

THEORY MATERIAL FOR
Ph.D. in Plant Molecular Biology & Biotechnology

**ABT 811 ADVANCES IN PLANT MOLECULAR
BIOLOGY (2+1)**

Prepared by

Dr. T. SABESAN, Ph.D.

**DEPARTMENT OF GENETICS & PLANT BREEDING
FACULTY OF AGRICULTURE
ANNAMALAI UNIVERSITY**

INTRODUCTION TO BIOTECHNOLOGY

Biotechnology broadly refers to development of techniques for the application of biological principles to make products by biological agents. For example, the production of wine by fermentation. The term biotechnology was coined by Karl Ereky, a Hungarian engineer in 1919. The term *plant biotechnology* refers to application of engineering techniques to modify plants for the benefit of humans.

In a narrow definition, *biotechnology* refers to genetic manipulation of organisms for specific purposes. Biotechnology has two branches namely, tissue culture and genetic engineering. *Tissue culture* refers to the *in vitro* cultivation of cells, tissues, organs, embryos, seeds and protoplasts on nutrient media under aseptic conditions in a laboratory. The term *genetic engineering* refers to manipulation of organisms at the molecular level directly involving the DNA.

Genetic engineering can modify bacterial cells to synthesize completely new substances, e.g. hormones, vaccines, monoclonal antibodies, etc., or introduce novel traits into plants or animals.

HISTORY OF PLANT TISSUE CULTURE

Gottlieb Haberlandt is regarded as the Father of plant tissue culture as he predicted the totipotency of plant cells and attempted *in vitro* culture of plant mesophyll cells as early as 1902. Totipotency is the ability of a plant cell to multiply, differentiate and grow into a complete plant. The first embryo culture was attempted by *Hanning* in 1904. In 1925, *Laibach* recovered hybrid progeny from an interspecific cross in *Linum* by using zygotic embryos of seeds as explant.

The first plant growth hormone indoleacetic acid (IAA) was discovered in 1930s by *F. Kogl et al.* In 1934, Professor *Philip White* successfully cultured tomato roots on the medium, later called as White's medium. In 1939, *Gautheret* successfully cultured carrot tissues and the possibility of cultivating plant tissues for an unlimited period was independently endorsed by *Gautheret, White and Nobecourt* in 1939. In 1955 *Miller and Skoog* published their discovery of the hormone kinetin, a cytokinin. *The first plant from a mature plant cell was regenerated by Braun in 1959.*

The focus of the scientists later shifted to preparation of single cell cultures. *Muir* (1953-54) demonstrated that, callus tissues in liquid medium when subjected to shaking, broke into single cells. In 1960, *Bergmann* developed the method for cloning of these single cells by filtering suspension cultures. This technique called Plating technique is widely used for cloning isolated single protoplasts.

In 1962, Toshio Murashige and Skoog published the composition a plant tissue culture medium known as MS medium, which became the most widely used medium for tissue culture now.

For the first time in the world, haploid plants from anthers of *Datura* were first produced by the Indians, *Guha, S and Maheshwari, S.C.* (Nature, 204:497 (1964) & Nature, 212:97-98 (1966)). This discovery received significant attention since, plants recovered from doubling of haploids are homozygous and express all recessive genes thus making them ideal for pure breeding lines.

The classical work of *Steward* (1966) on induction of somatic embryos from free cells in carrot suspension cultures finally demonstrated totipotency of somatic cells, thereby validating the ideas of *Haberlandt*. *Morel* utilized this application for rapid propagation of orchids and Dahlias. He was also the first scientist to develop virus free orchid and Dahlia plants by cultivation of the shoot meristem of infected plants.

Protoplast (a cell without cell wall) was produced by *Cocking* in 1960 by using cell wall degrading enzymes. The first somatic hybrid plants by fusing the protoplasts of *N. glauca x N. langsdorfl* was produced by *Carlson et al.* in 1972.

DEVELOPMENT OF BIOTECHNOLOGY IN INDIA

To promote biotechnology in India the Department of Biotechnology (DBT) was started in 1986. It was initially started as National Biotechnology Board (NBTB) in 1982 under Department of Science and Technology. Later, the International Center of Genetic Engineering and Biotechnology (ICGEB) was established by the United Nations to help the developing countries like India. ICGEB has two centers, one in New Delhi and the other in Trieste (Italy).

SCOPE AND IMPORTANCE OF BIOTECHNOLOGY

Biotechnology is controlled use of biological agents for beneficial use. It involves the integrated application of biochemistry, molecular biology and microbiology to develop technological application for enhancement of the capabilities of biological agents. Thus, biotechnology has emerged as a science with immense potential for human welfare which will be evident from following examples:

Biotechnology in Medicine: Production of monoclonal antibody, DNA and RNA probes for diagnosis of various diseases; synthesis of valuable drugs like insulin and interferon from bacteria for treatment of human diseases; DNA fingerprinting for identification of parents and criminals; Development of recombinant vaccines like human hepatitis B etc. by genetically engineered microbes are some of the notable achievements.

Industrial Biotechnology: It involves large scale production of alcohol and antibiotics by microorganisms. A variety of pharmaceutical chemicals, with better quality and quantity, like lactic acid, glycerine etc. are being produced by genetic engineering. Protein engineering i.e., remodeling of existing proteins and enzymes for a specific function or for increasing efficiency of their function is gaining momentum.

Biotechnology and Environment: Environmental problems like pollution control, depletion of non-renewable energy resources, conservation of biodiversity etc are being dealt with using biotechnology. For e.g. bacteria are being utilized for detoxification of industrial effluents, in removing oil spills, for treatment of sewage and for biogas production. Bio-pesticides give environmentally safer alternative to chemical pesticides for control of insect pests and diseases.

Biotechnology and Agriculture: In agriculture, plant cell, tissue and organ culture is used for rapid and economic clonal multiplication of fruit and forest trees, for production of virus free genetic stocks and planting material as well as in the creation of novel genetic variations through somaclonal variation.

Genetic engineering techniques are utilized to produce transgenic plants with desirable genes like disease resistance, herbicide resistance, increased shelf life of fruits etc. Also, molecular breeding has hastened the process of crop improvement for e.g. molecular markers like RFLP, SSRs provide powerful tools for indirect selection of both qualitative and quantitative traits and also for studying genotypic diversity.

Landmark discoveries in the field of molecular biology

YEAR	CONTRIBUTION
1970	Smith and Nathans discovered first restriction enzyme from <i>Haemophilus influenza</i> (HindIII)
1972	First recombinant DNA molecule produced by Berg <i>et al.</i> combining SV40 virus and λ virus.
1972	Discovery of reverse transcriptase by Termin
1974	Biotransformation in plant tissue culture by Reinhard
1974	Discovery of Ti plasmid as tumour inducing principle by Zaenen <i>et al.</i> , and Larebeke <i>et al.</i>
1977	Successful integration of Ti plasmid DNA from <i>Agrobacterium tumefaciens</i> in plants by Chilton <i>et al.</i>
1977	DNA sequencing methods based on degradation of DNA by Maxam and Gilbert.
1978	Crossing potato and tomato to produce pomato by somatic hybridization by Melchers <i>et al.</i>

1980	Commercial production of human insulin in bacterial cells through genetic engineering by Eli Lilly and Co.
1980	Restriction Fragment Length Polymorphism (RFLP) technique developed.
1981	Coining of the term <i>somaclonal variation</i> by Larkin and Scowcroft.
1983	Polymerase Chain Reaction (PCR), a chemical DNA amplification process conceived by Kary Mullis.
1984	Transformation of tobacco with <i>Agrobacterium</i> and development of transgenic plant by De Block <i>et al.</i> , and Horsch <i>et al.</i>
1984	Development of genetic fingerprinting technique for identification of individuals by analyzing DNA polymorphism by Jeffreys.
1987	Isolation of Bt gene from <i>Bacillus thuringiensis</i> bacterium by Barton <i>et al.</i>
1990	Launch of Human Genome Project formally.
1990	Development of Random Amplified Polymorphic DNA (RAPD) technique by Williams <i>et al.</i> , and Welsh & McClelland.
1991	Development of DNA microarray system by Fodor.
1995	Development of DNA fingerprinting by Amplified Fragment Length Polymorphism (AFLP) technique by Vos <i>et al.</i>
1997	Completion of DNA sequencing of <i>E.coli</i> genome by Blattner <i>et al.</i>
2001	Successful completion of Human Genome Project by Craig Venter <i>et al.</i> of Celera genomics

Questions

1. Define plant biotechnology.
2. Who coined the term biotechnology?
3. Haploid plants were first produced *in vitro* by _____.
4. _____ is considered as the father of plant tissue culture.
5. Polymerase chain reaction was conceived by _____.
6. Write an essay about the scope and applications of biotechnology?

THE STRUCTURE OF DNA

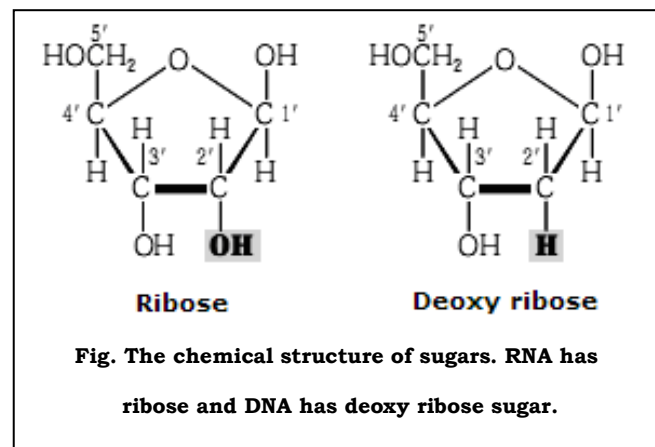
The structure of DNA can be studied in three levels as primary, secondary and tertiary structure. The primary structure refers to the structure of nucleotide and how they are joined together. The secondary structure refers to the double helical structure proposed by Watson and Crick. The tertiary arrangement refers to the complex packing of the double stranded DNA in chromosomes.

I. PRIMARY STRUCTURE OF DNA

Nucleic acids (DNA and RNA) are polymers made up of repeating units of **nucleotides**. A nucleotide consists of three basic components: 1. Pentose sugar, 2. Nitrogenous base, and 3. Phosphate group.

1. PENTOSE SUGAR

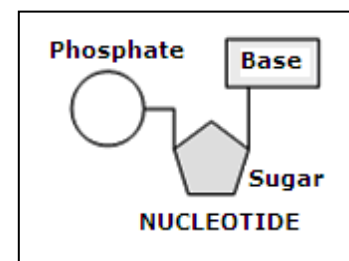
The sugar is a cyclic five-carbon structure. It contains a hydroxyl group attached to 2'-carbon atom and called as ribose in RNA. In DNA, the sugar lacks oxygen atom and hence called deoxyribose. This minor chemical difference is recognized by all the cellular enzymes that interact with DNA or RNA, thus yielding specific functions for each nucleic acid. Further, the additional oxygen atom in the RNA nucleotide makes it more reactive and less chemically stable than DNA.

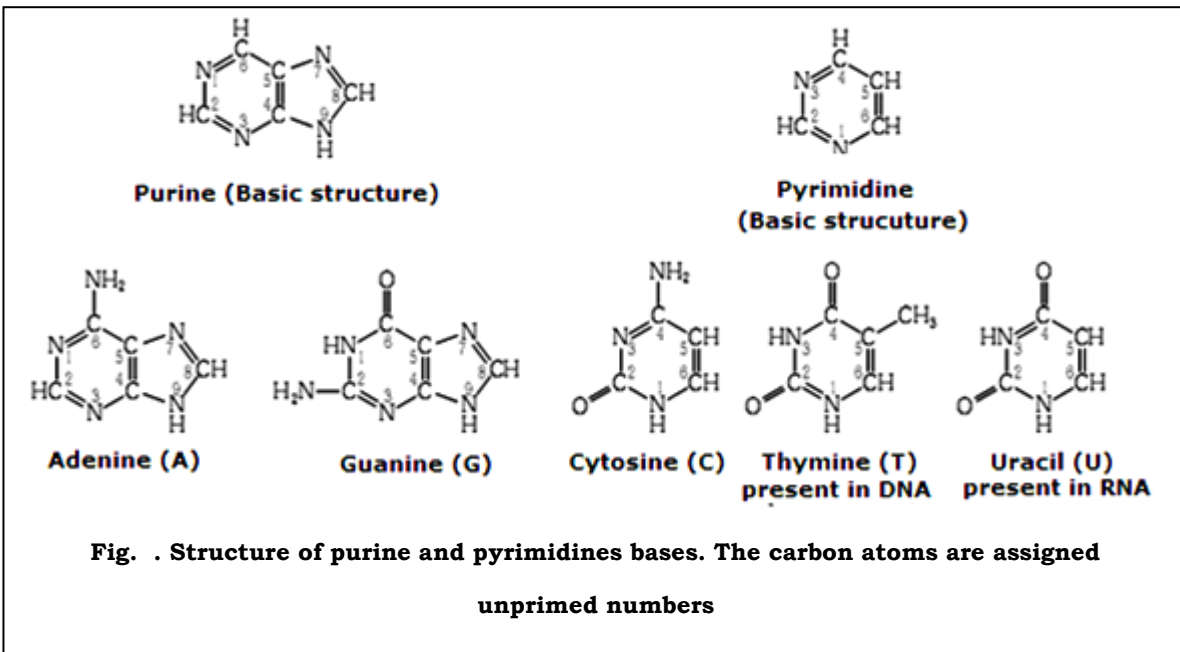


The carbon atoms are numbered as 1' (called one prime), 2', 3', 4' and 5' in order to differentiate them from the carbon atoms in the DNA and RNA bases. The 5' and 3' carbons of the pentoses forms the phosphodiester linkage, while the 1' carbon is always occupied by an organic base.

2. NITROGENOUS BASES

There are two kinds of bases: **purines** and **pyrimidines**. Purines are double ringed structure with a six-sided ring attached to a five-sided ring whereas, a pyrimidine consists of a six-sided ring only. There are two purines, adenine (A) and guanine (G), and three pyrimidines, cytosine (C), thymine (T), and uracil (U). Thymine occurs only in DNA, while uracil occurs only in RNA. The letters A, C, T, G, are usually referred to as the alphabets of life.





When a base is linked to a sugar, the product is called a **nucleoside**. A nucleoside linked to a phosphate is called as a **nucleotide** (Nucleotide = Nucleoside + Phosphate) (Fig.). In a nucleotide, the nitrogenous base always forms a covalent bond with the 1'-carbon atom of the sugar.

Components of Nucleic Acids

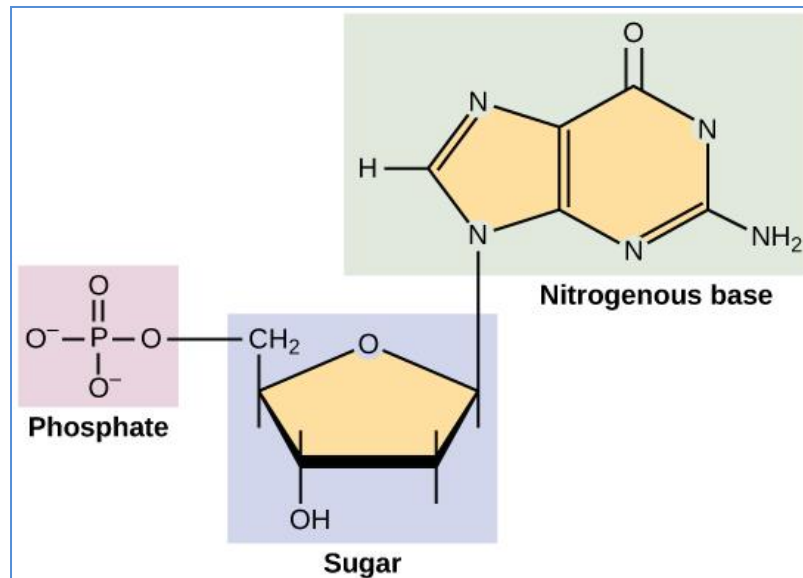
	DNA only	DNA & RNA			RNA only
Nitrogen Bases	<p>Thymine</p>	<p>Adenine</p>	<p>Guanine</p>	<p>Cytosine</p>	<p>Uracil</p>
Sugars & Phosphate	<p>2-Deoxyribose</p>	<p>Phosphate</p>			<p>Ribose</p>

3. PHOSPHATE GROUP

The third component of a nucleotide is the phosphate group, which consists of a phosphorus atom bonded to four oxygen atoms. Phosphate groups are found in every

nucleotide and frequently carry a negative charge, which makes DNA acidic. The phosphate is always bonded to the 5-carbon atom of the sugar in a nucleotide.

The DNA nucleotides are properly known as **deoxyribonucleotides** or deoxyribonucleoside 5-monophosphates. Because there are four types of bases, there are four different kinds of DNA nucleotides. The equivalent RNA nucleotides are termed **ribonucleotides** or ribonucleoside 5-monophosphates.



POLYNUCLEOTIDES

Two nucleotides are linked by a phosphodiester group i.e. the 5'-phosphate group of one nucleotide joins to the 3'-carbon atom of the next nucleotide. These bonds, called **phosphodiester linkages**, are relatively strong covalent bonds. Shorter chains (consisting of less than 20 nucleotides) are called oligonucleotides while longer chains are called **polynucleotides**.

An important characteristic of the polynucleotide strand is its direction or polarity. At one end of the strand a phosphate group is attached only to the 5'-carbon atom of the sugar in the nucleotide. This end of the strand is therefore referred to as the **5' end**. The other end of

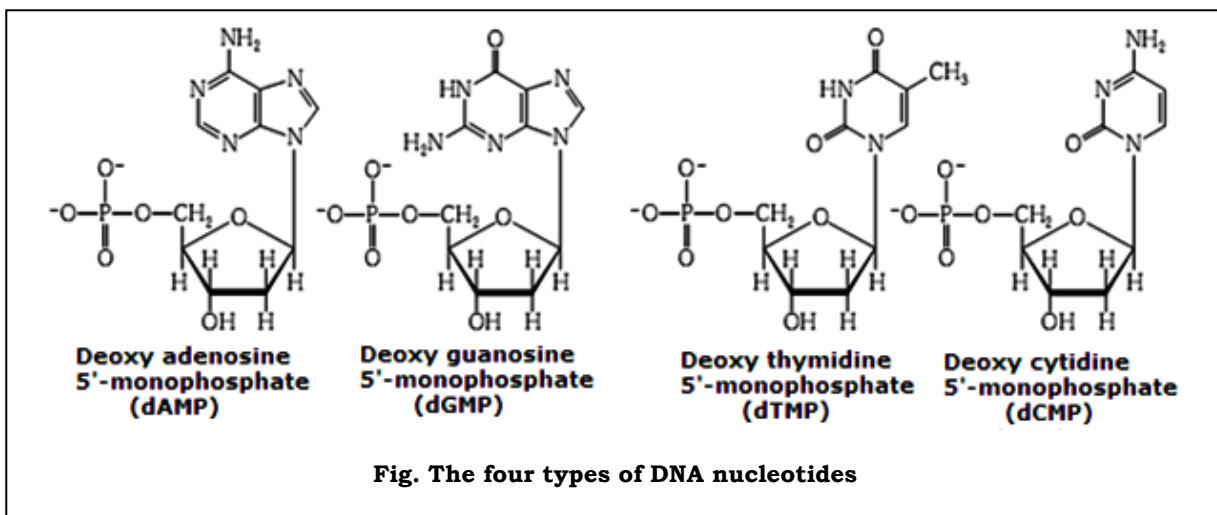


Fig. The four types of DNA nucleotides

the strand, referred to as the **3' end**, has an OH group attached to the 3-carbon atom of the sugar.

RNA nucleotides also are connected by phosphodiester linkages to form similar polynucleotide strands.

II. SECONDARY STRUCTURE OF DNA

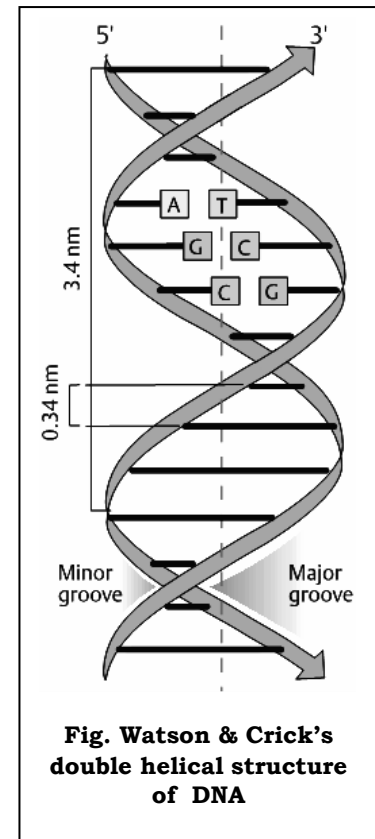
The secondary structure refers to the **double helical structure of DNA** proposed by Watson and Crick at Cambridge.

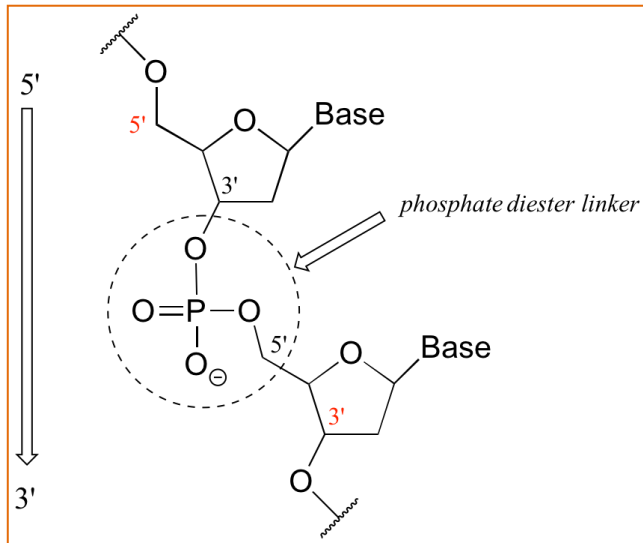
The X-ray diffraction analysis of Rosalind Franklin and Maurice Wilkins revealed that the DNA was a double helix with a width of 2 nm. The purine and pyrimidines bases were stacked 0.34 nm apart in a ladder. Since the width of the helix is 2 nm it can accommodate only two strands. Also, Erwin Chargaff found that $A=T$ and $G=C$ and that the number of purine bases ($A + G$) is equal to the number of pyrimidine bases ($C + T$).

Using the Chargaff's data and X ray diffraction studies Watson and Crick, built the molecular model of DNA using metal wires. Watson and Crick along with Maurice Wilkins, were awarded a Nobel Prize in 1962.

The key features about the double helical structure of DNA molecule are as follows:

1. The DNA double helix (DNAdh) consists of two polynucleotide chains coiled around a central axis in a spiral fashion.
2. The polynucleotide chains are antiparallel; one chain runs in the 5' to 3' orientation and the other in 3' to 5' direction. This anti-parallel orientation of the two strands is essential for the formation of hydrogen bonds between the pairs of DNA bases.
3. The two bases in each base pair lie in the same plane which is perpendicular to the axis of the helix. Neighbouring bases lie 3.4 Å apart. There are 10 base pairs per helical turn i.e. the helix repeats itself at an interval of 34 Å.
4. The helix has two kinds of alternating external grooves: a deep groove (called the major groove) and a shallow groove (called the minor groove).
5. The nitrogenous bases on one strand pair with those on the other strand in complementary fashion (A always pairs with T, while G pairs with C). The most common natural form of DNA is a right-handed double helix of diameter 2.0nm, called the B-DNA. DNA can also assume other forms like Z DNA, A DNA.





In addition to these features described above, certain implications deserve emphasis:

1. **Complementary base pairing** means that the replicate of each strand is given the base sequence of its complementary strand when DNA replicates.

2. Because the strands are **antiparallel**, when two nucleotides are paired, the sugar portions of these molecules lie in opposite directions (one upward and the other downward along the chain) (Fig.).

3. Because the strands are antiparallel, the convention for writing the sequence of bases in a strand is to start from the 5' P terminus at

the left (e.g., GAC refers to a trinucleotide 5' -P-GAC-3' -OH).

4. The conventional way of expressing the base composition of an organism is by the percentage of [G] + [C]. This value is approximately 50% for most eukaryotes with only

minor variations among species. In simpler organisms, there are significant variations (e.g., 27% for *Clostridium*, 50% for *Escherichia coli*, and 76% for *Sarcina*, all of these organisms being bacteria).

5 The chains of the double helix are held together by hydrogen bonds between base pairs in opposite strands. The bond between A and T is a double bond, while the bond between G and C is a triple hydrogen bond.

The DNA molecule satisfies the requirement of genetic material in the following ways:

1. It can replicate itself accurately during cell growth and division.

2. Its structure is sufficiently stable so that heritable changes i.e., mutations

can occur only very rarely.

3. It has a potential to carry all kinds of necessary biological information.

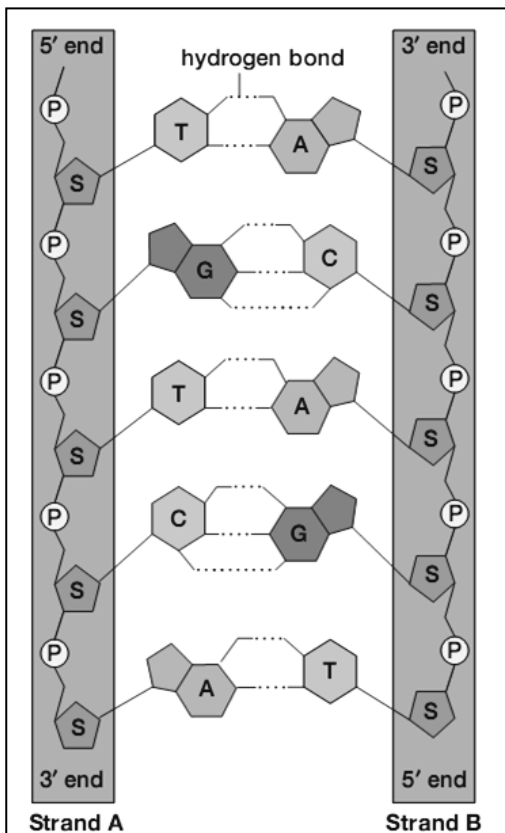


Fig. DNA consisting of two polynucleotide chains running antiparallel to each other & Composed of complementary strands. RNA has single nucleotide chain.

4. It transmits all the biological information to the daughter cells.

Thus the essential functions of DNA are the storage and transmission of genetic information and the expression of this information in the form of synthesis of cellular proteins.

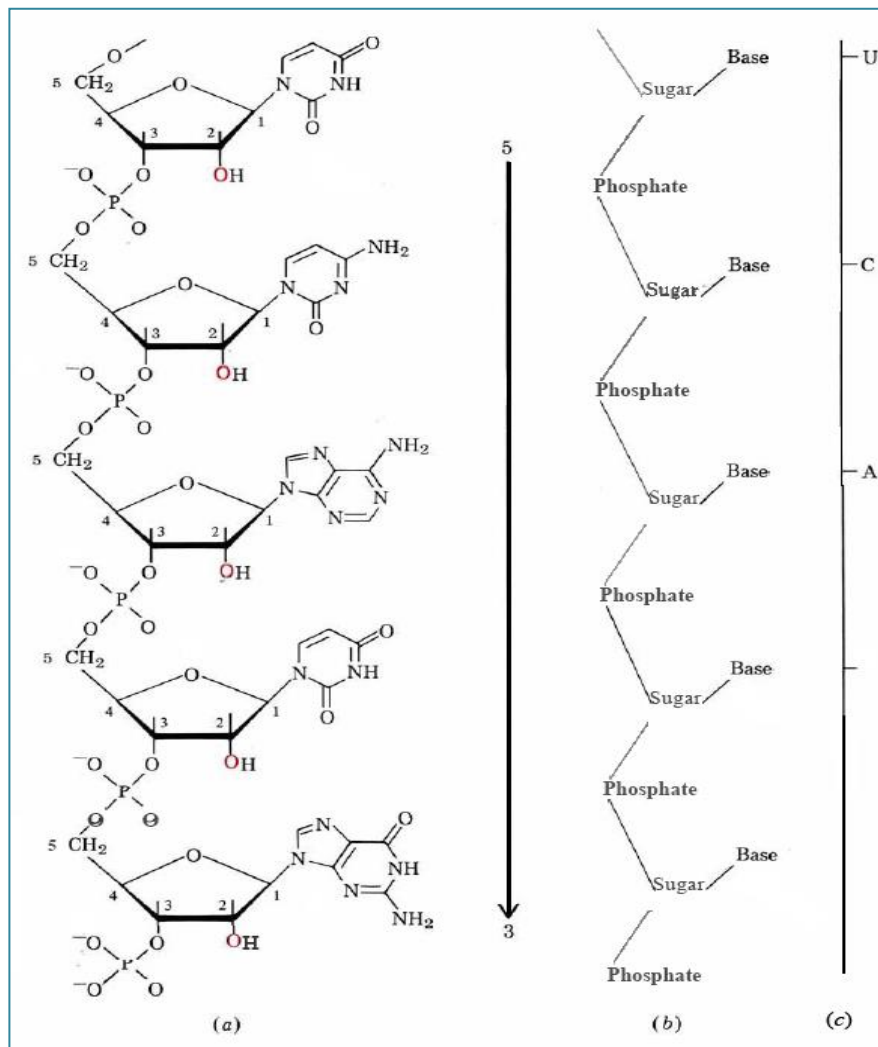


Nucleotide biosynthesis

Base	Nucleoside (base + ribose)	Nucleotide (nucleoside + phosphate)
Adenine	Adenosine	AMP (adenylic acid or adenosine monophosphate)
Guanine	Guanosine	GMP (guanylic acid or guanosine monophosphate)
Uracil	Uridine	UMP (uridylic acid or uridine monophosphate)
Thymine	Thymidine	TMP (thymidylic acid or thymidine monophosphate)
Cytosine	Cytidine	CMP (cytidylic acid or cytidine monophosphate)

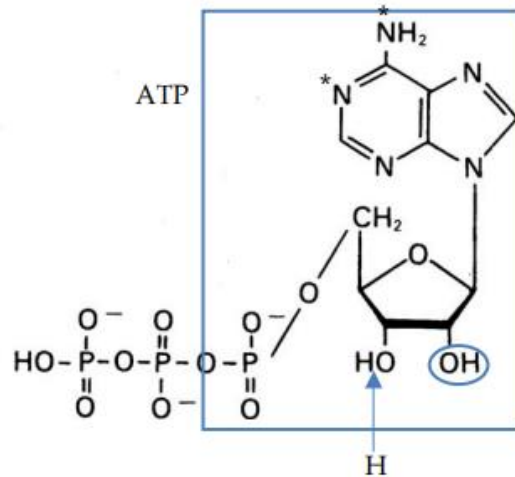
Denaturation: The hydrogen bonds between the DNA strands break on heating the DNA to high temperature (nearly 100 °C). The process of separation of DNA strands is known as denaturation.

Renaturation: Reunion of the separated or denatured DNA strands on cooling is called renaturation or annealing. The optimum temperature for renaturation is 20 – 25 °C.



Practice Problems in understanding the concept of DNA structure

1.



- Box the part that is added to a growing chain of nucleic acid.
- Star the **atom(s)** that can form a hydrogen bond with the complementary nitrogenous base.
- Circle the part of the molecule that decreases the stability of RNA as compared to DNA.
- Draw an arrow to the part of this molecule that you would modify to prevent further elongation. Indicate what change you would make next to the arrow drawn.

2. For the DNA strand 5'-TACGATCATAT-3' the correct complementary DNA strand is:

- 3' -TACGATCATAT-5'
- 3' -ATGCTAGTATA-5'
- 3' -AUGCUAGUAUA-5'
- 3' -GCATATACGCG-5'
- 3' -TATACTAGCAT-5'

DNA REPLICATION

The process by which a DNA molecule makes identical copies of itself is called as DNA replication.

THREE MODELS OF DNA REPLICATION

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

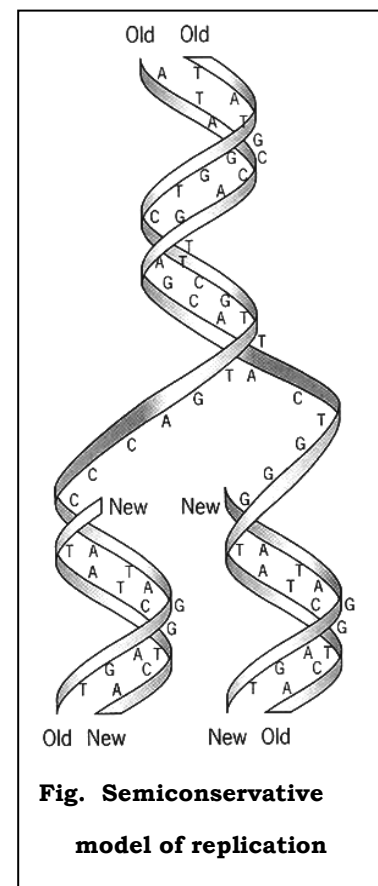
2. Conservative: According to conservative scheme, after replication the two newly synthesized strands would form a double helix, while the two old parental strands would form another double helix.

3. Semi conservative: In the semi conservative model of DNA replication, one of the two parental DNA strand serves as a template for the synthesis of new, complementary daughter strand. Each progeny DNA molecule would consist of one old and one newly synthesized strand. As one of the parental strand is conserved in the new daughter DNA molecule it is called as semi conservative model and is the universally accepted model.

EVIDENCE FOR SEMICONSERVATIVE REPLICATION

The experimental evidence for the semiconservative method of replication was provided by Meselson and Stahl (1958). They cultured *E. Coli* in a medium containing radioactive ^{15}N . and labeled its DNA with ^{15}N . This radioactively labelled *E.coli* was transferred to a medium containing normal nitrogen, ^{14}N , and allowed to divide.

The DNA of newly formed daughter cells contained one DNA strand labelled with ^{15}N and the other strand labelled with ^{14}N . This showed that each progeny DNA molecule would consist of one old and one newly synthesized strand.



REQUIREMENTS OF REPLICATION

Although the process of replication includes many components, they can be combined into three major groups:

1. A template consisting of single-stranded DNA,
2. Substrates for new nucleotide strand, and
3. Enzymes and other proteins that *read* the template and assemble the substrates into a DNA molecule.

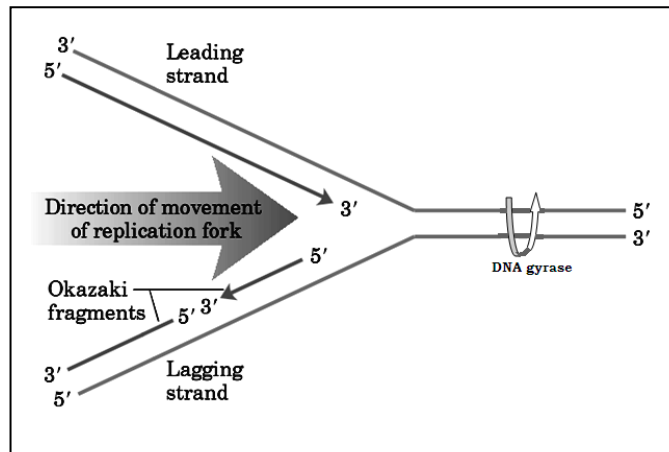
Table - Major components required for DNA replication in bacteria

Sl. No	COMPONENTS	FUNCTION
1.	Initiator protein	Initiates replication by binding to origin and unwinds DNA strands
2.	DNA helicase	Unwinds DNA at the replication fork
3.	Single-stranded binding proteins	Attach to single stranded DNA and prevent reannealing (re-joining)
4.	DNA gyrase	Moves before replication fork, cutting and resealing breaks in the DNA to release the torque developed in it due to unwinding of DNA
5.	DNA primase	Synthesizes short RNA primers
6.	DNA polymerase III	Elongates a new nucleotide strand
7.	DNA polymerase I	Replaces RNA primers with DNA
8.	DNA ligase	Joins Okazaki fragments by sealing sugar-PO ₄ backbone

MAJOR STEPS IN DNA REPLICATION

- 1). Replication begins when an initiator protein binds to origin of replication and unwinds a short stretch of DNA, to which DNA helicase attaches.
- 2). *DNA helicase* then unwinds the DNA at the replication fork and *single-strand-binding proteins* bind to single nucleotide strands to prevent them from reannealing.
- 3). *DNA gyrase* (a topoisomerase) removes the strain ahead of the replication fork that is generated by unwinding.
- 4). *Primase* synthesizes short primers of RNA nucleotides, providing a 3'-OH group to which *DNA polymerase* can add DNA nucleotides.

- 5). *DNA polymerase* adds new nucleotides to the 3' end of a growing polynucleotide strand.
- 6). The leading strand is continuously synthesized.
- 7). The other strand, which is synthesized discontinuously in short pieces is called as lagging strand or Okazaki fragments.
- 8). The Okazaki fragments are joined together by DNA ligase. In bacteria, Okazaki fragments are ~1,000 to 2,000 nucleotides long. In eukaryotic cells, they are 150 to 200 nucleotides long.
- 9). *DNA ligase* seals the nicks that remain in the sugar-phosphate backbones when the RNA primers replaced by DNA nucleotides.

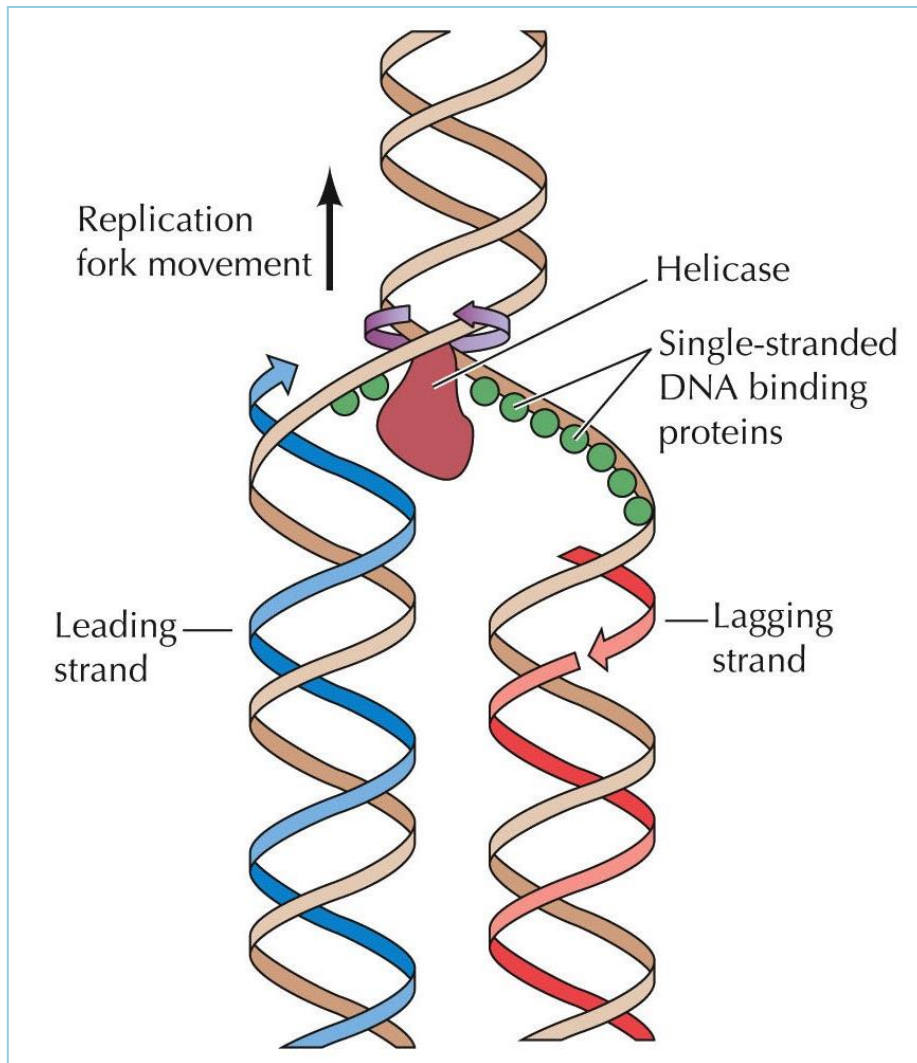


are

The high accuracy of DNA replication is maintained by (1) Base selection by the polymerase, (2) Proofreading by exonuclease from 3' to 5' activity and (3) Repair systems which replaces mismatches in replication.

EUKARYOTIC DNA REPLICATION

Although not as well understood, eukaryotic replication resembles bacterial replication in many respects. The most important differences are that eukaryotes have: (1) Multiple origins of replication in their chromosomes; (2) Different types of DNA polymerases, with different functions; and (3) Assembly of nucleosome immediately after DNA replication. Precise replication at multiple origins is ensured by a licensing factor that must attach to an origin before replication can begin. The licensing factor is removed after replication is initiated and renewed after cell division. The ends of linear eukaryotic DNA molecules are replicated by the enzyme *telomerase*.



Practice Problems in understanding the concept of DNA replication

Problem 1

Consider the following origin of replication that is found on a chromosome. The sequence of region 1 is shown.

Region 1: 5' ...CTGACTGACA...3'
3' ...GACTGACTGT...5'

a) Within Region 1, which strand will be the template for leading strand synthesis, the top or the bottom?
The bottom strand will be the template for leading strand synthesis.

b) If we assume that a lagging strand fragment is made from region 1, what will be its sequence?
5' IGTCAGTCAG 3'

Problem 2

Shown below is a schematic of replicating DNA in a bacterial cell.

Region 1 Origin Region 2

A). On the diagram, label the 5' and the 3' ends of the parental DNA strands.

B) Which parental DNA strand (top or bottom) serves as a template for the synthesis of the leading strand in Region 2?

c) To which site (A, B, or both) can the primer 5' UUGC 3' bind?

d) The replication of which strand (top, bottom, or both) in Region 2 would be affected in the absence of RNA primase? Assume that replication has not yet initiated on either strand. Explain.

e) You perform DNA replication in a test tube (in vitro) using a single-stranded linear DNA as the template and the appropriate DNA primer. From the list below, circle the proteins that are required for one round of replication.

Primase - DNAPolymerase - Ribonuclease - Topoisomerase - Ligase - Helicase

f) Why does the DNA of a eukaryotic cell require multiple origins of replication when some prokaryotic cell genomes have only one origin of replication?

g) The “leading strand” terminology refers to the fact that this strand is the first daughter DNA strand to be completed from a given replication fork. Explain why the leading strand is completed before the lagging strand.

h) Which of the following statements is true?

a. DNA polymerase moves along the template strand in the 3' to 5' direction.

b. DNA polymerase requires a primer to initiate replication.

c. DNA polymerase forms a covalent bond between a phosphate group and a hydroxyl group.

d. Polymerization of DNA for both the lagging and the leading strands is in the 5' to 3' direction.

e. Shortly after DNA replication has begun, you find two replication forks moving in opposite directions away for the Origin of Replication.

f. All of the above statements are true.

g. None of the above statements are true.

THE GENETIC CODE

A sequence of three nucleotides in mRNA, codes for an amino acid. This triplet is referred to as **codon**. For example, the mRNA sequence AUG codes for the amino acid methionine. The set of all the codons that specify the 20 amino acids is termed as **the genetic code**.

The set of bases in tRNA which base pair with a codon of mRNA is known as **anti-codon**. The sequence of bases in an anti-codon is exactly the opposite of that present in the codon. The codon is read in a 5' to 3' direction.

There are 20 amino acids involved in protein synthesis and there are only four bases (A,T,G,C) in the DNA coding for all the amino acids. Thus the 4 bases when arranged in the form of triplet code (4^3) can generate 64 codons. Of these, three codons, UAA, UAG and UGA do not code for any amino acid and serve as **stop codons** (*nonsense codons or termination codons*). One codon, AUG serves as **initiation or start codon** as it starts the synthesis of polypeptide chain. This codon also codes for amino acid methionine. In eukaryotes, the starting amino acid is methionine, while in prokaryotes it is N-formyl methionine.

It is proved that a sequence of three nucleotides in the mRNA codes for amino acid (**a triplet code** or codon), and the code is **non-overlapping** and **commaless** (Fig.). A commaless code means that all the bases in a polynucleotide are parts of codons and that no base serves as a punctuation mark.

The genetic code is said to be **degenerate** because nearly all amino acids are specified by at least two codons. Some (serine, arginine, leucine) are encoded by six different codons. Further, for a set of codons encoding the same amino acid, the first two letters in the figure are the same, with only the third being different (called the **wobble hypothesis**).

		Second base				
		U	C	A	G	
U	UUU	UCU	UAU	UGU	U	
	UUC	UCC	UAC	UGC	C	
	UUA	UCA	UAA	UGA	A	
	UUG	UCG	UAG	UGG	G	
C	CUU	CCU	CAU	CGU	U	
	CUC	CCC	CAC	CGC	C	
	CUA	CCA	CAA	CGA	A	
	CUG	CCG	CAG	CGG	G	
A	AUU	ACU	AAU	AGU	U	
	AUC	ACC	AAC	AGC	C	
	AUA	ACA	AAA	AGA	A	
	AUG	ACG	AAG	AGG	G	
G	GUU	GCU	GAU	GGU	U	
	GUC	GCC	GAC	GGC	C	
	GUA	GCA	GAA	GGA	A	
	GUG	GCG	GAG	GGG	G	

Fig. The genetic code

1. Given the following sense strand of DNA sequence, transcribe it into mRNA, showing the orientation of the mRNA [i.e. 3' and 5' ends]. Then translate this sequence into protein [indicating amino and carboxy termini, be sure to check for an open reading frame as well.]

5' GGGATCGATGCCCTTAAAGAGTTTACATATTGCTGGAGGCGTTAACCCCGGA 3'

2. You have just sequenced a short segment of DNA. You wish to analyze this DNA sequence to determine whether it could encode a protein.

5' TCAATGTAACGCGCTACCCGGAGCTCTGGGCCCAAATTCATCCACT 3'

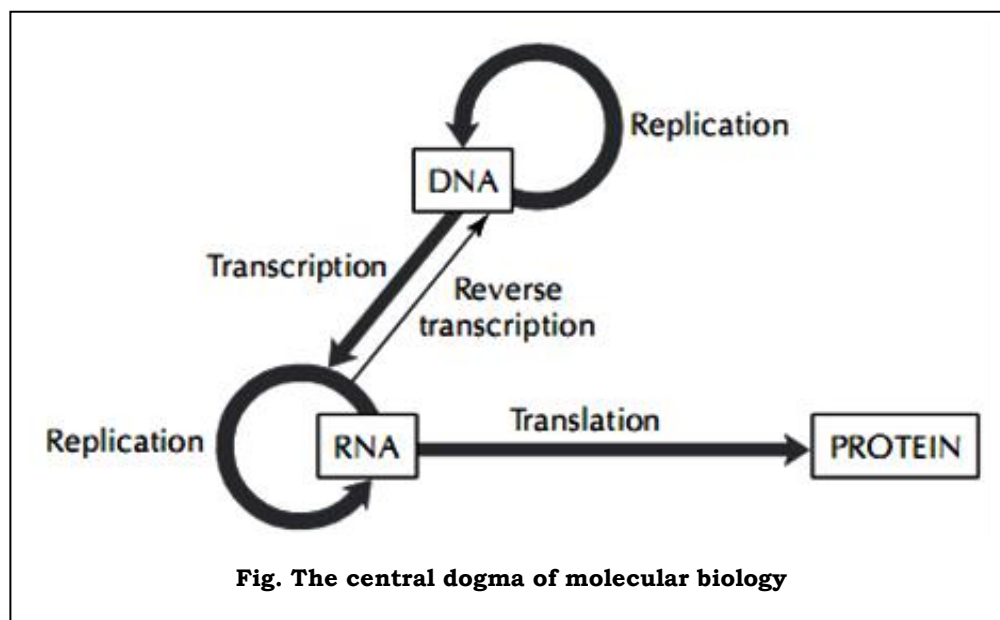
1. Find the longest open reading frame (ORF). Remember, there are six possibilities.
2. Label which strand on the DNA will be the sense strand, and which will be antisense when this DNA is transcribed.
3. Transcribe this ORF into mRNA, indicating the 5' and 3' ends.
4. Translate this mRNA into amino acids, indicating the amino (N) and carboxy (C) termini.

CENTRAL DOGMA OF MOLECULAR BIOLOGY

The flow of genetic information from DNA to RNA by the process of transcription and from RNA to proteins by the process of translation was formalized by Francis Crick in a concept called the **central dogma of molecular biology**.

It was now found that transfer of information from RNA to DNA, (via *reverse transcription*) and the transfer of information from RNA to RNA (by *RNA replication*) also occurs.

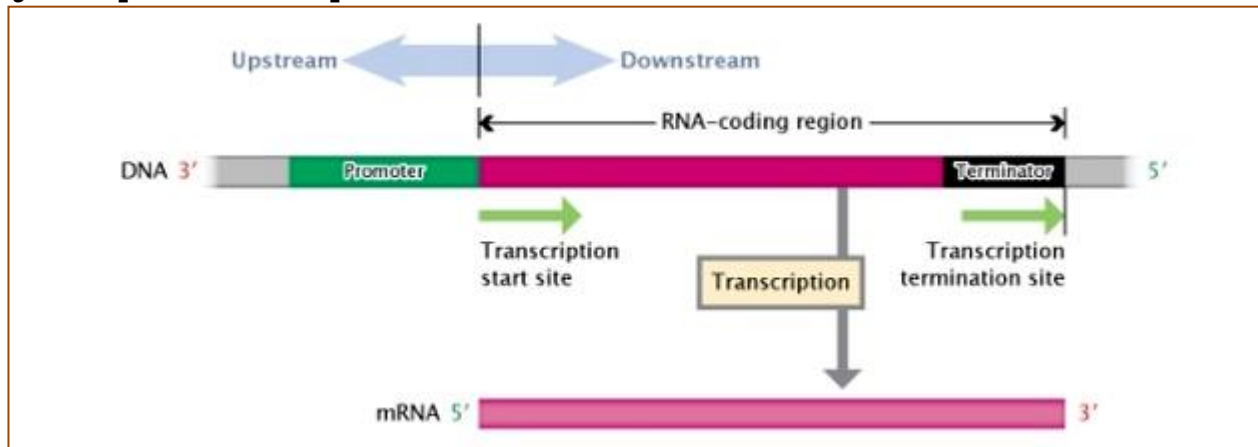
This flow of information can also be reversed. Thus, once a protein is known, the nucleotide sequence in the prescribing DNA strand can be determined and synthesized. The product is called a complementary DNA or cDNA.



TRANSCRIPTION (mRNA SYNTHESIS)

The process of copying the genetic information of the DNA to mRNA (messenger RNA) is called as transcription. The DNA strand which is transcribed is called as the *template strand* or *antisense strand* and the other strand is called as *sense strand*. The mRNA thus produced is complementary to the template stand and identical to the sense strand.

Major Steps in Transcription

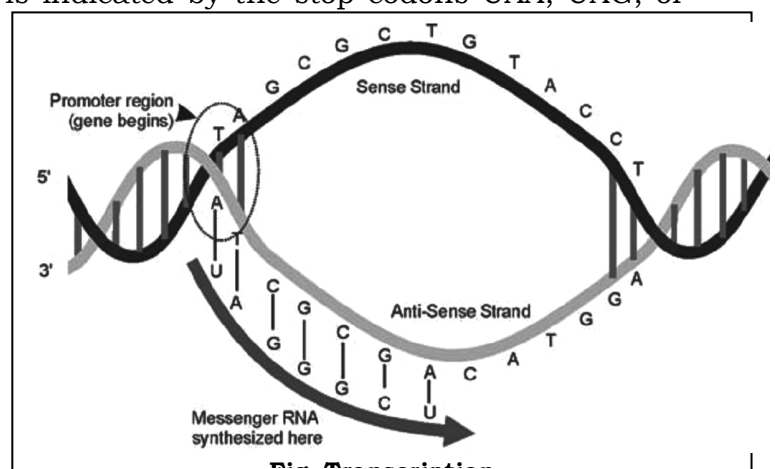


A DNA transcription unit.

A DNA transcription unit is composed, from its 3' to 5' end, of an RNA-coding region (pink rectangle) flanked by a promoter region (green rectangle) and a terminator region (black rectangle). Regions to the left, or moving towards the 3' end, of the transcription start site are considered "upstream;" regions to the right, or moving towards the 5' end, of the transcription start site are considered "downstream."

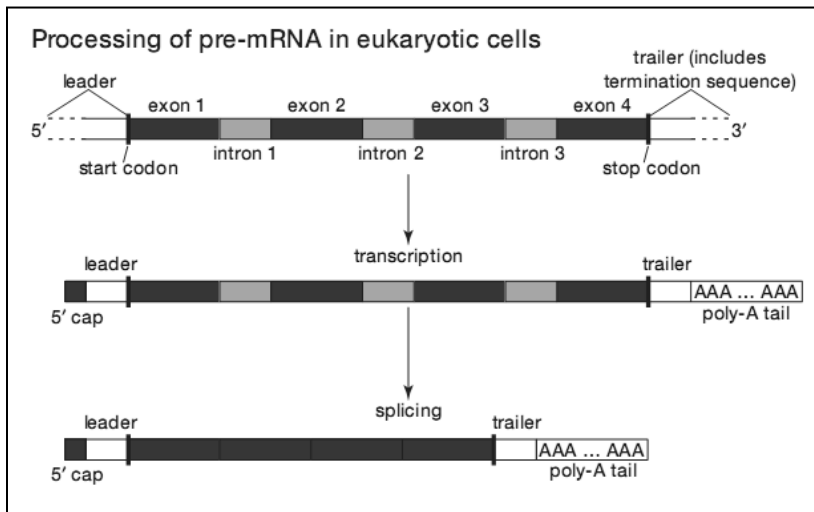
Source : Nature Education 2014. Adapted from Pierce, Benjamin. *Genetics: A Conceptual Approach*, 2nd ed.

- (i) The process of transcription is carried out by enzymes called *RNA polymerases* which unwinds a portion of DNA near a special DNA sequence called *promoter* region.
- (ii) Transcription always starts with the start codon AUG. As it codes for the amino acid methionine, all proteins begin with methionine.
- (iii) The end of the polypeptide sequence is indicated by the stop codons UAA, UAG, or UGA.
- (iv) The transcription always proceeds from 5' to 3' direction.
- (v) The newly synthesized mRNA remaining inside the nucleus is called *hn-RNA* (*heterogeneous nuclear RNA*).



(vi) This hnRNA is processed addition of methyl guanosine cap at the 5' end and by addition of several adenines (poly -A tail) at 3' end. This capping and addition of poly A-tail protects the mRNA from degradation by enzymes.

(vii) This molecule undergoes severe alterations to remove *non-coding parts called introns*, leaving only the *coding parts or exons* to produce the mRNA. This mRNA is about 25% of the original length.



(viii) The mRNA is then transported from the nucleus to the ribosomes in the cytoplasm for protein synthesis.

Problem 1.

Below is the double-stranded DNA sequence of part of a hypothetical yeast genome, which happens to contain a very small gene. Transcription starts at the Transcription Start Site (TSS) after the promoter (shown in yellow), and proceeds in the direction of the arrow. Transcription stops at the end of the Transcription Terminator (shown in blue).



- Which strand of DNA shown, the top or the bottom, is the template strand?
- What is the sequence of the mRNA produced from this gene? Label the 5' and 3' ends.
- What is the sequence of the protein produced from the mRNA in (b)? Label the N and C termini.
- If a mutation were found where a T/A (top/bottom) base pair were added immediately after the T/A base pair shown in bold, what would be the sequence of the mRNA? What would be the sequence of the protein?

2. Three types of RNA involved in comprising the structural and functional core for protein synthesis, serving as a template for translation, and transporting amino acid, respectively, are:

A [mRNA, tRNA, rRNA](#)

B [rRNA, tRNA, mRNA](#)

C [tRNA, mRNA, rRNA](#)

D [tRNA, rRNA, mRNA](#)

E [rRNA, mRNA, tRNA](#)

TRANSLATION (PROTEIN SYNTHESIS)

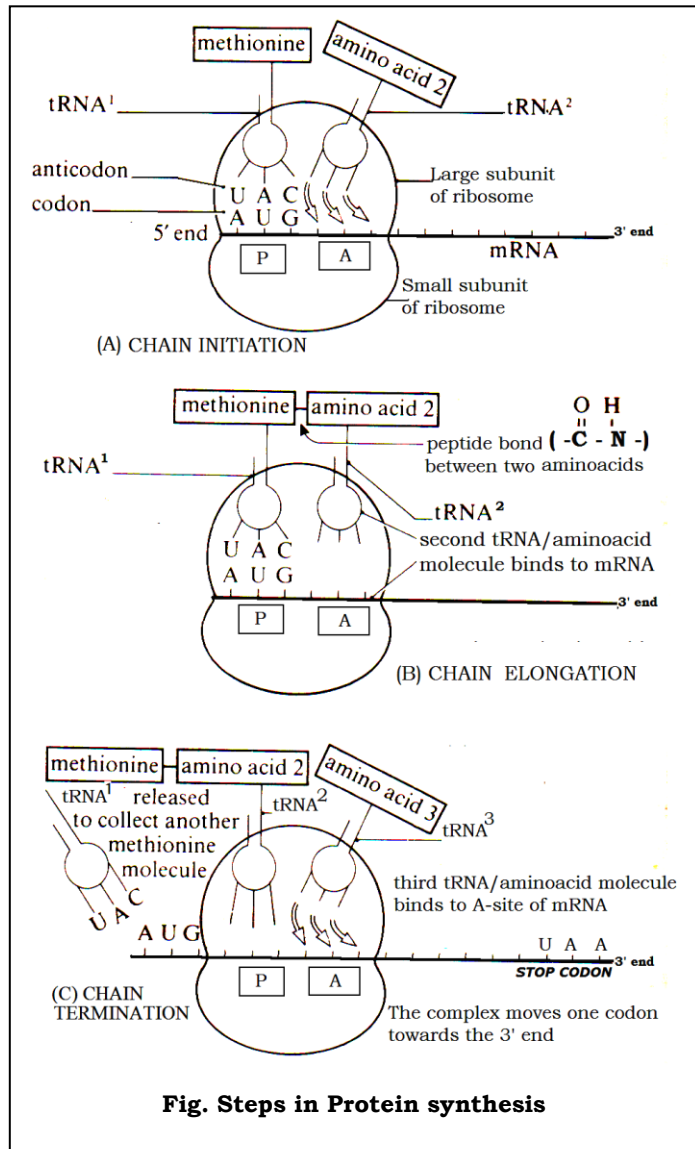
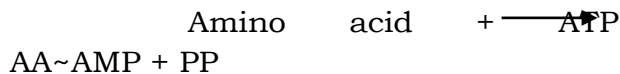
The process by which the nucleotides sequence present in the mRNA is translated into amino acid sequence is called as translation.

Ribosomes are the sites of protein synthesis. Translation requires mRNA, rRNA, ribosomes, 20 kinds of amino acids and their specific tRNAs and many translation factors. The process of translation (protein synthesis) consists of five major steps viz.,

- (1) Activation of amino acids
- (2) Transfer of activated amino acids to tRNA
- (3) Chain initiation
- (4) Chain elongation and
- (5) Chain termination.

1. Activation of amino acid

The mRNA moving out of the nucleus binds to ribosome in the cytoplasm. Ribosomes are the site for protein synthesis. The amino acids in the cytoplasm, are activated in the presence of ATP. Then the enzyme amino acyl synthetase links with amino acid ~ AMP to form amino acyl adenylate enzyme complex. This complex is called *activated amino acid*.



2. Transfer of activated aminoacid to tRNA

When an activated aminoacid collides with a specific tRNA, it binds with the A site (aminoacid attachment site) of tRNA, forming aminoacyl tRNA. The aminoacyl tRNA then moves towards the ribosome.

3. Chain initiation

Translation begins when an aminoacylated tRNA base pairs with the start codon (AUG) present in the mRNA located in the smaller subunit of ribosome. The initiation codon and the small sub unit forms the *initiation complex* to which the larger subunit joins.

4. Chain elongation

The ribosome has two distinct sites namely, A-site (acceptor or aminoacyl attachment site) and P-site (peptidyl site).

Each new aminoacyl-tRNA enters the ribosome and attaches to A-site. The mRNA codon of A-site determines which charged tRNA with aminoacid will attach next. As soon as the next tRNA attaches at the A-site, a peptide bond is formed between the aminoacid (-COOH) on the A-site and the polypeptide (-NH₂) on the P-site. The peptide bond formation is catalysed by the enzyme peptidyl transferase.

After the formation of peptide bond, the tRNA from P-site is released to the cytosol and the polypeptide chain is transferred to tRNA on A-site. Then, the tRNA on A-site is shifted to P-site, making A-site available for new tRNA. Then the ribosome complex moves one codon towards the 3' end on the mRNA, releasing the first tRNA from initiation point to pick another methionine. The free initiation point can now form a new initiation complex.

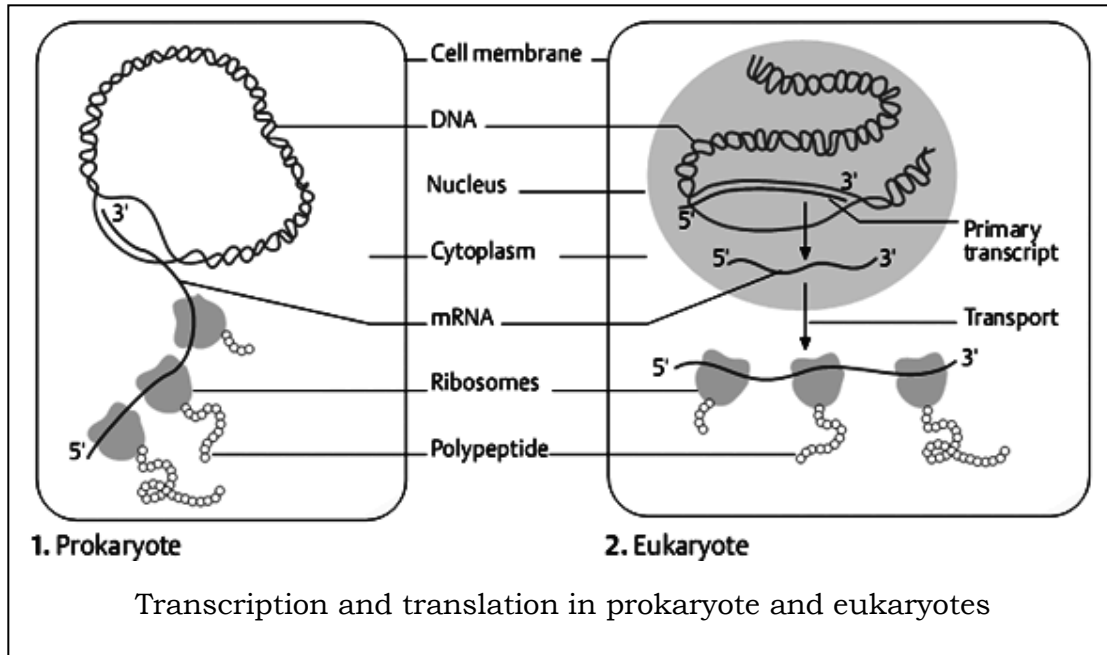
During protein synthesis a number of ribosomes are attached to a single mRNA molecule, each forming a different polypeptide chain. The complex thus formed is known as polyribosome. The process is repeated until the whole mRNA is translated, and adjacent amino acids are linked by peptide bonds.

5. Chain termination

The translation will proceed until a releasing factor binds to the stop codon (UAA, UAG, UGA) and terminates translation, as they does not code for any aminoacid. The interval between the start and stop codons is called the open reading frame (ORF).

The ribosome releases the polypeptide and mRNA and subsequently dissociates into two subunits. Further processing of polypeptide chain into proteins and enzymes is done in the cytoplasm itself depending on the bonding properties of the aminoacids in it. Most of the mRNA molecules are unstable and degraded after the release of polypeptide chain, but some mRNAs such as those coding for hemoglobin may be stable.

Differences between transcription in prokaryotes and eukaryotes



Prokaryotes have only exons and no introns, so the translation can begin even before transcription of the mRNA is complete. But, this is not possible in eukaryotes as the mRNA must be processed to remove the introns and the mRNA must leave the nucleus.

INTRON refers to the non-coding region of the eukaryotic genes that are transcribed into mRNA. They are removed by splicing of RNA.

EXON refers to the region of DNA that codes for a protein. In eukaryotes, the exons are separated by many introns.

1. A messenger acid is 336 nucleotides long, including the initiator and termination codons. The number of amino acids in the protein translated from this mRNA is:

A [999](#)

B [630](#)

C [330](#)

D [111](#)

E 110

2. A synthetic mRNA of repeating sequence 5'-CACACACACACACAC... is used for a cell-free protein synthesizing system like the one used by Nirenberg. If we assume that protein synthesis can begin without the need for an initiator codon, what product or products would you expect to occur after protein synthesis?

- A. one protein, consisting of a single amino acid
- B. three proteins, each consisting of a different, single amino acid
- C. two proteins, each with an alternating sequence of two different amino acids
- D. one protein, with an alternating sequence of three different amino acids
- E. one protein, with an alternating sequence of two different amino acids

3. Under conditions where methionine must be the first amino acid, what protein would be coded for by the following mRNA?

5'-CCUCAUAUGCGCCAUAUAAGUGACACACA-3'

- A. pro his met arg his tyr lys cys his thr
- B. met arg his tyr lys cys his thr
- C. met arg his tyr lys
- D. met pro his met arg his tyr lys cys his thr
- E. arg his ser glu tyr tyr arg leu tyr ser

4. Which mRNA codes for the following polypeptide?

met arg ser leu glu

		Second Base							
		U		C		A		G	
U	UUU	phe	UCU	ser	UAU	tyr	UGU	cys	U
	UUC	phe	UCC	ser	UAC	tyr	UGC	cys	C
	UUA	leu	UCA	ser	UAA	STOP	UGA	STOP	A
	UUG	leu	UCG	ser	UAG	STOP	UGG	trp	G
C	CUU	leu	CCU	pro	CAU	his	CGU	arg	U
	CUC	leu	CCC	pro	CAC	his	CGC	arg	C
	CUA	leu	CCA	pro	CAA	gln	CGA	arg	A
	CUG	leu	CCG	pro	CAG	gln	CGG	arg	G
A	AUU	ile	ACU	thr	AUU	asn	AGU	ser	U
	AUC	ile	ACC	thr	AAC	asn	AGC	ser	C
	AUA	ile	ACA	thr	AAA	lys	AGA	arg	A
	AUG	met	ACG	thr	AAG	lys	AGG	arg	G
G	GUU	val	GCU	ala	GAU	asp	GGU	gly	U
	GUC	val	GCC	ala	GAC	asp	GGC	gly	C
	GUA	val	GCA	ala	GAA	glu	GGA	gly	A
	GUG	val	GCG	ala	GAG	glu	GGG	gly	G

- A. 3' -AUGCGUAGCUUGGAGUGA-5'
- B. 3' -AGUGAGGUUCGAUGCGUA-5'
- C. 5' -AUGCGUAGCUUGGAGUGG-3'
- D. 1' -AUGCGUAGCUUGGAGUGA-3'
- E. 3' -AUGCGUAGCUUGGAGUGA-1'

5.

REGULATION OF GENE EXPRESSION IN PROKARYOTES

Regulation of gene expression refers to the turning on and turning off of the genes at appropriate time and to the desired level. It is important for proper growth and development. The gene expression may be regulated during transcription, mRNA processing, translation, and involve posttranslational modifications also.

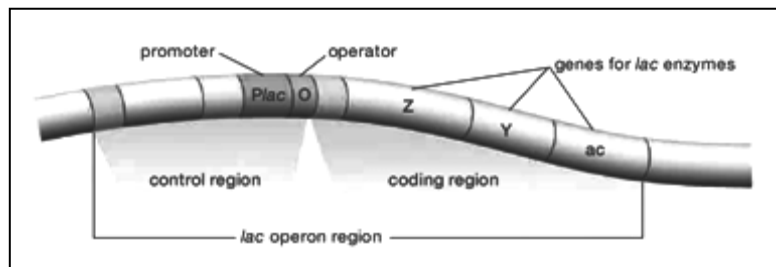
An Operon is a group of structural genes whose transcription is regulated by the action of a regulator gene 'r', a promoter gene 'p' and an operator gene 'o'.

REGULATION OF OPERON

There are two basic categories of gene regulation: *negative* and *positive*.

1. In negative regulation, an inhibitor which is bound to the DNA must be removed for transcription to occur. In this case the transcription is normally on and must be turned off. It is called as repressible operon.

2. In positive regulation, an activator has to bind to the DNA for transcription to occur. In this case, the transcription is normally off and must be turned on. It is called as inducible operon.



THE LAC OPERON OF *E. COLI*

The Operon concept was proposed by Jacob and Monod in 1961. The lac operon *E. coli* bacterium consists of three structural genes of namely,

- a). *lacZ* which codes for beta galactosidase,
- b). *lacY* which codes for the enzyme permease that allows the diffusion of lactose into bacterial cell and
- c). *lacA* which codes for the enzyme transacetylase.

Mechanism of lac operon when lactose is absent

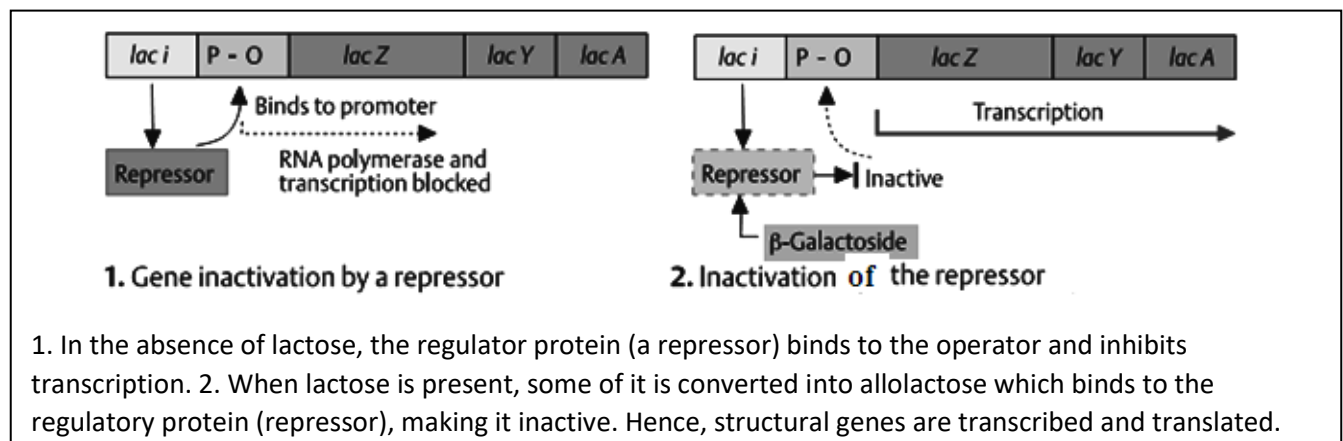
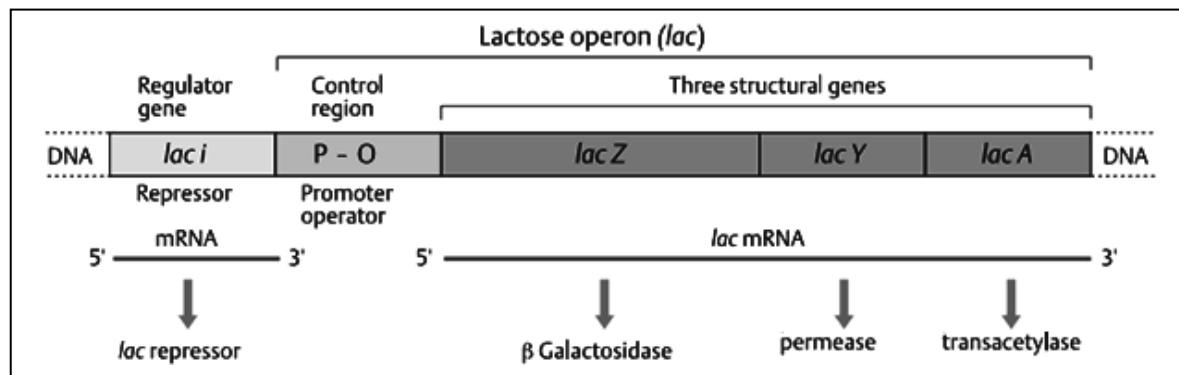
When *E. coli* is cultured in a medium devoid of lactose, the enzymes for lactose catabolism are not required and hence the lac-operon has to be switched off or repressed.

Under such conditions, the regulator gene produces a repressor protein which binds to the operator site. Therefore, RNA polymerase cannot bind to the promoter and transcription is stopped.

Mechanism of *lac* operon when lactose is present

When lactose is present in the medium, some of it is converted into allolactose which binds to the repressor, making it inactive.

Therefore, RNA polymerase binds to the promoter and hence, the transcription of *lacZ*, *lacY* and *lacA* takes place and *lac* enzymes are produced for the catabolism of lactose.



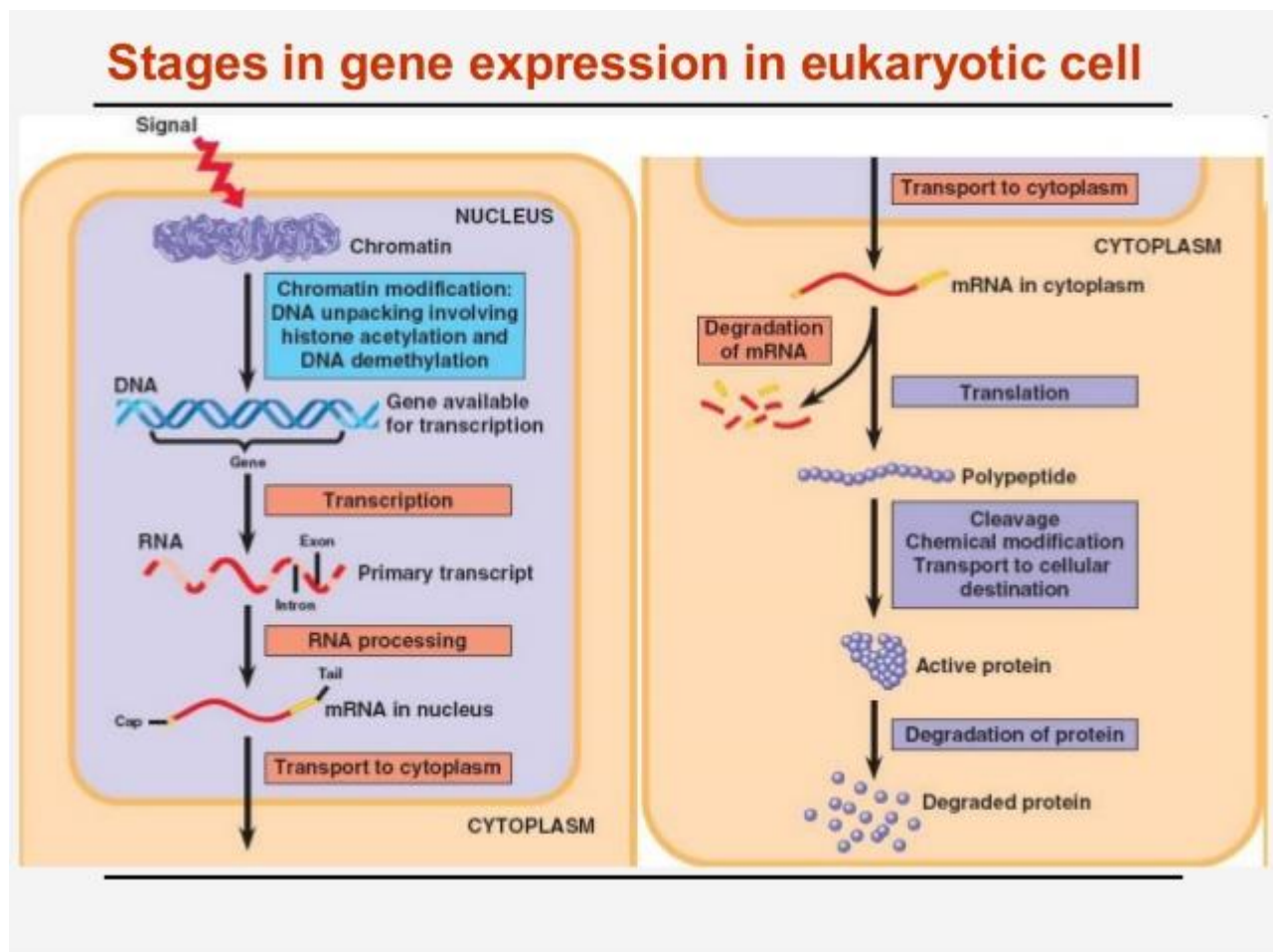
REGULATION OF GENE EXPRESSION IN EUKARYOTES

There are three major methods by which eukaryotic cells are known to regulate translation by:

- a) Altering the life of the mRNA,
- b) Controlling the initiation of translation, and
- c) Changing the overall rate of translation.

A typical eucaryotic mature mRNA consists of four major regions: A) a 5' noncoding region (leader), B) a coding region, C) a 3' noncoding region (trailer), and D) a poly-A tail. Each of the four segments may affect the half-life of mRNA molecules.

Eukaryotic genes contain introns (noncoding regions) interspersed among the coding regions (exons). Part of the process that converts primary transcripts to complete mRNA molecules involves removal of the introns and splicing the exons together. Variations in the excision and splicing jobs can lead to different mRNAs and, following translation, to different protein products.



Eukaryotic gene regulation is less well understood than bacterial regulation, partly owing to the larger genomes in eukaryotes, their greater sequence complexity, and the difficulty of isolating and manipulating mutations that can be used in the study of gene regulation. Nevertheless, great advances in our understanding of the regulation of eukaryotic genes have been made in recent years, and eukaryotic regulation continues to be one of the cutting-edge areas of research in genetics.

Comparison between Eukaryotic and Prokaryotic gene regulation

Many features of gene regulation are common to both bacterial and eukaryotic cells. For example, in both types of cells, DNA-binding proteins influence the ability of RNA polymerase to initiate transcription. However, there are also some differences.

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

Second, chromatin structure affects gene expression in eukaryotic cells; DNA must unwind from the histone proteins before transcription can take place.

Third, although both repressors and activators function in eukaryotic and bacterial gene regulation, activators seem to be more common in eukaryotic cells.

Finally, the regulation of gene expression in eukaryotic cells is characterized by a greater diversity of mechanisms that act at different points in the transfer of information from DNA to protein.

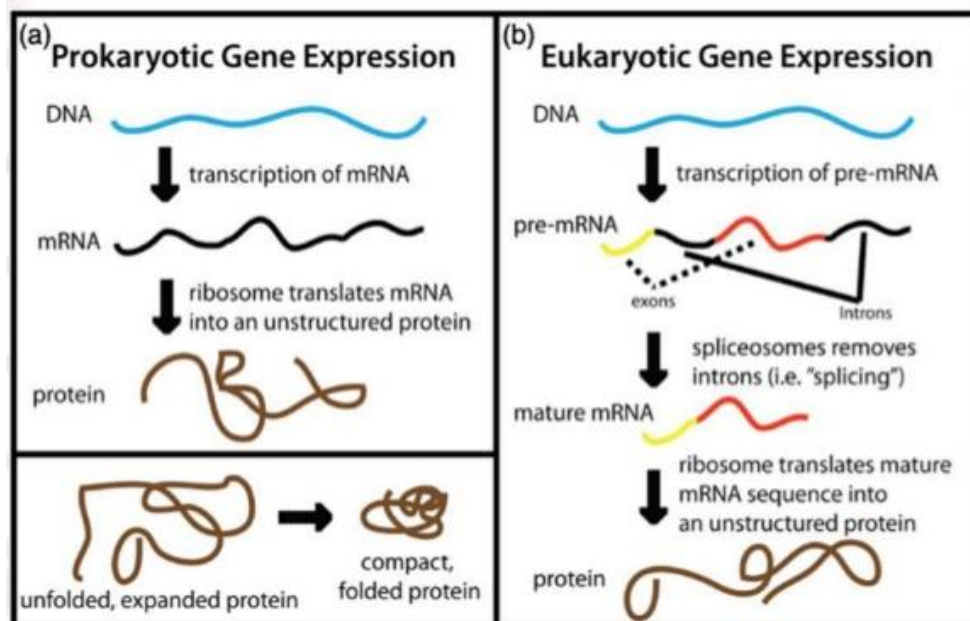


TABLE 26.1. Comparison of gene regulation model of prokaryotes and eukaryotes

<i>S.No. Particulars</i>	<i>Prokaryotes</i>	<i>Eukaryotes</i>
1. Gene regulation model developed by	Jacob and Monod	Britten and Davidson
2. Model known as	Operon model	Battery model
3. Genes involved	Structural, operator, promoter and regulator	Sensor, integrator, receptor and producer
4. Enzymes involved	beta-galactosidase, galactosidase permease, transacetylase and RNA polymerase.	mRNA polymerase
5. Gene regulation	Simple	More complex
6. Perception of mechanism	Better understood	Lesser understood
7. Basis of model	Based on empirical results	It is a theoretical model without proof
8. mRNA	Unstable Polycistronic	Stable Monocistronic
9. Matching gene	Structural genes, Operator gene Promoter gene Regulator gene	Producer gene Receptor gene Sensor gene Integrator gene
10. Popularity	Widely accepted	Widely accepted

POLYMERASE CHAIN REACTION

The polymerase chain reaction (PCR) is an enzymatic, *in vitro* method for rapid amplification or multiplication of specific DNA. It was developed in 1984 by Kary Mullis (Nobel prize, 1993). It is now a basic tool for molecular biology works. In this process, DNA is heated to separate the two strands, short primers attach to the target DNA, and DNA polymerase synthesizes new DNA strands from the primers. Each cycle of PCR doubles the amount of DNA.

Meet Kary Mullis, inventor of the PCR

“Sometimes a good idea comes to you when you are not looking for it. Through an improbable combination of coincidences, naïveté, and luck mistakes, such a revelation came to me one Friday night in April, 1983, as I gripped the steering wheel of my car and snaked along a moonlit mountain road into northern California’s redwood country.” *Sci. Am.* 1990 262:56-61, 64-5.



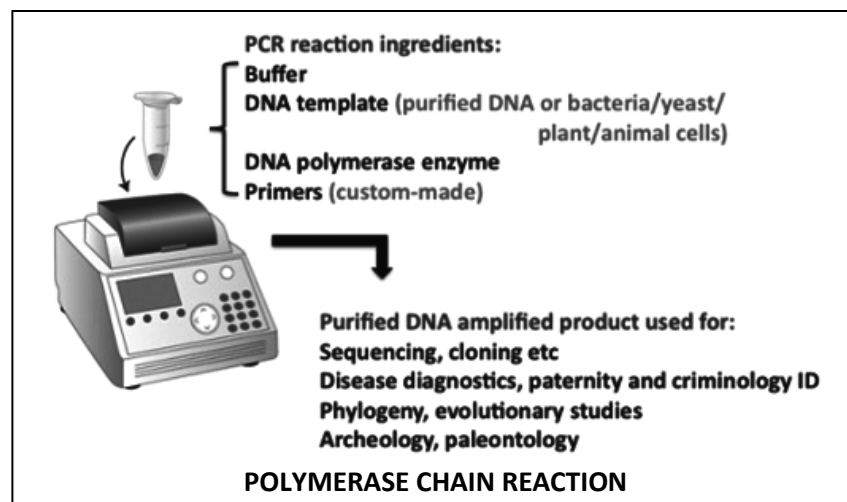
Requirements for PCR

1. A target DNA (100-35,000bp)
2. Primers complementary to flanking region of target DNA
3. dNTPs (Deoxyribonucleotides dATP, dCTP, dGTP, dTTP)
4. Taq polymerase from *Thermus aquaticus* which is stable at 95°C

PCR consists of three steps:

(i) Denaturation: The reaction mixture is heated to 95°C for a short time period (about 15-30 sec) to denature the target DNA into single strands that can act as templates for DNA synthesis.

(ii) Primer annealing / renaturation: The mixture is rapidly cooled to 55 to 65°C which allows the two primers to bind to the sequences on each of the two strands flanking the target DNA. This process is called annealing. One primer binds to each strand. The two parental strands do not reanneal with each other because the primers are in large excess over parental DNA.



(iii) Elongation: The temperature of the mixture is raised to 72° c (usually) and kept at this temperature for a pre-set period of time to allow DNA polymerase to elongate each primer by adding the dNTPs added in the mixture. Thus at the end of this incubation, both single-stranded template strands have been made partially double stranded. The new strand of each double-stranded DNA extends for a variable distance downstream.

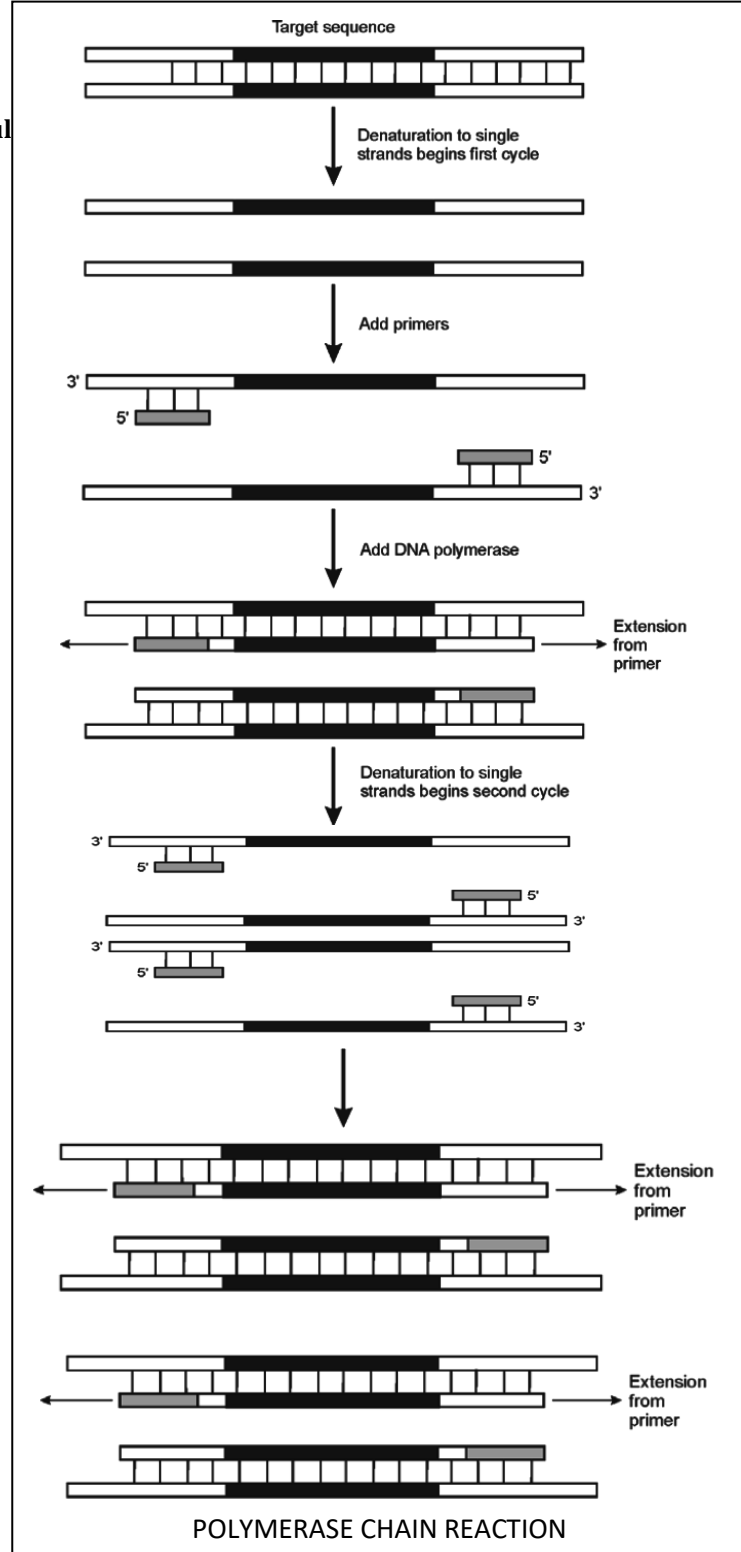
The three steps of the PCR cycle are repeated. Thus in the second cycle, the four strands denature, bind primers and are extended. No other reactants need to be added. The three steps are repeated again and again. By third cycle, the DNA sequences between the two primer sites alone are extended. After 20 cycles, the original DNA has been amplified a million-fold and this rises to a billion-fold (1000) million after 30 cycles. Automated thermocyclers are now routinely used to get a billion-fold amplification of the target DNA sequence (30 cycles) in less than one hour.

As the three steps—denaturation, primer annealing and primer extension—are carried out repetitively just by changing temperature of the reaction mixture PCR is also called as a thermocycler.

Applications of PCR

PCR already has very widespread applications, and new uses are being devised on a regular basis. Some of the applications of PCR are as follows:

- i. PCR can amplify a single DNA molecule from a complex mixture, largely avoiding the need to use DNA cloning to prepare that molecule. Also, it is now possible to amplify a specific single RNA molecule from a complex mixture.
- ii. PCR is used for DNA sequencing.
- iii. By using suitable primers, it is possible to use PCR to create point mutations, deletions and insertions of target DNA which greatly facilitates the analysis of gene expression and function.

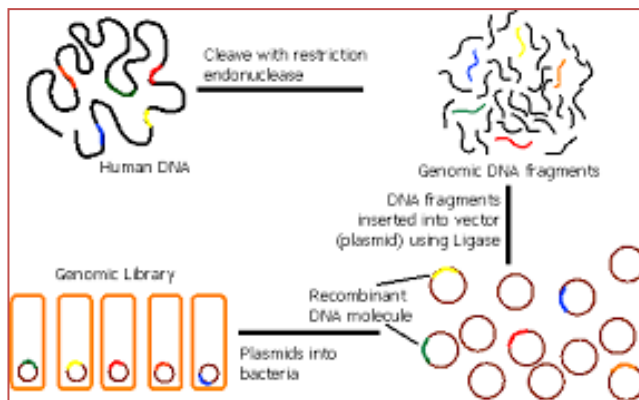


- iv. Using appropriate primers, very small amounts of specified bacteria and viruses can be detected in tissues, making PCR invaluable for medical diagnosis.
- v. PCR is now invaluable for characterizing medically important DNA samples. For example, in screening for human genetic diseases, it is rapidly replacing the use of RFLPs.
- vi. Because of its extreme sensitivity, PCR is now fundamentally important to forensic medicine. It is even possible to use PCR to amplify the DNA from a single human hair or a microscopic drop of blood left at the scene of a crime to allow detailed characterization.

DNA LIBRARY

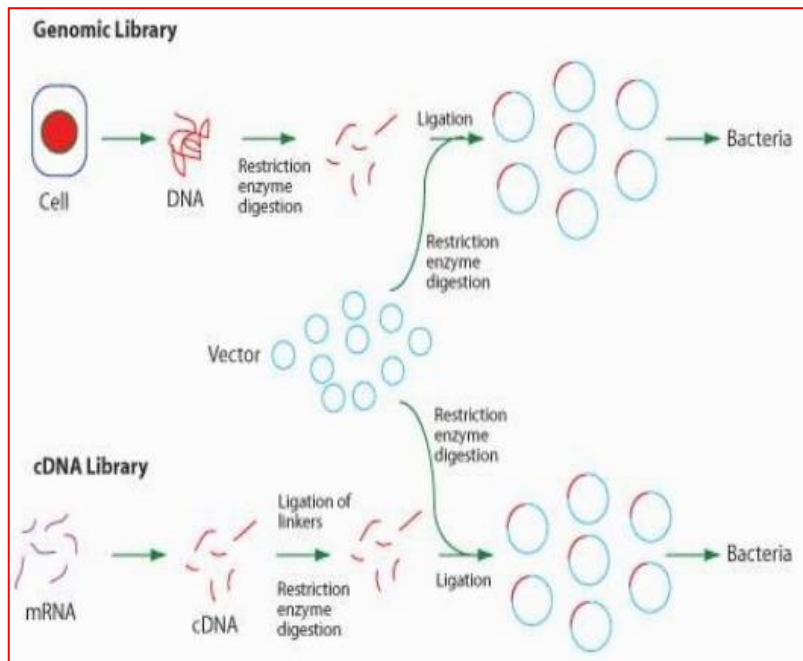
One method of finding a gene is to create and screen a DNA library. A genomic library is created by cutting genomic DNA into overlapping fragments and cloning each fragment in a separate bacterial cell. A cDNA library is created from mRNA that is converted into cDNA and cloned in bacteria.

GENOMIC LIBRARY :



It refers to a large collection of bacterial cells, each containing a random piece of human genomic DNA. For constructing a gene library, the entire DNA of cell is cleaved into small pieces by using different REs. All these cut fragments are then introduced into appropriate vectors. This forms a large collection of different recombinant clones, which are then introduced into the host bacterial cells to form the gene library. All the genes of an organism are represented in the gene library. In order to produce a complete gene library for *E. coli*, about 1500 fragments are required, whereas about 10 lakhs fragments are required for human gene library. Genomic libraries contain all the DNA of an organism, and cDNA libraries contain only expressed DNA.

cDNA Library:



In a genomic library, all the genomic material is represented but only about 3% of the cloned DNA codes for proteins. If a collection of only expressed DNAs were to be made, then RNA and not DNA would be the starting point. cDNA library is a collection of all the expressed DNA of a particular cell type or tissue. For example, a cDNA from pancreatic β -cell contains clones with cDNA for proinsulin. On the other hand, a cDNA library from bone marrow cell contains many clones with cDNA for α and β chains of haemoglobin. Thus, for a cDNA

library the tissue of origin is important. For a genomic library, the tissue of origin is unimportant because the genomic material is same in all cell types of an organism. Building of cDNA library: The mRNA is extracted from a specific tissue. It is used as a template for synthesis of complementary DNA strand; the enzyme catalyzing this synthesis is reverse transcriptase, which yields a single stranded cDNA. Double-stranded cDNA is then obtained from it by adding the DNA polymerase. The latter is incorporated in plasmid, λ phage or cosmid and introduced into host bacterial cell.

SCREENING OF GENOMIC LIBRARY:

It is possible to screen the genomic library and obtain a gene of interest from it. By the use of molecular markers. Then the identified DNA fragments can be inserted into cloning vectors. Cloning vectors must have an origin of replication, one or more unique restriction sites, and selectable markers. Plasmids are commonly used as cloning vectors.

SCREENING DNA LIBRARY TO IDENTIFY CLONED GENE BY COLONY HYBRIDIZATION

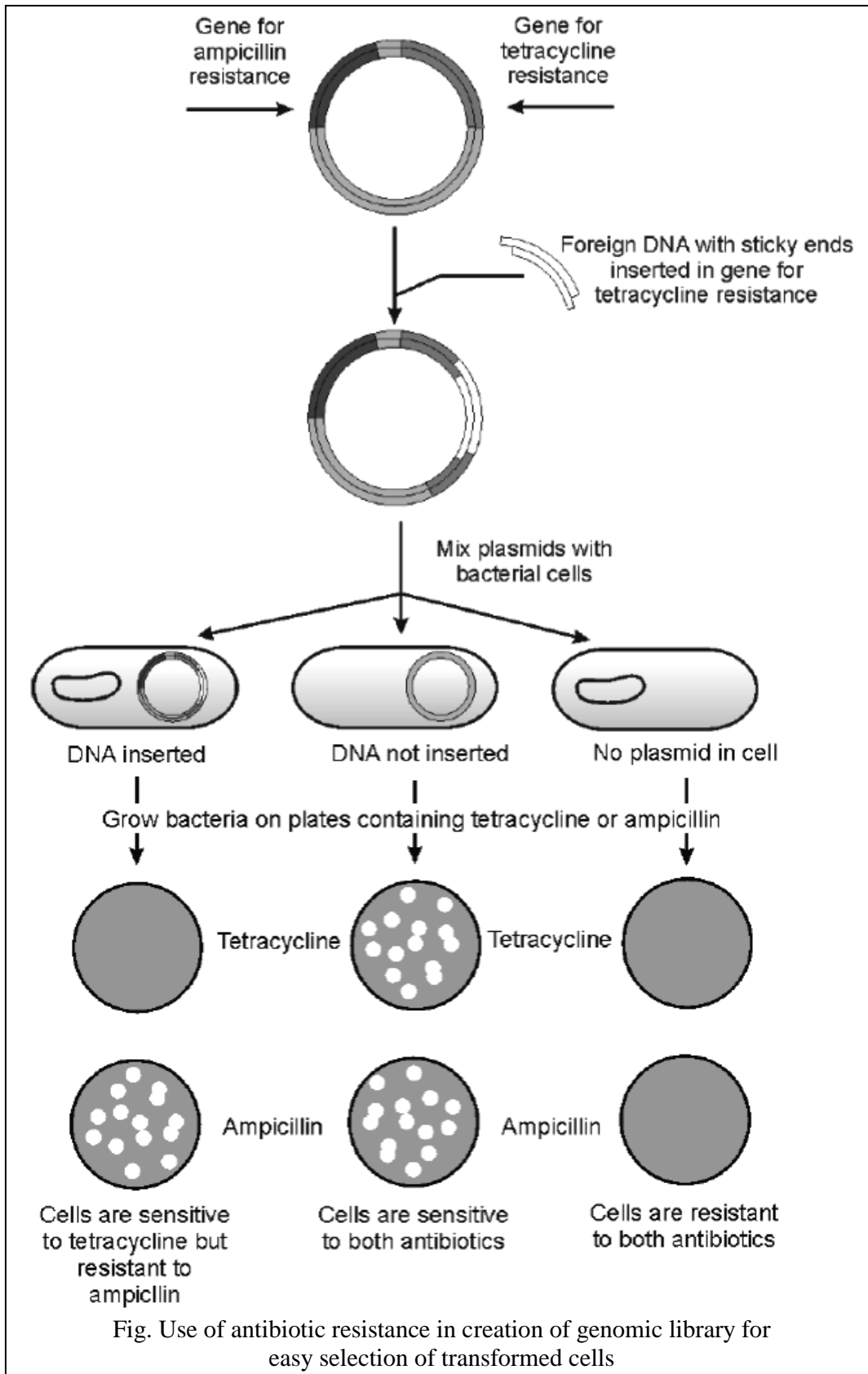
Libraries, once created, can be stored as permanent reference resources (which is why they're called libraries). Libraries are often used in the effort to isolate specific genes. There are two ways to find the specific clone within the library that contains a DNA fragment that includes a gene of interest. One is to look for a specific sequence of base pairs (if known). The other is to look for the specific protein encoded by the gene. In either case, the process is called *screening*.

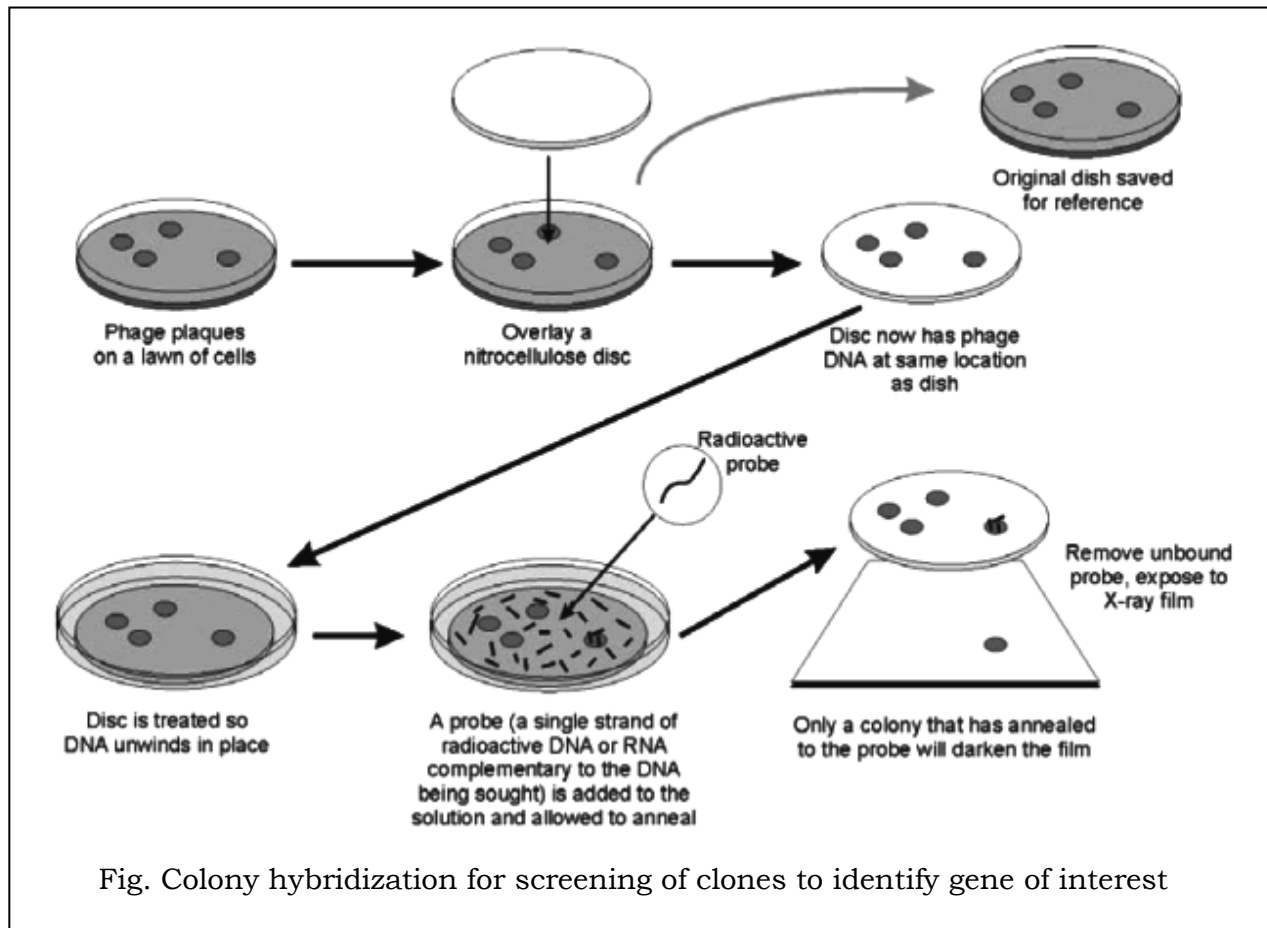
There are many techniques, depending on the vector used and the gene being sought. Bacterial colonies or phage plaques are transferred from an agar-filled Petri dish to a solid disk, usually made of nitrocellulose, by the simple method of laying the disk on the agar, then lifting it up again. This doesn't remove the colonies from the agar, but it does transfer enough material to the nitrocellulose for analysis.

Next, the DNA on the nitrocellulose disk is unwound by immersing the disk in a chemical solution. Then the disk is placed in a solution that contains a single strand of DNA or RNA tagged with a radioactive atom complementary to one of the strands in the clone. This tagged single strand is called a *probe*. Wherever the probe joins up with one of the recombinant strands, it leaves a radioactive spot on the disk which can be registered on photographic film, and directly correlated to a location on the original agar disk.

If the DNA sequence of the gene being sought is unknown, but a protein the gene codes for is known, another screening method is to create clones using vectors cells begin producing the protein the gene encodes, and researchers can then search the library for the protein in question.

Once a desired clone is located, it can be picked off the agar plate with a needle and allowed to multiply freely. The recombinant DNA can be chemically purified from the cells for use in the laboratory, and the clone that produces it can be stored and regrown as needed. This provides genetic engineers with an endless supply of a particular gene for insertion into other organisms.





NUCLEIC ACID BLOTTING TECHNIQUES

Blotting techniques are very widely used analytical tools for the specific identification of desired DNA or RNA fragments from thousands of molecules. Blotting refers to the process of immobilization of sample nucleic acids on solid support (nitrocellulose or nylon membranes). The blotted nucleic acids are then used as targets in the hybridization experiments for their specific detection. An outline of the nucleic acid blotting technique is depicted in fig.

Types of blotting techniques

The most commonly used blotting techniques are listed below

- (i) Southern blotting (for DNA)
- (ii) Northern blotting (for RNA)
- (iii) Dot blotting (DNA/RNA)
- (iv) Western blotting (for proteins)

SOUTHERN BLOTTING

Southern blotting techniques is the first nucleic acid blotting procedure developed in 1975 by Southern. The genomic DNA isolated from cells/tissues is digested with one or more restriction enzymes. This mixture is loaded into a well in an agarose or polyacrylamide gel and then subjected to electrophoresis. DNA, being negatively charged migrates towards the anode (positively charged electrode); smaller DNA fragments move faster.

The separated DNA molecules are denatured by exposure to a mild alkali and transferred to nitrocellulose or nylon paper. This results in an exact replica of the pattern of DNA fragments on the gel. The DNA can be annealed to the paper on exposure to heat (80°C). The nitrocellulose or nylon paper is then exposed to labeled cDNA probes. These probes hybridize with complementary DNA molecules on the paper.

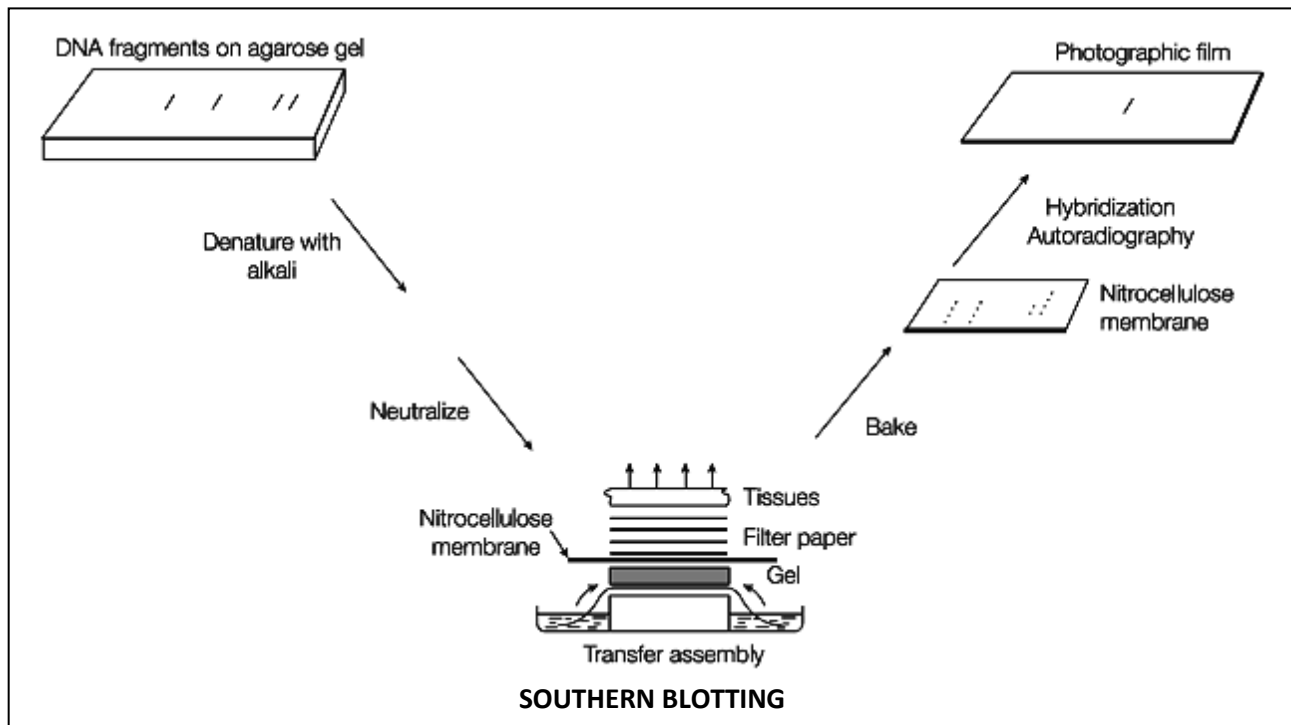
The paper after thorough washing is exposed to x-ray film to develop autoradiograph. This reveals specific bands corresponding to the DNA fragments recognized by cDNA probe. This procedure is called as Southern blotting.

Applications of Southern blotting

Southern blotting technique is extremely specific and sensitive, although it is a simple technique. Some of the applications are listed.

- (i) It is an invaluable method in gene analysis.
- (ii) Important for the confirmation of DNA cloning results.
- (iii) Forensically applied to detect minute quantities of DNA (to identify parenthood, thieves, rapists etc.).

- (iv) Highly useful for the determination of restriction fragment length polymorphism (RFLP) associated with pathological conditions.



NORTHERN BLOTTING

Northern blotting is the technique for the specific identification of RNA molecules. The procedure adopted is almost similar to that described for Southern blotting. RNA molecules are subjected to electrophoresis, followed by blot transfer, hybridization and autoradiography.

RNA molecules do not easily bind to nitrocellulose paper or nylon membranes. Blot-transfer of RNA molecules is carried out by using a chemically reactive paper prepared by diazobenzylxymethyl (DBM) paper. The RNA can covalently bind to DBM paper.

Northern blotting is theoretically, a good technique for determining the number of genes (through mRNA) present on a given DNA. But this is not really practicable since each gene may give rise to two or more RNA transcripts. Another drawback is the presence of exons and introns.

DOT-BLOTTING

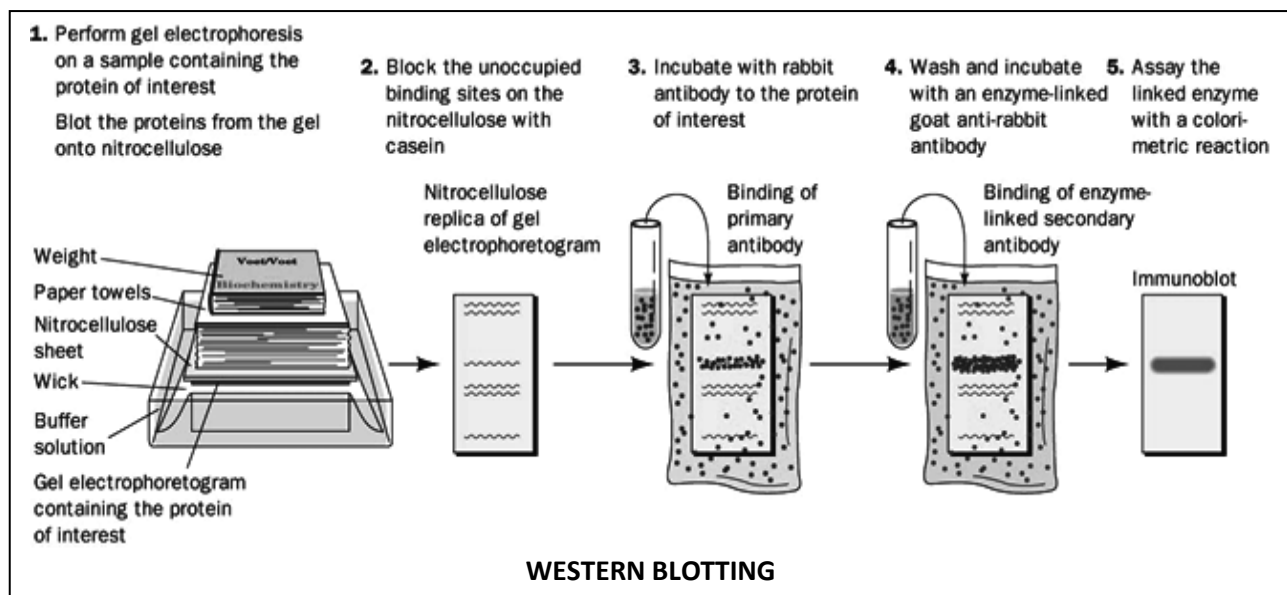
Dot-blotting is a modification of Southern and Northern blotting techniques described above. In this approach, the nucleic acids (DNA or RNA) are directly spotted onto the filters, and not subjected to electrophoresis. The hybridization procedure is the same as in original blotting techniques. Dot-blotting technique is particularly useful in obtaining quantitative data for the evaluation of gene expression.

WESTERN BLOTTING

Western blotting can be used for detection of one or more antigens in a mixture. The sample is electrophoresed on an SDS- polyacrylamide gel (SDS-PAGE) that separates the proteins on the basis of size, resulting in a series of protein bands down the gel. It is very useful to understand the nucleic acid functions, particularly during the course of gene manipulations. The technique of western blotting involves the transfer of electrophoresed protein bands from polyacrylamide gel to nylon or nitrocellulose membrane.

These proteins can be detected by specific protein-ligand interactions. The western blot is then reacted with labeled antibody, unreacted antibody is washed away and those protein bands that have bound the antibody become visible and hence are identified. The method of visualization depends on how the antibody was labeled. If it is labelled by radioactive probe (e.g. ^{125}I), then autoradiography is carried out to detect the radioactive protein bands. Alternatively, the antibody may be detected by incubating the sheet with a second antibody that recognizes the first antibody (e.g. if the first specific antibody was raised in rabbits, the second antibody could be a goat anti-rabbit antibody). The second antibody could be radiolabeled and its binding detected by autoradiography or it could be conjugated to an enzyme that generates a colored product as in ELISA.

Western blotting can also be used to analyze specific antigens after two-dimensional gel electrophoresis which resolves proteins as spots, separated on the basis of both charge and size.

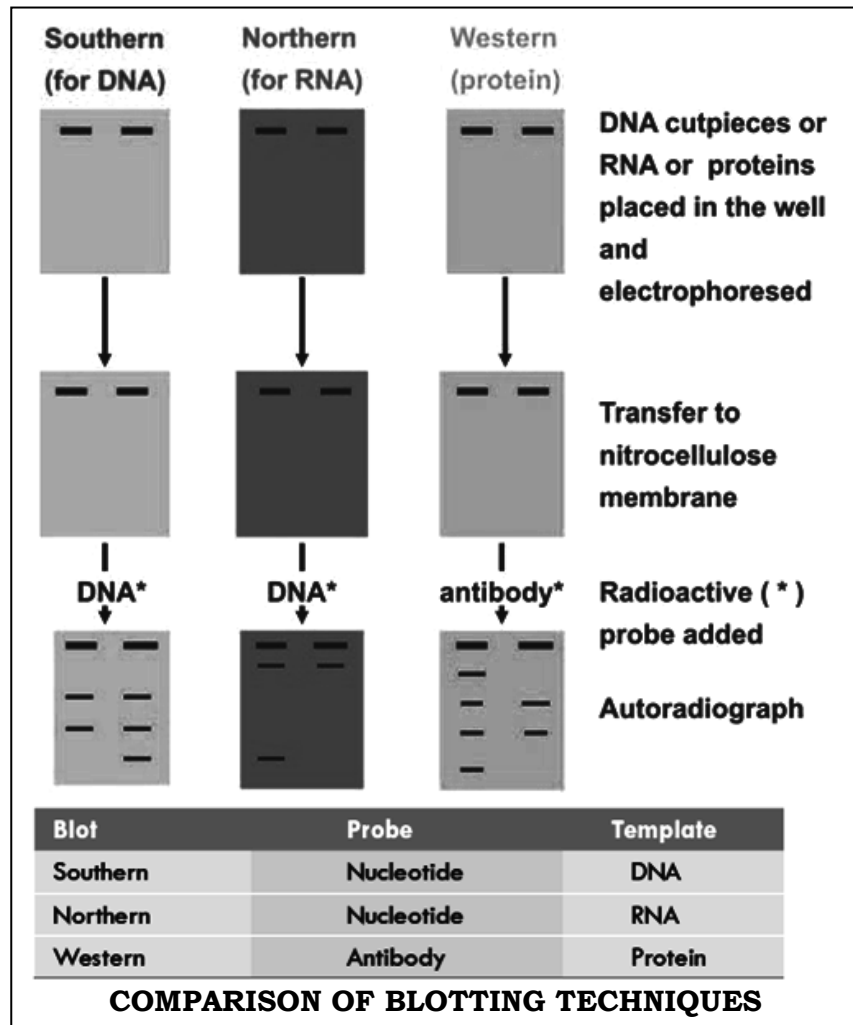


Autoradiography

Autoradiography is the process of localization and recording of a radiolabel within a solid specimen, with the production of an image in a photographic emulsion.

These emulsions are composed of silver halide crystals suspended in gelatin.

When a β -particle or a γ -ray from a radiolabel passes through the emulsions, silver ions are converted to metallic silver atoms. This results in the development of a visible image which can be easily detected.



RESTRICTION FRAGMENT LENGTH POLYMORPHISMS (RFLPs)

The human genome contains hundreds of variations in base sequences that do not affect the phenotype. The property of the molecules to exist in more than one form is known as polymorphism. Difference in mutation and DNA polymorphism: If more than 1% of the population has a particular alteration in the sequence, it is polymorphism. If only a few individuals have it, then it is mutation. Polymorphism is normal variation, and generally having no deleterious effect. Mutation is abnormal, and sometimes will have defective function, e.g., phenylketonuria. A polymorphic gene is one, in which the variant alleles are common in more than 1% of the total population. The existence of two or more types of restriction fragment patterns is called restriction fragment length polymorphism (RFLP). This can be used as a genetic marker. DNA is treated with restriction enzymes, which cleave DNA into fragments of defined lengths. Then electrophoresis is done in agarose gels, when the fragments are separated. Finally, the DNA from the agarose gel is transferred on to nitrocellulose paper (Southern blotting) and hybridized with labelled probe sequences. Genotypic changes can be recognised by the altered restriction fragments.

A RFLP represents a stretch of DNA that serves as a marker for mapping a specified gene. RFLPs are located randomly throughout a person's chromosomes and have no apparent function.

A DNA molecule can be cut into different fragments by a group of enzymes called restriction endonucleases (see table). These fragments are called polymorphisms (literally means many forms).

An outline of RFLP is depicted in fig. The DNA molecule 1 has three restriction sites (R1, R2, R3), and when cleaved by restriction endonucleases forms 4 fragments. Let us now consider DNA 2 with an inherited mutation (or a genetic change) that has altered some base pairs. As a result, the site (R2) for the recognition by restriction endonuclease is lost. This DNA molecule 2 when cut by restriction endonuclease forms only 3 fragments (instead of 4 in DNA 1).

As is evident from the above description, a stretch of DNA exists in fragments of various lengths (polymorphisms), derived by the action of restriction enzymes, hence the name restriction fragment length polymorphisms.

RFLPs in the diagnosis of diseases

If the RFLP lies within or even close to the locus of a gene that causes a particular disease, it is possible to trace the defective gene by the analysis of RFLP in DNA. The person's cellular DNA is isolated and treated with restriction enzymes. The DNA fragments so obtained are separated by electrophoresis. The RFLP patterns of the disease suspected individuals can be compared with that of normal people (preferably with the relatives in the same family). By this approach, it is possible to determine whether the individual has the marker RFLP and the disease gene. With 95% certainty, RFLPs can detect single gene-based diseases.

Methods of RFLP scoring

Two methods are in common use for the detection of RFLPs.

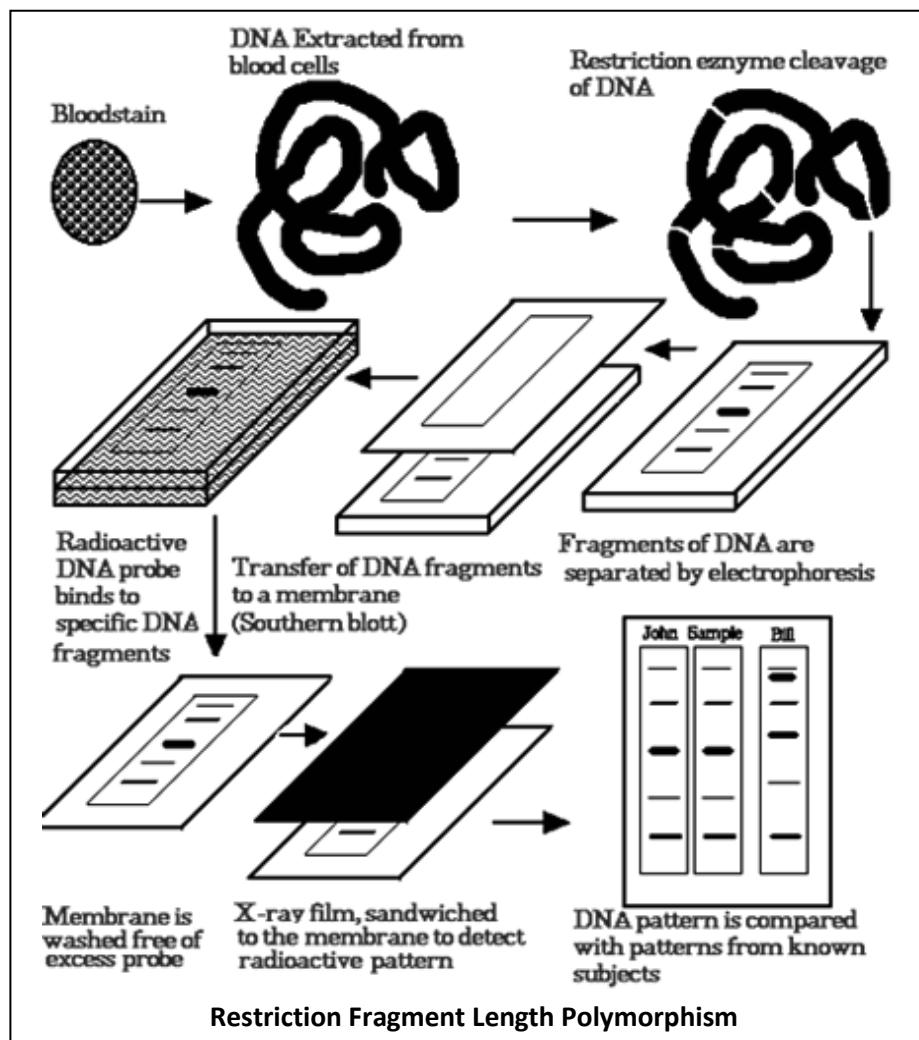
(i) Southern hybridization:

The DNA is digested with appropriate restriction enzyme, and separated by agarose gel electrophoresis. The so obtained DNA fragments are transferred to a nylon membrane. A DNA probe that spans the suspected restriction site is now added, and the hybridized bands are detected by autoradiograph. If the restriction site is absent, then only a single restriction fragment is detected. If the site is present, then two fragments are detected (fig).

(ii) Polymerase chain reaction:

RFLPs can also be scored by PCR. For this purpose, PCR primers that can anneal on

either side of the suspected restriction site used. After amplification by PCR, the DNA molecules are treated with restriction enzyme and then analysed by agarose gel electrophoresis. the restriction is absent only band is seen, while two bands found if the site found (fig).



are

If site one are is

Applications of RFLPs:

The approach by RFLP is very powerful and has helped many genes to be

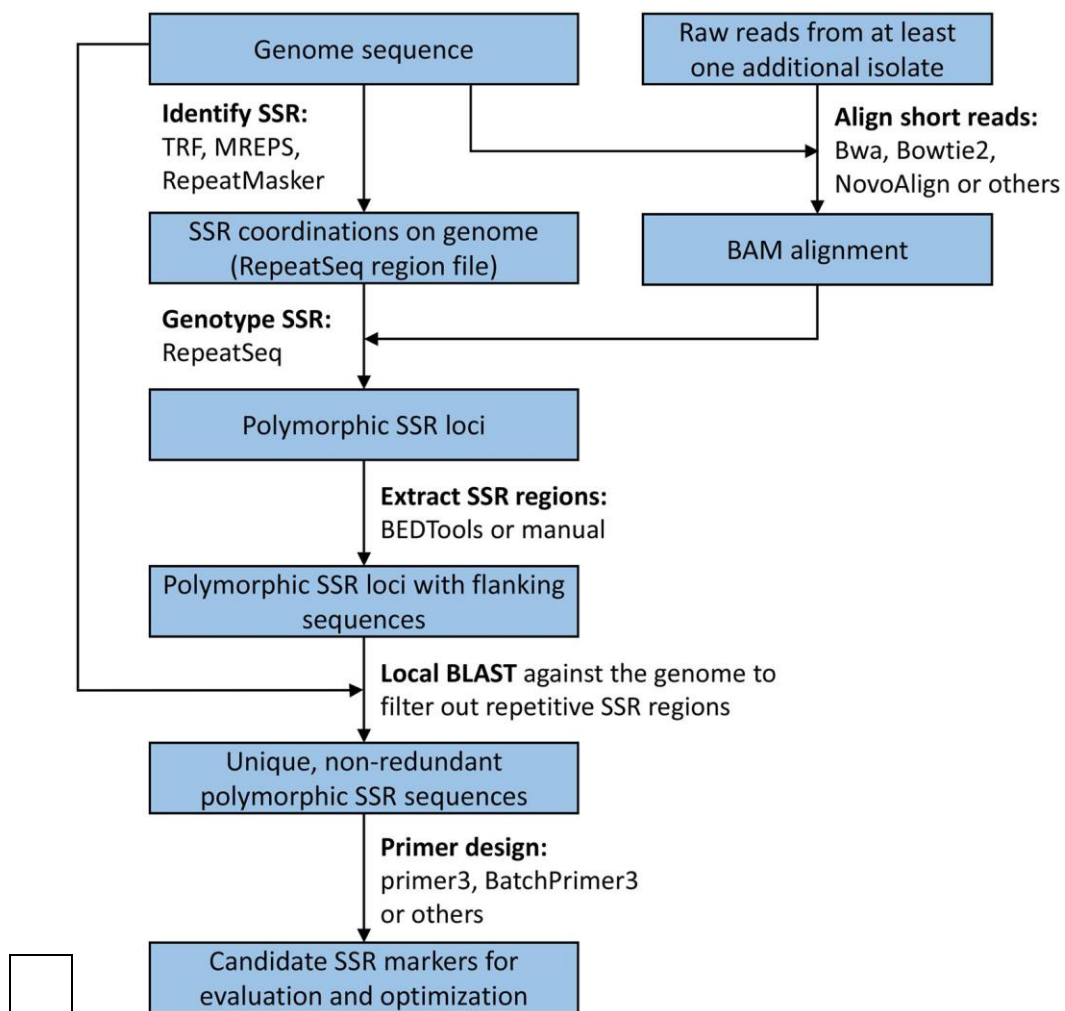
mapped on the chromosomes. E.g. sickle-cell anemia (chromosome 11), cystic fibrosis (chromosome 7), Huntington's disease (chromosome 4), retinoblastoma (chromosome 13), Alzheimer's disease (chromosome 21).

MICROSATELLITES/ SIMPLE SEQUENCE REPEATS

Microsatellites are short repeat units (10-30 copies) usually composed of dinucleotide or tetranucleotide units. These simple tandem repeats (STRs) are more popular than minisatellites (VNTRs) as DNA markers for two reasons.

1. Microsatellites are evenly distributed throughout the genome.
2. PCR can be effectively and conveniently used to identify the length of polymorphism.

Two variants (alleles) of DNA molecules with 5 and 10 repeating units of a dimer nucleotides(GA) are depicted in fig. By use of PCR, the region surrounding the microsatellites is amplified, separated by agarose gel electrophoresis are identified.



SSR (Simple sequence repeat)

Simple sequence repeat or Microsatellites, also known as Simple Sequence Repeats (SSRs) or Short Tandem Repeats (STRs), are repeating sequences of 2-6 base pairs of DNA. It is a type of Variable Number Tandem Repeat (VNTR). Microsatellites are typically co-dominant.

Sequence

Primer

ACTGTCG**ACACACACACACAC**GCTAGCT (AC)₇
TGACAGCT**TGTGTGTGTGTGT**GCGATCGA

ACTGTCG**ACACACACACACACAC**GCTAGCT (AC)₈
TGACAGCT**TGTGTGTGTGTGTGT**GCGATCGA

ACTGTCG**ACACACACACACACACACAC**GCTAGCT (AC)₁₀
TGACAGCT**TGTGTGTGTGTGTGTGTGT**GCGATCGA

ACTGTCG**ACACACACACACACACACACACAC**GCTAGCT (AC)₁₂
TGACAGCT**TGTGTGTGTGTGTGTGTGTGTGT**GCGATCGA





SINGLE NUCLEOTIDE POLYMORPHISMS (SNPs):

SNPs represent the positions in the genome where some individuals have one nucleotide (e.g. G) while others have a different nucleotide (e.g. C). There are large numbers of SNPs in genomes. It is estimated that the human genome contains at least 3 million SNPs. Some of these SNPs may give rise to RFLPs.

SNPs are highly useful as DNA markers since there is no need for gel electrophoresis and this saves a lot of time and labour. The detection of SNPs is based on the oligonucleotide hybridization analysis (fig).

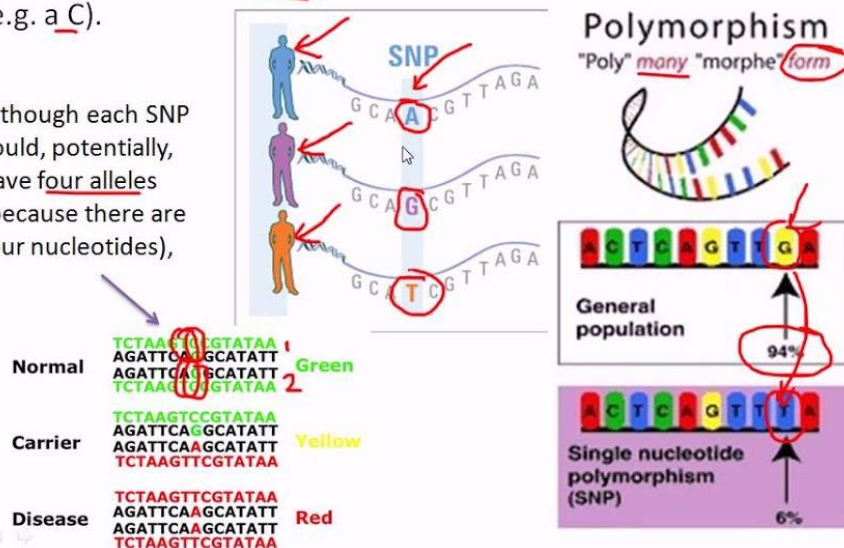
An oligonucleotide is a short single-stranded DNA molecule, synthesized in the laboratory with a length not usually exceeding 50 nucleotides. Under appropriate conditions, this nucleotide sequence will hybridize with a target DNA strand if both have completely base paired structure. Even a single mismatch in base pair will not allow the hybridization to occur.

DNA chip technology is most commonly used to screen SNPs hybridization with oligonucleotide.

Single nucleotide polymorphism (SNP)

These are positions in a genome where some individuals have one nucleotide (e.g. A) and others have a different nucleotide (e.g. C).

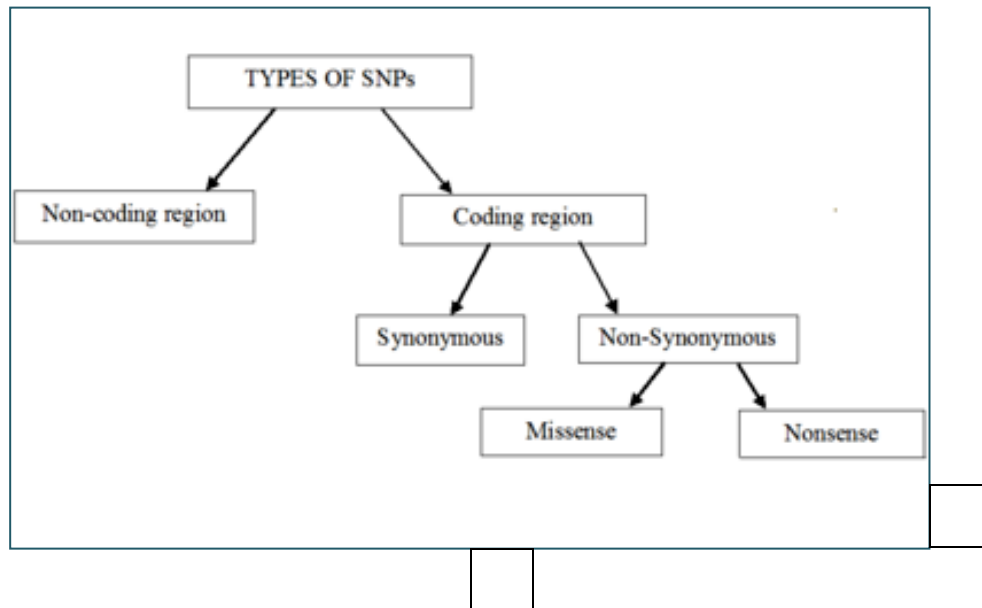
Although each SNP could, potentially, have four alleles (because there are four nucleotides),



CURRENT TECHNOLOGY OF DNA FINGERPRINTING

In the forensic analysis of DNA, the original techniques based on RFLPs and VNTRs are now largely replaced by microsatellites (short tandem repeats). The basic principle involves the amplification of microsatellites by polymerase chain reaction followed by their detection.

It is now possible to generate a DNA profile by automated DNA detection system (comparable to the DNA sequencing equipment).



Method – SNP development

BLAST

available flanking markers on GD sequence

Primer Design + PCR

ok

repeat

or

reject region

SEQ and SNP analysis

SNP found

ok

reject region

SNP confirmation of map position

with subset of resistant, susceptible and recombinant individuals closest to the gene

ok

reject region

SNP polymorphism

test with 8 founders

ok

SNP identified and validated

DNA FINGERPRINTING / DNA PROFILING /DNA TYPING

The use of DNA sequences to identify a person is called as DNA fingerprinting. It is a powerful tool for criminal investigations and other forensic applications. The technique was discovered in England by Alec Jeffreys. The basic principles of DNA fingerprinting are briefly described.

A sample of DNA from blood, semen, hair, or other body tissue is collected from the crime scene. If the sample is very small, PCR can be used to amplify it so that enough DNA is available for testing. Additional DNA samples are collected from one or more suspects. Each DNA sample is cut with one or more restriction enzymes, and the resulting DNA fragments are separated by gel electrophoresis. The fragments in the gel are denatured and transferred to nitrocellulose paper by Southern blotting. One or more radioactive probes is then hybridized to the nitrocellulose and detected by autoradiography. The pattern of bands produced by DNA from the sample collected at the crime scene is then compared with the patterns produced by DNA from the suspects. A match between the sample from the crime scene and one from the suspect can provide evidence that the suspect was present at the scene of the crime.

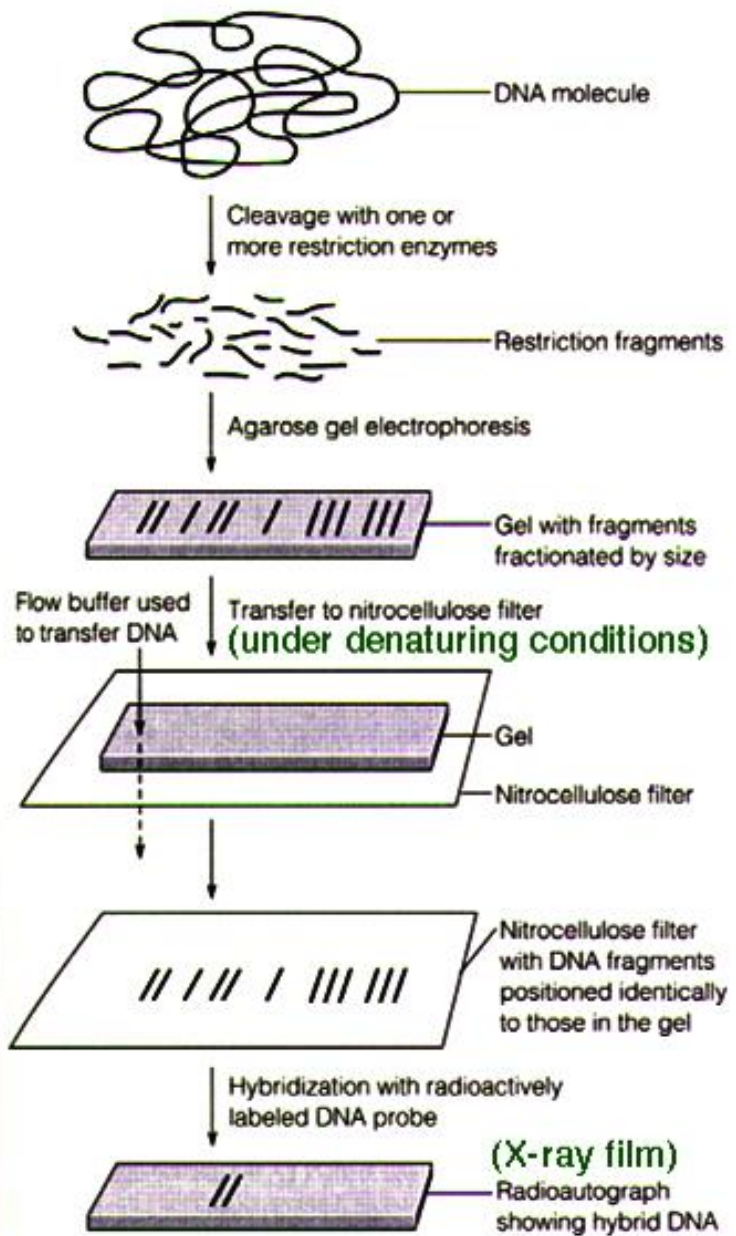
The probes most commonly used in DNA fingerprinting are complementary to short sequences repeated in tandem that are widely found in the human genome. People vary greatly in the number of copies of these repeats; thus, these polymorphisms are termed variable number of tandem repeats (VNTRs).

The structure of each person's genome is unique. The only exception being monozygotic identical twins. The unique nature of genome structure provides a good opportunity for the specific identification of an individual. Thus DNA fingerprinting is an analysis of the base sequence of DNA in an individual.

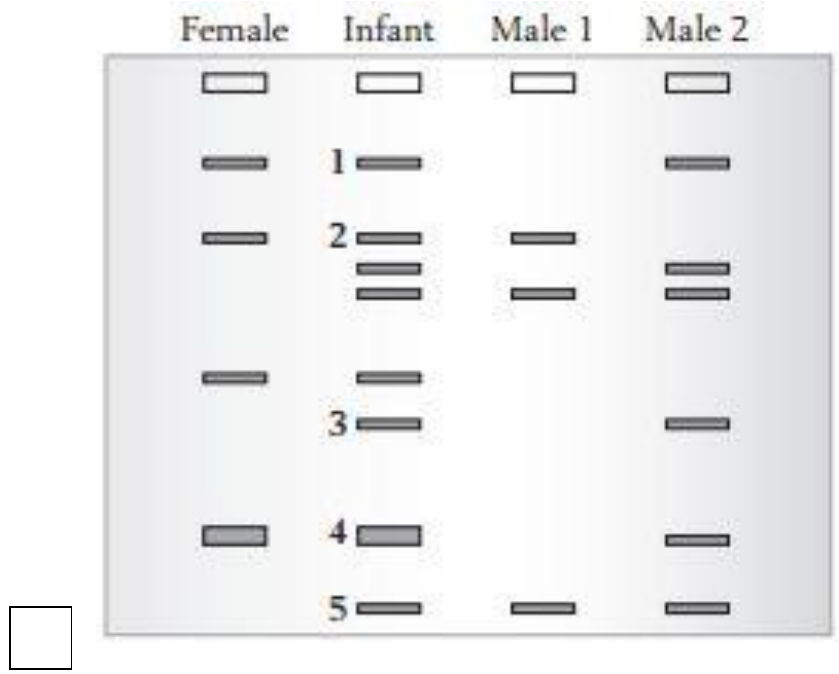
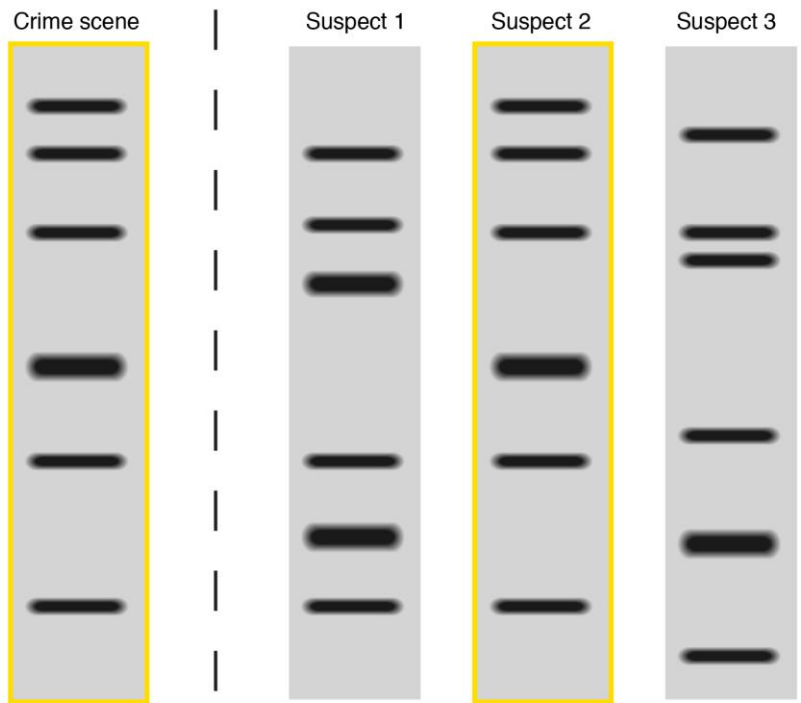
Applications of DNA fingerprinting:

The amount of DNA required for DNA fingerprint is remarkably small. The minute quantities of DNA from blood stains, body fluids, hair fiber or skin fragments are enough. Polymerase chain reaction is used to amplify this DNA for use in fingerprinting. DNA profiling has wide range of applications-most of them related to medical forensics. Some important ones are listed below.

- (i) Identification of criminals, rapists, thieves etc.
- (ii) Settlement of paternity disputes.
- (iii) Use in immigration test cases and disputes.
- (iv) Identification of crop varieties in agriculture.

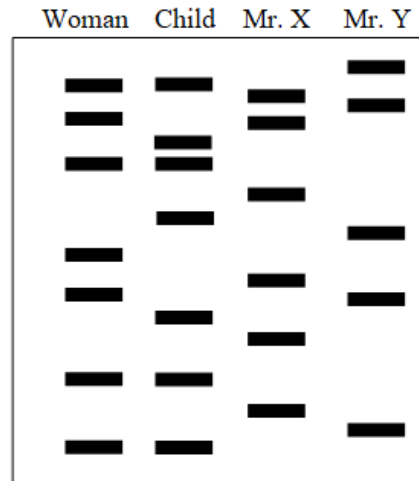


DNA FINGERPRINTING



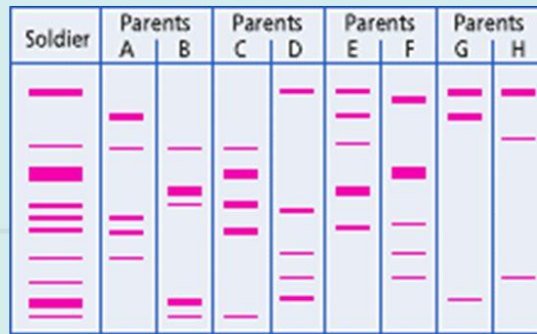


17. A woman claims that her former boyfriend, Mr. X, is the father of her child. Mr. X maintains that the child's father is Mr. Y and not him. DNA fingerprints of all four persons are shown below. What can be concluded about the child's paternity?

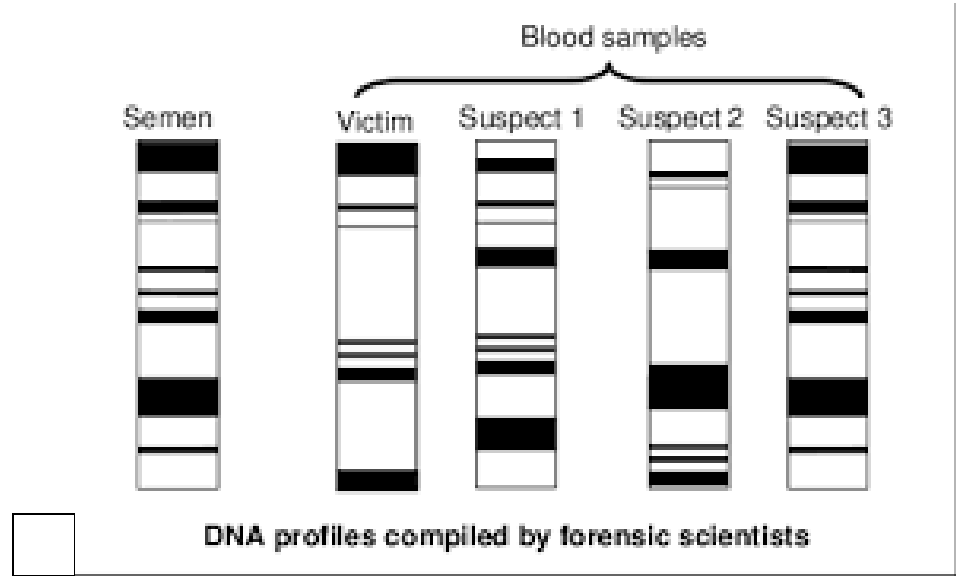


- Mr. X is the father.
- Mr. Y is the father.
- Neither Mr. X nor Mr. Y is the father.
- The woman is not this child's mother.
- Both c and d.

Interpreting a Gel - Who are the parents of this soldier?



http://www.glencoe.com/qa/images/b136/q4318/ch13_0_c.gif



MARKER ASSISTED SELECTION (MAS)

Selection of genotypes carrying desirable gene using linked-markers is called as marker assisted selection. In conventional breeding, breeders cross two different parents and select desirable genotypes in the F₂ generation by visual evaluation. The visual marker traits (eg. Black coloured stigma, purple stripes in leaves) that help in selection are called *morphological markers*. When biochemical characters (eg. Proline content) are considered for selection that character is called as *biochemical marker*. Both morphological and biochemical markers differ according to the environment. This defect can be overcome with the use of molecular markers.

Genetic markers are simply landmarks on chromosomes that serve as reference points to the location of other genes of interest when a genetic map is constructed. The markers that are currently widely used include RFLP, AFLP (amplified fragment length polymorphism), SNPs (single nucleotide polymorphisms), and microsatellites or simple sequence repeats (SSRs).

CLASSIFICATION OF MOLECULAR MARKERS:

Types of genetic markers

- **multilocus** markers (RAPD, AFLP, minisatellite DNA fingerprinting)
- **single-locus** markers (allozymes, microsatellites, SNPs)
- **dominant** markers – scored as present or absent (RAPD, AFLP, ...)
- **codominant** markers – identification of homologous alleles, i.e. scoring of homozygote and heterozygote states (allow estimation of allele frequencies – SNPs, microsatellites, ...)

1 **Single-locus, multiallelic, codominant markers.**

Examples are RFLPs and microsatellites (SSRs). Microsatellites are capable of detecting higher levels of polymorphisms than RFLPs.

2 **Multilocus, single-allelic, dominant markers.**

Examples are AFLPs and RAPD (random amplified polymorphic DNA).

POSITIVE AND NEGATIVE SELECTABLE MARKERS

- **Positive** selectable markers are selectable markers that confer selective advantage to the host organism. An example would be antibiotic resistance, which allows the host organism to survive antibiotic selection.
- **Negative** selectable markers are selectable markers that eliminate or inhibit growth of the host organism upon selection.^[6] An example would be thymidine kinase, which makes the host sensitive to ganciclovir selection.

A distinction can be made between selectable markers (which eliminate certain genotypes from the population) and screenable markers (which cause certain genotypes to be readily identifiable, at which point the experimenter must "score" or evaluate the population and act to retain the preferred genotypes). Most MAS uses screenable markers

rather than selectable markers.



POSITIVE AND NEGATIVE SELECTABLE MARKER

- **Positive** selectable markers are selectable markers that confers selective advantage to its host organism. An example would be antibiotics resistance, which allows the host organism to survive antibiotics selection.
- **Negative** selectable markers are selectable markers that would eliminate its host organism upon selection. An example would be thymidine kinase, which would make the host sensitive to ganciclovir selection.



PROPERTIES OF AN IDEAL MARKER:

- Easy recognition of all possible phenotypes (homo- and heterozygotes) from all different alleles
- Demonstrates measurable differences in expression between trait types or gene of interest alleles, early in the development of the organism
- Testing for the marker does not have variable success depending on the allele at the marker locus or the allele at the target locus (the gene of interest that determines the trait of interest).
- Low or null interaction among the markers allowing the use of many at the same time in a segregating population
- Abundant in number
- Polymorphic

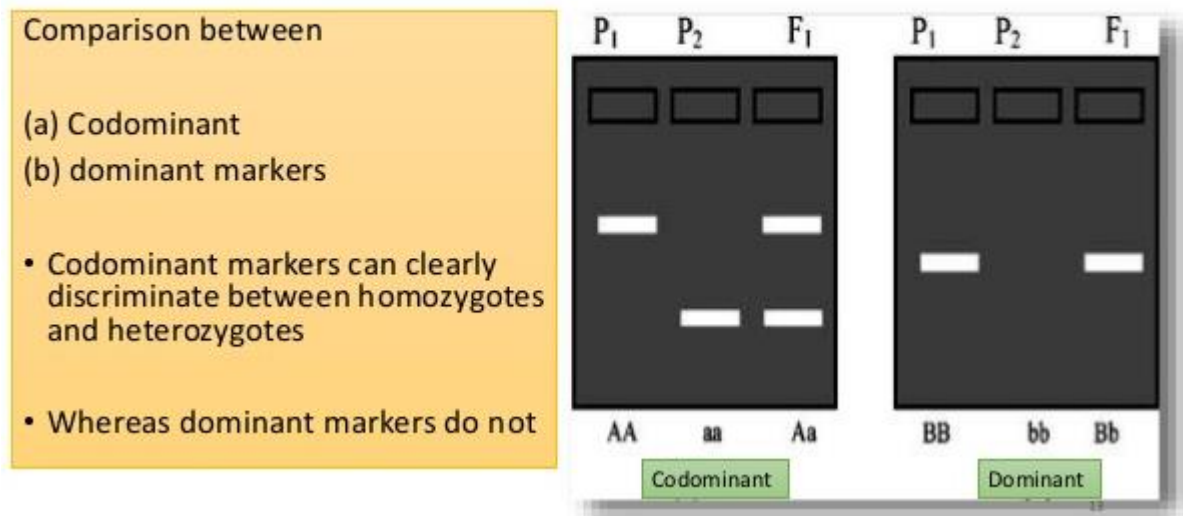
MARKER ASSISTED SELECTION

Broadly, marker assisted selection (MAS) can be divided into two categories, that of marker-assisted backcrossing or introgression and that of marker-assisted recurrent selection or population improvement. In the former, the goal is to incorporate one or few major genes or QTL into elite breeding lines (or in some situations, a breeding population). The second case involves using markers to improve the overall genetic value of a population with respect to some trait or suite of traits. Of the two, marker-assisted backcrossing, particularly of a single gene, is the easiest to put into practice; strategies to use markers in

recurrent selection are still being developed and the best strategy for a given situation is not clear at the current time.



Codominant and Dominant



FOREGROUND SELECTION VS. BACKGROUND SELECTION

Foreground selection refers to using markers that are tightly linked to the gene of interest in order to select for the target allele or gene. Background selection refers to using markers that are not tightly linked to the gene of interest in order to select against other DNA from the donor parent (i.e., to select for recurrent parent alleles at other loci than the target).

STEPS IN MARKER ASSISTED SELECTION

Generally the first step is to map the gene or quantitative trait locus (QTL) of interest first by using different techniques and then use this information for marker assisted selection. The markers to be used should be close to gene of interest (<5 recombination unit or cM). Use of two markers reduce the chances of an error due to homologous recombination. For example, if two flanking markers are used at same time with an interval between them of approximately 20cM, there is higher probability (99%) for recovery of the target gene.

The basic procedure for conducting MAS with DNA markers is as follows:

- a. Extract DNA from tissue of each individual or family in a population.
- b. Screen DNA samples via PCR for the molecular marker (SSR, SNP, SCAR, etc.) linked to the trait of interest.
- c. Separate and score PCR products, using an appropriate separation and detection technique.
- d. Identify individuals exhibiting the desired marker allele.
- e. Combine the marker results with other selection criteria (e.g., phenotypic data or other marker results), select the best fraction of the population, and advance those individuals in the breeding program.

MARKER ASSISTED BACKCROSSING OF QUANTITATIVE TRAITS

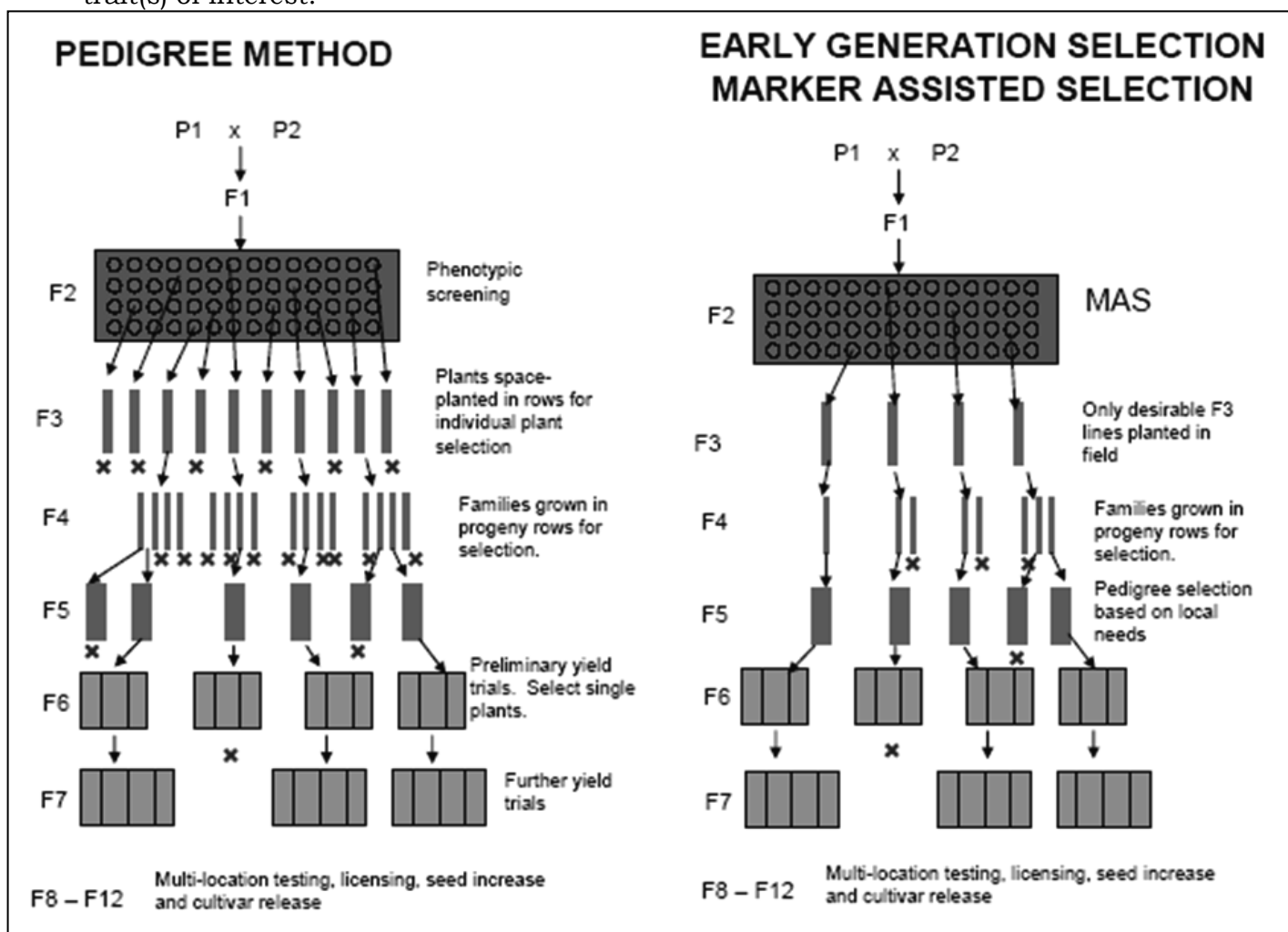
Marker assisted backcrossing uses DNA markers, which can be scored as a dominant or codominant trait prior to flowering, to facilitate the backcrossing program, saving time if progeny testing would need to be conducted and saving resources if phenotyping is difficult. Markers can be used to select for the gene being introgressed into the recurrent parent and to select against undesirable donor DNA. Markers enable the pyramiding of resistance genes; that is, they enable the incorporation of alleles at multiple loci each of which confers resistance to the same race of pathogen. This is difficult to do traditionally because one locus masks the presence of the others.

The general method is as follows:

Season 1: Cross two lines or genotypes to produce an F1.

Season 2: Generate ≥ 100 BC₁ individuals, select against clearly undesirable types but otherwise backcross all to the recurrent parent.

Season 3-4: Generate ≥ 200 BC₂ individuals (say 2-3 plants per BC₁), select against clearly undesirable types but otherwise self pollinate for one or two generations to develop BC₂F₂ and BC₂F₃ families for phenotypic evaluation. Develop a genetic map of the BC₂F₁ individuals; use the marker and phenotypic data to localize QTL controlling the trait(s) of interest.



Season 5: Identify BC₂F₂ lines (families) that are segregating for these QTL and continue backcrossing until a desired level of recurrent parent is reached followed by self-fertilization to recover homozygous plants for the introgressed QTL region. Ideally, the breeder could develop nearly isogenic lines for each QTL – called “QTL-NIL” – using markers to narrow the introgressed region and to recover the recurrent parent region.

The goal of this process is to develop a set of lines each carrying a single desirable QTL allele introgressed from the donor. With this set of lines, the effect of each QTL on the phenotype can be assessed. Later, after evaluating individual QTL, crosses among NILs carrying different QTL can be made to pyramid multiple QTL for the given trait into a single line.

Advantages of molecular markers:

- i. The desirable segregants can be scored at the seedling stage itself
- ii. It is possible to screen for difficult and expensive characters such as tolerance to drought, flood, salinity, mineral deficiency, toxicity, pest resistance, disease resistance etc.
- iii. Selection can be practiced for several traits.
- iv. Heterozygotes and homozygotes are easily identified without resorting for progeny testing.
- v. It speeds up breeding cycle by saving enormously on time.
- vi. Space required for screening is very low.

Disadvantages of MAS

Currently, one of the most important barriers for MAS is the prohibitive cost. Although there are only a small number of reports analyzing the economics of MAS versus conventional breeding in the literature, the cost-effectiveness of using MAS compared to conventional plant breeding varies considerably between studies.

Two additional factors need to be considered for cost-analysis:

- (1) the equipment and consumables required to establish and maintain a marker lab is considerable; and
- (2) there is a large initial cost in the development of markers which is seldom reported.

For marker assisted backcrossing, the initial cost of using markers would be more expensive compared to conventional breeding in the short term however time savings could lead to an accelerated variety release which could translate into greater profits in the medium to long term.

Another important factor obstructing the successful application of markers for line development is the low reliability of markers to determine phenotype. This is often attributable to the thoroughness of the primary QTL mapping study. Even QTLs that are detected with high LOD scores and explain a large proportion of the phenotype may be affected by sampling bias (especially in small populations), and therefore may not be useful for MAS. Furthermore, the effect of a QTL may depend on the genetic background. This emphasizes the importance of testing the QTL effects and the reliability of markers (i.e. QTL/marker validation) before MAS is undertaken.

Finally the level of integration between molecular geneticists and plant breeders (and scientists from other disciplines) may not be adequate to ensure that markers are effectively applied for line development.

Kindly go through the following article to know more on Markers, MAS and QTL.

Collard, Bertrand & Jahufer, Zulfi & Brouwer, J.B. & Pang, Edwin. (2005). An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica*. 142. 169-196. 10.1007/s10681-005-1681-5.

RECOMBINANT DNA TECHNOLOGY

Genetic recombination which refers to the exchange of genetic information between two individual organisms occurs in nature. When a gene (DNA fragment) of one species is transferred to another organism artificially it is called recombinant DNA technology/genetic engineering. Boyer and Cohen (1973) were the first to successfully recombine two plasmids (pSC 101 and pSC 102) and clone the recombinant plasmid in *E. coli*. This marked the beginning of modern rDNA technology. DNA manipulation is made possible by three key factors namely, the availability of restriction enzymes that cut DNA at specific sequences, the discovery of cloning vectors which carry the transgene into a host and the annealing (rejoining) properties of nucleic acids,.

BASIC PRINCIPLES OF DNA CLONING USING rDNA TECHNOLOGY

The basic principles of involved in DNA cloning are as follows:

1. Generation of DNA fragments by restriction enzymes and selection of the desired fragment of gene.
2. Insertion of the selected DNA fragment into a cloning vector to create a recombinant DNA.
3. Introduction of the recombinant vectors into the host cells (eg. bacteria).
4. Multiplication and selection of clones containing the transgene.
5. Expression of the gene to produce the desired product.

OBJECTIVES OF RECOMBINANT DNA TECHNOLOGY

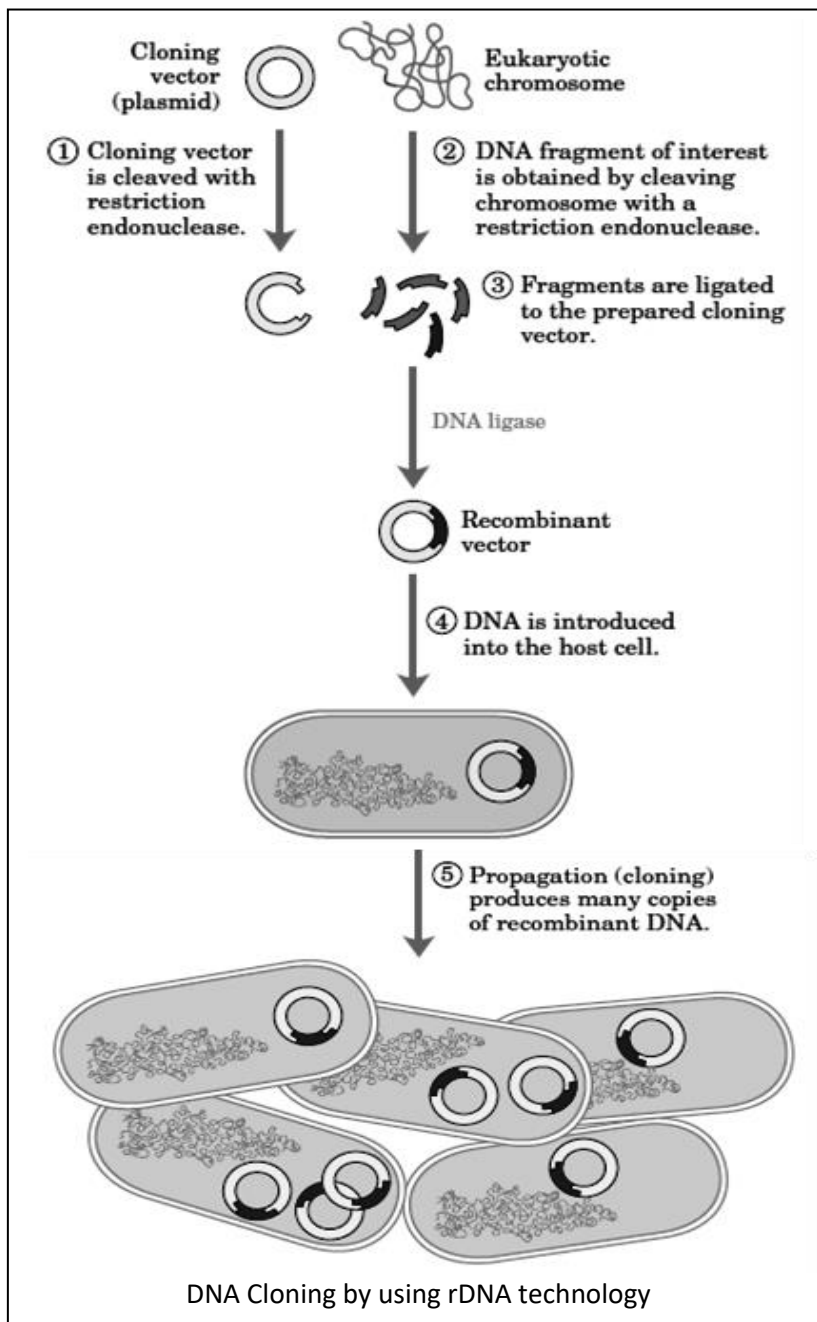
1. To obtain a large number of copies of specific DNA fragments
2. To recover large quantities of the protein production by the concerned gene
3. To integrate gene into the chromosome of target organism where it expresses itself.

Many enzymes which are involved in rDNA technology and their functions are given below.

Enzymes	Name of the enzyme
<i>Cleaving (cutting) DNA</i>	Nucleases a) Endonucleases b) Exonucleases S1 Nucleases DNA ases
<i>Joining the DNA fragments</i>	DNA ligases
<i>Synthesis of DNA</i>	DNA polymerase I Terminal transferase Reverse transcriptase
<i>Enzymes that modify ends of DNA molecule</i>	Alkaline phosphatase and Kinase
<i>Enzyme that degrade RNA</i>	RNA ases.

The various aspects involved in the recombinant DNA technology listed below are are described in the following chapters:

- Restriction endonucleases (DNA cutting enzymes)
- Cloning vectors (carriers)
- DNA amplification and cloning in the host
- Methods of gene transfer
- Molecular diagnostic techniques (Hybridization and blotting techniques).



I. RESTRICTION ENDONUCLEASES / DNA CUTTING ENZYMES

Restriction enzymes are bacterial enzymes which cleaves (cuts) the DNA at specific sites. Werner Arber discovered that certain enzymes protect the *E. coli* bacterium by cutting and destroying the invading viral DNA. The enzymes that restrict the viral replication are called as restriction enzymes eg. Eco RI. More than 800 enzymes are commercially available now.

Nomenclature of Restriction enzymes

Restriction enzymes have a three letter name with the first letter indicating the genus name, followed by the first two letters of the species name followed by the strain of the organism with a Roman numeral indicating the order of discovery. For example, Eco RI is derived from *Escherichia coli*, strain **Ry** 13 and the first endonuclease (**I**) to be discovered.

Restriction Enzyme	Source	Recognition Sequence and Site of Cleavage	Product
1. <i>Alu</i> I	<i>Arthrobacter luteus</i>	$\begin{array}{c} \downarrow \\ 5'-A-G-C-T-3' \\ 3'-T-C-G-A-5' \\ \uparrow \end{array}$	 blunt ends
2. <i>Bam</i> H I	<i>Bacillus amyloliquefaciens H</i>	$\begin{array}{c} \downarrow \\ 5'-G-G-A-T-C-C-3' \\ 3'-C-C-T-A-G-G-5' \\ \uparrow \end{array}$	 Sticky ends
3. <i>Eco</i> R I	<i>Escherichia coli</i> RY13	$\begin{array}{c} \downarrow \\ 5'-G-A-A-T-T-C-3' \\ 3'-C-T-T-A-A-G-5' \\ \uparrow \end{array}$	 Sticky ends
4. <i>Eco</i> R II	<i>Escherichia coli</i> R245	$\begin{array}{c} \downarrow \\ 5'-C-C-T-G-G-3' \\ 3'-G-G-A-C-C-5' \\ \uparrow \end{array}$	$\begin{array}{c} 5' \text{ C-C-T-G-G-} 3' \\ \\ 3' \text{ G-G-A-C-C-} 5' \end{array}$ blunt ends
5. <i>Hae</i> III	<i>Haemophilus aegyptius</i>	$\begin{array}{c} \downarrow \\ 5'-G-G-C-C-3' \\ 3'-C-C-G-G-5' \\ \uparrow \end{array}$	 blunt ends
6. <i>Hin</i> d III	<i>Haemophilus influenzae</i> Rd	$\begin{array}{c} \downarrow \\ 5'-A-A-G-C-T-T-3' \\ 3'-T-T-C-G-A-A-5' \\ \uparrow \end{array}$	 Sticky ends
7. <i>Hin</i> d II (first discovered Restriction endonuclease)	<i>Haemophilus influenzae</i> Rd	$\begin{array}{c} \downarrow \\ 5'-G-T-C-G-A-C-3' \\ 3'-C-A-G-C-T-G-5' \\ \uparrow \end{array}$	 blunt ends

There are three categories of restriction endonuclease:

1. Type I enzymes : They have different subunits for recognition, modification, and restriction or cleavage. The cleavage site is located more than 1,000 base pairs away from the recognition site. As a consequence, cleavage does not occur at a specific sequence even though certain regions are preferentially cleaved. It is not possible to define the recognition sites by characterizing the broken ends of the DNA.

2. Type II enzymes: It occur in about one in three bacterial strains. The enzymes are highly specific in action, as they are involved in only one act of restriction. The recognition sites or sequences are usually short (4 to 6 bp) and often palindromic. They cleave at or close to the target site and require no ATP for restriction. Some enzymes produce blunt ends while others produce staggered cuts (or sticky ends). The type II enzymes are the workhorses of recombinant DNA technology.

3. Type III enzymes: They have two subunits, one for recognition and methylation and the other for restriction. Like type I, type III restriction sites consist of asymmetrical sequences that may be 5 to 7 bp long. Cleavage occurs some 24 to 26 bp downstream from the recognition site.

Nomenclature of Restriction endonuclease

	Type II Enzyme	Type III Enzyme	Type I Enzyme
Protein structure	Separate endonuclease and methylase	Bifunctional enzyme of 2 subunits	Bifunctional enzyme of 3 subunits
Recognition site	4-6 bp sequence, often palindromic	5-7 bp Asymmetric sequence	Bipartite and asymmetric
Cleavage site	Same as or close to recognition site	24-26 bp downstream of recognition site	Nonspecific > 1000 bp from recognition site
Restriction & methylation	Separate reactions	Simultaneous	Mutually exclusive

Whereas all three types of enzymes previously described are proteins with a catalytic effect, there is a unique class of non-protein enzymes called ribozymes. These are RNA enzymes with the capacity for cleaving specific phosphodiester bonds. Recognition sequences

Restriction enzymes recognize specific palindromic sequences in the double stranded DNA which are 4-6 nucleotides long, and then cut both the strands of DNA at specific locations. These sequences are called as recognition sequences.

Cleavage patterns

Most REs cleave recognition sequence at one or two base pairs away from the center on both strands of DNA. This results in a double-stranded DNA with short, single-stranded

ends called sticky ends/cohesive ends. Examples Eco RI and Taq I. Other REs cleave the recognition sequence without any overhanging nucleotides resulting in DNA with blunt ends / flush ends eg. Hae III or Hpa I. The DNA fragments with sticky ends can join easily with other DNA fragments. The blunt ends can be joined with the help of adaptors. Adaptors are short, chemically synthesized DNA double strands which can be used to link the ends of two DNA molecules that have different sequences at their ends.

Restriction maps

Any double stranded DNA can be cut by a variety of restriction enzymes. After separating the restriction fragments by gel electrophoresis and measuring their size, the site where each restriction enzyme cleaved can be found out. A map showing the position of cut sites for a variety of restriction enzyme is called the restriction map for that DNA fragment. Restriction maps allows comparison between DNA molecules without the need to determine the nucleotide sequence.

DNA LIGASES / DNA JOINING ENZYMES

The cut DNA fragments are joined together by DNA ligases. DNA ligase joins the DNA fragments by forming phosphodiester bond between the 5' phosphate group and 3' hydroxyl group.

GENE ISOLATION

Gene isolation is one of the major activities of biotechnology. Before a gene can be genetically engineered, it must first be identified, isolated, and characterized (e.g., number and position of introns, the promoter and its elements). Isolating a gene enables researchers to determine its nucleotide sequence. From the DNA sequence several things can be deduced, including the amino acid sequence and the protein structure and function of the gene's product. In order to transfer a gene from one individual to another, it must first be identified and isolated. Isolation of a gene permits it to be amplified to obtain large quantities for studies.

A number of strategies may be used to isolate or clone a gene.

1. Activation tagging

This strategy requires the availability of a well-characterized transposon system, something that is lacking in many species, except species like corn. The gene to be isolated is first inactivated by transposon insertion, resulting in the formation of a mutant. The DNA sequence of the transposon is used to identify the clones that contain the gene of interest.

2. cDNA screening

A cDNA library is first created. A probe is then designed and used to screen the library to hybridize to the sequence of interest.

3. Map-based gene cloning

Map-based cloning or positional cloning is an rDNA-based method for identifying a gene without first knowing its product. The first step in this method is to produce a high-resolution genetic map (average distance of less than 5 centi Morgans). This is followed by the production of a physical map (a map of the location of identifiable landmarks on DNA regardless of their inheritance). The principal procedures include physical mapping by contig construction using BACs, YACs, STS-content mapping, DNA fingerprinting, and pulse-field gel electrophoresis. Once a physical map is in place, the target gene may be identified by chromosome walking, using RFLP or other molecular markers. This entails starting with a closely linked RFLP probe and isolating genomic clones that it corresponds with, and then walking from these clones to the target genes. Alternatively, molecular markers that are tightly linked with the gene of interest are first identified. The DNA markers are used to screen a genomic library to isolate clones that contain the target gene (called chromosome landing). Genetic complementation through transformation is also part of this process of gene identification.

4. Transformation-associated recombination

This method of gene isolation capitalizes on the natural ability of yeast cells to find and combine similar DNAs, regardless of their origin. Yeast cells are transformed with pieces of DNA along with a small fragment of the target DNA. As the yeast cells reproduce, only DNA that complements the small piece of DNA introduced into the cell are maintained (cloned).

II. VECTORS CARRIERS OF TRANSGENE

A vector is a DNA molecule which carries the foreign genetic material into another cell. A vector carrying a foreign gene is a chimera and is called as the recombinant DNA.

A vector must contain:

- (a) An ori site (Origin of replication)
- (b) Multiple cloning sites which are the active sites of the RE enzymes and they allow insertions of DNA into the vector to be targeted.
- (c) Selectable markers: such as an antibiotic resistance [e.g. tetracycline] is often carried by the vector to allow the selection of positively transformed cells.

Other desirable features that can be present in a suitable cloning vector may be

- (i) *vir* genes for plant transformation,
- (ii) intergrase sites for chromosomal insertion,
- (iii) *lacZ* fragment for complementation and blue -white selection,
- (iv) reporter genes flanking the MCS to facilitate the production of recombinant proteins.

Properties of a good vector

- It should be able to replicate independent of the replication of host chromosome
- It should be easy to isolate, purify and introduce into the host cells.
- The vector should have suitable marker genes that allow easy detection or/and selection of the transformed the host cell. Eg. Genes for ampicillin and Tetracycline resistance.
- A vector should contain unique target sites for as many restriction enzymes as possible into which the DNA insert can be integrated.
- When expression of the DNA insert is desired, the vector should contain suitable regulatory elements like promoter, operator, ribosome binding sites.

There are four major types of vectors namely, plasmids, bacteriophages, cosmids, and artificial chromosomes (BAC, YAC etc.), of which the first two are most commonly used ones.

a). Plasmid vectors

A plasmid is a circular, double-stranded, self-replicating extrachromosomal molecules present in some bacteria. In nature, plasmids confer antibiotic resistance to the host bacteria. They replicate independent of the bacterial DNA and can accept about 6-10 kb. Eg. pUC18. pUC refers to **p**lasmid from University of California.

b). Bacteriophages

The virus that infect and replicate inside the bacteria are called bacteriophages or phages. About one-third of the phage genome is nonessential and can be replaced with foreign DNA. These viral vectors can carry about 23 kb of DNA or RNA, and contain viral promoters for translating the target gene in the host cell. They often produce an identification mark to indicate successful transfection. The commonly used phages are phage λ , phage M₁₃.

c). Cosmids

A cosmid is a “hybrid” between a plasmid and a phage. It consists of the cos sequence of phage lambda (required for packing the phage DNA into the phage protein coat), the plasmid sequence for replication, and an antibiotic resistance gene to identify the host cell carrying the cosmids. They can handle about 40 kb of cloned DNA and can be maintained as either plasmids or bacteriophage λ vectors because they have an *E. coli* origin of replication. They also have cohesive ends (cos) sites found in phage. An example of a cosmid is the pJB8-5.

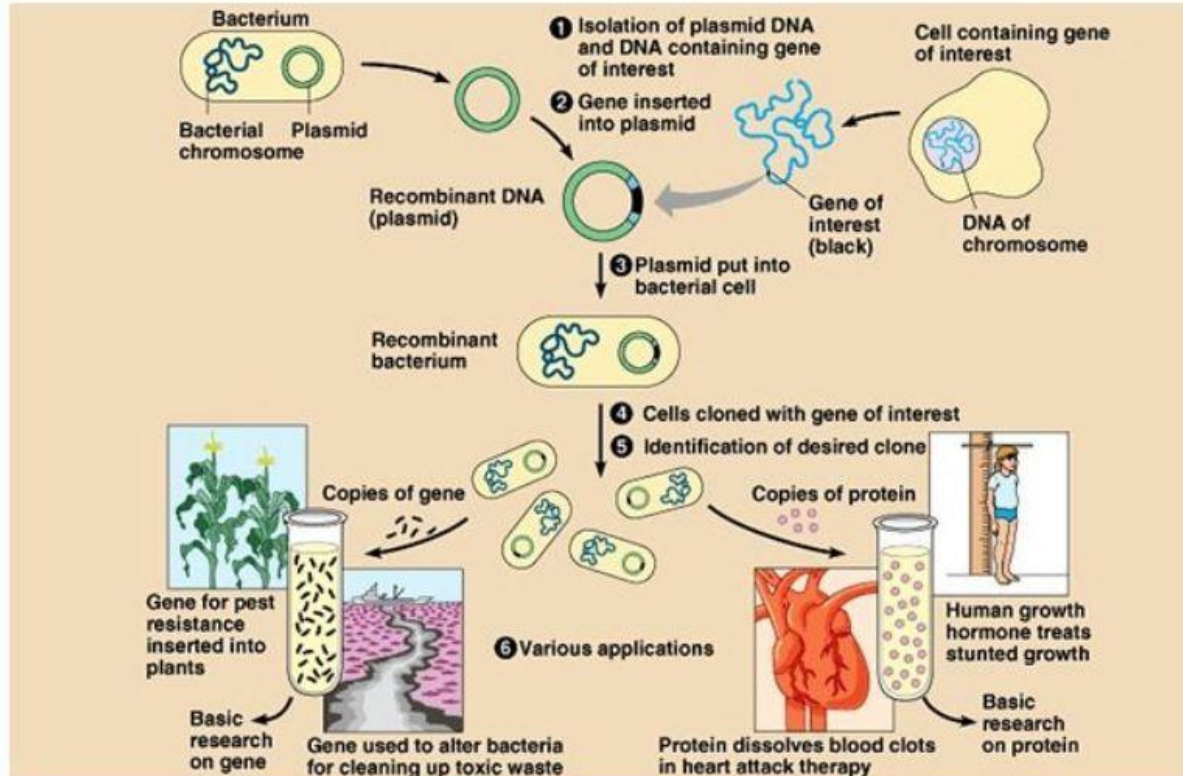
Cosmids are vectors possessing the characteristics of both plasmids and bacteriophages. It can be constructed by adding a fragment of phage λ DNA including cos site, to plasmids. They are capable of carrying about 42 kb of inserts. They replicate intensely and their success rate is high.

d). Bacterial Artificial Chromosomes (BACs): Bacterial artificial chromosomes are plasmids designed for the cloning of very long segments (100 to 300 kb) of DNA. They include selectable markers such as resistance to the antibiotic chloramphenicol, as well as a stable origin of replication (ori) that maintains the plasmid at one or two copies per cell. The large circular DNAs are then introduced into host bacteria by electroporation.

e). Yeast artificial chromosomes (YAC)

These are useful for cloning large fragments (100-300kb). They may be used to even clone whole chromosomes. But they are notorious for generating chimeras (recombinant molecules in which non-contiguous donor fragments are joined together), which are laborious to separate from the desired recombinants.

CATEGORIES OF VECTORS BY FUNCTIONS



Vectors in Molecular Biology

1. Cloning vectors

The vectors used for multiplication of DNA fragments in a suitable host are called cloning vectors. A vector is used because it provides an Ori (origin of replication site). For increased efficiency the original restriction sites of most cloning vectors are replaced by a synthetic multiple cloning site (contains many restriction sites). Other additional features that may be engineered into vectors include vir genes (for plant transformation), integrase sites (for chromosomal insertion), and lacZ α fragment (for a complementation).

The choice of cloning vector to use for a particular project depends on the insert size, copy number, cloning sites, selectable marker, and incompatibility. Two of the most common vectors used in genetic engineering are the Plasmids and the Phage DNA.

Features	Cloning Vector	Expression Vector
Definition	A cloning vector is a small piece of DNA, which is used to introduce the foreign gene of interest into the host cell. It can be stably maintained within the host organism.	An expression vector is a plasmid that not only introduces a gene of interest into host organism but also aids in the analysis of the gene of interest via relevant protein product expression.
Role	It is used to obtain multiple copies of the foreign gene of interest.	It is used to obtain or analyse the gene product, which may be RNA or protein, of the inserted gene of interest.
Types	They can be plasmids, cosmids, and phages, BACs, YACs or MACs.	Only a plasmid can be an expression vector.
Features	A typical cloning vector consists of an origin of replication, unique restriction sites, reporter gene and an antibiotic resistance site.	An expression vector consists of regulatory elements like enhancers, promoters, termination sequences, transcription initiation site and translation initiation sites, in addition to having all the features of a typical cloning vector.

2. Transcription vectors

Transcription is a necessary component of all vectors. Stable expression of an insert depends on stable transcription which depends on the promoters in the vector. Transcription vectors are designed to only be transcribed (replicated or amplified) but not translated (expressed). They are relatively simple in their construction. Plasmid transcription vectors cannot be used as transcription vector.

3. Expression vectors

When a vector is designed for the expression of i.e. production of the protein specified by the DNA insert, it is known as expression vector. Sometimes the goal in gene cloning is not just to replicate the gene, but also to produce the protein that it encodes. In addition to the usual origin of replication, restriction sites, and selectable markers, contains sequences required for transcription and translation in bacterial cells. These additional sequences may include:

1. A bacterial promoter, such as the *lac* promoter. The promoter precedes a restriction site where foreign DNA is to be inserted, allowing transcription of the foreign sequence to be regulated by adding substances that induce the promoter.
2. A DNA sequence that, when transcribed into RNA, produces a prokaryotic ribosome binding site.
3. Prokaryotic transcription initiation and termination sequences.
4. Sequences that control transcription initiation, such as regulator genes and operators.

The bacterial promoter and ribosome-binding site are usually placed upstream of the restriction site, which allows the foreign DNA to be inserted just downstream of the initiation codon. When the plasmid is placed in a bacterial cell, RNA polymerase binds to the promoter and transcribes the foreign DNA. Bacterial ribosomes attach to the ribosome-

binding site on the RNA and translate the sequence into a foreign protein. An expression vector contains a promoter, a ribosome-binding site, and other sequences that allow a cloned gene to be transcribed and translated in bacteria.

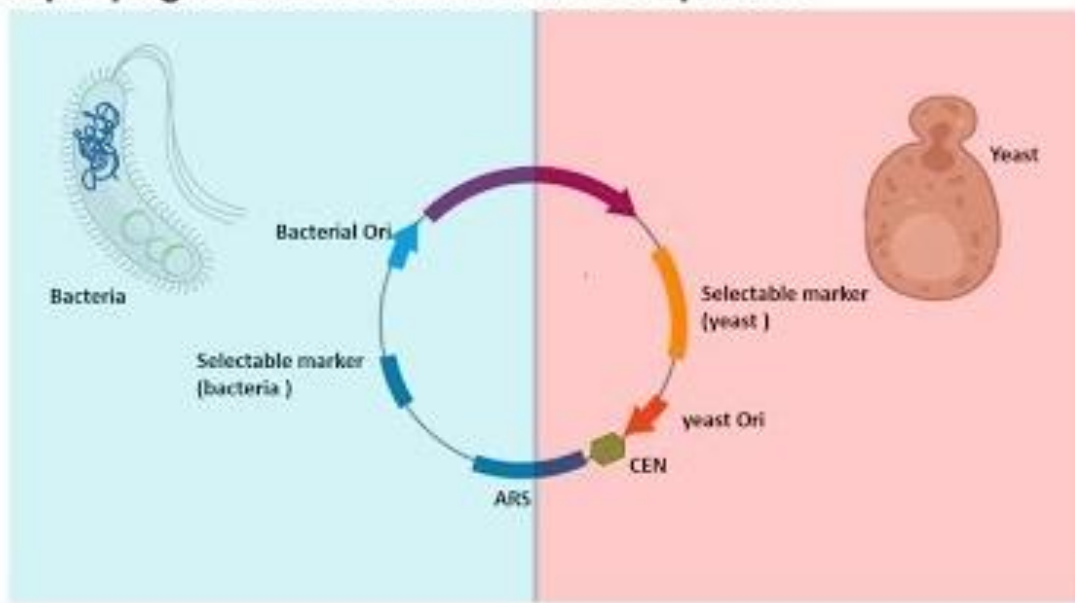
One of the first commercial products produced by recombinant DNA technology was the protein insulin. The gene for human insulin was isolated and inserted into bacteria, which were then multiplied and used to synthesize human insulin.

4. Shuttle vectors

Shuttle vectors are plasmids capable of propagating and transferring/shuttling genes between two different organisms. Hence, they contain two origins of replication, one specific for each host species, as well as those genes necessary for their replication and not provided by the host cells. These vectors are created by recombinant DNA technique. Some of them can be grown in two different prokaryotic species, while others can propagate in prokaryotic species (E.coli) and eukaryotic one (yeast, plants and animals) Since these vectors can be grown in one host and then moved into another without any extra manipulation they are called shuttle vectors.

Shuttle vectors have been designed to specifically satisfy the need i.e. for the initial cloning of DNA inserts in E.coli and sub-sequent functional tests in the species to which the DNA inserts belong. Most of the eukaryotic vectors are shuttle vectors.

A shuttle vector is a plasmid engineered in a way so that it can propagate in two different host species



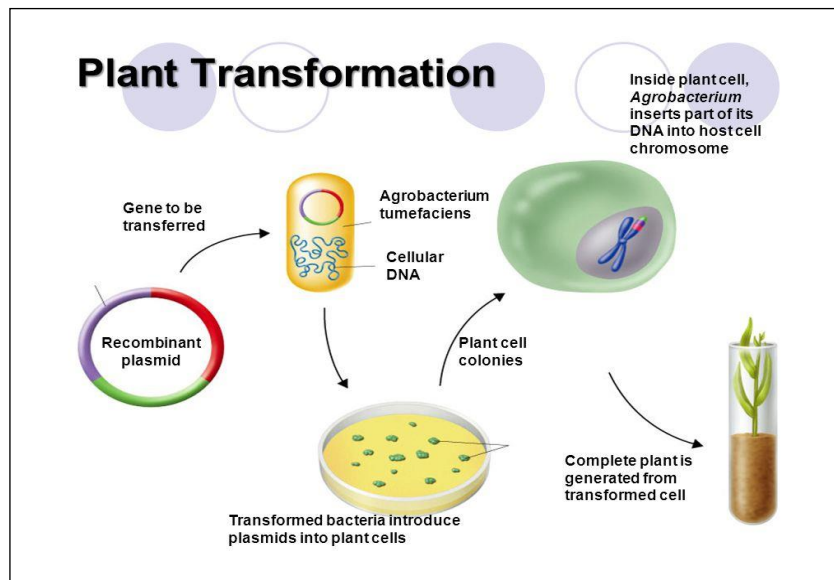
III. METHODS OF GENE TRANSFER

Introduction of a foreign gene into the host cell can be done indirectly or directly.

1. DIRECT GENE TRANSFER METHODS

The most commonly used methods includes transformation, conjugation, electroporation and liposome mediated gene transfer.

i. Transformation:

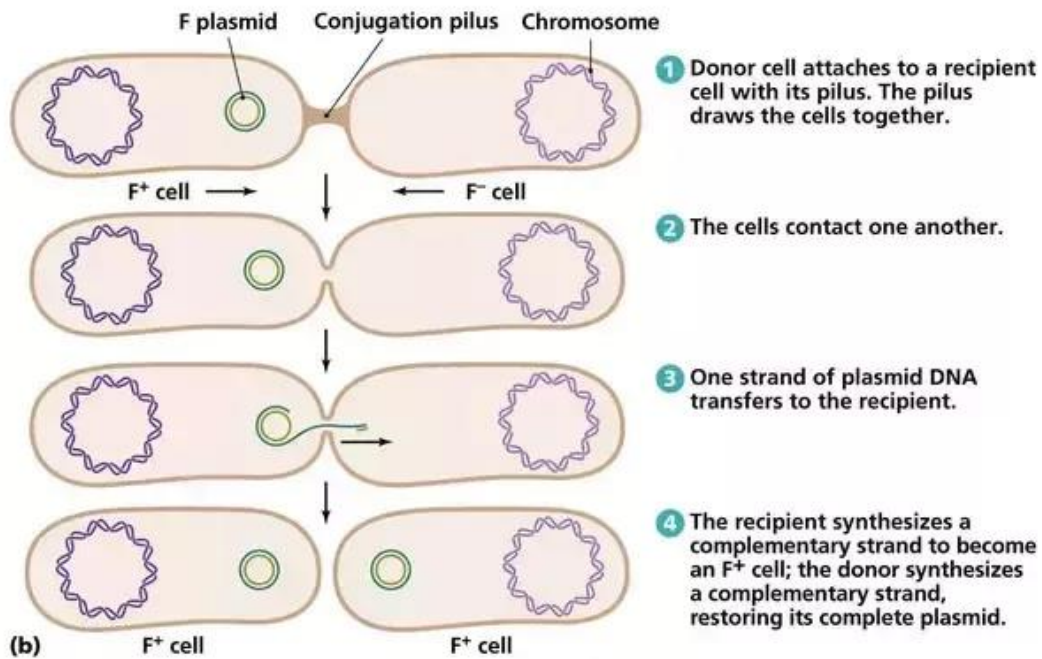


Transformation is the method of introducing foreign DNA into bacterial cells. The uptake of plasmid DNA by *E.coli* is carried out in 0 to 5°C ice-cold CaCl_2 and a subsequent heat shock (37-45°C for 90sec). the transformational frequency which is the fraction of transformed cells is good (one cell per 1000

cells). Eg. *Agrobacterium* mediated gene transfer.

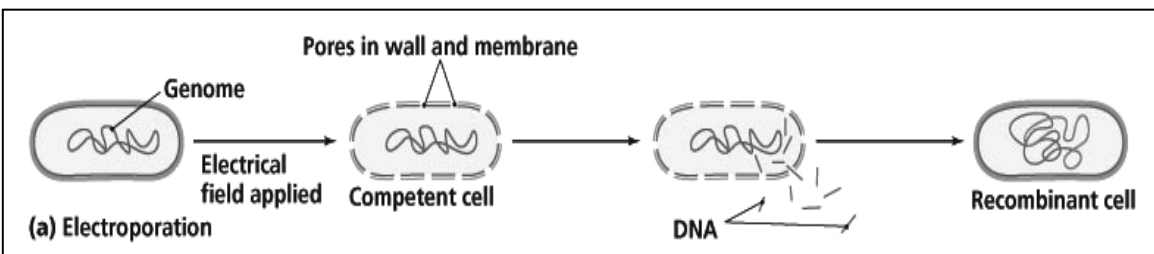
ii. Conjugation:

Conjugation is a natural microbial recombination process. During conjugation, two live bacteria (a donor and a recipient) come together, join by cytoplasmic bridges and transfer single stranded DNA from the donor to the recipient cell. The new DNA strand may integrate with the host chromosome or remain free. The natural phenomenon is exploited for transferring plasmid-insert DNA to new recipient cells.



iii. Electroporation:

Electric shocks can also induce cellular uptake of exogenous DNA from the suspended solution by means of pores formed by electric pulses. It is a simple and rapid technique for introducing genes into the cells from various organisms.



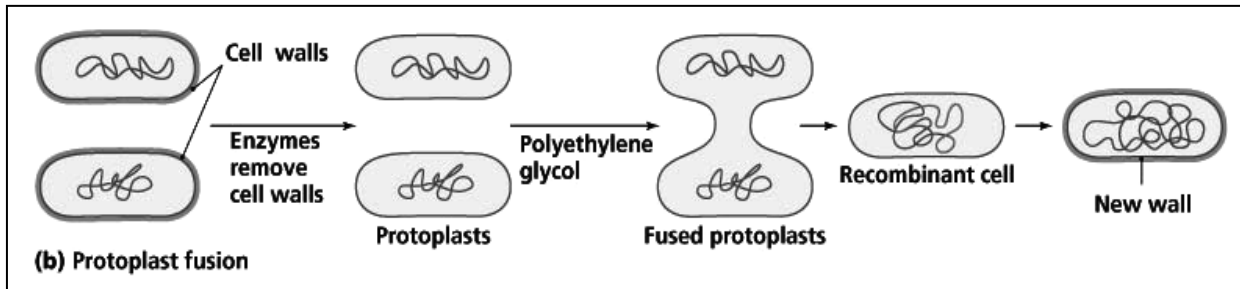
iv. Liposome-mediated gene transfer:

Liposomes are positively charged, circular lipid molecules, with an aqueous interior that can carry nucleic acids. Lipofection refers to the technique of encapsulating the DNA fragments in liposomes. These liposomes can adhere to cell membranes and fuse with them to transfer DNA fragments into the cells and then into the nucleus.

2. DIRECT GENE TRANSFER METHODS

The direct gene transfer methods include protoplast transformation, tissue/cell electroporation, silicon carbide fiber vortexing, and microprojectile bombardment.

i. Protoplast transformation was the first method used to demonstrate that direct gene delivery to plants was feasible. But, this method poses some technical challenges and hence is not widely used.

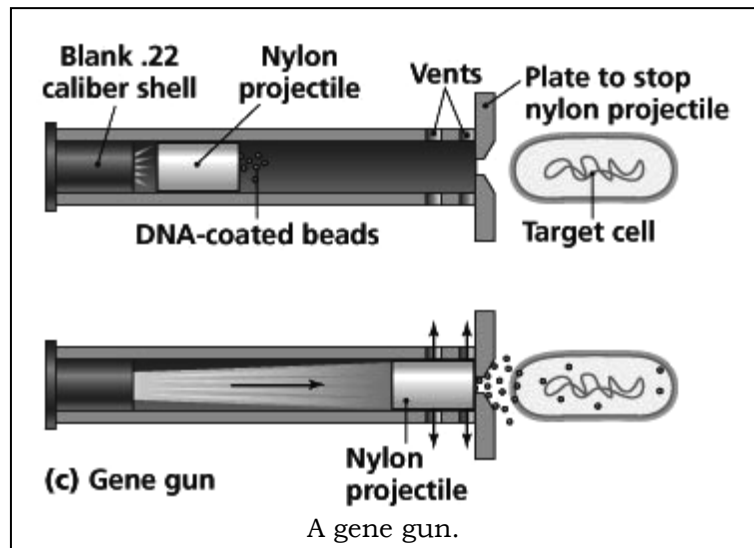


ii. Tissue/cell electroporation: Callus cultures or primary explants such as immature embryos or inflorescence may be used as target material. The transformation process occurs in an electroporator. Transformation efficiency levels by electroporation are sufficiently high.

iii. Silicon carbide fiber vortexing: Silicon carbide fibers are mixed with a suspension culture as explant along with plasmid DNA, and vortexed. The mixture is then cultured on a medium with selectable markers. To use this technique, there must be a regeneration system in place for regenerating plants from single cells.

iv. Microprojectile bombardment / Gene gun

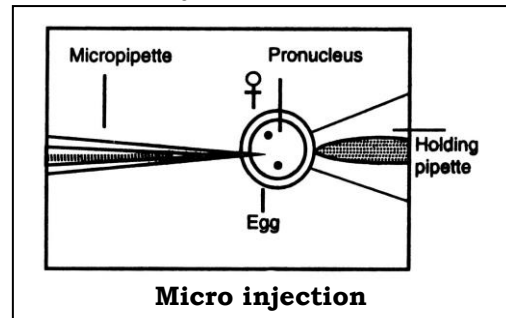
In this method the target DNA is coated in a carrier particle and transferred by shooting by the gene gun into the tissue. In this method 1 to 5 μ diameter tungsten or gold particles are coated with the DNA of interest. Compressed helium gas at the rate of 430m/s is used to propel the DNA coated particles through a barrel into the target cell. About 50 μ g of tungsten is required for each DNA transfer event. The carrier particles are placed on a support film and mounted in the particle acceleration device. The support film is accelerated by gas pressure and then stopped by a protective mesh. The carrier particles pass through the mesh, hitting the target tissue mounted in a petridish below the biolistic device. The survival of target cell is high when a low penetration number of projectiles (1 to 5 per cell) occurs.



Many factors such as chamber vacuum level, particle size range, and shot distance affect the success of a gene gun. Other factors include the amount of DNA per particle, explant type and physiological conditions, and gas type and pressure. Biolistics has been used to transform both dicots (e.g., soybean, peanuts, and tobacco) and monocots (e.g., corn, wheat, and rice). Organelle transformations have also been reported with this technique.

v. Micro-injection:

Microinjection refers to the process of injecting DNA directly into the cell, or even into the cell nucleus via an inserted cannula under a microscope. The target cell is held in place by two micromanipulators, one holding a pipette and one holding a microcapillary needle (0.5 to 5 μm diameter). The introduced DNA then integrates into the plant genome during its own DNA repair process. The main advantage is that the use of marker gene for identification of successful transformation is not needed.



BLUE WHITE SCREENING

A more sophisticated procedure for screening for the presence of recombinant screening plasmids, which can be carried out on a single transformation plate, is called blue–white screening. This method also involves the insertional inactivation of a gene and, as the name implies, uses the production of a blue compound as an indicator.

The blue–white screen is a screening technique that allows for the rapid and convenient detection of recombinant bacteria in vector-based molecular cloning experiments. DNA of interest is ligated into a vector. The vector is then inserted into a competent host cell viable for transformation, which are then grown in the presence of X-gal. Cells transformed with vectors containing recombinant DNA will produce white colonies; cells transformed with non-recombinant plasmids (i.e. only the vector) grow into blue colonies. This method of screening is usually performed using a suitable bacterial strain, but other organisms such as yeast may also be used.

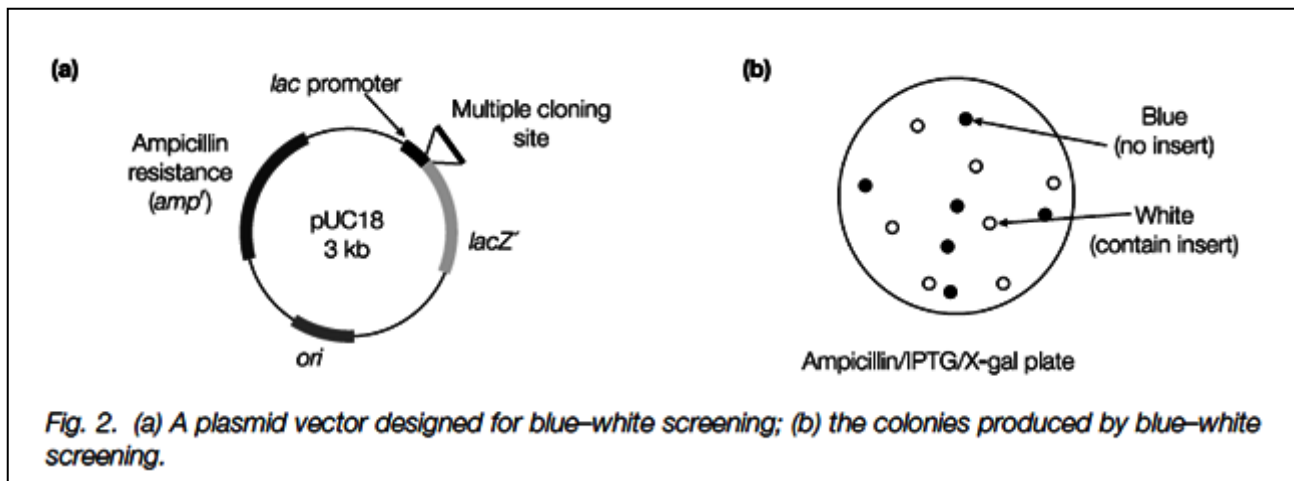


An LB agar plate showing the result of a blue white screen

Source: www.wikipedia.com

The gene in this case is *lacZ*, which encodes the enzyme β -galactosidase, and is under the control of the *lac* promoter (see Topic L1). If the host *E. coli* strain is expressing the *lac* repressor, then expression of a *lacZ* gene on the vector may be induced using isopropyl- β -D-thiogalactopyranoside (IPTG) (Topic L1), and the expressed enzyme can utilize the synthetic substrate 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-gal) to yield a blue product. Insertional inactivation of *lacZ* in the production of a recombinant plasmid would prevent the development of the blue color. In this method the transformed cells are spread on to a plate containing ampicillin (to select for transformants in the usual way), IPTG and X-gal, to yield a mixture of blue and white colonies. The white colonies have no expressed β -galactosidase and are hence likely to contain the inserted target fragment. The blue colonies probably contain religated vector.

In practice, the vectors used in this method have a shortened derivative of *lacZ*, *lacZ'*, which produces the N-terminal α -peptide of β -galactosidase. These vectors must be transformed into a special host strain which contains a mutant gene expressing only the C-terminal portion of β -galactosidase which can then complement the α -peptide to produce active enzyme. This reduces the size of the plasmid-borne gene, but does not alter the basis of the method.



IV. DNA SEQUENCING

Sequencing a genome requires breaking it up into small overlapping fragments whose DNA sequences can be determined in a sequencing reaction. The sequences can be ordered into the final genome sequence by a map-based approach (large fragments are ordered with the use of genetic and physical maps) or by whole-genome shotgun sequencing (overlap between the sequences of small fragments is compared by computers). DNA sequencing enables researchers to determine the base sequence of DNA found in genes and other chromosomal regions. It is one of the most important tools for exploring genetics at the molecular level.

During the 1970s, two methods for DNA sequencing were devised. One method, developed by Allan Maxam and Walter Gilbert, involves the base-specific chemical cleavage of DNA. Another method, developed by Frederick Sanger and colleagues, is known as dideoxy sequencing. Because it has become the more popular method of DNA sequencing, we consider the dideoxy method here.

The dideoxy procedure of DNA sequencing is based on DNA replication. DNA polymerase connects adjacent deoxyribonucleotides by forming a covalent bond between the 5 phosphate on one nucleotide and the 3' -OH group on the previous nucleotide. Chemists synthesize deoxyribonucleotides that are missing the -OH group at the 3' position. These synthetic nucleotides are called dideoxyribonucleotides (ddNTPs). The prefix *dideoxy-* indicates that two (*di*) oxygens (*oxy*) are removed (*de*) from this sugar compared with ribose.

Sanger reasoned that if a dideoxyribonucleotide is added to a growing DNA strand, the strand cannot grow because the dideoxyribo- nucleotide is missing the 3' -OH group. The incorporation of a dideoxyribonucleotide into a growing strand is therefore referred to as **chain termination**.

To detect the incorporation of dideoxynucleotides during DNA replication, the newly made DNA strands are labeled with a different colored fluorescent molecule: ddA is green, ddT is red, ddG is yellow, and ddC is blue.

Prior to automated DNA sequencing, the segment of DNA to be sequenced must usually be obtained in large amounts. This is accomplished by using gene cloning. For example the segment of DNA to be sequenced, called the target DNA, was cloned into a vector at a defined location. The target DNA was inserted next to a site in the vector where a primer will bind, which is called the primer-annealing site. The aim of the experiment is to determine the base sequence of the target DNA.

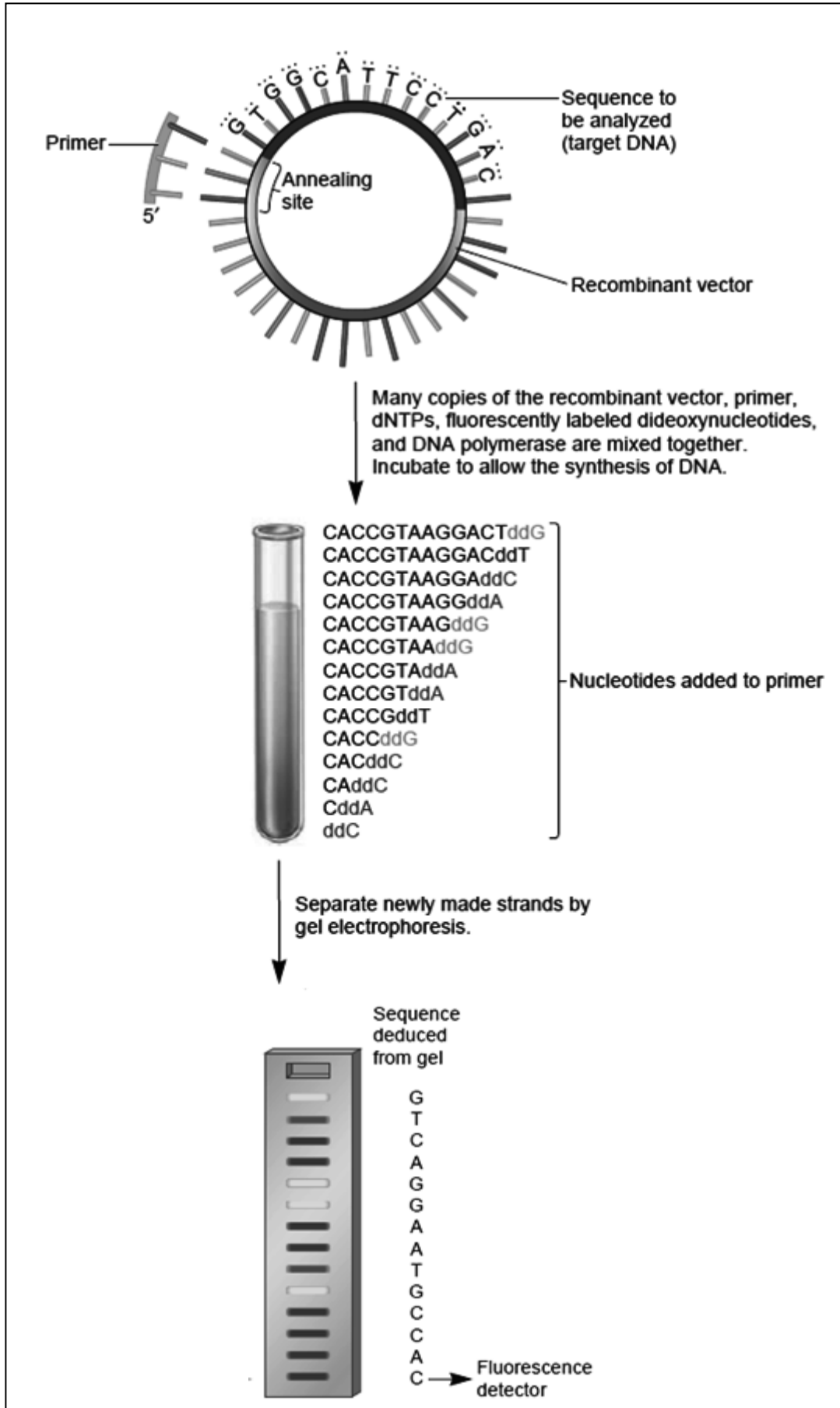
First, a sample containing many copies of the single-stranded DNA is mixed with many primers that will bind to the primer-annealing site. The primer binds to the DNA because the primer and primer-annealing site are complementary to each other. All four types of deoxyribonucleotides and DNA polymerase are then added to the annealed DNA fragments. The tube also has a low concentration of each dideoxyribonucleotide (ddG, ddA, ddT, or

ddC), which are fluorescently labeled. The tube is then incubated to allow DNA polymerase to make strands complementary to the target DNA sequence.

DNA synthesis continues until a dideoxynucleotide is incorporated into a growing strand. For example, chain termination can occasionally occur at the sixth or thirteenth position of the newly synthesized DNA strand if a ddT becomes incorporated at either of these sites. Note that the complementary A base is found at the sixth and thirteenth position in the target DNA. Therefore, we expect to make DNA strands that terminate at the sixth or thirteenth positions and have a ddT at their ends. Because these DNA strands contain a ddT, they are fluorescently labeled in red. Alternatively, ddA causes chain termination at the second, seventh, eighth, or eleventh positions because a complementary T base is found at the corresponding positions in the target strand. Strands that are terminated with ddA are fluorescently labeled in green.

After the samples have been incubated for several minutes, mixtures of DNA strands of different lengths are made, depending on the number of nucleotides attached to the primer. These DNA strands can be separated according to their lengths by running them on a slab gel or more commonly by running them through a gel-filled capillary tube. The shorter strands move to the bottom of the gel more quickly than the longer strands. Because we know the color of each dideoxyribonucleotide, we also know which base is at the very end of each DNA strand separated on the gel. Therefore, we can deduce the DNA sequence that is complementary to the target DNA by “reading” which base is at the end of every DNA strand and matching this sequence with the length of the strand. Reading the base sequence, from bottom to top, is much like climbing a ladder of bands. For this reason, the sequence obtained by this method is referred to as a sequencing ladder.

Theoretically, it is possible to read this sequence directly from the gel. From a practical perspective, however, it is faster and more efficient to automate the procedure using a laser and fluorescence detector. As the gel is running, each band passes the laser and the laser beam excites the fluorescent dye. The fluorescence detector records the amount of fluorescence emission from the excited dye. The detector reads the level of fluorescence at four wavelengths, corresponding to the four different colored dyes. An example of the printout from the fluorescence detector is shown in Figure 18.19b. As seen here, the peaks of fluorescence correspond to the DNA sequence that is complementary to the target DNA. Note that ddG is usually labeled with a yellow



TRANSGENIC PLANTS

Plants with desirable characters created through gene transfer methods are called transgenic plants. Though a number of methods have been developed to introduce the cloned genes into plant cells. Ti plasmid of *Agrobacterium tumefaciens* has been widely used as an effective vector for obtaining transgenic plants. Several transgenic plants have been produced to meet specific needs of agriculture and human life. Some of these are given below.

a. Transgenic crop plants having resistance to pathogens and pests:

1. Transgenic papaya is resistant to papaya ring spot virus
2. Bt. Cotton is resistant to insects.
3. Transgenic tomato plants are resistant to the bacterial pathogen *Pseudomonas*.
4. Transgenic potato plants are resistant to the fungus *Phytophthora*

b. Transgenic plants suitable for food processing technology:

Transgenic tomato 'Flavr Savr' is bruise resistant i.e., suitable for storage and transport due to delayed ripening and offers longer shelf-life.

c. Transgenic plants with improved nutritional value:

Transgenic golden rice obtained from 'Taipei' is rich in vitamin A and prevents blindness.

d. Transgenic plants useful for hybrid seed production:

Male sterile plants of *Brassica napus* are produced. This will eliminate the problem of manual emasculation and reduce the cost of hybrid seed production.

e. Transgenic plants tolerant to abiotic stresses caused by chemicals, cold, drought, salt, heat etc.

1. Basmati variety of rice was made resistant against biotic and abiotic stresses.
2. Round-up ready soybean is herbicide tolerant.

Besides these, genetically modified crops have evolved as alternative resources to industries, in the form of starches, fuels and pharmaceuticals.

Transgenic plants have been shown to express the genes of insulin, interferon, human growth hormones, antibiotics, antibodies etc. these biochemical produced by plants are as good as or sometimes better than those produced in bacteria.

Utilization of plants as biofactories or bioreactors for obtaining commercially useful products, specialized medicines, chemicals and antibodies on a large scale is described as molecular farming. In the near future this field is expected to revolutionise both the farming as well as biochemical industry.

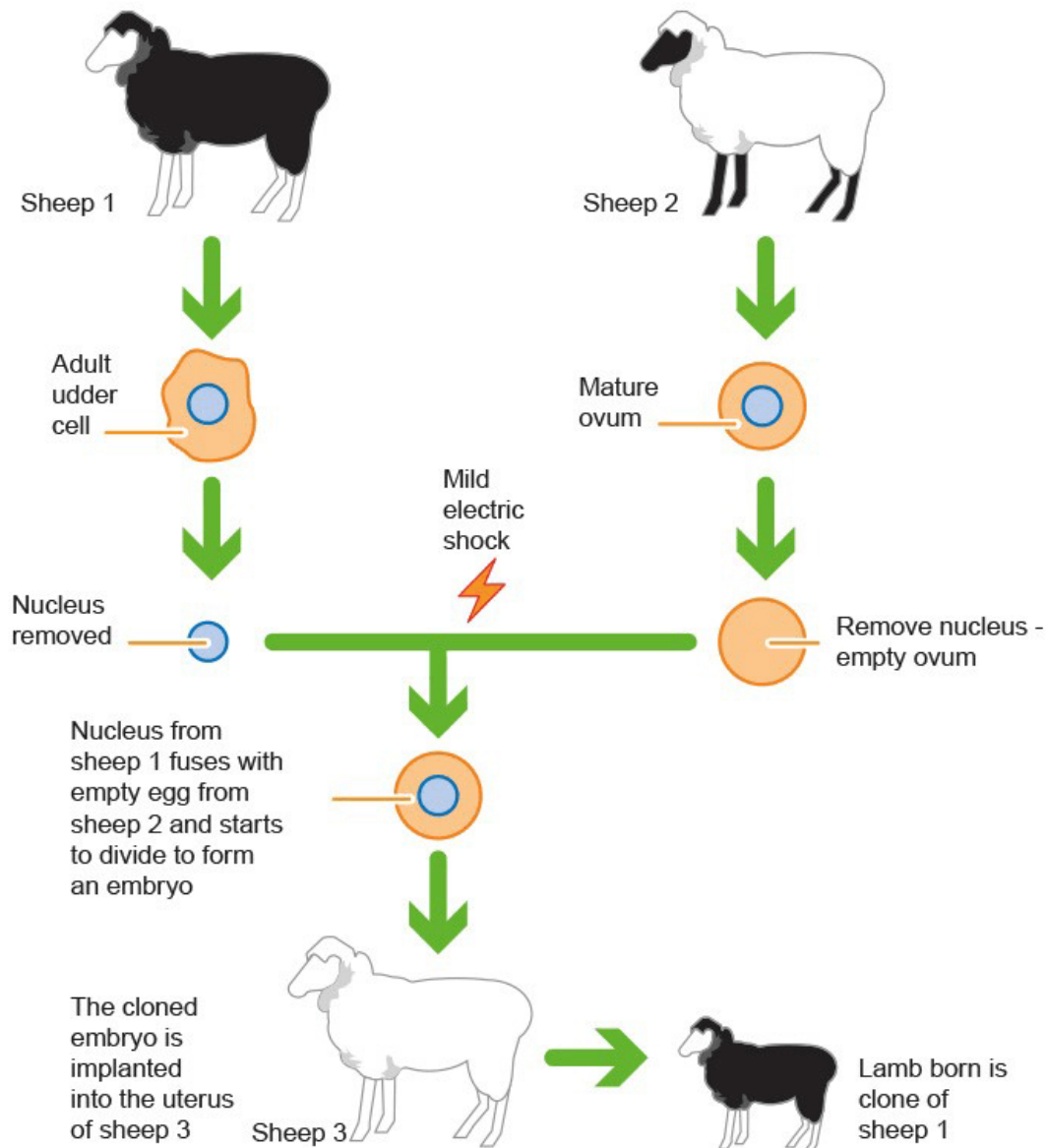
DOLLY- THE TRANSGENIC CLONE

Dolly, the first ever mammal clone was developed by Wilmut and Campbell in 1997. It is a sheep (female lamb) with a mother and no father.

The technique primarily involves nuclear transfer and the phenomenon of totipotency. The character of a cell to develop into different cells, tissues, organs, and finally an organism is referred to as totipotency or pluripotency. Totipotency is the basic character of embryonic cells. As the embryo develops, the cells specialize and finally give the whole organism. As such, the cells of an adult lack totipotency. Totipotency was induced into the adult cells for developing Dolly.

The cloning of sheep for producing Dolly, illustrated in fig. is briefly described here. The mammary gland cells from a donor ewe were isolated. They were subjected to total nutrient deprivation (starvation) for five days. By this process, the mammary cells abandon their normal growth cycle, enter a dormant stage and regain totipotency character. An ovum (egg cell) was taken from another ewe, and its nucleus was removed to form an enucleated ovum. The dormant mammary gland cell and the enucleated ovum were fused by pulse electricity. The mammary cell outer membrane was broken, allowing the ovum to envelop the nucleus. The fused cell, as it had gained totipotency, can multiply and develop into an embryo. This embryo was then implanted into another ewe which served as a surrogate/foster mother. Five months later, Dolly was born.

As reported by Wilmut and Campbell, they fused 277 ovum cells, achieved 13 pregnancies, and of these only one pregnancy resulted in live birth of the offspring-Dolly.



© ABPI 2017

CLONING OF PET ANIMALS

Some of the companies involved in transgenic experiments have started cloning pet animals like cats and dogs. Little Nicky was the first pet cat that was cloned at a cost of \$50,00 by an American company (in Dec. 2004). More cloned cats and dogs will be made available to interested parties (who can afford) in due course.

Some people who own pet animals are interested to continue the same pets which is possible through cloning. There is some opposition to this approach as the cloned animals are less healthy, and have shorter life span, besides the high cost factor.

GMOs IN AGRICULTURE

Genetically modified (GM) foods were first approved for human consumption in the United States in 1995, and by 1999 almost 50 percent of the corn, cotton, and soybeans planted in the United States were GM. By the end of 2010, GM crops covered more than 9.8 million square kilometres (3.8 million square miles) of land in 29 countries worldwide—one-tenth of the world's farmland.

GMOs in Agriculture

Bt crops

Engineered crops can dramatically increase per area crop yields and, in some cases, reduce the use of chemical insecticides. For example, the application of wide-spectrum insecticides declined in many areas growing plants, such as potatoes, cotton, and corn, that were endowed with a gene from the bacterium *Bacillus thuringiensis*, which produces a natural insecticide called Bt toxin. Field studies conducted in India in which Bt cotton was compared with non-Bt cotton demonstrated a 30–80 percent increase in yield from the GM crop. This increase was attributed to marked improvement in the GM plants' ability to overcome bollworm infestation, which was otherwise common. Studies of Bt cotton production in Arizona, U.S., demonstrated only small gains in yield—about 5 percent—with an estimated cost reduction of \$25–65 (USD) per acre due to decreased pesticide applications. In China, a seven-year study of farms planting Bt cotton demonstrated initial success of the GM crop, with farmers who had planted Bt cotton reducing their pesticide use by 70 percent and increasing their earnings by 36 percent. However, after four years, the benefits of Bt cotton eroded as populations of insect pests other than bollworm increased, and farmers once again were forced to spray broad-spectrum pesticides. While the problem was not Bt-resistant bollworms, as had been feared initially, it nonetheless became clear that much more research was needed for communities to realize sustainable and environmentally responsible benefits from planting GM crops.

Other GM plants were engineered for resistance to a specific chemical herbicide, rather than resistance to a natural predator or pest. Herbicide-resistant crops (HRC) have been available since the mid-1980s; these crops enable effective chemical control of weeds, since only the HRC plants can survive in fields treated with the corresponding herbicide. However, because these crops encourage increased application of chemicals to the soil, rather than decreased application, they remain controversial with regard to their environmental impact.

By 2002 more than 60 percent of processed foods consumed in the United States contained at least some GM ingredients. Despite the concerns of some consumer groups, especially in Europe, numerous scientific panels, including the U.S. Food and Drug Administration, have concluded that consumption of GM foods is safe, even in cases involving GM foods with genetic material from very distantly related organisms. Indeed, foods containing GM ingredients do not require special labeling in the United States, although some groups have continued to lobby to change this ruling. By 2006, although the majority of GM crops were still grown in the Americas, GM plants tailored for production

and consumption in other parts of the world were in field tests. For example, sweet potatoes intended for Africa were modified for resistance to sweet potato feathery mottle virus (SPFMV) by inserting into the sweet potato genome a gene encoding a viral coat protein from the strain of virus that causes SPFMV. The premise for this modification was based on earlier studies in other plants such as tobacco in which introduction of viral coat proteins rendered plants resistant to the virus.

Golden Rice (Ingo Potrykus)

The “golden” rice intended for Asia was genetically modified to produce almost 20 times the beta-carotene of previous varieties. Golden rice was created by modifying the rice genome to include a gene from the daffodil *Narcissus pseudonarcissus* that produces an enzyme known as phytylene synthase and a gene from the bacterium *Erwinia uredovora* that produces an enzyme called phytylene desaturase. The introduction of these genes enabled beta-carotene, which is converted to vitamin A in the human liver, to accumulate in the rice endosperm—the edible part of the rice plant—thereby increasing the amount of beta-carotene available for vitamin A synthesis in the body.





Golden rice (right) compared to white rice (left) (Source:

www.wikipedia.com)

Bio-fortification

Another form of modified rice was generated to help combat iron deficiency, which impacts close to 30 percent of the world population. This GM crop was engineered by introducing into the rice genome a ferritin gene from the common bean, *Phaseolus vulgaris*, that produces a protein capable of binding iron, as well as a gene from the fungus *Aspergillus fumigatus* that produces an enzyme capable of digesting compounds that increase iron bioavailability via digestion of phytate (an inhibitor of iron absorption). The iron-fortified GM rice was engineered to overexpress an existing rice gene that produces a cysteine-rich metallothioneinlike (metal-binding) protein that enhances iron absorption. A variety of other crops modified to endure the weather extremes common in other parts of the globe are also in production.

GMOs in medicine and research / Bio-farming / Biopharmaceuticals

GMOs have emerged as one of the mainstays of biomedical research since the 1980s. For example, GM animal models of human genetic diseases enabled researchers to test novel therapies and to explore the roles of candidate risk factors and modifiers of disease outcome. GM microbes, plants, and animals also revolutionized the production of complex pharmaceuticals by enabling the generation of safer and cheaper vaccines and therapeutics. Pharmaceutical products range from recombinant hepatitis B vaccine produced by GM baker's yeast to injectable insulin (for DIABETICS 🇨🇪) produced in GM *Escherichia coli* bacteria and to factor VIII (for hemophiliacs) and tissue plasminogen activator (tPA, for heart attack or strokepatients), both of which are produced in GM mammalian cells grown in laboratory culture. Furthermore, GM plants that produce "edible vaccines" are under development. Such plants, which are engineered to express antigens derived from microbes or parasites that infect the digestive tract, might someday offer a safe, cheap, and painless way to provide vaccines worldwide, without concern for the availability of refrigeration or sterile needles. Novel DNA vaccines may be useful in the struggle to prevent diseases that have proved resistant to traditional vaccination approaches, including HIV/AIDS, tuberculosis, and cancer.

GM Mosquitoes

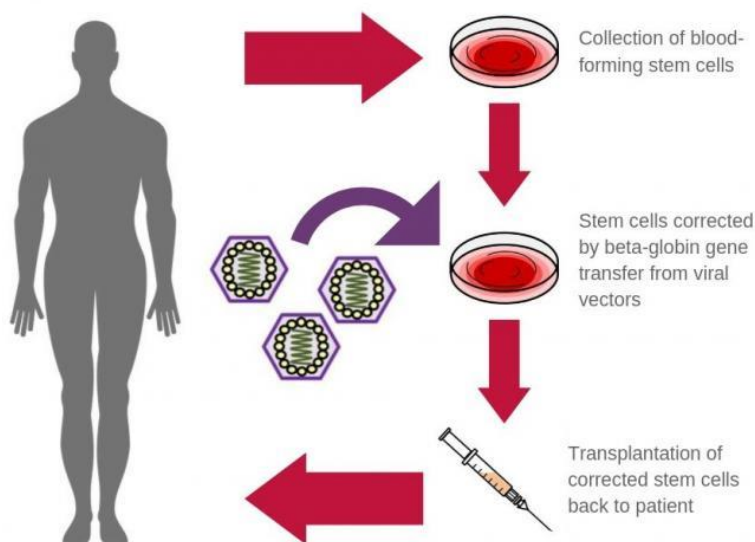
Genetic modification of insects has become an important area of research, especially in the struggle to prevent parasitic diseases. For example, GM mosquitoes have been developed that express a small protein called SM1, which blocks entry of

the malaria parasite, *Plasmodium*, into the mosquito's gut. This results in the disruption of the parasite's life cycle and renders the mosquito malaria-resistant. Introduction of these GM mosquitoes into the wild may someday help eradicate transmission of the malaria parasite without widespread use of harmful chemicals such as DDT or disruption of the normal food chain.

Gene therapy

Finally, genetic modification of humans, or so-called gene therapy, is becoming a treatment option for diseases ranging from rare metabolic disorders to cancer. Coupling stem cell technology with recombinant DNA methods may someday allow stem cells derived from a patient to be modified in the laboratory to introduce a desired gene. For example, a normal beta-globin gene may be introduced into the DNA of bone marrow-derived hematopoietic stem cells from a patient with sickle cell anemia, and introduction of these GM cells into the patient could cure the disease without the need for a matched donor.

Gene therapy for sickle cell disease



Role of GMOs in environmental management

Another application of GMOs is in the management of environmental issues. For example, some bacteria can produce biodegradable plastics, and the transfer of this ability to microbes that can be easily grown in the laboratory may enable the wide-scale "greening" of the plastics industry. Zeneca, a British company, developed a microbially produced biodegradable plastic called Biopol. This plastic is made using a GM bacterium, *Ralstonia eutropha*, to convert glucose and a variety of organic acids into a flexible polymer. GMOs endowed with the bacterially encoded ability to metabolize oil and heavy metals may provide efficient bioremediation strategies.

Genetic modification technologies may help save endangered species such as the giant panda, whose genome is being sequenced in an international effort led by the Beijing Genomics Institute at Shenzhen. Genetic studies of the panda genome may provide insight into why pandas have such low rates of reproductive success in captivity. A likely set of genes to consider for future genetic modification, should the goals of panda conservation warrant it, is the major histocompatibility complex (MHC). The MHC genes play an important role in regulating immune function and also influence behaviours and physiological patterns associated with reproduction.

Sociopolitical relevance of GMOs

While GMOs offer many potential benefits to society, the potential risks associated with them have fueled controversy, especially in the food industry. Many skeptics warn about the dangers that GM crops may pose to human HEALTH. For example, genetic manipulation may potentially alter the allergenic properties of crops. However, the more-established risk involves the potential spread of engineered crop genes to native flora and the possible evolution of insecticide-resistant “superbugs.” In 1998 the European Union (EU) addressed such concerns by implementing strict GMO labeling laws and a moratorium on the growth and import of GM crops. In addition, the stance of the EU on GM crops has led to TRADE disputes with the United States, which, by comparison, has accepted GM foods very openly. Other countries, such as Canada, China, Argentina, and Australia, also have open policies on GM foods, but some African states have rejected international food aid containing GM crops.

The use of GMOs in medicine and research has produced a debate that is more philosophical in nature. For example, while genetic researchers believe they are working to cure disease and ameliorate suffering, many people worry that current gene therapy approaches may one day be applied to produce “designer” children or to lengthen the natural human life span. Similar to many other technologies, gene therapy and the production and application of GMOs can be used to address and resolve complicated scientific, medical, and environmental issues, but they must be used wisely.

BIO-SAFETY AND ETHICAL ISSUES

Despite several advantages, genetic modifications of organisms can have unpredictable results when they are introduced into the natural ecosystem. Some of the apprehensions towards bio-safety issues of genetically engineered crops are:

The major concern about GM crops and GM foods are

1. There is fear of transfer of allergens from genetically modified food to humans and animals.
2. Due to molecular farming, there is a risk of changing the fundamental nature of vegetables.
3. The GM crops are manipulated artificially and are not naturally evolved which may affect their buffering capacity under adverse conditions.
4. GM crops pose a threat on biodiversity due to monoculture and impact environment.
5. There is a risk of gene pollution, which may result in the development of super-weeds.
6. Plants generally adapt the fluctuations occurring in nature and evolve gradually. GM plants may bring about changes in natural evolutionary pattern.
7. Use of herbicide-tolerant transgenic crops can lead to transfer of herbicide tolerance genes to sexually compatible wild relatives or weeds, and can create “super weeds”.
8. It would actually increase the dependence on a few herbicides rather than reducing herbicide usage.
9. Gene flow is the primary risk in releasing transgenic plants. Once released, the GM crops cannot be recalled back due to gene flow.
 10. There is fear of transferring allergens or toxins to humans and animals as side effects.

There is a risk of changing the fundamental nature of vegetables.

1. They may pose a harmful effect on biodiversity and have an adverse impact on environment.
2. There is a risk of gene pollution due to transfer of the new genes into related wild species through natural out-crossing. This may result in the development of super-weeds which may be fast-growing than the crops and may be resistant to weedicides.
3. They may bring about changes in natural evolutionary pattern.

Going beyond the morality of such issues, their biological significance is also important. The manipulation of living organisms by the human race cannot go on any further without regulation. Some ethical standards are required to evaluate the morality of all human activities that might help or harm living organisms. Therefore, the Indian Government has set up organisations such as GEAC (Genetic Engineering Approval Committee), which will make decisions regarding the validity of GM research and the safety of introducing GM-organisms for public services.

The modification/usage of living organisms for public services (as food and medicine sources, for example) has also created problems with patents granted for the same.

There is growing public anger that certain companies are being granted patents for products and technologies that make use of genetic materials, plants and other biological resources that have long been identified, developed and used by farmers and indigenous people of a specific region/country.

For example, rice is an important food grain, the presence of which goes back to thousands of years in Asia's agricultural history. There are an estimated 200,000 varieties of rice in India alone. The diversity of rice in India is one of the richest in the world. Basmati rice is distinct for its unique aroma and flavor and 27 documented varieties of Basmati are grown in India. There is reference to Basmati in ancient texts, folklore and poetry, as it has been grown for centuries. In 1997, an American company got patent rights on Basmati rice through the US Patent and Trademark Office. This allowed the company to sell a 'new' variety of Basmati, in the US and abroad. This 'new' variety of Basmati rice had actually been derived from Indian farmers' varieties. Indian Basmati was crossed with semi-dwarf varieties and claimed as an invention or a novelty. The patent extends to functional equivalents, implying that other people selling Basmati rice could be restricted by the patent. Several attempts have also been made to patent users, products and processes based on Indian traditional herbal medicines, such as turmeric and neem. If we are not vigilant and do not immediately counter these patent applications, other countries/individuals may encash on our rich legacy and we may not be able to do anything about it.

BIOSAFETY GUIDELINES AND REGULATION

Cultivation of genetically modified crops by the farmers is increasing rapidly throughout the world. In spite of impressive progress in this field, there is much uneasiness among the public towards biosafety issues of genetically engineered crops.

Due to the alarming concerns of these GMCs in Indian agriculture, the Govt. of India has evolved recombinant DNA safety guidelines for the manufacture, use, import, export and storage of hazardous microorganisms / genetically engineered organisms cells etc. These guidelines are being implemented through the Environmental Protection Act 1986 (EPA).

1. Institutional Biosafety Committees (IBCS) monitor research activity at institutional level.
2. Review Committee on Genetic Manipulation (RCGM) will monitor research activities in the laboratories and
3. Genetic Engineering Approval Committee (GEAC) of the Ministry of Environment and Forest has the power to permit large scale use of GMO's at commercial level and also monitors field trials of transgenic materials including agricultural crops, industrial products, health care products etc.

BIOPIRACY

Biopiracy is the term coined to refer to the use of bio-resources by multinational companies and other organizations without proper authorization from the countries and people concerned or without compensatory payment.

Most of the industrialized nations are rich financially but poor in biodiversity and traditional knowledge. In contrast, the developing and underdeveloped world is rich in biodiversity and traditional knowledge related to bio-resources. Traditional knowledge related to bi-resources can be exploited to develop modern applications and can also be used to save time, effort and expenditure during their commercialization.

There has been growing realization of the injustice, inadequate compensation and benefit sharing between developed and developing countries. Therefore, some nations are developing laws to prevent such unauthorized exploitation of their bio-resources and traditional knowledge.

The Indian Parliament has recently cleared the second amendment of the Indian Patents Bill, that takes such issues into consideration, including patent terms emergency provisions and research and development initiatives.

BIOREACTOR

Obtaining the Foreign Gene Product:

When a piece of alien DNA is inserted into a cloning vector and transferred into a host cell, the alien DNA gets multiplied. The ultimate aim is to produce a desirable protein. If any protein encoding gene is expressed in a heterologous host, it is called as a recombinant protein.

Any new biotechnological manufacturing process must first be tried on a laboratory scale. It is useful to make a pilot plan first and then to scale it up for industrial production.

A bioreactor may refer to any device or system that supports a biologically active environment. In one case, a bioreactor is a vessel in which is carried out a chemical process which involves organisms or biochemically active substances derived from such organisms. This chemical process carried out may be aerobic or anaerobic in nature. The bioreactors are frequently cylindrical in shape. They may vary in size. The body of a bioreactor is usually made up of stainless steel.

The bioreactor is also used to grow cell or tissues. The process of growing cells or tissues is employed in plant cell/tissue culture. An another field of the application of the bioreactor is the field of tissue engineering.

The classification of the bioreactors is based on their mechanism of working. They may be (a) *batch*, (b) *fed batch* or (c) *continuous*.

One of the biggest problems is to maintain the optimum conditions within the bioreactors and to maintain proper asepsis throughout the process. Any derivation from the optimum conditions and any contamination of the bioreactor would lead to a sub-optimum yield or no yield at all. Thus bioreactors have been considerably improved upon for ensuring these conditions.

Fouling of the bioreactor can decrease the efficiency of the bioreactors considerably. They may be especially detrimental to the heat exchangers that play a very important role in the bioreactors. During the designing, due care is taken to ensure that the surface of the bioreactors is smooth. Periodic cleaning of the bioreactors is mandatory.

A heat exchanger is an important component of bioreactors and it maintains the constant temperature required to carry out the process in the bioreactor. A large amount of heat can be produced in the fermentation carried out in a bioreactor. Therefore, to maintain a constant temperature, a bioreactor commonly employs a source of refrigeration. The refrigeration can be provided by an external mechanism such as an external jacket or by internal mechanisms such as internal cooling coils.

If a bioreactor is carrying out an aerobic biological process, adequate and uniform delivery of the oxygen is a major problem encountered. The delivery of oxygen is made difficult by the fact that oxygen is poorly soluble in water. The medium may be agitated to ensure uniform distribution of oxygen. Baffles are used for proper mixing of the medium. But this method is not very effective for oxygen transfer. A sparger (a perforated tube) is more commonly used to ensure adequate oxygen supply.

As the process of agitation may be detrimental to the microorganisms, care is taken in employing bacteria or other simple organisms that can withstand the forces of agitation. It is prudent to use organisms that have minimal nutritional requirements and those with a higher growth rates.

Downstream processing

It is referred to the stage of purifying and recovering the product from the bioreactor.

FERMENTER OR BIOREACTOR

Fermenter is a vessel in which the microorganisms are grown. It is a vessel designed to carryout fermentation process i.e. biological reactions under the controlled conditions, Hence it is called as bioreactor. Several criteria should be taken into account for designing a fermentor. They are

1. Long term operation in aseptic condition
2. Adequate aeration and agitation
3. pH controlling system
4. Sampling facility
5. Temperature controlling system
6. Minimum labour in operation, harvesting, cleaning and maintenance
7. Suitable for a variety of process
8. Fermentor is provided with limited amount of medium containing all the materials at optimum environmental condition

FERMENTATION PROCESS

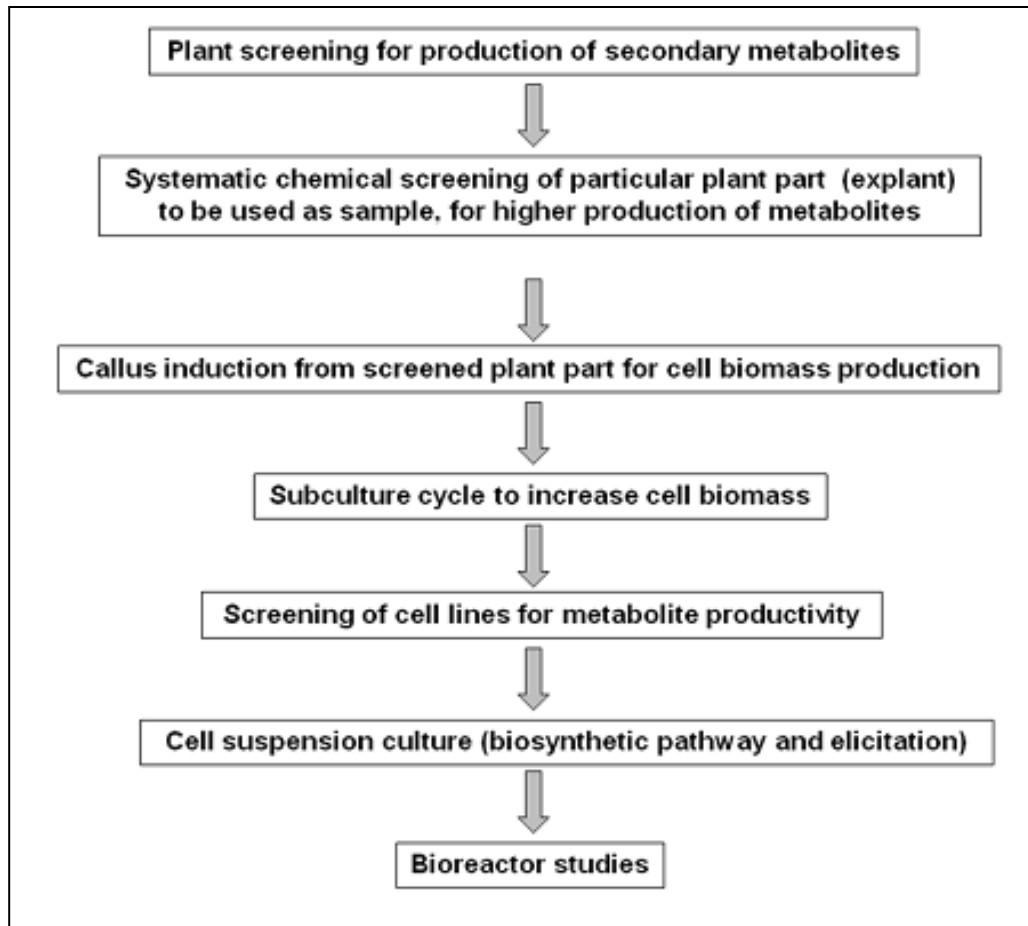
The word fermentation is derived from a latin verb “Fervete” means to boil. The definition of fermentation is the breakdown of larger molecules (metabolism). For example carbohydrates are breakdown into simple ones by micro organisms for their enzymes. In a microbiological way, fermentation is defined as “any process for the production of useful products through mass production of microorganisms. In a biochemical sense, the fermentation means the several oxidation-reduction reactions in which organic compounds are used as source of carbon and energy, act as acceptors or donars of hydrogen ions.

ANIMAL BIOREACTORS

Transgenesis is wonderfully utilized for production of proteins for pharmaceutical and medical use. In fact, any protein synthesized in the human body can be made in the transgenic animals, provided that the genes are correctly programmed. The advantage with transgenic animals is to produce scarce human proteins in huge quantities. Thus, the animals serving as factories for production of biologically important products are referred to as animal bioreactors or sometimes pharm animals. Some transgenic animals that serve as bioreactors are listed

- Transgenic cow for the production of lactoferrin and interferons.
- Transgenic goat to synthesize tissue plasminogen activator, and antithrombin III

- Transgenic mouse for the production of immunoglobulins, and urokinase.
- Transgenic pig to produce hemoglobin.



ANTIBODIES

Antibodies are proteins or Immunoglobulins (Igs) stimulated by specific antigen made up of two heavy and two light chains creates two functionally important sites, one recognize and bind the antigen (Fab) and the Fc portion at the other end of molecule which mediate the function of antibody. Normally β -cells secrete antibodies in correspondence to antigens.

Individual antibody possesses at least two sites i.e. antigen binding site. The result of an antibody antigen interaction depends on the nature of antigen and antibody and may include

1. Enhanced phagocytosis
2. Complement activation
3. Neutralization of toxin
4. Inactivation of proteins
5. Inhibition on binding of toxins or bacteria to surface

STRUCTURE OF ANTIBODIES

Human antibodies have two binding sites therefore, they are bivalent which is also called as monomer because it is the simplest antibody. Two monomers are interconnected by joining (J) chain. Similarly in a pentamer Ig molecules five monomers are held in position by a J-chain.

There are several types of antibody molecule, which differ in structure and function. There are five classes

- a. Ig G – gamma
- b. Ig M – Mu
- c. Ig A – Alpha
- d. Ig D – Delta
- e. Ig E – Epsilon

ANTIGEN

The antigens (Ag) or Immunogen is a large organic molecule capable of stimulating the production of specific antibody with which it may chemically combine. The ability of antigens to induce antibody formation is known as antigenicity.

Nature and properties of Antigens:

1. Chemically different types of antigens are found
 - a) Proteins
 - b) Nucleo proteins (Nucleic acid + protein)
 - c) Lipo proteins (lipid + protein)
 - d) Glycoproteins (carbohydrate + protein)
 - e) Large polysaccharides
2. Many antigens possess different types of determinants (site in which antibody binds) on their surface is the immune system may produce several antibodies against a single antigen.

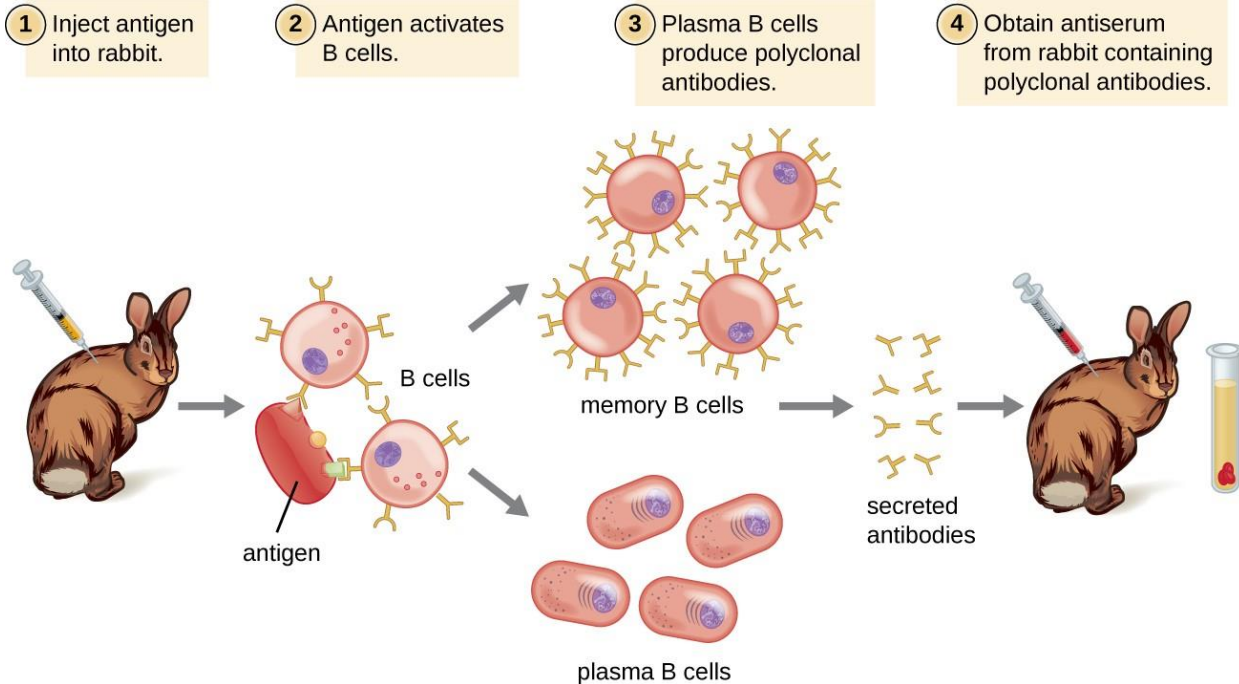
3. Most of antigens have molecular weight of 10000 daltons or more.

POLYCLONAL AND MONOCLONAL ANTIBODIES

POLYCLONAL ANTIBODIES

A preparation of antibody molecules that arises from several different clones of cells is called a polyclonal antibody. It is a mixture of antibody molecules that bind to different parts of the antigen and with different binding affinities.




If an antigen is injected into an animal, a number of antibody-producing cells will bind that antigen, albeit with varying degrees of affinity, and so the antibody which appears in the bloodstream will have arisen from several clones of cells, that is it will be a polyclonal antibody. Different antibody molecules in a preparation of polyclonal antibody will bind to different parts of the macromolecular antigen and will do so with different binding affinities. The binding region recognized by any one antibody molecule is called an epitope. Most antibodies recognize particular surface structures in a protein rather than specific amino acid sequences (i.e. the epitopes are defined by the conformation of the protein antigen). A preparation of polyclonal antibodies will bind to many episodes on the protein antigen.



MONOCLONAL ANTIBODIES

Antibody produced by a single clone of cells is a monoclonal antibody; all the antibody molecules are identical and bind to the same antigenic site with identical

binding affinities. Monoclonal antibodies can be generated in large amounts by creating a cell fusion (called a hybridoma) between an antibody producing cell and a myeloma cell.

<h2 style="color: red;">Polyclonal Antibody</h2>	<h2 style="color: blue;">Monoclonal Antibody</h2>
<ul style="list-style-type: none">• Cheap to produce• Mixed population of antibodies• May bind to different areas of the target molecule• Tolerant of small changes in protein structure	<ul style="list-style-type: none">• Expensive to produce• Single antibody species• Will only bind single specific site• May recognise a particular protein form
<p>Polyclonal antibody</p> 	<p>Monoclonal antibody</p> 
	

If a single clone of antibody-producing cells could be isolated, then all of the antibody produced from that clone would be identical; all antibody molecules in such a monoclonal antibody preparation would bind of the same antigen epitope.

The problem is that if an individual anti-body producing cell is isolated and grown in culture, its descendants have a limited lifespan that severely limits their use for the routine preparation of monoclonal antibodies. In 1975, Milstein and Kohler discovered how monoclonal antibodies of almost any desired antigen specificity can be produced indefinitely and in large quantities. Their method was to fuse a B lymphocyte producing antibody of the desired specificity with a cell derived from a cancerous lymphocyte tumor, called a myeloma cell, which is immortal. The cell fusion is called a hybridoma, which is both immortal and secretes the same specific antibody originally encoded by the B lymphocyte.

Monoclonal antibodies produced using this technology are now common tools in research because of their very high specificity. For example, they can be used to locate particular molecules within cells or particular amino acid sequences within proteins. If they are first bound to an insoluble matrix, they are also extremely useful for binding to and hence purifying the particular molecule from crude cell extracts or fractions. They are also increasingly of use in medicine, both for

diagnosis and as therapeutic tools, for example to inactivate bacterial toxins and to treat certain forms of cancer.

HYBRIDOMA TECHNOLOGY

Conventional methods adopted in the laboratory for the production of antisera against antigens lead to the formation of heterogeneous antibodies. Among these antibodies a few may have the desired properties but are found with many other antibodies which undoubtedly are not required. A simple, convenient and desirable method for the large scale production of specific antibodies remained a dream for immunologists for a long period. In 1975, George Kohler and Cesar Milstein (Nobel Prize 1984) made this dream a reality. They created hybrid cells that will make unlimited quantities of antibodies with defined specificities, which are termed as monoclonal antibodies (McAb). This discovery, often referred to as hybridoma technology, has revolutionized methods for antibody production.

Principle

This is based on the fusion between myeloma cells (malignant plasma cells) and spleen cells from a suitably immunized animal. Spleen cells die in a short period under ordinary tissue culture conditions while myeloma cells are adopted to grow permanently in culture. Mutants of myeloma cells lack the enzyme hypoxanthine guanine phosphoribosyltransferase (azaquinine resistant) or thymidine kinase (bromodeoxyuridine resistant). These mutant cannot grow in a medium containing aminopterin, supplemented with hypoxanthine and thymidine (HAT medium). Hybrids between the mutant myeloma cells and spleen cells can be selected and cultured in HAT medium.

From the growing hybrids, individual clones can be chosen that secrete the desired antibodies (monoclonal origin). The selected clones like ordinary myeloma cells can be maintained indefinitely. In short, the hybridoma technology for the production of monoclonal antibodies involves the following steps.

1. Immunization of appropriate animals with antigen (need not be pure) under study.
2. Fusion of suitable drug resistant myeloma cells with plasma cells, obtained from the spleen of the immunized animal.
3. Selection and cloning of the hybrid cells that grow in culture and produce antibody molecules of desired class and specificity against the antigen of interest.

Hybridoma technology can make available highly specific antibodies in abundant amounts. The clones once developed are far cheaper than the traditionally employed animals(horses, rabbits) for producing antibodies. The clones developed

from the hybrids will also ensure constancy of the quality of the product and will also avoid the batch to batch variation inherent in the conventional methods.

APPLICATIONS OF MONOCLONAL ANTIBODIES

The antibodies produced by hybridoma technology have been widely used for a variety of purposes. These include the early detection of pregnancy, detection and treatment of cancer, diagnosis of leprosy and treatment of autoimmune disease.

ENZYME LINKED IMMUNOSORBANT ASSAY (ELISA)

Specific antibodies can also be used to quantify the amount of the corresponding antigen in a biological sample. Several types of immunological assays exist. An increasingly popular version is enzyme-linked immunosorbent assay (ELISA) see fig. which can readily detect and quantify less than a nanogram of a specific antigenic protein. In ELISA, the specific antibody is coupled to a solid support. A convenient format for ELISA is to use a plastic tray that has molded wells in it (a microtiter tray) where the antibody has been coupled to the plastic forming the wells. Samples to be assayed are added to the wells. If antigen is present that is recognized by the antibody, it becomes bound (fig). the wells are then washed to remove unbound protein and incubated with a second antibody that recognizes the protein but a different epitope than the first antibody (fig). The second antibody is attached to an enzyme that can catalyze the conversion of a colorless or nonfluorescent substrate into a colored or fluorescent product. The intensity of the color or fluorescence produced for each sample is then measured to determine the amount of antigen present in each sample. Several machines are now commercially available that scan the wells of microtiter plates following ELISA and quantify the amount of antigen bound in each well.

Enzyme-linked immunosorbent assay (ELISA) is a non-isotopic immunoassay. An enzyme is used as a label in ELISA in place of radioactive isotope employed in RIA. ELISA is as sensitive as or even more sensitive than RIA. In addition, there is no risk of radiation hazards (as is the case with RIA) in ELISA.

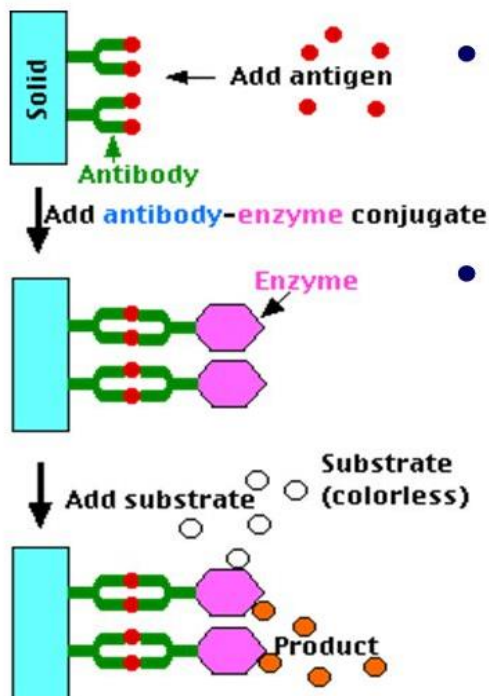
Principle

ELISA is based on the immunochemical principles of antigen-antibody reaction. The stages of ELISA, depicted in fig. are summarized.

1. The antibody against the protein to be determined is fixed on an inert solid such as polystyrene.
2. The biological sample containing the protein to be estimated is applied on the antibody coated surface.
3. The protein antibody complex is then reacted with a second protein specific antibody to which an enzyme is covalently linked. These enzymes must be easily assayable and produce preferably coloured products. Peroxidase, amylase and alkaline phosphatase are commonly used.
4. After washing the unbound antibody linked enzyme, the enzyme bound to the second antibody complex is assayed.
5. The enzyme activity is determined by this action on a substrate to form a product (usually coloured). This is related to the concentration of the protein being estimated.

The principle for the use of the enzyme peroxidase in ELISA is illustrated next.

Basic principles of ELISA



- the antibodies fixed to a solid surface, such as the inner surface of a test tube;
- a preparation of the same antibodies coupled to an enzyme. This is one (e.g., β -galactosidase) that produces a colored product from a colorless substrate.

Source: www.Wikipedia.com

Applications

ELISA is widely used for the determination of small quantities of proteins (hormones, antigens, antibodies) and other biological substances. The most commonly used pregnancy test for the detection of human chorionic gonadotropin

(hCG) in urine is based on ELISA. By this test, pregnancy can be detected within few days after conception. ELISA is also been used for the diagnosis of AIDS.

GENE TRANSFORMATION

AGROBACTERIUM MEDIATED GENE TRANSFER

The “Ti” plasmid, or tumor-inducing plasmid is the most commonly used plant-cloning vector used for cloning genes in plants. This plasmid is found in cells of the bacterium known as *Agrobacterium tumefaciens*. This bacterium normally lives in soil. *Agrobacterium* causes crown gall tumors in plants at the site of infection. The Ti plasmid is responsible for the development of the disease.

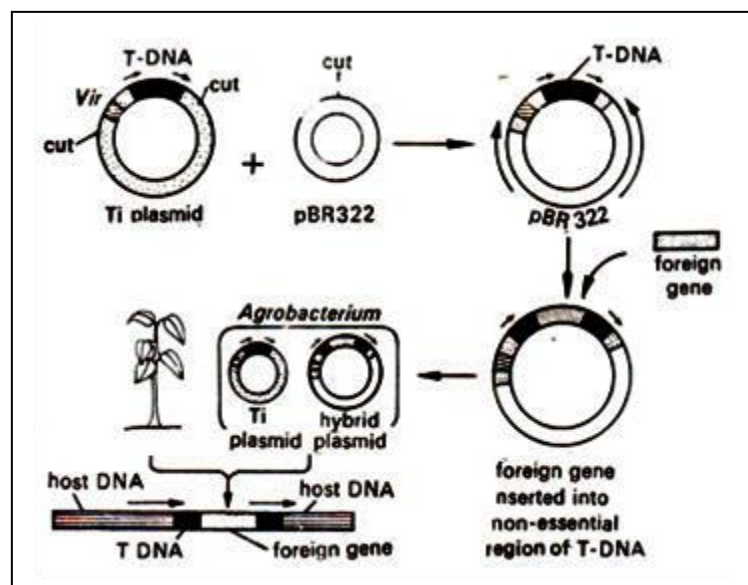
The Ti plasmid consists of a segment called as the T-DNA or the transforming DNA. When *Agrobacterium* infects a plant, it does not enter the cells of the plant. Instead, it lodges itself in the intercellular spaces. The T-DNA separates from the Ti plasmid and gets incorporated into the host genome.

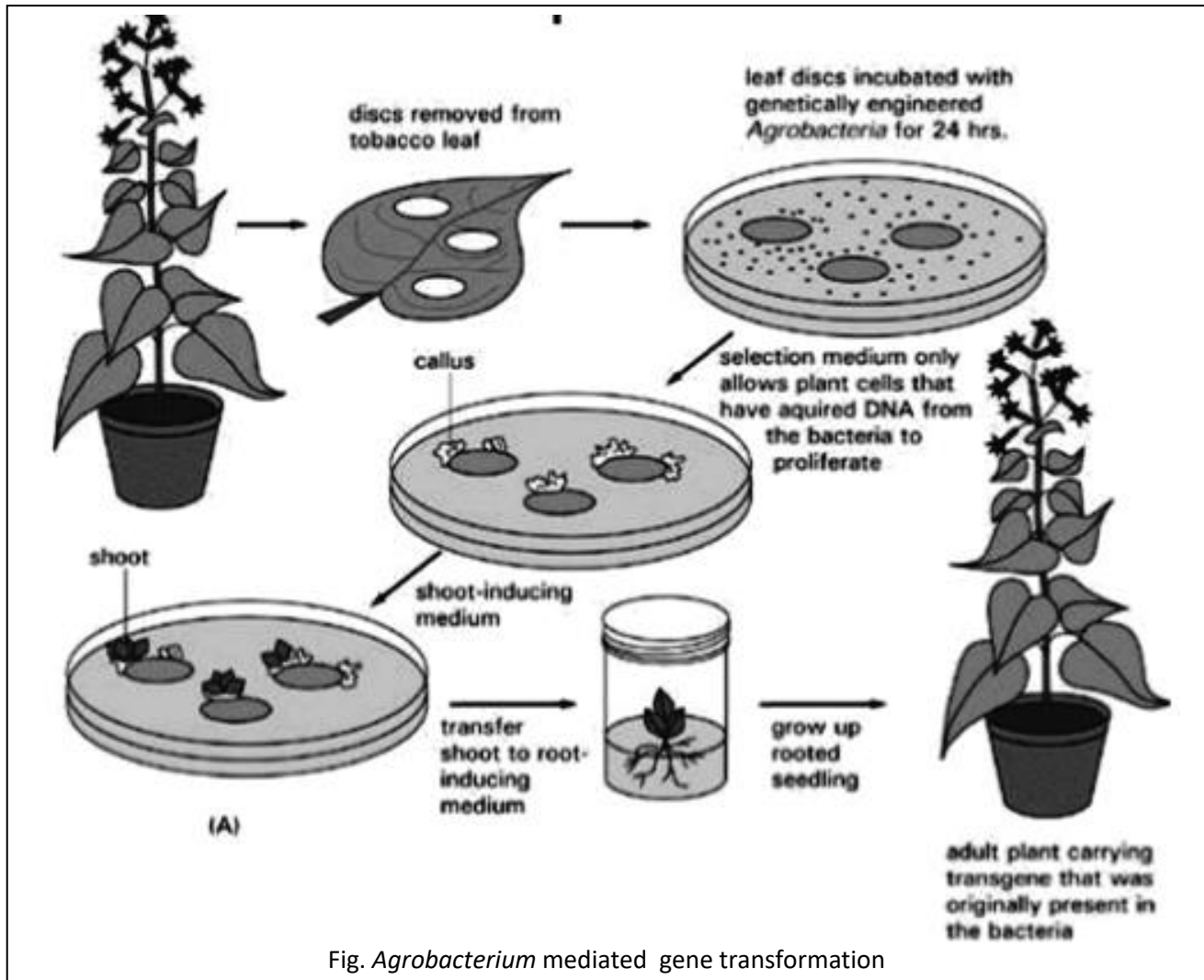
This property makes the Ti plasmid a useful plant-cloning vector and it is this property that is exploited to shuttle a foreign gene into the plant cells. The tumor causing genes of the Ti plasmid are excised so that the bacterium can no longer cause the disease. The desirable foreign gene that needs to be incorporated into the plant cells is inserted in the same place from where the disease causing genes were removed from the plasmid. The *Agrobacterium* is made to infect the plant by artificially creating an injury to the plant. The modified Ti plasmid, carrying the desirable exogenous genes and the intact T-DNA enters the protoplast of the plant cell.

The foreign DNA with the desirable genes and the T-DNA gets incorporated into the host genome. New plants can be generated through these engineered protoplasts. Each cell of the such generated plants will carry the desirable exogenous gene.

A major drawback of *Agrobacterium* was that in nature it infects only the dicots. Most of our important crops are monocots. Thus, it used to be difficult to modify our important crop plants.

Using the processes of microinjection, electroporation, particle bombardment and biolistics we can now insert the desirable exogenous genes into plant cell types that are not susceptible to *A. tumefaciens* transfection.

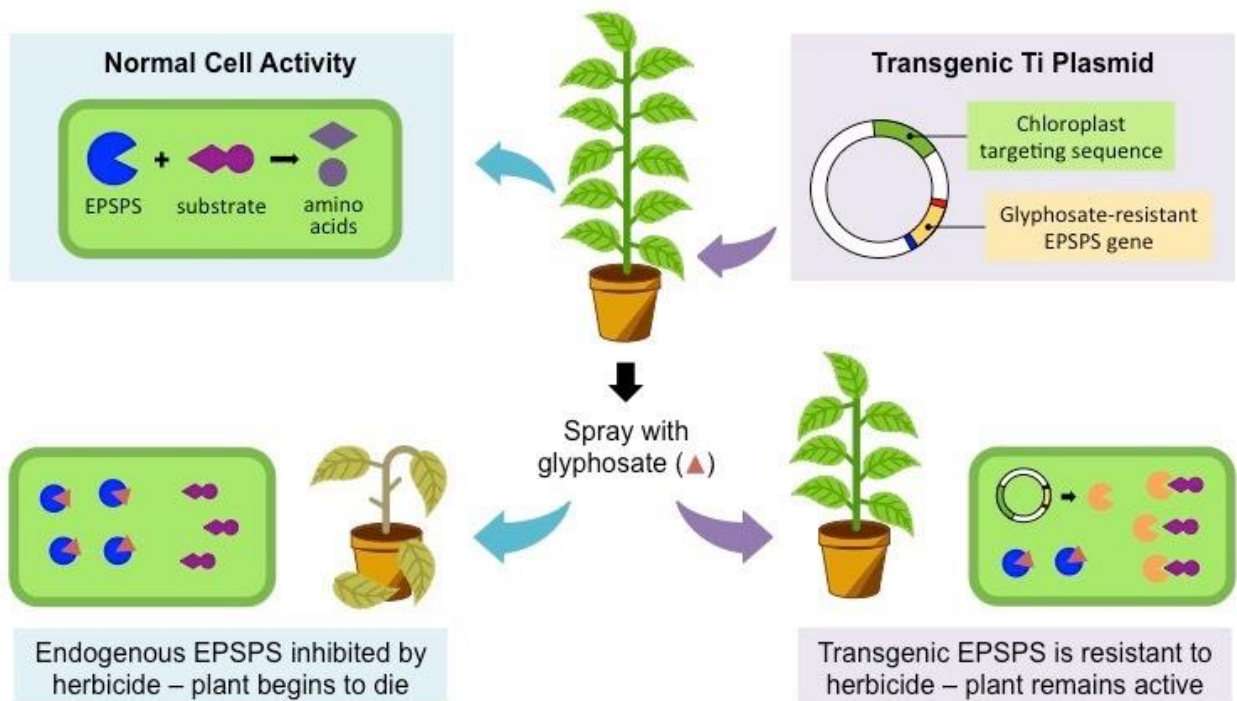


Fig. *Agrobacterium* mediated gene transformation

HERBICIDE TOLERANT CROPS

The most significant and widespread applications of recombinant DNA technology in practical plant breeding to date are the development of Bt cultivars and herbicide tolerance in plants. The soil-borne bacterium, *Bacillus thuringiensis*, is the source of the gene used in the development Bt products. The gene encodes the inactive form of a protein, Bt toxin, this is toxic to various herbivorous insects when ingested and converted to its toxic form (delta endotoxin) in the gut of the insect. Over 100 different variations of the Bt toxin have been identified, as well as a variety of associated target insect specificity. The toxins classified as Cry1a group target Lepidoptera or butterfly group, while the toxins in Cry3 group target beetles. Scientists have cloned the bacterial genes, which are then transferred into plants to provide resistance to target pest, thereby eliminating the need for pesticides. The major crops that have received such treatment include corn, cotton, and potato.

Researchers are pursuing additional naturally occurring insecticidal compounds as alternatives to the Bt technology. These included chitinase, lectins, alpha-amylase inhibitors, cystatin, and proteinase inhibitors.



Why engineer herbicide resistant crops?

A successful herbicide should destroy weeds only, leaving the economic plant unharmed. Broad-spectrum herbicides (non-selective) are attractive but their use in crop production can be problematic, especially in the production of broadleaf crops such as soybean and cotton. There is a general lack of herbicides that will discriminate between dicot weeds and crop plants.

Pre-plant applications may be practical to implement. However, once the crop is

established and too tall for the safe use of machinery, chemical pest management becomes impractical. Grass crops (e.g., wheat, corn) may tolerate broadleaf herbicides better than the reverse situation. Consequently, when cereal crops and broadleaf crops are grown in rotation or adjacent fields, the broadleaf plants are prone to damage from residual herbicides in the soil, or drift from herbicides applied to grasses. When a crop field is infested by weed species that are closely related to the crop (e.g., red rice in rice crop or nightshade in potato crop), herbicides lack sensitivity enough to distinguish between the plants.

To address these problems, one of two approaches may be pursued: (i) the development of new selective post emergent herbicides or (ii) genetic development of herbicide resistance in crops to existing broad-spectrum herbicides. The latter strategy would be advantageous to the agrochemical industry (increased market) and farmers (safer alternative of pesticides that are already in use). New herbicides are expensive to develop and take time.

Modes of action and herbicide resistance mechanisms

Most herbicides are designed to kill target plants by interrupting a metabolic stage in photosynthesis. Because all higher plants photosynthesize, most herbicides will kill both weeds and desirable plants. Plants resist phytotoxic compounds via one of several mechanisms:

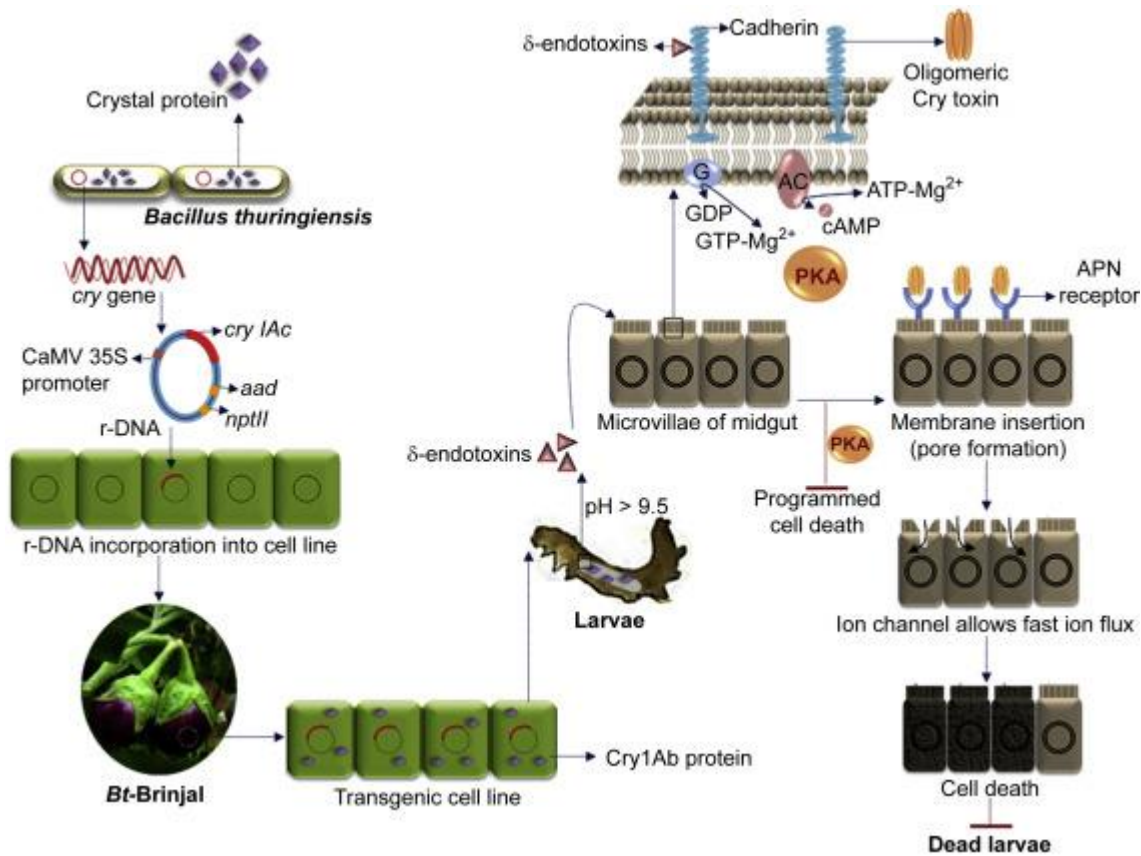
- The plant or cell does not take up toxic molecules because of external barriers such as cuticles.
- Toxic molecules are taken but sequestered in a sub-cellular compartment away from the target (e.g., protein) compounds the toxin was designed to attack.
- The plant or cell detoxifies the toxic compound by enzymatic processes into harmless compounds.
- The plant or cell equipped with resistance genes against the toxin may produce a modified target compound that is insensitive to the herbicide.
- The plant or cell overproduces the target compound for the phytotoxin in large amounts such that it would take a high concentration of the herbicide to overcome it.

Environmental impact

One of the complaints launched against the deployment of GM crops was that using plants as “pesticides” could have ecological consequence by being harmful to non-target organisms. This concern has not been significantly substantiated, at least in the case of Bt products. The tendency is to increase the dose rate of pesticides as pests develop resistance to them. However, this is counterproductive in that such an action only serves to intensify the selection pressure for resistant pests.

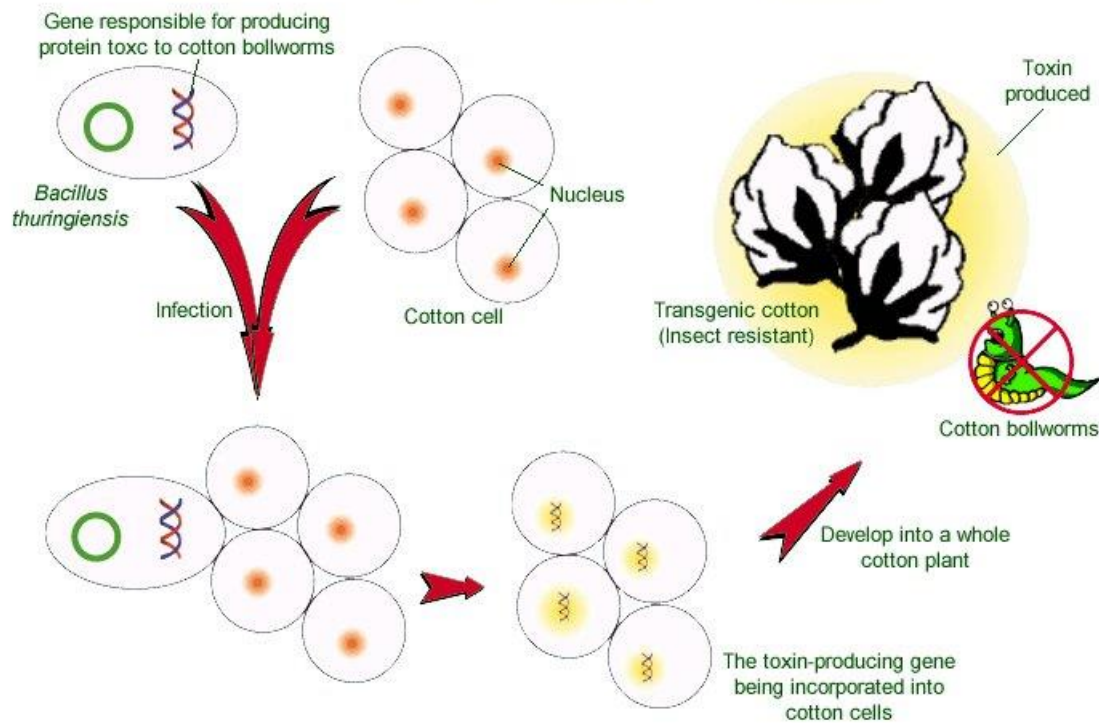
Pest resistance

Repeated and widespread use eventually results in resistance of pests to pesticides. The high rates of pesticide use create a situation for intense selective pressure that increases resistance to pesticides. Also, whereas older pesticides were designed to attack multiple sites in their target organism, modern pesticides are designed to be more specific in action (often one biochemical pathway), thereby increasing the chance for development of resistance.



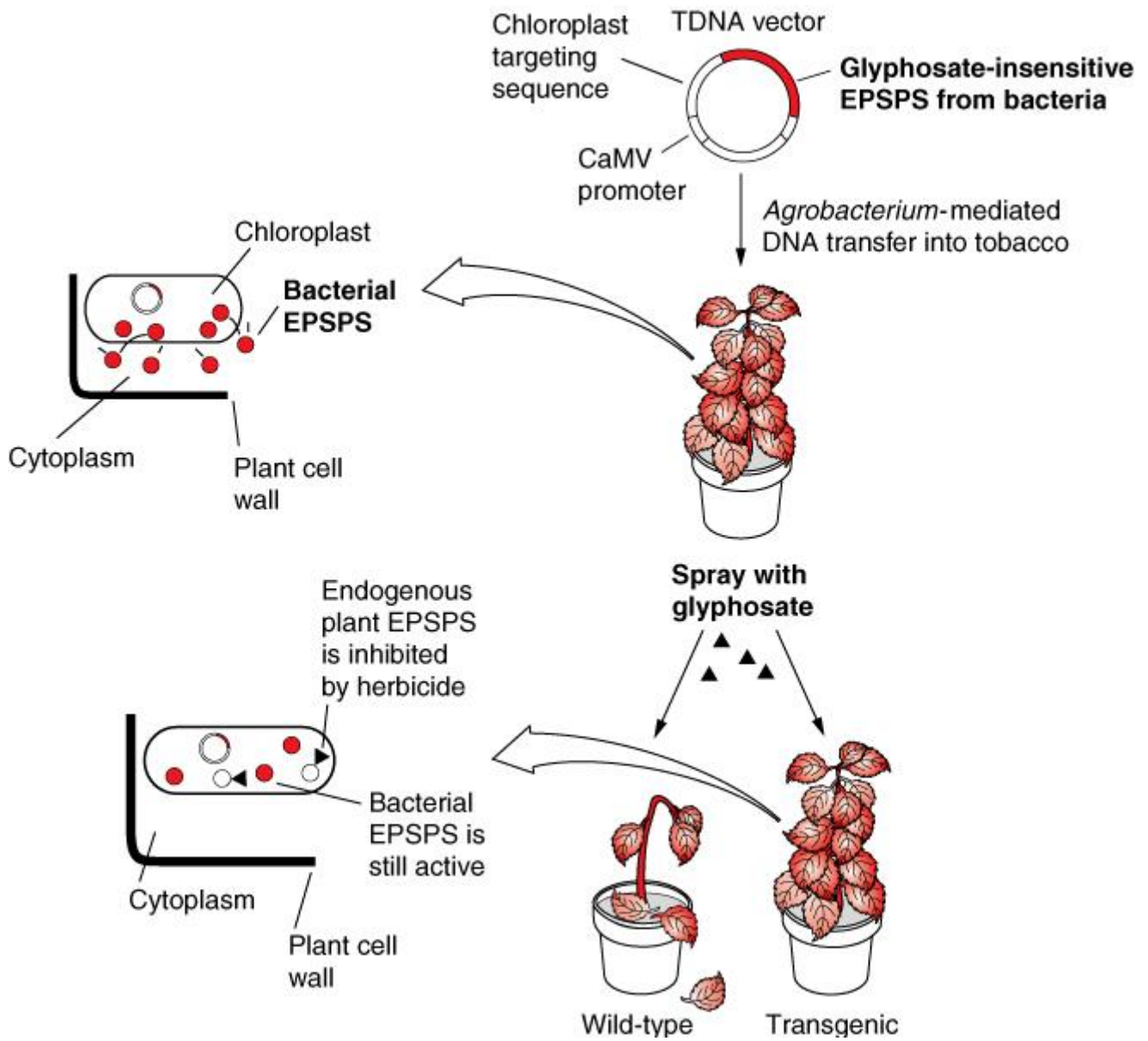
Resistance of insects to the defense mechanisms of plant is well-documented. Resistance to a number of commonly used insecticides has been reported in hundreds of insects and mites. There has been the concern that, sooner or later, resistance of insects to Bt cultivars would surface. Laboratory resistance to Bt has been demonstrated for some major pests, such as the tobacco budworm, Colorado potato beetle, and the diamondback moth. Some reports have indicated field resistance for the diamondback moth to Bt. Also, resistant populations of the bollworm (*Helicoverpa zea*) have been reported in some fields in Mississippi and Arkansas.

Production of Insect Resistant Cotton



Researchers are investigating how to extend the Bt durability in transgenic cultivars. Current approaches include the engineering of cultivars with very high levels of insecticidal crystal proteins. This way, only insects that have a high-level resistance gene can survive after feeding on these new cultivars. Another approach is to search for new insecticidal genes for developing new transgenic plants that can express multiple insecticidal genes that target different sites in the insect. Insects that can overcome this strategy are those with multiple resistance genes.

Resistance to herbicides is also growing, the confirmed cases approaching 300. The concern for growers and researchers is that some pests are resistance to multiple pesticides, while some are resistant to all the pesticides that are legally approved for their control!



BASIC CONCEPTS

- Aseptic culture** refers to *in vitro* culture of tissues free from micro-organisms like bacterium, fungi etc.
- Auxins** refers to a group of growth hormones which causes cell elongation, apical dominance, root initiation Eg : NAA, IAA, 2,4-D etc.
- Batch culture** refers to growth of cell suspension culture in fixed volume of liquid medium.
- Browning:** Discolouration of the *in vitro* culture due to a pathogen or nutritional problem or phenolic oxidation. Leads to death of the culture.
- Callus** refers to a unorganized mass of undifferentiated cells derived from explants.
- cDNA cloning:** A method of cloning the coding sequence of a gene, starting with its mRNA transcript. Normally used to clone a DNA copy of a eukaryotic mRNA.
- cDNA library :** A collection of cDNA clones
- cDNA:** Complementary DNA, a fragment of DNA produced from mRNA by reverse transcription.
- Cell culture** refers to growing of cells *in vitro*.
- Central dogma:** The concept that genetic information generally can flow only from DNA to RNA to protein. It is now known that information can flow back from RNA to DNA as in retroviruses.
- Clonal propagation** refers to asexual propagation of plants.
- Clone** refers to a group of cells, tissues or plants which are genetically identical to the mother cell or plant.
- Continuous culture :** When suspension culture is continuously supplied with nutrients by continuous flow of fresh medium it is called as continuous culture. The volume of the culture medium is normally constant.
- Cosmids** refers to plasmid vectors into which phage lambda cos sites were inserted. As a result the plasmid DNA can be packed in the phage coat. It can carry around 40kb of DNA fragment. Its replication is same as plasmid.
- Culture** refers to growing of cells or tissues of plant in nutrient medium under aseptic conditions. Depending upon the explant source it can be named as anther culture, pollen culture, embryo culture, protoplast culture, callus culture etc.
- Cybrid** refers to cytoplasmic hybrid. A somatic hybrid with nucleus from one parent and cytoplasm from both the parents.
- Cybrid:** It refers to a cytoplasmic hybrid produced by fusion of two protoplasts.
- Cytokinins** refers to a class of growth hormones which causes cell division, shoot differentiation, breaking of apical dominance etc. Kinetin, Zeatin etc.
- de novo** means from the beginning, arising anew.
- Dedifferentiation :** Reversion of differentiated cells to non-differentiated condition.
- Differentiation** is a process in which unspecialized cells develop into organs with specific function like roots, shoots etc.

DNA amplification : Multiplication of a piece of DNA into millions of copies by polymerase chain reaction (PCR).

DNA fingerprinting: A technique in which an individual's DNA is analysed

DNA ligase : The enzymes which joins DNA fragments

Embryo culture refers to the culture of embryos *in vitro* on nutrient medium.

Embryogenesis : The process by which an embryo develops from a fertilized egg cell or asexually from a group of cells.

Embryoids refers to embryo like structures formed in culture. They are also called as somatic embryos.

Excise: Cut out (with knife or scalpel) and prepare a tissue or organ for culture.

Explant refers to a piece of tissue or callus used to initiate tissue culture.

Genetic Engineering refers to changing the genetic architecture of an organism directly at the DNA level by molecular techniques.

Hardening refers to process of gradual acclimatization of tissue culture plants in green house.

In vitro refers to any process which is carried out in sterile condition in the laboratory. *In vitro* means "in glass".

In vivo refers to any process occur in a whole organism under field condition. *In vivo* means "in living".

Meristem refers a group of actively dividing cells from which other tissue such as root, shoot, leaf, flower etc. are derived.

Meristemoid refers to a group of cells in a callus giving rise to adventitious shoots or roots.

Micropropagation refers to clonal propagation of plants under *in vitro* conditions.

Nutrient medium : A solid or liquid medium containing nutrients such as inorganic salts, a carbon source, vitamins and growth regulators used for culturing plants or microorganisms is called as nutrient medium.

Organ culture refers to culture of plant organs *in vitro* leading to like root or shoot tips.

Organogenesis refers to formation of organs directly from a non-meristematic explant.

Parasexual hybridization refers to hybridization by non-sexual methods like protoplast fusion.

Passage time: The time interval between two successive sub-cultures is called as passage time.

Plant tissue culture refers to the *in vitro* cultivation of cells, tissues, organs, embryos, seeds and protoplasts on nutrient media under aseptic conditions in a laboratory.

Plasmids are extrachromosomal, autonomous circular DNA found in certain bacteria, capable of autonomous replication. They can transfer genes between bacteria and act an important tool in genetic engineering.

Protoplast refers to a cell without cell wall. The cell wall may be removed mechanically or

by using cell wall digesting enzymes.

Recognition sites refers to specific nucleotide sequence composed of 4-8 nucleotides to which restriction endonuclease binds.

Recombinant DNA or **rDNA** refers to the technique of cutting and recombining DNA from different organisms.

Restriction enzyme is an endonuclease which has the ability to cut (cleave) DNA at the point where a certain base sequence occurs.

Sub culture refers to establishment of a new culture by transfer of some of the cells from a previous culture to a fresh medium aseptically.

Suspension culture refers to culturing of cells in agitated liquid medium.

Totipotency is the ability of a single cell to develop into a whole organism.

Transformation: The process of transfer of foreign DNA to an individual using vectors.

Transgenic refers to organisms with foreign DNA inserted in its genome. Also called as Genetically Modified Organism (GMO). It may be a transgenic plant or animal or microbe.

Vector refers to a small self replicating DNA molecule in which DNA of interest is inserted. Vectors carry the inserted DNA into the host cell for multiplication. Eg plasmid, phage etc.

Solve these questions to gain more clarity on basic concepts.

1. In the Meselson-Stahl DNA replication experiment, if the cells were first grown for many generations in ^{15}N containing media, and then switched to ^{14}N containing media, what percent of the DNA had 1 light strand and 1 heavy strand after 2 generations of growth in ^{15}N growth media?

A 0

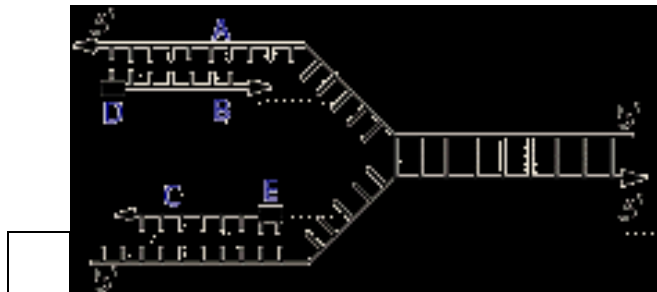
B 25

C 50

D 75

E 100

2.



In this diagram of the process of DNA replication at a replication fork, the black boxes labeled D and E are:

A RNA primers

B DNA template strands

C Okazaki fragments

D DNA polymerase

E Newly synthesized DNA strand

