Semester	Course	Hours	Credit	Sub. Code	Marks		
					Internal	External	Total
II	CC 6	6	4	18KP2B06	25	75	100

GENETICS AND MOLECULAR BIOLOGY

UNIT I: GENETICS

Mendelian genetics – Mendel's laws of inheritance – Monohybrid and Dihybrid Cross, Incomplete dominance and Complementary interaction of genes. Epistasis and Lethal alleles, Multiple alleles. General account of ABO blood group in man.

UNIT IV: MOLECULAR BIOLOGY

Structure of Gene- Cistron, Recon, Muton, Salient features of Genetic code, Structure of DNA and RNA – types of DNA (A, B, Z) and RNA (mRNA, t RNA, r RNA, hn RNA). Replication of DNA in Prokaryotes and Eukaryotes. Role of enzymes in DNA replication, DNA damage and repair mechanisms.

REFERENCE

1. Verma PS Agarwal VK. Molecular Biology (First edition),S.Chand and Company Ltd. New Delhi, (2009).
2. Gardner, E.J., 1972, Principles of genetics - John Wiley and sons, N.Y.
3. Freifelder, D.,1986,Molecular Biology. Jones and Bardett Publishin INC. Boston,Portola Valley.
4. Vasishta, P. C. and Gill, P. S. (1998). Genetics: Speciation and Plant Breeding. Pradeep Publications, Jalandhar.

PREPARED BY

UNIT- IV:

Dr. M. BOOMINATHAN

Assistant Professor in Botany

K.N. Government Arts College for Women (Autonomous),

Thanjavur -613 007.

UNIT-I

Dr. A. PAULINE FATHIMA MARY

Guest lecturer in Botany

K.N. Government Arts College for Women (Autonomous),

Thanjavur -613 007.

CORE COURSE - VII
GENETICS AND MOLECULAR BIOLOGY

UNIT I: GENETICS

Mendelian genetics – Mendel's laws of inheritance – Monohybrid and Dihybrid Cross, Incomplete dominance and Complementary interaction of genes. Epistasis and lethal alleles, multiple alleles. General account of ABO blood group in man.

INTRODUCTION TO GENETICS

The history of genetics started with the work of the Augustinian friar Gregor Johann Mendel.His work on pea plants, published in 1866, described what came to be known as Mendelian Inheritance. In the centuries before—and for several decades after—Mendel's work, a wide variety of theories of heredity proliferated.1900 marked the "rediscovery of Mendel" by Hugo de Vries, Carl Correns and Erich von Tschermak, and by 1915 the basic principles of Mendelian genetics had been applied to a wide variety of organisms—most notably the fruit fly Drosophila melanogaster. Led by Thomas Hunt Morgan and his fellow "drosophilists", geneticists developed the Mendelian model, which was widely accepted by 1925. Alongside experimental work, mathematicians developed the statistical framework of population genetics, bringing genetic explanations into the study of evolution. With the basic patterns of genetic inheritance established, many biologists turned to investigations of the physical nature of the gene. In the 1940s and early 1950s, experiments pointed to DNA as the portion of chromosomes (and perhaps other nucleoproteins) that held genes. A focus on new model organisms such as viruses and bacteria, along with the discovery of the double helical structure of DNA in 1953, marked the transition to the era of molecular genetics.

Johann Gregor Mendel (1822-1884) - Father of Genetics Gregor Mendel, through his work on pea plants, discovered the fundamental laws of inheritance. He deduced that genes come in pairs and are inherited as distinct units, one from each parent. Mendel tracked the segregation of parental genes and their appearance in the offspring as dominant or recessive traits. He recognized the mathematical patterns of inheritance from one generation to the next. Mendel's Laws of Heredity are usually stated as:
1) The Law of Segregation: Each inherited trait is defined by a gene pair. Parental genes are randomly separated to the sex cells so that sex cells contain only one gene of the pair. Offspring therefore inherit one genetic allele from each parent when sex cells unite in fertilization.

2) The Law of Independent Assortment: Genes for different traits are sorted separately from oneanother so that the inheritance of one trait is not dependent on the inheritance of another.

3) The Law of Dominance: An organism with alternate forms of a gene will express the form that isdominant.

The genetic experiments Mendel did with pea plants took him eight years (1856-1863) and he published his results in 1865. During this time, Mendel grew over 10,000 pea plants, keeping track of progeny number and type. Mendel's work and his Laws of Inheritance were not appreciated in his time. It wasn't until 1900, after the rediscovery of his Laws, that his experimental results were understood.

MENDEL’S LAWS OF INHERITANCE AND EXCEPTIONS TO THE LAWS

History

The assertion that life can instantaneously arise from non living matter is called spontaneous generation. Here are the critical experiments that busted the myth. Although today we understand that living things arise from other living things, the idea of spontaneous generation was entrenched in the minds of man throughout most of history. Spontaneous generation is the belief that, on a daily basis, living things arise from non living material. This

debunked belief is not the same as abiogenesis, the study of how life on earth could have arisen from inanimate matter billions of years ago.

Cell Theory(1838)

Schleiden and Schwann proposed cell theory 1838. They concluded that all plant and animal tissues were made of cells. It was also postulated that cell is the functional unit of living organism. In 1846 Negeli said that all cells originated from preexisting cells. Virchow 1853 elaborated this and referred it as cell linkage theory.

Mendelian concept of hereditary

The laws of inheritance were derived by Gregor Mendel, a 19th century monk conducting hybridization experiments in garden peas (Pisum sativum). Between 1856 and 1863, he cultivated and tested some 29,000 pea plants. From these experiments he deduced two generalizations which later became known as Mendel's Laws of Heredity or Mendelian inheritance. He described these laws in a two part paper, "Experiments on Plant Hybridization" that he read to the Natural History Society of Bruno on February 8 and March 8, 1865, and which was published in 1866. Mendel's findings allowed other scientists to predict the expression of traits on the basis of mathematical probabilities. A large contribution to Mendel's success can be traced to his decision to start his crosses only with plants he demonstrated were true-breeding. He also measured only absolute (binary) characteristics, such as color, shape, and position of the offspring, rather than quantitative characteristics. He expressed his results numerically and subjected them to statistical analysis. His method of data analysis and his large sample size gave credibility to his data. He also had the foresight to follow several successive generations (f2, f3) of his pea plants and record their variations. Finally, he performed "test crosses" (backcrossing descendants of the initial hybridization to the initial true breeding lines) to reveal the presence and proportion of recessive characters. Without his careful attention to procedure and detail, Mendel's work could not have had the impact it made on the world of genetics.

Mendel's Laws:

Mendel discovered that by crossing white flower and purple flower plants, the result was not a hybrid offspring. Rather than being a mix of the two, the offspring was purple flowered. He then conceived the idea of heredity units, which he called "factors", one which is a recessive characteristic and the other dominant. Mendel said that factors, later called genes, normally occur in pairs in ordinary body cells, yet segregate during the formation of sex cells. Each member of the pair becomes part of the separate sex cell. The dominant gene, such as the purple flower in Mendel's plants, will hide the recessive gene, the white flower. After Mendel self-fertilized the F1 generation and obtained the 3:1 ratio, he correctly theorized that genes can be paired in three different ways for each trait; AA, aa, and Aa. The capital A represents the dominant factor andlowercase a represents the recessive.

Mendel stated that each individual has two factors for each trait, one from each parent. The two factors may or may not contain the same information. If the two factors are identical, the individual is called homozygous for the trait. If the two factors have different information, the individual is called heterozygous. The alternative forms of a factor are called alleles. The genotype of an individual is made up of the many alleles it possesses. An individual's physical appearance, or phenotype, is determined by its alleles as well as by its environment. An individual possesses two alleles for each trait; one allele is given by the female parent and the other by the male parent. They are passed on when an individual matures and produces

gametes: egg and sperm. When gametes form the paired alleles separate randomly so that each gamete receives a copy of one of the two alleles. The presence of an allele doesn't promise that the trait will be expressed in the individual that possesses it. In heterozygous individuals the only allele that is expressed is the dominant. The recessive allele is present but its expression is hidden. Mendel summarized his findings in two laws; the Law of Segregation and the Law of Independent Assortment.

Law of Segregation (The "First Law")

The Law of Segregation states that when any individual produces gametes, the copies of a gene separate, so that each gamete receives only one copy. A gamete will receive one allele or the other. The direct proof of this was later found when the process of meiosis came to be known. In meiosis the paternal and maternal chromosomes get separated and the alleles with the characters are segregated into two different gametes.

Law of Independent Assortment (The "Second Law")

The Law of Independent Assortment, also known as "Inheritance Law", states that alleles of different genes assort independently of one another during gamete formation. While Mendel's experiments with mixing one trait always resulted in a 3:1 ratio between dominant and recessive phenotypes, his experiments with mixing two traits (dihybrid cross) showed 9:3:3:1 ratios. But the 9:3:3:1 table shows that each of the two genes are independently inherited with a 3:1 ratio. Mendel concluded that different traits are inherited independently of each other, so that there is no relation, for example, between a cat's color and tail length. This is actually only true for genes that are not linked to each other. Independent assortment occurs during meiosis I in eukaryotic organisms, specifically metaphase I of meiosis, to produce a gamete with a mixture of the organism's maternal and paternal chromosomes. Along with chromosomal crossover, this process aids in increasing geneticdiversity by producing novel genetic combinations.

In independent assortment the chromosomes that end up in a newly-formed gamete are randomly sorted from all possible combinations of maternal and paternal chromosomes. Because gametes end up with a random mix instead of a pre-defined "set" from either parent, gametes are therefore considered assorted independently. As such, the gamete can end up with any combination of paternal or maternal chromosomes. Any of the possible combinations of gametes formed from maternal and paternal chromosomes will occur with equal frequency. For human gametes, with 23 pairs of chromosomes, the number of possibilities is 2^23 or 8,388,608 possible combinations. The gametes will normally end up with 23 chromosomes, but the origin of any particular one will be randomly selected from paternal or maternal chromosomes. This contributes to the genetic variability of progeny.

Rediscovery of Mendelís work

Mendel's conclusions were largely ignored. Although they were not completely unknown to biologists of the time, they were not seen as generally applicable, even by Mendel himself, who thought they only applied to certain categories of species or traits. A major block to understanding their significance was the importance attached by 19th century biologists to the apparent blending of inherited traits in the overall appearance of the progeny, now known to be due to multigene interactions, in contrast to the organ-specific binary characters studied by Mendel. In 1900, however, his work was "re-discovered" by three European scientists, Hugo de Vries, Carl Correns, and Erich von Tschermak. The exact nature of the "re-discovery" has been somewhat debated: De Vries published first on the subject, mentioning Mendel in a footnote, while Correns pointed out Mendel's priority after having read De Vries's paper and realizing that he himself did not have priority. De Vries may not have acknowledged truthfully

how much of his knowledge of the laws came from his own work, or came only after reading Mendel's paper. Later scholars have accused Von Tschermak of not truly understanding the results at all. Regardless, the "re-discovery" made Mendelism an important but controversial theory. Its most vigorous promoter in Europe was William Bateson, who] coined the term "genetics", "gene", and "allele" to describe many of its tenets. The model of heredity was highly contested by other biologists because it implied that heredity was discontinuous, in opposition to the apparently continuous variation observable for many traits. Many biologists also dismissed the theory because they were not sure it would apply to all species, and there seemed to be very few true Mendelian characters in nature. However, later work by biologists and statisticians such as R.A. Fisher showed that if multiple Mendelian factors were involved in the expression of an individual trait, they could produce the diverse results observed. Thomas Hunt Morgan and his assistants later integrated the theoretical model of Mendel with the chromosome theory of inheritance, in which the chromosomes of cells were thought to hold the actual hereditary material, and create what is now known as classical genetics, which was extremely successful and cemented Mendel's place in history.

Mendel's Laws of Inheritance

Mendel postulated three laws, which are now called after his name as Mendel’s laws of heredity. These are:

1. Law of dominance and recessive
2. Law of segregation
3. Law of independent assortment

1. Law of Dominance

Definition: When two homozygous individuals with one or more sets of contrasting characters are crossed, the characters that appear in the F1 hybrids are dominant characters and those do not appear in F1 are recessive characters.

Law of dominance- If there are two alleles coding for the same trait and one is dominant it will show up in the organism while the other won't

Explanation : The dominance and recessive of genes can be explained on the basis of enzymatic functions of genes. The dominant genes - are capable of synthesizing active polypeptides or proteins that form functional enzymes, whereas the recessive genes (mutant genes) code for incomplete or non-functional polypeptides. Therefore, the dominant genes produce a specific phenotype while the recessive genes fail to do so. In the heterozygous condition also the dominant gene is able to express itself, so that the heterozygous and homozygous individuals have similar phenotype.

Critical appreciation of Law of Dominance Scientists conducted cross-breeding experiments to find out the applicability of law of dominance. The experiments were conducted by Correns on peas and maize, Tschermak on peas, by De Vries on maize etc., by Bateson and his collaborators on a variety of organisms, by Davenport on poultry,by Furst on rabbits, by Toyama on silk moth and by many others. These scientists observed that a largenumber of characters in various organisms are related as dominant and recessive.

Importance of law of dominance

The phenomenon of dominance is of practical importance as the harmful recessive characters are masked by the normal dominant characters in the hybrids. In Human beings a form of idiocy, diabetes, haemophilia etc. are recessive characters. A person hybrid for all these characteristics appears perfectly normal. Thus harmful recessive genes can exist for several generations without expressing themselves.

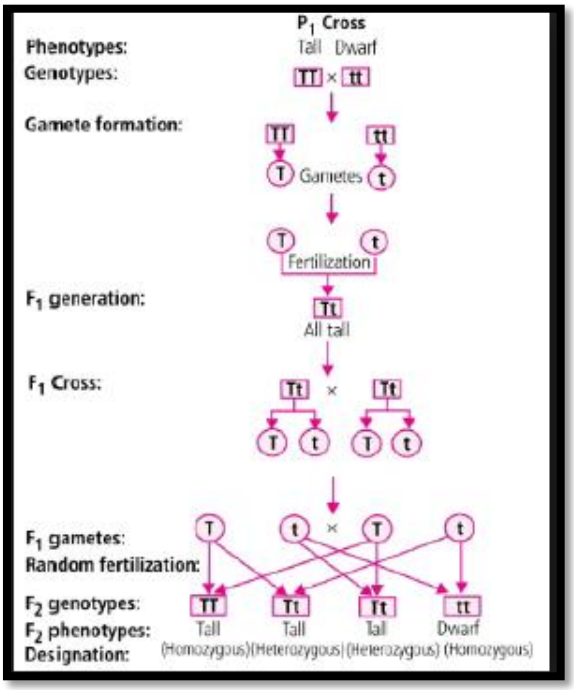
Exceptions to Law of Dominance is the Incomplete Dominance. After Mendel several cases were recorded by scientists, where F1 hybrids exhibited a blending of characters of two parents. These hybrids were found to be midway between the two parents. This is known as incomplete dominance or blending inheritance. It means that two genes of the allelomorphic pair are not related as dominant and recessive, but each of them expresses itself partially. As for example, in four-o'clock plant, Mirabilis jalapa, when plants with red flowers (RR) are crossed with plants having white flowers (rr), the hybrid F1 plants (Rr) bear pink flowers. When these F1 plants with pink flowers are self-pollinated they develop red (RR), pink (Rr) and white (IT) flowered plants in the ratio of 1 : 2 : 1 (F2 generation).

2. Law of Segregation (Purity of Gametes)
Explanation - The law of segregation states that when a pair of contrasting factors or genes or allelomorphs are brought together in a heterozygote (hybrid) the two members of the allelic pair remain together without being contaminated and when gametes are formed from the hybrid, the two separate out from each other and only one enters each gamete.
Example - Pure tall plants are homozygous and, therefore/possess genes (factors) TT; similarly dwarf possess genes tt. The tallness and dwarfness are two independent but contrasting factors or determiners. Pure tall plants produce gametes all of which possess gene T and dwarf plants t type of gametes.
During cross fertilization gametes with T and t unite to produce hybrids of F1 generation. These hybrids possess genotype Tt. It means F1 plants, though tall phenotypically, possess one gene for tallness and one gene for dwarfness. Apparently, the tall and dwarf characters appear to have become contaminated developing only tall character. But at the time of gamete formation, the genes T (for tallness) and t (for dwarfness) separate and are passed on to separate gametes. As a result, two types of gametes are produced from the heterozygote in equal numerosity. 50% of the gametes possess gene T and other 50% possess gene t. Therefore, these gametes are either pure for tallness or for dwarfness. (This is why the law of segregation is also described as Law of purity of gametes).

3. Law of Independent Assortment
Definition: The inheritance of more than one pair of characters (two pairs or more) is studied simultaneously, the factors or genes for each pair of characters assort out independently of the other pairs. Mendel formulated this law from the results of a dihybrid cross. Explanation: The cross was made between plants having yellow and round cotyledons and plants having green and wrinkled cotyledons.
The F1 hybrids all had yellow and round seeds. When these F1 plants were self fertilized they produced four types of plants in the following proportion:
(i) Yellow and round :9
(ii) Yellow and wrinkled : 3
(iii) Green and round :3
(iv) Green and wrinkled :1
The above results indicate that yellow and green seeds appear in the ratio of 9 + 3 : 3 + 1 = 3 : 1.
Similarly, the round and wrinkled seeds appear in the ratio of 9 + 3 : 3 +1 = 12:4 or 3 :1. This indicates that each of the two pairs of alternative characters viz. yellow-green cotyledon colour is inherited independent of the round-wrinkled character of the cotyledons. It means at the time of gamete formation the factor for yellow colour enters the gametes independent of R or r, i.e, gene Y can be passed on to the gametes either with gene R or r.
Cytological explanation of the results: In the above experiment yellow and round characters are

dominant over green and wrinkled characters which can be represented as follows:
(i) gene for yellow colour of cotyledons Y
(ii) gene for green colour of cotyledons y
(iii) gene for round character of cotyledons R
(iv) gene for wrinkled character of colyledons r
Therefore, plants with yellow and round cotyledons will have their genotype YYRR and those with
green and wrinkled cotyledons will have a genotype yyrr. These plants will produce gametes with gene YR and yr respectively. When these plants are cross pollinated, the union of these gametes will produce F1 hybrids with YyRr genes. When these produce gametes all the four genes have full freedom to assort independently and, therefore, there are possibilities of four combinations in both male and female gametes.
(i) RY (ii) Ry (iii) rY (iv) ry
This shows an excellent example of independent assortment. These gametes can unite at random producing in all 16 different combinations of genes, but presenting four phenotypes in the ratio of 9: 3: 3: 1.
Dihybrid ratio : RR yy - Round, yellow seeded ; Rr yy - Wrinkled and greed seeded
Test cross F1 Rr Yy x rr yy (recessive)1:1:1:1
Critical appreciation of law of Independent Assortment-
The law of independent assortment fails to have a universal applicability. Cytological studies have revealed that only those allelomorphs assort independently during meiosis, which are located in different homologous pairs of chromosomes. But, if the allelomorphs for different characters are present in the same homologous pair of chromosomes, these are passed on to the same gamete. Law of independent assortment does not apply to such cases.

MONOHYBRID CROSS
A cross is made between two true-breeding parents differing for a single trait, producing an F1 generation. These plants are intercrossed to produce an F2 generation.

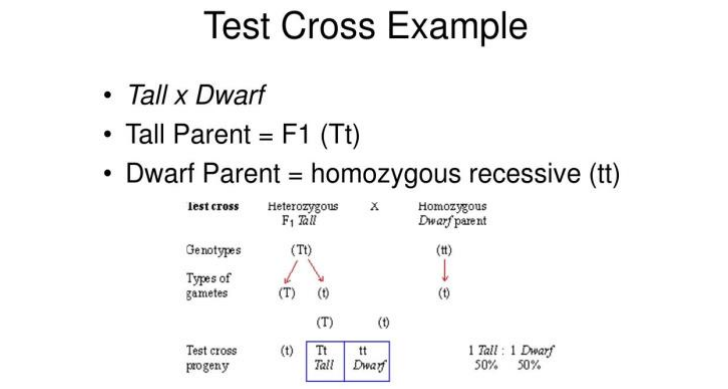


Test Cross:

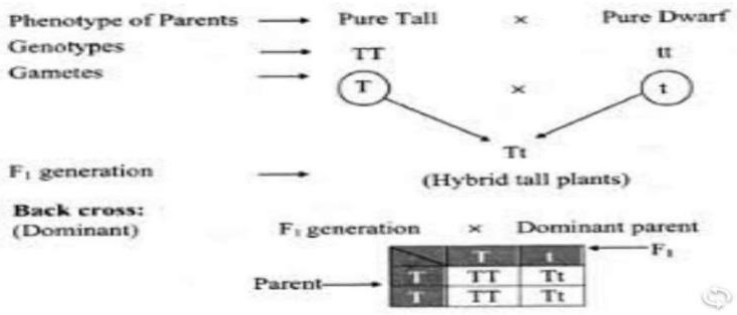
Under the law of dominance in genetics, an individual expressing a dominant phenotype could contain either two copies of the dominant allele (homozygous dominant) or one copy of each dominant and recessive allele (heterozygous dominant).[1] By performing a test cross, one can determine whether the individual is homozygous or heterozygous dominant.[1]

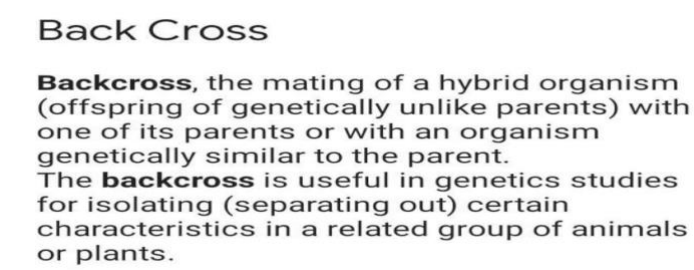
In a test cross, the individual in question is bred with another individual that is homozygous for the recessive trait and the offspring of the test cross are examined.[2] Since the homozygous recessive individual can only pass on recessive alleles, the allele the individual in question passes on determines the phenotype of the offspring.[3] Thus, this test yields 2 possible situations:

1. If any of the offspring produced express the recessive trait, the individual in question is heterozygous for the dominant allele.[1]
2. If all of the offspring produced express the dominant trait, the individual in question is homozygous for the dominant allele.



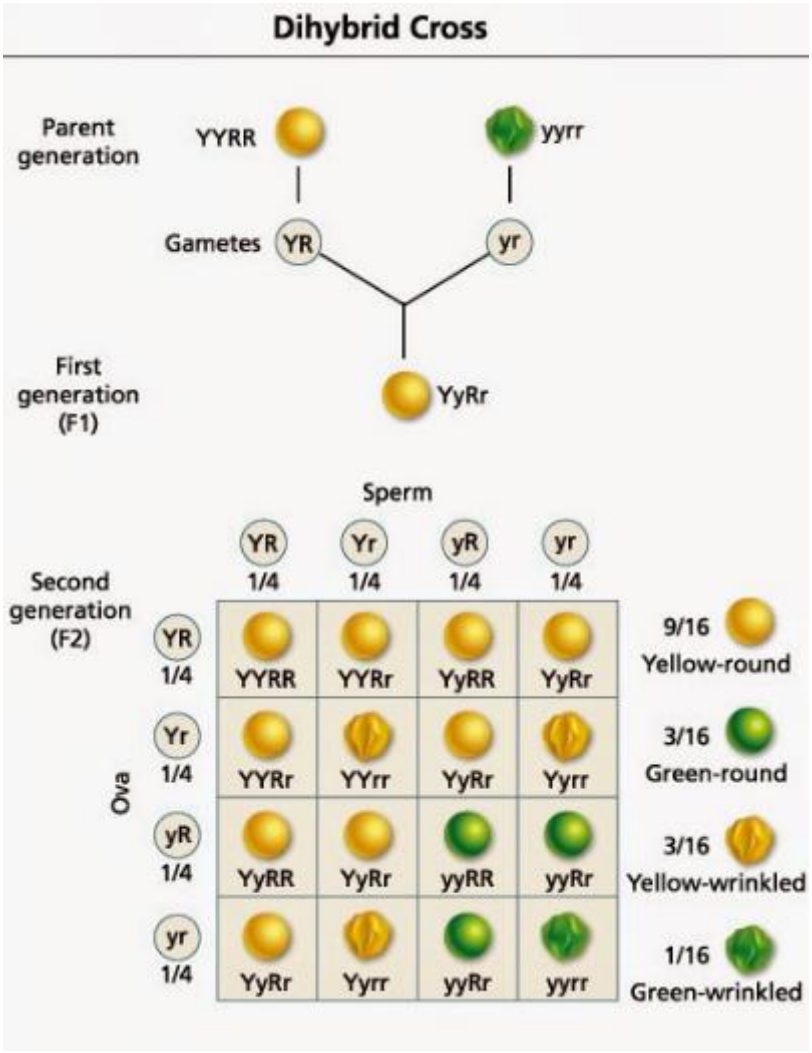
Back Cross:





Dihybrid Cross (9:3:3:1)

A classical case of two genes affecting the one and the same character and producing in the F2 four different phenotypes in the ratio of 9:3:3:1 was discovered in fowls by Bateson and Punnett. Each breed of poultry possesses characteristic type of comb. The Wyandotte breed has a comb known as the rose comb, the Brahma has a pea comb, and the leghorn has a single comb and the Malaya walnut comb. Each of these breeds true. Cross between rose comb and single combed types show that rose in dominant to single comb and that there is a segregation of 3 rose: 1 single comb in the F2. In mating between pea combed with single combed and 3:1 ratio appears in F2. In mating between pea combed with single combed bird, pea combed is found to be dominant over single comb and 3;1 ratio appears in F2. When a rose combed fowl is crossed with a pea combed one, all the F1 birds show a new comb know as walnut comb. When the walnut combs are inbred there appears in F2 walnut 3 rose pea single comb. As well in the ratio of 9:3:3:1. The rose comb is due to the presence of R gene and Pea due to P gene. Walnut comb is due to the presence of the dominant genes. R and P and single comb are due to the presence of recessive of r and p. The ratio expected in F2 is 9:3:3:1.



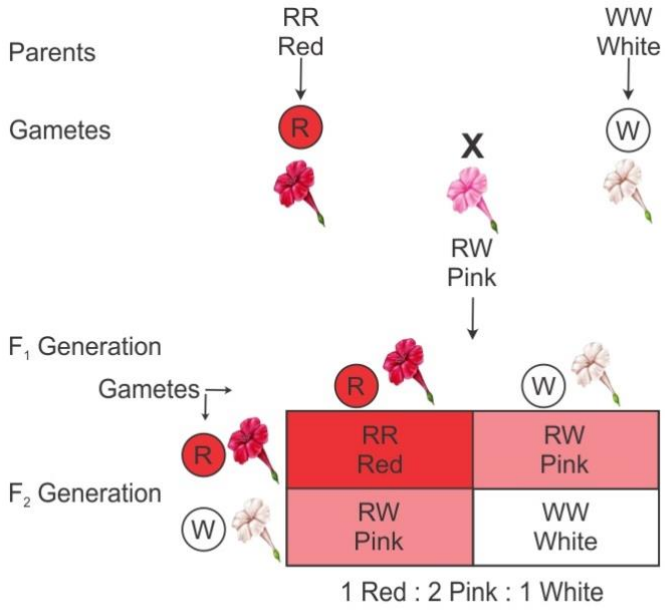
Genic Inteaction :

Some of the important forms of genic interactions are as follows:

1) Incomplete dominance
2) Complementary interaction of genes
3) Epistasis
4) Lethal alleles
5) Multiple Alleles

1. Incomplete dominance:

Mendel always observed complete dominance of one allele over the other for all the seven characters, which he studied, in garden pea. Later on cases of incomplete dominance were reported. For example, in four ëoí clock plant (Mirabilis jalapa) there are two types of flower viz., red and white. Generally, there are two phenotypic expressions of a particular character. For example, flower colour in four o'clock plant (*Mirabilis jalapa*) has two phenotypic expressions- red (dominant) and white (recessive). Only one of these is expressed- i.e., dominant phenotype in homozygous (AA) and heterozygous (Aa) conditions and recessive phenotype in only homozygous (aa) condition. However, in this case F1 (heterozygous) shows pink flowers which is intermediate between the dominant (red colour of flower) and the recessive (white colour of flower). This type of inheritance where heterozygote shows an intermediate phenotype, is termed as **incomplete dominance**. Further, in F2 generation both phenotypic and genotypic ratios are also similar, i.e.,1 red (RR): 2 pink (Rr): 1 white (rr).



Incomplete dominance in *Mirabilis jalapa*

A. COMPLEMENTARY GENE INTERACTION (9:7)

Ex: Flower color in *Lathyrus odoratus* (Sweet pea)

- Complementation between two non-allelic genes (C and P) are essential for production of a particular or special phenotype i.e., complementary factor.
- Two genes involved in a specific pathway and their functional products are required for gene expression, then one recessive allelic pair at either allelic pair would result in the mutant phenotype.
- When Dominant alleles are present together, they complement each other to yield complementary factor resulting in a special phenotype.
- They are called complementary genes.
- When either of gene loci have homozygous recessive alleles (i.e., genotypes of ccPP, ccPp, CCpp, Ccpp and ccpp), they produce identical phenotypes and change F2 ratio to 9:7.

Example: Flower color in *Lathyrus odoratus* (Sweet pea)

- In sweet pea (*Lathyrus odoratus*) two varieties of white flowering plants were seen.
- Each variety bred true and produced white flowers in successive generations.
- According to Bateson & Punnett, when two such white varieties of sweet pea were crossed, the offspring were found to have purple-coloured flowers in F1.
- But, in F2 generation 9 were purple and 7 white.
- Anthocyanin pigment synthesis in sweet pea (Lathyrus Odoratus).

PATHWAY OF ANTHOCYANIN PIGMENT SYNTHESIS

- The dominant allele or alleles [CC or Cc] of gene C are responsible for the presence of chromogen, while the homozygous recessive [cc] alleles of this gene are responsible for the absence of chromogen.
- Likewise, the dominant alleles of gene P in homozygous [PP] or heterozygous [Pp] conditions result in the production of an enzyme which is necessary for Anthocyanin (Complementary factor) from chromogen, while homozygous recessive [pp] condition does not produce any such enzyme.
- Thus, only the double dominant genotype has both enzymes functional and can make pigment.
- Blocking either of two steps prevents pigment formation.

P White flowers CCpp × White flowers ccPP

Gametes Cp, cP

Purple flowers CcPp

Gametes : CP.Cp, cP, cp Gametes

♀ \ ♂	CP	Cp	cP	cp
CP	CCPP Purple	CCPp Purple	CcPP Purple	CcPp Purple
Cp	CCPp Purple	CCpp White	CcPp Purple	Ccpp White
cP	CcPP Purple	CcPp Purple	ccPP White	ccPp White
cp	CcPp Purple	Ccpp Whie	ccPp White	ccpp White

Gametes

Phenotypic ratio : 9 Purple : 7 White

C. EPISTASIS GENE INTERACTION

- Epistasis is a Greek word that means standing over.
- BATESON used term epistasis to describe the masking effect in 1909
- The term epistasis describes a certain relationship between genes, where an allele of one gene hides or masks the visible output or phenotype of another gene.
- When two different genes which are not alleles, both affect the same character in such a way that the expression of one masks (inhibits or suppresses) the expression of the other gene, the phenomenon is said to be epistasis.
- The gene that suppresses other gene expression is known as **Epistatic gene**.
- The gene that is suppressed or remain obscure is called **Hypostatic gene**
- The classical phenotypic ratio of 9:3:3:1 F2 ratio becomes modified by epistasis.

Difference between Dominance and Epistasis		
	DOMINANCE	EPISTASIS
1.	Involves intra-allelic gene interaction	Involves inter-allelic gene interaction
2.	One allele hides the effect of other allele of the same gene	One allele hides the effect of other allele of the different gene

Epistasis is of two main types

i. Dominant Epistasis
ii. Recessive Epistasis

C {i} DOMINANT EPISTASIS (12:3:1)
Ex: Fruit Color in *Cucurbita pepo*

- When out of the two genes, the dominant allele (Example: A) of one gene masked the activity of alleles of another gene (Example: B), and expressed itself phenotypically, then A gene locus is said to be Epistatic to B gene locus.
- Because, the dominant allele A can express itself in the presence of either B or b allele, therefore, such type of epistasis is termed as dominant epistasis.
- The alleles of hypostatic locus or gene B will be able to express themselves phenotypically only when gene A locus may contain two recessive allele.
- The dominant epistasis modify the classical ratio of 9:3:3:1 into 12:3:1

Ex: FRUIT COLOR IN *Cucurbita pepo* (Summer squash)

- In fruit color in *Cucurbita Pepo*, commonly known as summer squash, is a standard example of dominant epistasis.
- **T**here are three types of fruit colors in this cucumber, viz., **white**, **yellow and green.**

White colour is controlled by dominant gene W	
Yellow colour by dominant gene Y.	
White is dominant over both yellow and green	

- The gene for white-colored squash is dominant to colored squash, and the gene symbols are W=White and w=colored.
- The gene for yellow-colored squash is dominant to green, and the gene symbols used are Y= yellow, y= green.
- Gene Y in homozygous or heterozygous condition converts green to yellow fruit color.
- Homozygous recessive yy cannot convert green to yellow, therefore results in green fruit color.

Parental Phenotype : **White fruit color** **X** **Yellow fruit color**

Parental Genotype : **WWyy** **X** **wwYY**

Parental Gametes : **Wy** **X** **wY**

F1 Generation : **WwYy** **White fruit color**

F1 selfing: F1 X F1 : **WwYy** **White fruit color** **X** **WwYy** **White fruit color**

F1 Gametes : WY Wy wY wy X WY Wy wY wy

F2 generation :

	WY	**Wy**	**wY**	**Wy**
WY	**WWYY White**	**WWYy White**	**WwYY White**	**WwYy White**
Wy	**WWYy White**	**WWyy White**	**WwYy White**	**Wwyy White**
wY	**WwYY White**	**WwYy White**	**wwYY Yellow**	**wwYy Yellow**
Wy	**WwYy White**	**Wwyy White**	**wwYy Yellow**	**wwyy Green**

Phenotypic Ratio = 12 White: 3 Yellow: 1 Green

F2 Analysis :

Genotype	Fruit color	Gene Actions
9W_Y_	White	Dominant white allele suppress effects of Y allele
3W_yy	White	Dominant white allele Suppress effect of y allele
3wwY_	Yellow	Recessive 'w' (color) allele allow yellow allele expression
1wwyy	Green	Recessive 'wy' allele allows green allele expression

- Gene W codes for inhibitor enzyme that stops conversion white to green.
- So, When Gene is homozygous or heterozygous [WW/Ww] fruit color is white
- Only in homozygous ww (absence of inhibitor) condition, it is colored (green or yellow)
- Therefore, if a Dihybrid is selfed, three phenotypes are produced in the ratio 12:3:1.
- Because the presence of the dominant W allele masks the effects of either the G or g allele, this type of interaction is called dominant epistasis

Other Examples of Dominant Epistasis are;

- Coat color in dogs
- Color of the hull in oats seeds
- Plumage color in poultry

Lethal Alleles:

Lethal alleles (also referred to as **lethal genes** or **lethals**) are alleles that cause the death of the organism that carries them. They are usually a result of mutations in genes that are essential for growth or development.[1] Lethal alleles may be recessive, dominant, or conditional depending on the gene or genes involved. Lethal alleles can cause death of an organism prenatally or any time after birth, though they commonly manifest early in development.

Recessive lethals

A pair of identical alleles that are both present in an organism that ultimately results in death of that organism are referred to as recessive lethal alleles. Though recessive lethals may code for dominant or recessive traits, they are only fatal in the homozygous condition. Heterozygotes will sometimes display a form of diseased phenotype, as in the case of achondroplasia. One mutant lethal allele can be tolerated, but having two results in death. In the case of homozygous achondroplasia, death almost invariably occurs before birth or in the perinatal period. Not all heterozygotes for recessive lethal alleles will show a mutant phenotype, as is the case for cystic fibrosis carriers. If two cystic fibrosis carriers have children, they have a 25 percent chance of producing offspring having two copies of the lethal allele, eventually resulting in the death of the child.

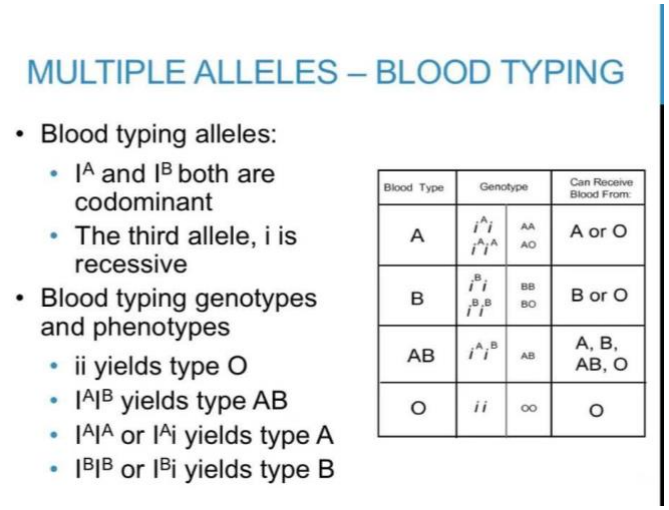
Another example of a recessive lethal allele occurs in the Manx cat. Manx cats possess a heterozygous mutation resulting in a shortened or missing tail. Crosses of two heterozygous Manx cats result in two-thirds of surviving offspring displaying the heterozygous shortened tail phenotype, and one-third of surviving offspring of normal tail length that is homozygous for a normal allele. Homozygous offspring for the mutant allele cannot survive birth and are therefore not seen in these crosses.

Dominant lethals

Alleles that need only be present in one copy in an organism to be fatal are referred to as dominant lethal alleles. These alleles are not commonly found in populations because they usually result in the death of an organism before it can transmit its lethal allele on to its offspring.[4] An example in humans of a dominant lethal allele is Huntington's disease, a rare neurodegenerative disorder that ultimately results in death. However, because of its late-onset (i.e., often after reproduction has already occurred), it is able to be maintained in populations. A person exhibits Huntington's disease when they carry a single copy of a repeat-expanded Huntington allele on chromosome 4.

Multiple alleles:

Multiple alleles exist in a population when there are many variations of a gene present. In organisms with two copies of every gene, also known as diploid organisms, each organism has the ability to express two **alleles** at the same time. They can be the same **allele**, which is called a homozygous genotype.



A **blood type** (also known as a **blood group**) is a classification of blood, based on the presence and absence of antibodies and inherited antigenic substances on the surface of red blood cells (RBCs). These antigens may be proteins, carbohydrates, glycoproteins, or glycolipids, depending on the blood group system. Some of these antigens are also present on the surface of other types of cells of various tissues. Several of these red blood cell surface antigens can stem from one allele (or an alternative version of a gene) and collectively form a blood group system.[1]

Blood types are inherited and represent contributions from both parents. As of 2019, a total of 38 human blood group systems are recognized by the International Society of Blood Transfusion (ISBT).[2] The two most important blood group systems are ABO and Rh; they determine someone's blood type (A, B, AB, and O, with +, − or null denoting RhD status) for suitability in blood transfusion.

ABO blood group system

The ABO blood group system involves two antigens and two antibodies found in human blood. The two antigens are antigen A and antigen B. The two antibodies are antibody A and antibody B. The antigens are present on the red blood cells and the antibodies in the serum. Regarding the antigen property of the blood all human beings can be classified into 4 groups, those with antigen A (group A), those with antigen B (group B), those with both antigen A and B (group AB) and those with neither antigen (group O). The antibodies present together with the antigens are found as follows:

1. Antigen A with antibody B
2. Antigen B with antibody A
3. Antigen AB has no antibodies
4. Antigen nil (group O) with antibody A and B.

There is an agglutination reaction between similar antigen and antibody (for example, antigen A agglutinates the antibody A and antigen B agglutinates the antibody B). Thus, transfusion can be considered safe as long as the serum of the recipient does not contain antibodies for the blood cell antigens of the donor.

The ABO system is the most important blood-group system in human-blood transfusion. The associated anti-A and anti-B antibodies are usually immunoglobulin M, abbreviated IgM, antibodies. It has been hypothesized that ABO IgM antibodies are produced in the first years of life by sensitization to environmental substances such as food, bacteria, and viruses, although blood group compatibility rules are applied to newborn and infants as a matter of practice.[10] The original terminology used by Karl Landsteiner in 1901 for the classification was A/B/C; in later publications "C" became "O".[11] Type O is often called *0* (*zero*, or *null*) in other languages.

GENE

- A gene is the basic physical and functional unit of heredity that controls a particular trait of an organism.

ALLELE

- An Allele is an alternative form of a same gene, located on same locus on the homologous chromosomes.
- Alleles were first defined by Gregor Mendel in the law of segregation.

NOTE:

- Most genes have two alleles, a dominant allele and a recessive allele. For example: tall (dominant) and dwarf (recessive).
- If an organism is heterozygous for that trait (possesses one of each allele), then usually the dominant trait is expressed. A recessive allele is only expressed if an organism is homozygous for that trait (possesses two recessive alleles).
- Although individual humans (and all diploid organisms) can only have two alleles for a given gene, multiple alleles may exist at the population level.

2.1 MULTIPLE ALLELES

- According to Altenburg, "Three or more kinds of gene which occupy the same locus are referred to as multiple alleles."
- Multiple alleles are defined as three or more alternative form of a same gene, located on same locus on the homologous chromosomes, coding for certain characteristic in a population.
- A gene controlled by more than two alleles and following Non-Mendelian pattern of inheritance and is described as MULTIPLE ALLELISM.

Examples;

1. ABO blood groups
2. Rh factor in Human
3. Coat Colour in Rabbit
4. Eye color in *Drosophila*

2.2 ABO BLOOD GROUPS

- ABO blood group in Human was discovered by LANDSTEINER in 1901.
- The ABO system is characterized by the presence or absence of antigens on the surface of Red Blood Cells.
- Individuals will naturally develop antibodies against the ABO antigens they do not have.
- For example, individuals with blood group A will have anti-B antibodies, and individuals with blood group O will have both anti-A and anti-B.

RBC Surface proteins (Ag)	Plasma Antibody (Ab)	Blood Type
A	b	A
B	a	B
AB	--	AB
--	a & b	O

- The ABO phenotype of any individual is ascertained by mixing a blood sample with antiserum containing type A or type B antibodies. If the antigen is present on the surface of Red Blood Cells of the person, then it will react with the corresponding antibody and cause clumping or agglutination of the Red Blood Cells.

GENETICS OF ABO BLOOD SYSTEM

- The ABO blood type is inherited in an Autosomal Co-Dominant fashion.
- The ABO locus is located on **chromosome 9** at 9q34.1-q34.2 consisting 18 kb of genomic DNA (Exon 7).
- The A and B alleles differ from each other by seven nucleotide substitutions
- The gene controlling ABO blood type is labeled as I.
- The alleles are designated as I^A, I^B and I^O (or i).
  Where, I → Isoagglutinogen (antigen)
  I^A → Allele for A antigen
  I^B → Allele for B antigen
  I^O → Allele for o antigen
- Dominance Hierarchy among alleles; $\mathbf{I^A = I^B > I^O}$
  i.e., Alleles I^A & I^B are dominant over I^O
  Alleles I^A & I^B are Co-dominant

Genotype	Antigen on RBC	Blood type
$I^A I^A$/ $I^A I^O$	A	A
$I^B I^B$/ $I^B I^O$	B	B
$I^A I^B$	A & B	AB
$I^O I^O$	--	O

- Further, studies suggest that I^A allele may occur in at least FOUR allelic forms; I^{A1}, I^{A2}, I^{A3} & I^{A4}
- Thus, Dominance Hierarchy among 6 alleles; $\mathbf{I^A}[I^{A1} > I^{A2} > I^{A3} > I^{A4}] = \mathbf{I^B > I^O}$

Note:
Recent data reports over 80 ABO alleles.

- The blood groups are defined by the presence of specific carbohydrate sugars [oligosaccharide chains] on the surface of red blood cells.
- The specificity of A and B antigens are based on the terminal sugars of carbohydrate group [i.e., precursor molecule - H antigen]
- The H locus is located on chromosome 19 at 19q13.3 (>5 kb of genomic DNA, three exons), and it encodes a fucosyltransferase that produces the H antigen on RBCs.
- H-antigen consists of 3 sugar molecules; galactose (Gal), N-acetylglucosamine (GlcNAc) and fucose (Fuc)
- I^A and I^B alleles each encode a specific glycosyl-transferring enzyme, which catalyzes the final step in the synthesis of the A and B antigen.
- The I^A allele encodes a glycosyltransferase (i.e., N-acetylgalactose transferase) that produces the A antigen (by adding terminal N-acetylgalactosamine - immunodominant sugar)
- The I^B allele encodes a glycosyltransferase (i.e., galactose transferase) that creates the B antigen (by adding terminal D-galactose - immunodominant sugar).
- The I^O allele encodes an enzyme with no function (*might be, not yet discovered*), and therefore neither A or B antigen is produced, leaving the underlying precursor (the H antigen) unchanged.

GENE

Definition:

"A functional segment of DNA which manufacture protein, regulate gene expression and renowned as a hereditary unit is known as a gene."

History:

Mendel first discovered the concept of the "inheritance of traits", despite, he fails to describe it. The term "gene" was coined and studied by Wilhelm Johannson. But he was unable to describe the chemical structure of it. In 1953, James Watson and F Crick defined the chemical structure of the DNA viz gene.

Structure of gene:

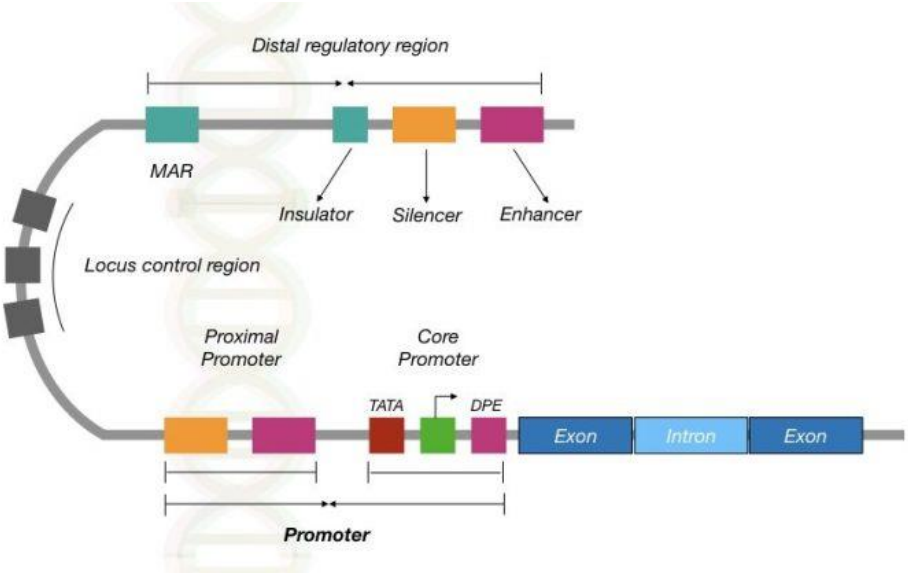
Genes are actually DNA strands thus are made up of the nucleotide chain. The chemical structure of a gene comprises nucleotides. A part of DNA- genes are made up of A, T, G and C nucleotides. With the nucleotides of the opposite strand, it binds with hydrogen bonds and with the adjacent nucleotide, it binds with phosphodiester bonds.

The nucleotides are the combination of nitrogenous bases (A, T, G and C), phosphate and pentose sugar. In general, the gene structure consists of two types of elements: core elements and regulatory elements. The core elements or sequences actually take parts in protein formation. While the regulatory elements maintain gene expression.

Exons are core elements. Sequences on the other side like promoters, enhancers and silencers are regulatory elements of a gene. The third type of element called maintenance elements possesses information for DNA repair, modification and replication. The functional or physical structure of a gene comprises introns, exons, promotes, enhancer and UTRs.

Introns are intervening non-coding sequences removed from the final transcript.

Exons are coding part of a gene which are joined after splicing and constructs the final transcript.



Promotes are non-coding sequences but facilitates binding sites for enzymes and transcriptional factors to work. The promoter consists of TATA box and CCAAT sequences for enzyme binding.

The entire promoter region is located on the 5' end and made up of core promoter and proximal promoter sequences (see the above image). Here, the core promoter facilitates RNA polymerase bindings (and other proteins) to start transcription. While the proximal promoter provides bindings for transcriptional factors.

The enhancer induces transcription while the silencer represses it. Collectively, enhancers and silencers located far away from exon, regulate gene expression. The 3' untranslated regions are non-coding regions of gene helps in aborting the process of transcription and to form the final transcript.

Once the RNA polymerase reaches the untranslated region it stops synthesizing RNA and detached from the strand. The eukaryotic gene structure consists of more regulatory sequences than prokaryotic genes. In addition to this, the entire machinery of transcription and translation is different in both.

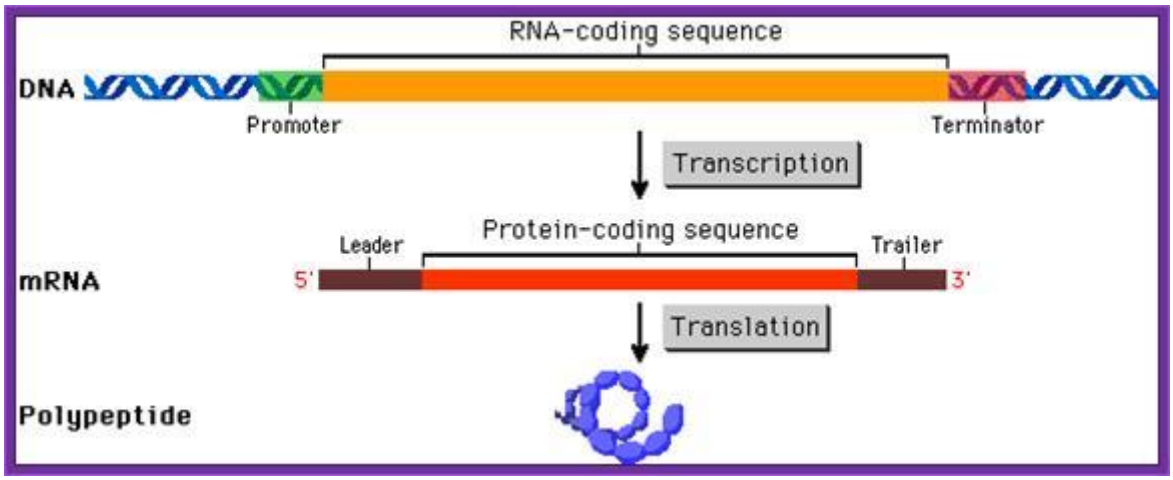
The operon concept of prokaryotic genes consists of a gene cluster of similar function. Introns are not a part of an operon. Contrary, the eukaryotic genes consist of introns (non-coding DNA) at regular intervals. Each and every gene have their own promoter region to facilitate transcription.

Prokaryotic genes

The overall organisation of prokaryotic genes is markedly different from that of the eukaryotes. The most obvious difference is that prokaryotic ORFs are often grouped into a polycistronic operon under the control of a shared set of regulatory sequences. These ORFs are all transcribed onto the same mRNA and so are co-regulated and often serve related functions. Each ORF typically has its own ribosome binding site (RBS) so that ribosomes simultaneously translate ORFs on the same mRNA. Some operons also display

translational coupling, where the translation rates of multiple ORFs within an operon are linked. This can occur when the ribosome remains attached at the end of an ORF and simply translocates along to the next without the need for a new RBS. Translational coupling is also observed when translation of an ORF affects the accessibility of the next RBS through changes in RNA secondary structure. Having multiple ORFs on a single mRNA is only possible in prokaryotes because their transcription and translation take place at the same time and in the same subcellular location.

The operator sequence next to the promoter is the main regulatory element in prokaryotes. Repressor proteins bound to the operator sequence physically obstructs the RNA polymerase enzyme, preventing transcription.[29][30] Riboswitches are another important regulatory sequence commonly present in prokaryotic UTRs. These sequences switch between alternative secondary structures in the RNA depending on the concentration of key metabolites. The secondary structures then either block or reveal important sequence regions such as RBSs. Introns are extremely rare in prokaryotes and therefore do not play a significant role in prokaryotic gene regulation.

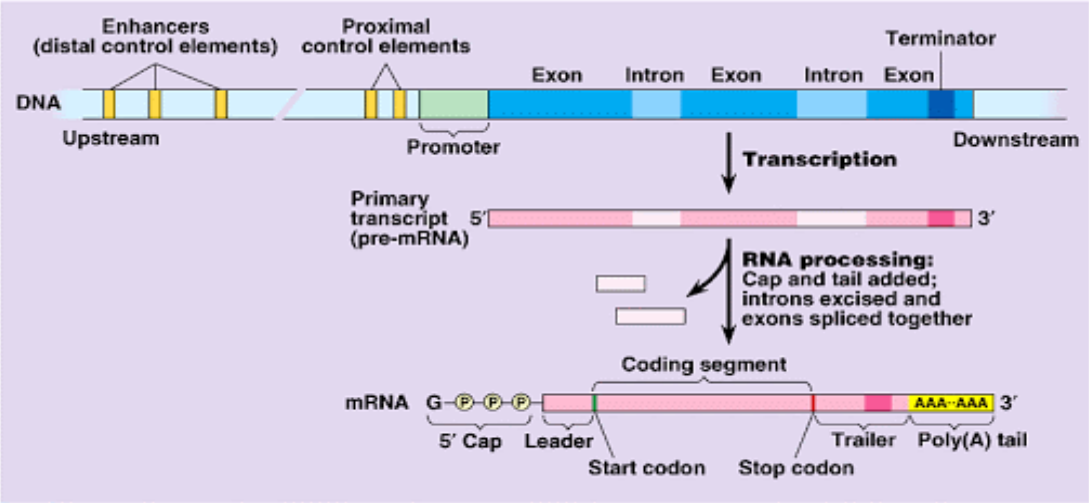


Eukaryotic genes

The structure of eukaryotic genes includes features not found in prokaryotes. Most of these relate to post-transcriptional modification of pre-mRNAs to produce mature mRNA ready for translation into protein. Eukaryotic genes typically have more regulatory elements to control gene expression compared to prokaryotes.[5] This is particularly true in multicellular eukaryotes, humans for example, where gene expression varies widely among different tissues.

A key feature of the structure of eukaryotic genes is that their transcripts are typically subdivided into exon and intron regions. Exon regions are retained in the final mature mRNA molecule, while intron regions are spliced out (excised) during post-transcriptional processing. Indeed, the intron regions of a gene can be considerably longer than the exon regions. Once spliced together, the exons form a single continuous protein-coding region, and the splice boundaries are not detectable. Eukaryotic post-transcriptional processing also adds a 5' cap to the start of the mRNA and a poly-adenosine tail to the end of the mRNA. These additions stabilise the mRNA and direct its transport from

the nucleus to the cytoplasm, although neither of these features are directly encoded in the structure of a gene.



Part of a Gene Can Function:

It was considered earlier that gene is the basic unit of function and parts of gene, if exist, cannot function. But this concept has been outdated now. Based on studies on rll locus of T4 phage, Benzer (1955) concluded that there are three sub divisions of a gene, viz., recon, muton and cistron.

These are briefly described below:

a. Recon:

Recons are the regions (units) within a gene between which recombination's can occur, but the recombination cannot occur within a recon. There is a minimum recombination distance within a gene which separates recons. The map of a gene is completely linear sequence of recons.

b. Muton :

It is the smallest element within a gene, which can give rise to a mutant phenotype or mutation. This indicates that part of a gene can mutate or change. This disproved the bead theory according to which the entire gene was to mutate or change.

c. Cistron:

It is the largest element within a gene which is the unit of function. This also knocked down the bead theory according to which entire gene was the unit of function. The name cistron has been derived from the test which is performed to know whether two mutants are within the same cistron on in different cistrons. It is called cis-trans test which is described below.

GENETIC CODE

The **genetic code** is the set of rules used by living cells to translate information encoded within genetic material (DNA or mRNA sequences of nucleotide triplets, or **codons**) into proteins. Translation is accomplished by the ribosome, which links proteinogenic amino acids in an order specified by messenger RNA (mRNA), using transfer RNA (tRNA) molecules to carry amino acids and to read the mRNA three nucleotides at a time. The genetic code is highly similar among all organisms and can be expressed in a simple table with 64 entries.

The code defines how codons specify which amino acid will be added next during protein synthesis. With some exceptions, a three-nucleotide codon in a nucleic acid sequence specifies a single amino acid. The vast majority of genes are encoded with a single scheme (see the RNA codon table). That scheme is often referred to as the canonical or standard genetic code, or simply *the* genetic code, though variant codes (such as in human mitochondria) exist.

While the "genetic code" is what determines a protein's amino acid sequence, other genomic regions determine when and where these proteins are produced according to various "gene regulatory codes".

History

Efforts to understand how proteins are encoded began after DNA's structure was discovered in 1953. George Gamow postulated that sets of three bases must be employed to encode the 20 standard amino acids used by living cells to build proteins, which would allow a maximum of $4^3 = 64$ amino acids.

Codons

The Crick, Brenner, Barnett and Watts-Tobin experiment first demonstrated that codons consist of three DNA bases. Marshall Nirenberg and Heinrich J. Matthaei were the first to reveal the nature of a codon in 1961.

They used a cell-free system to translate a poly-uracil RNA sequence (i.e., UUUUU...) and discovered that the polypeptide that they had synthesized consisted of only the amino acid phenylalanine. They thereby deduced that the codon UUU specified the amino acid phenylalanine.

This was followed by experiments in Severo Ochoa's laboratory that demonstrated that the poly-adenine RNA sequence (AAAAA...) coded for the polypeptide poly-lysine and that the poly-cytosine RNA sequence (CCCCC...) coded for the polypeptide poly-proline. Therefore, the codon AAA specified the amino acid lysine, and the codon CCC specified the amino acid proline. Using various copolymers most of the remaining codons were then determined.

Subsequent work by Har Gobind Khorana identified the rest of the genetic code. Shortly thereafter, Robert W. Holley determined the structure of transfer RNA (tRNA), the adapter molecule that facilitates the process of translating RNA into protein. This work was based upon Ochoa's earlier studies, yielding the latter the Nobel Prize in Physiology or Medicine in 1959 for work on the enzymology of RNA synthesis.

Extending this work, Nirenberg and Philip Leder revealed the code's triplet nature and deciphered its codons. In these experiments, various combinations of mRNA were passed through a filter that contained ribosomes, the components of cells

that translate RNA into protein. Unique triplets promoted the binding of specific tRNAs to the ribosome. Leder and Nirenberg were able to determine the sequences of 54 out of 64 codons in their experiments. Khorana, Holley and Nirenberg received the 1968 Nobel for their work.[10]

The three stop codons were named by discoverers Richard Epstein and Charles Steinberg. "Amber" was named after their friend Harris Bernstein, whose last name means "amber" in German. The other two stop codons were named "ochre" and "opal" in order to keep the "color names" theme.

Features of Genetic Code

Reading frame

A reading frame is defined by the initial triplet of nucleotides from which translation starts. It sets the frame for a run of successive, non-overlapping codons, which is known as an "open reading frame" (ORF). For example, the string 5'-AAATGAACG-3' (see figure), if read from the first position, contains the codons AAA, TGA, and ACG; if read from the second position, it contains the codons AAT and GAA ; and if read from the third position, it contains the codons ATG and AAC.

Every sequence can, thus, be read in its 5' → 3' direction in three reading frames, each producing a possibly distinct amino acid sequence: in the given example, Lys (K)-Trp (W)-Thr (T), Asn (N)-Glu (E), or Met (M)-Asn (N), respectively (when translating with the vertebrate mitochondrial code).

When DNA is double-stranded, six possible reading frames are defined, three in the forward orientation on one strand and three reverse on the opposite strand. Protein-coding frames are defined by a start codon, usually the first AUG (ATG) codon in the RNA (DNA) sequence. In eukaryotes, ORFs in exons are often interrupted by introns.

Start/stop codons

Translation starts with a chain-initiation codon or start codon. The start codon alone is not sufficient to begin the process. Nearby sequences such as the Shine-Dalgarno sequence in *E. coli* and initiation factors are also required to start translation. The most common start codon is AUG, which is read as methionine or, in bacteria, as formylmethionine. Alternative start codons depending on the organism include "GUG" or "UUG"; these codons normally represent valine and leucine, respectively, but as start codons they are translated as methionine or formylmethionine.

The three stop codons have names: UAG is *amber*, UGA is *opal* (sometimes also called *umber*), and UAA is *ochre*. Stop codons are also called "termination" or "nonsense" codons. They signal release of the nascent polypeptide from the ribosome because no cognate tRNA has anticodons complementary to these stop signals, allowing a release factor to bind to the ribosome instead.

Genetic Codon Chart

	U		C		A		G		
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys	U
	UUC	Phe	UCC	Ser	UAC	Tyr	UGC	Cys	C
	UUA	Leu	UCA	Ser	UAA	Stop	UGA	Stop	A
	UUG	Leu	UCG	Ser	UAG	Stop	UGG	Trp	G
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	U
	CUC	Leu	CCC	Pro	CAC	His	CGC	Arg	C
	CUA	Leu	CCA	Pro	CAA	Gln	CGA	Arg	A
	CUG	Leu	CCG	Pro	CAG	Gln	CGG	Arg	G
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser	U
	AUC	Ile	ACC	Thr	AAC	Asn	AGC	Ser	C
	AUA	Ile	ACA	Thr	AAA	Lys	AGA	Arg	A
	AUG	Met	ACG	Thr	AAG	Lys	AGG	Arg	G
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly	U
	GUC	Val	GCC	Ala	GAC	Asp	GGC	Gly	C
	GUA	Val	GCA	Ala	GAA	Glu	GGA	Gly	A
	GUG	Val	GCG	Ala	GAG	Glu	GGG	Gly	G

Translation START codon

Translation STOP codon

Positively charged amino acids

Negatively charged amino acids

Hydrophobic amnio acids

Hydrophilic non-charged amino acids

Cysteine

DNA AND RNA

STRUCTURE, FUNCTION, TYPES, MODES OF REPLICATION AND REPAIR

The discovery that DNA is the prime genetic molecule, carrying all the hereditary information within chromosomes, immediately had its attention focused on its structure. It was hoped that knowledge of the structure would reveal how DNA carries the genetic messages that are replicated when chromosomes divide to produce two identical copies of themselves. During the late 1940s and early 1950s, several research groups in the United States and in Europe engaged in serious efforts— both cooperative and rival—to understand how the atoms of DNA are linked together by covalent bonds and how the resulting molecules are arranged in three-dimensional space. Not surprisingly, it was feared that DNA might have very complicated and perhaps bizarre structures that differed radically from one gene to another. Great relief, if not general elation, was thus expressed when the fundamental DNA structure was found to be the double helix. It told us that all genes have roughly the same three-dimensional form and that the differences between two genes reside in the order and number of their four nucleotide building blocks along the complementary strands.

What is DNA?

The work of many scientists paved the way for the exploration of DNA. Way back in 1868, almost a century before the Nobel Prize was awarded to Watson, Crick and Wilkins, a young Swiss physician named Friedrich Miescher, isolated something no one had ever seen before from the nuclei of cells. He called the compound "nuclein." This is today called nucleic acid, the "NA" in DNA (deoxyribo-nucleic-acid) and RNA (ribo- nucleic-acid)

. Two years earlier, the Czech monk Gregor Mendel, had finished a series of experiments with peas. His observations turned out to be closely connected to the finding of nuclein. Mendel was able to show that certain traits in the peas, such as their shape or colour, were inherited in different packages. These packages are what we now call genes.

For a long time the connection between nucleic acid and genes was not known. But in 1944 the American scientist Oswald Avery managed to transfer the ability to cause disease from one strain of bacteria to another. But not only that: the previously harmless bacteria could also pass the trait along to the next generation. What Avery had moved was nucleic acid. This proved that genes were made up of nucleic acid. **Solving the Puzzle**

In the late 1940's, the members of the scientific community were aware that DNA was most likely the molecule of life, even though many were skeptical since it was so "simple". They also knew that DNA included different amounts of the four bases adenine, thymine, guanine and cytosine (usually abbreviated A, T, G and C), but nobody had the slightest idea of what the molecule might look like.

In order to solve the elusive structure of DNA, a couple of distinct pieces of information needed to be put together. One was that the phosphate backbone was on the outside with bases on the inside; another that the molecule was a double helix. It was also important to figure out that the two strands run in opposite directions and that the molecule had a specific base pairing.

Watson and Crick

In 1951, the then 23-year old biologist James Watson travelled from the United States to work with Francis Crick, an English physicist at the University of Cambridge. Crick was already using the process of X- ray crystallography to study the structure of protein molecules. Together, Watson and Crick used X-ray crystallography data, produced by Rosalind Franklin and Maurice Wilkins at King's College in London, to decipher DNA's structure.

This is what they already knew from the work of many scientists, about the DNA molecule:

1. DNA is made up of subunits which scientists called nucleotides.
2. Each nucleotide is made up of a sugar, a phosphate and a base.
3. There are 4 different bases in a DNA molecule: adenine (a urine) cytosine (a pyrimidine) guanine (a purine) thymine (a pyrimidine)
4. The number of purine bases equals the number of pyrimidine bases
5. The number of adenine bases equals the number of thymine bases
6. The number of guanine bases equals the number of cytosine bases
7. The basic structure of the DNA molecule is helical, with the bases being stacked on top of each other

Components of DNA

DNA is a polymer. The monomer units of DNA are nucleotides, and the polymer is known as a "polynucleotide". Each nucleotide consists of a 5-carbon sugar (deoxyribose), a nitrogen containing base attached to the sugar, and a phosphate group. There are four different types of nucleotides found in DNA, differing only in the nitrogenous base. The four nucleotides are given one letter abbreviations as shorthand for the four bases.

- A is for adenine
- G is for guanine
- C is for cytosine
- T is for thymine

Purine Bases

Adenine and guanine are purines. Purines are the larger of the two types of bases found in DNA. Structures are shown below:

Pyrimidine Bases

Cytosine and thymine are pyrimidines. The 6 stoms (4 carbon, 2 nitrogen) are numbered 1-6. Like purines, all pyrimidine ring atoms lie in the same plane.

Deoxyribose Sugar

The deoxyribose sugar of the DNA backbone has 5 carbons and 3 oxygens. The carbon atoms are numbered 1', 2', 3', 4', and 5' to distinguish from the numbering of the atoms of the purine and pyrmidine rings. The hydroxyl groups on the 5'- and 3'- carbons link to the phosphate groups to form the DNA backbone. Deoxyribose lacks an hydroxyl group at the 2'-position when compared to ribose, the sugar component of RNA.

Nucleosides

A nucleoside is one of the four DNA bases covalently attached to the C1' position of a sugar. The sugar in deoxynucleosides is 2'- deoxyribose. The sugar in ribonucleosides is ribose. Nucleosides differ from nucleotides in that they lack phosphate groups. The four different nucleosides of DNA are deoxyadenosine (dA), deoxyguanosine (dG), deoxycytosine (dC), and (deoxy)thymidine (dT, or T).

Nucleotides

A nucleotide is a nucleoside with one or more phosphate groups covalently attached to the 3'- and/or 5'- hydroxyl group(s).

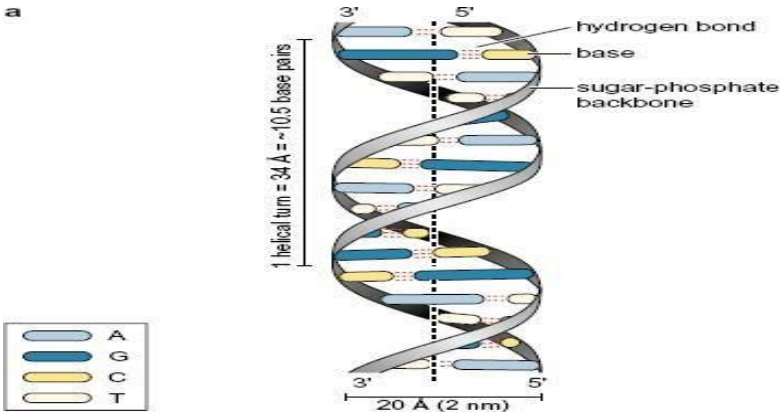
DNA Backbone

The DNA backbone is a polymer with an alternating sugar- phosphate sequence. The deoxyribose sugars are joined at both the 3'- hydroxyl and 5'-hydroxyl groups to phosphate groups in ester links, also known as "phosphodiester" bonds.

DNA Double Helix

DNA is a normally double stranded macromolecule. Two polynucleotide chains, held together by weak thermodynamic forces, form a DNA molecule.

Structure of DNA Double Helix

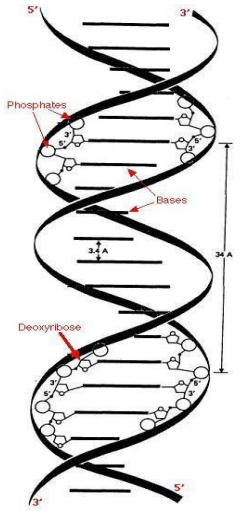


The Double Helix

The double helix of DNA has these features:

- It contains two polynucleotide strands wound around each other.
- The backbone of each consists of alternating deoxyribose and phosphate groups.
- The phosphate group bonded to the 5' carbon atom of one deoxyribose is covalently bonded to the 3' carbon of thenext.
- The two strands are "antiparallel"; that is, one strand runs 5′ to 3′ while the other runs 3′ to 5′.

- The DNA strands are assembled in the 5′ to 3′ direction and, by convention, we "read" them the same way.
- The purine or pyrimidine attached to each deoxyribose projects in toward the axis of the helix.
- Each base forms hydrogen bonds with the one directly opposite it, forming **base pairs** (also called nucleotide pairs).
- 3.4 Å separate the planes in which adjacent base pairs arelocated.
- The double helix makes a complete turn in just over 10 nucleotide pairs, so each turn takes a little more (35.7 Å to be exact) than the 34 Å shown in the diagram.
- There is an average of 25 hydrogen bonds within each complete turn of the double helix providing a stability of binding about as strong as what a covalent bond wouldprovide.
- The diameter of the helix is 20 Å.
- The helix can be virtually any length; when fully stretched, some DNA molecules are as much as 5 cm (2 inches!) long.
- The path taken by the two backbones forms a major (wider) groove (from "34 A" to the top of the arrow) and a minor (narrower) groove (the one below).



Nucleic acids (DNA and RNA) are the polymers i.e. long chain compounds. The molecular structure of DNA has two aspects

1) its chemical sub units and
2) the way in which these chemical sub units are arranged to form a long chain molecule.

The second aspect is very significant as the accepted DNA model should be such that it explains biochemically the various aspects (function) of a gene such as stability to metabolic and external agents, the capacity for replication (self duplication) the capacity to store vast hereditary information in coded form and the capacity to express the phenotypes they control.

FUNCTIONS OF DNA

DNA carries the genetic information of a cell and consists of thousands of genes. Each gene serves as a recipe on how to build a protein molecule. Proteins perform important tasks for the cell functions or serve as building blocks. The flow of information from the genes determines the protein composition and thereby the functions of thecell.

The DNA is situated in the nucleus, organized into chromosomes. Every cell must contain the genetic information and the DNA is therefore duplicated before a cell divides (**replication**). When proteins are needed, the corresponding genes are transcribed into RNA (**transcription**). The RNA is first processed so that non-coding parts are removed (**processing**) and is then transported out of the nucleus (**transport**). Outside the nucleus, the proteins are built based upon the code in the RNA (**translation**).

Types of DNA

DNA can be classified in various ways based on 1. number of base pair per turn. 2. coiling pattern, 3. location

4. structure, 5. nucleotide sequence and 6. number of strands.

1. Number of base per turn. Depending upon the nucleotide base per turn of the helix, tilt of the base pair and humidity of the sample, the DNA can be observed in four different forms namely A,B, C and D.

2. Coiling pattern. On the basis of coiling pattern of the helix DNA is of two types viz right handed and left handed. Most of the DNA molecules are right handed i.e. coiling of helix is in the right direction. It is also called positive coiling. All the four forms of DNA viz A, B, C and D are right handed. The Z DNA has left handed double helical structure. This DNA is considered to be associated with gene regulation.

3. Location. Based on the location in the cell DNA is of three types. Viz., chromosomal DNA cytoplasm DNA and promiscuous DNA. Chromosomal DNA is found in chromosomes. And are called as chromosomal DNA or nuclear DNA.

4. Cytoplasmic DNA is found in the cytoplasm especially in mitochondria and chloroplasts. Such DNA plays an important role in cytoplasmic inheritance and has circular structure. Promiscuous DNA.

5. Some DNA segments with common base sequence are found in the chloroplasts, mitochondria and nucleus. This suggests that some DNA sequences move from one organelle to other. Such DNA is referred to as promiscuous DNA.

6. Structure of RNA: It contains ribose sugar, nitrogen bases and phosphate group. The nitrogen bases include adenine, guanine, cytosine and uracil. In DNA thymine is present in place of uracil and deoxyribose sugar is found in place of ribose sugar. In RNA, the pairing occurs between adenine and uracil. It has usually single strand. However, some viruses have double stranded RNA.

The DNA molecule that Watson and Crick described was in the B form. It is now known that DNA can exist in several other forms. The primary difference between the forms is the direction that the helix spirals.

A, B, C = right-handed helix Z = left-handed helix (found in vitro under high salt)

B is the major form that is found in the cell. Z-DNA was initially found only under high salt conditions, but the cellular environment is actually a low-salt environment. The question then is whether type Z exist under cellular conditions. Several features have been discovered that can stablize Z-DNA under in a low salt environment.

Differences between DNA and RNA

S. No	**Particulars**	**DNA**	**RNA**
1.	Strands	Usually two, rarely one	Usually one, rarely two
2.	Sugar	Deoxyribose	Ribose
3.	Base	Adenine guanine cytosine and thymine	Adenine guanines cytosine
4.	Pairing	AT and GC	AU and GC
5.	Location	Mostly in chromosomes some in mitochondria and chloroplasts	In chromosomes and ribosomes

MODES OF REPLICATION

There are three possible modes of DNA replication:

(1) Dispersive

(2) Conservative

(3) Semiconservative

1. **In dispersive replication**, the old DNA molecule would break into several pieces, each fragment would replicate and the old and new sesgments would recombine randomly to yield progeny DNA molecules. Each progeny molecule would have both old and new segments along its length.

2. According to the **conservative scheme**, the two newly synthesized strands (following the replication of a DNA molecule) would associate to form one double helix,

while the two old strands would remain together as one double helix.

3. In contrast, in the **semi conservative mode** of DNA replication, each newly synthesized strand would remain associated with the old strand against which it was synthesized. Thus each progeny DNA molecule would consist of one old and one newly synthesized strand.

Semi Conservative Replication

The semi conservative mode of DNA replication was postulated by Watson and Crick along with the double helix model of DNA. The main features of this mode of DNA replication are as follows:

1. A progressive separation of the two strands of a DNA molecule.
2. Complementary base pairing of the bases located in the single stranded regions thus produced with the appropriate freedeoxyribonulceotides.
3. Formation of phosphodiester linkages between the neighbouring deoxyribonucleotides that have base paired with the single stranded regions, thereby producing regions the new strand.
4. This ensures that the base sequence of the new strands are strictly complementary top those of the old strands.
5. The base sequence of a newly synthesized strand is dictated by the base sequence of the old strand, since the old strand serves as a template or lould for the synthesis of the new strand.

DNA Replication

It is proposed by Watson and Crick. According to this method, both the strands of parental DNA separate from one another. Each old strand synthesizes a new strand. Thus, each of the two resulting DNA has one parental and one new strand. This method of DNA replication is universally accepted because there are several evidences in support of semi conservative method and it consists of several steps.

1. **Initiation of Replication** DNA replication starts at a specific point on the chromosome. This unique site is known as origin. The site of initiation differs from organism to organism. Sometime replication starts with an incision made by an incision enzyme known as endonuclease.
2. **Unwinding of strands.** The two stands of DNA double helix unwind. The opening of DNA stands take's places with the help of DNA unwinding protein.
3. **Formation of RNA Primer.** Synthesis of RNA primer is essential for initiation of DNA synthesis RNA primer is synthesized by the DNA template near the origin with the help of a special type of RNA polymerase.
4. **Synthesis of DNA on primer.** After formation of RNA primer, DNA synthesis starts on the RNA primer. Deoxyribose nucleotides are added to the 3e end position of RNA primer. The main DNA strand is synthesized on the DNA template with help of DNA polymerase. The DNA synthesis takes place in short pieces. Which are known as Okazaki fragments.
5. **Removal of RNA Primer:** DNA polymerase degrades the RNA primer

1. This enzyme also catalyzes the synthesis of short DNA segment to replace the prime. The newly synthesized segment is joined to the main DNA strand with the help of DNA ligase enzyme.

6. Union of Okazaki Fragments. The discontinuous fragment of Okazaki is joined to make continuous strands. The union of Okazaki fragments takes place with the help of a joining enzyme called polynucleotide ligase. The replication may take place either in one direction or in both the directions from the point of origin.

Evidence for semi conservation replication

Various experiments have demonstrated the semi-conservative mode of DNA replication. Now it is universally accepted that DNA replicates in a semi-conservative manner. There are three important experiments, which support that DNA replication is semi-conservative. These include (1) Meselson and Stahl experiment (2) Cairns experiment and (3) Taylor.s experiment.

Taylor.s experiment: Taylor (1969) conducted his experiments with root tip cells of *vicia faba*. He treated root tips with radioactive thymidine to label the DNA. The root tips were grown in the normal medium. In the first generation both chromatids were labeled.

In the second generation of cell division, one chromatid of each chromosome was labeled and the other one was normal. This demonstrated semi conservative mode of chromosome replication. The DNA replication is associated with chromosome replication.

Enzymes involved in DNA / RNA replication

DNA replication involves several proteins and enzymes, which together form the multienzymes complex, rep0lication apparatus or replisome. In E coli at lest two dozen gene products are involved in DNA replication. Many of these protein were first identified through studies of mutants e.g. Genes dna E, dna N, dna x etc of E colic code for the four of the seven polypeptides of the complete DNA polymerase III enzyme, and DNA G specifies the primase enzyme. Some enzymes like ligase, DNA polymerase 1 etc were discovered biochemically.

DNA repair systems

Damages to the genetic material, i.e., DNA are taken care of by the DNA repair systems. The various damages to DNA may be grouped into the following two types:

(1) Single base changes: Such changes affect a single base of a DNA molecule they do not produce structural distortions and do not affect either replication or transcription of the affected molecules. These changes ar represented by the conversion of one base into another, eg; deamination of 5 methylcytosine results in thymine and by the covalent addition of a small group to a base which affects its pairing behavior. As a result, the affected base does not pair properly with its partner base.

(2) Structural distortations: These changes generally adversely affect the replication and or transcription of the affected DNA molecule. They are represented by a single strand nick, removal of a base, covalent links between bases in the same or in the opposite strands (eg) Pyrimidine dimmers and addition of a bulky adduct to a base which may distort the configuration of the double helix.

The repair systems recognize a variety of changes in DNA to initiate action. Each cell possesses several repair systems in order to be able to deal with the various types of DNA damage; these systems may be grouped into the following general categories

1. Direct repair
2. Excision repair
3. Mismatch repair
4. Tolerance systems
5. Retrieval systems

1. Direct repair of DNA

The reversal or simple removal of the damage to the DNA is known as direct repair, eg., removal of the covalent bonds between the two 4 and two 5 carbons of the two thymine residues participating in the formation of thymine dimmers. Thymine dimers are generally formed due to UV radiation and interfere with replication and transcription. A specific enzyme mediates the splitting of the covalent bonds between the two T residues, which specifically recognizes to thymine dimmers. The enzyme can bind to the thymine dimmers in the dark, but requires the energy from blue light for removal of the covalent bonds between the T residues; that is why this process is known as photoreactivation. The direct repair system is wide spread in nature and is especially important in plants.

2. Excision repair

In this repair pathway, the damaged or mispaired segment of the DNA strand is exercised and new stretch of DNA is synthesized in its place. The various excision repair systems vary in their specificity. The repair process consists of the following steps:

a. Recognition and incision: The damaged section of a strand recognized by an endonuclease; this enzyme then cuts the affected strand on both the sides of damage.

b. Excision: After the incision, a 5' to 3 ' exonulcease digests away the damage/ mispaired section; this generates a single stranded region in the DNA double helix.

c. Synthesis: In this step, the single stranded region produced by excision serves as a template for a DNA polymerase which synthesis the replacement for the excised segment. DNA ligase then seals the nick that remains after the synthesis of the replacement for the excisedsection.

3. Mismatch repair: When single bases in the DNA are mismatched, either due to alterations in the existing bases or due to errors during replication, structural distortions result in the DNA double helix.

4. Tolerance systems: These systems deal with the damages that block normal replication at the damaged sites possibly by permitting the replication of the damaged sites possibly with a high frequency of errors. These systems may be particularly important in the eukaryotes where the genome size is very large and hence a complete repair of the damage is rather unlikely.

5. Retrieval systems: These systems are also known as post replication repair or recombination repair.

RNA AND ITS STRUCTURE, FUNCTION AND TYPES

With the discovery of the molecular structure of the DNA double helix in 1953, researchers turned to the structure of ribonucleic acid (RNA) as the next critical puzzle to be solved on the road to understanding the molecular basis of life. Ribonucleic acid (RNA) is a type of molecule that consists of a long chain of nucleotide units. Each nucleotide consists of a nitrogenous base, a ribose sugar, and a phosphate. RNA is very similar to DNA, but differs in a few important structural details: in the cell, RNA is usually single-stranded, while DNA is usually double-stranded; RNA nucleotides contain ribose while DNA contains deoxyribose (a type of ribose that lacks one oxygen atom); and RNA has the base uracil rather than thymine that is present in DNA.

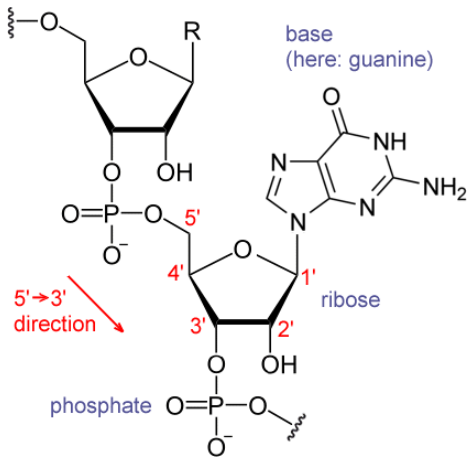
RNA is transcribed from DNA by enzymes called RNA polymerases and is generally further processed by other enzymes. RNA is central to the synthesis of proteins. Here, a type of RNA called messenger RNA carries information from DNA to structures called ribosomes. These ribosomes are made from proteins and ribosomal RNAs, which come together to form a molecular machine that can read messenger RNAs and translate the information they carry into proteins. There are many RNAs with other roles – in particular regulating which genes are expressed, but also as the genomes of most viruses.

Ribose Nucleic Acids

Most cellular RNA is single stranded, although some viruses have double stranded RNA. The single RNA strand is folded upon itself, either entirely or in certain regions. In the folded region a majority of the bases are complementary and are joined by hydrogen bonds. This helps in the stability of the molecule. In the unfolded region the bases have no complements. Because of this RNA does not have the purine, pyrimidine equality that is found in DNA.

RNA also differs from DNA in having ribose as the sugar instead of deoxyribose. The common nitrogenous bases of RNA are adenine, guanine, cytosine and uracil. Thus the pyrimidine uracil substitutes thymine of DNA. In regions where purine pyrimidine pairing takes place, adenine pairs with uracil and guanine with cytosine. In addition to the four bases mentioned above, RNA also has some unusual bases.

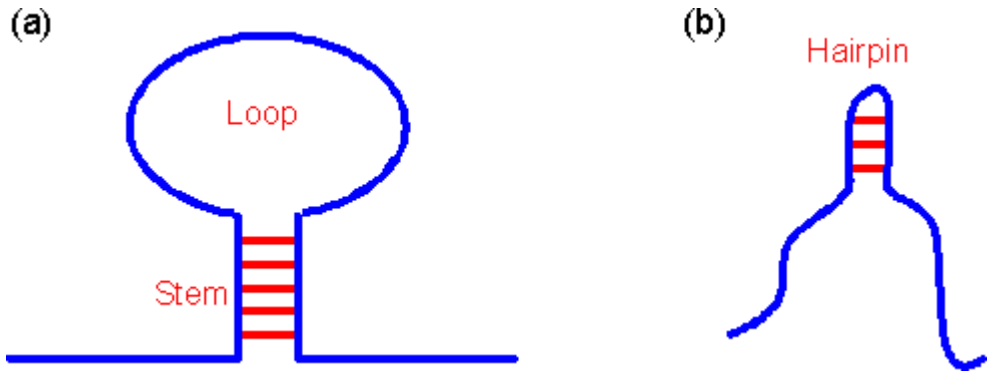
CHEMICAL STRUCTURE OF RNA



group at the 2' position of the ribose sugar. The presence of this functional group causes the helix to adopt the A-form geometry rather than the B-form most commonly observed in DNA. This results in a very deep and narrow major groove and a shallow and wide minor groove. A second consequence of the presence of the 2'-hydroxyl group is that in conformationally flexible regions of an RNA molecule (that is, not involved in formation of a double helix), it can chemically attack the adjacent phosphodiester bond to cleave the backbone.

Most cellular RNA molecules are single stranded. They may form secondary structures such as stem-loop and hairpin.

Secondary structure of RNA. (a) stem-loop. (b) hairpin.



There are more unusual bases in RNA than in DNA. All normal RNA chains either start with adenine or guanine: Three types of cellular RNA have been distinguished:

Messenger RNA (mRNA) or template RNA Ribosomal RNA (rRNA) and
Soluble RNA (sRNA) or transfer RNA (tRNA)

Ribosomal and transfer RNA comprise about 98% of all RNA. All three forms of RNA are made on a DNA template.

Transfer RNA and messenger RNA are synthesized on DNA templates of the chromosomes, while ribosomal RNA is derived from nucleolar DNA. The three types of RNA are synthesized during different stages in early development. Most of the RNA synthesized during cleavage is mRNA. Synthesis of tRNA occurs at the end or cleavage, and rRNA synthesis begins during gastrulation.

Comparison between DNA and RNA

	DNA	**RNA**
1.	DNA is the usual genetic material	RNA is the genetic material of some viruses.
2.	DNA is usually double-stranded, (In certain viruses DNA is single stranded, e.g. φ X 174).	Most cellular RNA is single stranded. (Some viruses e.g. retrovirus, have double stranded RNA).
3.	The pentose sugar is deoxyribose.	The pentose sugar is ribose.
4.	The common organic bases are adenine, guanine, cytosine and thymine.	The common organic bases are adenine, guanine, cytosine and uracil.
5.	Base pairing: adenine pairs with thymine and guanine with cytosine.	Adenine pairs with uracil and guanine with cytosine.
6.	Pairing of bases is throughout the length of the molecule.	Pairing of bases is only in the helical region
7.	There are fewer uncommon bases	There are more uncommon bases.
8.	DNA is only of one type	There are three types of RNA messenger, ribosomal and transfer RNA.
9.	Most of the DNA is found in the chromosomes. Some DNA is also found in the cytoplasm e.g. in mitochondria and chloroplasts.	Messenger RNA is formed on the chromosomes, and is found in the nucleolus and cytoplasm. rRNA and tRNA are also formed on the chromosomes, and are found in cytoplasm.
10.	Denaturation (melting) is partially reversible only under certain conditions of slow cooling (renaturation).	Complete and practically in stantaneous reversibility of the process of melting.
11.	Sharp, narrow temperature interval of transition in melting.	Broad temperature interval of transition in melting.
12.	DNA on replication forms DNA, and on transcription forms RNA.	Usually RNA does not replicate or transcribe. (In certain viruses RNA can synthesize an RNA chain).

13.	Genetic messages are usually encoded in DNA.	The usual function of RNA is translating messages encoded in DNA into proteins.
14.	DNA consists of a large number of nucleotides, up to 4.3 million	RNA consists of fewer nucleotides, up to 12,000.

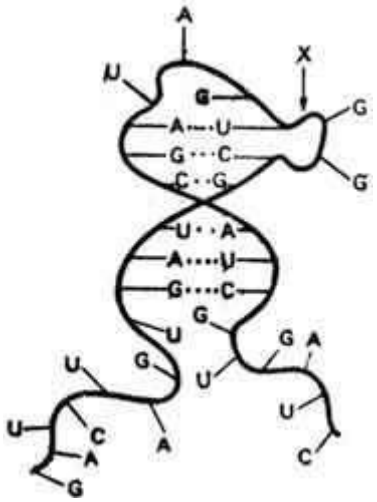
Ribosomal RNA – rRNA

Ribosomal RNA, as the name suggests, is found in the ribosomes. It comprises about 80% of the total RNA of the cell. The base sequence of rRNA is complementary to that of the region of DNA where it is synthesized.

In eukaryotes ribosomes are formed on the nucleolus. Ribosomal RNA is formed from only a small section of the DNA molecule, and hence there is no definite base relationship between rRNA and DNA as a whole.

Ribosomal RNA consists of a single strand twisted upon itself in some regions. It has helical regions connected by intervening single strand regions. The helical regions may show presence or absence of positive interaction. In the helical region most of the base pairs are complementary, and are joined by hydrogen bonds. In the unfolded single strand regions the bases have no complements.

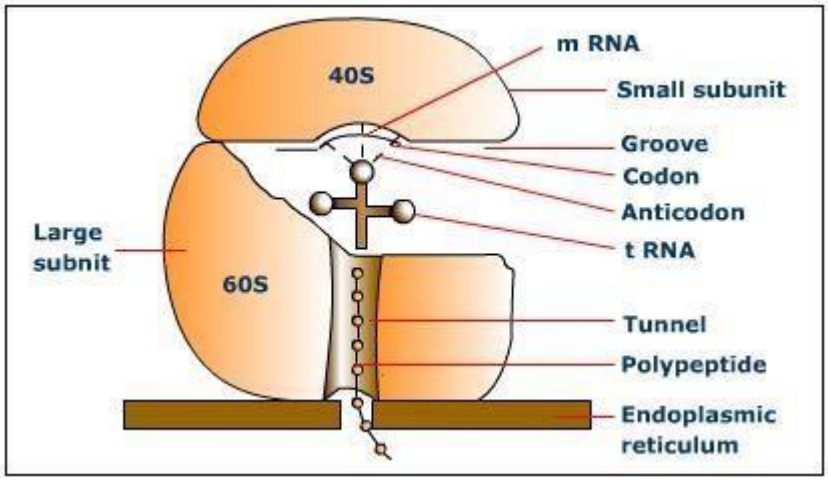
Ribosomal RNA contains the four major RNA bases with a slight degree of methylation, and shows differences in the relative proportions of the bases between species. Its molecules appear to be single polynucleotide strands which are unbranched and flexible. At low ionic strength rRNA behaves as a random coil, but with increasing ionic strength the molecule shows helical regions produced by base pairing between adenine and uracil and guanine and cytosine.



Hence rRNA does not show purine-pyrimidine equality. The rRNA strands unfold upon heating and refold upon cooling. Ribosomal RNA is stable for at least two generations. The ribosome consists of proteins and RNA. The 70S ribosome of prokaryotes consists of a 30S subunit and a 50S subunit. The 30S subunit contains 16S rRNA, while the 50S subunit contains 23S and 5S rRNA.

The 80S eukaryote ribosome consists of a 40S and a 60S subunit. In vertebrates the 40S subunit contains 18S rRNA, while the 60S subunit contains 28- 29S, 5.8S and 5S rRNA. In plants and invertebrates the 40S subunit contains 16- 18S RNA, while the 60S subunit contains 25S and 58 and 5.8S rRNA. There are three types of ribosomal RNA on the basis of sedimentation and molecular weight.

Two of these classes are high molecular weight RNAs, while the third is a low molecular weight RNA. The three classes are: (I) high molecular weight rRNA with molecular weight of over a million, e.g. 21s-29s RNA, (2) high molecular weight rRNA with molecular weight below a million e. g. 12-8-188 rRNA, (3) low molecular weight rRNA e. g. 58 rRNA.



Messenger RNA - mRNA - Jacob and Monod (1961) proposed the name messenger RNA for the RNA carrying information for protein synthesis from the DNA (genes) to the sites of protein formation (ribosomes). It consists of only 3 to 5% of the total cellular RNA.

Size of Messenger RNA - mRNA - The molecular weight of an average sized mRNA molecule is about 500,000, and its sedimentation coefficient is 8S. It should be noted however, that mRNA varies greatly in length and molecular weight. Since most proteins contain at least a hundred amino acid residues, mRNA must have at least 100 X 3= 300 nucleotides on the basis of the triplet code.

Stability of Messenger RNA - mRNA - The cell does not contain large quantities of mRNA. This is because mRNA, unlike other RNAs is constantly undergoing breakdown. It is broken down to its constituent ribonucleotides by ribonucleases.

Structure of Messenger RNA - mRNA

Messenger RNA is always single stranded. It contains mostly the bases adenine, guanine, cytosine and uracil. There are few unusual substituted bases. Although there is a certain amount of random coiling in extracted mRNA, there is no base pairing. In fact base pairing in the mRNA strand destroys its biological activity

Since mRNA is transcribed on DNA (genes), its base sequence is complementary to that of the segment of DNA on which it is transcribed. This has been demonstrated by hybridization experiments in which artificial RNA, DNA double strands are produced. Hydrization takes place only if the DNA and RNA strands are complementary.

Usually each gene transcribes its own mRNA. Therefore, there are approximately as many types of mRNA molecules as there are genes. There may be 1,000 to 10.000 different species of mRNA in a cell. These mRNA types differ only in the sequence of their bases and in length.

When one gene (cistron) codes for a single mRNA strand the mRNA is said to be monocistronic. In many cases, however, several adjacent cistrons may transcribe an mRNA molecule, which is then said to be polycistronic orpolygenic.

The mRNA molecule has the following structural features:

1. **Cap**. At the 5' end of the mRNA molecule in most eukaryote cells and animal virus molecules is found a 'cap'. This is blocked methylated structure, m7Gpp Nmp Np or m7Gpp Nmp Nmp Np. where: N = any of the four nucleotides and Nmp = 20 methyl ribose. The rate of protein synthesis depends upon the presence of the cap. Without the cap mRNA molecules bind very poorly to the ribosomes.
2. **Noncoding region 1 (NC1)**. The cap is followed by a region of 10 to 100 nucleotides. This region is rich in A and U residues, and does not translateprotein.
3. **The initiation codon** is AUG in both prokaryotes and eukaryotes
4. **The coding region** consists of about 1,500 nucleotides on the average and translates protein It is made up of 73-93 nucleotides (Rich and RajBhandary, 1976). Each bacterial cell probably contains about a hundred or more different types of tRNA. The function of tRNA is to carry amino acids to mRNA during protein synthesis. Each amino acid is carried by a specific tRNA. Since 20 amino acids are coded to form proteins, it follows that there must be at least 20 types oftRNA.

It was formerly thought that only 20 tRNA molecular types exist, one for each amino acid. It has, however, been shown that in several cases there are at least two types of tRNA for each amino acid. Thus there are many more tRNA molecules than amino acid types. These are probably coded by one gene.

Transfer RNA is synthesized in the nucleus on a DNA template. Only 0.025% of DNA codes for tRNA. Synthesis of tRNA occurs near the end of cleavage stages. Transfer RNA is an exception to other cellular RNAs in that a part of its ribonucleotide sequence (-CCA) is added after it comes off the DNA template. Like rRNA, tRNA is also formed from only a small section of the DNAmolecule.

Therefore, it does not show any obvious base relationships to DNA. The tRNA molecule consists of a single strand looped about it self. The 3' end always terminates in a -C-C-A (cytosine- cytosine-adenine) sequence. The 5' end terminates in G (guanine) or C (cytosine). Many of the bases are bonded to each other, but there are also unpaired bases.

Transfer RNA - tRNA OR Soluble RNA – sRNA

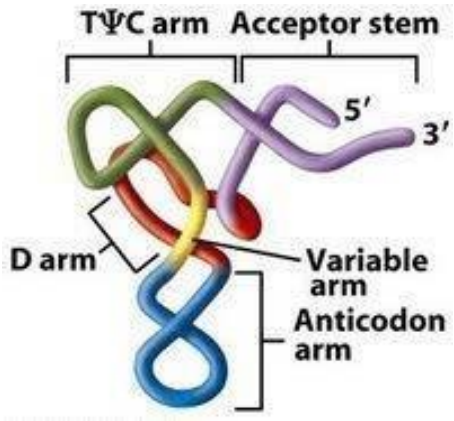
After rRNA the second most common RNA in the cell is transfer RNA. It is also called soluble RNA because it is too small to be precipitated by ultracentrifugation at 100,000 g. It constitutes about 10-20% of the total RNA of the cell. Transfer RNA is a relatively small

RNA having a molecular weight of about 25,000 to 30,000 and the sedimentation coefficient of mature eukaryote tRNA is 3.8S.

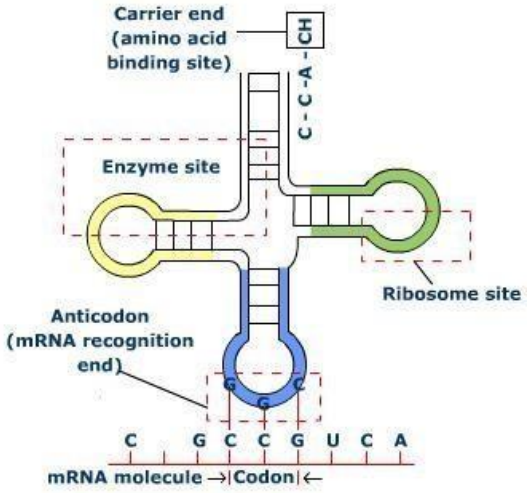
Structure of Transfer RNA – tRNA

The nucleotide sequence (primary structure) of tRNA was first worked out by Holley et al (1965) for yeast alanine tRNA. Since then the sequence of about 75 different tRNAs, ranging from bacteria to mammals, has been established. The different tRNAs are all minor variants of the same basic type of structure. Several models of the secondary structure of tRNA have been proposed, and of these the cloverleaf model of Holley is the most widely accepted.

Transfer RNA (tRNA) is an essential component of the protein synthesis reaction. There are at least twenty different kinds of tRNA in the cell[1] and each one serves as the carrier of a specific amino acid to the site of translation. tRNA's are L-shaped molecules. The amino acid is attached to one end and the other end consists of three anticodon nucleotides. The anticodon pairs with a codon in messenger RNA (mRNA) ensuring that the correct amino acid is incorporated into the growing polypeptide chain. The L-shaped tRNA is formed from a small single-stranded RNA molecule that folds into the proper conformation. Four different regions of double-stranded RNA are formed during the folding process.



The two ends of the molecule form the *acceptor stem* region where the amino acid is attached. The anticodon is an exposed single-stranded region in a loop at the end of the *anticodon arm*. The two other stem/loop structures are named after the modified nucleotides that are found in those parts of the molecule. The *D arm* contains dihydrouridylate residues while the *TΨC arm* contains a ribothymidylate residue (T), a pseudouridylate residue (Ψ) and a cytidylate (C) residue in that order. All tRNA's have a similar *TΨC* sequence. The *variable arm* is variable, just as you would expect. In some tRNA's it is barely noticable while in others it is the largest arm. tRNA's are usually drawn in the "cloverleaf" form (below) to emphasize the base- pairs in the secondary structure.



Clover leaf model of tRNA
Unusual Bases in tRNA

In addition to the usual bases A, U, G and C, tRNA contain a number of unusual bases, and in this respect differs from mRNA and rRNA. The unusual bases of tRNA account for 15-20% of the total RNA of the cell. Most of the unusual bases are formed by methylation (addition of -CHa or methyl group to the usual bases), e.g. cytosine and guanine on methylation yield methylcytosine and methyl/guanine, respectively.
Precursor tRNA molecules transcribed on the DNA template contains the usual bases. These are then modified to unusual bases. The unusual bases are important because they protect the tRNA molecule against degradation by RNase. This protection is necessary because RNA is found floating freely in the cell.

Some of the unusual bases of tRNA are methyl guanine (GMe), dimethylguanine(GMe2), methylcytosine (Me), ribothymine (T), pseudouridine (ψ), dihydrouridine (DHU, H2U, UH2), inosine (I) and methylinosine (IMe, MeI). In general, organisms high in the evolutionary scale contain more modified bases than lower organisms.

Classification of tRNA - A Study of different tRNAs shows that the structure of the acceptor stem, the anticodon arm and the TψC arm are constant. The differences in the tRNAs lie in the D arm and the variable arm. Based on the differences in these two variable regions, three classes of tRNA have been recognized.

Class I (D4-V4-5), with 4 base pairs in the D stem and 4-5 bases in the variable loop.

Class II (DS-V4-5), with 3 base pairs in the D stem and 4-5 bases pairs in the variable loop.

Class III (D3-VN), with 3 base pairs in the D stem and a large variable arm.

A simpler classification based only on the variable arm recognizes two types of tRNA. Class I with 4-5 bases in the variable loop
Class II with a large variable arm of 13-21 bases.

Tertiary Structure of Transfer - tRNA - Electron density maps have revealed that tRNA has a tertiary structure. This structure is due to hydrogen bonds

(i) between bases,

(ii) between bases and ribose phosphate backbone and

(iii) between the backbone residues. (The hydrogen bonding in the double helical stem regions of the tRNA molecular are considered to be in the secondary structure).

Initiator Transfer RNA - tRNA

The starting amino acid in eukaryote protein synthesis is methionine, while in prokaryotes it is N- formyl methionine. The tRNA molecule3 specific for these two amino acids are methionyl tRNA (tRNAmet) and N-formyl- methionyl IRNA (tRNAfmet) respectively.

These tRNAs are called initiator tRNAs, because they initiate protein synthesis. Initiator tRNAs have certain features which distinguish them from other tRNAs, and the initiator tRNAs of prokaryotes' and eukaryotes also differ.

In most prokaryotes the 5' terminal nucleoside is C. It has opposite it (i.e. in the fifth position from the 3' end) an A nucleotide. There is no Watson-Crick base pairing between the two. In the blue green 'alga' Anacystis nidulans, however, the fifth nucleotide from the 3' end is C. In eukaryotes there is an A.U base pair at the acceptor stem.

As noted previously, prokaryotes use tRNAf-met for initiation of protein synthesis, while eukaryotes use tRNAmet. The prokaryote Halo bacterium cutirubrum is, however, reported to initiate protein synthesis with tRNA met and has an A.U base pair at the end of the accept or stem. In these respects it resembles eukaryotes

The D loop of prokaryote initiator tRNAs contains an A11, U24 base pair. All other tRNAs have a Y11, R24 base pair. Eukaryotic cytoplasmic initiator tRNAs have AU or AU* instead of TΨ in the TΨC loop. Also, in eukaryotes instead of a pyrimidine nucleotide (Y) there is A at the 3' end of the TΨC loop.

In some eukaryotic cytoplasmic initiator tRNAs the anticodon sequence CAU is preceeded by C instead of U as in all other tRNAs. In prokaryotes the purine nucleotide following C in the TΨC loop is A, while in eukaryotes it is G. In tRNA f- met the nucleotide adjacent to the 3' side of the anticodon triplet is adenosine while in tRNA met it is alkylated adenosine

Specificity of Tranfer RNA - tRNA

Two important steps in translation during protein synthesis are the activation of amino acids and the transfer of amino acids to tRNAs. Each amino acid has a specific activating enzyme tRNA aminoacyl synthetase. Thus there are 20 different tRNA aminoacyl synthetases for the 20 common amino acids found in proteins.

Some tRNA synthetases can activate more than one amino acid, i.e. they show only a limited substrate specificity. Thus isoleucine tRNA synthetase can also activate L valine, and valine tRNA synthetase can also react with threonine. The enzymes, however, recognize only a specific set of, tRNAs as substratesL isolecine tRNA synthetase recognizes only tRNAileu and valine tRNA synthetase recognizes only tRNAval. Thus specificity is involved at two stages, activation of the amino acid and transfer of the amino acid to tRNA. Another group of enzymes, the tRNA aminoacyl transferases catalyse the transfer of an amino acid .