07

Molecular Biology and Recombinant DNA Technology

Nucleic acids are capable in directing their own replication from their monomers. In most organisms, DNA is found as the genetic material. However , some viruses such as influenza virus, consist RNA as their genetic material. Accurate replication of DNA, its transmission from one generation to the other and its ability to store and express hereditary information facilitates DNA for acting as vital genetic material in organisms.

DNA Double IIelix Model: James Watson and Francis Crick, based on the X-ray crystallography of a DNA molecule obtained by Rosalind Franklin, proposed the double helix model to describe how the six molecules; de-oxy ribose sugar, phosphate group and four different nitrogenous bases, are organized into DNA and explain its properties. According to this model, DNA is a twisted ladder (spiral staircase) in which the rails are composed of alternating phosphate and sugar molecules forming the back bone. The steps of the ladder are paired nitrogenous bases. A purine is paired with a pyrimidine as per the base pairing rule with either two hydrogen bonds (A=T) or three (G ≡ C). T.H. Morgan and his group, through their experiments concluded that chrosomes are made up of DNA and proteins and genes as regions of chromosomes.

Architecture of Chromosomes

Architecture of chromosomes is the way the DNA molecules are arranged in the nucleoid or nuclear region of the cytoplasm of the prokaryotic cells or in the nucleus of eukaryotic cells.

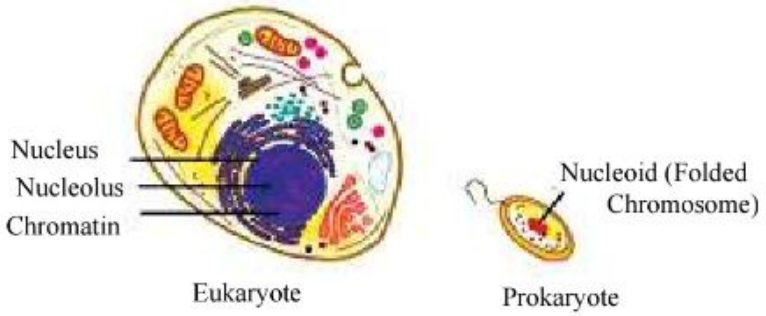


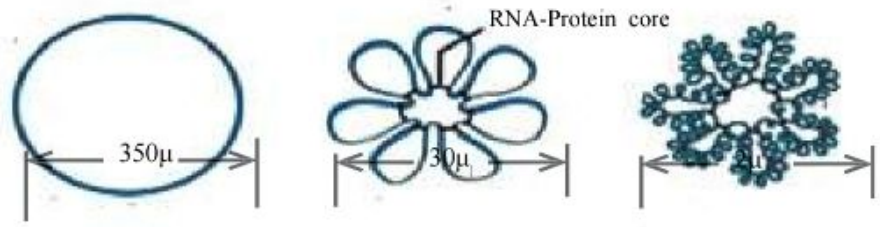
Figure 7.1: DNA is packaged into a nucleus in eukaryotes and a nucleoid in prokaryotes.

DNA of both prokaryotes and eukaryotes are called chromosomes. However, true chromosomes are present only in eukaryotes.

Prokaryotic (bacterial) chromosome is a single double stranded circular DNA molecule associated with a few protein molecules. Eukaryotes have several chromosomes, each consisting of a single double stranded linear DNA associated with histones and other protein molecules.

When the size of all chromosomes of an organism are considered, it is an enormous amount of DNA. These facts illustrate that, a cell has a huge problem in having its DNA either in the nucleoid of prokaryotes or the nucleus of eukaryotes. The containment of the genome/DNA in the nucleoid or in the nucleus is called DNA packaging.

Packaging of DNA in prokaryotes is facilitated by the proteins associated with the DNA molecule. The protein molecules cause the DNA to coil (or fold or loop) and supercoil, so the DNA molecule is compacted into the nucleoid. The DNA molecule is initially coiled into loops and these loops then independently supercoil into domains identifiable in electron micrographs. The loops of compacted mass of DNA are bound to a 'core' consisting of RNA and protein. The 'core' also attaches the chromosome to the membrane. The supercoiled DNA can be relaxed by introducing single strand nicks. Since the chromosomes are attached to the membrane and held together by the 'core' acting as a barrier for rotation, the domains can relax and supercoil independently. This is important in the transcription of specific genes. Removal of RNA will lead to loss of independence of the loops.



a) Circular, unfolded Chromosome

b) Folded chromosome 40 to 50 loops

c) Supercoiled, chromosome

Figure 7.2:Folding and supercoiling compact the prokaryotic chromosomes

In addition to the chromosomal DNA, certain prokaryotes have extra chromosomal genetic elements called plasmids. They are also circular DNA and are coiled and supercoiled.

Eukaryotic chromosomes are associated with a large number of proteins called histones that helps to organize the DNA inside the nucleus. This DNA-protein complex is known as chromatin. The chromatin may be lightly packed as in euchromatin or tightly packed as in heterochromatin. Euchromatin is rich in genes and is probably active in transcription. Heterochromatin consists nucleotide sequences which are mostly inactive. They may contribute in gene regulation, epigenetic inheritance and protection from chromosomal integrity.

In the first level, the double helix winds around a complex of eight histone molecules. These are called nucleosomes and they look like beads of a necklace. The adjoining beads of nucleosomes are linked together by a stretch of DNA: linker DNA. (Figure 7.3)

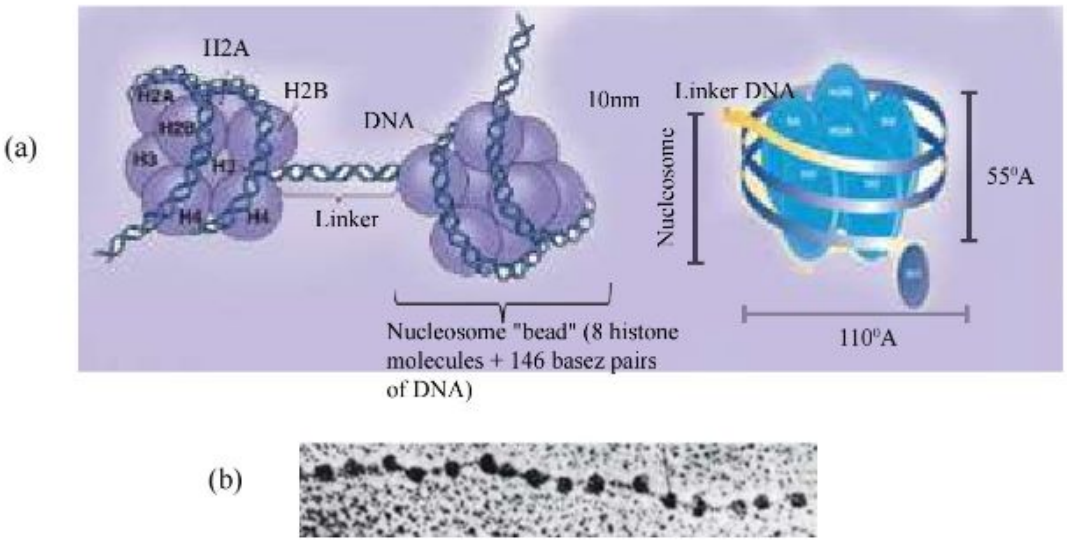


Figure 7.3: a) First level of packing: Forming beads of nucleosomes joined together by linker DNA, b) Electron micrograph showing nucleosome (beads) and linkers (strings)

In the second level, the nucleosomes twist and pack in a spiral fashion to form a chromatin fibre of roughly 30 nm in diameter:30 nm fibres from 10 nm fibres. (Figure 7.4)

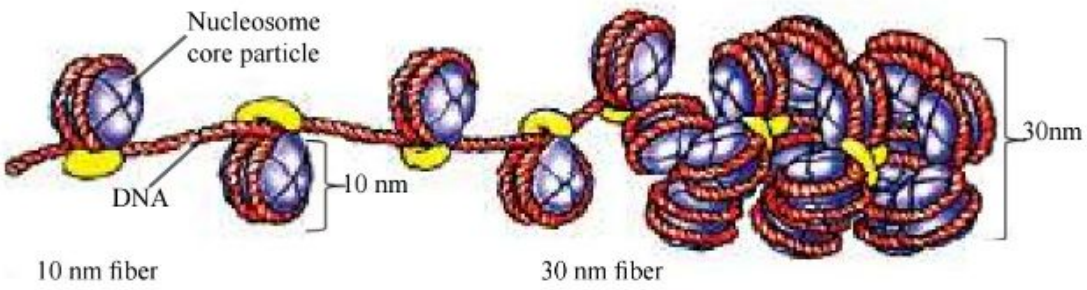


Figure 7.4: Formation of a 30 nm fiber (Cylindrically coiled)

At the third level, the 30 nm fiber forms loops, called looped domains, attached to a protein scaffold. This structure has a thickness of 300 nm. (Figure 7.5)

Finally, at the fourth level, the looped domains coil, fold and further compact to form the mitotic chromosome. The diameter of a chromatid is about 700 nm. In the metaphase chromosome, the chromatids are already replicated (Figure 7.5).

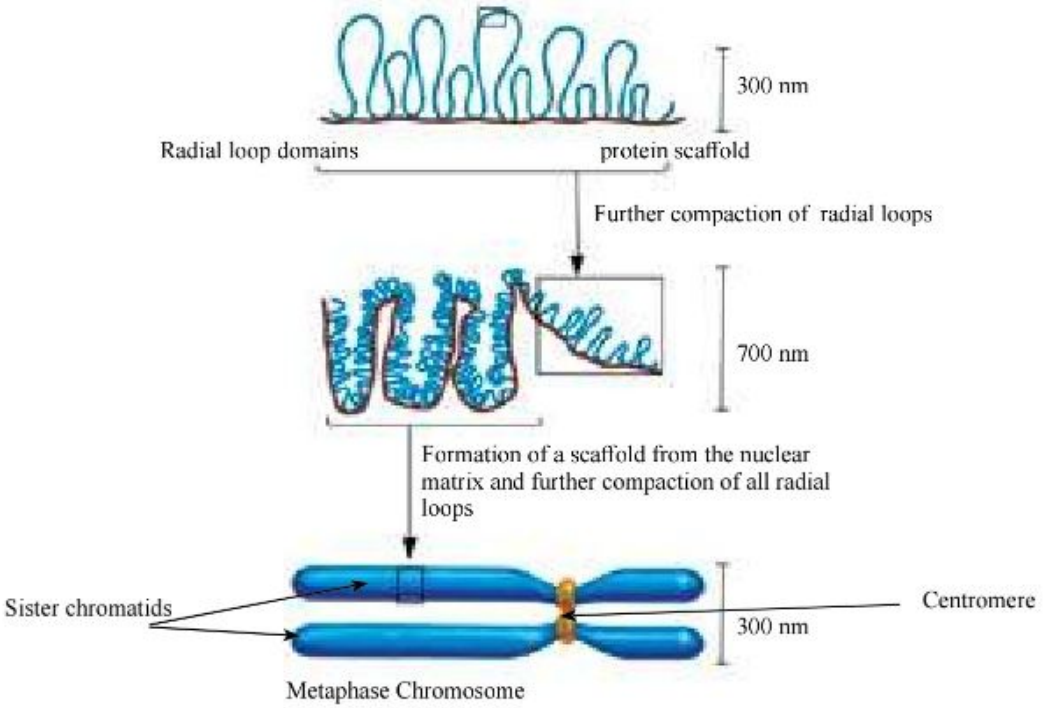


Figure 7.5: Compaction into looped domains on scaffold and into chromatids

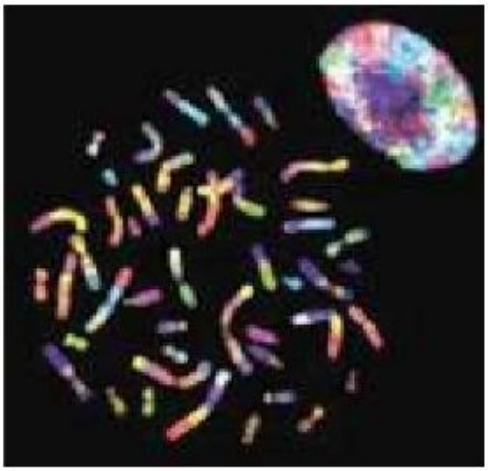


Figure 7.6: Metaphase chromosomes (discrete units) and chromatin in the interphase

DNA Replication -

This is the process which copies a double stranded DNA molecule to produce two identical copies. Because, DNA replication process is basically similar in prokaryotes and eukaryotes. However, there are differences in enzymes involved. This is because the Eukaryotic nuclear DNA is organized into chromosomes and contain histones in their structure for packaging, and the prokaryotic DNA are usually found as circular molecules, which are supercoiled for packaging.

Importance of DNA replication

- The essential information for life is stored in DNA. Therefore, the new cells produced must receive the DNA from their parent cells as, each cell in the body of a diploid organism should contain the same genetic information as was in the zygote. A multicellular organism grows by addition of new cells.
- Damaged or dead cells are also replaced by new cells.
- In asexual reproduction, the offspring is identical to the parent cell. This is possible only because DNA has the ability to replicate and provide an identical set of genetic information stored in DNA to each daughter cell by mitosis.
- In sexually reproducing organisms, meiosis occurs at some point of life cycle in order to maintain a constant number of chromosomes. DNA replication occurs prior to meiosis.
- DNA replication is a very accurate process, so that it makes identical copies. However, rare errors occur in DNA replication, introducing mutations which results in variation. Variation leads to evolution of organisms.
- Therefore, DNA replication is important in maintenance of the life of an individual organism and continuation of a species.

The process of DNA Replication in prokaryotes

The entire replication process is carefully controlled and coordinated by a number of enzymes and other proteins.

DNA is synthesized on the existing DNA strands of the double helix of the DNA molecule. The parent DNA strands are used as the templates. Thus, the newly synthesized DNA double helix contains one parent DNA strand and a new complimentary strand.

First of all, the tightly packed DNA (supercoiled DNA in prokaryotes and chromatin in eukaryotes) has to relax, so that the DNA replication machinery can have access to the place where the replication begins.

The separation of the double helix occurs at the origin of replication. Origin of replication or 'Ori' is a specific sequence of DNA to which proteins that initiate DNA replication bind. Starting from this, the entire circular DNA is replicated on both directions. Since the enzyme synthesizing the new DNA strand can move only in one direction (5' to 3' direction) one of the new strands can be synthesized continuously while the other will be synthesized as small fragments. The

small fragments of lagging strand are called okazaki fragments. These are called leading and lagging strands respectively. Replication of large DNA molecules may begin at several origins of replication to speed up the process.

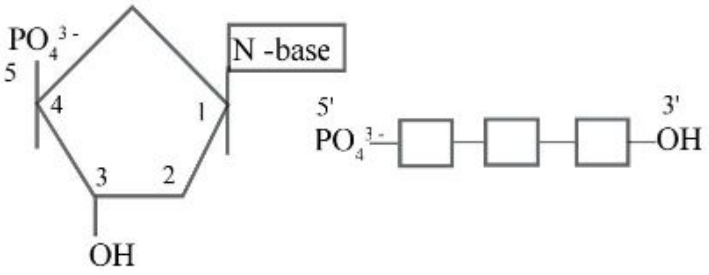


Figure 7.7: 5'- phosphate and 3'- OH of a DNA Molecule

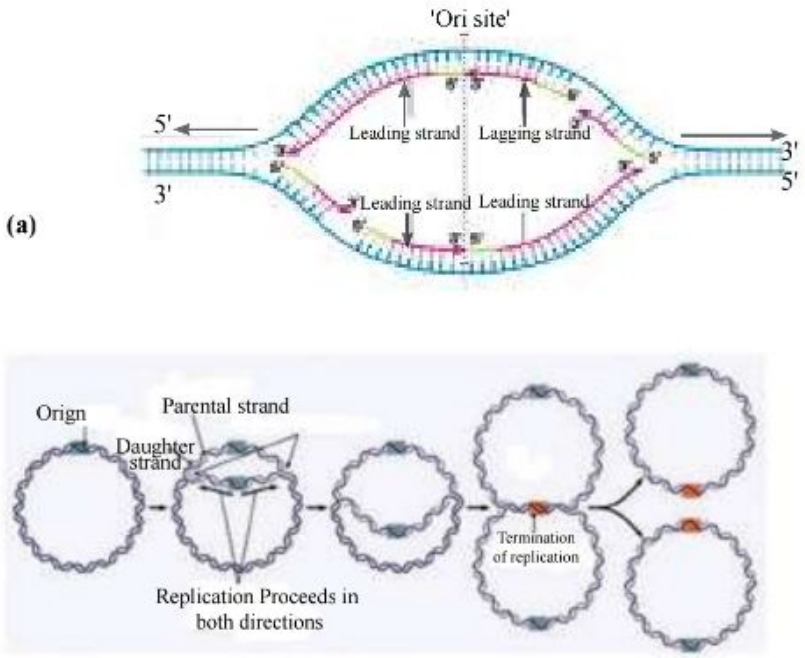


Figure 7.8: (a) Details of DNA a sreplication (b) DNA replication of small circular DNA

Functions of major enzymes and other proteins of replication machinery

A number of d enzymes and other proteins are required in the replication of DNA. These proteins assemble at the orign of replication. The major enzymes involved in the DNA replication are: helicase, topoisomerase, primase, DNA polymerases and DNA ligase. There are several other proteins in the replication machinery including Single Strand Binding (SSB) proteins.

Helicase: These are the enzymes that unwind the double helix and separate the two strands of a DNA molecule, with the expense of energy as ATP. The separation occurs by breaking hydrogen bonds between the base pairs of the two strands. This is important for exposing the two strands to function as template for new DNA synthesis (replication).

Topoisomerase: These enzymes are working ahead of the direction of DNA synthesis. The untwisting of the strands at one place causes further twisting and strain on other places. Topoisomerase enzymes, introduce breaks on one or both strands of DNA, twist the molecule to relieve the strain and then reseal the cut ends.

Single Strand Binding Proteins (SSB): These protein molecules bind to exposed single stranded DNA to prevent re-pairing of the seperated DNA strands and to stabilize them. If the two strands re-pair, then they cannot serve as templates for new DNA synthesis.

Primase: In order to synthesize a new DNA strand on a template, complementary deoxyribonucleotides should be added one after the other in the correct sequence. This is done by DNA polymerase, but DNA polymerases can add nucleotides only to the 3' end of an already formed nucleic acid chain. For this, a small fragment of nucleic acid chain is sufficient and this fragment is called a primer. Primase is a type of RNA polymerase which can initiate synthesis of RNA on a DNA template by adding ribonucleotides. Primase adds a short RNA primer on the DNA template forming DNA-RNA hybrid to facilitate the action of DNA polymerase (figure 7.9).

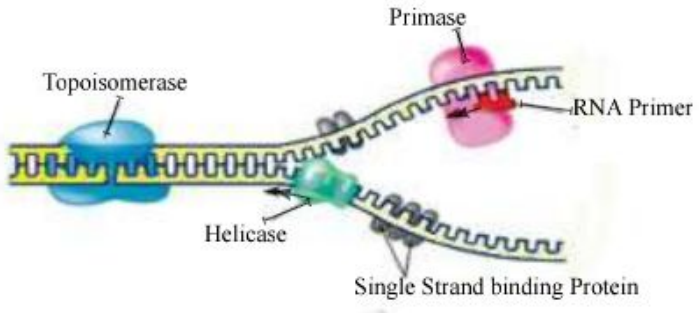


Figure 7.9 : ds DNA replication at the origin of replication forming replication bubble

DNA polymerases: There are several types of DNA polymerases. One type of DNA polymerase initiates DNA polymerization by adding a deoxyribonucleotide to the 3' end of a primer and continues polymerization on 5' to 3' direction elongating the new DNA strand with deoxyribonucleotides with bases complementary to template (figure 7.10).

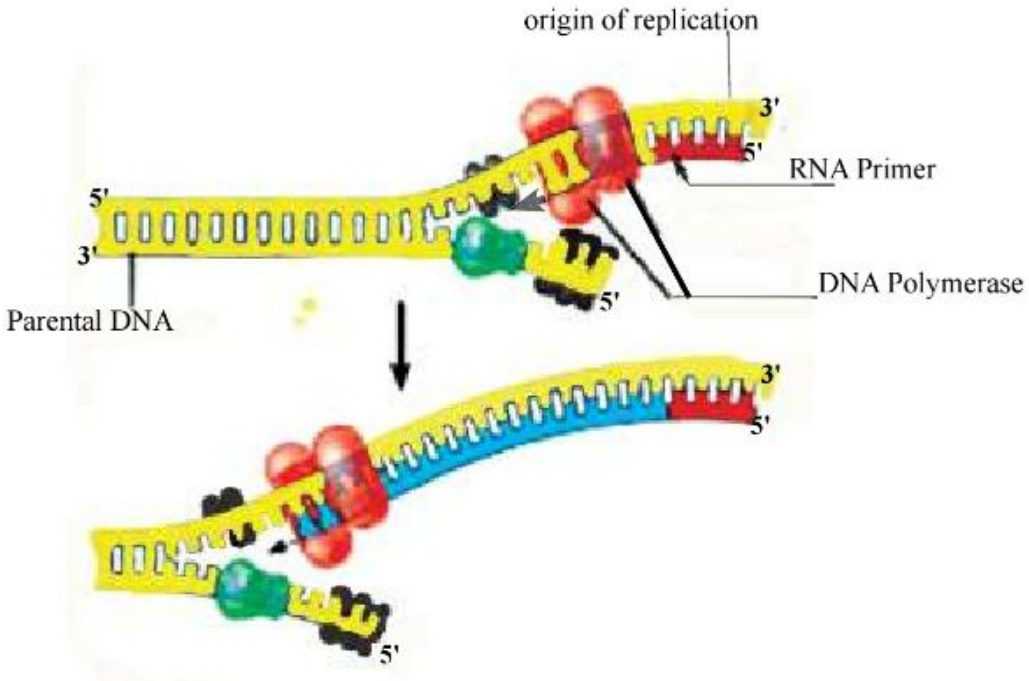


Figure 7.10: DNA Polymerase extends the new strand of DNA starting from 3' end of RNA primer

This DNA polymerase is almost 100% error-free in adding correct complementary nucleotides to growing strand, according to the nucleotide sequence of the parent DNA strands. However one error in 10^5 added nucleotides is possible, has a proofreading mechanism and therefore can correct its own mistakes and the error rate will be reduced by 100,000 times to one in 10^{10}. Therefore, the daughter DNA molecules are almost identical to the parent DNA molecule, and of course to each other.

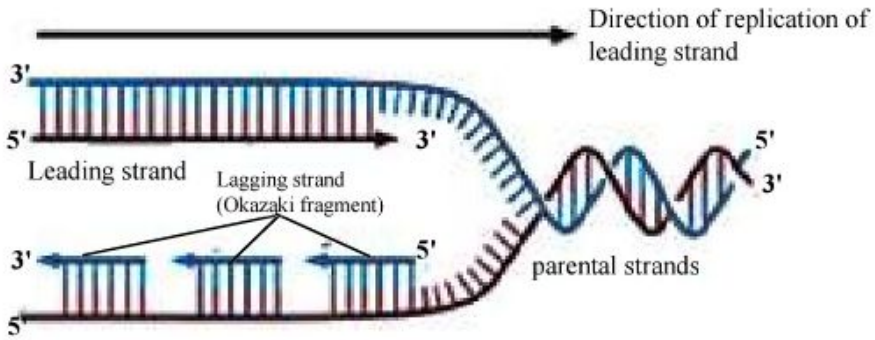


Figure 7.11: How DNA Polymerase solves the problem of antiparallel nature of DNA molecule

When a wrong nucleotide is added to the growing chain of DNA by DNA Polymerase this mismatch is identified by the same enzyme, stops adding the next nucleotide, removes the incorrect nucleotide exonuclease activity and continues again its polymerase activity. This is called as proofreading activity of DNA polymerase.

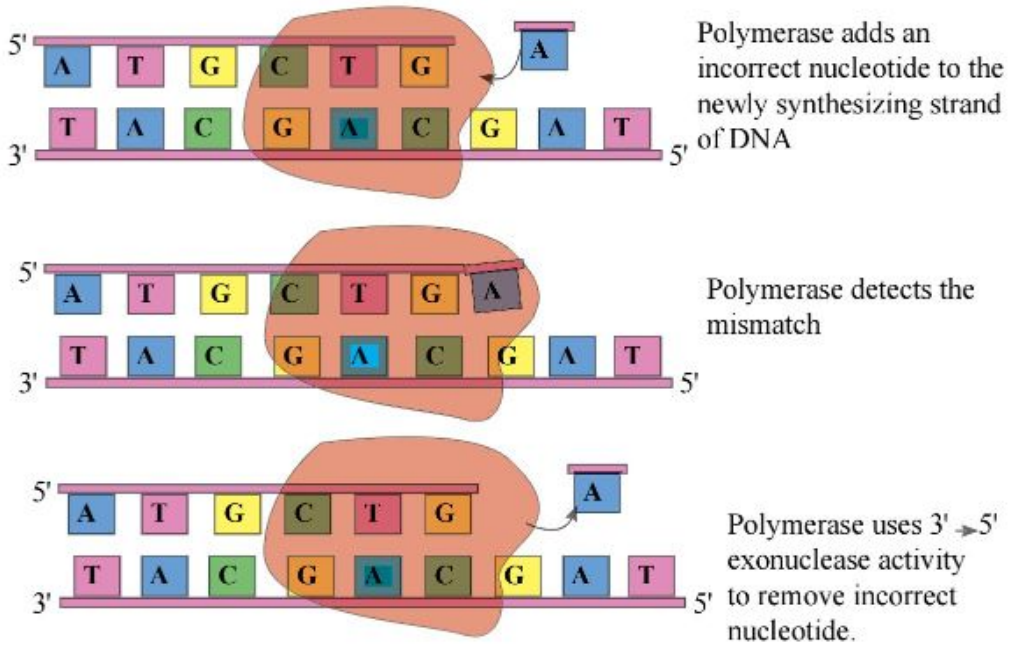


Figure 7.12. Proofreading activity of DNA Polymerase

Another type of DNA polymerase identifies DNA-RNA hybrid, removes ribonucleotides and replaces those with deoxyribonucleotides and so replaces the RNA primer with DNA. Now the DNA segment is incomplete, but the DNA polymerase unable to correct the ends of okazaki fragments, results gaps.

DNA ligase: In DNA synthesis, it joins the newly synthesized adjacent fragments making a complete strand by the formation of a phosphodiester bond. It seals the gaps of the newly synthesized DNA strand.

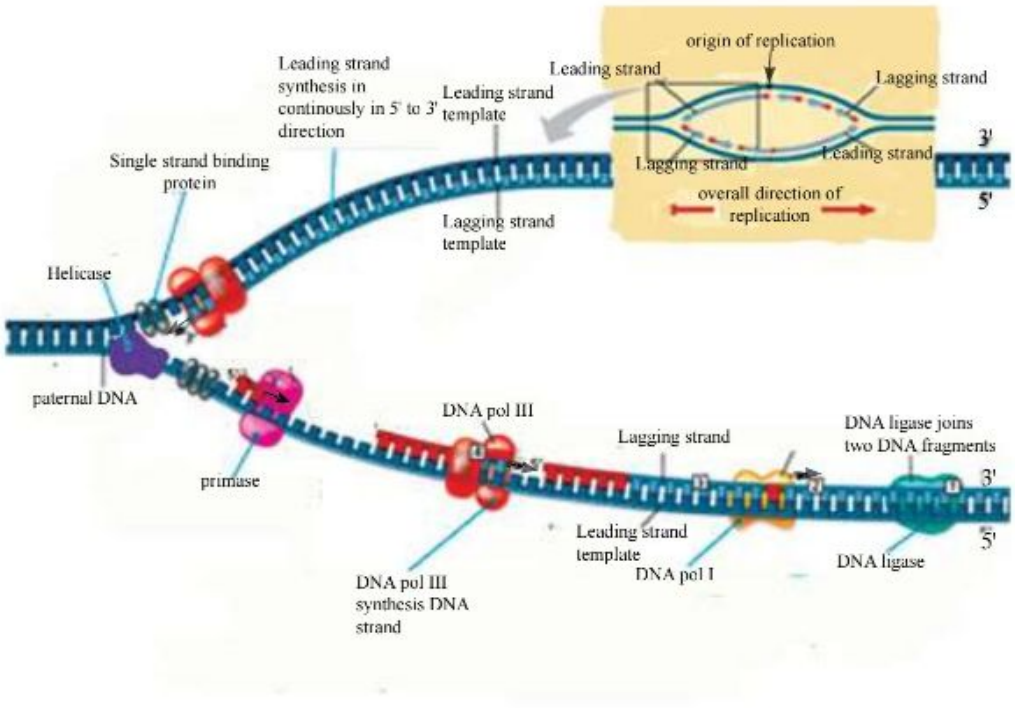


Figure 7.13: Overall process of DNA replication

Overall Process of DNA Replication

Relaxation of tightly wound DNA

Unwinding/ unzipping of the double helix

Stabilization of single stranded DNA

Priming of DNA synthesis by a RNA primer

Extension of new DNA strand - Leading strand: Continuous

- Lagging strand – Discontinuous

Removal of RNA primer and replacing RNA (ribonucleotides) with DNA (deoxyribonucleotides)

Sealing gap between the adjacent nucleotides

Similarities and Differences between Prokaryotic and Eukaryotic DNA Replication:

Eukaryotes share many characteristics of DNA replication of the Prokaryotes. They also use helicase to unwind the dsDNA, and use DNA polymerases for the polymerization reaction. The replication is originated in both types at specific sequences (origin of replication – Ori) to initiate DNA replication, packed DNA is relaxed with topoisomerase and replication process move in the same manner and therefore having leading and lagging strand. RNA primers are formed and replaced. The gaps are sealed by ligase.

Although the process is superficially similar, there are several significant differences too. The size of the DNA molecule in a eukaryotic chromosome is much larger than that of a circular DNA molecule of a bacterium. Therefore prokaryotes usually have one Ori while an Eukaryotic chromosome has several Ori. The DNA polymerases in eukaryotes and prokaryotes are different from each other in their structure, while having the same functions. The DNA replication of the prokaryotes occurs continuously, whereas in eukaryotes it happens only in the S-phase of cell cycle.

DNA repair and its significance

Damage to DNA can be caused by various chemical and physical agents introducing mismatches in DNA double helix and may lead to a permanent change in the sequence of DNA. This may also be caused by errors in DNA replication that goes unnoticed by proofreading. This is called mutation. A mutation or collection of mutations can make a cell malignant and result in a cancer. Mutations also alter the phenotype, and very often they are either lethal or at least produce unfavorable phenotypes. When the mutations occur in cells producing gametes, they may be passed onto the next generations resulting in variations among offspring.

When such a mismatch is there, the shape of the double helix is distorted. For example, UV radiation can covalently link two adjacent thymine bases, distorting the DNA molecule. One of the two copies of DNA molecules generated by these distorted sequences carries a permanent change in the sequence, a mutation. Such distorted locations are usually recognized and repaired by the DNA repair mechanism of the cell before it becomes permanent, and that reduces the risk of accumulation of mutations. Because DNA repair is very important for the survival of organisms, a large number of DNA repairing enzymes are found in various organisms.

Such enzymes may cut off the mismatched sequences in the damaged strand and replace with correct nucleotides. The cutting (excision) is done by a nuclease and filling the gap, using the correct strand as the template, is done by a type of DNA polymerase. This is called nucleotide excision repair. DNA ligase completes the DNA strand by sealing it with a phosphodiester bond.

Genes and how they work

The nature of prokaryotic and eukaryotic genes

When Gregor Mendel put forward his laws of inheritance in 1860, he used the term hereditary factors to explain what control the characters shown in a phenotype and which are transmitted from generation to generation. At that time, these were imaginary units and where they were located in cellular structure was not known. Today, these physical and functional units of

heredity have been identified as genes that are located on chromosomes as discrete units. This disclosure began with the development of cytology and the ability to observe the behavior of chromosomes in mitosis and meiosis. Behavior of chromosomes and the behavior of the Mendel's hereditary factors show the same pattern. In eukaryotes, chromosomes occur in pairs in diploid somatic cells, and hence genes also occur in pairs. A pair of chromosomes, coming from the two parents, containing the same genes is called homologous chromosomes. Generally prokaryotes have one chromosome in each cell, and hence may be considered haploid.

Gene
Gene is the fundamental physical and functional unit of heredity.
A gene is comprised of a segment of DNA on a specific site on a chromosome. It specifies the sequence of RNA.

The location or the site of a gene on the chromosome is called a locus (loci in plural). The alteration versions of genes that are located at the same locus on different chromosomes are called alleles of the genes. Prokaryotic genes are housed in loci on the circular DNA molecule, as discrete segments of DNA.

There are many steps in a biochemical pathway and each step is controlled by a gene, and therefore many genes are involved in controlling a particular phenotype. In eukaryotes these genes are scattered among several chromosomes, while in prokaryotes, they are arranged in the same region of the chromosome one after the other as clusters. These clusters are expressed together by a single control region and are transcribed into one mRNA. Such mRNA molecules will be translated into several different peptides. These organized gene clusters in prokaryotes are called operons.

Operon: A group of genes that functions as a single transcription unit. It is comprised of a control region (an operator, a promoter) and structural genes that are transcribed into one mRNA. Coding for several peptides.

In prokaryotes, all DNA segments in the chromosome are functional (either transcribed into mRNA or act as control regions), while in eukaryotes, a large proportion of DNA has no identified function. Such DNA segments that are present in between the genes are referred to as intergenic DNA. There are also sequences within the genes that are transcribed, but are not translated into polypeptides. That means there are coding sequences as well as noncoding sequences in the transcript of the gene. The noncoding sequences within the genes are called introns while the sequences that code for polypeptides are exons. Accordingly the transcript contains exons as well as introns. The transcript is a pre-mRNA and it undergoes processing where introns are excised and exons are joined inorder to produce mRNA.

Chromosome Theory of Inheritance

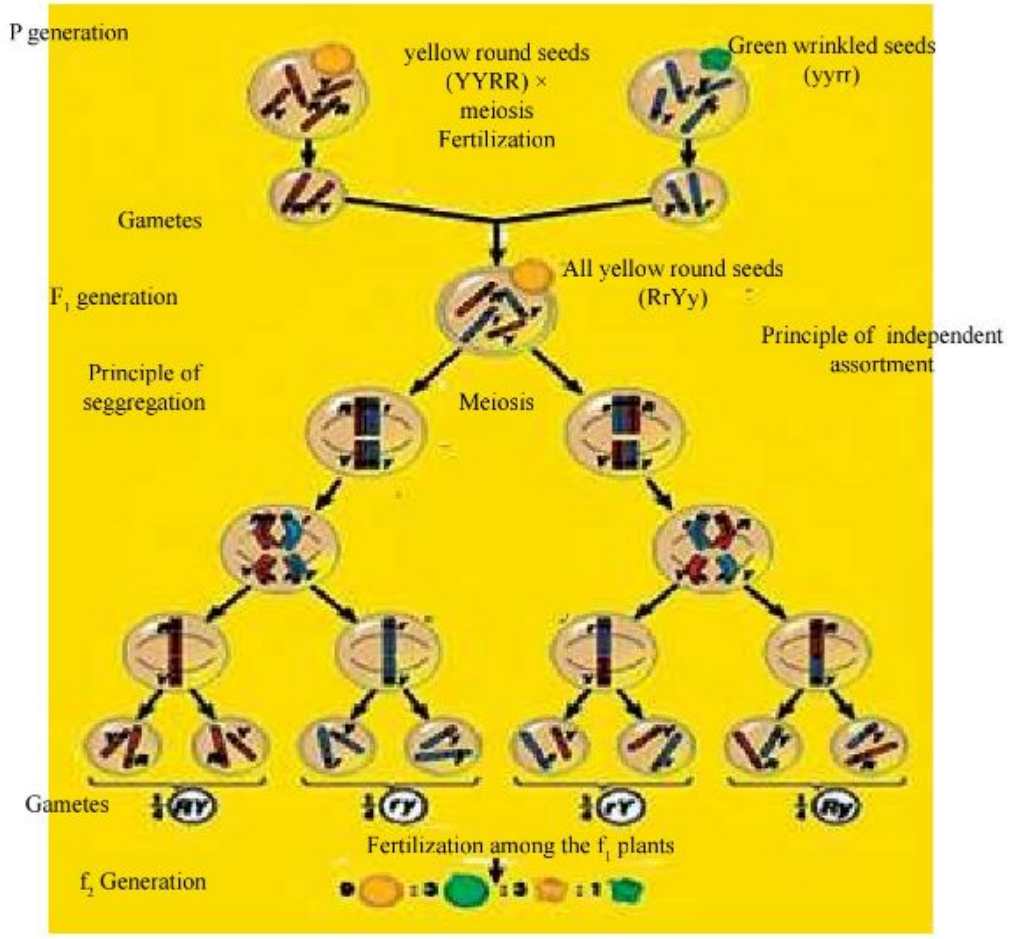


Figure 7.14: Chromosomal basis of Mendel's Laws: How the alleles of genes on homologous chromosomes behave during meiosis

Figure 7.14 shows the parallelism between the behavior of Mendel's hereditary factors and genes and their alleles located on chromosomes. Evidences collected in genetic studies were noted by scientists and several of them independently developed chromosome theory of inheritance. The Mendelian hereditary factors or genes are located on the chromosomes at specific loci. Therefore chromosomes and genes housed on them occur in pairs in diploid cells. The paring of homologous chromosomes in metaphase I occurs at random and therefore this is independent assortment (i.e. no order in arranging maternal and paternal chromosomes). In anaphase I the independently assorted homologous chromosomes separate reducing the number of chromosomes to half. This is segregation. With the independent assortment and segregation of chromosomes, alleles of genes on nonhomologous chromosomes are assorted independently in different combinations in Metaphase I. The alleles segregate after completion of anaphase I to form four haploid cells having different allele combinations in equal proportions. The same reasoning can explain the phenotypic ratios observed by Mendel in F_2 generation after random crossing of F_1

Gene Expression

Genes control characters when a gene is at work and, it is said that the gene is expressed. Gene expression is the process by which the information stored in a gene is used to make a functional gene product. The final product of a gene is usually a polypeptide, which after suitable modifications becomes a protein. However several RNAs also function as final gene products. For example, ribosomal RNA (rRNA) and transfer RNA (tRNA) are such RNA molecules, having direct functions.

In examining how genes control characters, it was revealed that (the first suggestion came in 1902 from Archibald Garrod) inherited diseases are caused by the inability to produce related enzymes as a result of inborn errors in metabolism. The symptoms of a hereditary disease condition, known as alkaptonuria is due to inability to make the metabolic enzyme which metabolizes the chemical alkapton. In patients alkapton remains in urea and its oxidation results in black colouration of urine.

The term copying is used in transcription, as in linguistics, since the information written in the nucleotide sequence of DNA by four letters (AGCT) is copied to the same chemical language as a nucleotide sequence of RNA written in four letters (AGCU), with only difference that T is replaced by U. The amino acid sequence in the polypeptide chain is also linear and parallel to the base sequence in the gene or mRNA, but the chemical languages are different: language written in four letters to a language written in 20 letters. This is the reason why the second step is called translation.

The gene expression begins by copying the information stored in a segment of DNA or a gene into sequence of RNA. In synthesis of polypeptide the gene is not directly converted into polypeptide, but involves an RNA molecule to pass the message in DNA to a message in polypeptide. Since the RNA molecule acts as a messenger to communicate the information from DNA to polypeptide, it is called messenger RNA (mRNA). The two steps in polypeptide synthesis are:

1. Transcription - Copying a sequence of DNA into mRNA
2. Translation – Converting the information in mRNA to a sequence of amino acids

Transcription is similar to replication since a DNA strand is acting as a template to make a complementary mRNA strand. The difference in transcription is that the copy is a mRNA molecule and only one DNA strand is copied. The major enzyme catalyzing polymerization is named as RNA polymerase. This is mRNA, because it transmits the message stored in the gene to the place where the polypeptide chain is actually being assembled.

Message in mRNA is translated to a sequence of amino acids. This process occurs in association with ribosomes, which are present in the cytosol.

In addition to mRNA, other types of RNAs and enzymes are involved in the polypeptide synthesis. The basic mechanism of the process of polypeptide synthesis is similar in prokaryotes and eukaryotes, with some important differences.

Genetic Code

In transcription, each letter of the template will be copied to the corresponding letter in mRNA. Note that the mRNA is complementary to template, and hence it is a copy of the other DNA strand. This seems to be straight forward, one to one copying. On the other hand, language of nucleic acid has four letters (nucleotides), whereas the language of protein has 20 letters (amino acids). If each kind of nucleotide is translated into amino acid, only four amino acids would be specified or coded. Therefore a combination of nucleotides would be needed to code for one amino acid. It has been proven by experiment that amino acids are coded by triplets of nucleotide bases and the protein synthesis is based on a triplet code. Therefore, the genetic code is a triplet code. If three-letter combinations or triplets are considered, $4^3 = 64$ possibilities are there. The three letter words or triplets are read one after the other, and hence are non-overlapping. Since all the words have three letters, no space is required to delimit the words. The genetic code stored in a gene in non-overlapping three-letter word is copied to non-overlapping, three-letter word in complementary mRNA. This is translated by reading three letters at a time identifying an amino acid corresponding to each three-letter word (Figure 7.16). The triplet of nucleotide bases of mRNA or coding for an amino acid formation is called a codon. Therefore the genetic code has 64 codons.

Out of 64, 61 triplets code for 20 amino acids, and the other three are used to signal the "stop" of translation or as "termination" codons (UAA, UAG and UGA). AUG codon codes for methionine (Met) as "start codon" signaling the protein synthesizing machinery to start translation of mRNA at that codon. Therefore, the first amino acid of all protein is methionine, but this may be enzymatically removed after the translation. Figure 7.16 shows all 64 codons and what each of them codes for. It is easily noticeable that certain amino acids are coded by more than one codon.

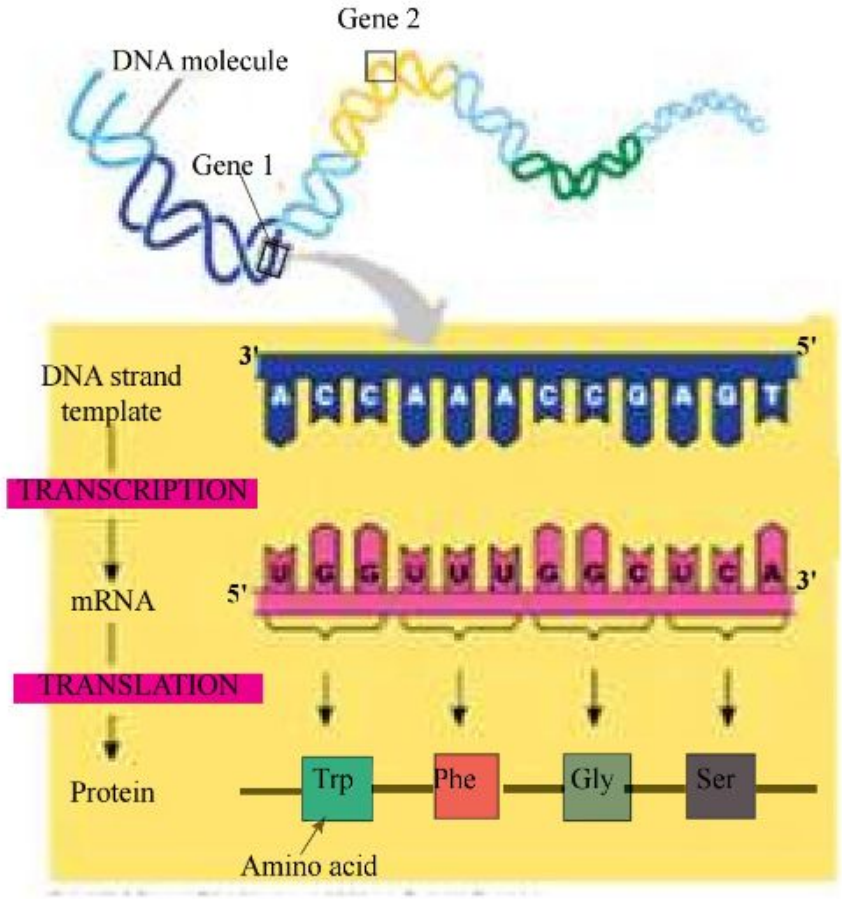


Figure 7.15: The genetic code is copied to mRNA and translated into amino acid in polypeptide chain using triplet codons

To read a message accurately, a starting point and a termination point as well as correct sequence of letters should be identified. This is called a reading frame. The protein synthesizing machinery starts and ends reading at precise positions and reads triplets one after the other in a non-overlapping fashion. Since all words are three letter ones, no spaces between words are required. If the reading starts at a wrong place, entirely wrong message will be read and a wrong polypeptide will be synthesized. If one letter is missing or one letter is added in the reading frame, the wrong message will be read from that point onwards, once again making a wrong polypeptide, if at all. The reading of the message is always from left to right by convention.

Another interesting feature of the genetic code is that its universality. This means nearly all organisms share the same genetic code. Accordingly, a gene isolated from one organism and inserted in to another related or nonrelated organism should express the same protein. This is how human insulin is produced by bacteria. The reading frame for insulin protein will be translated

in exactly the same way in both human and bacterial cells. A firefly gene can be expressed in tobacco plant, so that the plant emits light.

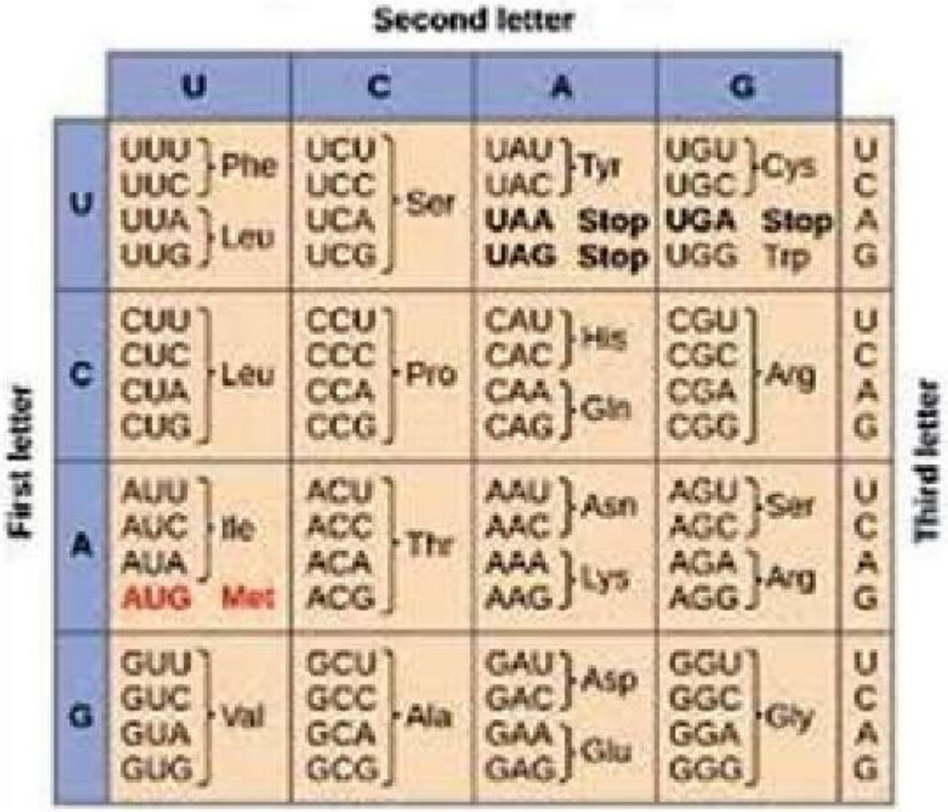


Figure 7.16: The codon table for mRNA

Mechanism of Polypeptide Synthesis

I. Transcription

Transcription is DNA directed RNA synthesis. This is completed in three steps.

1. Initiation: The process of transcription is initiated at a specific site called promoter. The promoter site includes the transcription initiation site and several other nucleotides. Only one strand of the double stranded DNA acts as a template for transcription. This is because only the template strand will have the promoter sequence in the correct orientation, which facillitates the binding of RNA polymerase. The enzyme polymerizing the RNA is RNA polymerase. This enzyme binds to the promoter site in correct orientation. The RNA polymerase then unwinds the two DNA strands and begins the transcription at start point. A component of RNA polymerase has the helicase activity, and hence DNA helicase is not involved in transcription.

2. Elongation: RNA polymerase enzymes can start adding complementary ribonucleotides against the template DNA. RNA polymerase continues to add nucleotides in 5' to 3' direction until it reaches the transcription termination site. The DNA strands unwinds as the RNA polymerase moves forward, exposing template DNA and allowing pairing with ribonucleotide. The two strands rewind at the other end (figure 7.17)

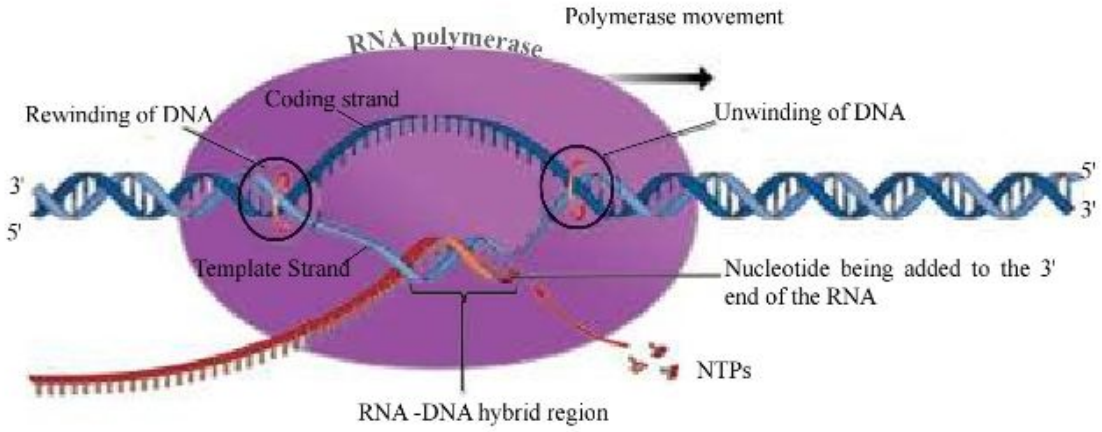


Figure 7.17: Elongation of the newly forming RNA transcript

3. Termination: In prokaryotes, the polymerization continues passing the termination sequence of DNA and RNA polymerase enzyme falls off ending the transcription..

After termination, the newly synthesized pre mRNA in eukaryotes is subjected to RNA processing. The mature RNA leaves the nucleus.

II. Translation

Once the mRNA is in the cytosol, translation process is initiated. Ribosomes read the message written as a sequence of triplet codons in an mRNA and translate it into a sequence of amino acids in a polypeptide with the assistance of transfer RNAs (tRNA). A tRNA attached to the correct amino acid from a pool in the cytosol, transports it to the ribosome which adds the amino acid to the growing end of the polypeptide chain by making a peptide bond. The key players in translation are tRNAs .

A specific tRNA molecule binds a specific amino acid to its one end. Its structure also carries, at a specific location, a triplet of nucleotides, which is complementary to the codon in the mRNA that code for the amino acid it carries. This triplet is known as anticodon and it can base-pair with the codon (figure 7.18). This is how tRNA does the translation, by acting as an adapter molecule between the triplet codon and the amino acid specified by it.

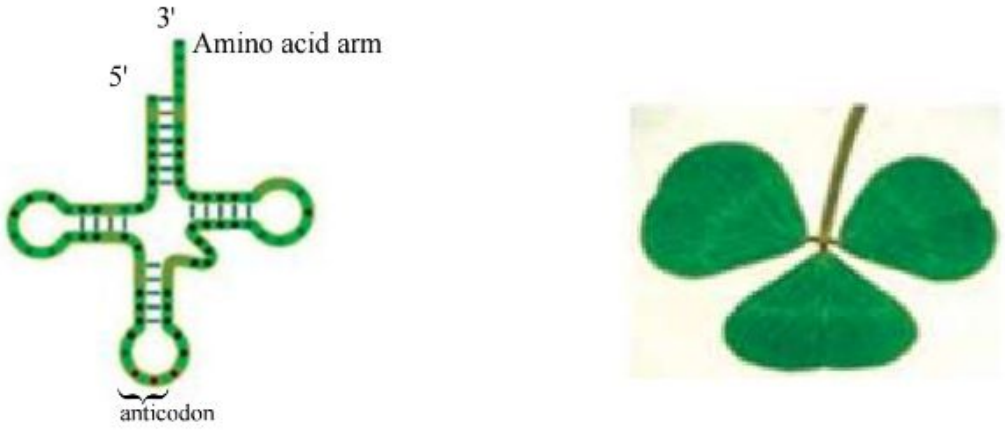


Figure 7.18: Two dimensional model of tRNA showing clover leaf structure

Structure of ribosome

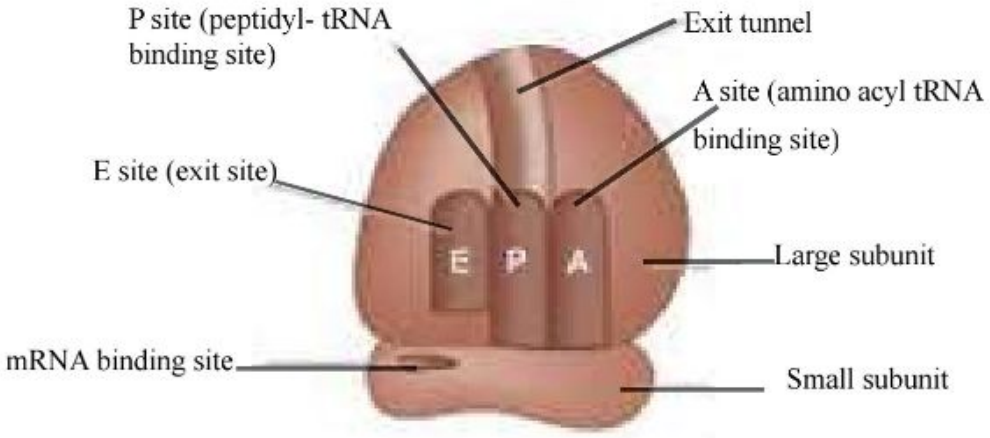


Figure 7.19: Structure of ribosome

Process of translation

The translation is also completed in three stages.

1. **Initiation:** The first step in initiation is binding of small subunit of ribosome to mRNA and to the initiator tRNA which carries methionine as the first amino acid. Then, the two subunits of the ribosome combine to form the functional ribosome. This complex of ribosomal subunits, mRNA and initiator tRNA, is called transcription initiation complex (Figure 7.20). Then the mRNA moves until the AUG start codon, aligns with the P site of the large subunit. Then, the anticodon of the initiator tRNA forms hydrogen bonds with AUG start codon. This signals the initiation of the translation.

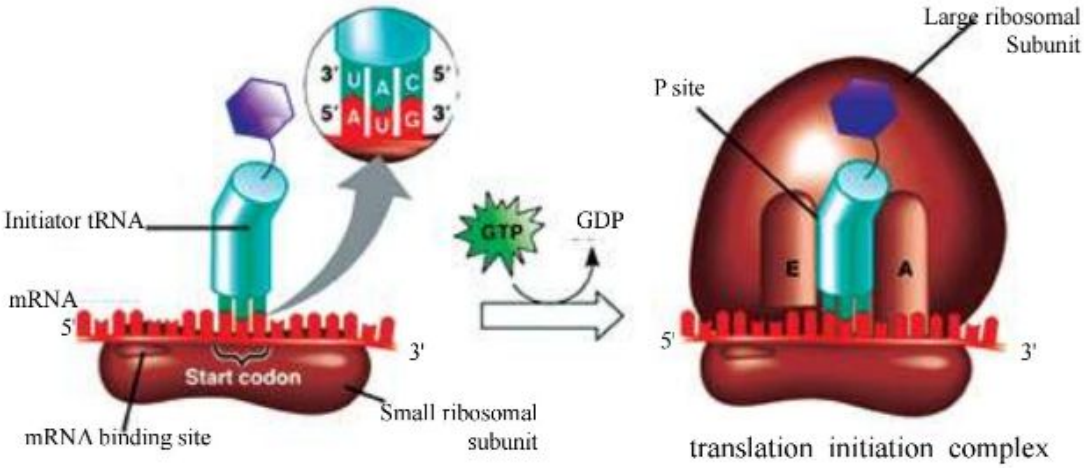


Figure 7.20. Formation of translation initiation complex

2. **Elongation:** In this stage, amino acids are added to the C- terminus of the growing polypeptide chain by peptide bonds, one after the other governed by the triplet codons. Elongation is completed by a three-step cycle. At the end of the initiation stage, P site is occupied with tRNA attached to methionine and A site is empty and is aligned with the next codon. The second tRNA, charged with the corresponding amino acid is brought to the A site, matching the codon and anticodon. This first step in the cycle is the codon recognition. As the second step, a peptide bond is formed between the carboxyl group of the growing polypeptide chain in the P site with the amino group of the amino acid in the A site. An rRNA catalyses this reaction. The third step is the translocation of the mRNA. The mRNA moves codon by codon unidirectionally. In the process, the tRNA with the growing polypeptide chain in the A site is moved to the P site. The released tRNA in the P site will move simultaneously to the E site, from where it is released to the cytosol. The A site is now aligned with the next codon, so that this cyclic process can continue. GTP is used for the energy requirement of the elongation process. See Figure 7.21 for an illustration of elongation stage in translation.

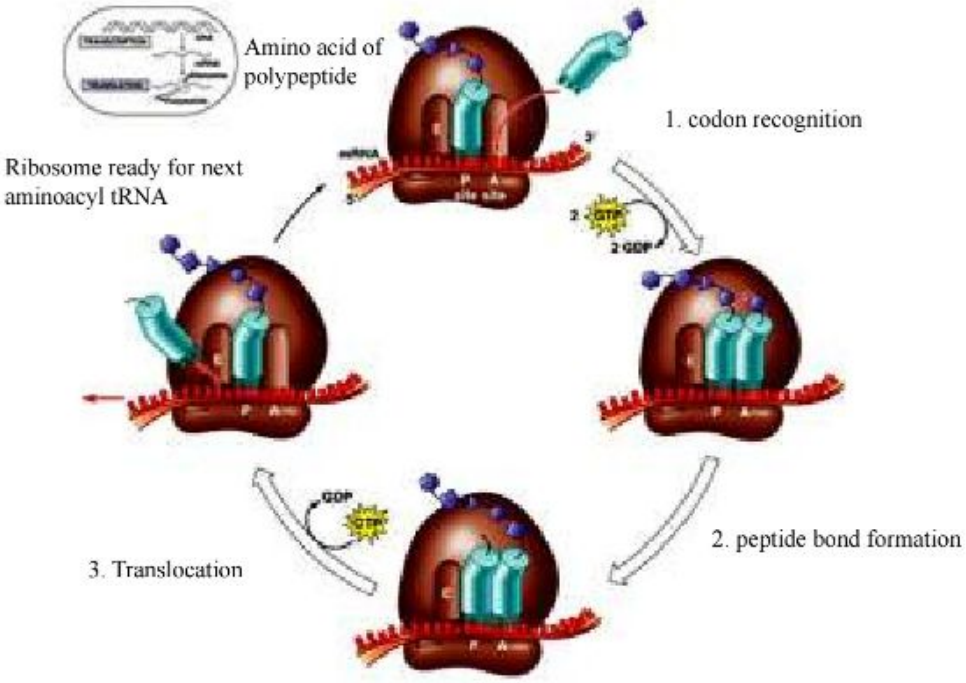


Figure7.21: Three-steps that occur in cyclic manner in elongation stage in translation.

3. **Termination:** When the mRNA moves, it finally aligns one of the stop codons: UAG, UAA or UGA with A site. They do not code for any amino acid, hence there is no tRNA which comes into A site. This releases the completed polypeptide chain to the cytosol. The ribosome and the remainder of the translation assemble then fall apart (Figure 7.22).

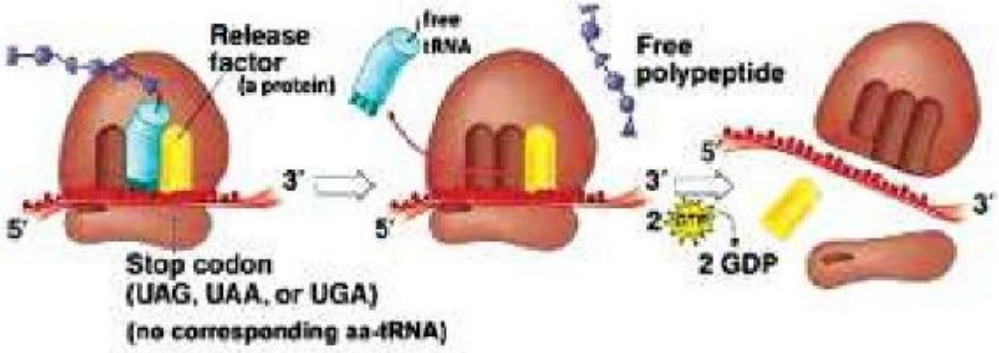


Figure 7.22: Termination of the protein synthesis

Polyribosomes/ Polysomes:

When the mRNA moves away for a sufficient distance, a second ribosome can bind to it. Depending on the length of mRNA, a number of ribosomes can be attached to the mRNA simultaneously. As such, mRNAs that are being translated actively are associated with multiple ribosomes attached to a string of mRNA, forming polyribosomes or polysomes. Forming polyribosomes increases the rate of translation since they allow simultaneous translation by several ribosomes.

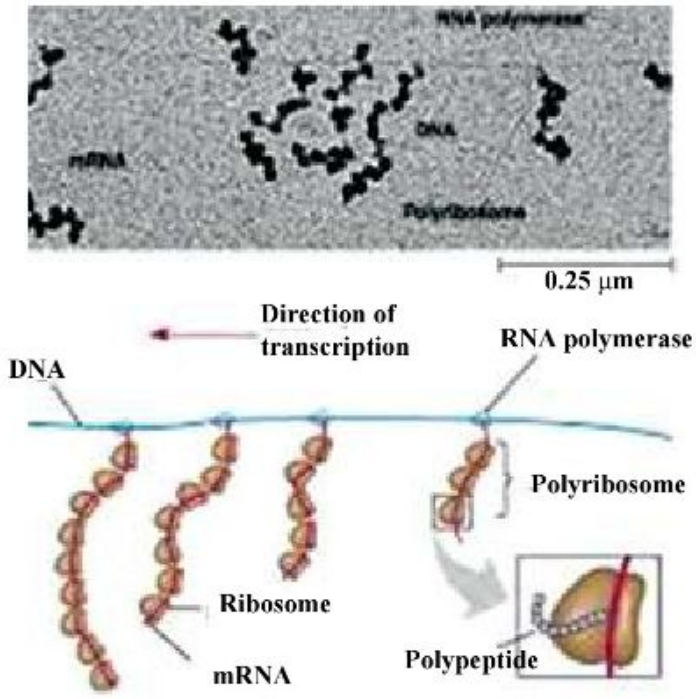


Figure 7.23: Polyribosomes of prokaryotic organism - mRNA is still growing and attached to the DNA

Fate of the proteins

The newly synthesized polypeptide is the primary structure of the polypeptide. This is not functional as a protein as it is. The polypeptide has to assume its functional form by folding (refer unit 02) and sometimes with post-translational modifications.

Certain polypeptides have additional segments than what is required for its function. For example, a short segment of amino acids are present in certain polypeptides to act as signal peptide. Signal peptide guides the polypeptide to a particular location in the cell or to be secreted. This is referred to as protein trafficking. Once the polypeptide is in place, the extra piece of the peptide chain is no longer needed and maybe removed enzymatically.

The post-translational modifications include chemical modification of certain amino acids by attaching sugars (glycoproteins), lipids (lipoproteins), phosphate groups (phosphorylated proteins) and other additions. The first amino acid, methionine, maybe removed enzymatically. Enzymes may also cut the initial polypeptide into two or more pieces which form a functional protein, by joining different combinations. For example, the protein insulin is produced as a single polypeptide, but cut at two places to remove the central piece. The other two polypeptide chains are joined together to form the functional insulin.

Selective Degradation of Proteins

The amount of a protein in a cell is determined on one hand by the rate of synthesis and on the other by the rate of degradation. The selective degradation of proteins is an essential mechanism in regulation of cellular activities. Certain proteins are degraded in response to specific signals. Faulty or damaged proteins are recognized and rapidly degraded to avoid bad effects due to mistakes in polypeptide synthesis or errors in folding. Some proteins for example regulatory proteins, need to be rapidly degraded after their function . The structural proteins may remain for a longer period.

Mutations

The phenotype of an organism basically depends on the genetic information stored in DNA and the final outcome is a result of the interaction between genetics of the organisms and the effects of the environment.

Alterations in DNA may bring about certain changes in the characters of individuals in a species, which result in phenotypic variation among organisms. These changes occur permanently, and are called mutations.

Mutation
An alteration of the nucleotide sequence of the genome of an organism

Mutations are the sources of variation seen among individuals of a species.

The effect of a mutation may be neutral, beneficial or deleterious. Deleterious mutations may be lethal or at least they are less favourable than the original phenotype. A mutation may also cause complete loss of a function. In rare occasions, the function of a polypeptide may be enhanced due to a mutation. These are beneficial mutations. Entirely new functions may also be resulted by mutations. For example, an enzyme specific to a substrate may change due to mutation in such a way that it changes its specificity to act on another substrate. The product due to the mutation is capable of catalyzing a new biochemical reaction.

There are two major types of mutations based on the scale of alterations in the genetic material; small-scale changes occurring in the sequence of nucleotides in a gene or large-scale changes in the number or the structure of chromosomes. These are known as gene mutations and chromosome aberrations or chromosome mutations, respectively.

Gene mutations

The permanent alterations of DNA sequence of a gene are called gene mutations. They may be occurred due to rare errors occur during DNA replication. These are called spontaneous mutations. In addition, certain external factors can also cause mutations at a higher rate. Since these factors generate mutations, they are called mutagens. Mutagenic agents can be classified as chemical or physical factors. X-rays and UV rays are examples for mutagenic physical agents.

The mutagenic agents can introduce mutations to replicating DNA in a cell. Mutations are also the cause of carcinogenesis. Therefore, mutagens are carcinogens and carcinogens are mutagens. These chemicals and radiations need to be handled with utmost care.

Types of Gene Mutations

These are small-scale mutations involving only one nucleotide pair or more than one nucleotide pair. If only one pair is altered, these are called point mutations. There are three types of gene mutations. They are:

1. A single nucleotide pair substitution – change one nucleotide pair with another
2. Nucleotide pair insertions – addition of one or more nucleotide pairs
3. Nucleotide pair deletions – removal of one or more nucleotide pairs

The substitution of a pair of nucleotide is a point mutation. Insertions or deletions can be point mutations, or may involve more than one nucleotide pair.

Substitution

Here one nucleotide pair is replaced by another pair (Figure 7.23). There is no change in the lenghth of the gene. Because the same amino acid may be coded by more than one codon. Some substitutions are silent mutations. Substitution of one nucleotide pair of a gene may not have an effect on the polypeptide it codes for. Because the same amino acid may be coded by more than one codons. The third letter in the codon triplet has a wobble, meaning that even if the third letter of a codon is replaced by another letter the same amino acid is coded in the third letter (for example, the codon 5'- GGC-3' on the mRNA will be modified to 5'-GGU-3', if 3'-CCG-5' triplet on the DNA template strand be altered as 3'-CCA-5' by substitution G with A). A substitution may also change one amino acid in the polypeptide. Therefore, the meaning of the primary structure of the polypeptide is changed slightly; hence, these mutations are missense mutations. This substitution of the amino acid with another amino acid may or may not have significant impact on the tertiary or quartnery structure (functional form) of the protein. Occasionaly, this may render the protein a higher activity or even with new properties. However, very often, these changes are either neutral or detrimental, making the protein useless or less efficient.

It is also, possible that a point mutation converts a codon coding for an amino acid to a stop codon. This causes premature termination of the protein synthesis, and therefore it is called non-sense mutation (figure 7.23). The result is a shorter polypeptide than original, which is usually non functional.

Insertion and deletion

These mutations bring about drastic changes in polypeptides compared to substitutions (note: nonsense mutations due to substitutions may also result in major changes). Insertion or deletion of a nucleotide or a couple of nucleotides will result in shift in the reading frame, reading wrong codons after the point of mutation. Therefore, such mutations are called frameshift mutations (figure 7.23), resulting in extensive missense. The polypeptide may not be functional at all, unless the insertion or deletion occurred is very close to the termination codon. It may also introduce a new stop codon, which was not in the original sequence. If that is the case, the translation terminates causing nonsense mutation. However, if the insertion or deletion is a triplet or a multiple triplet, the reading frame, will be back to the original reading frame immedietly after the point mutation (Figure 7.23). In such cases, one or a small number of amino acids will be added or removed, respectively from the entire sequence. The message will only be slightly changed and the polypeptide may be functional, depending on the importance of the mutated region of the polypeptide in its correct folding.

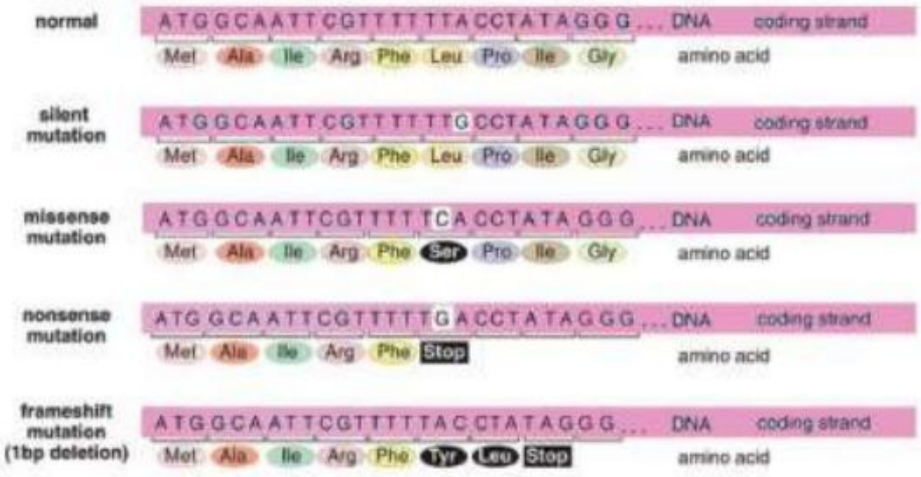


Figure 7.23: Types of Gene mutations

Chromosome Aberrations/ Chromosome Mutations

Since many genes are involved, most of the chromosomal mutations are lethal, while others are deleterious. Spontaneous abortions or miscarriages in mammals occur due to unusual numbers or structures of chromosomes. Such mutations also bring about various developmental disorders. Beneficial chromosome mutations are extremely rare. In plants, some chromosomal mutations cause beneficial variations.

I. Mutations due to alterations in chromosome structure:

Large chunks of chromosomes, which may include several genes to several hundred genes, may be lost, moved to another chromosome (cut and paste), copied and moved to another chromosome (copy and paste), or the orientation changed in chromosome mutations. These four types of chromosome mutations are called deletion, translocation, duplication and inversion (Figure 7.24).

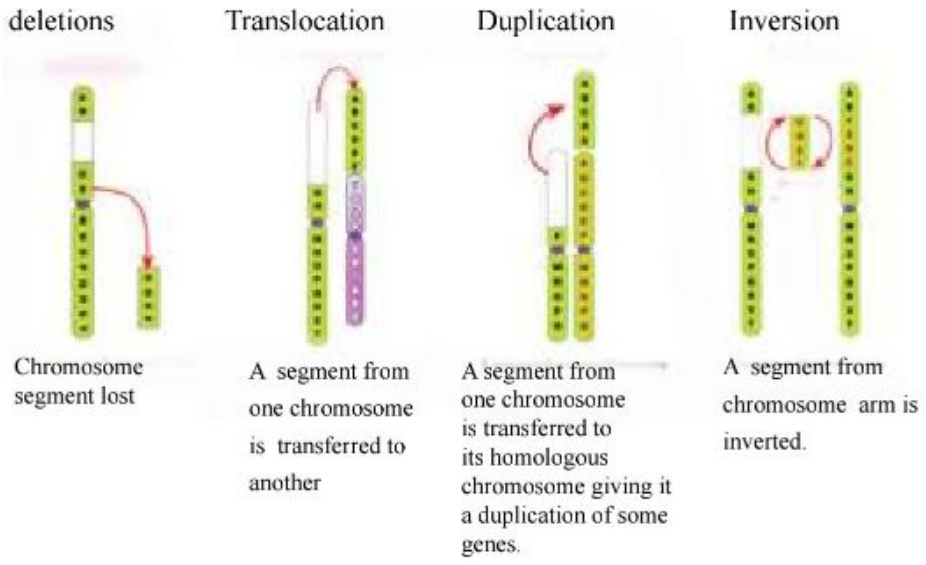


Figure 7.24: Four types of chromosomal mutations

The loss of a part of a chromosome removes several genes, and hence often these changes are lethal. In translocation, there is no loss of total amount of DNA. However, the expression of genes can be altered due to the change in the environment at the new location. The cutting of the chromosome may occur within a gene, and if this happens, that gene cannot function. In duplication, an extra piece of DNA carrying many genes is present in another location of the genome. This situation may also alter the gene expression, usually leading to deleterious impacts on the phenotype. The change of the orientation of a chunk of chromosome or inversions can also result in changes in gene expression; most of these are adverse variations.

II. Mutations due to alterations in chromosome number:

In addition to the structural changes of the chromosomes, a cell may contain an entire chromosome or even a set of chromosomes in excess to the normal number of chromosomes. A cell may also get one chromosome less than the normal number. When there is one more or one less chromosome in a cell, that situation is referred to as aneuploidy. Here the ploidy level has not changed, but when a complete set of chromosomes are present in excess, then ploidy level is said to be increased, for example to triploid, tetraploid, hexaploid etc.

Aneuploidy is resulted in by the mistakes in meiosis. During meiosis I, the two sets of chromosomes in a diploid cell must separate and move towards the two poles of the cell. However, due to abnormal arrangement of homologous chromosomes, both chromosomes of a pair may migrate to one pole, so the other end has one less chromosomes. The resulting cells or gametes in sexual reproduction also have either one more or one less chromosomes than the haploid number of chromosomes. A similar result is obtained when chromatids of a chromosome do not separate and migrate to the opposite poles of the cell in meiosis II. The inability to separate a pair or pairs of chromosomes in meiosis is called nondysjunction (Figure 7.25). When a gamete with one less chromosomes, unites with a normal gamete, the resulting zygote is an aneuploid with 2n-1 chromosomes. This type of cells is called monosomic, since a specific chromosome has only one of it. A gamete with one more than the usual haploid set of chromosomes may fuse with a normal gamete. The zygote will carry a chromosome in triplicate, and hence 2n+1. This aneuploidy status is called trisomic for that chromosome.

Such abnormalities can take place also in mitosis. Ploidy level can also increase by abnormal separation of chromosomes. A triploid (3n) can result in fertilization of an abnormal diploid egg. If a zygote does not divide after the first mitotic division, it will carry four sets of chromosomes and be developed into a tetraploid (4n).

Animals with higher ploidy level are very rare. Plants on the other hand, can tolerate higher ploidy levels, and often perform better than their diploid individuals.

Examples of plants having higher ploidy levels:
Banana – Triploid (3n); Wheat – hexaploid (6n); Strawberries – Octaploid (8n)

Polyploids are more common in invertebrates than in vertebrates. Among vertebrates, polyploidy is observed in a few fishes and amphibians.

Polyploids are rather normal compared to aneuploids. Although having higher number of chromosomes than usual, polyploids maintain the genetic balance, whereas genetic balance is lost in aneuploids.

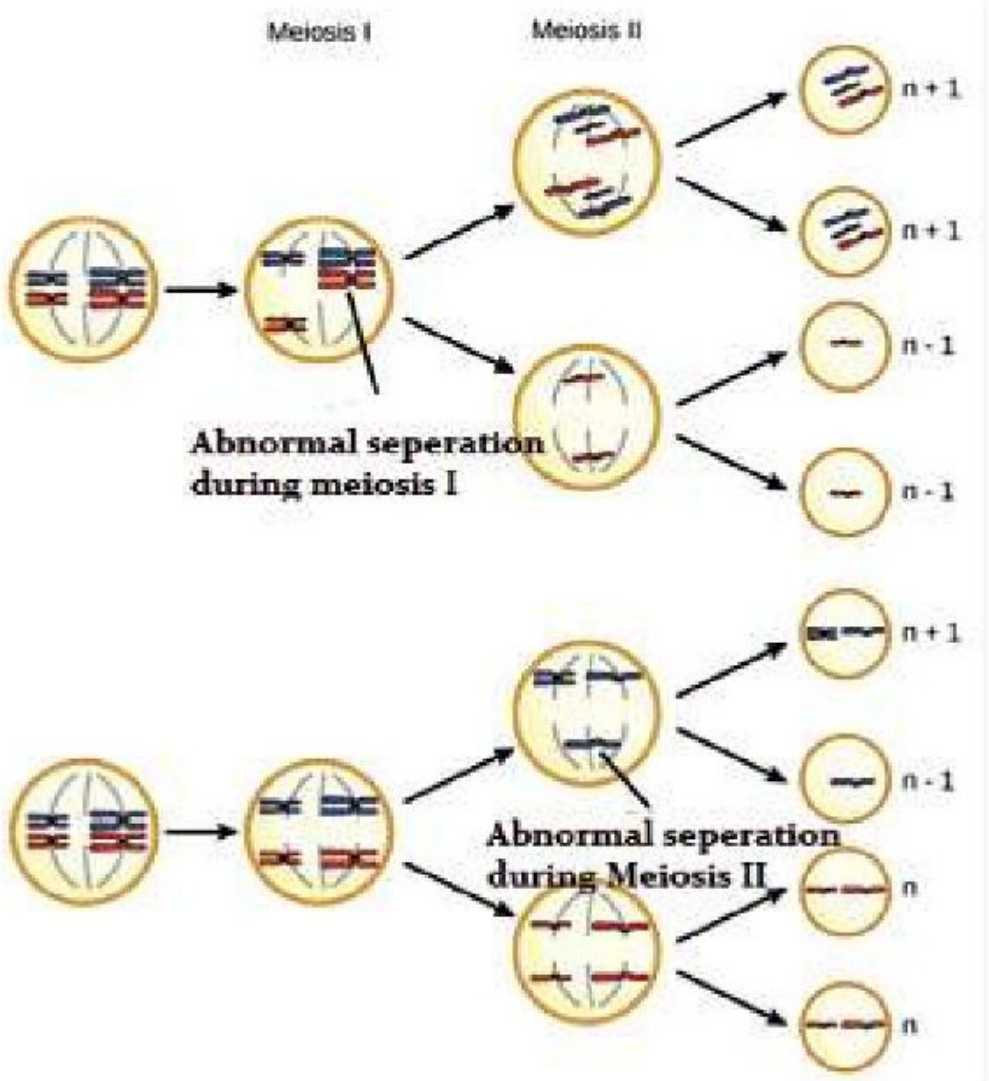


Figure 7.25: Abnormalities in meiotic division causing aneuploidy

Human Genetic Disorders

I. *due to Gene Mutations*

Two examples of human genetic disorders by gene mutations are described below.

Colour Blindness

Colour blindness or Colour Vision Deficiency is a genetic disorder, more common in males than in females, due to mutations on one or more genes located in X chromosome. These genes code for proteins that absorb different wavelengths of visible light. These are visual pigments called photopsins, which are categorized as red, green or blue. A person with a normal colour vision can distinguish different colours and their hues because they have all three categories of pigments in the retina. The absorption of different proportions of different wavelengths by the pigments is resolved by the brain as the colour of the object. In human genes coding red and green pigments are located in X chromosomes and gene for blue pigment is on the chromosome 7. Since males have only one X chromosome and the relevant genes are absent in Y chromosome, any defect in one or both of these genes will make the phenotype. In females at heterozygous condition a defective allele in one X chromosome is masked by the good allele in the other. Therefore colour vision deficiency is more common in males (5-8% of males) than in females (less than 1% of females). Colour blindness nearly always affects the perception of red or green colours, because these are the genes that are sex linked.

Sickle Cell Aneamia

Sickle cell aneamia is a genetic disease prevalent in human populations in Africa and other warmer regions of the world. A mutated allele of a gene coding for β -globin subunit of hemoglobin, the oxygen carrying pigment, causes an abnormality of the hemoglobin molecules. The presence of this abnormal hemoglobin in red blood cells changes the shape of the RBCs to curve as a sickle from its disk-shape. Individuals with this disorder have fewer numbers of RBCs, and hence develop aneamia. This is because the sickled RBCs breakdown prematurely. The mutation substitutes glutamic acid at a particular place in the primary structure of the β globin with valine (Figure 7.26), which results in abnormal folding of hemoglobin. The mutated allele is codominant, which means both normal β globin and mutated β globin are produced in individuals who are the heterozygous for this locus. Therefore, they have both good and bad hemoglobin and hence both normal and sickled RBCs are present. They are normally healthy and are the carriers of the mutant allele. Since the mutated alleles causes severely determintal effects in homozygous individuals, normally they should have been eliminated from the human population by natural selection. However, there is prevelence of malaria in warmer countries like in Africa, the heterozygous individuals will survive malaria attacks better than individuals with homozygous wild type alleles. This is because, the malaria parasite can not survive in sickeled RBCs. Therefore, at heterozygous individuals, parasite density remains at a lower level.

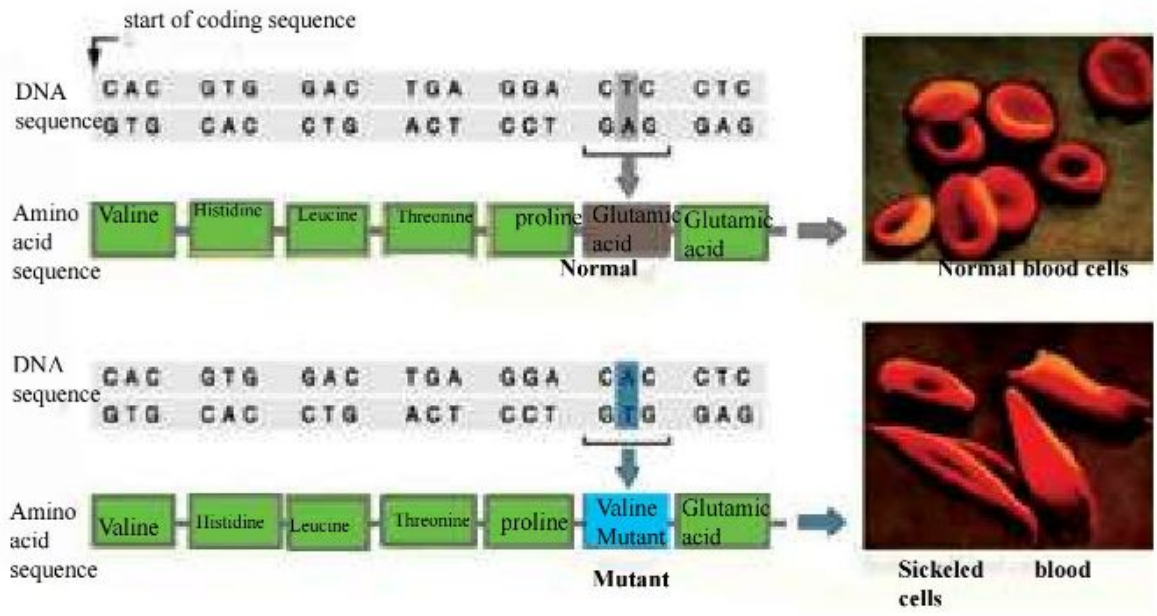


Figure7.26: The molecular basis of sickle cell aneamia

II. Due to Chromosomal Mutations

Chromosomal mutations bring about drastic changes in amount of genetic material or chromosome structure, leading to abortions of the fetus in mammals. If they survive, they show a peculiar set of abnormal characteristics in the phenotype which are called syndromes.

Three genetic disorders of humans due to aneuploidy are described below.

Down syndrome:

Down syndrome is also known as trisomy -21, because the cells of the affected individual contain an extra copy of chromosome 21. The syndrome shows characteristic facial features, short body, heart defects (which can be corrected) and developmental delays. They have a high risk of developing leukemia and Alzheimer disease. Almost all males and half of the females with Down syndrome are sexually under developed and sterile. Their life span is shorter than normal, but can live past middle age with proper medical treatment. However, they have a lower risk of high blood pressure, artherosclerosis (hardening of the arteries), strokes and many types of solid tumors. Despite their abnormalities, many live independently and are employed. The risk of having a baby with Down syndrome increases with the age of the mother. This is caused by nondisjunction in meiosis-I. While Down syndrome is due to trisomy of an autosome, there are human genetic disorders due to aneuploidy in sex chromosomes as well. The aneuploidy conditions in sex chromosomes cause Turner syndrome due to monosomy and Klinefelter syndrome due to trisomy.

Turner syndrome:

Turner syndrome is due to monosomy in X chromosome. In very rare cases, there are females with only one X chromosome, and hence their genotype is XO. This is the only known viable monosomy in humans. These individuals are phenotypically female, but they are sterile because their sex organs do not mature. When girls affected with Turner syndrome are treated with estrogen replacement therapy, they develop secondary sex characters.

They have a short stature and some may have an extra skin on the neck (webbed neck). Other characters are puffiness or swelling (lymphedema) of the hands and feet, skeletal abnormalities, heart defects, high blood pressure and kidney problems.Most of them have normal intelligence.

Klinefelter syndrome:

This is due to rare condition having an extra X chromosome in the genotype of XXY. Since it carries a Y chromosome, these individuals are males. Although these people have male sex organs, they are sterile. There testes are abnormally small. One out of the two X chromosomes is inactivated. Yet these males may have enlarged breasts and may develop other female body characteristics. They have subnormal (less than normal) intelligence.

XYY trisomy will make males and XXX trisomy will make female in females. They do not show any syndrome and have normal male and female characteristics, respectively. They are fertile and slightly taller than the average.

Genetic Counseling

Genetic counseling is a service which is important for families having genetic disorders or families having a risk of genetic disorders. This service is intended to estimate the risk a couple may have in conceiving a child with genetic disorders and providing necessary advice to avoid such instances. Genetic counseling is a profession that requires, on one side, a good knowledge of human genetics to understand, how traits obeying the rules of simple Mendelian inheritance would behave and on the other side, an ability to provide guidance to minimize the risk of having a child with a genetic disorder. If the family already has such a child, the genetic counselor would also provide advice how to manage the situation and how to plan the birth of next child.

Some of the genetic disorders are multifactorial, which means several factors including multigene inheritance and that is also affected by environment. For example, heart attacks and diabetes may be inherited, but the risk of developing the disease is affected by external environmental factors such as life style and food habits. As such, a clear pattern of inheritance of the disease cannot be traced. The risk of the effect of a trait that follows simple Mendelian laws of inheritance on a child to be conceived can be estimated by studying the history of the family with regard to a

particular disorder, and hence, becomes the subject of genetic counseling. If the disorder is caused by a dominant allele, it can easily be observed in the potential parents. However, if the allele is recessive, a parent or both the parents with normal phenotype may be homozygous for dominant allele or heterozygous carrier. Tracing the history of the family with regards to the disease using pedigree analysis, would allow estimation of probability of parents being carriers and accordingly would estimate the probability of the risk of having a child with the disorder. The information available through pedigree analysis, sometimes is sufficient to determine exactly the genotype of one or both parents. The genetic counselor explains the situation to the potential parents and guides them to choose the most suitable option in having a child.

Techniques are now available to determine whether the already conceived fetus carries the mutant alleles. For this, the cells of the early fetus are sampled and DNA is sequenced to find out the presence or absence of the mutant allele and, if present, whether the fetus is homozygous or heterozygous. This information is very important to take well informed decision to keep the fetus or to abort it. Legislations of some countries allow abortions of such fetuses which would otherwise be born with genetic disorders. This however is a difficult decision for parents to make. Therefore genetic counselor has a task of guiding the potential parents to make the best decision possible.

Gene Technology

Tools, Techniques and Methods in Gene Technology

This section examines the process of gene technology starting from DNA isolation, through identification of desired DNA sequences to gene technology or recombinant DNA technology. The isolated DNA needs to be cut different pieces of DNA must be joined and at times DNA must be copied *in vitro*. These involve several enzymes acting on DNA. Distinguishing a unique DNA sequence from the rest of the DNA, requires separation of DNA based on the size of the fragments and identification of it. In making genetically modified organisms, this DNA needs to be transferred to a recipient organism using a suitable method. Copying of DNA is possible *in vivo* using cloning and *in vitro* using Polymerase Chain Reaction (PCR). Sequencing of DNA has become a very important technique in many studies in DNA.

Isolation of DNA

Gene technology begins with isolation of the target DNA sequence from the total genome of the donor cells. Purified DNA is required for many applications such as studying DNA structure and chemistry, examining DNA-protein interactions, carrying out DNA hybridizations, sequencing, PCR, performing various genetic studies or gene cloning.

Since DNA molecules are very long, the entire length of a DNA molecule cannot be isolated, except for shorter DNA molecules such as plasmid DNA and viral DNA. However, breaking of DNA or shearing has to be minimized during the extraction process.

The basic principles and major steps of DNA isolation can be identified as follows.

- **Homogenization or disruption of cells:** DNA is located in the nucleus of the eukaryotic cells and it is concentrated in nucleoid in prokaryotic cells. The first step in DNA isolation is releasing the DNA from the cells by breaking cells or lysis of the cells. Cells can be lysed by mechanical methods such as grinding and homogenization or enzymatically by lysozyme to break bacterial cell wall.
- **Inhibition of DNAse:** When the cells are broken, the DNA may get in touch with DNA degrading enzymes such as Deoxyribonuclease (DNAse). Therefore, DNA must be protected from such enzymes causing shearing. Chelating agents are added to remove metal ions required for nuclease activity.
- **Dissociation of nucleoprotein complexes:** DNA needs to be freed from the proteins that bind to DNA. DNA-protein interactions are disrupted with SDS, phenol, or proteolytic enzymes.
- **Removal of contaminating materials:** All other molecules in a cell are contaminants of DNA. Removal of this contaminants is required for some applications.
- **Precipitation of DNA:** Here the DNA dissolved in aqueous phase is precipitated with cold (0 °C) ethanol. The precipitate is usually re-dissolved in a buffer. RNA is removed by limited treatment with DNAse free RNAase (ribonuclease).

Enzymes Reacting with DNA

Enzymes are required in cutting, joining and copying of DNA *in vitro*.

1. **Restriction endonucleases:** There are several different types of nucleases in cells having different functions. In gene technology, cutting DNA at precise locations is important. The enzymes that recognize specific sequences of DNA and cuts at or near these sites are called restriction endonucleases. The position where the DNA sequence is cut is called as restriction site or cleavage site (Figure 7.28).

   eg. *EcoRI*- Source: *E. coli*

2. **DNA Ligase:** Cut DNA fragments from different sources are joined together by forming a phosphodiester bond by DNA ligase to obtain recombinant DNA molecules (Figure 7.27). T4 DNA ligase is the most commonly used DNA joining enzyme in gene technology. The source of this enzyme is T4 bacteriophage.

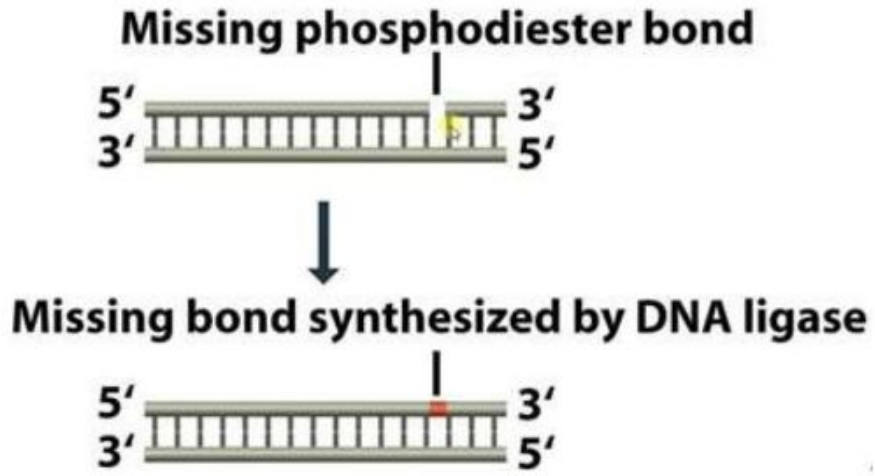


Figure7.27: Formation of a new phosphodiester to fill the gap between adjacent nucleotides

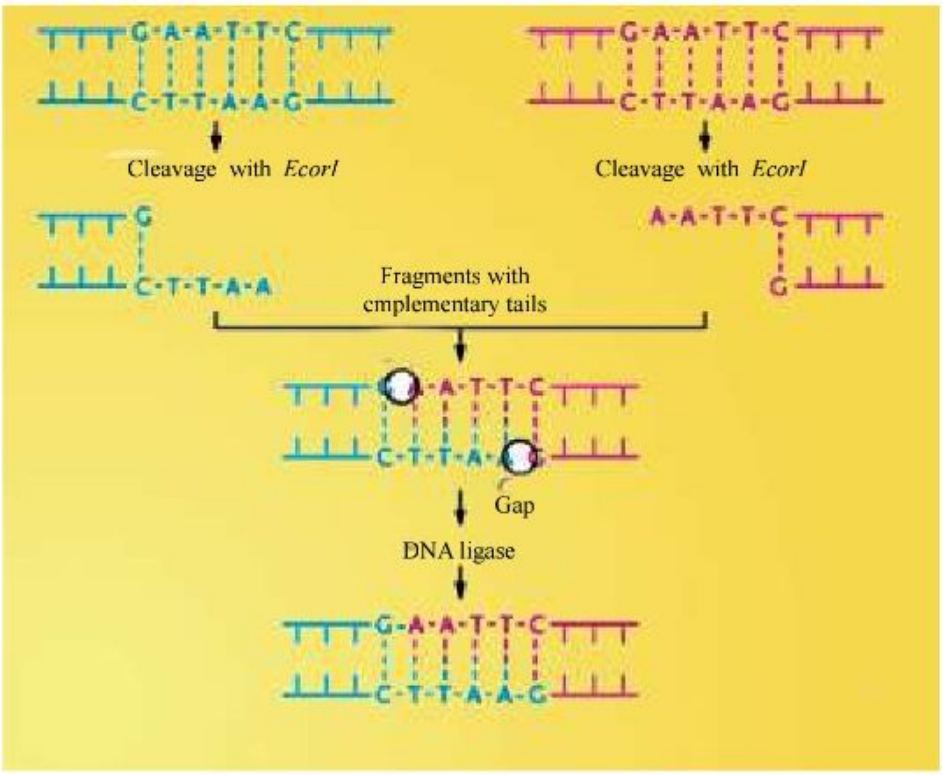


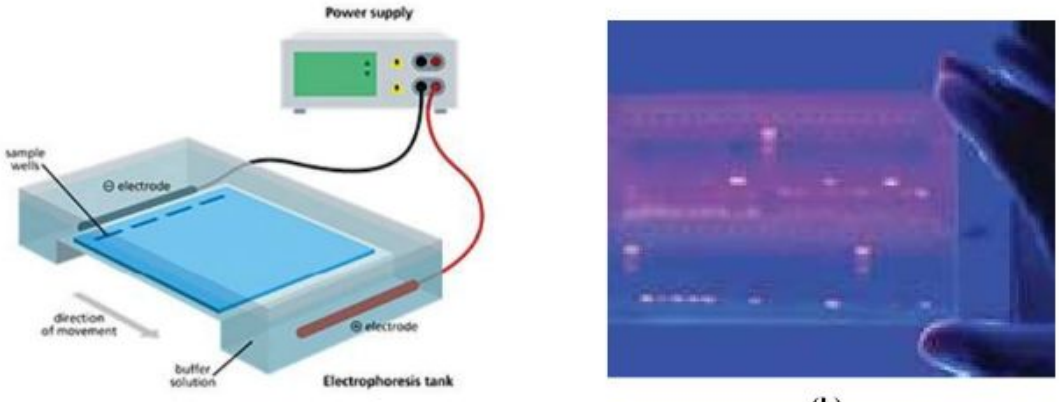
Figure 7.28. Cutting of DNA of different origins with *Eco R I* restriction enzyme and joining fragments of the different DNA by DNA ligase to form a recombinant DNA molecule.

3. **DNA Polymerases**: These are enzymes that add complementary deoxyribonucleotides to the template strand in a growing DNA strand, so that they copy DNA. These are therefore very important in gene technology especially in PCR and DNA sequencing. The most widely used DNA polymerase is Taq DNA polymerase. This is a heat stable enzyme isolated originally from the thermophilic bacterium *Thermus aquaticus.*

In addition to the enzymes reacting with DNA, enzymes which can make DNA on an RNA template is also very useful in gene technology. These enzymes are called reverse transcriptase, because their action is the reverse of the transcription. This is used to make cDNA (copy DNA or complementary DNA) on an mRNA template.

The DNA shearing or cutting with restriction enzymes makes a mixture of DNA fragments of different sizes. In DNA fingerprinting, using PCR, DNA strands having different sizes are obtained. Therefore in many applications with DNA, separation of DNA molecules becomes important. Separation of fragments having different size on a gel matrix is the most practical way of doing this.

Agarose gel electrophoresis: Electrophoresis is a technique that separates large charged molecules, (such as DNA, RNA and proteins) according to their mobility in an electric field. The speed of a molecule moving in an electric field depends on its net charge and the size. In gel electrophoresis, the molecules are moving through small pores of a matrix of gel. This restricts the movement of the molecules and helps in separation according to the size; larger molecules move slowly compared to smaller molecules. In the case of nucleic acids, the net charge depends on the length of the molecule, and therefore the separation is based on the size. For the separation of DNA, Agarose Gel Electrophoresis is the most frequently used technique. Agarose is purified agar, obtained from a type of seaweed. This forms a polysaccharide matrix. In the Agarose gel electrophoresis equipment, the gel is placed in buffer and a cathode and an anode are placed on either end of the gel (Figure 7.29(a)). When a current is applied using a power supply, the negatively charged DNA molecules migrate towards the anode through the gel. When preparing the gel, wells are formed, and the DNA is loaded into these wells. The separated DNA can be stained by ethidium bromide and be visualized by exposing to UV light (Figure 7.29(b)).



(a) (b)

Figure 7.29: (a) Agarose gel electrophoresis equipment (b) Visualizing separated bands of DNA on a gel using UV light

Although stains like ethidium bromide show the presence of double stranded DNA band on an agarose gel, those stains cannot differentiate a band with a specific nucleotide sequence from the others. In order to identify such a band among many other bands, a DNA probe is used.

DNA Probes and hybridization: A DNA probe is a fragment of single stranded labeled DNA used to detect the presence of complementary nucleic acid sequences by hybridization. Labeling is a modification of the DNA strand in such a way that it gives a signal which enables the detection of that DNA strand. The labeling can be done by, for example, incorporation of a radioactive isotope, or addition of a fluorescent molecule to the structure of the probe. This single stranded DNA stretch can hybridize to complementary single strand of DNA or RNA. Therefore before hybridizing with the probe, the double stranded DNA needs to be denatured to make room for the probe. The denatured bands on the gel need to be transferred to a nitrocellulose or nylon filter membrane by a process called Southern Blotting. The bands are then fixed to the membrane. Then the labeled probe is added to the membrane and allowed to renature. The probes will bind strongly only to the complementary sequences, which are fixed to the membrane. When the membrane is washed, the probe is removed except the probes bound to the bands with target nucleotide sequence. If the probe is radioactively labelled, the band with the target sequence can be identified by autoradiography of the membrane. If the probe is labeled with a fluorescent dye, the band can be identified using UV light.

Recombinant DNA Technology

All organisms on Earth have evolved from a common ancestor and their genetic information is stored in DNA, except in some viruses. At the chemical level, DNA is the same in all organisms. Further, all organisms share the same genetic code, so the same peptide is encoded by a gene whether it is expressed in a bacterium, a plant or an animal. This forms the basis of recombinant DNA technology, where DNA from two or more different species are joined together and inserted into a host to obtain a new genetic combination which is of value to science, medicine, agriculture, industry and environmental applications.

Recombinant DNA molecules:
DNA molecules made by laboratory methods of genetic recombination which brings together DNA from different sources, creating sequences that are not found naturally.

All the techniques mentioned below are needed to make a recombinant DNA molecule (rDNA).

They are;

- Isolation of DNA from different sources.
- Restriction digestion of isolated DNA with restriction enzymmes.
- Seperation of DNA fragments by gel electrophoresis.
- Identification of the correct fragments with desired nucleotide sequences using probes.
- joining the DNA fragments from multiple sources using DNA ligase.

Inserting DNA molecules to host cells is a difficult step. Cells show resistance to DNA uptake. This is important for the survival of organisms because invading DNA can cause genetic alterations which are usually harmful.

Therefore, a large number of copies of the recombinant DNA molecules are necessary to make sure at least a few host cells will receive a copy. If the desired DNA fragment is a shortone, *in vitro* multiplication can be done by a technique called DNA cloning.

DNA Clonning

In DNA cloning, the DNA replication machinery of a host cell is used to copy the desired DNA. However, inserting a fragment of DNA into a host cell will not be copied, unless an origin of replication (Ori) is present in that. Therefore, in order to replicate the recombinant DNA molecule or DNA of interest, it has to be combined with DNA with Ori and that can be replicated independent of the chromosomal DNA (chromosomal DNA replicates only once in a cell division). In a bacterial host, there are several copies of plasmids and also when infected with a bacteriophage, a large number of copies of viral DNA per bacterial cell are present. As such the DNA molecule of interest can be integrated into these self-replication units, which are called vectors.

Vectors: The vectors are vehicles to carry the DNA of interest into a host for multiplication or cloning.The vectors used in cloning of DNA are called cloning vectors.When the vector is carrying the foreign DNA, it is called a recombinant vector.

Making a recombinant vector also follows the same procedure for making the recombinant DNA molecule. Here the gene of interest has to be cut with a restriction enzyme and the vector (plasmid or the viral DNA) should also be cut with the same enzyme. These two should be mixed and allowed to integrate and ligated using DNA ligase (Figure 7.30). The site in the vector, where the DNA to be cloned is inserted is a called cloning site. In order to be able to use several restriction enzymes to cut DNA (both the vector and the DNA to be cloned), the cloning site contains a sequence which carry sites for many restriction enzymes. Therefore, this is called multiple cloning site. A host cell, usually a bacterial host which can copy the vector, is then transformed by the recombinant vector. The host then will copy the plasmid carrying the DNA of interest. A number of recombinant plasmid is present in each descendent cell of the colony of bacterial host.

Types of vectors and their differences

Any self-replicating unit in a particular host cell can be used as a vector. In bacteria, plasmids and bacteriophages are used as vectors. Plasmids are present in yeast cells also. Therefore they can also be used as vectors in yeats. The yeast cloning vectors are called yeast artificial chromosomes or YACs. they are plasmids, but are called chromosomes, because they carry sequences of centromeres. They work as chromosomes when they are lineraized. In addition, they also have autonomously replicating sequences (ARS) to enable them to replicate independent of the cell division. All of these vectors carry genes which are not necessary for a vector. These are removed and this space is used to insert the DNA of intrest. The yeast vectors should contain sequence of centromere and autonomously replicating sequences (ARS).

As described above, main purpose of cloning vector is to copy DNA in an *in vivo* system. For this, number of copies in a single host has to be high. So this condition is fullfilled by bacterial plasmids, bacteriophages and YACs. The transformation of cells is a very inefficient process. However, using bacteriophages as vectors overcome this problem, because the vector can be inserted to the host cell by the mechanism of infection of the bacteriophage. An advantage of YACs is that they are large, therefore, large amount of DNA can be copied using them. It has also another advantage, because they operate in the eukaryotic system.

Transformation:
Direct uptake of exogenous DNA from the surroundings of a host through its cell membrane and incorporation, resulting in a genetic alteration

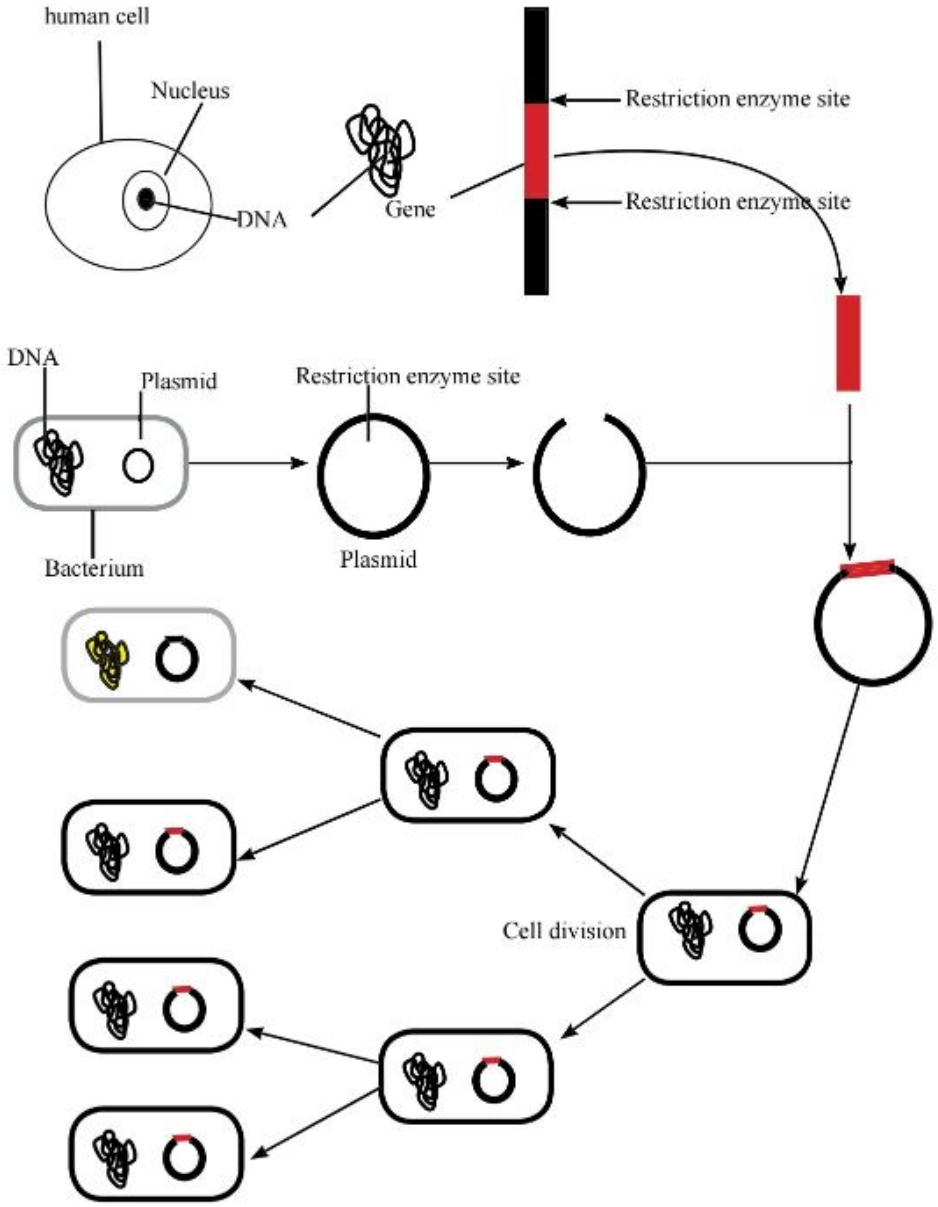


Figure 7.30: Cloning of a gene of interest using a plasmid vector and bacterial host

The copies of the DNA of interest or recombined DNA is retrieved by harvesting the hostcells, releasing the vectors by lysis, isolation of the vectors and cutting the DNA with the same restriction enzymes that was used originally to isolate the fragment of DNA. The recombinant DNA is then isolated after separation and detection on an agarose gel after electrophoresis.

Use of marker genes: The efficiency of the transformation of the host cells with the recombinant plasmid vector is very low. This means, that for one transformed host cell, there are millions or billions of cells that are not transformed. Both transformed and untransformed cells will make

colonies in a suitable medium, but these cannot be distinguished. Therefore, some sort of marker gene should be engineered into the cloning vector, so that the few colonies originating from the transformed cells can be screened among many untransformed cells. Most common markers are antibiotic resistance genes. The host cell is sensitive to the particular antibiotic, and hence will not grow on a medium containing the antibiotic. Since the vector is carrying the gene for resistance, the transformed cells can grow in the medium with antibiotic. Such markers are called selectable markers, since they allow the growth of only the transformed cells.

There is another problem to be solved. The transformation does not necessarily mean that the insert is there. Not all vectors are recombined with the gene of interest. Therefore, another marker is required to distinguish the colonies with the vector containing the insert from those with vector only.The essential features that should be present in a cloning vector are shown in Figure 7.31.

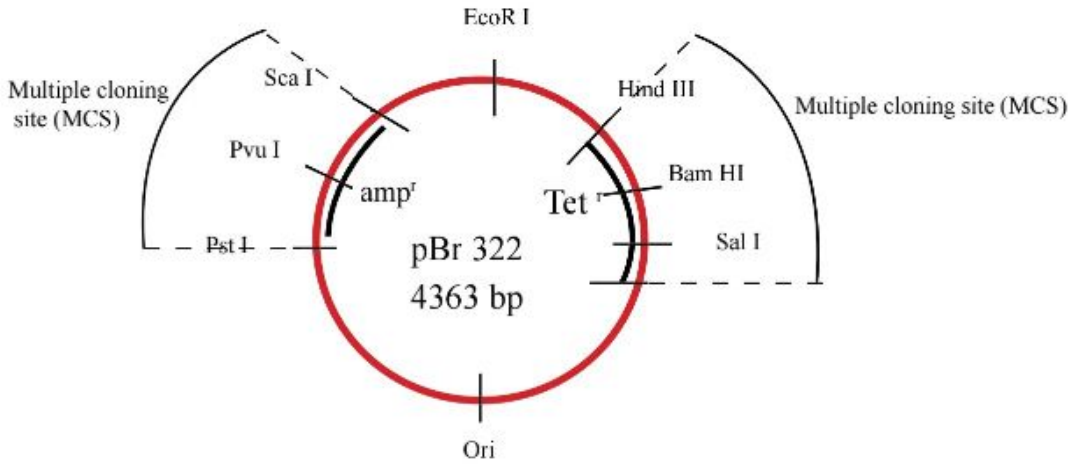


Figure 7.31: An example of a cloning vector (pBR322) showing essential features (ori, multiple cloning site, markers)

DNA Libraries

When a genome is cut into random pieces by shearing using mechanical force or by restriction enzymes, it creates an extremely large number of different sequences, depending on the size of the genome. All these pieces can be integrated into cloning vectors and the recombinant vectors can be used to transform bacterial hosts. These hosts can then be cultivated in a suitable medium for selection of transformed cells and screening for transformed cells with vectors carrying an insert. Since there was no selection for a particular segment of DNA in this case, each transformed cell with an insert may carry a different piece of DNA of the genome selected earlier. When all colonies are isolated and cultivated separately, the collection of these colonies is referred to as a genomic DNA library (Figure 7.32). DNA libraries are, a collection of microbial cultures each propagating a different fragments of a total genomic DNA which are cloned in a population of

identical vectors.Inserts of each colony can be sequenced separately in order to have a complete sequence of the genome. This is how the elucidation of sequence human genome under the Human Genome Project started.

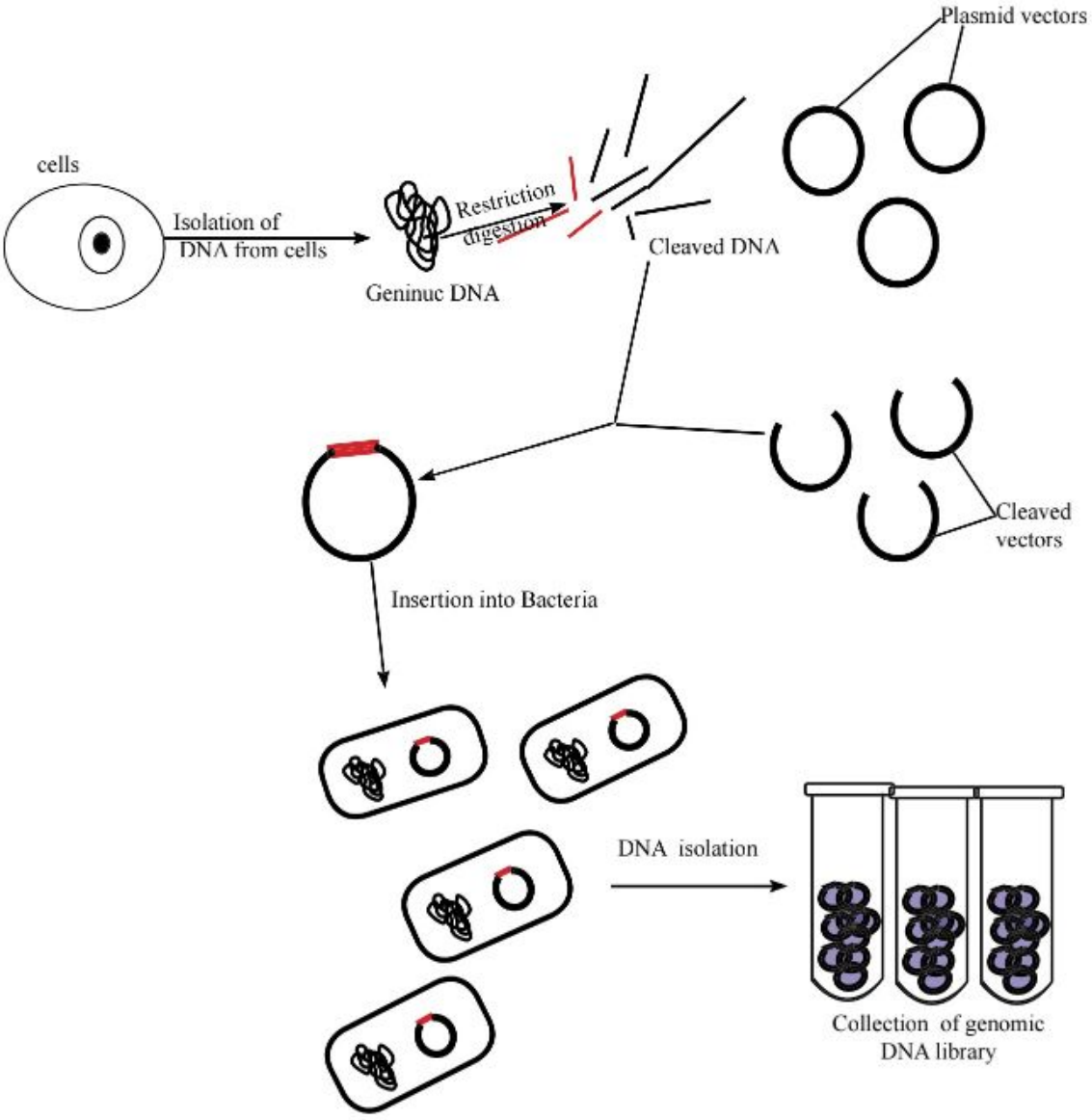


Figure 7.32: Steps in construction of a genomic DNA library

There is another type of DNA libraries. These are called cDNA libraries. These libraries contain complementary DNA obtained by reverse transcription of the mRNAs isolated from cells/ tissues. Collection of mRNAs of a cell is called transcriptome. mRNAs are isolated and reverse transcribed into a DNA strand complementary to the mRNAs. The enzyme used here is reverse transcriptase. The second DNA strand is replicated on the first DNA template using DNA polymerase to obtain double stranded cDNA. The DNA fragments are cloned and similar procedure is followed as for genomic DNA library in order to construct a cDNA library.

DNA libraries are basically used as sources of DNA fragments for sequencing. cDNA libraries also reflect the pattern of expression of genes.

DNA Delivery Systems

A cell contacting foreign DNA is called a transformed cell. Getting the foreign DNA in a cell can be done using several methods.

- Transformation: In this method a large number of copies of DNA of interest is mixed (e.g. a recombined vector) with host cells. This is based on the ability of a cell to take up DNA from its surroundings through the cell membrane. The efficiency is very low in getting the DNA into the cells. Various treatments can increase the competency (ability to take up DNA from outside) of host cells.
- Transduction: This method is based on the ability of bacteriophages to infect host cells. Viruses infecting plants and animals can also be used as vectors to deliver foreign DNA into plant and animal hosts. The gene of interest is integrated into the modified viral genome and packaged into the protein capsid. This viral particle can pass the recombined DNA as in its normal infection process. The capsid protects the DNA and this method is more efficient than transformation.
- Gene Gun: In this method, small particles of a heavy metal such as gold are coated with a large number of copies of the DNA of interest and the particles are shot at high velocity into the cells to be transformed. The equipment used is the gene gun (Figure 7.33).

Figure 7.33: Gene Gun

Agrobacterium **mediated gene transfer:** *Agrobacterium* is a soil bacterium which can infect plants. Their mode of infection is very special. The infection causes a tumour on the plant, and the bacterium is living inside it. The disease is called crown gall disease. The cells in the tumour or gall are actually genetically transformed by a segment of a plasmid of *Agrobacterium*. This plasmid is calleded Ti (tumour inducing) plasmid (Figure 7.34). A part of this plasmid can actually be transferred to plant genome, and hence called transfer DNA or T-DNA. T-DNA contains genes which induce formation of a tumour and also characters associated with pathogenicity. What is needed for transfer of the DNA is the right and left boarder sequences of the T-DNA. Therefore

scientists have removed most of the bacterial genes, including virulent genes from T-DNA and have made room for inserting genes of interest into this space between the two border sequences. *Agrobacterium* can deliver this modified T-DNA with inserted gene to plant cells through its infection ability. However plant cells will not become sick because the virulent genes have been removed from T-DNA. This is called disarming the T-DNA.

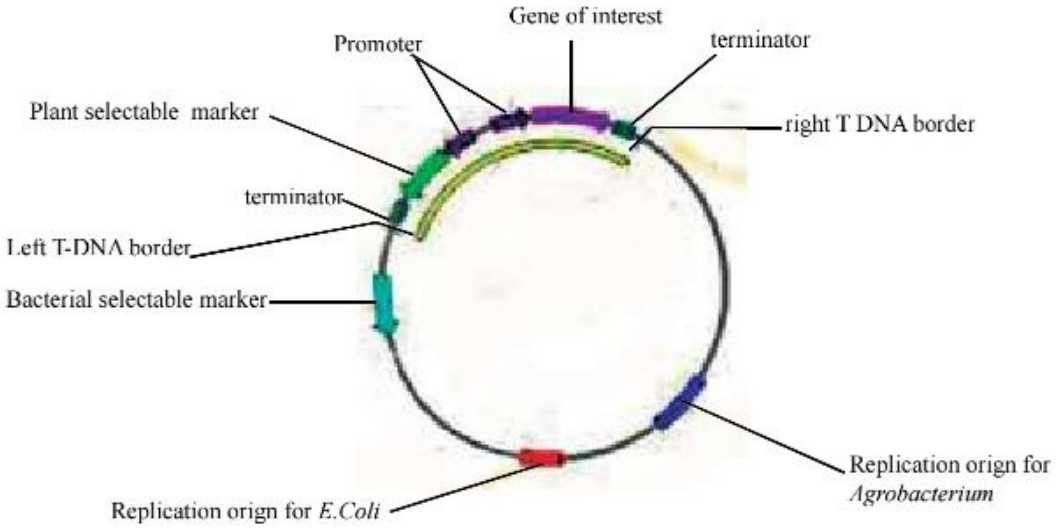


Figure 7.34: Ti Plasmid vector

DNA Analysis

When the morphological characters are used to classify, the number of characters available is limited, so that usually the smallest group identified is a species. When more characters are available, more divisions such as subspecies, strains, varieties etc. become possible. Biochemical properties (i.e. enzyme activities) are useful characters in classification to separate organisms into smaller groups. Since characters are controlled by the combination of the genetics of the organism and its environment, above mentioned characters may vary depending on the environment. Hence if one needs to look how two groups of organisms are similar or different genetically, he has to investigate at the DNA level.

Various techniques for DNA analysis have been developed to facilitate the identification of genetic similarities and differences among organisms and some of these can be used even to identify an individual. These techniques are used in conjunction with techniques such as DNA isolation, gel electrophoresis and use of probes that were discussed above.

Restriction Maps

As shown earlier, restriction enzymes cut dsDNA into fragments at specific sequences. Depending on the number of restriction sites and where they are located, a number of fragments of varying sizes will be generated. Different restriction enzymes cut at different sites and generate different numbers of fragments having different sizes. A restriction map is a diagram showing the position of each restriction site with respect to each other and the distance between these sites (Figure 7.35).

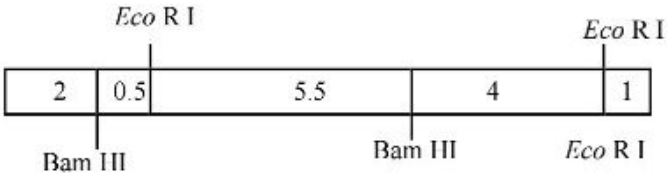


Figure7.35:Restriction map of a small segment of DNA

Restriction maps are very important in construction of cloning vectors: Cloning vectors are cut by restriction enzymes at the cloning site to insert the fragments of DNA from other sources to the cloning site. Restriction map of a commonly used plasmid vector is shown in Figure 7.36.

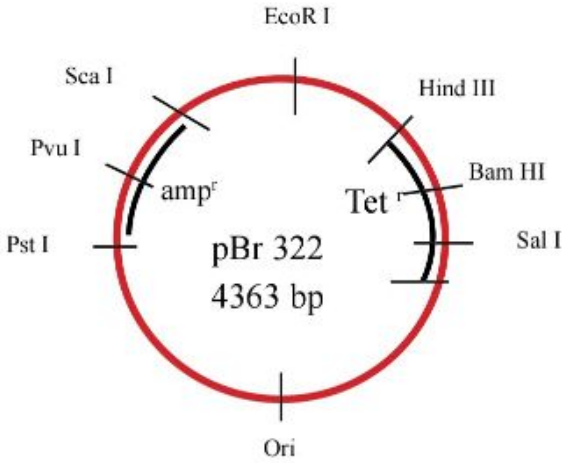


Figure7.36: Restriction map of a pBR322 plasmid vector

DNA Sequencing

A DNA molecule is made of two complementary and antiparallel strands each comprising four bases; adenine, guanine, cytosine and thymine, arranged linearly in a sequence. DNA sequencing is a process to determine the precise order of these bases in a DNA molecule.

The techniques have been improved greatly since the introduction of DNA sequencing in 1977. The technology of DNA sequencing that was available by the time of obtaining the sequence of entire human genome in 2003, under the human genome project was called first generation sequencing technology. Those methods were time consuming and it was possible to determine the sequence only of short DNA fragments. The technology developed since then to next generation sequencing or second generation sequencing and to the most modern third generation sequencing technology. The most modern technology can sequence strands of lengths of millions of nucleotides, and therefore, the time required for sequencing has reduced greatly. Although the Human Genome Project took 15 years to complete the sequencing of human genome, one can today obtain his/ her genome sequenced in several hours at a cost 1000 US$ (2018).

With the development of the technology of DNA sequencing, its applications also have been widened.

Applications of DNA sequencing

Molecular Biology: Information on base sequence of DNA is important in understanding the functions of DNA. Genes coding for a polypeptide can be located by studying the DNA sequence. Certain domains in the DNA sequence of a gene specify the function of the protein. For example, whether the protein will be a transmembrane protein or whether it is a DNA binding protein. DNA sequencing has revealed the presence of multiple copies of genes in human genome. Although amino acid sequence of a peptide can be determined using an amino acid sequencer, it has now become easier to work out the amino acid sequence via DNA sequence.

Evolutionary biology: DNA is passed from generation to generation. Over time, changes occur and accumulated in the DNA. Therefore the similarities and differences in the sequences of DNA of individuals within a species and among different species reveal their evolutionary relationships. Sequencing of DNA samples obtained from ancient humans remains preserved, for example, as mummies or buried in ice or as fossilized remains has given insight into when the *Homo sapience* evolved, and how they migrated to conquer the world.

Medicine: Certain families inherit some genetic disorders. Sequencing of DNA reveals if a healthy person is a carrier or not. How a particular disease causing allele is distributed among family members is very important in assessing the risk and to plan a management. Similarly diagnosis of cancer is also possible through DNA sequencing. The response to a drug given for cancer can be followed by sequencing DNA carried in the blood of a patient. If a drug is responding, then the DNA sequences related to cancer must be reduced in blood. DNA isolated from a placenta of a fetus is useful in early diagnosis of presence of genetic disorders.

Forensics: Except for identical twins, it is extremely rare that two individuals share exactly the same sequence of DNA. That means DNA sequencing can identify an individual having the same DNA sequences as DNA material from (blood, hair, sperm, saliva etc) found in a crime scene. Similarly paternity testing is another use of DNA sequencing.

Metagenomics: Microbiome is the totality of microorganisms in a particular habitat including human body and different environments. Conventional methods for studying the microorganisms in a microbiome are based on culturing in pure cultures. However, a large number of microorganisms are unculturable and hence, are largely neglected. Metagenomics is a science in which DNA present in an environment is extracted as community DNA and studying this sample as a whole. Sequencing some specific sequences in this community DNA and analysis using suitable software, will reveal the number of different species and their identity. Some of these are presently known and many others may be new species. Therefore, metagenomics studies are important in ecology, epidemiology and other fields.

DNA Fingerprinting

The unique set of genetic markers of an individual makes its DNA fingerprint or genetic profile. The presence/ absence of markers are nowadays determined by mostly PCR using specific primers for the marker (see below). These markers are called Small Tandem Repeats (STR markers) or Microsatellite DNA. Eukaryotic DNA contains some non-coding sequences where two to six base pairs are repeated tandemly (one after the other) 100 to 1000 times, so that the lengths of these repeats vary. Since they are non-coding, variations do not have an impact on phenotype. These are variable in individuals, so that can be used as makers. The advantages of using STR markers are

- they occur frequently in genome
- easily amplified by PCR
- highly variable polymorphisms and
- a large number of characterized STRs are available.

The method used earlier was probing specific sequences using labeled markers as describes earlier (see DNA Probes and hybridization).

In DNA profiling, a set of markers (probes or PCR primers) is used. A DNA fingerprint cannot be obtained by using one marker, since there are many individuals with the same banding pattern. When more and more markers are used in combination, the probability of finding the same pattern reduced. It has been calculated that, if 13 markers are used the probability becomes between 10 billion to several trillions. Since the world population is in the region of seven billion, it is highly unlikely that two individuals to have the same genetic profile/ fingerprint.

Applications of DNA Fingerprinting

Criminal identification and victim identification (Figure 7.37) – The fingerprints of the suspects are matched with fingerprints obtained from biological material obtained from a crime scene. The expert's opinion on the identity of the criminal is accepted by the Courts of law.

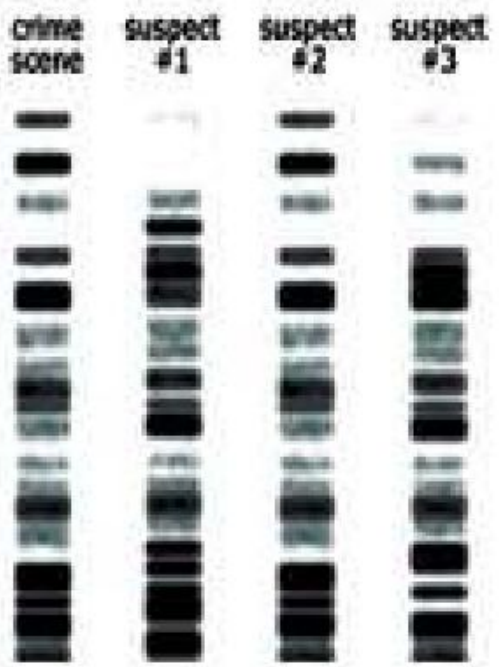


Figure7.37: Comparison of DNA fingerprints of a sample from the crime scene and three suspects.

The profile of suspect #2 matches with that of the crime scene sample.

Paternity testing (Figure 7.38). The DNA fingerprint of a child is never identical to that of either the father or the mother. However, the child shares some markers from the father and others from the mother. Therefore, when the paternity of a child is in question, DNA profiling can be accurately used to confirm or rule out a person as the father of that particular child (Figure 7.38).

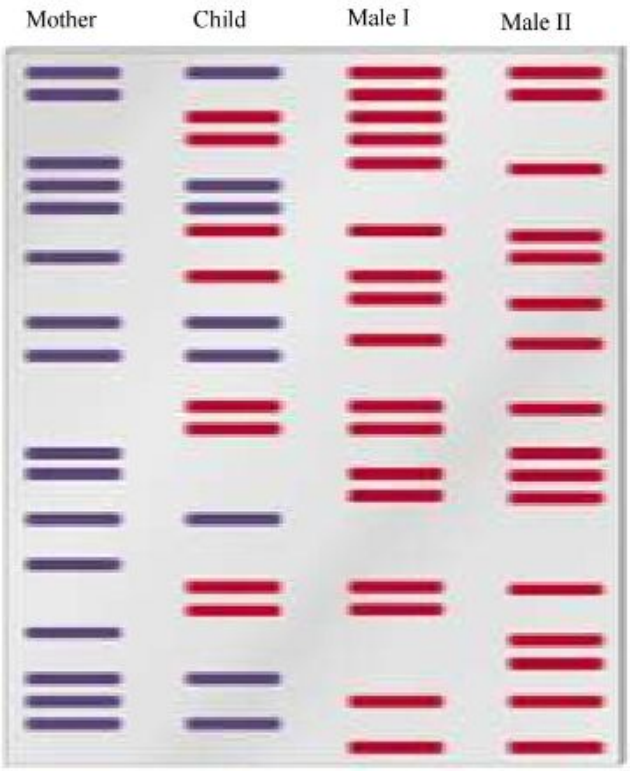


Figure7.38: Use of DNA fingerprinting for paternity testing

Identifying infectious agents: When probes or primers are known to obtain the fingerprint of a pathogenic infectious organism, DNA fingerprinting can detect the presence or absence of this organism in the patient or in food or water.

Polymerase Chain Reaction (PCR)

Polymerase Chain Reaction (PCR) is used to copy sequences of DNA *in vitro,* mimicking the DNA replication. As in replication, a DNA polymerase enzyme is needed to catalyze the reaction of elongation of the new DNA strand. The raw materials needed are the deoxyribonucleotide triphosphate (dNTPs), which are the four deoxyribonucleotides (dATP, dGTP, dTTP and dCTP). A single stranded template DNA is required. Since the DNA polymerase cannot initiate DNA replication, a primer is also needed. In PCR, a primer is a specific sequence of DNA with a small number of nucleotides (oligonucleotide) which is complementary to a sequence at 3' end of the target DNA to be copied. To copy both the strands two primers, each binding to either of the two strands at 3' end, are used. In the cell, the primer is an RNA sequence. In addition to these, Mg^{2+} is also required. These become the ingredients in PCR mix.

The DNA fragment in which the sequence to be copied is present in dsDNA, and hence it needs to be denatured. Denaturation is done by heating the PCR mix to 95 °C. At this temperature most of the enzymes will denature, and therefore it may be necessary to add DNA polymerase after the denaturation. However, the enzymes of the thermophilic organisms are resistant to high temperature. Therefore, the common heat resistant DNA polymerase used in PCR is Taq DNA polymerase which is obtained from thermophilic bacterium *Thermus aquaticus*. The primer will bind to the complementary sequence of the denatured template DNA. This happens at a lower temperature and this step is called annealing. The annealing temperature will depend on the sequence and the length of the primer. Once the annealing of the primer is achieved, then the primer extension (DNA synthesis) will take place at a different temperature. This is the optimum temperature of the DNA polymerase used. When sufficient time is given, a complementary copy of the template DNA is completed. At the end of this first thermal cycle (denaturation, annealing and extension temperatures), one copy of each strand is produced. HIowever, these are longer than the intended copy of the target DNA sequence (Figure 7.39). After a couple of PCR cycle, an exact copy of the target DNA is synthesized. After this, copies of the target DNA will be produced in an exponential manner (i.e. 2, 4, 8, 16, etc.) after each cycle. A typical PCR will have 35 to 40 cycles. At the end, millions of copies of the desired DNA sequence will be produced from a single template DNA molecule. The PCR product formation in the first three cycles is shown in Figure 7.39.

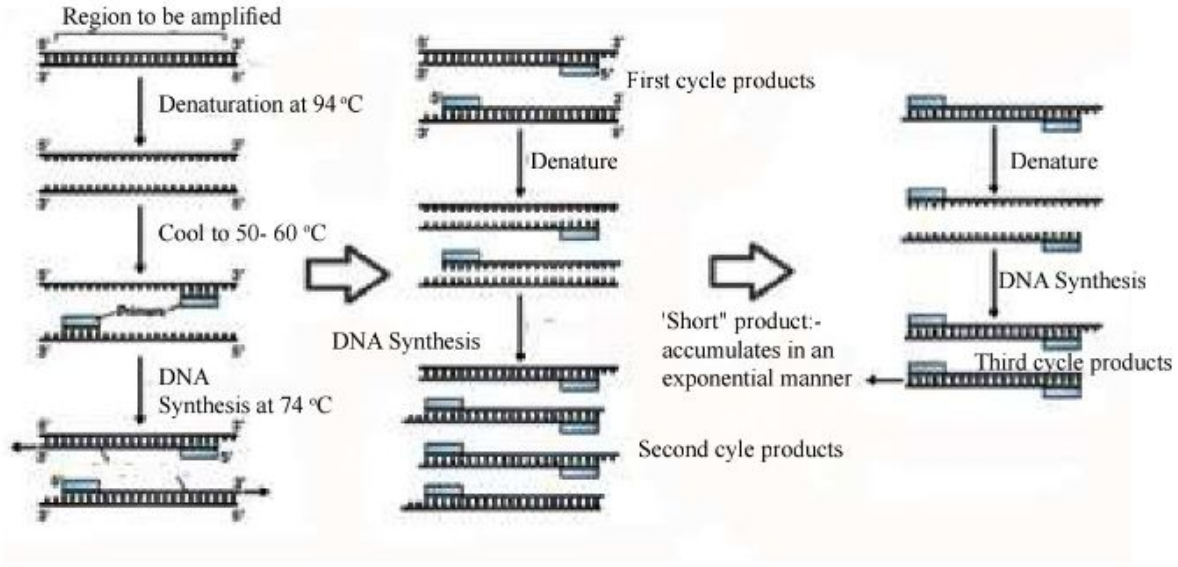


Figure 7.39: PCR products formed at the end of first to third cycles

The repeating cycles are automated and are done in a PCR machine (thermal cycler) (Figure 7.40). PCR mixture is prepared in PCR tubes and these are inserted in to the wells in PCR machine.

Figure 7.40:. A PCR machine showing the wells where the PCR tubes are inserted

PCR is a rapid method of obtaining a large number of copies of DNA with high accuracy.

PCR is an essential technique in cloning procedure which allows generation of large amounts of pure DNA from tiny amount of template strand and further study of a particular gene.

Applications of PCR

- Analyzing clinical specimens for the presence of infectious agents (e.g. HIV, hepatitis, malaria)
- Analysis of mutations causing genetic diseases. (e.g. cystic fibrosis, sickle cell aneamia, phenylketonuria)
- Used in forensics laboratories. Since PCR can make a large number of copies from a small number of template DNA, it is especially useful because only a tiny amount of original DNA is required (e.g. droplet of blood or a single hair).
- PCR is an essential technique in cloning procedure which allows generation of large amounts of pure DNA from tiny amount of template strand and further study of a particular gene.
- Sequencing of DNA depends on PCR.
- PCR has been used to identify, and to explore relationships among species in the field of evolutionary biology.
- In anthropology, it is also used to understand the ancient human migration patterns. In archaeology, it has been used to spot the ancient human race.
- PCR commonly used by paleontologists to amplify DNA from extinct species or cryopreserved fossils of millions of years old, and can thus be further studied to elucidate on their evolutionary relationships.

Applications of Genetically Modified Organisms (GMOs)

The host with the extra trait introduced by genetic engineering techniques is called a genetically modified organism (GMO). Some related terms are Genetically Engineered Microorganisms (GEMs), transgenic organisms and living modified organisms (LMOs). The food and feed obtained from GMOs are called Genetically Modified Food (GMF). Since majority of the crop plants, farm animals and pets animals used today are genetically modified by domestication, it is important to note that what is referred to as GMOs in today's context are essentially the organisms derived as a result of rDNA technology.

(Note: It is important to realize that animal cloning is not a gene technology. It does not include any of the above mentioned steps.)

The steps in the process making a genetically modified plant or an animal are as follows.

1. Identification of a suitable gene
2. Isolation and purification of the gene
3. Amplification of the gene by cloning
4. *In vitro* modifications of the gene of interest
5. Amplification of modified gene by cloning
6. Transformation of the recipient cells (microbial cells, cells or protoplasts of plants or fertilized eggs of animals)
7. Screening for expression of the inserted gene
8. Monitoring for stable integration of the modified gene
9. Backcrossing to introduce the new trait to other varieties of crops and breeds of animals

The applications of gene technology can be found in many fields, including agriculture, medicine and industry.

Applications of GMO in Agriculture

With the growing human population and decrease in agriculture land, it becomes obvious that , the crop yield per unit area needs to be increased. In order to achieve an economically sustainable agriculture, a higher crop yield should be obtained with a low cost of production. Improving the quality of foods is also a major concern in agriculture, in addition to the quantity. The green revolution from 1930s to 1960s increased the crop yield by introducing high yielding crops and use of artificial fertilizer and pesticides. However, impact of green revolution is also limited and is now being overcome by the genetically modified plants (GM crops). When a single plant cell is genetically modified, it can be regenerated in a plant, as plant cells are totipotent. When a useful

trait is introduced into a crop, it can be introduced to other varieties of the same crop by plant breeding.

Most significant contributions in crop yield enhancement by gene technology in agriculture are due to the production of GM crops that are resistant to

- pests and diseases
- herbicides and
- environmental stresses.

In addition, crops with higher nutritional values are also available, e.g. golden rice enriched with vitamin A, increased triglycerides content in Canola oil.

Plants Resistant to Pests: Several GM crops have been engineered with a gene producing a toxic protein that kills the plant eating larval stages of certain lepidopteron and coleopteron insects. Most widely cultivated pest resistant GM plants are cotton, maize, canola and potato. A GM rice variety resistant to lepidopteron insects is also available.

The protein is called Bt toxin, since the protein originally came from a bacterium; *Bacillus thuringiensis*. Different strains of this bacterium produce several different Bt toxins. When larvae feed on plant parts that express Bt toxins, they die due to the ingestion of the toxin.

Bt toxins are not harmful to mammals and are therefore considered safe for human consumption. However Bt maize (Bt corn) is grown mostly for biofuel and for animal feed. Since the toxin is present in the plant tissues, the only insects that are killed are plant pests and therefore Bt crops are also considered safe for beneficial insects. Figure 7.41 show some insect pests of corn and cotton. Bt toxins are natural products and therefore, are biodegradable.

However, the insects, when exposed to the same toxin over the time, develop resistance to that toxin, rendering the GM crop useless. Several solutions to delay the resistance development in insects have been proposed. Since the pollen grains containing the toxin can escape from the field where Bt crop is planted, insects that are not feeding on the crop may also be killed by accidental ingestion of such pollen. Therefore, Bt crops has a potential danger for non-target insects.

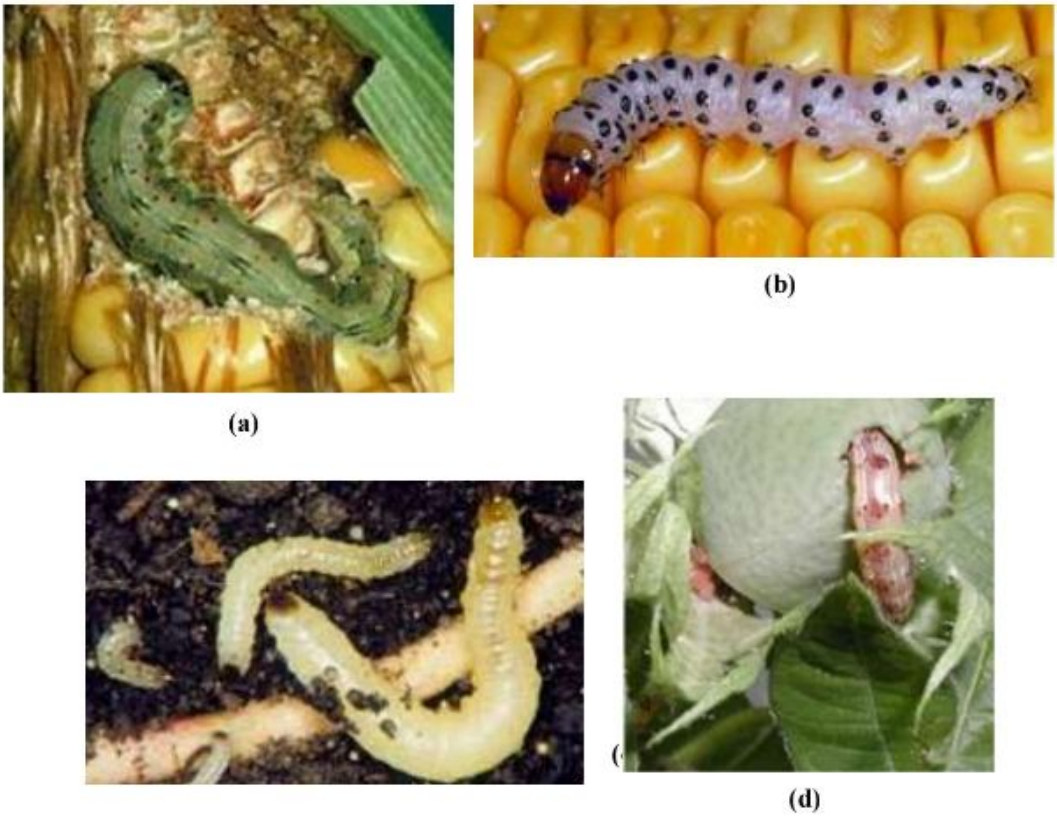


Figure 7.41: Some insect pests of maize: a) Corn earworm, b) European corn borer and c) Corn rootworm. (d) Cotton bollworm

Plants Resistant to Diseases

A well known example of disease resistant crops developed by genetic engineering is the availability of new varieties of papaya resistant to papaya ringspot virus (PRSV). This virus is limiting the success of papaya cultivation worldwide. The same virus also attacks the cucurbits. Squash with resistance to the virus has been successfully developed and cultivated.

Other examples for disease resistant crops are potato varieties resistant to Potato Virus Y (PVY), Potato Leaf Roll Virus (PLRV) and late blight disease.

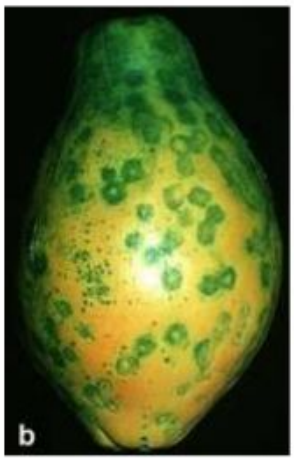


Figure 7.43: Infected papaya fruit with papaya ring spot virus

Plants Resistant to Herbicides: The herbicide tolerant crops (HTCs) are developed in order to be able to spray a broadspectrum herbicide to control weeds when the crop has been established in the field. When the crop is resistant to the herbicide that is sprayed, it can kill all the weeds without harming the crop. The advantage in this is that the farmers can wait and see if the weeds become a problem and can use herbicide only when it is necessary. This reduces the herbicide usage. However if the same herbicide is used repeatedly, the weeds resistant to that particular herbicide will develop. These are called superweeds. When a crop with resistance to one herbicide is planted after a GM crop with resistance to the same herbicide, leftover seeds of the previous crop may germinate and become a weed which cannot be controlled by the same herbicide. In order to avoid these problems, rotation of crops with tolerance to different herbicides can be practiced.

The best known examples of herbicide tolerant crops (HTCs) are crops modified to tolerate the herbicide glyphosate. These plants are known as "RoundUP Ready" crops because the commercial name of glyphosate is RoundUP.. The examples of commercially available RoundUP Ready crops are cotton, maize, canola, soybean, sugar beet and wheat.

Another type of popular herbicide tolerant crops are called "Liberty Link" and "InVigor" and are resistant to glufosinate. Examples of glufosinate resistant crops include cotton, maize, canola, soybean, sugar beet and rice.

Cotton modified to tolerate Bromoxinol are called BXN Cotton.

GM Plants with other Important Traits

Other agriculturally important traits include improving product quality.

One of the priorities in this area is to increase the nutritional value of crops.

- GM canola varieties with increased triglyceride component and with increased phytase enzyme to breakdown indigestible plant phytate to release phosphorus, GM potato with reduced amylase and increased amylopectin content and soybean with increased oleic acid content in the seed are commercially available.
- A rice variety named Golden rice with increased levels of provitamin A is an interesting example of product quality enhancement. The commercially available variety is modified with genes from *Pantoea ananatis*, a plant pathogenic bacterium containing a yellow pigment.
- Tomato with delayed fruit ripening to increase the flavour and reduce the rate of softening is another interesting development in GM crops. The source of the gene is tomato itself. A part of the gene is copied in reverse direction by changing the orientation of the promoter.
- Non-browning apples due to reduced polyphenol oxidation and maize with increased thermostability of amylase which helps in bioethanol production are other examples.
-

Among plants that can be genetically modified to tolerate the environmental stresses, only maize and soybean having drought resistance have been commercialized.

Applications in Medicine

Human insulin, vaccines and other therapeutics are being produced using GMOs. These pharmaceuticals produced by GMO are cheaper, since mass production at low cost is possible and are considered safer.

Producing human insulin by genetically engineered *E. coli* is the best known example of the use of genetically modified organisms in producing pharmaceuticals. Insulin extracted from animals had various side effects on diabetic patients. The cost of production was also very high due to the limited amount of source to extract insulin. Since those insulins are not the same as human insulin, they were less effective. Nowadays the entire insulin supply is from *E.coli* genetically engineered by inserting insulin gene from human. Therefore the bacteria derived insulin is exactly the same as its original product.

Hepatitis B vaccine used today is a recombinant vaccine produced in yeast. Engineering plants to produce vaccines in their edible parts is a concept being experimented. The idea is to express an antigenic protein in the plant cells. When the edible part with the antigen (e.g. fruit) is consumed by someone; the person will develop antibodies against the antigen, so that immunity is developed against a particular disease. These are called edible vaccines, and if this becomes successful, low cost, safe vaccines can be produced and the delivery of the vaccine would be painless. Storage will also not be a main problem. Therefore these are more important in less developed regions of the world.

GM mammalian cells grown in cell culture are used to extract factor VIII used to treat haemophiliacs and tissue plasminogen activator (tPA) used to treat heart attack and stroke patients.

After the completion of human genome project, the causes for various genetic diseases are being identified with ease and speed. When the cause is identified, how to correct the mistake in the defective genes can also be worked out. The incorrect gene then can be replaced with the correct one by gene technology. If the problem is the expression of a specific gene, this technology can also influence the expression of that gene. This treatment is called gene therapy or human gene transfer. The DNA is delivered to the target cells extracted from the patient by using a viral vector or as naked DNA using various techniques. The cells with corrected gene are reintroduced into the appropriate tissue of the patient. Although the concept of gene therapy has a long history since 1972, only a few applications are available to date.

A gene therapy to treat leukemia is the first of that type in the USA. Another example of gene therapy is to replace the mutated beta-globin gene causing sickle cell anemia with the correct gene. During this procedure, bone marrow derived haematopoietic stem cells are extracted from the patient, the normal beta globin gene is inserted into these cells and return the modified cells back into the patient. The corrected bone marrow stem cells will produce normal erythrocytes.
The concept of personalized medicine is being developed, which is based on genetic information of the patient to treat or prevent diseases.
GM insects have been employed to control insect vector born disease. GM mosquitoes have been engineered which do not allow entry of malaria parasite into the gut, therefore breaking the life cycle of the parasite. Release of these mosquitoes into the wild would reduce incidence of malaria. Another example is the making of GM male mosquitoes carrying a male sterile gene. Mass release of these male sterile insects results in mating with females, but not producing offspring. This technology is call 'sterile insect technology' (SIT). In field trials in Brazil, the Aedes aegypti populations have reduced by 95% by the introduction of sterile GM males.

Applications in Industry

Use of GMOs in industry has made it possible to produce novel products with low cost while minimizing environmental impacts. Industries based on organisms or their products run at ambient temperatures and pressures, demanding less energy.

Apart from the booming pharmaceutical industry based on GMOs and GM crop industry, some products originated from GMOs are produced industrially. Some enzymes for food processing and detergents are products of GEMs. Chymosin (rennin or rennet) is the first approved enzyme produced by GMO. This is used in cheese industry to coagulate milk to separate whey. Chymosin has been obtained by extracting it from stomachs of slaughtered calves. Since the supply was limited, the cost was high and it affected dairy industry. The chymosin gene from cattle has been engineered in yeast cells and these recombinant yeasts are the source of chymosin today. The price has gone down considerably and the product is purer and devoid of contaminants of animal origin. Amylomaltase is another enzyme produced by a GM *Bacillus* sp. This enzyme modifies starch to be used as an ingredient in dairy industry.

Aspartame, a strong sweetener, is a food additives produced by GM *E. coli.*

Concerns of using GMOs

The most significant risk factor in use of GMOs is their possible unintended effects. Since this is a relatively new technology, public is reluctant to accept it readily. However the public also accepts the unlimited potentials of the GMOs. The debate between pro-GMO and anti-GMO groups, organizations and individuals is intensive. Some of the controversial issues about GMOs are mentioned below.

Health Issues

Health Issues

1. Data of some experiments involving rats, mice and other animals showed some health implications after feeding with genetically modified food such as potato, maize, tomato and soybean. The reports include damage to various tissues such as stomach, liver and kidney and even increase in deaths. However, many other scientists question the methodology used in such experiments and they also claim that they could not reproduce the results. Therefore, more independent research is needed to confirm or to reject such claims.

2. Allergy development due to consumption of GM food or inhaling pollen of GM crops is discussed as another health issue. The integration of foreign DNA in the host cells may change gene expression or cause mutations, leading to unpredictable products. Some of them may be allergenic, toxic or even carcinogenic. However, concrete scientific research findings are not available or the findings are doubtful. The technology is now developed so that the insertion can be done at precise locations without disturbing other functions of the host.

3. The possibility of horizontal gene transfer of antibiotic resistant genes used as marker gene is also highlighted as a potential health issue. The GM foods contain such genes and are ingested in large quantities by consumers. Although this is a possibility, all organisms have barriers to horizontal gene transfer and therefore, chance of horizontal gene transfer into humans is very low. Horizontal gene transfer among bacteria is more probable, and hence transfer of antibiotic resistance to pathogenic organisms may have some health related issues. However, the antibiotics used in rDNA technology are not used in chemotherapy.

   On the other hand, the humans as well as all other animals have been eating food of animal or plant origin ever since they were evolved, but there is no evidence to show genes transferred due to eating food.

Environmental Issues

1. The development of insect tolerant crops may harm the non-target insect by accidentally ingesting the toxin produced in the GM crops. The toxin may be dispersed in pollen and deposit on non-crop plants which the insects feed on. An experiment has shown that Monarch butterfly larvae feeding on milk weed leaves dusted with pollen from GM crops die. However, proponents of GM argue that the amount of pollen dusted on the leaves is very much higher than the amount that could be deposited naturally.

2. Cross pollination will transfer the transgene (the foreign gene) to other non-GM varieties of the same crop as well as wild relatives of the crop. It can therefore, contaminate organic or non-GM farming.

3. When Bt gene is transferred to wild plants, the insects feeding on them will die causing environmental imbalances.

4. When the herbicide resistance genes are transferred to weeds, they cannot be controlled using the same herbicide, and therefore become super weeds.
5. The spread of a foreign gene in naturally growing plants is called gene pollution.

6. Since the herbicide tolerant crops are resistant to a particular weedicide, farmers may tend to over use this herbicide to keep their fields clean. If that is the case herbicide tolerant weeds will develop by frequent exposure of the weeds to the same herbicide. It is however argued that the farmers will spend money just to have a clean field, and they can use a "wait and see approach" in order to spray herbicides only when necessary rather than spraying as a precaution. Rotation of crops with tolerance to different herbicides will avoid development of such super weeds.

7. As GM crops are being accepted by farmers as well as consumers, the extent of land under cultivation will be dominated by GM crops limited to very few varieties. When the crop diversity is reduced to such a small number, the tolerance to environmental impacts become very low and a single environmental event may wipe out entire crop fields, leading to food scarcity.

8. Narrowing down of crop diversity will contribute to loss of genes from the crop gene pool.

Socio-economic Issues

1. The newly developed GM crop varieties will be owned by the developers and patented. Therefore the farmers are compelled to purchase their seeds every year from giant seed companies with a monopoly, spending a lot of money. Poor farmers may not be able afford to buy the seeds, and therefore, there is a risk of widening the gap between the rich and the poor farmers.

2. There is a growing public concern whether it is ethically correct to patent crops and biological resources including genes that are found in nature. Some crops and products that have been developed traditionally and used by indigenous people have been patented under some biotechnology companies.

3. The consumer has the right to decide whether he/ she purchases GM or non-GM food. To secure this right, the regulatory agencies need to implement a system of labeling to indicate clearly whether the product is GM or not and if it is GM, then what are the changes that have been made. Labeling is mandatory in some countries. However products labeled non-GM have often been found to be contaminated with GM when tested.

4. Biological resources of a country or region with high biodiversity and traditional knowledge are taken away by biotechnology companies without any authorization by the countries and people or without paying compensation for product development. This is called biopiracy.

5. The manipulation of the nature in making GMOs contradicts the beliefs of certain religions.

In order to address potential risks and hazards of GMO, GMF and associated processes, a thorough testing and screening procedures have come into effect. As such the process of producing and commercialization of GMO and GMF are under strict control of several legislations and authorities in order to safeguard the consumer, society and environment. Approval of some GMOs has taken 25 years from the creation to market (e.g. GM Atlantic Salmon which grow twice as faster than non GM).

An example of an international agreement is Cartagena Protocol. Many countries also have their own legislations, such as National Biosafety Framework of Sri Lanka.

Cartagena Protocol on Biosafety

Cartagena Protocol on Biosafety to the Convention on Biological Diversity is an international agreement signed on 15th May, of 2000 in Montreal, Canada as a supplement to the Convention on Biological Diversity (CBD) signed in 1992, at the Rio- Earth Summit and is effective from 1993. Cartagena Protocol on Biosafety came into effect from 11th September, of 2003. It is called Cartagena Protocol on Biosafety since it had been originally scheduled to sign in Cartagena in Colombia. The provisions in CBD were not sufficient to cover many aspects in biodiversity related to the GMOs. The number of signatories to the Cartagena Protocol has exceeded 100. This includes Sri Lanka which gave formal consent to the protocol on 28th April, of 2004.

The aim of the Cartagena Protocol on Biosafety is to protect the biological diversity from the

potential risks caused by genetically modified organisms or living modified organisms (LMOs) created as a result of modern biotechnology. CBD defined biotechnology as "any technology that uses biological systems, living organisms, or derivatives thereof, to make or modify products or processes for specific use". The Protocol is based on 'precautionary principle' of CBD so that products of new biotechnologies should follow rigorous control measures to avoid any potential risk to the environment or human health and is applied to transboundary movement, transit, handling and use of LMOs that may have adverse effect on the conservation and sustainable use of biological diversity. This also applies to risks to human health. The provisions of the Protocol are intended to allow developing nations to balance public health against economic benefits. Countries and states may take appropriate actions to limit entering of an LMO to their territory if they feel that there is lack of scientific information to ensure that the LMO is safe on environment or human health. The LMOs maybe intended to be introduced into the environment or to be used as food or feed. The shipment should accompany with appropriate documentation which identify the LMO and whom to contact for further information. Sufficient information should be supplied by the importer or exporter to the importing party to arrive at an informed decision on accepting or rejecting the LMO imports and how to handle them in a safe manner.

The Protocol established a 'Biosafety Clearing House' (BCH) to assist parties to implement the protocol by facilitation exchange of scientific, technical environmental and legal information on and experience with movement of LMOs.

Sri Lanka signed the Protocol in May, 2000 and it is in force in Sri Lanka from July, 2004. Ministry of Environment and Natural Resources was identified as the responsible institution for coordinating the activities related to the Protocol.

National Biosafety Framework of Sri Lanka

The drafting of the National Biosafety Framework of Sri Lanka (NBFSL) was completed in 2005 by Ministry of Environment and Natural Resources (now Ministry of Mahaweli Development and Environment, MoMDE). This is based on a precautionary approach in accordance with Cartagena Protocol on Biosafety with the objective of ensuring that the risks likely to be caused by modern biotechnology and its products will be minimized and biodiversity, human health and environment will be protected in a maximum way regulating the transboundary movements through formulation of relevant policies, regulations, technical guidelines and establishment of management bodies and supervisory mechanisms.

National Biosafety Framework of Sri Lanka is the starting point for a permanent legislative framework on Biosafety in Sri Lanka. Based on the NBFSL two policies have been formulated.

- National Biosafety, Policy 2005 is one of those. The policy on biosafety sets the overall framework in which adequate safety measures will be developed and put into force to minimize possible risks to human health and the environment while extracting maximum benefits from any potential that modern biotechnology may offer.

- A National Policy on "Access to Biological Resources, sustainable Use and Benefit Sharing" has been formulated in 2013 by MoMDE with the goal of "ensuring conservation and sustainable use of biological resources and the fair and equitable sharing of benefits arising from them" in accordance with Cartagena Protocol on Biosafety and the National Biosafety Framework of Sri Lanka.

However these policies have not yet been enacted by law.

The Target 12 of the National Biodiversity Strategic Action Plan for 2016-2022 of Biodiversity Secretariat of the MoMDE states that "by 2022 biosafety is assured". The actions to be taken in this direction are

(1) strengthening policy on biosafety

(2) implementing biosafety master plan and formulate biosafety legislature

(3) strengthening risk assessment procedures for new technologies

(4) strengthening capacity for risk assessment

(5) developing and implementing legal instruments to protect native biodiversity and indigenous crops from contamination from GMOs and

(6) enhancing Sri Lanka's scientific capacity on biosafety.

References

1. Biology, A Global Approach, 10th edition, Campbell Reece, Urry, Cain,Wasserman. Minorsky, Jackson.
2. NCBI