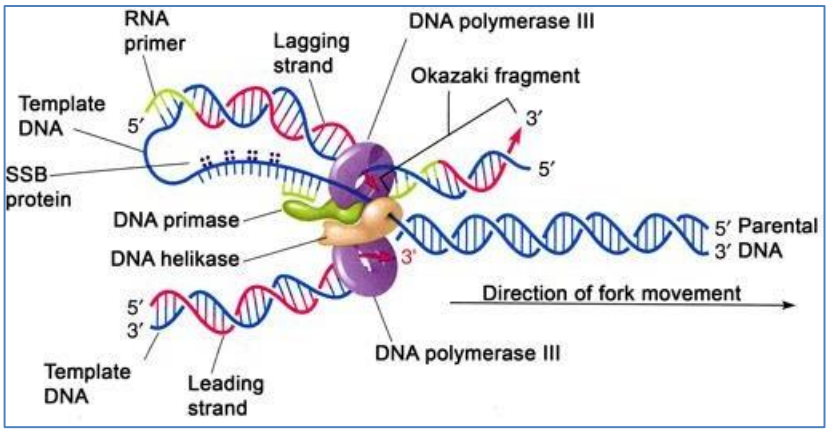



**Figure 7-17:** Summary of the process of DNA replication

## DNA Replication Repair

Damage to DNA may occur in the DNA by physical, chemical or biological factors in the environment such as UV light and many domestic chemicals. Cells often spend much more energy repairing DNA than synthesizing it. All living cells have a variety of DNA polymerase enzymes; some serve for new strands and others for DNA replication repair. When new DNA is synthesized, occasional errors in base pairing occur with frequency of ~ 1 in 10,000 base pairs. If not corrected, it could lead to mutations, loss of functions, loss of competitiveness, evolutionary weeding out. DNA repair is a proof reading process made by the enzyme DNA ligase that seals any gaps and removes mispaired nucleotides in the new strands. The process of DNA replication repair is diagrammed in Figure 7-18.

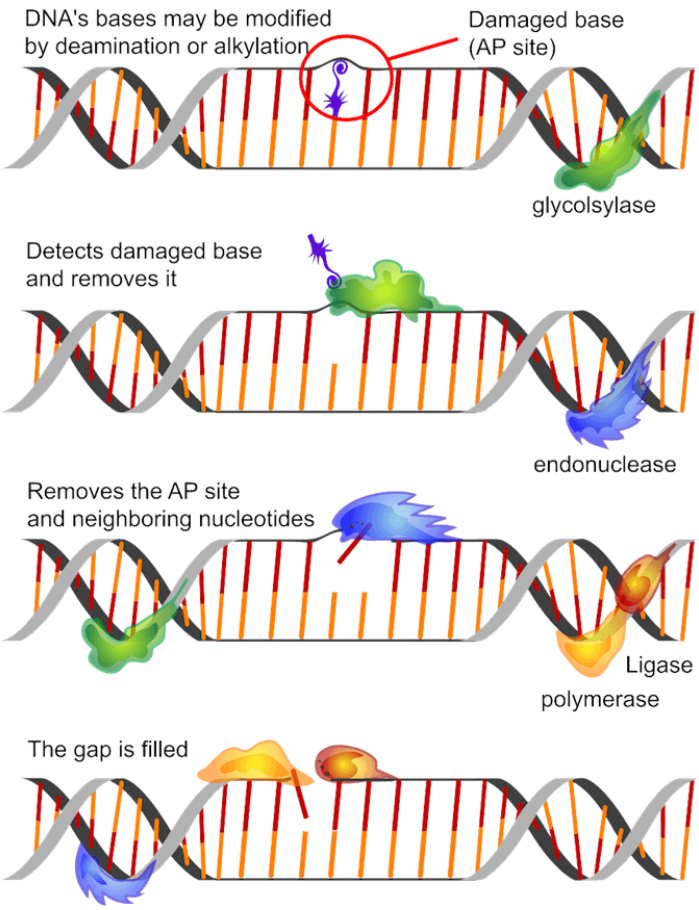



**Figure 7-18:** Diagram illustrating simple process of DNA replication repair that may occur by deamination or alkylation

## DNA Replication in prokaryotes and eukaryotes

In prokaryotes the genome is made of a single circular molecule that opens at a single replication origin with two growing ends of the replication forks that meet again at the origin of replication. In eukaryotes replication starts at numerous replication origins, each proceeds in two directions to replicate parts called **replication unit** or **replicon**. The arrangement of DNA molecules in the chromosomes of eukaryotes into replication units is necessary so that the enormous amount of DNA, relative to the prokaryotic genome, can be replicated within reasonable time. This has been demonstrated by electron microscope and autoradiography studies that showed that replication starts at many different sites along the chromosomes.

The multiple replication units' hypothesis for DNA replication in eukaryotic DNA was demonstrated by Taylor in *Vicia faba* chromosomes. The spacing between the origins of replication units varies between 10 and 100 nm ($3\times10^4$ to $3\times10^5$ base pairs) of nucleotides. Cairns (1963) followed the replication of the $^3H$ labeled DNA in the *E. coli* genome using autoradiography; he was able to demonstrate that the replication in bacterial genome has a single replication origin. The replication then proceeds in two directions until two new DNA molecules are produced. A diagrammatic representation of the single origin of replication in bacteria and the numerous origins of replications in eukaryotes is shown in Figures 7-19 & 7-20.

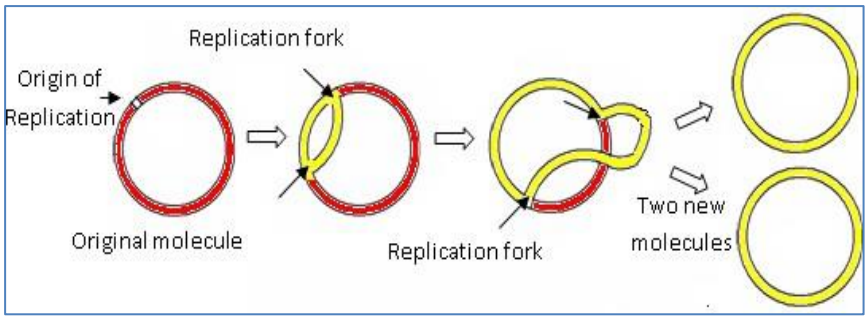



**Figure 7-19:** Diagrammatic representation of DNA replication in bacteria with a single replication origin

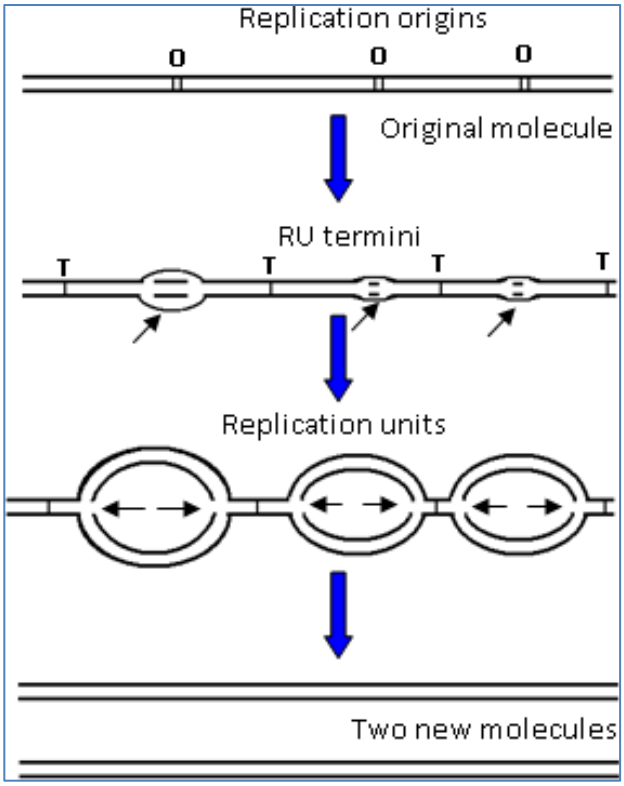



**Figure 7-20:** Diagrammatic representation of DNA replication in eukaryotes with many replications origins

## Gene Expression

### The central dogma

The genetic information is contained in specific sequences of nucleotides in DNA. These sequences are translated into proteins, which are responsible for the phenotypic traits of the living organisms and thus form the link between genotype and phenotype**.** The pathway from DNA to RNA to proteins is known as **central dogma of genetics**; the information is passed as triplet code instructions (codons) from the DNA in a process called transcription to a sequence of codons in a messenger RNA (mRNA) which is translated into a sequence of amino acids in a polypeptide of a protein. Molecules of another RNA; the transfer RNA (tRNA), carry the amino acids and match their anticodons to the mRNA codons. The translation of mRNA codons in amino acids takes place on the ribosomes,

which is comprised of a third type of RNA, the ribosomal RNA (rRNA) and proteins.

The tRNAs and rRNAs are transcribed from DNA but their sequences are not translated into proteins. In prokaryotes, which lack a nuclear membrane, transcription of DNA to mRNA and translation of mRNA into protein occur almost simultaneously. In eukaryotes, the mRNA must exit the nucleus before translation can occur. RNA processing, the modification of pre-RNA or the primary transcript within the nucleus, occurs only in eukaryotes. The steps of gene expression as illustrated in protein synthesis are shown in Figure 7-21.

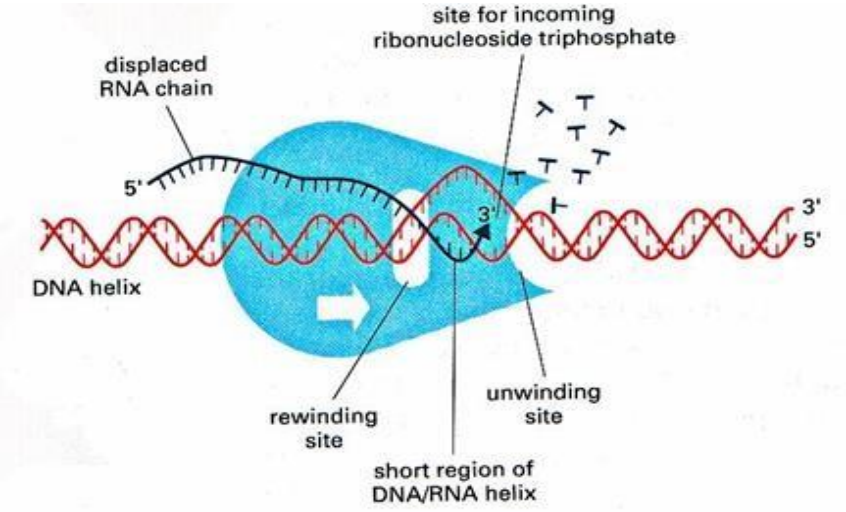



**Figure 7-21**: Diagrammatic representation of the steps of gene expression illustrated in protein synthesis in eukaryotes

## The differences between DNA and RNA

RNA is the link between a gene and the protein for which it codes. RNA differs from DNA in three ways: the sugar component of its nucleotides is ribose, rather than deoxyribose, the pyrimidine base uracil (U) replaces thymine as one of its nitrogenous bases, and RNA is usually single stranded. The structural features that distinguish RNA from the DNA are shown in Fig. 7-22.

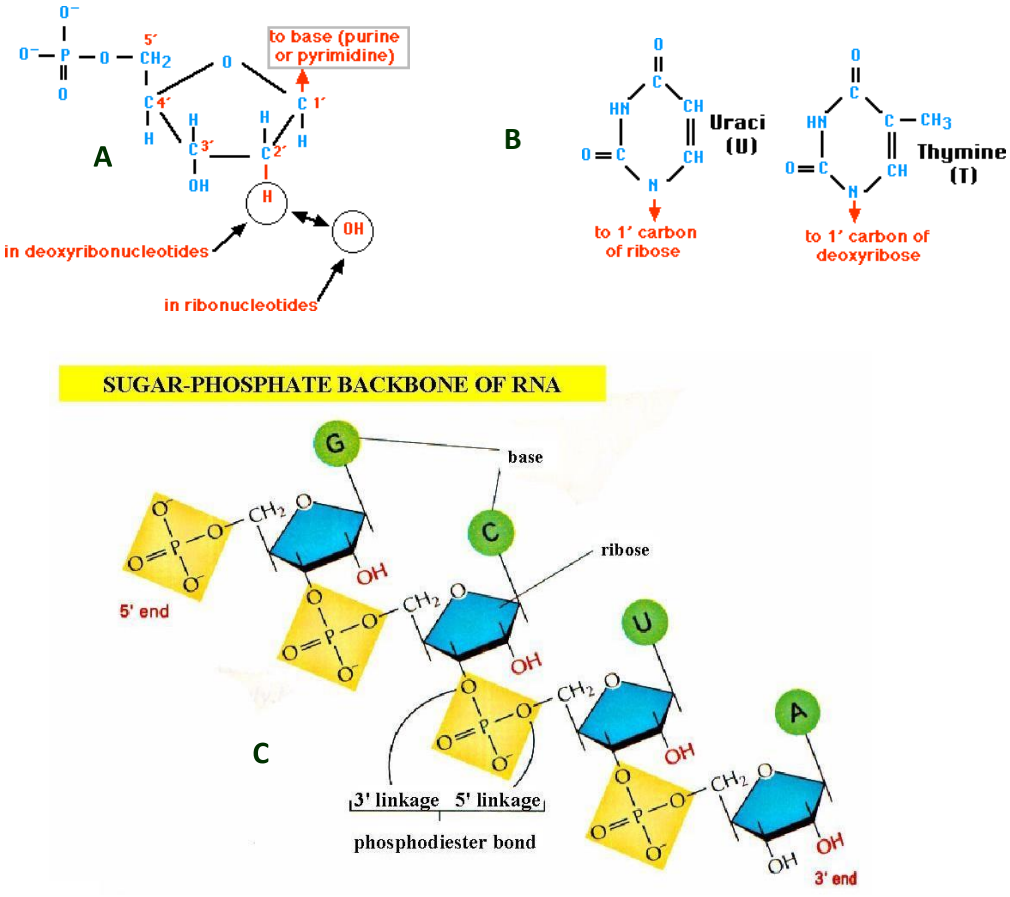



**Figure 7-22:** Structural features that distinguish RNA from the DNA: A = ribose vs deoxyribose B = uracil vs thymine
and C = single-stranded RNA

## Transcription of RNAs

In genetic context, transcription is the DNA-directed synthesis of RNAs. A transcription unit is the sequence of DNA that is transcribed into one RNA molecule. The enzymes that control RNA transcription are generally known as RNA polymerases. Transcription begins when an RNA polymerase separates the two strands of DNA and links RNA nucleotides that base-pair along the template strand of DNA to the 3' end of the growing RNA polynucleotide. Bacteria have one type of RNA polymerase while eukaryotes have three types; the one that synthesizes mRNA is called RNA polymerase II. These enzymes are named **polymerase I**, **polymerase II**, and **polymerase III**. Furthermore, these three enzymes have distinct roles in the cell. Polymerase I makes ribosomal RNA (the 18S, 28S, and 5.8S varieties); polymerase II makes messenger RNA; polymerase III makes transfer RNA.

RNA polymerases bind to promoters in the DNA, which include an initiation site for mRNA synthesis and recognition sequences such as the TATA box in eukaryotes i.e. a nucleotide sequence that is upstream from the initiation site (Figure 7-23). Transcription factors are proteins that first bind to the promoter and aid RNA polymerase II in locating and binding to this region, producing an assembly known as a transcription initiation complex. Nucleotide sequences in the promoter determine the orientation of RNA polymerase and thus which DNA strand is used as the template.

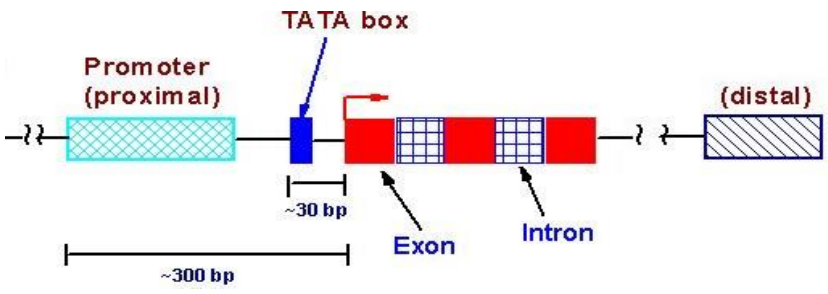



**Figure 7-23: Diagrammatic representation of gene structure in eukaryotic DNA**

RNA polymerase untwists and separates the DNA double helix, exposing DNA nucleotides for base-pairing with RNA nucleotides, and joins the activated RNA nucleotides, (also called ribonucleoside triphosphate) to the 3' end of the growing RNA chain. The growing RNA chain is displaced from the DNA template and the DNA double helix rewinds again to its original form. Several molecules of mRNA polymerase may be transcribing simultaneously from a single gene, enabling the cell to produce large quantities of a protein. Transcription ends after RNA polymerase transcribes a sequence called a terminator. In eukaryotes, polymerase continues past the termination sequence AAUAAA before it cuts loose the pre-mRNA. The process of mRNA transcription is illustrated in Figure 7-23.

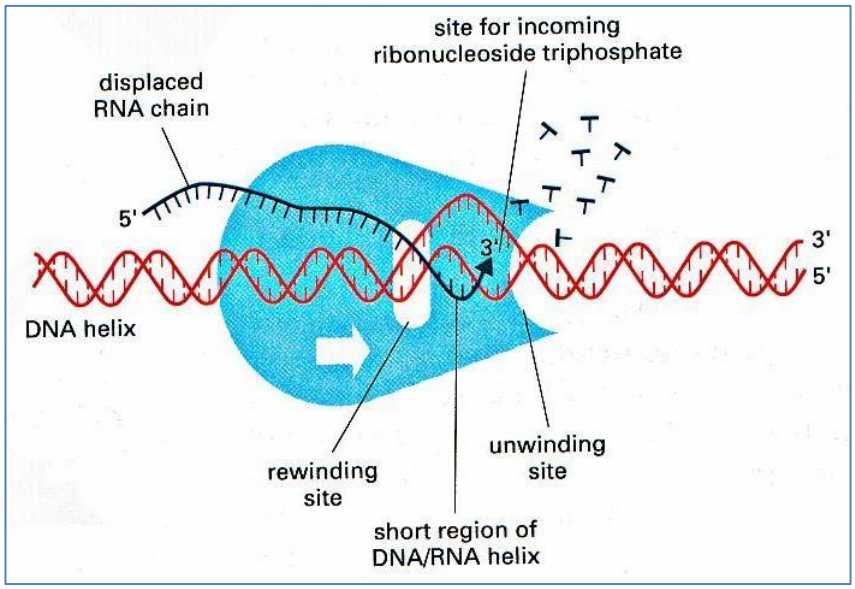



**Figure 7-24:** Illustration of RNA transcription from DNA

## Post-transcription modifications of mRNA in eukaryotes

In eukaryotes, mRNA is transcribed as a pre-RNA or the primary transcript within the nucleus, where is processed to mature mRNA molecule. The pre-mRNA is preceded by a sequence known as the leader and followed by a sequence known as the trailer. The post-transcription modifications in the pre- mRNA include the following:

### 1. Alteration of mRNA ends

These alterations includes (i) Capping the 5' end of a pre- mRNA by the addition of a modified guanine nucleotide (7- methyle guanidine). (ii) Attaching a string of adenine nucleotides, called a poly (A) tail, to the 3' end. The 5' cap serves as part of a recognition site for ribosomes, and both the 5' cap and poly (A) tail protect the ends of the mRNA from hydrolytic enzymes. The poly (A) tail also aids transport of the mRNA from the nucleus and ribosome attachment.

### 2. Removing introns and joining exons

Long segments of non-coding nucleotide sequences, known as introns or intervening sequences, occur within the boundaries of eukaryotic genes. The remaining coding regions are called exons, since they are expressed in protein synthesis. A primary mRNA transcript is made of the gene, but introns are removed and exons joined before they leave the nucleus, in a process called RNA splicing (Fig. 1-22).

Signals for RNA splicing are sets of a few nucleotides at either end of each intron. Small nuclear ribonucleoproteins (snRNPs), composed of proteins and small nuclear RNA

(snRNA), are components of a molecular complex called a spliceosome. The spliceosome snips an intron out of the RNA transcript and connects the adjoining exons. In addition to splice- site recognition and spliceosome assembly, a function of snRNA may be catalytic in intron removal. RNA molecules that act as enzymes are called ribozymes.

Some introns are involved in regulating gene activity, and splicing may help regulate the export of mRNA from the nucleus. Alternative RNA splicing allows some genes to produce different proteins. Exons may code for polypeptide domains, functional segments of a protein, such as binding and active sites. Introns may facilitate recombination of exons between different alleles or even between different genes to create novel proteins.

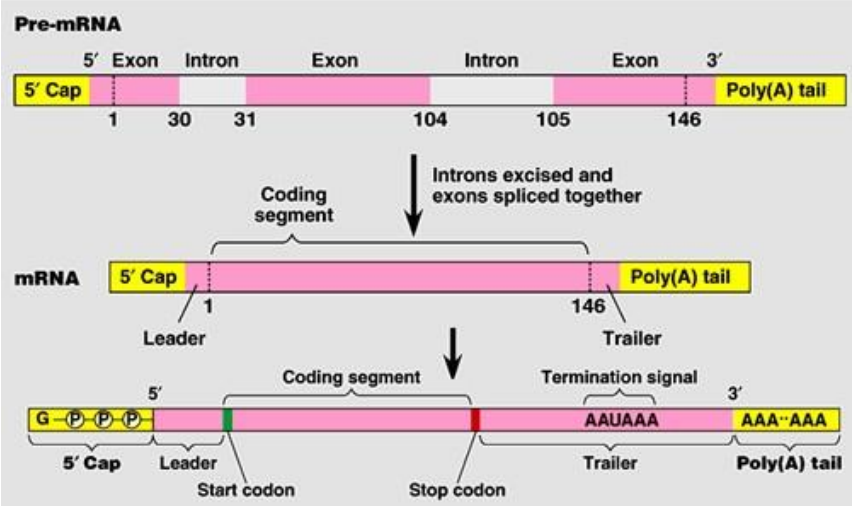



**Figure 7-25:** Illustration of mRNA post-transcription modifications in eukaryotes

**Transfer RNA (tRNA)**

In both prokaryotes and eukaryotes, the RNA to make tRNA and to make rRNA are considerably processed and in some cases cut out from much longer strands of RNA. In its final form, the tRNA molecule is a small single strand of nucleotides (~780b) and is looped on itself in some parts. There are different molecules of tRNA, each with a specific base triplet, called an anticodon that binds to a complementary codon on mRNA, thus assuring those amino acids are arranged in the sequence prescribed by the transcription from DNA. However, all tRNAs share general resemblances in size, shape and structure.

All are in the form of a clover leaf structure and show double stranded regions and have a considerable proportion of unusual bases in the single strand regions. These single-stranded, short RNA molecules are arranged into a clover-leaf shape by hydrogen bonding between complementary base sequences and then folded into a three-dimensional, roughly L-shaped structure. The anticodon is at one end of the L; the 3' end is the attachment site for its amino acid. The 3' end of the molecules is the amino acids attachment site and always has the sequence ACC. The three bases at the bottom loop are the anticodon. The general features of the tRNA molecules are illustrated in Fig. 7-25. As with other RNAs, transfer RNA is transcribed in the nucleus of eukaryotic cells and moves into the cytoplasm where it can be used repeatedly.

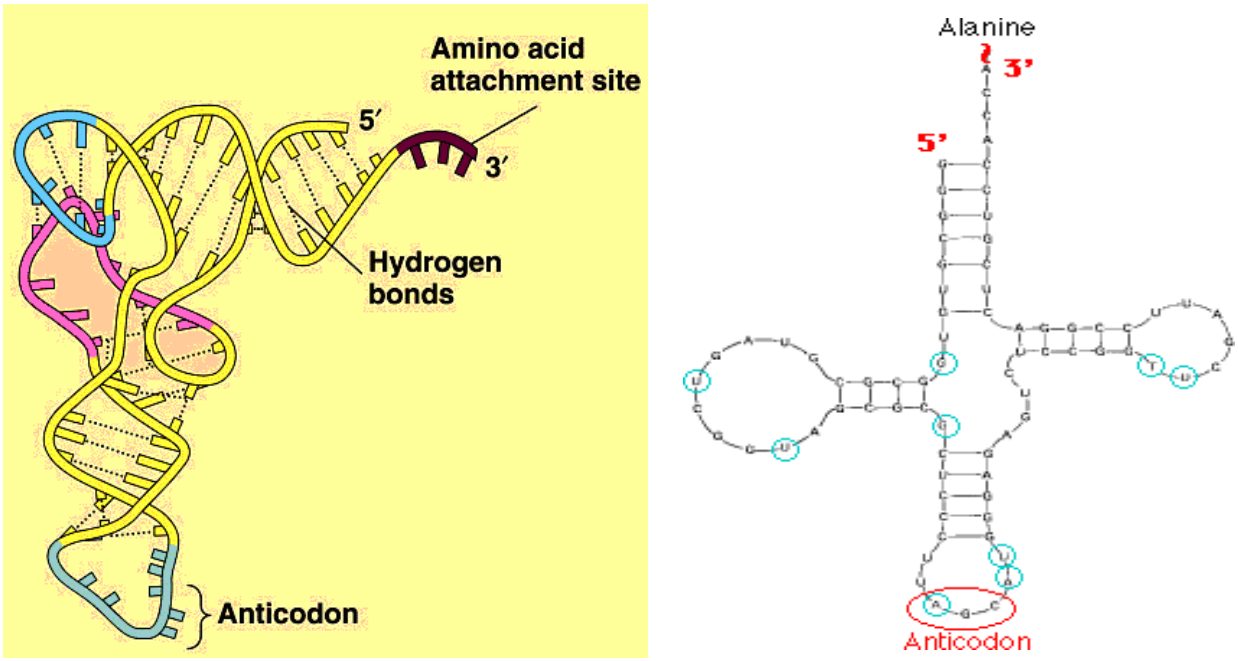



**Figure 7-26:** Diagram illustrating the two dimensional structure of the alanine tRNA (above) and the three dimensional structure of the tRNA molecule (below)

## Ribosomal RNA (rRNA)

Transcription of rDNA leads to a large precursor called a pre-rRNA in both prokaryotes and eukaryotes, which is processed to form smaller mature rRNA molecules that participate in forming the ribosomes. Ribosomes facilitate the specific coupling of tRNA anticodons with mRNA codons during protein synthesis. The ribosome consists of a large and a small subunits, each composed of proteins and RNA molecules. Its subunits are constructed in the nucleolus in eukaryotes. Prokaryotic ribosomes are smaller and differ enough in molecular composition that some antibiotics can inhibit them without affecting eukaryotic ribosomes. In *E. coli*, the pre-rRNA is made as a 30S (Svedberg units) and is cleaved into three pieces having 23S, 5S and 16S; the 23S and 5S pieces are linked to 31 proteins to produce a large 50S ribosome subunit and the 16S piece is linked to 21 proteins to produce the small 30 ribosome subunit (Fig. 7-27).

The ribosome subunits combine together to form the whole ribosome, which has 70S in the prokaryotes. In the eukaryotes, the pre-rRNA is larger than the prokaryotes; it is transcribed as 45S precursor and is cleaved into three pieces with 18S, 28S and 5.8S, the first piece takes part in the formation of the small 40S ribosome subunit and the other two pieces take part in the formation of the large 60S subunit, the whole ribosome size in eukaryotes is 80S.

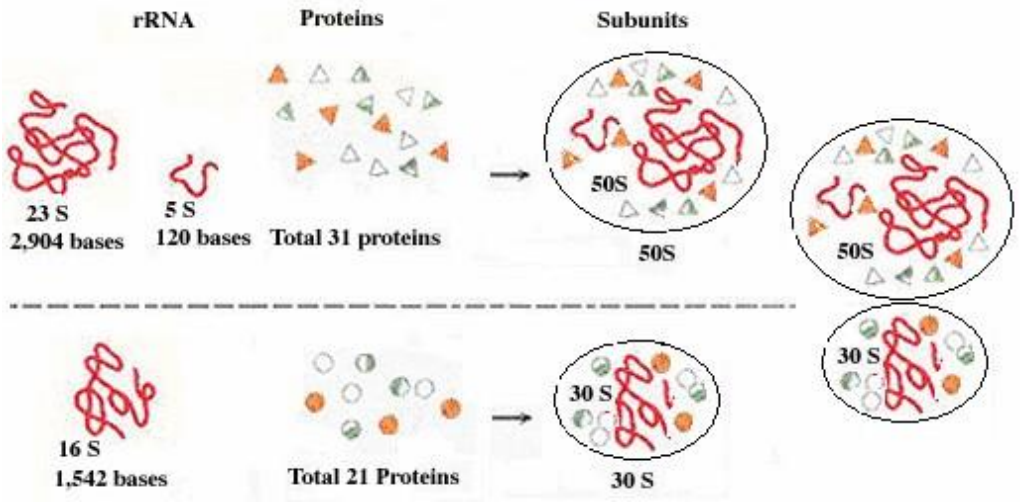



**Figure 7-27:** Diagrammatic illustration of the steps of ribosome formation in *E. coli*

## Translation of genetic information and the genetic code

In the genetic context, translation is the RNA-directed synthesis of a polypeptide in protein synthesis. The reactions of translation are comprised of three stages: polypeptide chain initiation, elongation, and termination, all require the aid of initiation factors. The first two stages also require energy, which is provided by GTP (guanosine triphosphate).

## Polypeptide chain initiation

The initiation stage begins as the small subunit of the ribosome binds to the leader segment of mRNA; this reaction is controlled by IF3 as follows:-

mRNA+30S ribosome+IF3 $\longrightarrow$ mRNA.30S.IF3

An initiator tRNA, carrying the amino acid methionine (formyl methionine in prokaryotes), attaches to the start codon AUG on the mRNA with the aid of

IF2. Each amino acid has a specific aminoacyl-tRNA synthetase that attaches it to its appropriate tRNA molecule to create an aminoacyl tRNA. The hydrolysis of GTP drives this process.

$$IF2.GTP + tRNA + fmet \longrightarrow tRNA.fmet.IF2.GTP$$

The third initiation factor (IF1) drives the combination of the products of the above two reactions forming a 30S initiation complex as in the following reaction.

$$mRNA.30S.IF3 + tRNA.f.met.IF2.GT \longrightarrow (30S.mRNA.tRNA.fmet) + IF3 + IF2 + IF1$$

The large subunit of the ribosome attaches to the small one, forming a translation initiation complex. The steps of polypeptide chain initiation are illustrated in Fig. 7-28.

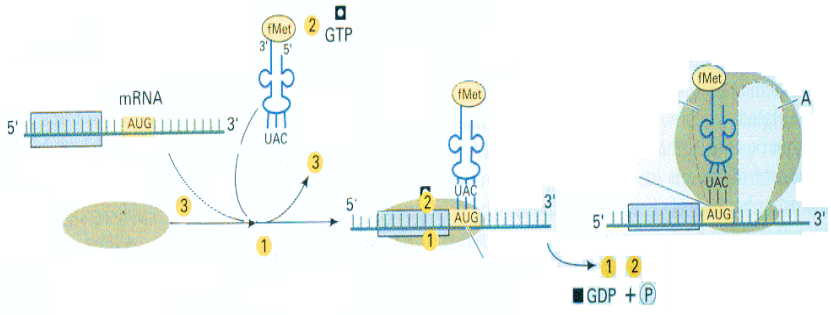



**Figure 7-28:** Diagrammatic illustration of the steps of polypeptide chain initiation

## Polypeptide chain elongation

The addition of amino acids in the elongation stage of polypeptide synthesis involves several enzymes called elongation factors and occurs in a three-step cycle: codon recognition, peptide bond formation and translocation.

In the codon recognition step, an elongation factor brings the correct tRNA into the A binding site of the ribosome, where the anticodon

hydrogen-bonds with the mRNA codon. This step requires energy from the hydrolysis, of GTP (Fig. 1-26).

In the second step, an enzyme called peptidyle transferase and RNA molecule of the large subunit catalyze the formation of a peptide bond between the carboxyl end of the methionine in the P site and the amino acid alanine in the A site. The polypeptide is now held by the amino acid in the A site (Fig. 7-28).

The tRNA from the A site then moves to the P site in the step of translocation and then leaves the ribosome. The tRNA carrying the polypeptide is now translocated to the P site, a process requiring energy from the hydrolysis of another GTP molecule. The next mRNA codon moves into the A site through the ribosome to receive the aminoacyl-tRNA of the thrionine. The steps of chain elongation are illustrated in Figure 7-28.

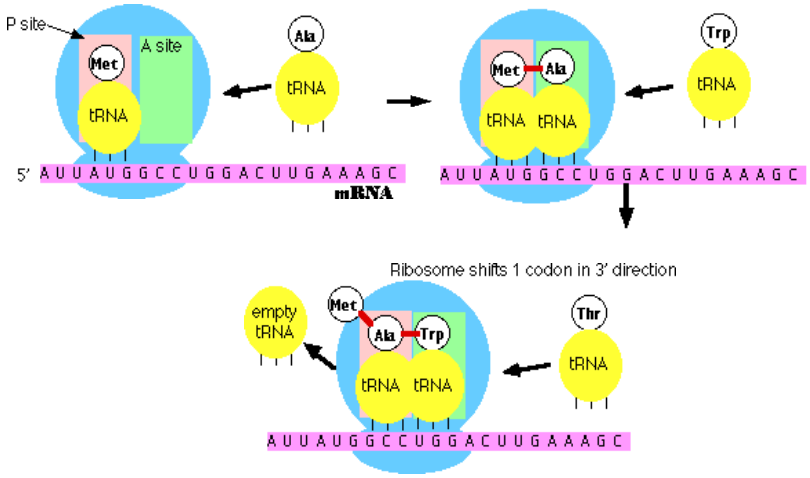



**Figure 1-26:** Diagrammatic illustration of the steps of polypeptide chain elongation

## Polypeptide chain termination

Termination occurs when one of three termination (stop) codons in the mRNA i.e. UAA, UAG, or UGA reaches the A site of the ribosome. A release factor binds to the stop codon and hydrolyzes the bond between the polypeptide and the tRNA in the P site, freeing the completed polypeptide. The steps of chain termination are shown in Fig. 7-29. The two ribosomal subunits and other components then dissociate.

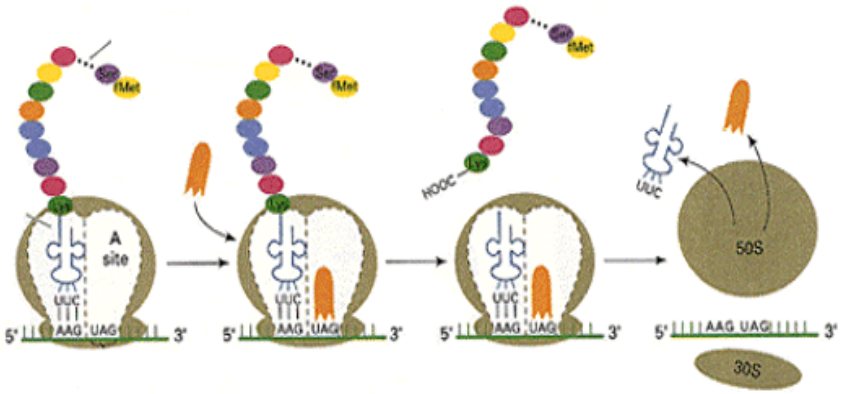



**Figure 7-29:** Diagrammatic illustration of the steps of polypeptide chain termination

During translation, mRNA may be translated simultaneously by several ribosomes in strings called polyribosome. During and following translation, a polypeptide folds spontaneously into its secondary and tertiary structure. Chaperone proteins often facilitate the correct folding. The protein may need to undergo posttranslational modifications: Amino acids may be chemically modified; one or more amino acids at the beginning of the chain may be enzymatically removed; segments of the polypeptide may be excised; or several polypeptides may associate into a quaternary structure.

## Eukaryotic polypeptides transport

In eukaryotes, protein synthesis takes place in the cytoplasm. The polypeptides in these organisms are targeted to specific destinations in the cell. If a protein is destined for the endomembrane system or for secretion,

its polypeptide chain will begin with a signal peptide that is recognized by a protein-RNA complex called a signal-recognition particle, (SRP), which attaches the ribosome to a receptor protein that is part of a multiprotein complex on the ER membrane. As the growing polypeptide threads into the ER, the signal peptide is usually removed. Other signal peptides direct some proteins made in the cytosol to specific sites such as mitochondria, chloroplasts, or the interior of the nucleus.

## The Genetic Code

The nucleotide sequence on mRNA is read in the correct reading frame, starting at the start codon and reading each triplet sequentially. A sequence of three nucleotides provides $4^3$ or 64 possible unique sequences of nucleotides, more than enough to code for the 20 amino acids. In the early 1960s, Nirenberg and his co-workers attempted to discover the codons for all amino acids, by synthesizing artificial mRNA of one type of nucleotides. By linking uracil RNA nucleotides and adding this "poly U' to a test tube containing all the biochemical ingredients necessary for protein synthesis, they obtained a polypeptide containing the single amino acid phenylalanine. Nirenberg and his group realized that such approach would not lead to deciphering the possible 64 codons. By *in vitro* synthesis of mini mRNAs that contains one of the 64 possible codons, Nirenberg and Mathai deciphered all 64 codons by 1966. Like the discovery of DNA structure in the 1950s, cracking of the genetic code in the 1960s, is considered one of the major achievements of science in the 20$^{th}$ century.

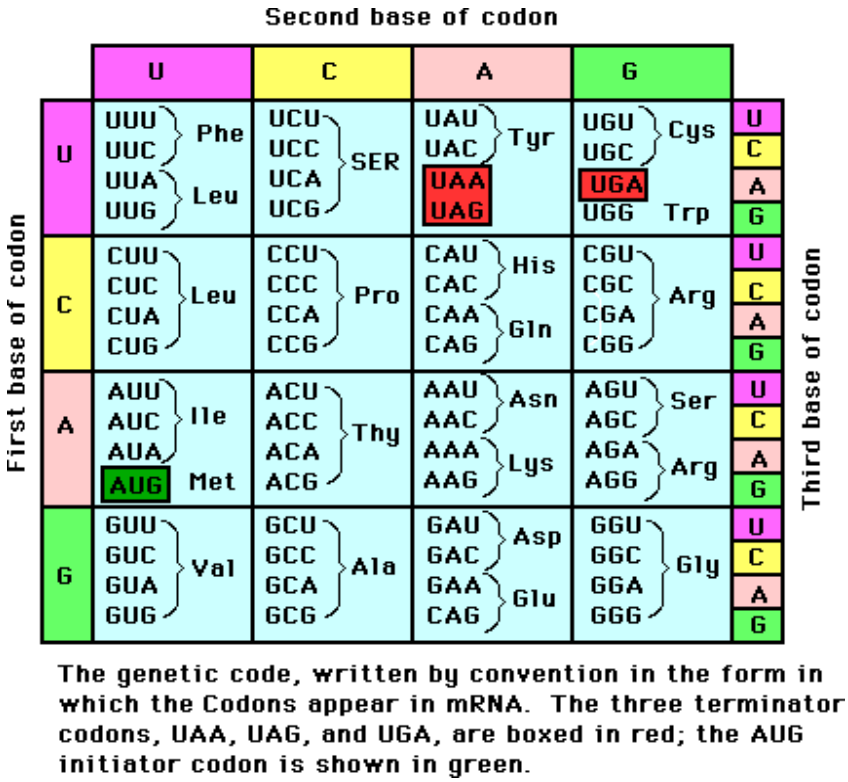



**Figure 7-30:** Diagrammatic illustration the genetic code represented as mRNA 64 codons

The genetic code of codons and their corresponding amino acids is almost universal. A bacterial cell can translate the genetic messages of human cells. The near universality of a common genetic language lends compelling evidence to the antiquity of the code and the evolutionary connection of all living organisms. The code is often redundant, meaning that more than one codon may specify a single amino acid. The code is never ambiguous; no codon specifies two different amino acids or termination codons. The codon AUG codes for methionine and functions as an initiation codon, a start signal for translation. The genetic code is illustrated in the Figure 7-30.

## Wobble

Research on the genetic code revealed 61 codons for amino acids that can be read from mRNA, but there are only about 45 different tRNA molecules. A phenomenon known as wobble enables the third nucleotide

of some tRNA anticodons to pair with more than one kind of base in the codon. Thus, one tRNA can recognize more than one mRNA codon, all of which code for the same amino acid carried by that tRNA. This is demonstrated by the presence of two, four or even six codons for the same amino acid (see the genetic code). This means that the first two bases in the mRNA codons are enough to read the anticodons of tRNA that fix the correct amino acid in its position. In Wobble, the base in the third position plays no role in pairing codon with anticodon. The third base may be a modified base such as inosine (I) that has been found in the third position on several tRNAs and can pair with U, C.

# Gene Mutations

## Definitions and terms

Heredity is a conservative process; the information encoded in the nucleotide sequence of the DNA is faithfully reproduced during replication so that each replication results in two DNA molecules identical to each other and to the parental one. Occasionally, however mistakes lead to different nucleotide sequences in parental and daughter DNA molecules; these changes are called gene mutations. A gene mutation is an alteration in the nucleotide sequence of the gene that affects the structure, and therefore, the function of the proteins expressed from the mutant gene. Mutations often occur during DNA replication, repair, or recombination. These are changes that are not corrected by DNA polymerase or other DNA repair mechanisms. The term mutation is used to designate both the processes by which hereditary changes arise and the outcomes or end products of such process “mutants’. In the most inclusive sense, genetic mutations are changes in the hereditary materials not due to genetic recombination.

Mutations can occur in somatic or germ cells, somatic mutation are not inherited and therefore are not regarded as true mutations. Mutations are also classified as chromosomal mutations and gene mutations; chromosomal mutations include the numerical and structural changes described in the part dealing with cytogenetics. Gene mutations generally arise as a result of base-pair substitutions, insertions, deletions, of a single nucleotide but longer mutations may involve the loss of a whole gene or even some genes. Mutations caused by base insertion or deletion are called frame shift mutations.

## Substitution Mutations

Nucleotide substitutions can be either of transitions or transversions. Transitions are replacements of a purine by another purine (A by G or vice

versa) and of a pyrimidine by another pyrimidine (C by T or vice versa). Whereas, transversions are replacements of a purine by a pyrimidine and vice versa (C or T for G or A or vice versa). Substitutions in the nucleotide sequence of a structural gene may or may not result in changes in the amino acid sequence of the polypeptide encoded by the gene and consequently severely affect or not the biological function of the protein. Based on the effect of base substitution of the mutant gene product, the following types of substitution mutations are recognized:

### 1. Silent mutations

A base-pair substitution replaces one nucleotide and its complementary partner with another pair of nucleotides. Due to the redundancy of the genetic code, some base-pair substitutions in the third nucleotide of a codon do not change gene translation and are called silent mutations. This happens if the nucleotide sequence is changed and different codon codes for the same amino acid. Silent mutations do not affect the functioning of the gene product (Figure 7-31).

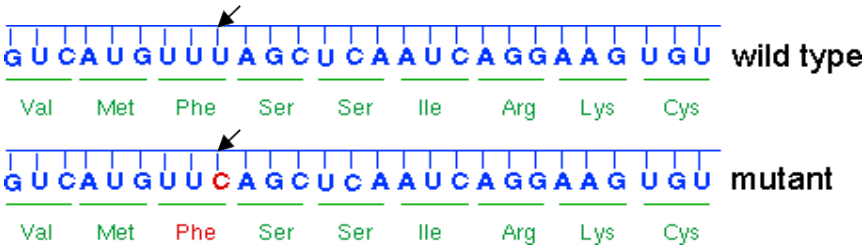



**Figure 7-31:** Diagrammatic representation of a silent mutation; the change of U to C, in the 3rd codon, does not change coding for the amino acid phenyl alanine.

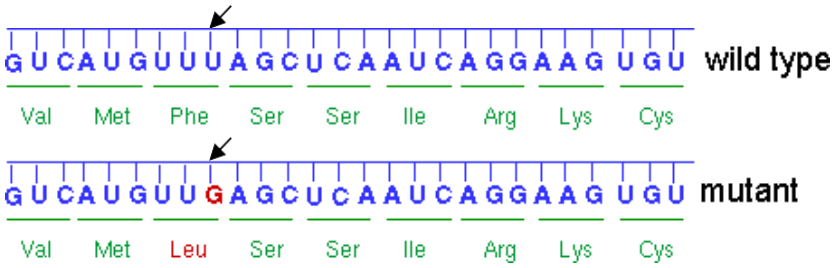
### 2. Mis-sense mutations

A substitution of a nucleotide in the gene sequence that result in the insertion of a different amino acid without altering the character of the protein is called mis-sense mutation (Fig. 1- 29). This occurs when the new amino acid has similar properties or is not located in a region crucial to that protein's function. Mis-sense mutations may not affect the functioning of the gene product. However, if the new amino acid is inserted into a crucial

part of the protein such as the active site of an enzyme, it may significantly impair protein function.

**Figure 7-32:** Diagrammatic representation of a mis-sense mutation; the change of U to G, in the 3rd codon, changes coding for the amino acid phenyl alanine to leucine

### 3. Non-sense mutations




Non-sense mutation is a substitution of a nucleotide in the gene sequence that results in generation of a stop codon prematurely halting the translation of the polypeptide chain and usually creating a nonfunctional protein (Figure 7-33). Only truncated protein can be produced from the mutant gene. The non- sense mutation usually leads to the production of inactive gene product. A base-pair substitution that results in a different amino acid in a critical portion of a protein,

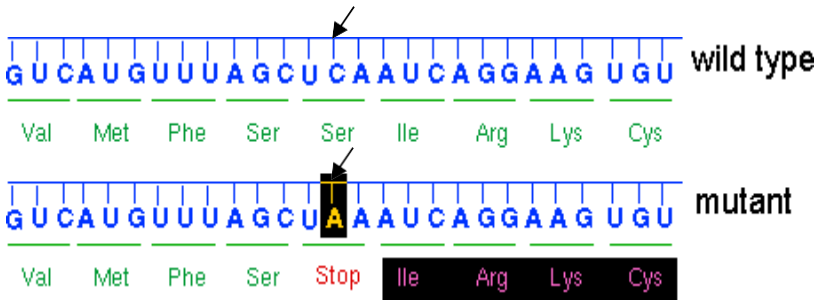



**Figure 7-33:** Diagrammatic representation of a non-sense mutation; the change of C to A changes coding for the amino acid serine to stop codon

## Frame-shift mutations

Base-pair insertions or deletions that are not in multiples of three nucleotides alter the reading frame of the mRNA sequence. All nucleotides downstream from the mutation will be improperly grouped into codons, creating extensive mis-sense and usually prematurely ending in nonsense. In frame-shift mutations, a significant part of the encoded protein can have an incorrect amino acid sequence (Figure 7-34). Frame-shift mutations usually lead to the production of inactive gene product that almost always produces nonfunctional proteins.

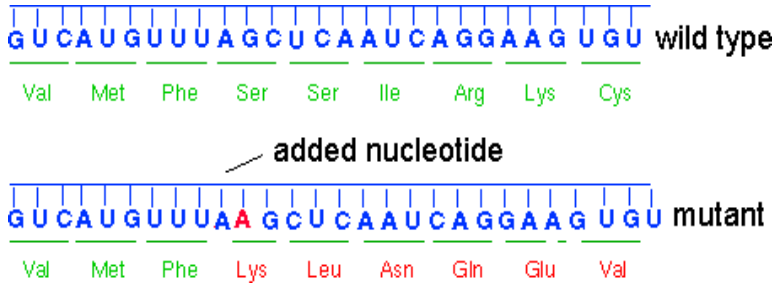



**Figure 7-34:** Diagrammatic representation of a frame shift mutation; the addition of A, in the 4th codon, changes all codons of the next codons in the mRNA strand

The addition or deletion of one or two nucleotide pairs shifts the "reading frame" of the nucleotide sequence from the point of the insertion or deletion to the end of the molecule (Fig. 31). A simple explanatory situation is to assume that a DNA segment is read as five codons of CAT coding for five valine amino acids:

**CAT .... CAT .... CAT …. CAT CAT**

If the T is inserted in the second position of the first codon in this segment, it would then be read as five codons; one CTA coding five aspartic acid and four TCA coding for four serine amino acids.

**↙ CTA …. TCA .... TCA ….TCA ….TCA**

The work on back or reverse mutations by Crick and his coworkers in the 1960s proved that if a total of three nucleotide pairs are either added or deleted, the original reading frame is restored in the rest of the sequence.

Additions or deletions of nucleotide pairs in numbers other than three or multiples of three cause frame mutation. If one nucleotide is inserted at some point and another is deleted at some other point, the original reading frame and the corresponding amino acid sequence will be restored after the second mutational change. An instance of this situation has been worked out in the bacteriophage $T_4$ for the gene coding for a lysozyme, an enzyme that digests bacterial walls.

Some mutant strains of $T_4$ lack a functional lysozyme. Back mutations that occasionally restore the function of the lysozyme in the back mutated strains are called “pseudo-wild” because the amino acid sequence in the lysozyme is different from that of wild strains, as shown in the following figure:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| WT Lysosome | Thr | Lys | Ser | Pro | Ser | Leu | Asn | Ala |
| WT mRNA | **ACX** | **AAY** | **AGU** | **CCA** | **UCA** | **CUU** | **AAU** | **GCX** |
| | | | delete **A** | | | insert **G** | | |
| Pseudo-WT mRNA | **ACX** | **AAY** | **GUC** | **CAU** | **CAC** | **UUA** | **AUG** | **GCX** |
| Pseudo-WT Lysozome | Thr | Lys | Val | His | His | Leu | Met | Ala |

The effect of the above two compensating frame-shift mutations in the gene coding for 1ysozyme in the bacteriophage $T_4$ is limited to the fraction of the messenger RNA shown. A nucleotide deletion has occurred in the third codon and an addition in the seventh codon. The amino acids sequence between the deletion of adenine in the third codon and insertion of guanine in the seventh codon is changed. But the normal sequence is restored after the seventh amino acid shown above; X= A, G, C or U; Y= A or G.

## Causes of Gene Mutations

Based on their cause, mutations are classified as two types, spontaneous and induced.

### Spontaneous mutations

Spontaneous mutations occur for no visible or known agent. It is estimated that one gamete in 10 human gametes carries a new mutant, but spontaneous mutations are recorded at a very low rate. Most genes mutate at a rate of one in a million, but some genes occur at a much higher rate. For example, rate for achondroplastic dwarfism is 40-120 in a million and for color blindness is 30 per million.

### Induced mutations

Induced mutations are caused by physical agents, such as ionizing radiations, X-rays and UV light, and various chemical agents, which include many laboratory and domestic chemicals. The agents that induce mutations are called mutagens.

**Radiations and UV cause mutation by:**

1. Formation of thymine dimers (T~T) in DNA by inducing the formation of a kink that causes covalent linkage of adjacent T bases; this causes errors in replication (Figure 7-35).
2. Induction of chromosomal structural changes. e.g. deletions or translocations.
3. UV irradiation causes photo-reactivation due to excision repair damage
    - Xeroderma pigmentation is a genetic disease caused by a repair defect

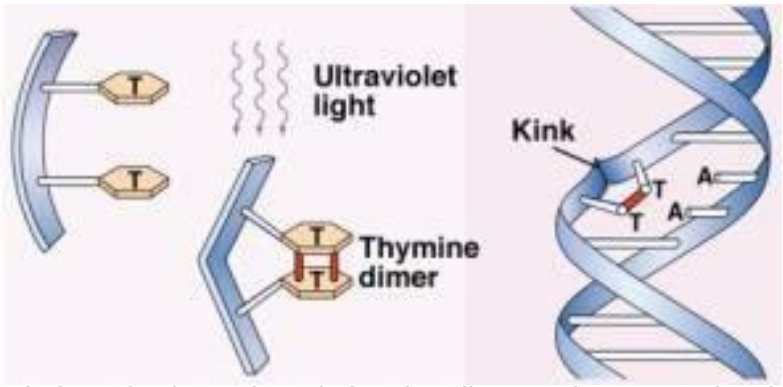



**Figure 1-35:** UV induce the formation of Thymine dimmers in DNA molecules that causes the covalently joined H-bonds converted to covalent bonds

Chemical mutagens are of three types:-

(i) Base analogs that substitute for normal bases and then pair incorrectly in DNA synthesis.

(ii) Chemicals that insert into and distort the double helix. Agents that chemically change DNA bases. These types of chemical mutagens exert their action by one of the following methods:-

1. **Alkylation** - addition of alkyl group (methyl $CH_3$- or ethyl $CH_3$-$CH_2$-) e.g. ethyl methane sulfonate (EMS). This modifies G to 6-ethyl G that pairs with T (Figure 7-36).
2. **Deamination** - conversion of amino to keto group e.g. nitrous acid ($HNO_2$) (common food additive). This modifies C to U by loss of $NH_2$ U then pairs with A. A hypoxanthine by loss of $NH_2$, pairs with C.
3. **Depurination** - loss of purine (A, G), loss of C or T.
4. **Intercalation**- base analogue such as acridine dyes and ethidium bromide wedge into double helix of DNA. Intercalated analogue read as 'extra' base in the DNA strands.

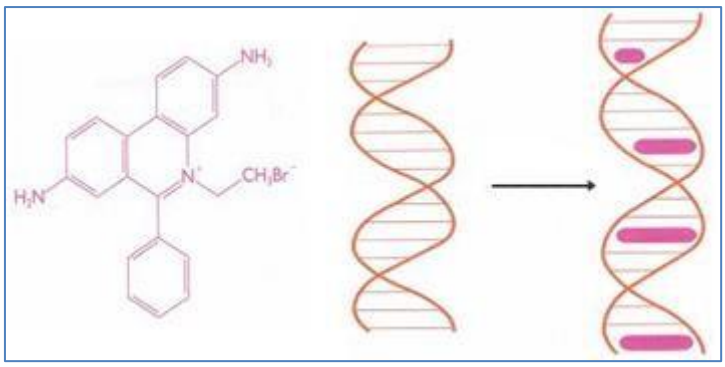



**Figure 7-36:** Chemical structure of ethidium bromide and the way it intercalate in the DNA molecule as a base analogue

## Tautomeric shift transitions

A heritable change in base sequence can occur if one of the strands in a parental DNA duplex molecule undergoes a temporary change in sequence just at the time the chromosome is being replicated so that the daughter strand comes to contain one or more wrong bases.

Since Watson and Crick, in their original conception of the structure of DNA, mutations were predicted to occur by one such process, now known as tautomeric shift. We now know that the purines and pyrimidines of DNA usually exist in particular toutomeric forms and that both the stability of an AT or GC base pairs and the accuracy of base pairing during replication are dependent on the existence of these stable chemical configurations. The most known tautomers are base analogs such as the 5-bromoracil and 2-amino purine; 5-bromouracil in its keto form replace T and pairs with A and in its enol form replace C and pairs with resulting in substitution mutations (Fig. 1-34). Tautomeric shifts are also caused by some other mutagens that cause base modification; the well-known base modifying mutagens are Nitrous acid, Hydroxyl amine, ethyl methane sulfonate (EMS) and MMS methyl ethane sulfonate (MMS).

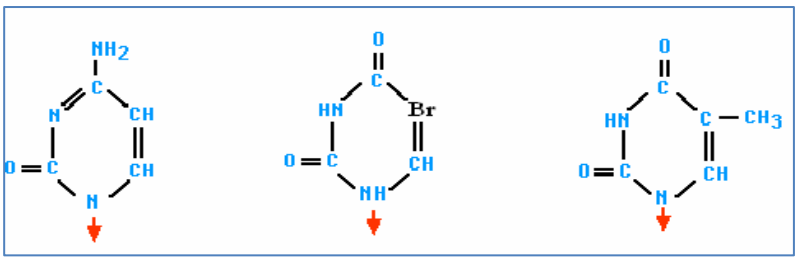



**Figure 7-37:** Chemical structure of 5-bromouracil (middle) compared to cytosine (left) and thymine (right)

### *Effects of mutations*

Most mutations are recessive (not expressed), the effects of expressed mutations may be:-

- **Harmful**: may be lethal and may cause minor or major damage to organisms
- **No effect:** include :- Neutral mutation

- Mutation in non-functional DNA
- Silent mutation

- **Beneficial: -** Increases fitness

- Produce variability

**Chapter 5 post test**

**A- Mark the correct answers with the sign (√) and the wrong answers with the sign (x); correct the wrong statements**

1. The T2 phages have a head and tail shape and contains DNA in the head.
2. In the DNA molecule, adenine pairs with cytosine.
3. In DNA replication, the synthesis of new strands starts with RNA primer
4. Transfer RNA is the link between a gene and the protein for which it codes.
5. The processing of mRNA in eukaryotes involves capping the 5'end with A-tail.
6. In the genetic code, nucleotide triplets in mRNA specify amino acids.
7. In protein synthesis, ribosomal RNA molecules carry amino acids to the ribosomes.
8. Nonsense mutations occur when the codon for an amino acid is changed into a stop codon.
9. In Eukaryotes, protein synthesis occurs in the ribosomes on the rough ER.
10. RNA is made of a single strand of amino acids connected in a polypeptide chain.
11. In DNA replication, the lagging strands is made of segments called Okazaki fragments.
12. The processing of mRNA in eukaryotes involves capping the 5'end with 7mG.
13. Missencee mutations occur when a codon for an amino acid is changed into a stop codon.
14. In the genetic code, AUG is the start codon in mRNA and codes for methionine.
15. Transduction is the transfer of genes between bacterial strains through

plasmids.

16. RNA serves as the genetic material in the P22 viruses.

**B- Circle the correct answer for the following questions**

1. In transduction, genes are transferred by:-

a. Plasmids, b. Phages, c. Plastids, d. Cosmids

2. Compared to DNA, RNA contains:-

a. Single strand, b. Uracil, c. Ribose, d. All the preceding

3. The enzyme that links DNA fragments together is called:-

a. Lipase, b. Ligase, c. Helicase, d. DNAse

4. The number of stop codons in the genetic code is:

a. One, b. Two, c. Three, d. Four

5. In transduction, genes are transferred by:

a. Plasmids, b. Phages, c. Plastids, d. Ribosomes

6. Which of the following codons is not a stop codon:

a. UAA, b.UAG, c. UGA, d.UGG

7. In the DNA molecule, adenine pairs with:

a. Cytosine, b. Glycine, c. Thymine, d. Leucine

8. The number of nucleotides in the codons is: code:

a. One, b. Two, c. Three d. Four

9. The number of codons for tryptophane is:

a. One, b. Two, c. Three, d. Four

10. The enzyme that transcribes RNA from DNA is called:

a. Transaminase, b. RNA polymerase c.DNA helicase d,DNAse

11. UV causes mutation by inducing:

a. Thymine dimmers b.Deamination c.Alkylation d. Depurination

12. By convention, the sequence of bases in a nucleic acid is usually expressed in the_____direction.

a. 3' to 1', b. 3' to 5', c. 1' to 3', d. 5' to 3,' e. clockwise

13. The first DNA nucleotides in a newly synthesized strand in DNA replication is linked to:-

a. DNA primer, b. DNA polymerase, c. ligase, d. RNA primer

14. Each unit of a nucleic acid consisting of a sugar, attached phosphate group and base is called:-

a. Nucleolus, b. Nucleotide, c. Nucleosome, d. histone

15. In a nucleic acid, the bases are always attached to the carbon number of the deoxyribose sugar.

a. 5' , b 4' , c. 3' , d 2' , e 1'

**C- Answer the following questions as required:-**

**1.** Mention four conclusive evidences that had confirmed DNA as the genetic material, Describe, with illustrative drawing, experiments that confirmed this basic fact of biology

**2.** Replication is a basic function of DNA. Describe one experiment that had illustrated the semi-conservative nature of DNA replication

**3.** Mention four types of gene vectors and describe the basic differences between plasmids and phages as gene cloning vectors?

**4.** Based on the following DNA sequence, answer the questions that follow:-

5'…GGC TAC GAC CCC T**C**A TTG ATC CTG TAC TAC CAG CTG … 3'

**a.** Write the DNA complementary strand for this sequence.

**b.** Write the RNA transcript of this strand.

**c.** Write the amino acid sequence for this RNA transcript.

**d.** What is the kind of mutation if **c** is added before the highlighted a. in codon 3

**e.** What is the kind of mutation if the highlighted **c** in codon 5 is replaced by **a**.

**D- Explain the differences between each of the following:-**

1. Missense and non-sense mutations.
2. Replication and transcription.
3. Start codon and stop codon.
4. Exons and introns sequences.

5. DNA replication in prokaryotes and eukaryotes
6. DNA replication and RNA transcription
7. Transformation and transduction
8. Leading DNA strand and lagging DNA strand
9. Start codon and stop codons
10. Messenger RNA and transfer RNA
11. Substitution mutations and frame shift mutation
12. The leading and lagging DNA strands
13. Codons and anticodons
14. Ribosomes in prokaryotes and eukaryotes
15. Spontaneous mutations and induced mutation

**E- Answer the questions below focusing in subtle points**

1. What is the specific sugar found in DNA molecules?
2. What are the complementary pairs of nitrogen bases found in DNA molecules?
3. What is the only nucleotide found in DNA but ***not*** found in RNA? What is the only nucleotide found in RNA but ***not*** found in DNA?
4. What term refers to the original DNA molecule being copied during DNA replication?
5. Why would a cell undergo DNA replication?
6. What does the term “replication fork” refer to when describing DNA replication?
7. Which enzyme causes replication forks to form
8. Why hundreds of replication forks on a single DNA molecule undergoing DNA replication in eukaryotes? What term refers to the two molecules that result during DNA replication?
9. How similar are the DNA molecules produced during DNA replication?
10. What happens to the two strands of the parent (original) DNA molecule being copied during DNA replication?

*Chapter 6*

# Principles of Gene Technology

## Background

In 1967 the enzyme DNA ligase was isolated; this enzyme joins two strands of DNA together, a prerequisite for the construction of recombinant DNA molecule. This was followed by the isolation of the first restriction enzyme in 1970, a major milestone in the development of gene technology. Restriction enzymes are essentially molecular scissors, which cut DNA at precisely defined sequences. Such enzymes can be used to produce fragments of DNA that are suitable for joining by DNA ligase. Another milestone in gene technology was the use of plasmids and viral DNA as gene vectors that act as vehicles to carry genes in DNA fragments. The stage was then set for the appearance of a number of developments that provided the necessary stimulus for gene manipulation to become a reality.

## Gene cloning (The Birth of Gene Technology)

In 1973 Stanley Cohen and Paul Berg formulated six main experimental steps for gene cloning. Gene cloning begins with the isolation of a gene of interest. The gene is then inserted into a vector and cloned. A vector is a piece of DNA that is capable of independent growth; commonly used vectors are bacterial plasmids and viral phages. The gene of interest (foreign DNA) is integrated into the plasmid or phage to form the recombinant DNA. Cloning is necessary to produce numerous copies of the recombinant DNA. Once the vector is isolated in large quantities, it can be introduced into the host cells such as special bacterial cells or mammalian, plant or yeast cells. The host cells will then synthesize the foreign gene product from the recombinant DNA. When the cells are grown in vast quantities, the foreign or recombinant gene product can be isolated and purified in large amounts.

**Steps of gene cloning**

- Fragmentation of cellular DNA using restriction enzymes.
- Separation of DNA fragments and isolation of the gene of interest for further manupulation.
- Vector linearization using the same restriction enzymes.
- Ligation, connecting the DNA of the gene of interest with the cloning vector to form recombinant vector.
- Transformation / transfection, introducing the recombinant vector to the host cells.
- Screening for target gene(s), by:
  - Loss of antibiotic resistance
  - Insertional inactivation (Colony colour)
  - DNA hybridization.
- Clone characterization, by electrophoresis, southern transfer, DNA hybridization or other methods.

The practical procedures for gene cloning in a plasmid or phage vector and the transformation of host bacterial cells are summarized in the following steps; the procedures for gene cloning in a plasmid are illustrated in Figure 8-1.

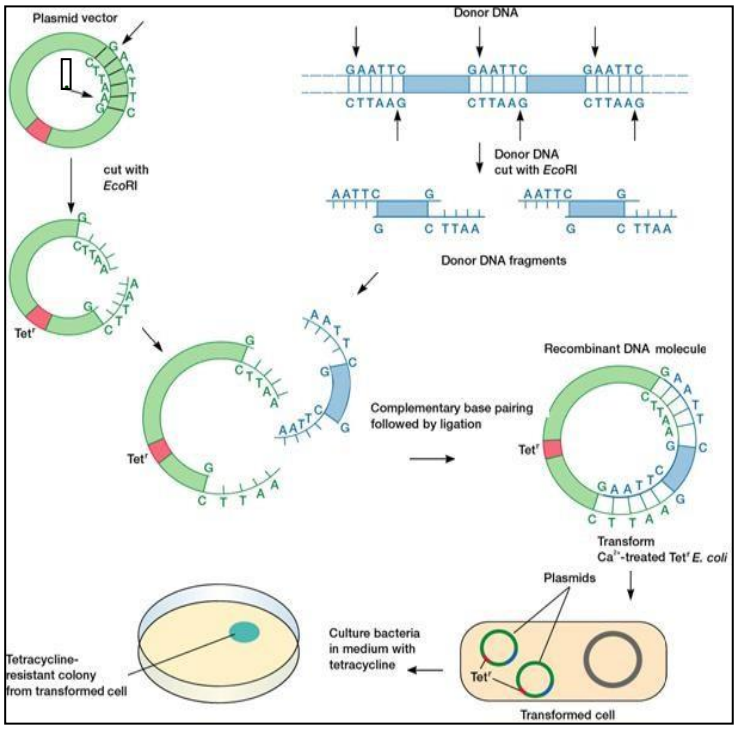



**Figure 8-1:** The basic steps of gene cloning in *E. coli* using a plasmid vector

## Main objectives of gene cloning

1. Identifying genes and the proteins they encode.
2. Storing genes and genomes in DNA libraries.
3. Isolating genes for sequencing, transfer and modifications.
4. Modifying genes and correct endogenous genetic defects.
5. Manufacture of large quantities of specific gene products such as hormones, vaccines.
6. Re-expressing genes in other hosts or transgenic organisms.

## Tools of gene cloning

## Restriction enzymes

- **Basic biology of restriction enzymes**
  - Act as bacterial defense system to attack invading DNA
  - Three types (type II endonucleases are used for DNA cloning)
  - >200 different restriction enzymes have been identified
- ***Recognition sequences***
  - Each recognizes a specific sequence, usually 4- 6 bp
  - Cut both strands of DNA
- ***Produce sticky ends or blunt ends***
  - If cut is asymmetric then generates sticky ends
  - If cut is symmetric then generates blunt ends

Examples of the restriction enzymes produced by different species of bacteria and their recognition sites in the DNA stand are given in table 2-1. Figure 2-2 illustrates that the enzyme *Eco*R1 cuts at six base sequence and produce sticky ends.

**Table 2-1:** Examples of restriction enzymes produced by species of bacteria and their recognition sites in DNA.

| Enzyme | Bacterium | Recognition site | Type of cut ends |
|---|---|---|---|
| *Eco*RI | *Escherichia coli* RY | 5'-G↓AATTC-3'<br>3'-CTTAA↑G-5' | 5'-AATT sticky 5'<br>TTAA-5' |

| | | | |
|---|---|---|---|
| *Bam*HI | *Bacillus amyloliquefaciens* H | 5'-G↓GATCC-3'<br>3'-CCTAG↑G-5' | 5'-GATC sticky 5'<br>CTAG-5' |
| *Pst*I | *Providencia stuartil* 164 | 5'-CTGCA↓G-3'<br>3'-G↑ACGTC-5' | TCGA-3' sticky 3'<br>3'-AGCT |
| *Hpa*1 | *Haemophilus parainfluenza* | 5'-GTT↓AAC-3'<br>3'-CAA↑TTG-5' | GTT-3'blunt5'-AAC<br>CAA-5' 3'-TTG |
| *Hae*I | *Haemophilus haemolyticus* | 5'-GCG↓C-3'<br>3'-C↑GCG-5' | GCG-3' sticky 3'<br>3'CGC |
| *Taq*I | *Thermus aquaticus* YTI | 5'-T↑CGA-3'<br>3'-AGC↑T-5' | 5'-CG sticky 5'<br>GC-5' |

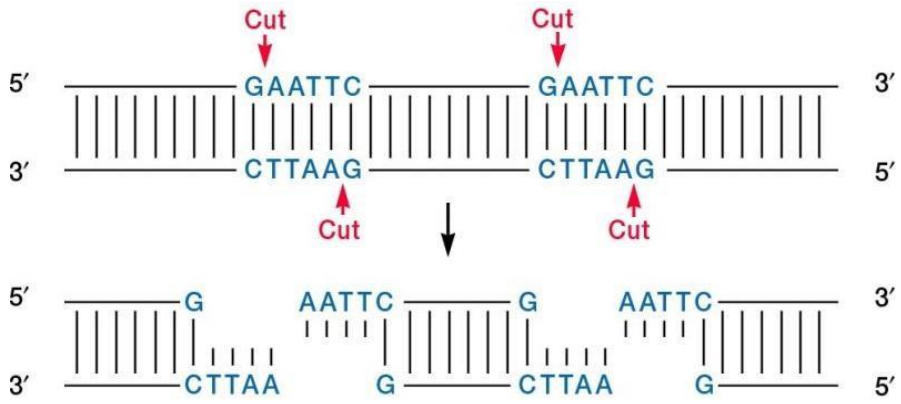



**Figure 8-2:** Diagram illustrating the production of sticky ends by *Eco*RI at a six base sequence in DNA

## *Gene cloning vectors*

Gene vectors are DNA molecules that act as vehicles to carry genes in DNA fragments. It is used to clone genes in a host organism (obtain millions of copies of the gene). A cloning vector has a single origin of replication and is capable of replicating in a host cell. Most vectors are genetically engineered plasmids or phages. There are also other types of gene vectors known as cosmid vectors, bacterial artificial chromosomes, and yeast artificial chromosomes.

- Plasmids are circular, double-stranded DNA molecules that exist in bacterial cells (Figure 8-3) and in the cells of some eukaryotic organisms such as yeast, plants and animals.
- Plasmids replicate independently of the host cell.

- The size of plasmids ranges from a few kb to near 100 kb
- Plasmids are of different types, the most well-known examples include the following types:
  - F-plasmid: Fertility F-plasmids contain tra genes that makes bacteria capable of conjugation resulting in the expression of sex pilli
  - Resistance plasmids: contain genes that provide resistance against antibiotics or poisons. They were historically known as R-factors, before the nature of plasmids was understood.
  - Col plasmids: contain genes that code for bacteriocins; proteins that can kill other bacteria.
  - Degradative plasmids: enable the digestion of unusual substances, e.g. toluene and salicylic acid.
  - Virulence plasmids: turn the bacterium into a pathogen.

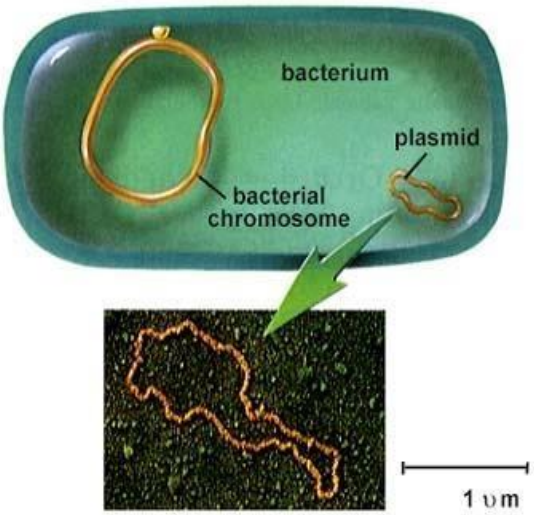



**Figure 8-3:** Diagram illustrating the symbiotic existence of plasmids in bacterial cells and enlarged photograph of plasmid DNA.

# REFERENCES


1. **Badr A. (2014):** General Genetics. faculty of science, Helwan university.
2. **Benjamin A. Pierce (2010):** GENETICS ESSENTIALS, Concepts and Connections. W. H. Freeman and Company / New York.
3. **Snustad, D. Peter. & Michael J. Simmons (2012):** Principles of genetics - 6th ed. John Wiley & Sons, Inc.

# Index