Cell Biology and Genetics

CHROMOSOME ORGANIZATION AND COMPOSITION

D. C. Gautam
Department of Bio-Sciences,
Himachal Pradesh University,
Shimla 171 005 India

CONTENTS

INTRODUCTION

In eukaryotes, the nucleus of every cell has several thread-like structures called chromosomes which are made up of DNA and proteins. However, in prokaryotes, the chromosome consists of a molecule of DNA which is in close contact with other components of the cell. DNA stands for deoxyribonucleic acid and this is one of the most important kinds of molecules found in living organisms. DNA macromolecules contain instructions that determine the chemistry of the cell, as the production of different proteins in the cell is decided by the information contained in the DNA.

Individuals of a species have a unique set of inherited characteristics that make them different from the individuals of other species. For example, all traits of human beings make them unique to be known as *Homo sapiens* and distinguish them from apes. Even individuals within a species are by no means identical. Many traits or phenotypes differ from individual to individual in a species. For example, in human beings, there is great deal of variation in the traits such as colour of hair, eye or skin as well as in other characteristics such as height, weight and personality features etc. Some traits are transmitted biologically, while others are acquired as they are influenced by the environment. Colour of eyes results from biological inheritance but the personality development is influenced by the environmental factors. Some traits are influenced jointly by biological inheritance and environmental factors.

Every organism has a specific genome. The haploid set of DNA in an organism is called a genome. In fact, a genome is composed of long DNA molecules that are, in turn, the main components of chromosomes. Each chromosome contains one DNA molecule carrying many genes, which are responsible for characteristics of the organisms. For all practical purposes, the genes are regions of chromosomal DNA that can be transcribed.

In most of prokaryotic organisms, the genome consists of only one chromosome, whereas, in eukaryotes, the genome consists of a set of many linear chromosomes. The eukaryotic chromosomes are found in the nucleus although mitochondria and chloroplasts, in eukaryotic cell, also have their own DNA. In addition to these, cells of many prokaryotes and some eukaryotes have plasmid genomes that contain genes necessary for their own propagation. Viruses are not considered as organisms but these contain their own sets of genes.

At present, the sequence of entire genome is known in a number of viruses and bacteria. In unicellular eukaryote, *Saccharomyces cerevisiae* also, the genome sequence is known. In multicellular eukaryotes, the vast majority of genome sequence is now known for *Arabidopsis thaliana,* round worm *Caenorhabditis elegans*, the fruitfly

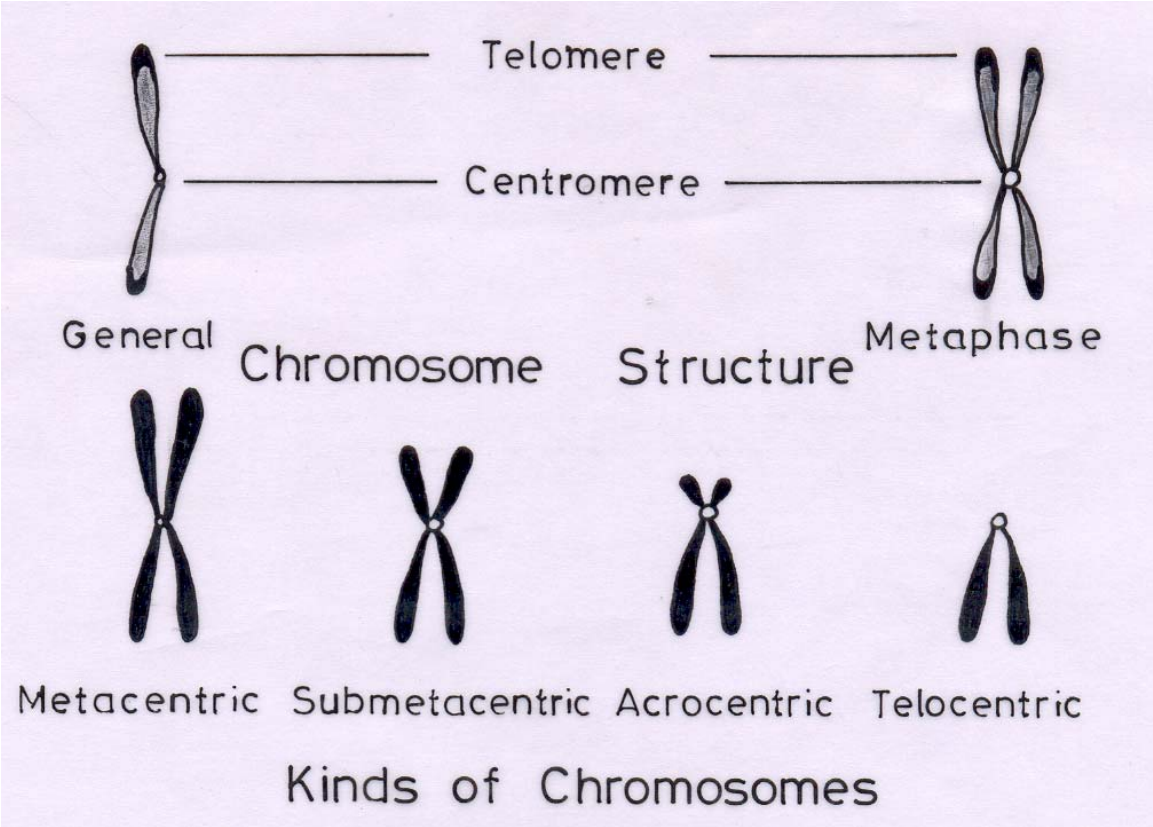
Drosophila melanogaster and humans. Studies on the sequence data in these organisms revealed that a large portion of their genome does not encode mRNA or any other RNA required by the organism. It is important to note that such non-coding DNA constitutes most of the chromosomal DNA.

CHROMOSOME MORPHOLOGY

The chromosome literally means coloured body. The nuclei of eukaryotes contain chromosomes which are long, linear molecules that carry all the genes present in the genome of an organism. In species, chromosomes may differ considerably in size and may be categorized in groups. For example, in humans, the chromosomes are placed into seven groups from A to G. The largest chromosome is placed in group A and the smallest chromosome in group G. Between organisms, the size differences of chromosomes can be many fold and such differences even exist between species.

Chromosomes can be observed inside the nucleus as thread like structures that become visible under the light microscope when the cell is stained with certain dyes. The number of chromosomes in each cell of a particular species is always constant. Chromosomes are found to exhibit a characteristic splitting behaviour in which each daughter cell formed by cell division receives an identical complement of chromosomes. There is close relationship between the chromosomes and the DNA as the chromosomes contain DNA and proteins. While the amount of DNA per cell is constant, the amount and kinds of chromosomal proteins differ greatly from one cell type to another.

Metaphase chromosomes contain the two DNA molecules (replicated DNA) linked together at a structure called the centromere. Position of the centromere is characteristic for a particular chromosome and is also a feature which can be used to distinguish individual members of the entire complement of chromosomes in the nucleus. Centromere is the point at which the chromosome attaches to the microtubules that draw the daughter chromatids into their respective nuclei during cell division. The centromeric region of chromosome contains special proteins and is characterized by specific DNA sequences.



The spindle fibres act as molecular strings and are attached to chromosomes during cell division at a specialized region called centromere. It appears as a constriction that divides the chromosome into two arms; the shorter is called p and the larger is called q. The position of the constriction decides the ratio between the lengths of the two chromosome arms. The arm ratio is a useful characteristic for identifying individual chromosomes. Centromeric position can also be characterized as telocentric (centromere at one end), acrocenric (close to one end), submetacentric (close to middle) or metacentric (exactly in the middle). However, there are species in plants and animals that have holocentric chromosomes which have kinetic activity diffused throughout the length of the chromosome.

Tips of the chromosomes are called telomeres. These are distinct from the other parts of the chromosome. The telomeres and centromeres have unique molecular structures that are crucial to normal chromosome behaviour.

Microscopic study of chromosomes and analysis of their genetic properties is called cytogenetics. The components of a chromosome including its DNA become condensed from the extended form found in non-dividing cells, into a shorter, thicker form that can be easily handled by the division apparatus of the cell.

KARYOTYPE

A karyotype is a micrograph in which all the chromosomes within a single cell are arranged in a standard fashion to give pictorial or photographic representation. The chromosomes are usually arranged in order of size and numbered from largest to smallest. Sex chromosomes are designated with letters. For example, in humans, there are 23 pairs of chromosomes. Of these, 22 pairs are of autosomes and a pair is of sex chromosomes. The autosomes are numbered from 1 to 22 according to their size and shape. The sex chromosomes are designated as X and Y. Generally, the chromosomes in actively dividing cells can be karyotyped.

Over a period of time, the cytogeneticists have devised various ways to identify and classify chromosomes. Three commonly used features are size, location of centromere and banding patterns that are revealed when the chromosomes are treated with stains. We know that chromosomes differ in size and if their sizes are the same even then, these may differ in the position of the centromere. The centromere is identified as a large constriction where the chromosome appears to be constricted. The term kinetochore is usually used to designate a structure around the centromeric region.

The centromere divides the chromosome into arms. If the centromere is in the middle, the chromosome is metacentric and has arms of equal lengths. If the centromere is towards one end, the chromosome is either acrocentric or submetacentric. Acrocentric and submetacentric chromosomes have one short arm and one long arm. If the centromere is towards one end, the chromosome is telocentric. In case of telocentric chromosome, the short arm is non-existent. Long arm of chromosome is designated with letter q and a short arm is designated with letter p. Each arm is divided into regions and then classified by bands starting from the centromere to the distal tip of the arm. The location of gene for ABO blood group is designated as $9q^{34}$. This means long arm of chromosome 9, region 3 and band 4.

While preparing a karyotype, the chromosomes are aligned with short arms on top and long arm on the bottom. In 1956, Tijo and Levan determined correct diploid number of human chromosomes (2n=46). In humans, a system is used for identifying chromosome based on chromosomal size, position of centromere and banding patterns. Autosomes are numbered first on the basis of length, with X and Y chromosomes identified separately.

Even if the sizes of the chromosomes and their centromeric locations are the same, these can be distinguished from each other through the banding pattern which is unique for every chromosome. Banding patterns are also

used to detect changes in chromosome structure. Thus, the chromosome banding may reveal evolutionary relationships among the chromosomes of closely related species.

Using all the available chromosomal landmarks together, cytogeneticists can distinguish each and every chromosome of any species. The features like size, arm ratio, position of centromere, heterochromatin, location of nucleolar organizers and banding patterns identify the individual chromosomes within the set that characterizes a species.

There are six levels of karyology used to study the characteristics of chromosomes in different species. In alpha karyology, only chromosome number and approximate sizes are determined. This is the only practical level of analysis in yeast and most protozoa, fungi and algae. In beta karyology, chromosome numbers and lengths of chromosome arms are known and positions of centromeres are accurately located. Most of the karyotype data falls in this category. In gamma karyology, Giemsa and fluorescent banding is done to map the chromosomes. In delta karyology, locations of satellite DNAs and NORs are determined. Epsilon karyology is involved in the analysis of lampbrush chromosomes and zeta karyology deals with the polytene chromosome analysis.

In case of monocentric chromosomes, the two distinguishing features prominently clear in mitosis are the actual length of chromosome and the position of centromere. From these, the centromeric index, the arm ratio and relative length of the chromosome can be calculated. The centromeric index is defined as the length of the shorter of two chromosome arms multiplied by 100 and divided by the whole length of the chromosome. Arm ratio is the length of the longer arm of the chromosome divided by the length of the shorter one. Its value is always greater than 1. Relative length is defined as the length of the whole chromosome multiplied by 100 and divided by the total length of all the chromosomes in the complement. It is expressed as a percentage. In holocentric chromosomes, as the centromere is diffused, so only the actual lengths of the chromosomes are taken into consideration. From actual lengths, the total complement length is calculated. Then relative length of a chromosome is expressed as a percentage of total complement length.

CHROMOSOME BANDING

Special staining techniques have revealed intricate sets of bands (transverse stripes) on chromosomes in many different organisms. The bands represent useful landmarks because their positions and sizes are highly chromosome-specific. There are Q bands produced by quinacrine hydrochloride, G bands produced by Giemsa stain and R bands produced by reversed Giemsa. Chromosome banding is of fundamental importance for

chromosome identification on the basis of longitudinal differentiation and is thus very useful for elucidating the evolutionary relationships in organisms.

For chromosome characterization, four distinct kinds of bands can be recognised. These bands are specific to euchromatin, heterochromatin, nucleolar organizer regions and kinetochores. Euchromatic bands (including Q-, G- and R-bands) form a pattern of positive (darkly staining or brightly fluorescent) and negative (weakly staining or dimly fluorescent) bands throughout the length of the chromosome of higher vertebrates. Q, G and R bands probably reflect the degree of compactness of the DNA and are never associated with centromeric heterochromatin. Darker bands are found near the centromeres or on the ends (telomeres) of chromosomes. Dark staining areas are heterochromatic while light staining regions are euchromatic. These regions generally remain constant in different cells or individuals of a given species. Euchromatic regions often undergo a regular cycle of contraction and extension.

Heterochromatic bands (C-bands and various more specific types) are specific to heterochromatin and help in chromosome identification in majority of those species that lack euchromatic bands. Heterochromatic bands are highly localized, usually around the centromeres, but also occur elsewhere on the chromosomes. They correspond to classically defined heterochromatin, that is, regions of chromosomes that normally remain condensed throughout interphase. Nucleolar organizer regions (NORs) are the regions of chromosome that contain genes for ribosomal RNA and are responsible for the formation of nucleoli in the interphase nuclei. Kinetochores are the special regions by which chromosomes are attached to the spindle during cell division and these can be demonstrated through immunolabelling.

Banding patterns have revealed that chromosomes are segmented into a series of regions having distinctive properties in respect of DNA base composition, time of DNA replication and gene content. Thus, the chromosome bands are, in fact, visible expression of the functional and compositional compartmentalization of chromosomes.

Natural banding patterns become readily visible in polytene chromosomes and can serve as landmarks such as in fruitly, *Drosophila melanogaster*. The diploid chromosome number in this species is eight (2n=8) and these eight chromosomes are present in most of the cells. However, in the cells of the special organs that contain the polytene chromosomes, certain interesting peculiarities exist. The banding pattern of each chromosome is unique. The bands do not represent genes as in any chromosomal region of *Drosophila,* there are more genes than the number of polytene bands.

CENTROMERE

Centromere refers to the most prominent region of condensed mitotic chromosomes called the primary constriction. This region was initially called centromere as it was invariably located in the middle between the ends of two chromosome arms. Later, the term was extended to describe the primary constriction of all mitotic chromosomes even when it is not located in a central position. Centromeres have multiple roles during mitosis enabling the equal distribution of genetic material during cell division.

On either side of the centromere, is a trilaminar plate structure called kinetochore. It is a multiprotein complex located at the surface of the chromosomes that binds spindle microtubules and regulates chromosome movement in mitosis. It is also the final site of sister chromatid pairing before segregation takes place. The centromere and the kinetochore help in ensuring the proper orientation of chromosomes at metaphase. Microtubules are attached to the kinetochores.

Though centromeres have been conserved throughout evolution, they show structural variability and are classified into two different types. Centromeres may be diffused as is found in many arthropods (especially aphids) and plants (*Luzula*) or localized as are generally present in eukaryotes. In diffused centromeres, spindle microtubules attach along the entire length of the chromatids while in localized centromeres, there is single region of attachment for spindle microtubules.

NUCLEOLAR ORGANIZERS

In addition to centromere, there may also be a secondary constriction on some chromosomes. The part of chromosome located distal to the secondary constriction is called a satellite. The secondary constriction usually contains genes that help form the nucleolus, the structure that is apparent in various stages of cell division. Secondary constriction regions are also called nucleolar organizer regions (NORs) as they are the sites for the organization of the nucleolus. Nucleoli are spherical structures found associated with secondary constriction of chromosomes. Diploid cells of many organisms have just a pair of nucleoli. The number of nucleolar organizers in a haploid set of chromosomes varies in different organisms which range in number from one to many per chromosome set.

Nucleolar organizers contain numerous tandem copies of genes that code for ribosomal RNA, which is component of the ribosomes. Ribosomal RNA is synthesized at the nucleolar organizers, deposited into the nucleoli, and later exposed to the cytoplasm to be incorporated into ribosomes.

TELOMERES

The ends of linear chromosomes are called telomeres. These specialized structures play important role in protecting the ends of chromosomes from attack by nuclease enzymes. They also prevent the chromosomes from joining together as the broken chromosomes attach immediately. Moreover, extreme 5'-terminus of a linear DNA molecule comprises an RNA primer that is not replaced with DNA. This ought to reduce chromosome length which does not happen.

A telomere is specialized to make the natural end of a linear chromosome behave differently from a simple double-stranded DNA break. Muller's experiments on effects of X irradiation on *Drosophila* cells proved that the natural ends of chromosomes have special structure which is required for chromosome stability. Muller termed this structure, the telomere. The ends of eukaryotic nuclear chromosomes have a special DNA sequence to which specific proteins are bound and this DNA-protein complex is responsible for many aspects of telomere functions.

Eukaryotes have linear chromosome and one problem faced by them is that once the first primer on each strand is removed, it appears that there is no way to fill in the gaps, since DNA can not be extended in the 3'-5' direction, and there is no 3'-end upstream as would be in a circle. If this were actually the situation, the DNA strands would get shorter every time they replicated, and the genes would be lost for ever. But the cells have their own mechanism to solve this problem. The ends of the eukaryotic chromosomes are formed by an enzyme called telomerase, which adds DNA to the 3'-ends of the chromosomes according to the instructions of its own RNA template. The renewal of chromosome ends counteracts the tendency of the chromosomes to lose telomeric DNA with every cell division.

It is actually the inability to replicate completely the end of a linear DNA molecule with DNA polymerase that has led to evolution of special DNA sequences called telomeres at the ends of eukaryotic chromosome. These sequences are different in diverse groups of organisms. They consist of many tandem repeats of a short sequence that contains a block of neighbouring G-nucleotides. The G-rich strand of telomere always forms the 3' end of DNA molecule and it seems to fold to form a special structure that protects the chromosome end.

Telomeres, or ends of eukaryotic chromosomes, contain no genes. Instead they are composed of many repeats of short, GC rich sequences. The exact sequence of the repeat in a telomere is species-specific. These repeats are added to the very 3'-ends of DNA strands, by semi-conservative replication by telomerase.

Telomeric DNA is made up of multiple copies of a short sequence, 5'-AGGGTT-3' in humans which is repeated possibly a thousand times or more at the extreme ends of each chromosomal DNA molecule. Actual structure is different for each telomere. 5'-3' strand is G rich while 3'-5' strand is C rich. The repeat sequences act as binding sites for telomeric-specific proteins. The bound proteins probably act as a cap, preventing the ends of the chromosomes being degraded or fusing with other chromosomes. The telomerase enzyme consists of a protein subunit and an RNA molecule. Telomerase RNA contains at one position a short sequence identical to one or more repeat sequences of the C-rich strand of telomere. It acts as a template for synthesis of repeat sequence of G-rich strand. Thus the shortening effect of DNA replication can be counterbalanced by repeatedly extending the G-rich strand.

There is relationship between telomeres and cellular life span. Primary human cells have a finite division capacity *in vitro,* at the end of which they enter a viable, though, non-dividing state. This is also known as Hayflick limit and provides a potent barrier to the uncontrolled and unlimited cell division.. Telomerase is expressed at low or undetectable levels in most human cells, although there are exceptions such as stem cells, whereas, it is expressed at high levels in the germline. Telomere DNA length decreases with successive rounds of division and this leads to ageing.

EUCHROMATIN AND HETEROCHROMATIN

There are two fundamental classes of chromatin structure; euchromatin and heterochromatin. Euchromatin is relatively extended and most of the active genes are present in euchromatin. When chromosomes are treated with certain chemicals that react with DNA, such as Feulgen reagent, distinct regions with different staining characteristics are revealed. Densely staining regions are called heterochromatin and reflect a high degree of compactness and poorly staining regions are called euchromatin and indicate less tightly packed regions.

Heterochromatin is very condensed and its DNA is inaccessible. It even appears as clump when viewed microscopically. Heterochromatin is found at or near the centromeres of chromosomes. The DNA associated with the centromere belongs to the class of DNA that is repeated many times in the genome. The tightly coiled DNA in such a state is not accessible to RNA polymerase. Heterochromatin is of two types; constitutive and facultative. Constitutive heterochromatin is a permanent part of genome and is not convertible to euchromatin and is generally

inactive. Facultative heterochromatin consists of euchromatin that takes on staining and compactness characteristic of heterochromatin during some phase of development.

DNA - THE GENETIC MATERIAL

Frederick Griffith, in 1928, demonstrated that DNA is the genetic material. He was working with two strains of bacterium *Streptococcus pneumoniae* identified as S and R. When a bacterial cell is grown on a solid medium, it undergoes repeated cell divisions to form a large colony.

Of the two strains of *Streptococcus pneumoniae,* the S type synthesizes a gelatinous capsule composed of polysaccharide which makes each colony large and gives it a glistening or smooth appearance. This capsule protects the bacteria from the defense mechanisms of an infected animal which suffers from pneumonia. However, bacteria of the R strain of *Streptococcus pneumoniae* are unable to synthesize the capsular polysaccharide so they form small colonies that have a rough ® surface. Without the capsule, the bacteria are inactivated by the immune system of the host, so this strain of bacteria does not cause pneumonia to the infected individual.

The S and R bacterial strains breed true as the progeny formed by cell division have either S or R capsule type similar to that of parent. When the mice are injected with living S cells, they get pneumonia but those injected either with living R cells or with heat killed S cells remain healthy. However, when mice are injected with a mixture of living R cells and heat killed S cells, they get the disease and often die of pneumonia.

Isolation of bacteria from blood samples of infected or dead mice showed that they produce S cultures, with a capsule typical of the injected S cells even though the injected S cells were heat killed. This showed that R bacteria can be transformed into S bacteria. The transformation was discovered but the chemical substance that caused it was still unknown.

Avery, Macleod and McCarty in 1944 demonstrated that the chemical substance that caused transformation of R cells into S cells was actually DNA. These workers developed chemical procedures to isolate pure DNA from cells. When they added DNA isolated from S cells to growing cultures of R cells, they observed transformation. Treatments that destroyed DNA, eliminated the transforming ability. These experiments proved that the substance responsible for genetic transformation was the DNA of the cell and, therefore, DNA is the genetic material.

DNA FINGERPRINT

The non-coding DNA in multicellular organisms has great significance as this contains certain conserved repeated sequences that are not identical among the species. Variations within such stretches of this repetitious DNA are so

great that each person can be distinguished by a DNA fingerprint based on these sequence variations. So fingerprinting of DNA sequences in an individual is a molecular manifestation differentiating one individual from other.

Jeffrey and his colleagues in 1988 coined this term while working on simultaneous detection of variable DNA loci by hybridization of specific multilocus probes with restriction fragments. This involves digestion of DNA with specific enzymes followed by electrophoresis, blotting and detection. In fact, the hybridization of DNA fragments with specific oligonucleotide probes has made DNA fingerprinting an important technique. DNA fingerprinting can be done through 'DNA typing' or 'DNA profiling'. In DNA typing, genomic DNA is cut with restriction enzymes, followed by electrophoresis and hybridization of the DNA fragments with multilocus probes. This yields a specific band pattern in the gel. However, in DNA profiling, specific DNA sequences with oligonucleotide primers are first amplified and then separated. Staining is then done to detect the bands.

In humans, there are minisatellites, the satellite DNA that exists in relatively short 1 to 5 kb regions made up of 20 to 50 repeat units each containing about 100 base pairs. These are distinguished from the more common regions of tandemly repeated satellite DNA, which are 20 to 100 kb in length. Even slight differences from different individuals can be detected by Southern blotting.

DNA CONTENT

Genome size is normally expressed as amount of DNA per haploid set of chromosomes. It is referred to as the 'C value' of the organism. The amount of DNA per diploid cell is the '2C value'. Since genome refers to all DNA present in the haploid set of chromosomes, so 'C' refers to the amount of DNA present in picograms (i.e. grams X 10^{-12}) per haploid set of chromosomes.

It is quite difficult to estimate the number of genes in an organism if its genome has not been completely sequenced. Even in those cases where sequence information is available, there can be problems distinguishing between small genes and non-functional open reading frames that have arisen by chance.

In molecular terms, a gene is the entire DNA sequence required for the synthesis of a functional protein or RNA molecule. In addition to coding regions (exons), a gene includes control regions and sometimes introns (non coding regions). In vertebrate genome, duplicated genes probably constitute half of the protein-coding DNA sequences. A set of duplicated genes that encode proteins with similar but non-identical amino acid sequence is called a gene family. Examples of a few gene families are protein kinases and vertebrate immunoglobins.

The amount of DNA per gene is greater in the human genome due to introns which may range from one to many, even more than 20. There is intergenic DNA comprising functional motifs such as gene promoters, upstream regulatory sequences and DNA replication origins. Human genome incorporates nearly 20 times as much DNA as that of *Drosophila melanogaster*.

A substantial part of most eukaryotic nuclear genomes is made up of repetitive DNA composed of individual sequence elements, that are repeated many times, either in tandem arrays or interspersed throughout the genome. Amount of repetitive DNA mainly determines the size of an organism's genome. Single copy DNA constitutes most genes and is made up of sequences that are not repeated elsewhere. Several amphibians and flowering plants have genome 10 times bigger than that of humans, even though the number of genes is almost similar. This shows that the large stretches of intergenic regions occur in these organisms.

DNA STRUCTURE

In organisms, it is possible to study differences between species through the comparison and analysis of DNA. It was Friedrich Miescher who discovered a weak acid in the nuclei of white blood cells in 1869 and this weak acid turned out to be the chemical substance called DNA.

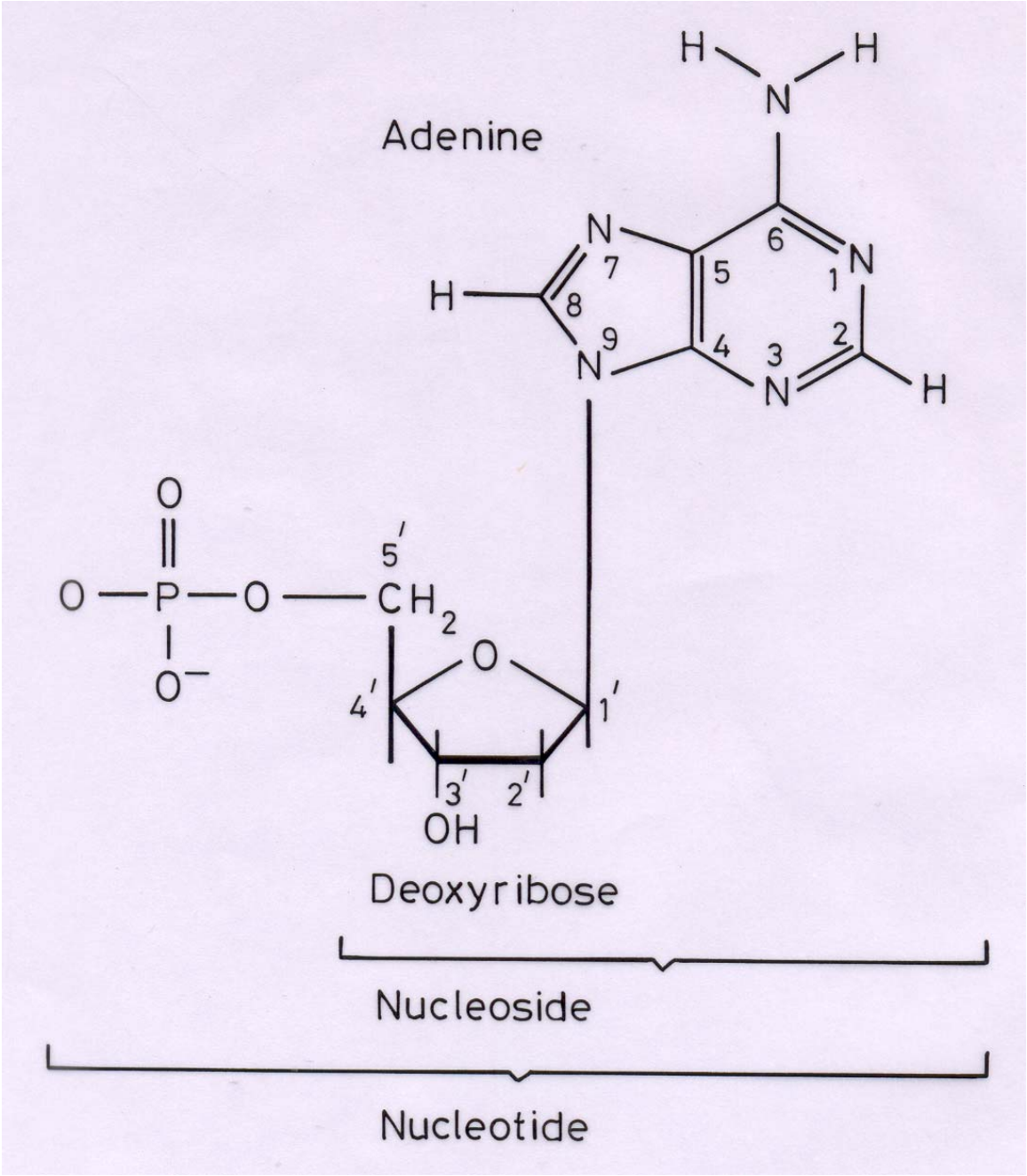
Feulgen, in 1912, discovered that when DNA is subjected to hydrolysis and Schiffs reagent is added to it, then a staining reaction occurs resulting in reddish purplish colour (Feulgen reaction). This specificity of Feulgen reaction paved the way for a number of quantitative studies, especially cytophotometry, involving the measurement of amount of light transmitted through Feulgen stained preparations. Through such studies, it has been established that the amount of DNA in the nuclei of cells of an organism is always constant, except for gametes (n) and polyploid cells.

Deoxyribonucleic acid (DNA) is a long polymeric molecule consisting of numerous basic units called monomers that are linked in a series and organized in a helix. The basic unit of DNA molecule is the nucleotide. DNA has three types of chemical components: a sugar called deoxyribose, four nitrogenous bases and phosphate group. The chemical components of DNA are arranged into groups called nucleotides.

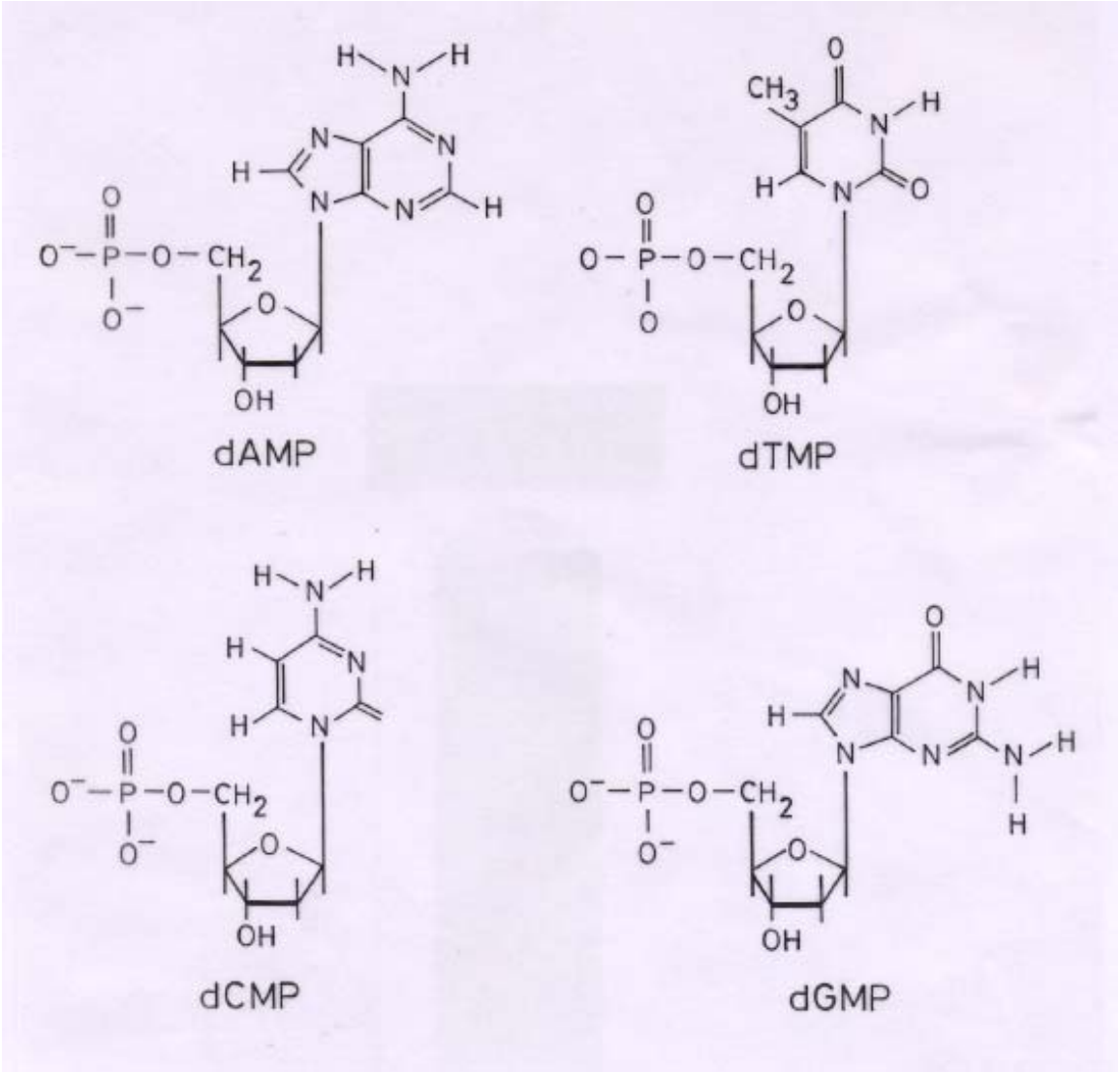
In a nucleotide, the sugar component is a pentose sugar called 2'-deoxyribose which has five carbon atoms. The pentose sugar can exist in two forms either as the straight chain (Fisher structure) or as a ring (Haworth structure) and it is the ring form of 2'-deoxyribose that occurs in the nucleotide of DNA. Four bases in DNA are adenine, guanine, cytosine and thymine. These bases can be grouped into two categories; purine and pyrimidine. Two

purine bases i.e. adenine and guanine have a double ring structure and the pyrimidine bases i.e. cytosine and thymine have a single ring structure.

The term used to describe the nucleic acid unit is based on three structural features i.e. the sugar, the type of base and the number of phosphate groups. The four different nucleotides that polymerize to form DNA are called 2'-deoxyadenosine 5'-triphosphate (dATP), 2'-deoxyguanosine 5'-triphosphate (dGTP), 2'-deoxycytidine 5'-triphosphate (dCTP) and 2'-deoxythymidine 5'-triphosphate (dTTP). In writing out DNA sequences of nucleotides found in a DNA molecule, these are abbreviated to just A, G, C and T. The individual nucleotides are joined together to form a polymer, called polynucleotide and is formed by attaching one nucleotide to another by way of the phosphate groups.



When a base is joined to sugar, the molecule is called a nucleoside. When phosphoric acid is attached to the 5'-carbon of sugar, it is called nucleotide. Upto three individual phosphate groups can be attached in series giving a nucleoside monophosphate (NMP), nucleoside diphosphate (NDP) and nucleoside triphosphate (NTP). These individual phosphate groups are designated as α, β, and γ. The α-phosphate group is the one directly attached to the sugar.



Deoxyadenosine monophosphate (dAMP), deoxythymidine monophosphate (dTMP), deoxycytidine monophosphate (dCMP) and deoxyguanosine monophosphate (dGMP)

In the DNA nucleotide, the carbon atoms in the sugar molecule are numbered in clockwise direction. They are always numbered in the same way with the carbon of carboxyl group (-C=O) occurring at one end of the chain form, numbered as 1'. The numbering of the carbon atoms is important as it indicates at what position on the sugar other components of the nucleotides are attached. The numbers are called 1' (one-prime'), 2' (two-prime'), 3' (three-prime) and so on. In fact, the prime is used to distinguish the carbon of the sugar from the carbon and nitrogen atoms in the nitrogenous base which are simple forms 1, 2, 3 and so on.

The nucleotides in the strand of DNA are covalently attached to each other in a linear fashion. The fifth carbon atom is always outside the ring. In a single nucleotide, the base (purine or pyrimidine) is always attached to 1'-carbon atom and the phosphate groups are attached to 5'-carbon of the pentose sugar. The number 3'-carbon

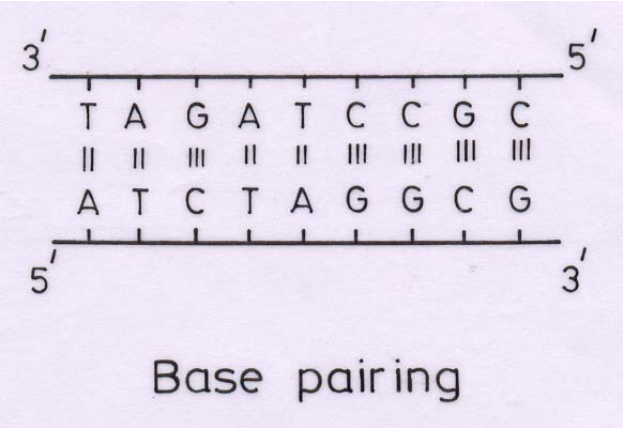
atom has –OH group which is important for allowing nucleotides to form covalent linkages with each other. It is the phosphate group that connects two sugar molecules together. For this reason, the linkage in DNA strands is referred to as phosphodiester linkage. In fact, the backbone of a DNA strand is formed by sugar and phosphate group. There are two ester (C-O-P) bonds in each linkage. Precisely it is referred to as 3'-5' phosphodiester bond.

The nucleotides are linked together by joining the α-phosphate group, attached to 5'-carbon of one nucleotide, to the 3'- carbon of the next nucleotide in the chain. Normally a polynucleotide is build up from nucleoside triphosphate subunits, so during polymerization the β- and γ-phosphate are removed. The hydroxyl group attached to 3'-carbon of the second nucleotide is also lost.

There are two distinct ends of the polynucleotide. At one end, 5'-carbon has not participated in a phosphodiester bond and β- and γ- phosphates are still in place. This end is called 5' or 5'-P terminus. The other end has unreacted hydroxyl group (3'-hydroxyl). This end is called the 3' or 3-OH terminus. Due to this, the polynucleotides have a direction which can be 5'-3' (down) or 3'-5' (up). The direction of the polynucleotide is very important in molecular genetics.

DNA molecules in chromosomes are much longer, possibly several million nucleotides in length. At any point in the chain the nucleotide could be A, G, C, T. If a chain has n nucleotides in length, it could have any one of 4^n different sequences. Nothing restricts the sequence of bases in a single strand, so any sequence could be present along one strand. This principle explains how only four bases in DNA can code for huge amount of information needed to make an organism.

In the four bases of DNA i.e. adenine (A), guanine (G), cytosine (C) and thymine (T), there is pairing between A and T and between G and C which is said to be complementary (Watson-Crick pairing). The base pairs lie almost flat, stacked on top of one another perpendicular to the long axis of double helix. Stacking adds to stability of DNA molecule by excluding water molecules from the spaces between the base pairs. While discussing a DNA molecule, biologists frequently refer to the individual strands as single stranded DNA and to the double helix as double stranded DNA or duplex DNA.



DNA differs from RNA in some respects. The sugar in RNA is ribose. In DNA, the deoxyribose has hydrogen (-H), while in RNA, the ribose has hydroxyl group (-OH) at number 2' carbon atom in the molecule. RNA contains uracil instead of thymine. RNA in the cell usually exists as a single polynucleotide chain whereas DNA is invariably in the form of two polynucleotides wrapped around one another to form double helix.

THE DOUBLE HELIX

DNA is composed of two side-by-side chains or strands of nucleotides twisted into the shape of a double helix. The two nucleotide strands are held together by weak associations between the bases of each strand, forming a structure like a spiral staircase. The backbone of each strand is a repeating phosphate-deoxyribose sugar polymer. The sugar-phosphate bonds in this backbone are called phosphodiester bonds. One part of the phosphodiester bond is between the phosphate and the 5' carbon of the deoxyribose and the other is between the phosphate and the 3' carbon of deoxyribose. Thus, there is polarity in each sugar-phosphate backbone and understanding this polarity is essential for knowing the functional aspects of DNA.

Thus double helix executes a turn every ten base pairs (10bp), the pitch of the double helix is 34 Å. The helix is 20 Å in diameter. The two strands of the double helix run antiparallel. One polynucleotide runs in 3'-5' direction while the other runs in 5'-3'. This antiparallel pattern provides stability to the helix.

In the double helix, the sugar-phosphate backbone of molecule is on the outside while the nitrogenous bases are stacked on the inside of the helix. Adenine on one strand pairs with thymine of the other strand with two hydrogen bonds and G of one strand pairs with C of the other strand with three hydrogen bonds. Helical shape of DNA depends entirely on the pairing and stacking of the bases in antiparallel strands. The most stable form that results from base stacking is a double helix with two distinct sizes of grooves. The helix has a major and a minor groove. Interaction between pairs of bases, one from each strand, holds the two strands of DNA molecule together.

It has been found that the DNA containing many GC base pairs is more stable than the DNA containing AT base pairs.

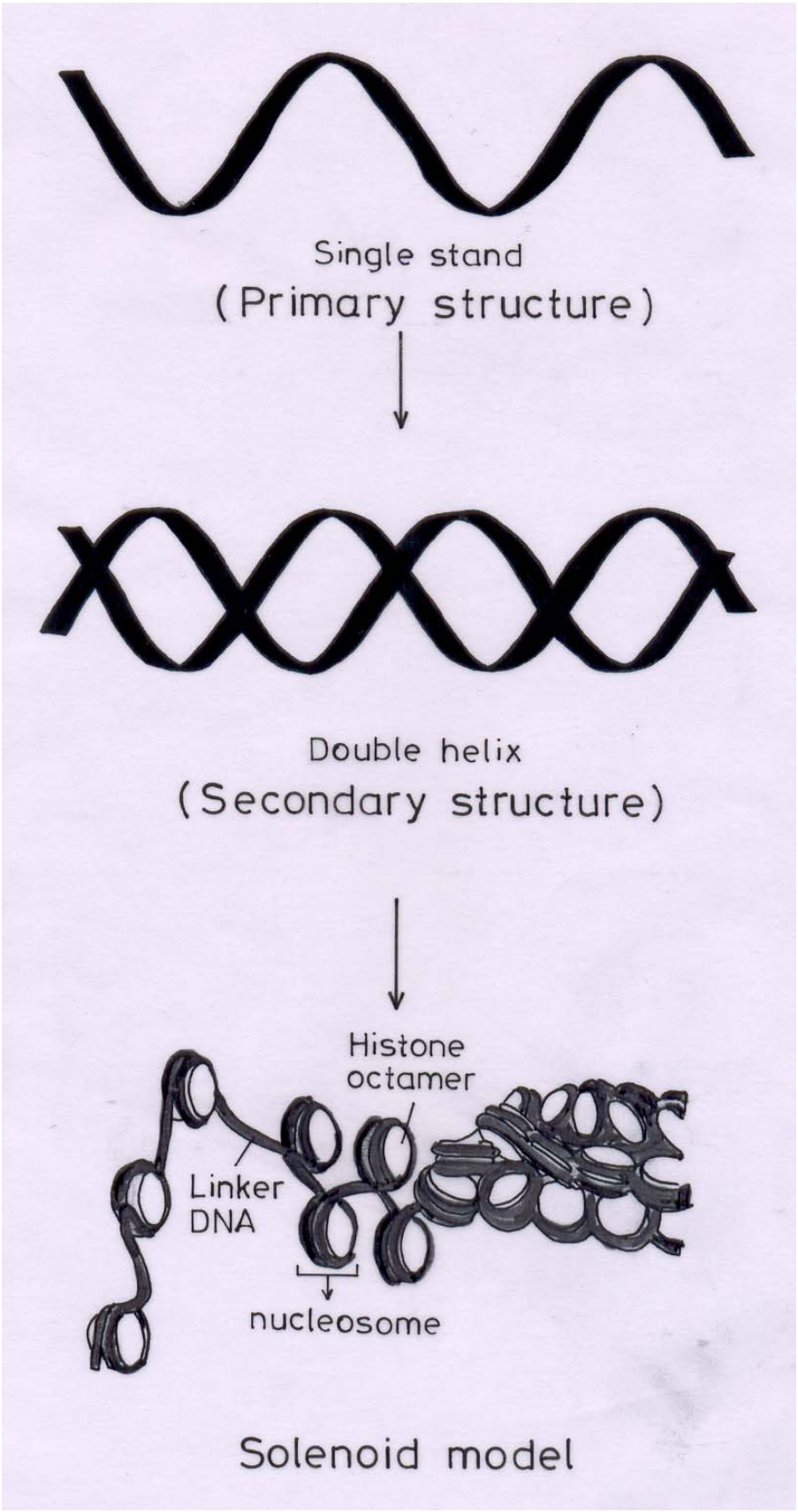
Hydrogen bonds are individually weak. The two strands of DNA molecule are held together in a relatively stable manner due to numerous such bonds. It is important that the strands be associated through such weak interactions since they have to be separated during DNA replication and transcription. Heat causes the two strands of DNA double helix to separate, a process called DNA melting or DNA denaturation. DNA with higher GC content requires higher temperature to melt down.

The discovery of double helix by Watson and Crick at Cambridge in 1953 was one of the greatest events in the history of molecular biology. The double helix satisfies certain conditions such as if the polynucleotide is coiled or folded in any other way then the various atoms must not be placed too close together and the new chemical should occur between atoms at the appropriate distance apart. The double helix model also took into account the results obtained by scientists Erwin Chargaff in USA and Rosalind Franklin in UK.

Chargaff carried out the analysis of chemical composition of DNA. Chargaff's base ratios paved the way for correct structure. He determined the exact amounts of each of four nitrogenous bases in samples of DNA purified from different tissues and different organisms. Chargaff and his colleagues revealed that in any sample of DNA, the number of adenine residues equals the number of thymine residues and the number of guanine residues equals number of cytosines, that is A=T and G=C. So the total purines (A+G) are equal to total pyrimidines (C+T). However, GC and AT contents of an organism's DNA vary considerably from species to species.

Rosalind Franklin at King's College, London carried X-ray diffraction analysis of DNA using technique developed by Maurice Wilkins. X-ray pattern obtained when a crystallized DNA fibre is bombarded with X-rays revealed that DNA is a helix with two regular periodicities of 3.4 Å and 34 Å along the axis of the molecule.

Taking into consideration all these aspects, Watson and Crick deduced that the only structure that fitted the facts, was "the double helix". For their discovery of 'the double helix' Watson and Crick shared the 1962 Nobel Prize with Maurice Wilkins.


Single stand
(Primary structure)
Double helix
(Secondary structure)
Histone
octamer
Linker
DNA
nucleosome
Solenoid model

Nucleotides are linked together to form strands of DNA and the linear sequence of nucleotides within a strand is known as the primary structure of DNA. DNA double helix can adopt several secondary structures known as A-DNA, B-DNA and Z-DNA. The folding and bending of secondary structures into a final three-dimensional structure is known as tertiary structure. Within living cells, DNA is associated with a variety of proteins that influence its final tertiary structure.

The double helix is right handed. Watson-Crick structure refers to the B-form. This is, in fact, the structure that DNA takes up in the cell. The two polynucleotides of the double helix are complementary, the sequence of one determining the sequence of the other. Complementary base pairing is also vital for the expression of biological information in a form utilizable by the cell.

A form differs from B form in relatively minor though important ways. A-DNA is still a double helix but it is more compact than B-DNA with 11 bp per turn of helix and a diameter of 23 Å. Z-DNA is the most strikingly different as in this structure the helix is left handed. Z-DNA can be observed in a double helix if that has a sequence made up of alternating purine and pyrimidine nucleotides. These forms of DNA received increasing attention as it is now clear that nucleotide sequence is one of the factors that influence the form taken by a segment of the double helix. The different forms of DNA are distinguished from one another by different dimensions to their double helices.

PACKAGING OF DNA INTO CHROMOSOMES - THE NUCLEOSOME MODEL

Genes in a prokaryotic cell, with the exception of few carried by plasmids, are organized into a single circular DNA molecule. However in eukaryotes, the nuclear genome is split into a number of individual DNA molecules, each of which is contained in a different chromosome.

A human cell contains about 2 meters of DNA (1m per chromosome set). This length of 2 meters of DNA is packed into 46 chromosomes, all inside a nucleus. This means all the DNA molecules in a human cell must be contained within cell with diameter of less than 10 μm, a compaction ratio of greater than 10^5.

DNA is folded into bead like structures called nucleosomes, which are repeating units of chromosome structure. These units are remarkably uniform in size, about 11 nm in diameter. Nucleosomes contain 200 base pairs of DNA, and a set of basic proteins called histones. Most eukaryotic cells contain five different kinds of histones called H1, H2A, H2B, H3 and H4. All these can be separated by gel electrophoresis. Histones are basic

proteins as these contain high proportion of basic amino acids. Non-histone proteins associated with DNA are acidic in nature and include various enzymes and proteins involved in the process of replication and transcription.

In a nucleosome, eight histones molecules (a pair each of H2A, H2B, H3 and H4) form a ball like structure around which DNA is wound almost twice. The centre of this structure has a tetramer formed by two molecules each of H3 and H4. The tetramer is flanked by dimers of H2A and H2B at top left and lower right. A single histone molecule of histone H1 binds to 'linker' DNA outside the nucleosome to keep the DNA in place. When treated with DNase, the linker DNA is digested and a nucleosome core particle with about 150 bp of DNA is released.

There is universality of nucleosomes in eukaryotes. Histones are remarkable in that they are highly conserved between different species. For example, the histone H4 from cow differs from H4 from pea plant by only two amino acids out of 102 and these are conservative substitutions that do not change the protein structure significantly. The most likely role for nucleosomes is, of course, a structural one. The nucleosomes provide the first order of condensing or coiling of extremely long, thin chromosomal fibre.

The overall mixture of material that comprises chromosomes is given the general name chromatin and it stains most strongly with chromosome specific dyes. DNA in a chromosome is extremely longer in comparison to its width (more than ten million to one). DNA molecule is complexed with proteins to form a structure called chromatin and is coiled up in a highly organized way. Electron microscope observations of chromatin also showed linear arrays of structures called nucleosomes. Packaging of DNA around nucleosome reduces its length while forming chromatin fibre. Higher level of packaging comes into play during cell division when metaphase chromosomes, the most highly organized structures, become visible in condensed state.

Chromatin in interphase nuclei takes on various thicknesses. One of these, 11 nm in thickness, simply represents a string of nucleosomes, or nucleosome fibre. The next thickness, about 25 nm, is due to further winding of nucleosomes to form a hollow coil called solenoid. Formation of solenoid was traced by making electron micrographs of chromatin in solutions of increasing salt concentration. At very low concentration, the chromatin appears as a string of nucleosomes. As the salt concentration increases, the coiling takes place to form a solenoid structures. Interaction of H1 of one nucleosome with H1 of the other causes chromatin folding which is followed by supercoiling. DNA in a chromosome is packaged in an organized fashion so that the genes are accessible and molecules do not get tangled up during replication. There is a highly organized packaging system to fit lengthy DNA molecules into such small structures.

In first order of coiling, DNA binds onto histones which act somewhat like spools. In further coiling, solenoid contractions occur. Further size reduction is achieved by supercoiled loops radiating from a core called scaffold comprising non-histone proteins. The coiling converts solenoids into the three dimensional structure called chromosome.

The diameter of supercoils is the same as the diameter of chromosomes during cell division. It is evident from observation on chromosomes from which histones have been removed chemically. After such treatment, the chromosomes have a densely staining central core of non-histone protein called the scaffold. Projecting from this protein scaffold are loops of DNA. The fibres forming these loops are believed to be the solenoids. In electron micrograph, each DNA loop begins and ends at the scaffold. Central scaffold is largely composed of the enzyme topoisomerase II, which has the ability to pass a strand of DNA through another strand.

Solenoids arrange in loops emanating from the central scaffold matrix, which itself is in the form of a spiral. The loops attach to scaffold by special regions along the DNA called scaffold attachment regions. The progressive levels of chromosome packing are as follows. DNA binds onto nucleosome spools. The nucleosome chain coils into a solenoid. The solenoid forms loops that attach to a central scaffold. The scaffold plus loops arrange into a giant supercoil.

During early stages of meiotic division, there are bead like localized thickenings found along the chromosomes. These unique meiotic structures are called chromomeres. They are assigned as Giemsa-positive bands, whereas interchromomere regions are numbered as Giemsa-negative bands. The positions of the chromomeres are the same in all homologous chromosomes.

SATELLITE AND REPETITIVE DNA

Satellite DNA refers to the highly repeated sequences with such a uniform nucleotide composition that upon fractionalization of the genomic DNA and separation by density gradient centrifugation, they form one or more bands that are clearly different from the main band of DNA. So the term satellite DNA is derived from the way in which repetitive DNA is prepared as a pure fraction separate from the rest of DNA. DNA molecule migrates to the position where the density of the solute (CsCl) is equal to its buoyant density. The buoyant density of a DNA molecule depends on its GC content. The base composition of satellite DNA differs from that of the majority of DNA in a eukaryotic species i.e. it is either AT rich or GC rich.

In human genome, extensive tracts of repeat sequences are arranged into long tandem arrays. This type of repetitive DNA is called satellite DNA. Human DNA has GC content that forms a main band in a density gradient. Three satellite bands also appear at other positions in the density gradient. These additional bands contain repetitive DNA. When the chromosome molecule is cleaved into fragments, the single copy DNA, rich in GC content, is close to main DNA band while the repetitive DNA migrates to satellite band position. Based on length of the cluster, satellite DNA is of different types. Classical satellite DNA refers to cluster between 100 to 5000 kb length e.g. centromeric clusters of alpha repeats. Minisatellite DNA refers to shorter clusters of 100bp to 20 kb such as telomeric repeats of 10 to 15 kb. And the microsatellite DNA has generally less than 4 base pairs repeats. The dinucleotide CA/GT repeats are also very common.

There are some repetitious DNA sequences that have no function in the life cycle but play an important role in evolution. These are not found in constant positions in the DNA and are called as mobile elements. Such mobile DNA elements which are present both in prokaryotes and eukaryotes can cause mutations when they move to new sites in the genome. The process by which these sequences are copied and inserted into a new site in the genome is called transposition.

CHROMOSOMAL ALTERATIONS

Alterations in chromosome structure play an important role in the evolution of new species. We know that many chromosomal aberrations are distinctly harmful. But certain chromosomal changes have an important commercial value as is observed in cultivated plants. They appear to play an important role in the process of evolution through the origin of new species as well as through changes in the chromosome pattern within the species.

Any variation in chromosome structure can have major effects on the phenotype of an organism. For example, several human diseases can be caused by changes in chromosome structure. Chromosomal changes are often the result of chromosomal breaks, in which same broken ends do not reunite.

The change in the chromosomal constitution of an organism may occur in two ways; either by altering the chromosome structure or by altering the chromosome number. Besides their immediate effect on the chromosomes, both types of aberrations may have major consequences. For example, individual heterozygous for chromosomes with different structures often have lowered fertility and individuals with altered number of chromosomes may even be inviable or stertile.

CHANGES IN CHROMOSOMAL STRUCTURE

There are diverse mechanisms by which the structure of chromosome is altered. Either the total amount of genetic material within a chromosome increases or decreases significantly or the genetic material in one or more chromosomes is rearranged without affecting the total amount of genetic material. Mutations have great impact in altering chromosome structure. Some of changes in chromosomal structure are deletions, duplications, inversions and translocations. Of these, inversions and translocations involve the changes in the order of genes that result in the position effect. As the position effect causes phenotypic changes, it acts as a source of genetic variation. Alterations in chromosomal structure are clearly visible in the polytene chromosomes in the salivary glands of *Drosophila.* As the polytene chromosomes remain in somatic synapse, it is easy to recognize different kinds of modifications and identify their locations on chromosomes

Deletions

If a chromosomal segment is missing, it is termed a deletion or a deficiency. It may be interstitial deletion if an interstitial part of a chromosome is missing or it may be terminal if there is one break at one end of the chromosome. Generally, the absence of the interstitial part of chromosome is called deletion and that of the terminal part is called deficiency.

As deletions result in loss of genes, these are often lethal when individual is homozygous. In case of heterozygous individual, these can cause abnormal development. For example, in heterozygous humans, the deletion of substantial part of short arm of chromosome 5 (5p), causes the cri-du-chat (cry-of-the-cat) syndrome. This syndrome shows mental retardation and makes a peculiar cat like sound. Deletions can be identified and used to map the sequence of genes on the chromosomes. If the deletion occurs, recessive alleles on normal chromosome are expressed.

Duplications

Duplications arise when a segment of chromosome occurs more than once in a single chromosome. Duplication may occur adjacent to original chromosomal region. When this occurs, the order may either be the same as the original order, called tandem duplication, or there may be the opposite order, called a reverse duplication. The duplicated region may not be adjacent to original segment resulting in a displaced duplication which may be on the same chromosome or on another chromosome. Bar eye in *Drosophila* is due to duplication in the X chromosome at 16A segment. In this case, the eye of heterozygous female is somewhat smaller than the

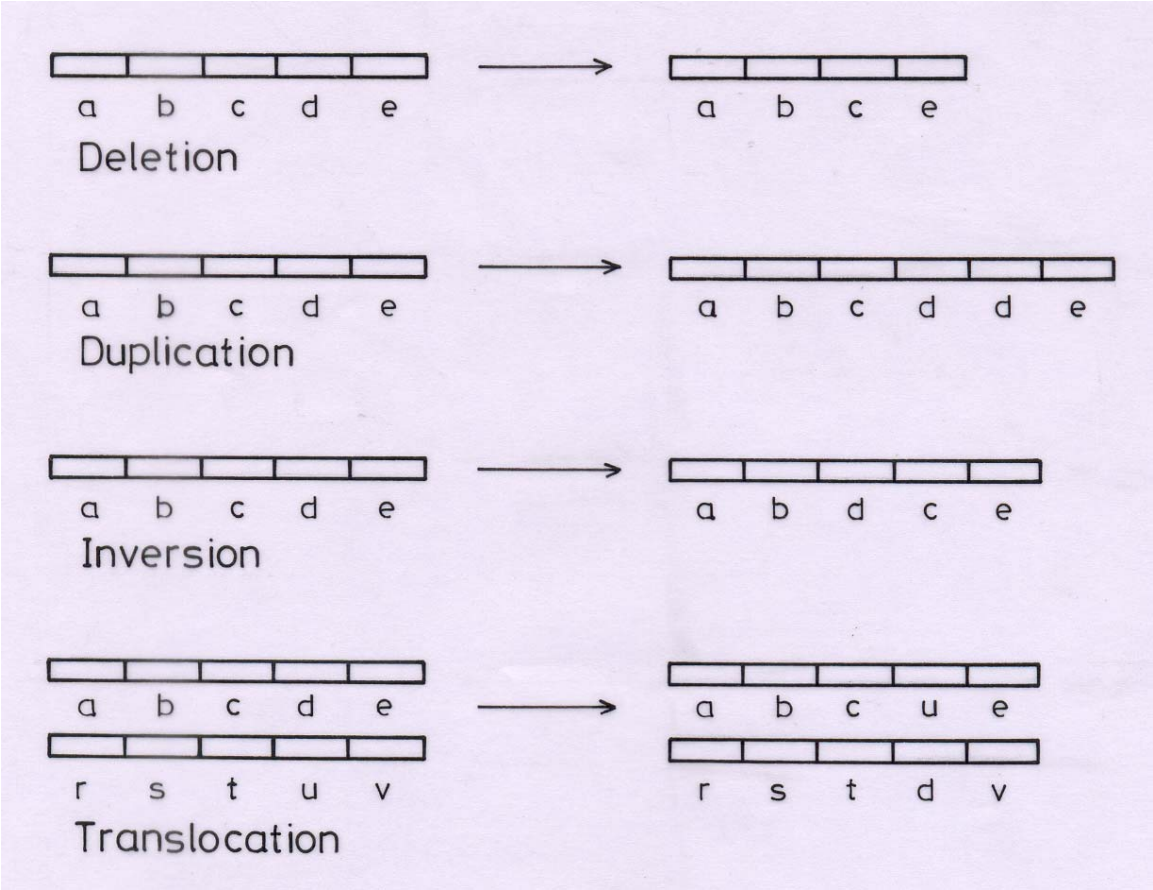
normal eye and the sides are straighter, giving an oblong or bar appearance. In the homozygous or hemizygous condition, the eye is considerably smaller.

In a chromosome, tandem duplication can be generated when homologues overlap and there are simultaneous breaks in the two homologous chromosomes at different points. If the homologues reunite, one chromosome will have tandem duplication and the other a deletion. Unequal crossing may also lead to duplications or deletions. If a heterozygous individual has one homologous chromosome with duplication and other a normal chromosome, the duplicated region does not have a homologous segment to pair with in meiosis. Thus, formation of loop like structure of the duplicated region occurs when such homologues pair.

If the duplicated segments are small, these may not have any effect on the viability of the individual but these exhibit some phenotypic effects. Viable individuals provide a potential for further evolutionary changes in these extra genes. In fact, it is thought that this happened with the different globin genes, the genes that code for the components of the protein haemoglobin. Duplications of the ancestral genes may have led to the divergence of duplicated genes in their function during the course of evolution.

Inversions

If the sequence of the genes on the chromosomes is inverted, it is called inversion. Homologous chromosomes have genes in the same sequence. However, in some cases, the sequence may differ on different chromosomes. Inversion can be generated by a simultaneous break at two points in a chromosome followed by joining of the fragment in inverted orientation. Based on the position of the centromere, inversions may be of two different types. It may be pericentric if the inverted segment involves the centromere or it may be paracentric if it does not involve the centromere. When the homologues pair, one twists on itself and makes a loop and other makes a loop without a twist. Generally, inversions survive the meiotic process and viable gametes are produced.



Translocations

When a part of one chromosome is joined to a part of a different chromosome, the aberration is called a translocation. In fact, a translocation is the movement of a chromosomal segment from one chromosome to another non-homologous chromosome. This occurs as a result of breaking and rejoining. When two breaks occur in one chromosome and one in the non-homologous chromosome the fragment from the first chromosome may join in intercalary position in the second chromosome. The size of the chromosome and the position of the centromere may change as a result of such alteration. Even the action of the gene is depressed if it is translocated to heterochromatic region.

Reciprocal translocations can be homozygotic or heterozygotic. As in reciprocal homozygote, both pairs of each of the non-homologous chromosomes carry same reciprocal translocation, so there is no problem in meiosis. However, in case of translocation heterozygote, where only one member of homologous pair is involved in translocation, considerable irregularities are seen in meiosis. As the products of meiosis are unbalanced, so in gametes, either duplications or deletions occur. Several human diseases are caused as a result of translocation.

Some cases of Down's syndrome are due to Robertsonian translocation between the long arm of chromosome 21 and any other acrocentric, usually chromosome 14.

VARIATION IN CHROMOSOME NUMBER

Variation in chromosome numbers in organisms occurs frequently in nature. The number of chromosomes may vary in two basic ways i.e. either the whole set of chromosomes differs resulting in euploid individual or the number of a particular chromosome is not diploid resulting in aneuploid individual. Changes in chromosome number in the form of euploidy or aneuploidy generally have an even greater effect on survival than do changes in chromosome structure. In humans, more than half of the spontaneous abortions that occur in the early months of pregnancy involve foetuses with aneuploidy, polyploidy or other large chromosomal aberrations.

There are other mechanisms also by which chromosome number can change. For example, if reciprocal translocations occur between two acrocentric chromosomes such that the large segments reattach, the result is large metacentric chromosome and a small chromosome that can be lost during cell division. Even two acrocentric chromosomes fuse to form one metacentric chromosome resulting in variation in chromosome number. Generally two non-homologous chromosomes join at their centromeres to form a single metacentric chromosome. Sometimes, a chromosome may also split at the centromere perpendicular to the length of chromosome, resulting in two smaller chromosomes. Chromosomal fusions are more common than chromosomal fissions.

Aneuploidy

In case of aneuploidy, a set of chromosomes has variable complement. It is due to non-disjunction. The term non-disjunction refers to the condition when chromosomes do not separate properly during anaphase of meiosis or mitosis. So it is either meiotic non-disjunction which produces unusual gametes or mitotic non-disjunction which occurs in somatic cells after fertilization. In meiosis, both the homologous chromosomes go to same pole leaving one daughter cell with an extra chromosome and other daughter cell without that chromosome. When these gametes are fertilized by normal gamete, they either have an extra chromosome (n+1) and this is termed as trisomy or may not have a chromosome (n-1) and this is called monosomy. Non-disjunction may occur either in meiosis I or II but generally it occurs in meiosis I. In mitotic non-disjunction, the organism has normal and aneuploid cells that are genetically different from the rest of body. This condition occurs more frequently in *Drosophila* and is referred to as mosaicism.

Non disjunction may occur in chromosome pairs, both autosomes as well as sex chromosomes. This results in the formation of monosomics and trisomics. The term monosomic refers to those cases where there is

only one chromosome rather than normal two in the cell. Trisomic refers to the condition where one chromosome is in triplicate. Aneuploid variations form a series as is found in species of *Crepis* (x=3, 4, 5, 6 and 7). Species with x=7 is the most primitive while advanced species have lower numbers.

Take the example of human beings. There are many cases of aneuploids such as Down's syndrome (trisomy in chromosome 21), Edward's syndrome (trisomy in chromosome 18), Patau syndrome (trisomy in chromosome 13), in addition to Klinefelter (47, XXY) and Turner (45, XO) syndrome. Klinefelter syndrome with XXY (or 47), is sterile male with some female characteristics. Turner syndrome (XO or 45) is sterile female, short in stature with some neck webbing.

In Down's syndrome, the smallest chromosome (number 21) of group G of autosomes exists in triplicate. So in cytological preparation, 47 chromosomes are seen. In some cases, extra chromosome exists as a translocated attachment to chromosome 14 and in these cases the cytological preparations show 46 chromosomes. However, phenotypes of both types are the same. The loss of an autosome chromosome is generally detrimental as generally such monosomics in humans do not survive. In plants, monosomics and even nullisomics (2n-2) are viable.

In many mammalian species including human, X chromosomes are different from other chromosomes in that only one is active in a given cell. In normal females, only one X is active in a given cell and the other X is heterochromatic and remains in a condensed state throughout the interphase. The inactivated X forms a structure called Barr body (named after its discoverer Murray Barr) that can be identified in a cell. Normal males and XO individuals have no Barr body while normal females have one Barr body and XXX individuals have two. By counting the number of Barr bodies in a cell, chromosomal abnormalities involving X chromosomes can be determined.

Polyploidy

New species arise through polyploidy. In polyploid organisms, there are three or more complete sets of chromosomes. If x is the haploid number of chromosomes, the organisms with three sets of chromosomes (3x) are called triploids. Similarly, 4x are called tetraploids and 6x are called hexaploids. Triploids are common in grasses, vegetables and flower varieties. Seedless fruits such as watermelon and banana are also triploids. These are propagated through grafting and budding. Tulips are also triploid (3n) and propagated vegetatively. Usually, 'x' refers to the number of chromosomes in a set and 'n' to the number of chromosomes in a gamete. If we take into consideration a hexaploid organism with 42 chromosomes then the basic number (x) in this species is 7 and the number (n) in the gametes is 21.

Polyploidy is relatively common in plants but rare in animals. Nearly half of the flowering plants are polyploids and so are many important crops. Potatoes are tetraploids (4x=48) and bread wheat is hexaploid (6x=42). In animals, it is reported to occur only in certain groups such as beetles, earthworms, salamanders and fishes.

There are several reasons of polyploidy being less frequent in animals than plants. Firstly, sex determination is more sensitive to polyploidy in animals than in plants. Plants can often self-fertilize, so a single new polyploid plant with an even number of chromosomal sets (tetraploid, hexaploid etc.) can still reproduce. Plants generally hybridize more easily with other related species. This is an important attribute as different sets of chromosomes in a polyploid often have different origins.

There are two distinct types of polyploids. The polyploids that receive all their chromosomal sets from the same species are called autopolyploids. And the polyploids that obtain their chromosomal sets from different species are called allopolyploids. If diploid or unreduced pollen from a diploid organism fertilizes a diploid egg of the same species, the offspring are autotetraploids AAAA where A is complete chromosomal set of type A. All chromosomal sets in an organism are homologous, just as they are in a diploid.

Polyploids occur naturally but their frequency is very low in nature. If in mitosis, all chromosomes go to one pole this will result in autotetraploid chromosome number. Polyploids are often large than related diploids. Many food crops are autotetraploids or other types of polyploids. Autotetraploids may have a normal meiosis if they form either bivalent or quadrivalents.

Polyploids can be produced artificially using colchicine, a chemical that interferes with the formation of spindle fibres. As a result of colchicine application, chromosomes do not move to poles and autotetraploids are often formed. Colcichine is extremely valuable in horticulture as polyploids plants often yield products which are commercially superior.

If diploid pollen of one species fertilizes a diploid egg of another related species, the offspring are allotetraploid or AABB where A indicates the genome of species A and B indicates the genome of B species. Triploids are generally sterile as all the gametes formed are unbalanced and generally non-functional. Bananas are triploids, they produce unbalanced gametes and as a result are seedless.

Allopolyploids occur more frequently in nature and thus result in production of new species. For example, bread wheat *Triticum aestivum* is an allohexaploid with 42 chromosomes. It descended from three

different diploid ancenstors, each of which contributed two sets of chromosomes. Pairing occurs only between the homologous sets so that meiosis is normal and that results in formation of viable gametes with n=21.

A haploid pollen grain with genome A may pollinate a flower of a species with genome B resulting in sterile hybrids. If mitotic failure takes place even in one branch, it results in AABB allopolyploid. The cultivated species of genus *Brassica* i.e. *Brassica napus* has 2n=38. This number has arisen by the duplication of entire chromosome complement of a hybrid between *Brassica campestris* (2n=20) and *Brassica oleracea* (2n=18). Even intergeneric cross between radish genus *Raphanus* (2n=18) and cabbage genus *Brassica* (2n=18 resulted in hybrids which are sterile. However, tetraploid forms (*Raphanobrassica)* are fertile and viable.

SEX CHROMOSOMES

In an organism, the complement has autosomes and the sex chromosomes. The chromosomes that determine the sex of the individual are called sex chromosomes and the other pairs in the complement are autosomes. In most sex chromosomes, one chromosome is called X and the other is called Y chromosome. Sex determining characteristics are determined by the genes located on sex chromosomes.

In animals, the sex determination mechanism may be XX-XY (mammals, humans), XX-XO (aphids, grasshoppers), ZZ-ZW (birds) or haplodiploid (honeybee). Sex differences observed in the morphology, physiology and other characteristics, are controlled by the genes present on the sex chromosomes. In majority of animals, there is distinct sexual dimorphism. The X chromosmes contain genes that are essential for both the sexes. At least one copy of an X chromosome is required for the development. A female needs two copies of X chromosome to be fertile. The male determining genes are located on the Y chromosome.

Most of the plants are monoecious, so in such cases, sexual differentiation is limited to the male and female reproductive organs only which are present in the flowers. In dioecious plants, the male and female plants exist and generally there is XY system of sex determination.