SRINIVASAN COLLEGE OF ARTS & SCIENCE

(Affiliated to Bharathidasan University, Trichy)

PERAMBALUR – 621 212.

DEPARTMENT OF MICROBIOLOGY

Course : B.Sc Year: **III** Semester: **VI**

Course Material on:

rDNA TECHNOLOGY

Sub. Code : 16SMBEMB2

Prepared by :

Dr. S.RAJA, M.Sc., M.Phil., Ph.D.

ASSISTANT PROFESSOR / MB

Month & Year : APRIL – 2020

SUBJECT: RECOMBINANT DNA TECHNOLOGY

CLASS: III B.SC., MICROBIOLOGY

SUBJECT CODE: 16 SMBEMB 2

Dr.S.RAJA *Msc.,M.Phil., P.hD.,*

RECOMBINANT DNA TECHNOLOGY

OBJECTIVES

To educate the learners with the fundamental knowledge and importance of recombinant DNA (rDNA) technology. To teach the jargons of genetic engineering/ rDNA technology as well as the basic tools, techniques and methods employed in gene cloning and gene expression strategies.

UNIT I

Milestones in rDNA technology - Definition of gene manipulation - Major steps involved in gene cloning - Isolation and Purification of Chromosomal and Plasmid DNA, Isolation and Purification of RNA - Chemical Synthesis of DNA, Genomic Library and cDNA Library - applications.

UNIT II

Restriction endonucleases: Discovery, Type I, II and III and Mode of action, Applications of type II restriction endonucleases, Ligases, DNA polymerases, DNA modifying enzymes and topoisomerases.

UNIT III

Cloning vectors: Definition and properties – Plasmid based vectors: Natural vectors (pSC101, pSF2124, pMB1), Artificial vectors (pBR322 and pUC) – Phage based vectors- ג (Lamda) phage vectors and its derivatives - Hybrid Vectors- Phagemid and Cosmid, BAC and YAC – Expression systems – *E. coli*.

UNIT IV

Gene/ DNA transfer techniques: Physical – Biolistic Method (Gene gun), Chemical- Calcium chloride and DEAE Methods, Biological *in vitro* packaging method in viruses - Selection and Screening of recombinants: Direct Method: Selection by Complementation, Marker inactivation methods – Indirect methods: Immunological and Genetic methods.

UNIT V

Blotting (Southern, Western, Northern and North- eastern) techniques – PCR - basic steps in DNA amplification, RAPD, RFLP and their applications – DNA finger printing - DNA microarray analysis – Applications of recombinant DNA technology.

Definition of Recombinant DNA:

For centuries humans have been altering the genetic makeup of organisms by selective breeding of plant and animals. The deliberate modification in genetic material of an organism by changing the nucleic acid directly is called genetic engineering or gene cloning or gene manipulation and is accomplished by several methods which are collectively known as recombinant DNA (rDNA) technology.

Recombinant DNA technology begins a new area of research and applied aspects of biology. Therefore, it is a part of biotechnology which is gaining momentum and much boost in recent years.

In breeding programmes much work has been done on alteration of nucleotides by several para-sexual or conjugation methods in different group of organisms. Now-a-day, a large number of mutagenic agents are available that mutate the genes. It is likely that the changed genes may be beneficial, neutral or lethal.

Moreover, the conventional breeding programmes are time taking to ensure that the genes have been altered. In contrast, the rDNA technology has solved several problems which hardly or never are possible through the conventional methods.

Gene cloning or genetic engineering can be defined as "changing of genes by using in vitro processes". A unified definition of genetic engineering has been given by Smith (1996) as the formation of new combinations of heritable material by the insertion of nucleic acid molecules produced by whatever means outside the cell, into any virus, bacterial plasmid or other vector system so as to allow their incorporation into a host organism in which they do not naturally occur but in which they are capable of continued propagation.

In brief, gene technology is the modification of the genetic properties of an organism by using rDNA technology. Genes are like the biological software filled with programmes that govern the growth, development and function of organism. By changing in programme of the software it is possible to bring about alteration in the characters of a given organism.

A gene of known function can be isolated from its normal location by biochemical methods in vitro. Moreover, a gene can be synthesized by using gene machine.

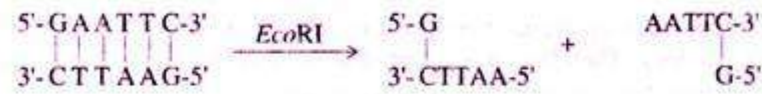
The isolated genes can be transferred into the microbial cells (that of course do not contain) via a suitable vector. The transfened gene replicates normally and is handed over to the next progeny over generations. After confirmation for its presence through biochemical procedures clone of the same cell is produced.

History of Recombinant DNA:

The first break through of rDNA technology occurred with the discovery of restriction endonucleases (restriction enzyme) during the late 1960s by Werner, Arber and Hamilton Smith. The restriction

enzymes were discovered in microorganisms. These enzymes protect the host cell from the bacteriophage. The restriction enzymes are described in the preceding section.

1. In 1969, Herbert Boyer isolated restriction enzyme EcoRI from E. coli that cleaves the DNA between G and A in the base sequence GAATTC as below:



2. In 1970 Howard Temin and Davin Baltimore independently discovered the enzyme reverse transcriptase from retroviruses. Later on this enzyme was used to construct a DNA called complementary DNA (cDNA) from any mRNA.
3. In, 1972 David Jackson, Robert Symons and Paul Berg successfully generated rDNA molecules. They allowed the stickly ends of complementary DNA by using an enzyme DNA ligase.
4. In 1973 for the first time S.Cohen and H. Boyer developed a recombinant plasmid (p^{SC101}) which after using as vector replicated well within a bacterial host.
5. In, 1975, Edwin M.Southern developed a method for detection of specific DNA fragments for isolation of a gene from complex mixture of DNA. This method is known as the Southern blotting technique.

Milestones of Recombinant DNA:

1. 1976 – First prenatal diagnosis by using gene specific probe.
2. 1977 – Methods for rapid DNA sequencing, discovery of split genes and somatostanin by rDNA.
3. 1979 – Insulin synthesized by using rDNA; first human viral antigen.
4. 1981 – Foot and mouth disease viral antigen cloned.
5. 1982 – Commercial production of coli of genetically engineered human insulin, Isolation, cloning and characterization of human cancer gene.
6. 1983 – Engineered Ti-plasmid used to transform plants.
7. 1985 – Insertion of cloned gene from Salmonella into tobacco plant to make resistant to herbicide glyphosphate; Development of PCR technique.
8. 1986 – Development of gene gun.
9. 1989 – First field test of genetically engineered virus (baculovirus) that kills cabbage looper caterpillars.
10. 1990 – Production of first transformed com.

11. 1991 – Production of first transgenic pigs and goats, manufacture of human haemoglobin, first test of gene therapy on human cancer patients.

12. 1994 – The Flavr Savr tomato introduced; the first genetically engineered whole food approved for sale. Fully human monoclonal antibodies produced in genetically engineered mice.

13. 1997 – World's first mammalian clone (Dolly) developed from a non-reproductive cell of an adult animal through cloning by nuclear transplantation.

Gene manipulation:

Gene manipulation is also sometimes called *the genetic engineering*. It is a general term for any method which manipulate with the genetic material. Gene manipulation includes *gene splicing*, use of *recombinant DNA*, forming of the *monoclonal antibodies* or *PCR* (polymerase chain reaction).
At first the gene manipulation was used in the agriculture – to improve the quality of the plants. It was discovered in 2001 in native Mexican corn. The use in human genetic is still very complicated. Nowadays we use especially the bacteria to produce some human hormones. The gene that codes the hormone is isolated and put into the genome of the bacteria – the most common is *E. coli*. Then they are cultivate on the suitable cultures. The bacteria becomes capable of the hormone production – *insulin*, *human interferon*, *human growth hormone* and other.
Another method is the recombinant DNA which allows us to produce the antibodies. Thanks to recombinant DNA we can also analyse the human genes in detail. We use the ability of produce the large number of clones. These clone we then screen by the probe. The probe is a radioactively RNA or DNA fragment. The gene manipulation allowed us to create the genes maps.

Major steps involved in gene cloning:

1. **Isolation of DNA (Gene of Interest) Fragments to be Cloned:**
   Before we carry out the operation of gene cloning we need two basic things in their purified state – the gene of our interest (GI) and the vector. A GI is a fragment of gene whose product (a protein, enzyme or a hormone) interests us. For example, gene encoding for the hormone insulin.
   Similarly, the vector is a carrier molecule which can carry our GI into a host, replicate there along with the GI making its multiple copies. In this state the GI can also be expressed in the host cell producing the product of the gene which is needed by us.
2. **Insertion of Isolated DNA into the a Suitable Vector to Form the Recombinant DNA:**
   Once the ingredients are ready we can start the operation. Our next step will be to cut both the vectors as well as the GI by using a special type of enzyme, called restriction endonuclease. A restriction endonuclease is an enzyme that cuts double-stranded or single-stranded DNA at

specific recognition nucleotide sequences known as restriction sites towards the inner region (hence endonuclease).

They are also regarded as molecular scissors as they cut open the DNA strands. After this cutting step we move to pasting. Here the GI is taken and pasted to the cut vector. This procedure also needs an enzyme, called DNA ligase. They are also considered as molecular glue.

The resulting DNA molecule is a hybrid of two DNA molecules – our GI and the vector. In the terminology of genetics this intermixing of different DNA strands is called recombination (which naturally takes place in the prophase 1 of meiosis 1). Hence, this new hybrid DNA molecule is also called a recombinant DNA molecule and this technology is called recombinant DNA technology (RDT).

3. **Introduction of the Recombinant DNA into a Suitable Organism known as Host:**

   When our recombinant DNA molecule is ready we need to introduce it into a living system known as host.

   This is done either for one or both of the following reasons:

   (a) To replicate the recombinant DNA molecule in order to get the multiple copies of our GI.

   (b) To let our GI get express and produce the protein which is needed by us.

   Introduction of the recombinant DNA into the host cell is done by various ways and strictly depends upon the size of the DNA molecule and the nature of GI. Some of the methods followed to carry out this step includes electroporation, micro-injection, lipofection, etc.

   When we carry out this process some of the host cells will take up the recombinant DNA and some will not. The host cells which have taken up the recombinant DNA are called transformed cells and the process is called transformation.

4. **Selection of the Transformed Host Cells and Identification of the Clone Containing the Gene of Interest:**

   The transformation process generates a mixed population of transformed and non-transformed host cells. As we are interested only in transformed host cells it becomes necessary to filter them out. This is exactly what is done in the selection process. There are many existing selection strategies some of which include taking the help of reporter genes, colony hybridization technique, etc.

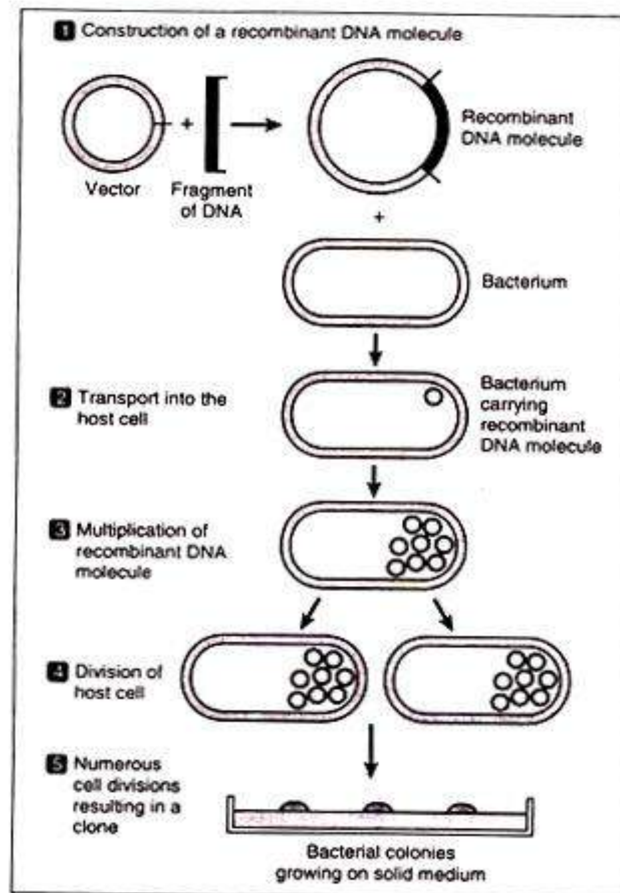
5. **Multiplication/Expression of the Introduced Gene in the Host:**

   Once we have purified our transformed host cells by the screening process; it is now our job to provide them optimum parameters to grow and multiply. In this step the transformed host cells are introduced into fresh culture media which provide them rich nourishment followed by an incubation in the oven at right temperature.

   At this stage the host cells divide and re-divide along with the replication of the recombinant DNA carried by them. Now at this point we have two choices.

When the aim of the cloning process is to generate a gene library, then our target will be obtaining numerous copies of GI. So with this plan in our mind we will simply go for the replication of the recombinant DNA and not beyond that.

If the aim of the cloning experiment is to obtain the product of GI, then we will go for a step ahead where we will provide favourable conditions to the host cells in which the GI sitting in the vector can express our product of interest (PI).



6. **Isolation of the Multiplied Gene Copies/Protein Expressed by the Introduced Gene:**

   In this step we isolate our multiplied GI which is present attached with the vector or the protein encoded by it. This can be rightly compared with the process of harvesting where we collect the crop from the field. There are many processes of isolation, the selection of which varies from case to case.

   The recombinant host cell is then grown in culture media but the culture may contains colonies both recombinant cell and non-recombinant cell. For isolation of recombinant cell from non-recombinant cell, marker gene of plasmid vector is employed. For examples, PBR322 plasmid vector contains different marker gene (Ampicillin resistant gene and Tetracycline resistant gene. When pst1 RE is used it knock out Ampicillin resistant gene from the plasmid, so that the recombinant cell become sensitive to Ampicillin.

ISOLATION AND PURIFICATION OF GENOMIC/PLASMID DNA AND RNA

Isolation and Purification of Genomic DNA

Genomic DNA is found in the nucleus of all living cells with the structure of doublestranded DNA remaining unchanged (helical ribbon). The isolation of genomic DNA differs in animals and plant cells. DNA isolation from plant cells is difficult due to the presence of cell wall, as compared to animal cells. The amount and purity of extracted DNA depends on the nature of the cell. The method of isolation of genomic DNA from a bacterium comprises following steps

1. Bacterial culture growth and harvest.
2. Cell wall rupture and cell extract preparation
3. DNA Purification from the cell extract.
4. Concentration of DNA solution.

Growth and harvest of bacterial culture: Bacterial cell culture is more convenient than any other microbe, as it requires only liquid medium (broth) containing essential nutrients at optimal concentrations, for the growth and division of bacterial cells. The bacterial cells are usually grown on a complex medium like Luria-Bertani (LB), in which the medium composition is difficult to decipher. Later, the cells are separated by centrifugation and resuspended in 1% or less of the initial culture volume.

Preparation of cell extract: Bacterial cell is surrounded by an additional layer called cell wall, apart from plasma membrane with some species of E. coli comprising multilayered cell wall. The lysis of cell wall to release the genetic material i.e. DNA can be achieved by following ways-

• Physical method by mechanical forces.

• Chemical method by metal chelating agents i.e. EDTA and surfactant i.e. SDS or enzyme (e.g. lysozyme).

Lysozyme

• present in egg-white, salivary secretion and tears.

• catalyzes the breakdown of cell wall i.e. the peptidoglycan layer.

EDTA (Ethylene diamine tetra-acetic acid)

• a chelating agent necessary for destabilizing the integrity of cell wall.

• inhibits the cellular enzymes that degrade DNA.

SDS (Sodium dodecyl sulphate)

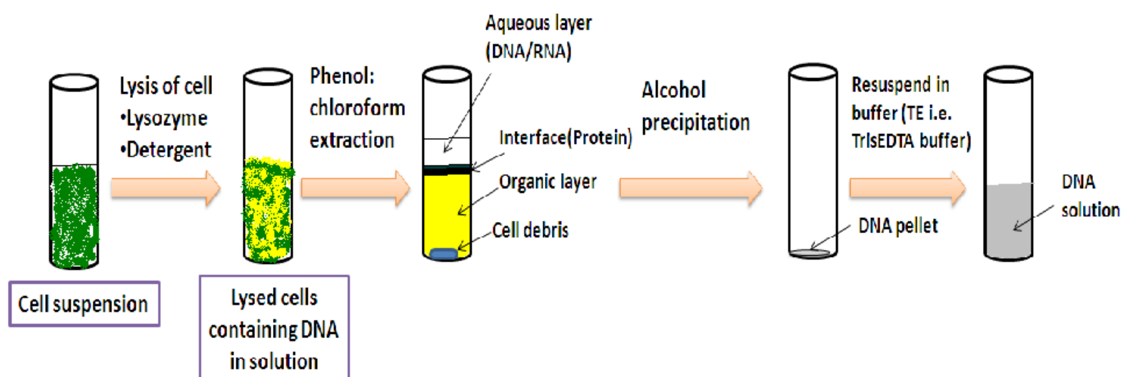
• helps in removal of lipid molecules and denaturation of membrane proteins.

Purification of DNA :

In addition to DNA, a cell extract contains significant quantities of protein and RNA which can be further purified by following methods

Organic extraction and enzymatic digestion for the removal of contaminants: It involves the addition of a mixture of phenol and chloroform (1:1) to the cell lysate for protein separation. The proteins aggregate as a white mass in between the aqueous phase containing DNA and RNA, and the organic layer. Treatment of lysate with pronase or protease, in addition to phenol/chloroform, ensures complete removal of proteins from the extract. The RNA can be effectively removed by using Ribonuclease, an enzyme which rapidly degrades RNA into its ribonucleotide subunits. Repeated phenol extraction is not desirable, as it damages the DNA.

Using ion-exchange chromatography: This involves the separation of ions and polar molecules (proteins, small nucleotides and amino acids) based on their charge. DNA carrying negative charge binds to the cationic resin or matrix which can be eluted from the column by salt gradient. Gradual increase in salt concentration detaches molecules from the resin one after another.



Isolation and Purification of Plasmid DNA

Plasmids are circular, double stranded extra cellular DNA molecules of bacterium and most commonly used in recombinant DNA technology. The isolation of plasmid DNA involves three major steps

1. Growth of the bacterial cell.
2. Harvesting and lysis of the bacteria.
3. Purification of the plasmid DNA..

Growth of the bacterial cell: It involves growth of the bacterial cells in a media containing essential nutrients.

Harvest and lysis of bacteria:

Lysis of bacteria results in the precipitation of DNA and cellular proteins. Addition of acetate-containing neutralization buffer results in the precipitation of large and less supercoiled chromosomal DNA and proteins leaving the small bacterial DNA plasmids in solution.

Purification of Plasmid DNA: This step is same for both plasmid and genomic but former involves an additional step i.e. the separation of plasmid DNA from the large bacterial chromosomal DNA.

Methods for separation of plasmid DNA:

Separation of plasmid DNA is based on the several features like size and conformation of plasmid DNA and bacterial DNA. Plasmids are much smaller than the bacterial main chromosomes, the largest plasmids being only 8% of the size of the E. coli chromosome. The separation of small molecules (i.e. plasmids) from larger ones (i.e. bacterial chromosome) is based on the fact that plasmids and the bacterial chromosomes are circular but bacterial chromosomes break into linear fragments during the preparation of the cell extract resulting in separation of pure plasmids. The methods of separation of plasmid DNA are described as below

Separation based on size difference

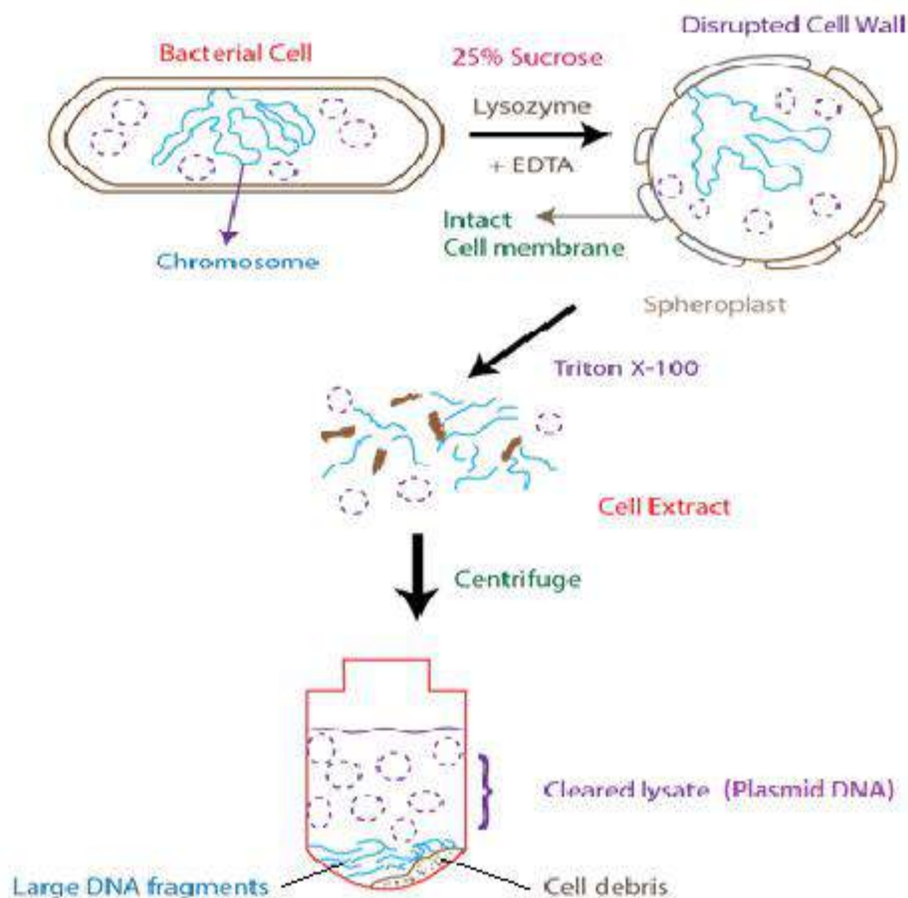
• It involves lysis of cells with lysozyme and

• Cells with partially degraded cell walls are formed that retain an intact cytoplasmic membrane called as sphaeroplasts.

• Cell lysis is then induced by the addition of a non-ionic detergent (e.g. Triton X100) or ionic detergents (e.g. SDS) causing chromosomal breakage.

• Bacterial chromosome attached to cell membrane, upon lysis gets removed with the cell debris.

• A cleared lysate consisting almost entirely of plasmid DNA is formed with very little breakage of the bacterial DNA.

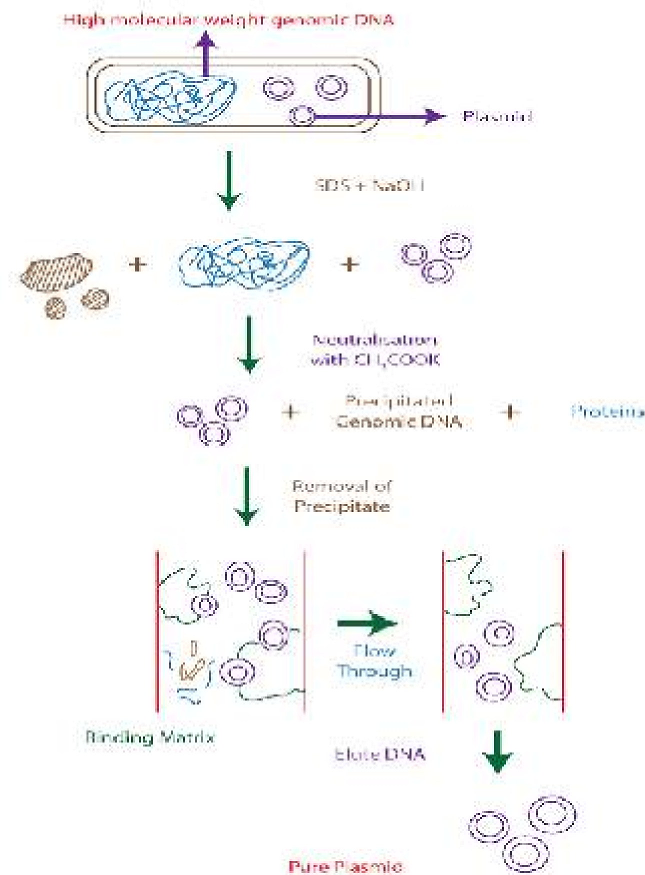


Separation based on conformation

Plasmids are supercoiled molecules formed by partial unwinding of double helix of the plasmid DNA during the plasmid replication process by enzymes called topoisomerases. The supercoiled conformation can be maintained when both polynucleotide strands are intact, hence called covalently closed-circular (ccc) DNA. If one of the polynucleotide strands is broken, the double helix reverts to its normal relaxed state taking an alternative conformation, called open-circular (oc). Super coiling is important in plasmid preparation due to the easy separation of supercoiled molecules from non-supercoiled ones. The commonly used methods of separation based on conformation are as follows

1. Alkaline denaturation method
2. Ethidium bromide-cesium chloride density gradient centrifugation

Separation of plasmid DNA by Alkaline denaturation method



Isolation and Purification of RNA:

RNA (Ribonucleic acid) is a polymeric substance consisting of a long single-stranded chain of phosphate and ribose units with the nitrogen bases adenine, guanine, cytosine and uracil bonded to the ribose sugar present in living cells and many viruses. The steps for preparation of RNA involve homogenization, phase separation, RNA precipitation, washing and re-dissolving RNA. The method for isolation and purification of RNA are as follows

1) Organic extraction method

2) Filter-based, spin basket formats

3) Magnetic particle methods

4) Direct lysis method.

Organic extraction method: This method involves phase separation by addition and centrifugation of a mixture of a solution containing phenol, chloroform and a chaotropic agent (guanidinium thiocyanate) and aqueous sample. Guanidium thiocyanate results in the denaturation of proteins and RNases, separating rRNA from ribosomes. Addition of chloroform forms a colorless upper aqueous phase containing RNA, an interphase containing DNA and a lower phenol-chloroform phase containing protein. RNA is collected from the

upper aqueous phase by alcohol (2-propanol or ethanol) precipitation followed by rehydration. One of the advantages of this method is the stabilization of RNA and rapid denaturation of nucleases. Besides advantages, it has several drawbacks such as it is difficult to automate, needs labor and manual intensive processing, and use of chlorinated organic reagents.

Direct lysis methods: This method involves use of lysis buffer under specified conditions for the disruption of sample and stabilization of nucleic acids. If desired, samples can also be purified from stabilized lysates. This method eliminates the need of binding and elution from solid surfaces and thus avoids bias and recovery efficiency effects.

cDNA library

A cDNA library is a combination of cloned cDNA fragments constituting some portion of the transcriptome of an organism which are inserted into a number of host cells. In eukaryotic cells, the mRNA is spliced before translation into protein. The DNA synthesized from the spliced mRNA doesn't have introns or non-coding regions of the gene. As a result, the protein under expression can be sequenced from the DNA which is the main advantage of cDNA cloning over genomic DNA cloning.

Construction of a cDNA Library : The construction of cDNA library involves following steps

1. Isolation of mRNA
2. Synthesis of first and second strand of cDNA
3. Incorporation of cDNA into a vector
4. Cloning of cDNAs

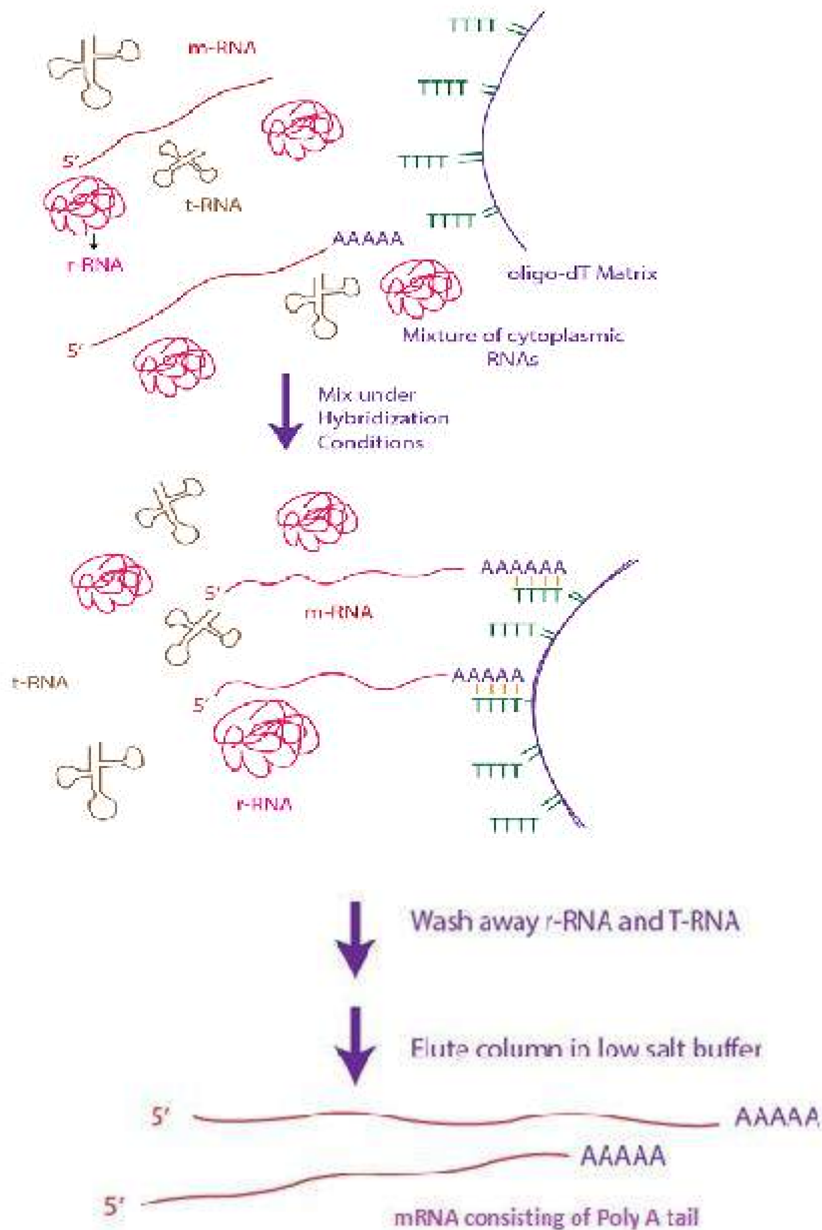
Isolation of mRNA: It involves the isolation of total mRNA from a cell type or tissue of interest. The amount of desired mRNA can be increased by following ways-

- Chromatographic purification of mRNA using oligo-dT column, which retains mRNA molecules, resulting in their enrichment.
- Spinning down mRNA by density gradient centrifugation.
- mRNA preparation from specialized cell types, e.g. developing seeds, chicken oviduct, erythrocytes, β cells of pancreas etc.

The 3′ ends of eukaryotic mRNA consist of a string of 50 – 250 adenylate residues (poly A Tail) which makes the separation easy from the much more prevalent rRNAs and tRNAs in a cell extract using a column containing oligo-dTs tagged onto its matrix.

When a cell extract is passed through an oligo-dT column, the mRNAs bind to the column due to the complementary base-pairing between poly (A) tail and oligo-dT. Other RNAs (ribosomal RNAs and transfer RNAs) flow through as unbound fraction. The bound mRNAs can then be eluted using a low-salt buffer.

Isolation of mRNA using oligo-dT column chromatography



Synthesis of first and second strand of cDNA

• mRNA being single-stranded cannot be cloned as such and is not a substrate for DNA ligase. It is first converted into DNA before insertion into a suitable vector which can be achieved using reverse transcriptase (RNA-dependent DNA polymerase or RTase) obtained from avian myeloblastosis virus (AMV).

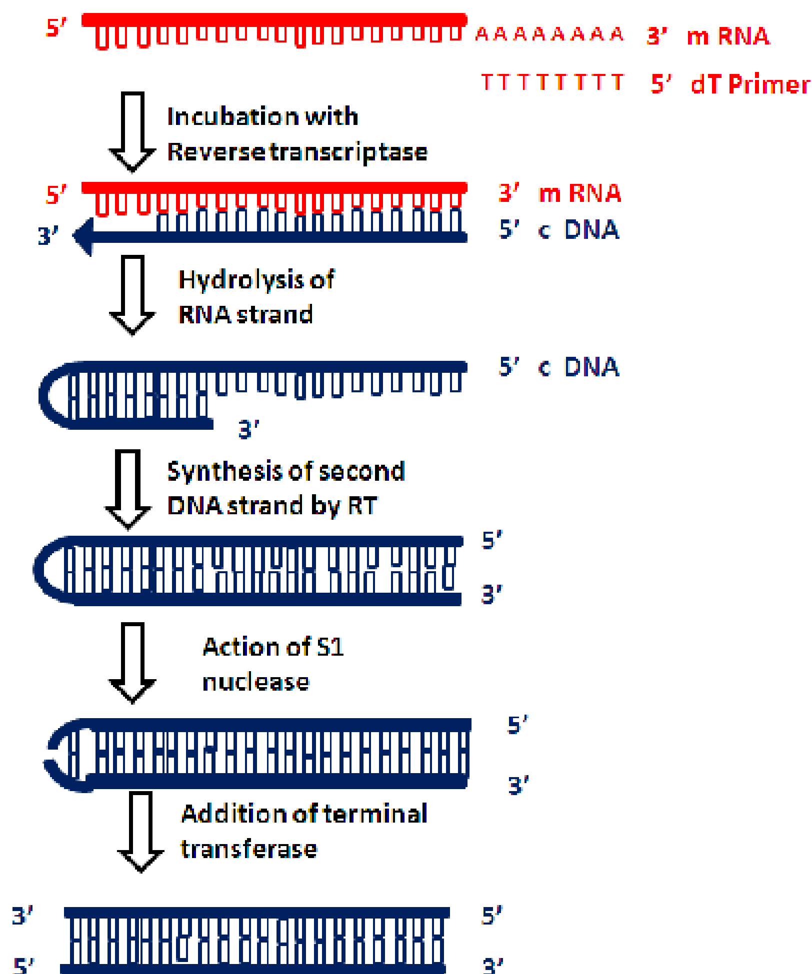
• A short oligo (dT) primer is annealed to the Poly (A) tail on the mRNA.

• Reverse transcriptase extends the 3´-end of the primer using mRNA molecule as a template producing a cDNA: mRNA hybrid. • The mRNA from the cDNA: mRNA hybrid can be removed by RNase H or Alkaline hydrolysis to give a ss-cDNA molecule.

• No primer is required as the 3´end of this ss-cDNA serves as its own primer generating a short hairpin loop at this end. This free 3´-OH is required for the synthesis of its complementary strand.

• The single stranded (ss) cDNA is then converted into double stranded (ds) cDNA by either RTase or E. coli DNA polymerase.

• The ds-cDNA can be trimmed with S1 nuclease to obtain blunt–ended ds-cDNA molecule followed by addition of terminal transferase to tail the cDNA with C's and ligation into a vector.



Incorporation of cDNA into a vector

The blunt-ended cDNA termini are modified in order to ligate into a vector to prepare dscDNA for cloning. Since blunt-end ligation is inefficient, short restriction-site linkers are first ligated to both ends.

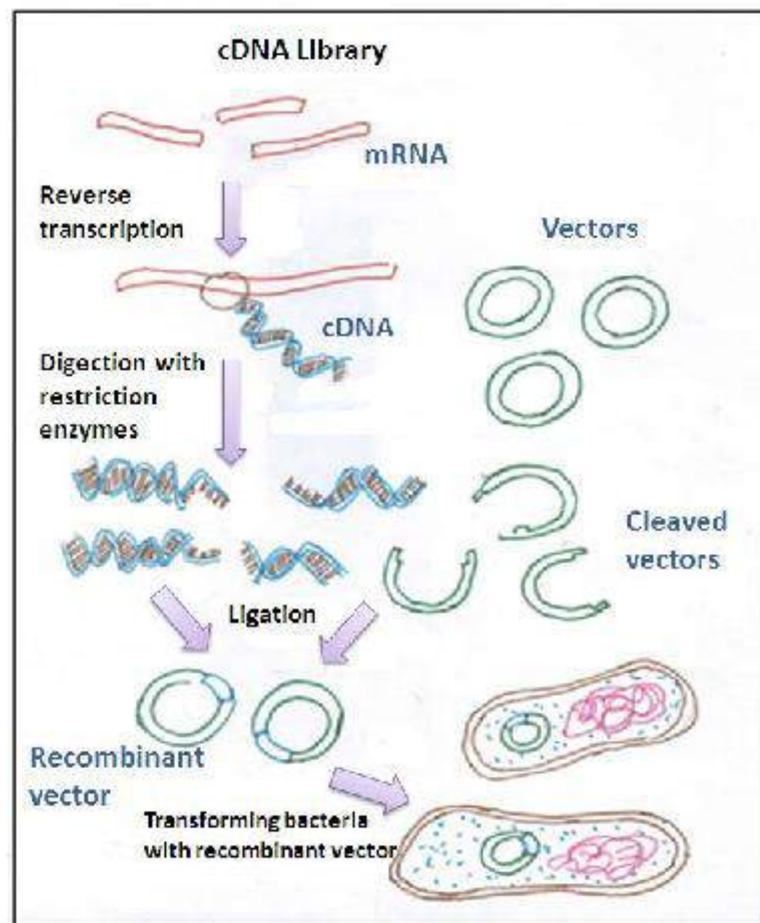
Cloning of cDNAs

cDNAs are usually cloned in phage insertion vectors. Bacteriophage vectors offer the following advantageous over plasmid vectors,

• are more suitable when a large number of recombinants are required for cloning low-abundant mRNAs as recombinant phages are produced by in vitro packaging.

• can easily store and handle large numbers of phage clones, as compared to the bacterial colonies carrying plasmids.

Plasmid vectors are used extensively for cDNA cloning, particularly in the isolation of the desired cDNA sequence involving the screening of a relatively small number of clones.



Applications of cDNA libraries/cloning

• Discovery of novel genes.

• in vitro study of gene function by cloning full-length cDNA.

• Determination of alternative splicing in various cell types/tissues.

• They are commonly used for the removal of various non-coding regions from the library.

• Expression of eukaryotic genes in prokaryotes as they lack introns in their DNA and therefore do not have any enzymes to cut it out in transcription process. Gene expression

required either for the detection of the clone or the polypeptide product may be the primary objective of cloning.

• To study the expression of mRNA.

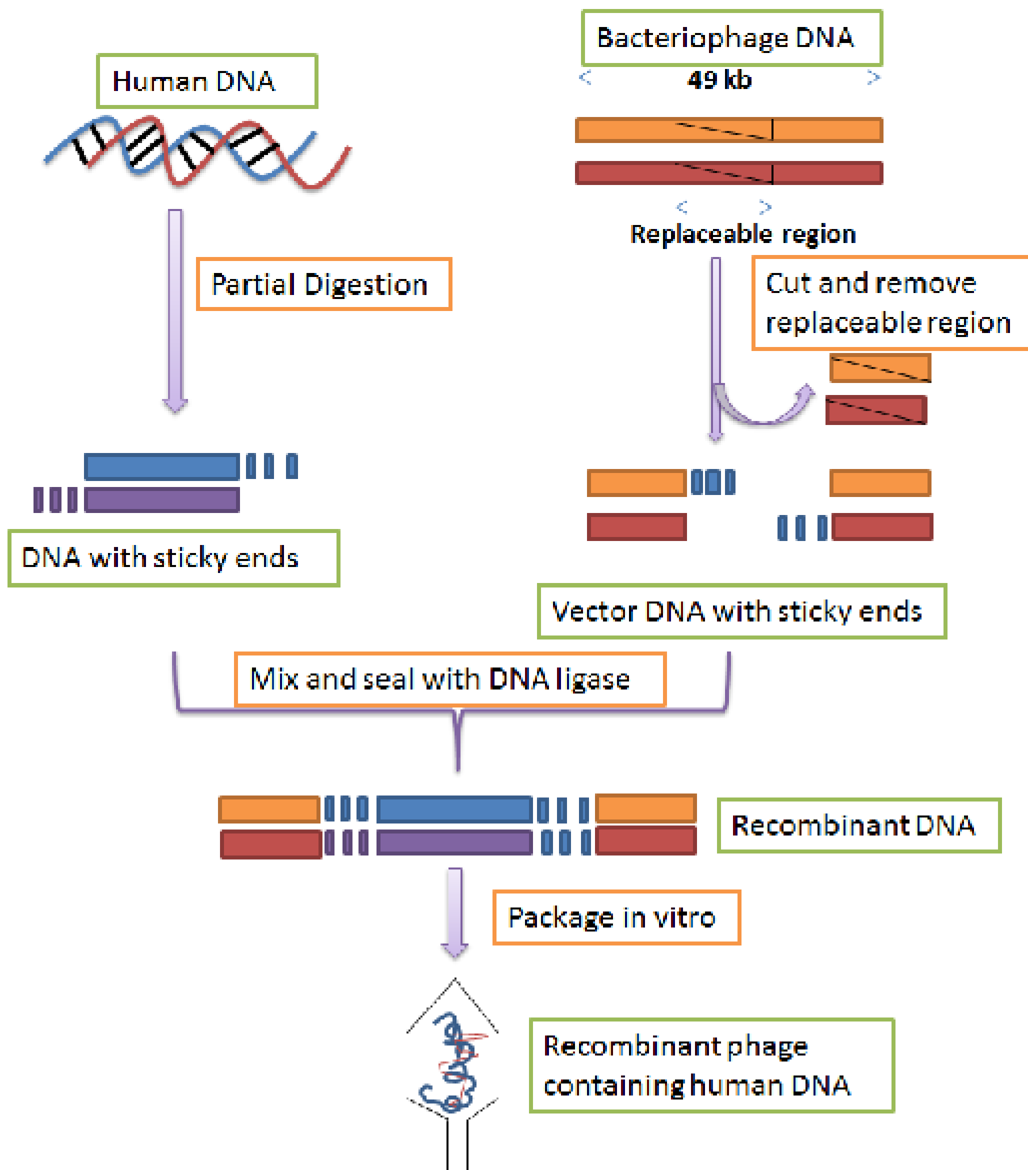
Genomic library

A genomic library is an organism specific collection of DNA covering the entire genome of an organism. It contains all DNA sequences such as expressed genes, nonexpressed genes, exons and introns, promoter and terminator regions and intervening DNA sequences.

Construction of genomic library

Construction of a genomic DNA library involves isolation, purification and fragmentation of genomic DNA followed by cloning of the fragmented DNA using suitable vectors. The eukaryotic cell nuclei are purified by digestion with protease and organic (phenol-chloroform) extraction. The derived genomic DNA is too large to incorporate into a vector and needs to be broken up into desirable fragment sizes. Fragmentation of DNA can be achieved by physical method and enzymatic method. The library created contains representative copies of all DNA fragments present within the genome.

- It involves use of restriction enzyme for the fragmentation of purified DNA.
- This method is limited by distribution probability of site prone to the action of restriction enzymes which will generate shorter DNA fragments than the desired size.
- If, a gene to be cloned contains multiple recognition sites for a particular restriction enzyme, the complete digestion will generate fragments that are generally too small to clone. As a consequence, the gene may not be represented within a library.
- To overcome this problem, partial digestion of the DNA molecule is usually carried out using known quantity of restriction enzyme to obtain fragments of ideal size.
- The two factors which govern the selection of the restriction enzymes are- type of ends (blunt or sticky) generated by the enzyme action and susceptibility of the enzyme to chemical modification of bases like methylation which can inhibit the enzyme activity.
- The fragments of desired size can be recovered by either agarose gel electrophoresis or sucrose gradient technique and ligated to suitable vectors.



Cloning of genomic DNA

Various vectors are available for cloning large DNA fragments. λ phage, yeast artificial chromosome, bacterial artificial chromosome etc. are considered as suitable vectors for larger DNA and λ replacement vectors like λDASH and EMBL3 are preferred for construction of genomic DNA library. T4 DNA ligase is used to ligate the selected DNA sequence into the vector.

Applications: • To determine the complete genome sequence of a given organism. • To study and generate transgenic animals through genetic engineering, serving as a source of genomic sequence. • To study the function of regulatory sequences in vitro. • To study the genetic mutations. • Used for genome mapping, sequencing and the assembly of clone contigs.

Restriction enzyme

A restriction enzyme is a nuclease enzyme that cleaves DNA sequence at a random or specific recognition sites known as restriction sites. In bacteria, restriction enzymes form a combined system (restriction + modification system) with modification enzymes that methylate the bacterial DNA. Methylation of bacterial DNA at the recognition sequence typically protects the own DNA of the bacteria from being cleaved by restriction enzyme.

There are two different kinds of restriction enzymes:

(1) **Exonucleases** catalyses hydrolysis of terminal nucleotides from the end of DNA or RNA molecule either 5'to 3' direction or 3' to 5' direction. Example: exonuclease I, exonuclease II etc.

(2) **Endonucleases** can recognize specific base sequence (restriction site) within DNA or RNA molecule and cleave internal phosphodiester bonds within a DNA molecule. Example: EcoRI, Hind III, BamHI etc.

The Discovery of Restriction Endonucleases

Prior to 1968 the existence of restriction enzymes was unknown. However, the phenomenon of restriction was well known. Restriction was the term given to the ability of bacteria to recognize and attack any foreign DNA source whether it came from a virus or from another strain of bacteria.

In 1968 Matthew Meselson and Robert Yuan reported that they had identified an enzyme in the bacterium Escherichia coli, strain K12, that appeared to be able to recognize and digest foreign DNAs. This enzyme, they concluded, could be the agent responsible for restriction. They coined the term restriction endonuclease to refer to this enzyme. They further determined that such enzymes would be ubiquitous among bacteria and that they would recognize and digest any double-stranded DNA that was not protected by a specific pattern of DNA base methylation . Methylation of DNA involves adding a methyl-group (CH3) to the DNA base such that the restriction enzyme will not recognize it. The process of methylation has been shown to be carried out by DNA sequence-specific methyltransferase enzymes. The restriction endonucleases found by Meselson and Yuan in E. coli required the

presence of Mg2+, SAM, and ATP for it to carry out its function. Thus, the first restriction enzyme to be identified was a Class I enzyme.

The cleavage pattern of both enzymes was limited and consistently reproducible, suggesting that there was a specific DNA sequence that was recognized by the enzymes and that the enzyme would bind to this sequence prior to cleavage.

Kelly and Smith offered evidence that the recognition site of their enzyme was a run of six specific nucleotides in the form,

5'… G T Py | Pu A C … 3'

3'… C A Pu | Py T G … 5'

where Py refers to either pyrimidine (T or C), Pu refers to either purine (A or G), and the vertical line indicates the cleavage site of the enzyme. Note that the symmetry of this recognition sequence is in the form of a palindrome, a nucleotide sequence in which the 5' to 3' sequence of one strand of a segment of DNA is the same as that of its complementary strand. This feature did not escape notice, "It is unlikely that the symmetry of this sequence is fortuitous, since the number of possible asymmetric sequences of this type is about 30 times the number of possible symmetric sequences…" They concluded that symmetry in the recognition sequence was a basic feature of the action of the enzyme.

Properties of Restriction Enzymes

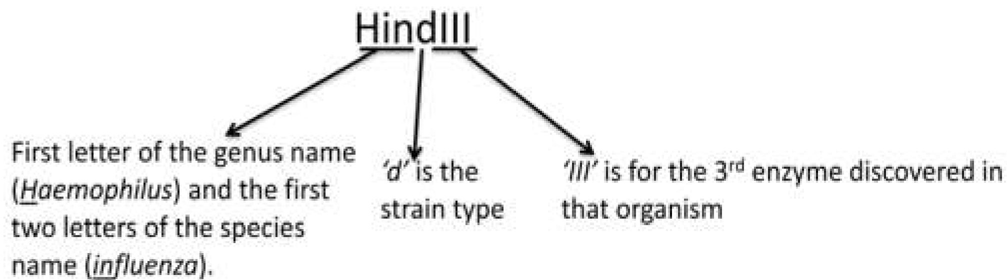
1. The restriction enzyme designated Bam HI was the first enzyme found in the bacterium *Bacillus amyloliquifaciens*, strain H while the restriction enzyme Hae III was the third enzyme found in *Haemophilus aegyptius*.
2. Most restriction enzyme recognition sequences are from four to eight bases long and most are palindromic. Additional diversity was found among the isoschizomers.

For example, the enzymes Sma I and Xma I both recognize the six base sequence CCCGGG but give different fragments with the former cutting CCC|GGG and the latter cutting C|CCGGG. Similarly, the isoschizomeric pair Hha I and Hin PI both recognize the sequence GCGC but the former cuts GCG|C and the latter G|CGC. Further differences were found in relation to sensitivity to methylation.

Restriction Endonuclease Nomenclature:

Restriction endonucleases are named according to the organism in which they were discovered, using a system of letters and numbers. For example, HindIII (pronounced "hindee-three") was discovered in Haemophilus influenza (strain d). The Roman numerals

are used to identify specific enzymes from bacteria that contain multiple restriction enzymes indicating the order in which restriction enzymes were discovered in a particular strain.



Restriction enzymes are named based on the bacteria in which they are isolated in the following manner:

E - Escherichia (genus)

co - coli (species)

R - RY13 (strain)

I - First identified Order ID'd in bacterium

Classification of Restriction Endonucleases: There are three major classes of restriction endonucleases based on the types of sequences recognized, the nature of the cut made in the DNA, and the enzyme structure:

1. Type I restriction enzymes
2. Type II restriction enzymes
3. Type III restriction enzymes

Type I restriction enzymes:

- These enzymes have both restriction and modification activities. Restriction depends upon the methylation status of the target DNA.
- Cleavage occurs approximately 1000 bp away from the recognition site.
- The recognition site is asymmetrical and is composed of two specific portions in which one portion contain 3–4 nucleotides while another portion contain 4–5 nucleotides and both the parts are separated by a non-specific spacer of about 6–8 nucleotides.
- They require S-adenosylmethionine (SAM), ATP, and magnesium ions (Mg2+) for activity.
- These enzymes are composed of mainly three subunits, a specificity subunit that determines the DNA recognition site, a restriction subunit, and a modification subunit.

Type II restriction enzymes:

- Restriction and modification are mediated by separate enzymes so it is possible to cleave DNA in the absence of modification. Although the two enzymes recognize the same target sequence, they can be purified separately from each other.
- Cleavage of nucleotide sequence occurs at the restriction site.
- These enzymes are used to recognize rotationally symmetrical sequence which is often referred as palindromic sequence.
- These palindromic binding site may either be interrupted (e.g. BstEII recognizes the sequence 5´-GGTNACC-3´, where N can be any nucleotide) or continuous (e.g. KpnI recognizes the sequence 5´-GGTACC-3´).
- They require only Mg2+ as a cofactor and ATP is not needed for their activity.
- Type II endonucleases are widely used for mapping and reconstructing DNA in vitro because they recognize specific sites and cleave just at these sites.

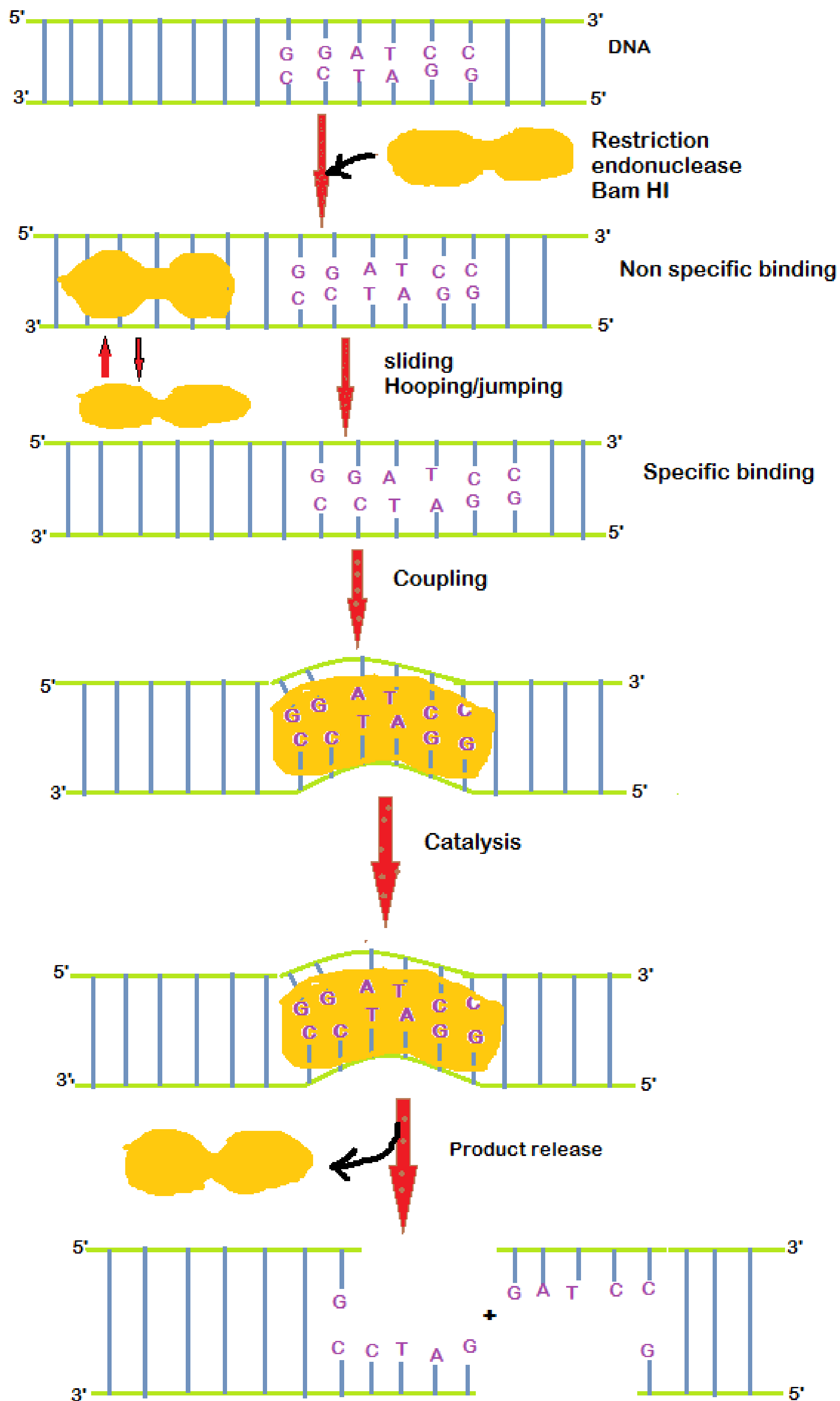
Type III restriction enzymes:

- These enzymes recognize and methylate the same DNA sequence but cleave 24–26 bp away.
- They have two different subunits, in which one subunit (M) is responsible for recognition and modification of DNA sequence and other subunit (R) has nuclease action.
- Mg+2 ions, ATP are needed for DNA cleavage and process of cleavage is stimulated by SAM.
- Cleave only one strand. Two recognition sites in opposite orientation are necessary to break the DNA duplex.

The steps involved in DNA binding and cleavage by a type II restriction endonuclease (Or) Mechanism of Action of Restriction Enzymes

- These enzymes have nonspecific contact with DNA and initially bind to DNA as dimmers.
- The target site is then located by a combination of linear diffusion or "sliding" of the enzyme along the DNA over short distances, and hopping/jumping over longer distances.
- Once the target restriction site is located, the recognition process (coupling) triggers large conformational changes of the enzyme and the DNA, which leads to activation of the catalytic center.

• Catalysis results in hydrolysis of phosphodiester bond and product release.

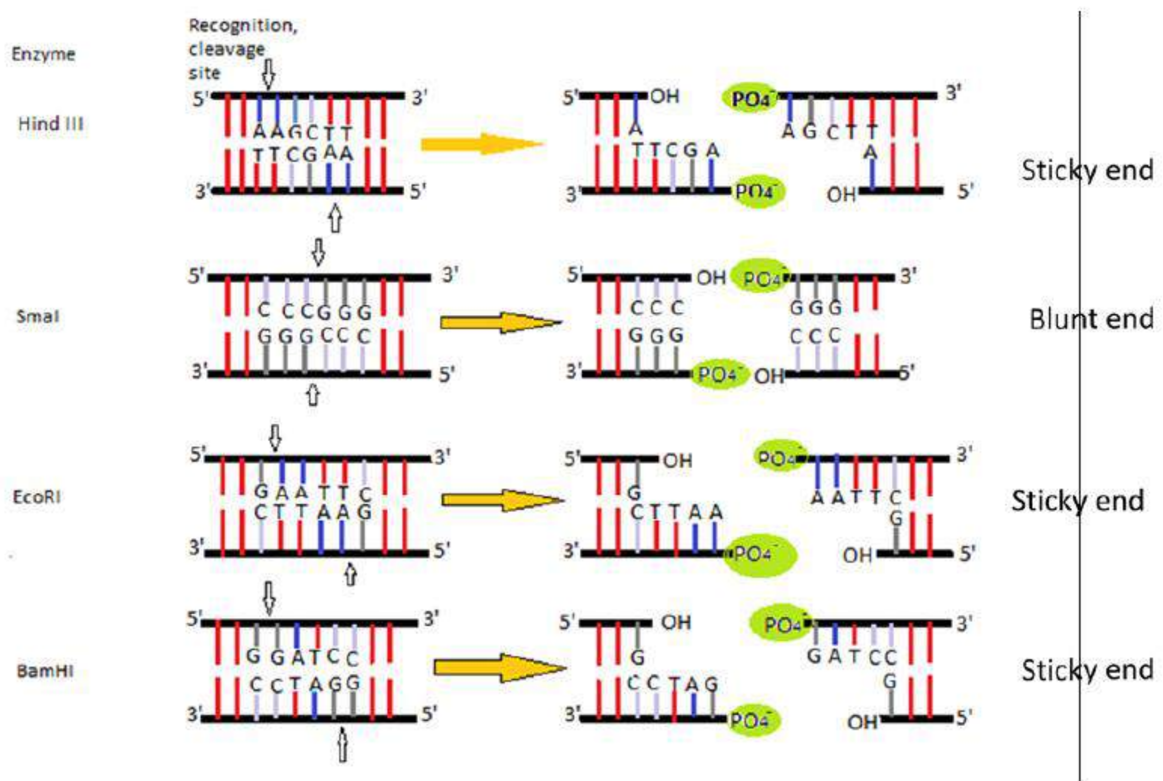


Structures of free, nonspecific, and specific DNA-bound forms of BamHI.

The two dimers are shown in brown, the DNA backbone is in green and the bases in gray. *BamH*I becomes progressively more closed around the DNA as it goes from the nonspecific to specific DNA binding mode.

Cleavage Patterns of Some Common Restriction Endonucleases:

The recognition and cleavage sites and cleavage patterns of *HindIII, SmaI, EcoRI*, and *BamHI* are shown. Cleavage by an endonuclease creates DNA sequence with either a sticky end or blunt end. The blunt ended fragments can be joined to any other DNA fragment with blunt ends using linkers/adapters, making these enzymes useful for certain types of DNA cloning experiments.



Cleavage patterns of *HindIII, SmaI, EcoRI* and *BamHI*

Application of Restriction Enzymes

(A) **Gene cloning and protein expression:** Restriction enzymes in combination with DNA ligase help in insertion of genes into plasmid vectors during gene cloning and protein expression. For this, both plasmid DNA containing multiple cloning sites and gene insert are treated with the same restriction enzymes and then glued together

with the help of DNA ligase. Restriction enzymes are also used after cloning to confirm that insertion of gene has taken place correctly.

(B) **DNA mapping** : A DNA map generated by restriction digestion can be used to find the relative positions of the genes. The restriction digestion generates different lengths of DNA which gives a specific pattern of bands after gel electrophoresis. This can be used for DNA fingerprinting. Additionally, it helps in diagnosis of single nucleotide polymorphisms (SNP) and insertions/deletions, identification of genetic disorder loci, assessment of the genetic diversity of populations, and parental testing.

(C) **Restriction fragment length polymorphism (RFLP)**: Restriction enzymes are used to digest genomic DNA for gene analysis using Southern blot. With the help of this technique, researchers can find copy number of a gene present in the genome of one individual as well as number of gene mutations (i.e., polymorphisms) within a population. The latter is called restriction fragment length polymorphism (RFLP).

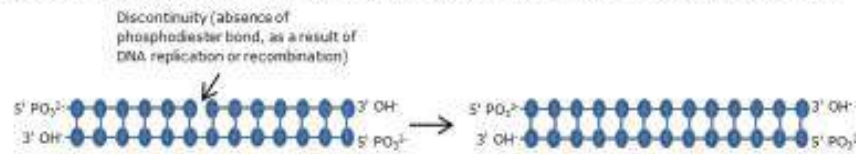
(D) **Studying epigenetic modification:** The sensitivity of restriction endonucleases toward methylated bases has been used to map modified bases within genomes.

(E) **Preparation of DNA libraries**: Restriction enzymes are used in SAGE (serial analysis of gene expression) for identification and quantification of a large number of mRNA transcripts in cancer research to diagnose mutations and study gene expression.

Ligases:

Ligases are enzymes that join the nucleic acid molecules together. These nucleic acids can either be DNA or RNA, and the enzymes are thus called DNA ligase and RNA ligase, respectively. DNA ligase catalyses the formation of a phosphodiester bond between the 5' phosphate of one strand and the 3' hydroxyl group of another. In nature the function of DNA ligase is to repair single strand breaks (discontinuities) that arise as a result of DNA replication and/or recombination. In recombinant DNA technology, ligases catalyse the joining of DNA of interest called as ‘insert’, with the vector molecule and the reaction is known as ligation.

Role of DNA Ligase in nature (to repair the discontinuities in DNA)



Role of DNA Ligase in genetic engineering (to join two DNA molecules)

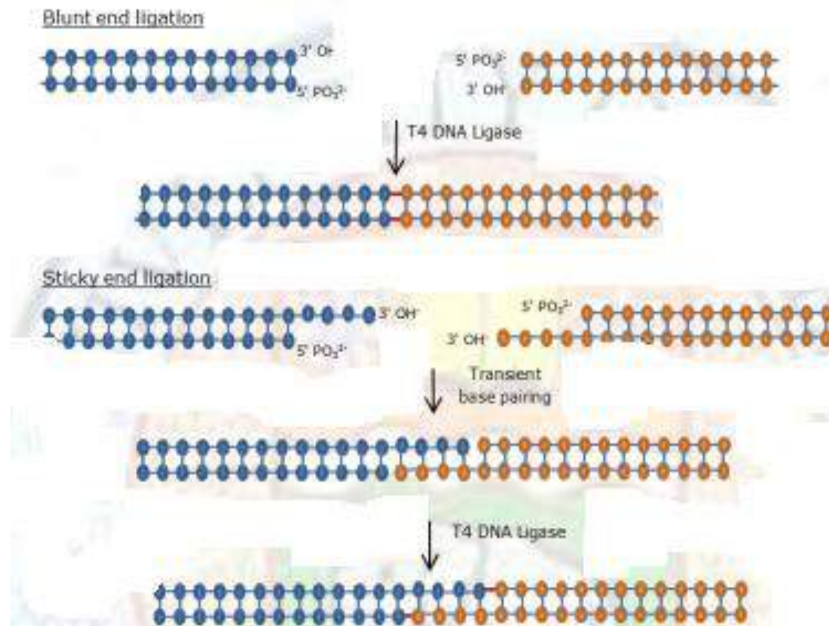


Figure:Diagrammatic depiction of reactions catalysed by DNA ligase. In nature they repair the missing phosphodiester bonds in one strand of DNA. In recombinant DNA technology, they are used to join two DNA molecules.

- DNA ligase catalyses the formation of phosphodiester bond between two deoxynucleotide residues of two DNA strands.
- DNA ligase enzyme requires a free hydroxyl group at the 3´ -end of one DNA chain and a phosphate group at the 5´-end of the other and requires energy in the process.
- *E.coli* and other bacterial DNA ligase utilizes NAD+ as energy donor, whereas in T4 bacteriophage, T4 DNA ligase uses ATP as cofactor.
- The role of DNA ligase is to seal nicks in the backbone of double-stranded DNA after lagging strand formation to join the okazaki fragments.
- This joining process is essential for the normal synthesis of DNA and for repairing damaged DNA. It has been exploited by genetic engineers to join DNA chains to form recombinant DNA molecules. Usually single stranded break are repaired using

the complimentary strand as the template but sometimes double stranded breaks can also be repaired with the help of DNA ligase IV.

• The most widely used DNA ligase is isolated from T4 bacteriophage. T4 DNA ligase needs ATP as a cofactor. The enzyme from *E. coli* uses cofactor NAD. Except this, the catalysis mechanism is somewhat similar for both the ligases. The role of cofactor is splitting and forming an enzyme-AMP complex which further aids in formation of phosphodiester bonds between hydroxyl and phosphate groups by exposing them.

Mechanism of Action of DNA Ligases:

• ATP, or NAD+, reacts with the ligase enzyme to form a covalent enzyme–AMP complex in which the AMP is linked to ε-amino group of a lysine residue in the active site of the enzyme through a phosphoamide bond.

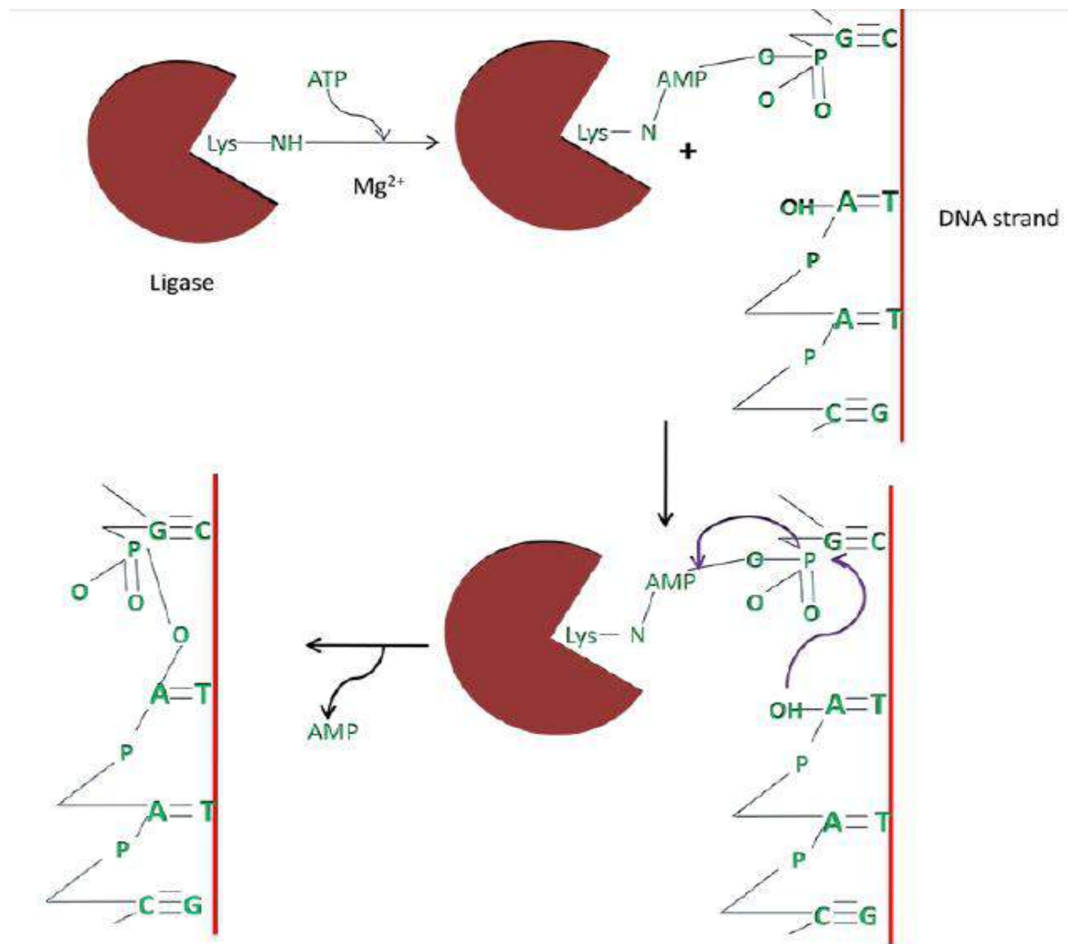
• The AMP moiety activates the phosphate group at the 5´-end of the DNA molecule to be joined. It is called as the donor.

• The final step is a nucleophilic attack by the 3´-hydroxyl group on this activated phosphorus atom which acts as the acceptor. A phosphodiester bond is formed and AMP is released.

• The reaction is driven by the hydrolysis of the pyrophosphate released during the formation of the enzyme–adenylate complex. Two high-energy phosphate bonds are spent in forming a phosphodiester bond in the DNA backbone with ATP serving as energy source.

• The temperature optimum for T4 DNA ligase mediated ligation *in vitro* is 16˚C. However ligation is also achieved by incubation at 4˚C by incubating over night or at room temperature condition by incubating for 30 minutes.

• Adenylate and DNA-adenylate are the important intermediates of the phosphodiester bond forming pathway.



The mechanism of DNA joining by DNA ligase.

Application:

• DNA ligase enzyme is used by cells to join the "okazaki fragments" during DNA replication process. In molecular cloning, ligase enzyme has been routinely used to construct a recombinant DNA. Followings are some of the examples of application of ligase enzyme in molecular cloning.Joining of adapters and linkers to blunt end DNA molecule.

• Cloning of restricted DNA to vector to construct recombinant vector.

DNA Polymerases

DNA polymerases are enzymes that catalyse the synthesis of a new DNA strand from a pre-existing strand.The enzyme adds deoxyribonucleotides to the free 3’-OH of the chainundergoing elongation. The direction of synthesis is 5’-3’.Ithas three major requirements for its activity;

(1) a template strand for which the enzyme synthesizes a complementary strand;

(2) a primer with a free 3’-OH group that hybridizes with the template to form a double stranded region that initiates the polymerization and

(3) a pool of all the four dNTPs that are used to synthesize the new DNA strand. In addition, some cofactors like Mg2+ ions may be required in a buffer solution with correct pH for optimum activity.

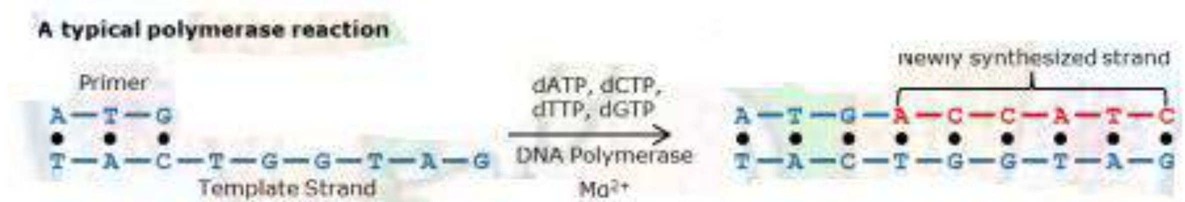
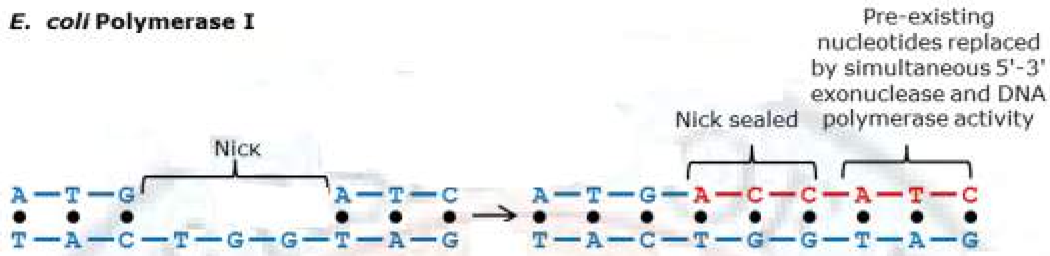


Figure:A typical reaction catalysed by DNA polymerase.

Different types of DNA polymerases are used in recombinant DNA technology.

1. E. coliDNA Polymerase I
2. Klenow Fragment
3. Thermostable DNA Polymerase
4. Reverse Transciptase

E. coli DNA polymerase I (PolI) is an enzyme that has both DNA polymerase as well as DNA nuclease activity. This enzyme binds to the ‘nick’ region (region of a double stranded DNA where one or more nucleotides of one strand are missing, making it single stranded). The polymerase activity of the enzyme synthesizes the complementary strand for the nick and continues synthesizing the complete new strand by simultaneously degrading the pre-existing strand by its 5'-3'exonuclease activity.



Reaction catalysed by E. coliDNA polymerase I.

Different domains of the E. coli PolI are responsible for different catalytic activities. The C-terminal is responsible for the polymerase activity whereas the N-terminal of the enzyme catalyses the 5'-3'exonuclease activity. The central region of the enzyme is responsible for 3'-5' exonuclease activity that can remove any misread nucleotide and hence acts as a proofreading mechanism.

If the E. coli Pol I holoenzyme is treated with amild protease, it results in the formation of two fragments. A larger fragment retaining both 5'-3' polymerase and 3'-5' exonuclease activities; while the smaller one has only the 5'-3' exonuclease activity. The larger fragment is known as 'Klenow fragment'.

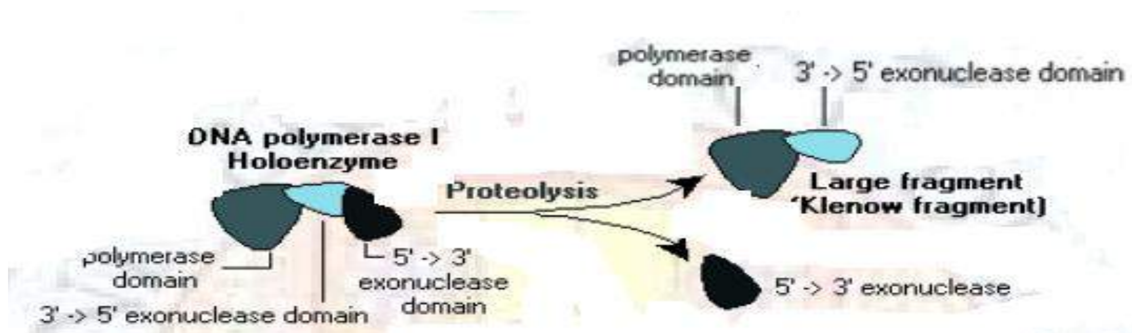


Figure: Diagrammatic representation of proteolytic cleavage of E. coli Pol I holoenzyme into two fragments having different catalytic activities.

This Klenow fragment can synthesize the new DNA strand complementary to the template but cannot degrade the existing strand. Klenow fragment is predominantly used in DNA sequencing. Other uses in recombinant DNA technology where Klenow fragment is used are

• Synthesis of double stranded DNA from single stranded template.

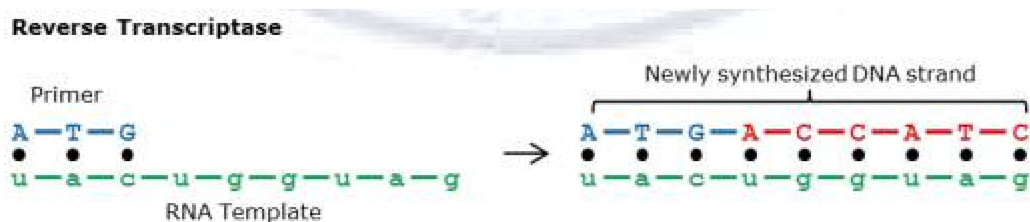
• Filling of 5' overhangscreated by restrictionenzymesto create blunt ends.

• Digestion of protruding 3' overhangs to produce blunt ends.

Figure: Diagrammatic representation of reaction catalyzed by Klenow fragment. The enzyme fills up nicks in double stranded DNA molecules.

Thermostable DNA polymerases are a class of DNA polymerases that remain functional even at high temperatures. In other words, they are resistant to denaturation by heat treatment. They are isolated from the bacterium Thermusaquaticus that lives in hot springs. The enzyme isolated from this bacterium is known as'TaqPolymerase'. Major application of Taq polymerase is in the polymerase chain reaction (PCR) technique which is used to amplify DNA fragments. PCR is covered in detail in another chapter that will take care of the applications of Taq polymerase. The readers can visit the following link to know more about PCR.

Reverse transcriptase (RT) is aRNA dependent DNA polymerase found in RNA viruses also called as retroviruses. This enzyme is involved in the replication of retroviruses, where the RNA genome is first converted into DNA and then integrated into the host. RT uses mRNA template instead of DNA for synthesizing new DNA strand.The complementary DNA strand formed on the mRNA template is called the complementary DNA (cDNA). RT also shows RNAseH activity that degrades the RNA molecule from a DNA-RNA hybrid. Formation of a double stranded cDNA from the mRNA molecule using RT finds applications in genetic engineering. The cDNA thus formed from any mRNA can be cloned in an expression vector and its respective protein can be made to express in large quantities.



Diagrammatic representation of reaction catalyzed by reverse transcriptase.

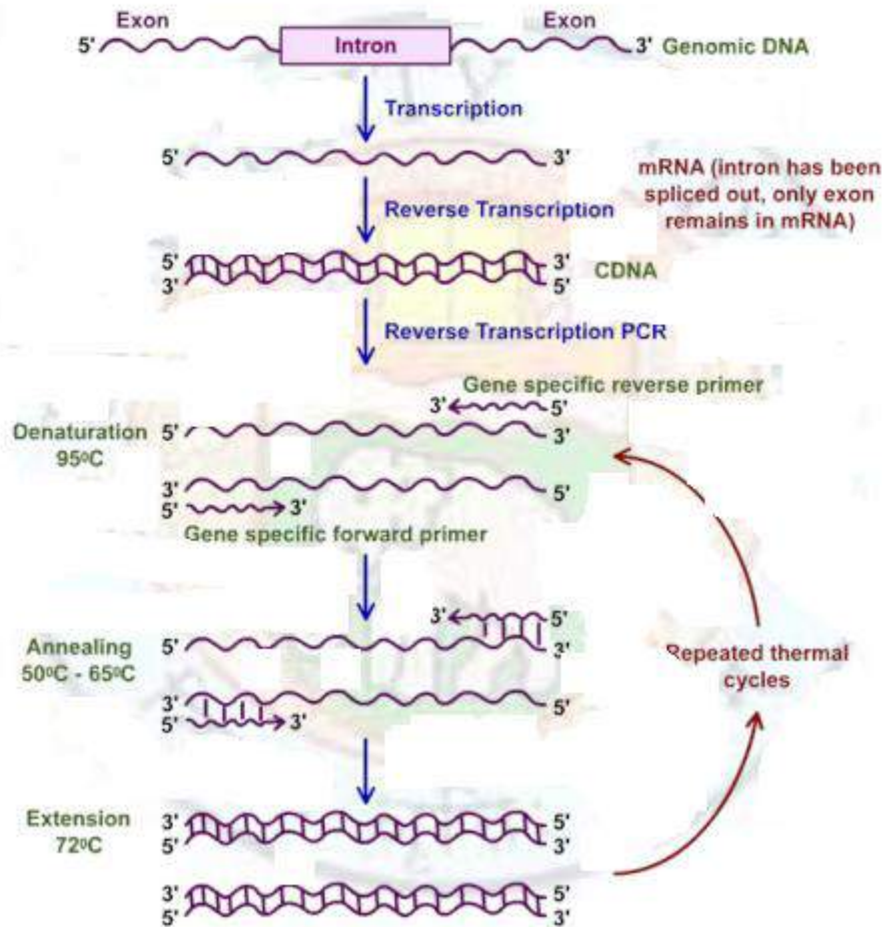


Figure: Diagrammatic representation showing reverse transcriptase PCR.

DNA modifying enzymes

Addition or deletion of different chemical groups from DNA molecules comes in handy for molecular cloning purposes. Three such enzymes known for performing such modifications are:

1. **Alkaline phosphatase (AP):** This group of enzymes removes the phosphate group (PO3 2-) from 5' terminus of the DNA molecule. It is active at alkaline pH, hence the name 'alkaline phosphatase'. Commercially, it is obtained from three major sources, viz., E. coli (bacteria), calf intestine and arctic shrimp.

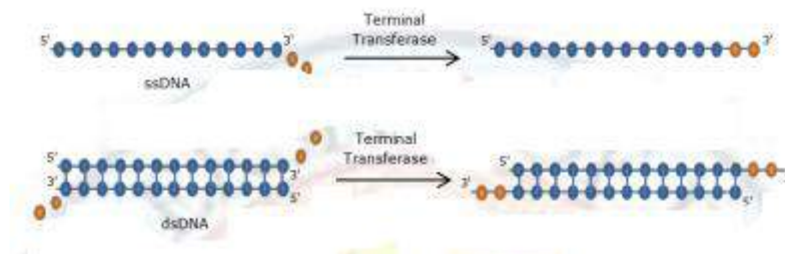
Diagrammatic representation of reaction catalyzed by alkaline phosphatase.

Treatment of vector DNA with AP is important in cloning experiments, as removal of 5'phosphate prevents self-annealing of the digested vector and increases the possibility of ligating with the insert DNA fragment in the presence of ligase. Also, radiolabeled DNA probes are prepared by initially removing the5'PO3 2- by AP treatment, followed by polynucleotide kinase treatment in the presence of radioactive.

2.Phosphate.Polynucleotide kinase (PNK): This group of enzymes perform a role completely opposite to the one performed by AP. PNK catalyses the transfer of a phosphate group from ATP to the 5' terminus of the DNA molecule. This enzyme is obtained from E. coli infected with T4 phage.

Diagrammatic representation of reaction catalysed by polynucleotide kinase.

3.**Terminal transferase:** This group of enzymes catalyses the addition of one or more deoxyribonucleotides to the 3' terminus of the DNA molecule.The enzyme can work on both double stranded as well as single stranded DNA molecules without the need of any primers. The enzyme is obtained from calf thymus tissue.



Diagrammatic representation of the reaction catalysed by terminal transferase.

Topoisomerases:

Topoisomerases are enzymes that participate in the overwinding or underwinding of DNA. The winding problem of DNA arises due to the intertwined nature of its double-helical structure. During DNA replication and transcription, DNA becomes overwound ahead of a replication fork. If left unabated, this torsion would eventually stop the ability of DNA or RNA polymerases involved in these processes to continue down the DNA strand.

In order to prevent and correct these types of topological problems caused by the double helix, topoisomerases bind to DNA and cut the phosphate backbone of either one or both the DNA strands. This intermediate break allows the DNA to be untangled or unwound, and, at the end of these processes, the DNA backbone is resealed again. Since the overall chemical composition and connectivity of the DNA do not change, the DNA substrate and product are chemical isomers, differing only in their global topology, resulting in the name for these enzymes. Topoisomerases are isomerase enzymes that act on the topology of DNA.

- DNAs that differ only in linking number are known as topoisomers.
- Topoisomerases bind to DNA and cut the phosphate backbone of either one or both the strands of DNA
- There are at least four different topoisomerases in *E. coli*, distinguished by Roman numerals I through IV.
- Topoisomerases are divided into **two classes** (type 1 and type 2) depending on the number of strands cut in one round of action (Both these classes of enzyme utilize a conserved tyrosine for their catalytic activity).

Type 1 topoisomerase: Topoisomerase I and topoisomerase III

- They cuts one strand of the DNA double helix, relaxation of DNA occurs, finally the cut strand is religated.
- Catalyze the relaxation of supercoils in DNA by changing their linking number in increments of one turn until the supercoil is entirely relaxed.
- The action of topoisomerase I change the linking number by decreasing the number of twists.
- Topoisomerase I and Topoisomerase III of *E. coli* belong to type 1 topoisomerases.
- Generally relax DNA by removing negative supercoils. Relaxation of supercoiled DNA is a thermodynamically favorable process (do not require ATP hydrolysis)

Type2 topoisomerase: Topoisomerase II

- Cuts both strands of DNA double helix, passes the unbroken DNA helix through the break, and then re-ligates the cut strands.
- Change *Lk* in increments of 2.
- Use ATP hydrolysis free energy to add negative supercoils to DNA.

  Eukaryotic cells also have type 1 and type 2 topoisomerases.
- The **type 1** enzymes are topoisomerases I and topoisomerase III
- The **type 2** enzymes are topoisomerases IIα and topoisomerase IIβ.

Topoisomerase function

1) Topoisomerases play important role in **growing fork movement** and in **resolving (untangling) finished chromosomes after DNA duplication**. Type II topoisomerases separate both replicated circular and linear DNA chromosomes.

a) Function of Topoisomerase II in DNA Replication (growing fork movement)

- Most cellular DNA is negatively supercoiled, as a result the separation of nucleotide strands during replication is easy.
- Unwinding of duplex DNA is necessary step to start replication because then only each single strand of DNA can serve as a template for the synthesis of a complementary strand.
- When helicase enzyme separates the annealed DNA helix, the region ahead of the ezyme complex will experience a stress. The reason for the developement of this stress is due to the cyclic nature of the DNA.
- This stress is offset by positive supercoils in front of the complex. DNA is rewound behind the complex and compensatory negative supercoils will occur
- DNA gyrase is the only one that can introduce negative supercoils into the structure of DNA to relieve the strain from the supercoiling. by doing so it will allow the helicase to continue splitting the DNA strands
- In eukaryotes the positive supercoil generated as a result of unwinding is removed by topoisomerase IIα.
- If they are not functioning, DNA helicase would not be able to continue splitting the DNA strands. This will lead to the haulting of transcription especially elongation step.

Glossary

Bacteriophage: They are bacteria infecting viruses.

DNA ladder: A mixture of DNA fragments, used in gel electrophoresis as size markers.

DNA ligase:An enzyme that, in the cell, repairs single-stranded discontinuities in double-stranded DNA molecules. Purified DNA ligase is used in gene cloning to join DNA molecules together.

DNA polymerase: An enzyme that synthesizes DNA on a DNA or RNA template.

Double digestion: Cleavage of a DNA molecule with two different restriction endonucleases simultaneously.

DNA sequencing: Determination of the order of nucleotides in a DNA molecule.

Endonuclease: An enzyme that breaks phosphodiester bonds within a nucleic acid molecule.

Exonuclease: An enzyme that removes nucleotides from the ends of a nucleic acid molecule.

Gel electrophoresis: Electrophoresis performed in a gel matrix so that molecules of similar electric charge can be separated on the basis of size.

Homopolymer tailing: Attachment of a sequence of identical nucleotides (e.g., CCCCC) to the end of a nucleic acid molecule. It is generally done by terminal transferase.

Host-controlled restriction: A mechanism by which some bacteria prevent phage attack through the synthesis of a restriction endonuclease that cleaves the non-bacterial DNA.

Indel: A position where a DNA sequence has been inserted into or deleted from a genome.

Insert: DNA fragment that is inserted in the cloning vector

Klenow fragment (of DNA polymerase I): A DNA polymerase enzyme, obtained byproteolysis of E. coli DNA polymerase I, used mostly in chain termination DNA sequencing.

Methyl Transferase: An enzyme that adds methyl group to the nucleotides in a DNA molecule.

Nick:A single-strand break, involving the absence of one or more nucleotides, in a double-stranded DNA molecule.

PCR (Polymerase Chain Reaction):A technique that enables amplification of a DNAmolecule using thermostable polymerases.

Plasmid: A usually circular piece of DNA, primarily independent of the host chromosome, often found in bacteria and some other types of cells. It is generally used in recombinant DNA technology as cloning vector.

Polylinker: A synthetic double-stranded piece of DNA carrying a number of restriction sites.

Primer: A short single-stranded oligonucleotide that attaches to the template strand by base pairing and initiates the synthesis of the DNA strand by a DNA polymerase enzyme.

Recombinant DNA: A DNA molecule created in the test tube by ligatingtogether pieces of DNA that are not normally contiguous.

Recombinant DNA technology: All of the techniques involved in the construction, studyand use of recombinant DNA molecules.

Restriction endonuclease: An endonuclease that cuts DNA molecules only at specific nucleotide sequences.

Restriction fragment length polymorphism (RFLP): A mutation that results in alteration of a restriction site and hence a change in the pattern of fragments obtained when a DNA molecule is cut with a restriction endonuclease.

Restriction Map:A map showing the positions of different restriction sites in a DNAmolecule.

Reverse transcriptase: An RNA-dependent DNA polymerase that uses RNA template to synthesize a complementary DNA molecule.

Taq DNA polymerase: The thermostable DNA polymerase used in PCR, isolated from Thermus aquaticus.

Template:A single-stranded polynucleotide (or region of a polynucleotide) that directssynthesis of a complementary polynucleotide.

Vector: A DNA molecule, capable of replication in a host organism, into which a geneis inserted to construct a recombinant DNA molecule.

A vector is a DNA molecule which is used for transporting exogenous DNA into the host cell. A vector is capable of self-replication and stable integration inside the host cell.

General characteristics of a vector:

- It should have an Origin of Replication, known as ori, so that the vector is capable of autonomous replication inside the host organism.
- It must possess a compatible restriction site for insertion of DNA molecule.
- A vector should always harbour a selectable marker to screen the recombinant organism. This selectable marker can be an antibiotic resistance gene.
- For easy incorporation into the host machinery, a vector should itself be small in size and be able to integrate large size of the insert.

CLONING VECTOR

A cloning vector is also a fragment of DNA which is capable of self-replication and stable maintenance inside the host organism. It can be extracted from a virus, plasmid or cells of a higher organism. Most of the cloning vectors are genetically engineered. It is selected based upon the size and the kind of DNA segment to be cloned.

The cloning vectors must possess the following general characteristics:

- It should small in size.
- It must have an origin of replication.
- It must also be compatible with the host organism.
- It must possess a restriction site.
- The introduction of donor fragment must not intervene with the self-replicating property of the cloning vector.
- A selectable marker, possibly an antibiotic resistance gene, must be present to screen the recombinant cells.

- It should be capable of working under the prokaryotic as well as the eukaryotic system.
- Multiple cloning sites should be present.

Importance of Cloning Vectors

Cloning Vectors are used as the vehicle for transporting foreign genetic material into another cell. This foreign segment of DNA is replicated and expressed using the machinery of the host organism.

A cloning vector facilitates amplification of a single copy DNA molecule into many copies. Molecular gene cloning is difficult without the use of the cloning vectors.

Features of Cloning Vectors

1. Origin of Replication (ori)

- A specific set/ sequence of nucleotides where replication initiates.
- For autonomous replication inside the host cell.
- Foreign DNA attached to ori also begins to replicate.

2. Cloning Site

- Point of entry or analysis for genetic engineering.
- Vector DNA at this site is digested and foreign DNA is inserted with the aid of restriction enzymes.
- Recent works have discovered plasmids with multiple cloning sites (MCS) which harbour up to 20 restriction sites.

3. Selectable Marker

- Gene that confers resistance to particular antibiotics or selective agent which, under normal conditions, is fatal for the host organism.
- Confers the host cell the property to survive and propagate in culture medium containing the particular antibiotics.

4. Marker or Reporter Gene

- Permits the screening of successful clones or recombinant cells.
- Utilised extensively in blue-white selection.

5. Inability to Transfer via Conjugation

- Vectors must not enable recombinant DNA to escape to the natural population of bacterial cells.

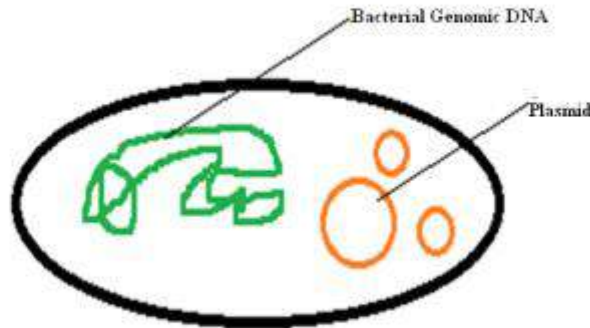
Types of Cloning Vectors

- Plasmids
- Bacteriophage
- Phagemids
- Cosmids
- Bacterial Artificial Chromosome (BAC)
- Yeast Artificial Chromosome (YAC)
- Human Artificial Chromosome (HAC)
- Retroviral Vectors

Plasmids

Plasmids are naturally occurring extra chromosomal double-stranded circular DNA molecules which can autonomously replicate inside bacterial cells. Plasmids range in size from about 1.0 kb to over 250 kb. • Plasmids encode only few proteins required for their own replication (replication proteins) and these proteins encoding genes are located very close to the ori. All the other proteins required for replication, e.g. DNA polymerases, DNA ligase, helicase, etc., are provided by the host cell.Thus, only a small region surrounding the ori site is required for replication. Other parts of the plasmid can be deleted and foreign sequences can be added to the plasmid without compromising replication.

• The host range of a plasmid is determined by its ori region. Plasmids whose ori region is derived from plasmid Col E1 have a restricted host range. They only replicate in enteric bacteria, such as E. coli, Salmonella, etc. Plasmids of the RP4 type will replicate in most gram negative bacteria, to which they are readily transmitted by conjugation. Plasmids like RSF1010 are not conjugative but can be transformed into a wide range of gram -ve bacteria. Plasmids isolated from Staphylococcus aureus have a broad host range and can replicate in many other gram-positive bacteria.



- Plasmids were the first vectors to be used in gene cloning.
- They are naturally occurring and autonomously replicating extra-chromosomal double-stranded circular DNA molecules. However, not all plasmids are circular in origin.
- They are present in bacteria, archaea, and eukaryotes.
- The size of plasmids ranges from 1.0 kb to 250 kb.
- DNA insert of up to 10 kb can be cloned in the plasmids.
- The plasmids have high copy number which is useful for production of greater yield of recombinant plasmid for subsequent experiments.
- The low copy number plasmids are exploited under certain conditions like the cloned gene produces the protein which is toxic to the cells.
- Plasmids only encode those proteins which are essential for their own replication. These protein-encoding genes are located near the ori.

Examples: pBR322, pUC18, F plasmid, Col plasmid.

The plasmids are divided into 6 major classes as described below depending on the phenotype:

i) **Resistance or R plasmids** carry genes which give resistance to the bacteria from one or more chemical agents, such as antibacterial agents. R plasmids are very important in clinical microbiology as they can have profound consequences in the treatment of bacterial infections. Eg: RP4 plasmid, which is commonly found in Pseudomonas and in many other bacteria.

ii) **Fertility or F plasmids** are conjugative plasmid found in F+ bacterium with higher frequency of conjugation. F plasmid carries transfer gene (tra) and has the ability to form Conjugation Bridge (F pilus) with Fbacterium. Eg: F plasmid of E. coli.

iii) **Col plasmids** have genes that code for colicins, proteins that kill other bacteria. Eg: ColE1 of E. coli.

iv) **Degradative plasmids** allow the host bacterium to metabolize unusual molecules such as toluene and salicylic acid. Eg TOL of Pseudomonas putida.

v) **Virulence plasmids** confer pathogenicity on the host bacterium. Eg: Ti plasmids of Agrobacterium tumefaciens, which induce crown gall disease on dicotyledonous plants.

vi) **Cryptic Plasmids** do not have any apparent effect on the phenotype of the cell harboring them. They just code for enzymes required for their replication and maintenance in the host cell.

Based on the origin or source of plasmids, they have been divided into two major classes: such as natural and artificial.

i) Natural plasmids: They occur naturally in prokaryotes or eukaryotes. Example: ColE1.

ii) Artificial plasmids: They are constructed in-vitro by re-combining selected segments of two or more other plasmids (natural or artificial). Example: pBR322.

Natural vectors

pSC101:

pSC101 is a DNA plasmid that is used as a cloning vector in genetic cloning experiments. pSC101 was the first cloning vector, used in 1973 by Herbert Boyer and Stanley Norman Cohen. They are demonstrated that a gene from a frog could be transferred into bacterial cells and then expressed by the bacterial cells. The plasmid is a natural plasmid from *Salmonella typhimurium*.

Natural plasmid vectors for E. coli pSC101

• First used for in vitro cloning of eukaryotic DNA

• 9kbp in size • Low copy number(1-2 copies)

• Has advantage of a single Eco RI site at which DNA can be inserted

• Has selectable marker for tetracycline resistance

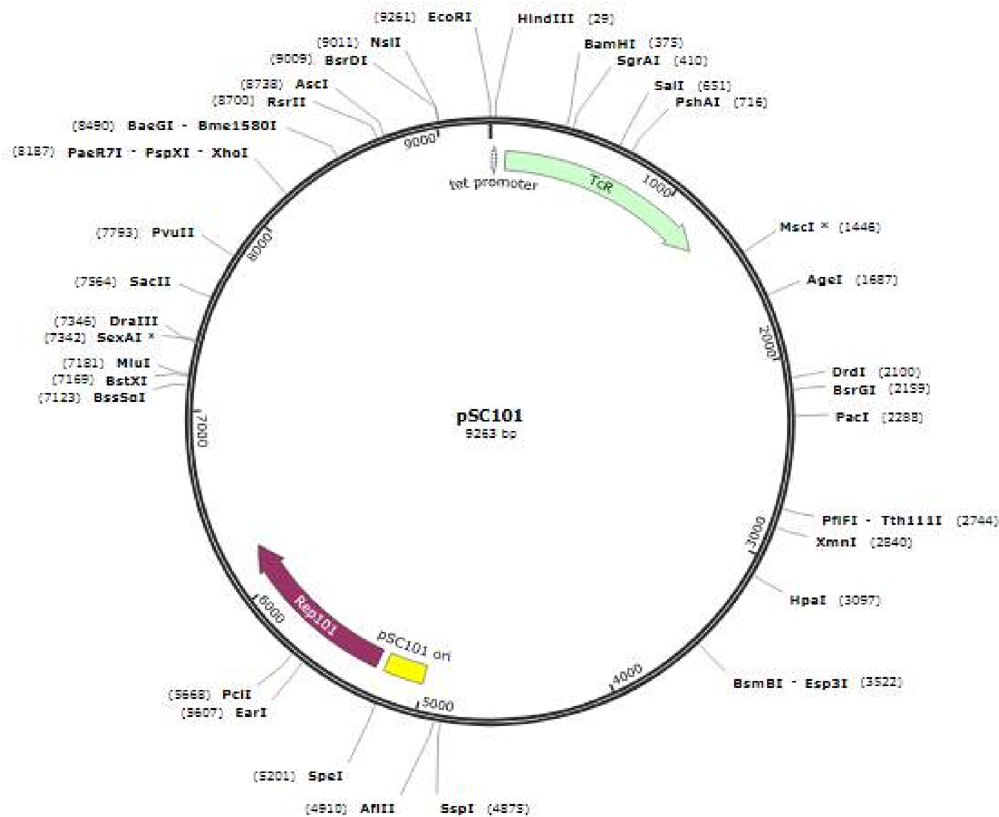
• Derived from the conjugative plasmid R6-5 Disadvantages are

• Large size

• Stringent replicative control

• Low copy number

• Low insert capability



pSF 2124(RSF2124)

• Produced by transfer of the ampicillin resistance gene

•Has ability for colicin biosynthesis •Has high copy number

•Has single sites for Bam H1 and EcoR1

• Not currently used as vector as it does not provide easy selection by insertional inactivation

•Mobilizable plasmid

pMB1: Initiating replication

pMB1 (*ori*) contains regions that promote the synthesis of RNAI and RNAII. Replication if the plasmid is usually initiated by RNAII, which is transcribed from the plasmid 550 nucleotides (nt) upstream from *ori*, and hybridizes strongly to the plasmid. The formation of this hybrid at the origin is a critical prerequisite for plasmid replication. The RNA part of this hybrid becomes a substrate for RNaseH, which digests away RNA II yielding a 550 nt molecule that can be used as a primer for DNA polymerase I for the initiation of replication of the entire plasmid.

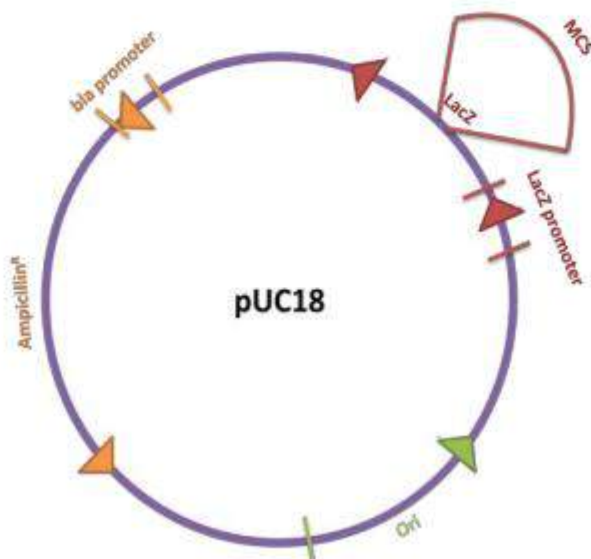
Artificial Vectors

Artificial plasmids are designed by combining different elements from diverse sources. Artificial plasmids vectors are classified into two broad types based on their use:

1. Cloning vector
2. Expression vector

Apart from the following, there is another class of vectors known as shuttle vector. Shuttle vectors can be propagated in two or more different host species (both in prokaryotes and eukaryotes). Hence, inserted DNA can be manipulated and replicated in two different cellular systems. Cloning vectors are designed for efficient transfer of foreign DNA into the host. Expression vectors have efficient machinery for cloning and expression of foreign gene in the host system. Selection of a vector depends upon various criteria decided by the experimental goal.

CloningVector: A cloning vector is defined as a vector used for replication of a cloned DNA fragment in a host cell. These vectors are frequently engineered to contain "ori" – origin of replication sites particular to the host organism. Examples of commonly used cloning vectors are: pUC18, pUC19, pBluescript vectors etc.



Important features of a cloning vector used to carry DNA molecules are as follows:-

• **Stability in host cell**: Vectors should be stabile in host cell after introduction and should not get lost in subsequent generations. This permits replication of vectors producing large copies of gene of interest.

• **Ability to control their own replication**: This property enables them to multiply and exist in high copy number.

• **Small size**: Ideal vector should be less than or equal to 10kb. The small size is essential for easy introduction in cell by transformation, transduction and electroporation.

• **Multiple cloning sites**: This property permits the insertion of gene of interest and plasmid re-circularization.

• **Should not be transferred by conjugation**: This property of vector molecule prevents recombinant DNA to escape to natural population of bacteria.

• **Selectable make gene**: Vector molecules should have some detectable traits. These traits enable the transformed cells to be identified among the nontransformed ones. eg. antibiotic resistance gene.

Vector	Insert size	Source	Application
Plasmid	$\leq$ 15 kb	Bacteria	Subcloning and downstream manipulation, cDNA cloning and expression assays
Phage	5-20 kb	Bacteriophage λ	Genomic DNA cloning, cDNA cloning and expression library
Cosmid	35-45 kb	Plasmid containing a bacteriophage λ *cos* site	Genomic library construction
BAC (bacterial artificial chromosome)	75-300 kb	Plasmid ocntaining *ori* from *E.coli* F- plasmid	Analysis of large genomes
YAC (yeast artificial chromosome)	100-1000 kb (1 Mb)	*Saccharomyces cerevisiae* centromere, telomere and autonomously replicating sequence	Analysis of large genome, YAC transgenic mice
MAC (mammalian artificial chromosome)	100 kb to $>$ 1 Mb	Mammalian centromere, telomere and origin of replication	Under development for use in animal biotechnology and human gene therapy

Examples of Cloning Vector:

pBR322 is a widely-used E. coli cloning vector.

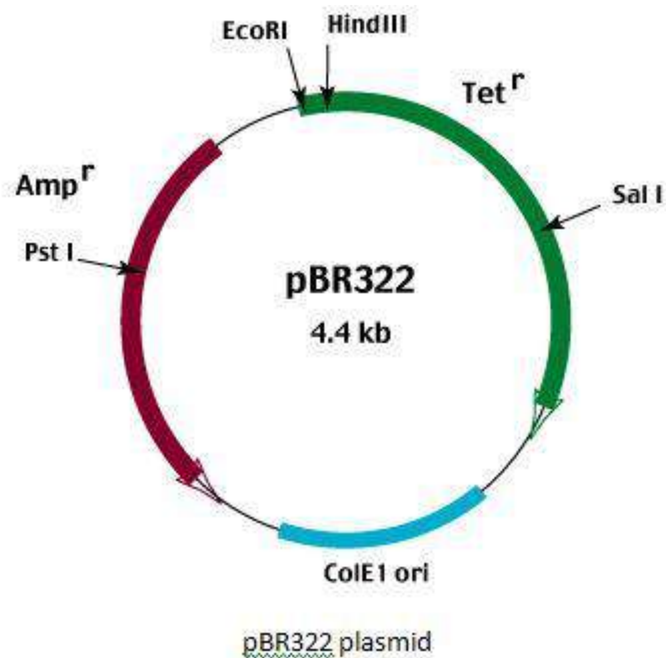
It was created in 1977 in the laboratory of Herbert Boyer at the University of California San Francisco. The p stands for "plasmid" and BR for "Bolivar" and "Rodriguez", researchers who constructed it.

• pBR322 is 4361 base pairs in length.

• pBR322 plasmid has the following elements: } “rep” replicon from plasmid pMB1 which is responsible for replication of the plasmid. } “rop” gene encoding Rop protein. Rop proteins are associated with stability of RNAI-RNAII complex and also decrease copy number. The source of “rop” gene is pMB1plasmid. } “tet” gene encoding tetracycline resistance derived from pSC101 plasmid. } “bla” gene encoding β lactamase which provide ampicillin resistance (source: transposon Tn3)

Nomenclature of plasmid cloning vector: pBR322 cloning vector has the following elements:

- p= plasmid
- B= Bolivar (name of the scientist)
- R= Rodriguez (name of the scientist)
- 322= number of plasmid discovered in the same lab



pBR322 plasmid

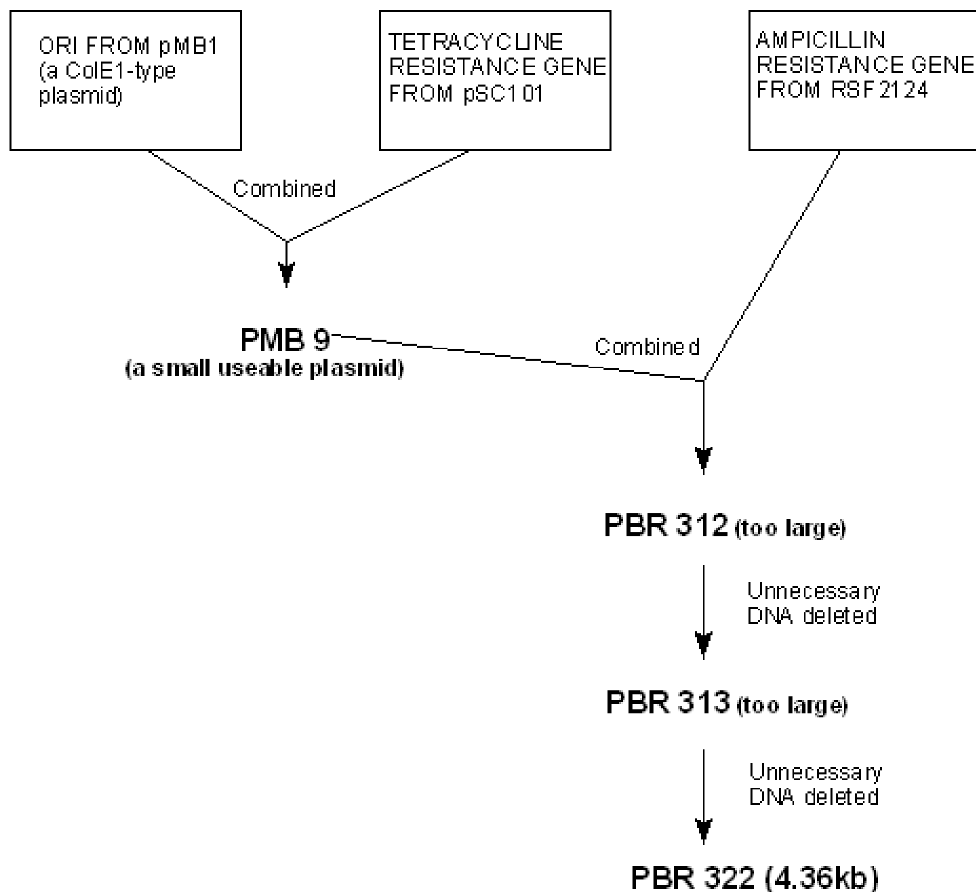
Advantages of using Plasmids as vectors:

- Easy to manipulate and isolate because of small size.

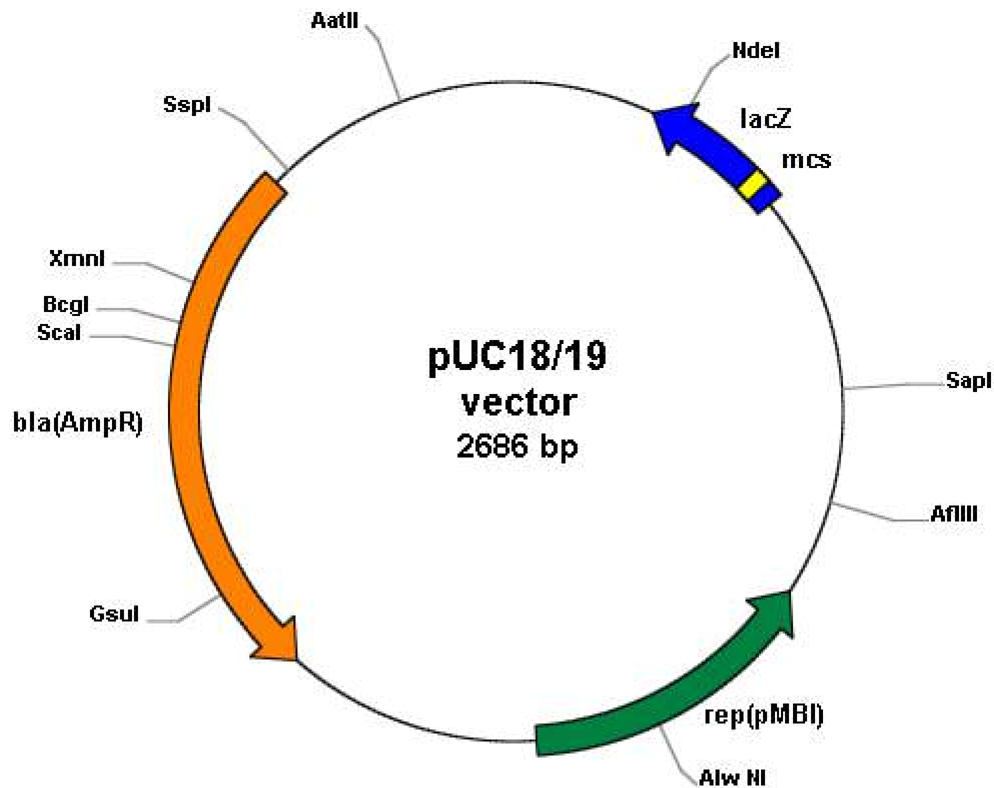
- More stable because of circular configuration.
- Replicate independent of the host.
- High copy number.
- Detection easy because of antibiotic-resistant genes.

Disadvantages of using Plasmids as vectors:

- Large fragments cannot be cloned.
- Size range is only 0 to 10kb.
- Standard methods of transformation are inefficient.



pUC plasmids:



• pUC plasmids are small, high copy number plasmids of size 2686bp.

• This series of cloning vectors were developed by Messing and co-workers in the University of California. The p in its name stands for plasmid and UC represents the University of California.

• pUC vectors contain a lacZ sequence and multiple cloning site (MCS) within lacZ. This helps in use of broad spectrum of restriction endonucleases and permits rapid visual detection of an insert.

• pUC18 and pUC19 vectors are identical apart from the fact that the MCS is arranged in opposite orientation.

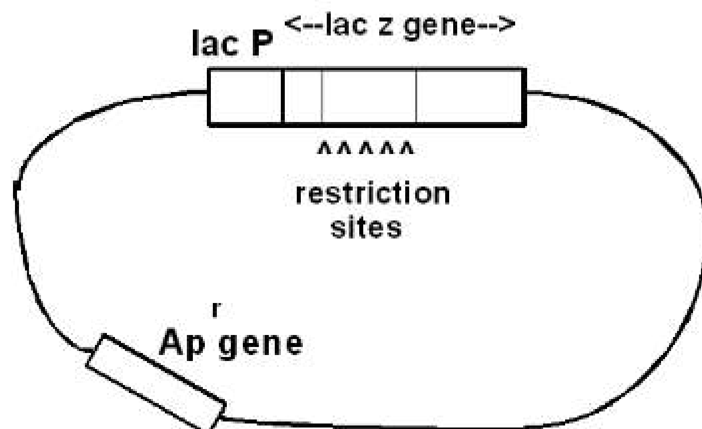
• pUC vectors consists of following elements: } pMB1 "rep" replicon region derived from plasmid pBR322 with single point mutation (to increase copy number). } "bla" gene encoding β lactamase which provide ampicillin resistance which is derived from pBR322. This site is different from pBR322 by two point mutations. } E.coli lac operon system.

• "rop" gene is removed from this vector which leads to an increase in copy number. An MCS is a short DNA sequence consisting of restriction sites for many different restriction endonucleases. MCS escalates the number of potential cloning strategies

available by extending the range of enzymes that can be used to generate a restriction fragment suitable for cloning. By combining them within a MCS, the sites are made contiguous, so that any two sites within it can be cleaved simultaneously without excising vector sequences. The MCS is inserted into the lacZ sequence, which encodes the promoter and the α-peptide of β-galactosidase. Insertion of the MCS into the lacZ fragment does not affect the ability of the α-peptide to mediate complementation, while cloning DNA fragments into the MCS does. Therefore, recombinants can be detected by blue/white screening on growth medium containing X gal in presence of IPTG as an inducer.

pUC plasmids

These contain several restriction sites within a ß-galactosidase [**lac z**] gene from the lac operon (actually only part of it to save space). These plasmids also contain an ampicillin resistance gene (Ap^r) for selection of host cells containing plasmid from those not containing plasmid.



Insertion of foreign DNA into this gene leads to loss of ß-galactosidase activity by insertional inactivation.

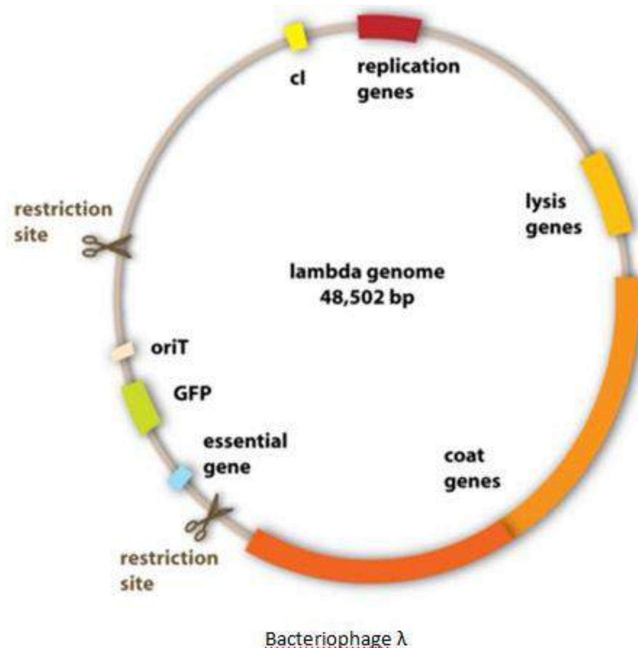
Bacteriophage

- Bacteriophages or phages are viruses which infect bacterial cells.
- The most common bacteriophages utilized in gene cloning are Phage λ and M13 Phage.
- A maximum of 53 kb DNA can be packaged into the phage.
- If the vector DNA is too small, it cannot be packaged properly into the phage.

Examples: Phage Lambda, M13 Phage, etc.

Phage Lambda λ

- It has head, tail, and tail fibers.
- Its genome consists of 48.5 kb of DNA and 12 bp ss DNA which comprise of sticky ends at both the terminals. Since these ends are complementary, they are cohesive and also referred to as cos sites.
- Infection by λ phage requires adsorption of tail fibers on the cell surface, contraction of the tail, and injection of the DNA inside the cell.



Bacteriophage λ

Phage Vectors:

To insert DNA fragments of more than 10 kb, normally plasmids are not the suitable vehicles, as large inserts may trigger plasmid rearrangement or affect plasmid replication. This leads to development of a new class of vectors based on bacteriophages. Amongst various bacteriophages available such as λ, T4, T5, and T7 phages; the λ phage gained favourable attention due to its unique life cycle.

λ phage

Bacteriophage λ contains ~49kb of DNA and has a very efficient mechanism for delivering its genome into a bacterium. Two key features contribute to its utility as a vector to clone larger DNA fragments:

1. One-third λ genome is nonessential and could be replaced with foreign DNA. Approximately 24.6kb of λ genome can be deleted, hence maximum insert size could be upto 26 kb.

2. Packing of DNA in phage could only take place if the size is between 40 and 52 kb long, a constraint that can be used to ensure packaging.

Two problems had to be addressed before λ -based cloning vectors could be developed:

- • The size limitation of the insert is determined by the genome size of phage λ (distance between the cos sites). The size range of the modified genome size should be within the range between 78-105% of the genome size for proper packaging. If >2.4kb is inserted to full length λ vector, the packaging efficiency is reduced. Hence λ vector should be smaller in size than wild type λ genome.
- • The large λ genome has a few unique recognition sequences for bacterial restriction endonuclease. Bacterial restriction digestion system may target the modified vector and cleave the λ DNA molecule. This limitation can be overcome by replacing or mutating the restriction sites.

Two types of vector have been developed using λ genome:

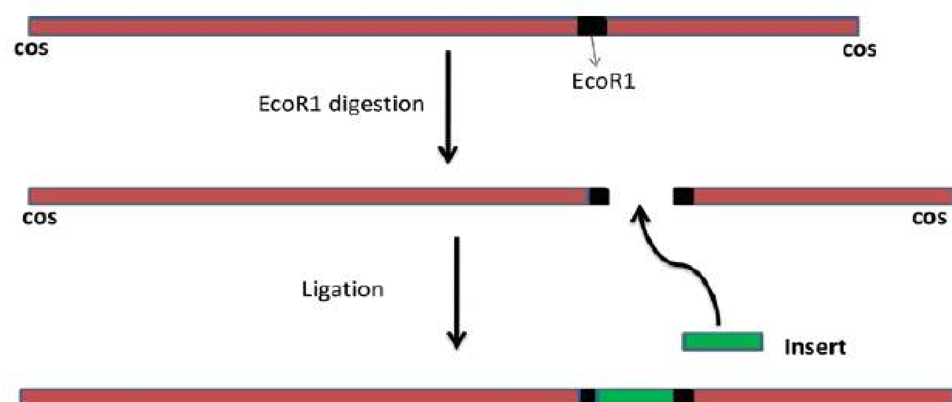
1) Insertion vectors:

• Foreign DNA sequence is inserted into the λ genome without any significant change of the wild type genome.

• Smaller insert size (upto ~10kb).

• They may contain a multiple cloning site inserted in *lacZ* system for screening of recombinant bacterial colonies.

• Can be used to clone smaller DNA molecule. Eg: λ ZAP, λ gt etc.



2) Replacement vectors:

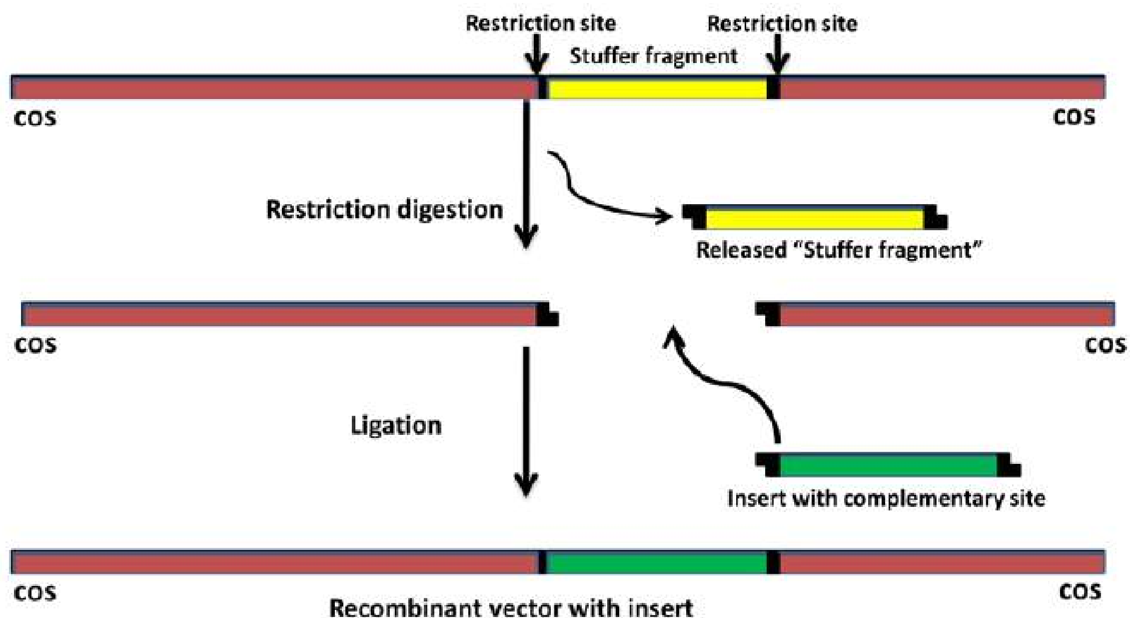
• Full length λ molecule having two identical restriction sites flanked by “stuffer fragment”.

• Stuffer fragment is replaced by foreign DNA during restriction cloning.

• The vector without the foreign insert cannot be packaged due to the size limitation (smaller than the required).

• Insert size ranges between 10-23 kb.

• Example: λ EMBL 3, λ EMBL 4, λ DASH etc.

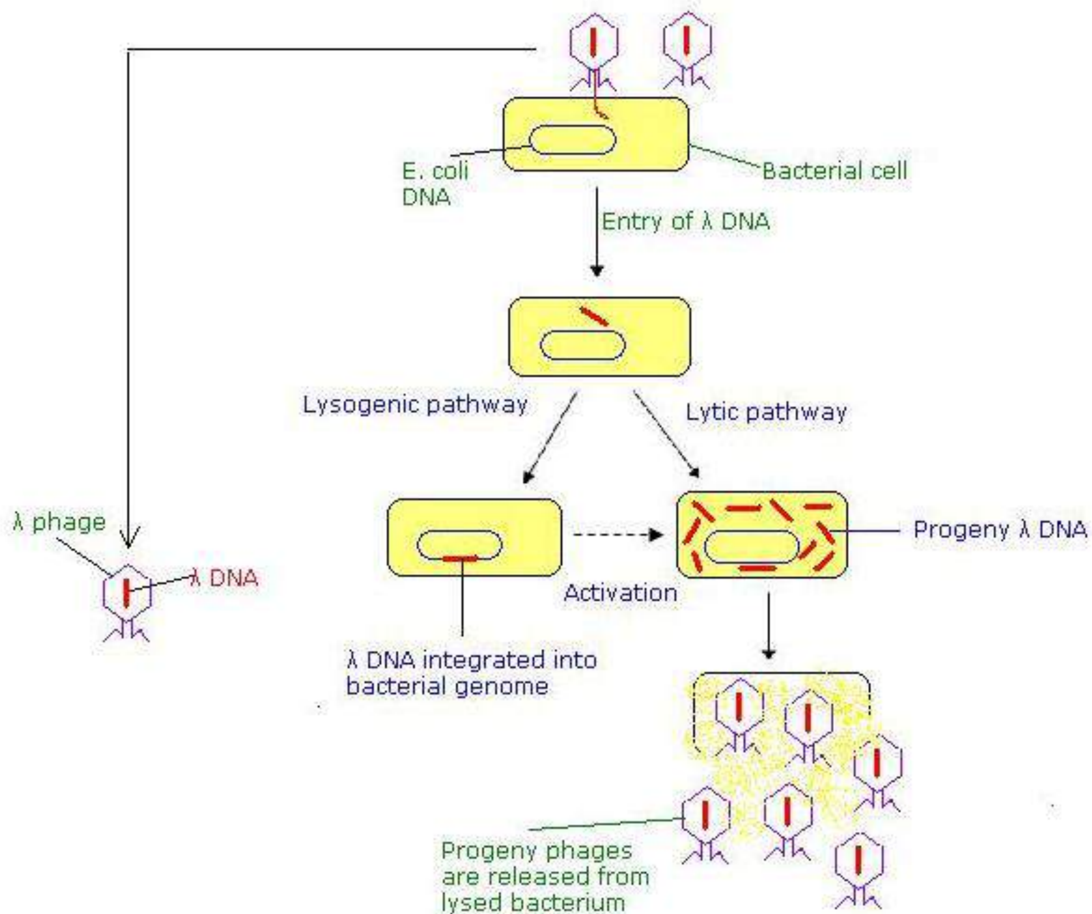


Phase vectors are a more efficient introduction of DNA into bacteria. Phage vectors such as those derived from bacteriophage l can carry **larger inserts** and can be **introduced into bacteria more efficiently**. l phage has a duplex DNA genome of about 50 kb. The internal 20 kb can be replaced with foreign DNA and still retain the lytic functions. Hence restriction fragments up to 20 kb can replace the l sequences, allowing larger genomic DNA fragments to be cloned.

Recombinant bacteriophage can be introduction into *E. coli* by **infection**. DNA that has the cohesive ends of l can be packaged in vitro into infective phage particles. Being in a viral particle brings the efficiency of infection reliably over 108 plaque forming units per mg of recombinant DNA.

Some other bacteriphage vectors for cloning are derived from the virus M13. One can obtain **single stranded DNA** from M13 vectors and recombinants. M13 is a virus with a genome of single stranded DNA. It has a nonessential region into which foreign genes can be inserted. It has been modified to carry a gene for b-galactosidase as a way to screen for recombinants. Introduction of recombinant M13 DNA into E. coli will lead to an infection of

the host, and the progeny viral particles will contain single-stranded DNA. The replicative form is duplex, allowing one to cleave with restriction enzymes and insert foreign DNA.



Lambda vectors for cloning

M13 Phage Vectors:

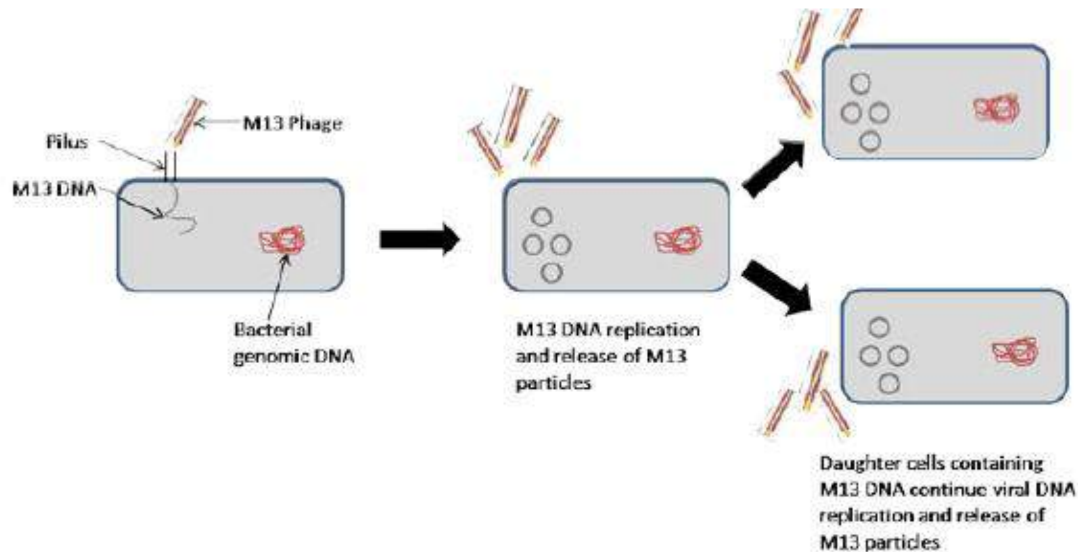
M13 phage is filamentous phage that infects *E. coli* via F-pilus. The genome is a single stranded circular DNA of size ~6.4kb surrounded by a proteinaceous coat. The DNA strand present in phage is called plus (+) strand. After entering to *E. coli* host, it converts into double stranded DNA molecule called **replicative form (RF)** by utilizing bacterial machinery. M13 phage as cloning vector can be obtained in both single stranded as well as double stranded form. Replicative form double stranded vector are modified and replicated inside *E. coli* host similar to a plasmid vector. Single stranded vectors can be isolated by collecting M13 phage. M13 vectors have useful application in following areas:

• DNA sequencing

• Mutagenesis study

• probe generation

• Phage display



Phagemid

Although M13 vectors are very useful for the production of single-stranded recombinant genes, they have certain disadvantages. There is a limit to the size of DNA fragment that can be cloned in an M13vector, with 1.5 kb being the ideal capacity, although fragments up to 3 kb have occasionally been cloned. To overcome this limitation, phagemid vectors were developed by combining a part of the M13 genome with plasmid DNA.

• Phagemids are cloning vectors developed as a hybrid of the filamentous phage M13 and plasmids to produce a vector that can get packed as a phage particle but also can propagate as a plasmid.

• They contain an origin of replication (ori) for double stranded replication inside *E. coli* host, as well as an "f1 ori" to enable single stranded replication and packaging into phage particles. Many commonly used plasmids contain an f1 ori and are thus phagemids.

• Phagemid generally encode no or only one of the capsid proteins of virus. Other structural and functional proteins necessary for phage lifecycle are provided by the helper phage.

• The components present in a phagemid vector are:

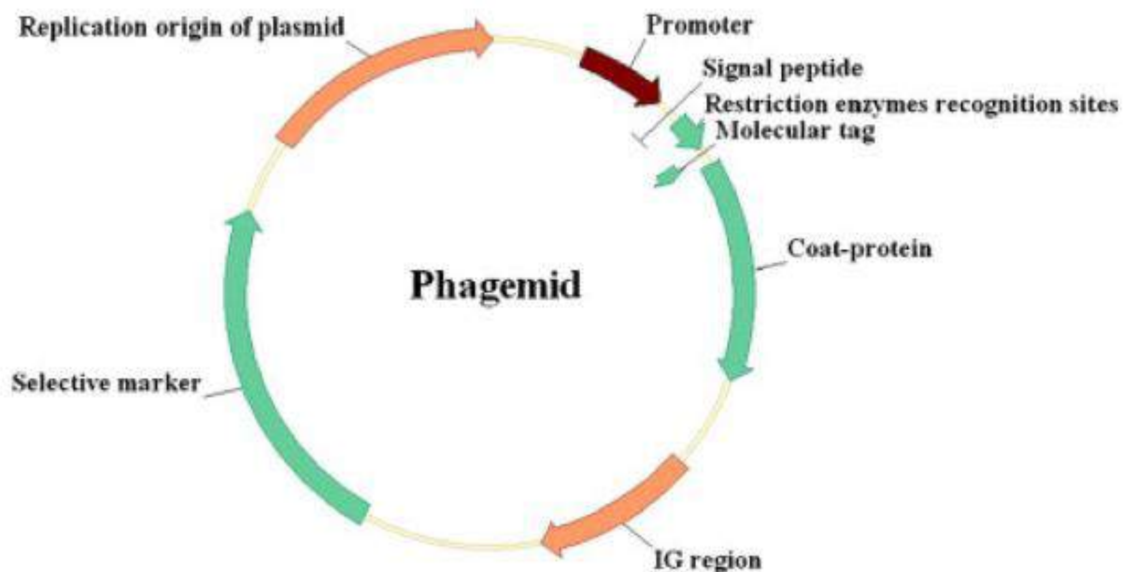
- Origin of replication (ori) of a plasmid.
- Intergenic region (IG region) which contains the packaging signal for the phage particle and also has replication origin inside phage.
- A gene encoding phage coat protein.
- A selection marker.
- Restriction enzyme recognition sites.

Phagemid vectors are commonly used for "phage display technology" by which a broad range of proteins and peptides can be expressed as fusions to phage coat proteins and displayed on the viral surface. The advantage of phagemid vectors is that double stranded

phagemid vectors can be converted into single stranded vectors and packaged into virion particles by infecting the cells with helper phage.

Phagemid has certain advantages over phage vectors:

- The carrying capacity of phagemid is higher than phage vectors.
- Phagemid has higher efficiency in transformation than phage vectors.
- Phagemids are genetically more stable than recombinant phage vectors.
- Phagemids can be exploited to generate single stranded DNA template for sequencing purposes.
- Single stranded phagemid vectors inside the phage can be targeted for site-directed mutagenesis.
- Single stranded vectors can be used to generate hybridization probes for mRNA or cDNA.



Cosmid:

• A cosmid, first described by **Collins** and **Hohn** in 1978, is a type of hybrid plasmid with a bacterial "ori" sequence and a "*cos*" sequences derived from the lambda phage.

• *Cos* site is the sequence required by a DNA molecule in order to be recognized as a 'λ genome' by the proteins that package DNA into λ phage particles.

• Cosmid DNA containing particles are as transmittable as real λ phages, but once inside the cell, the cosmid cannot control synthesis of new phage particles and instead replicates as a plasmid.

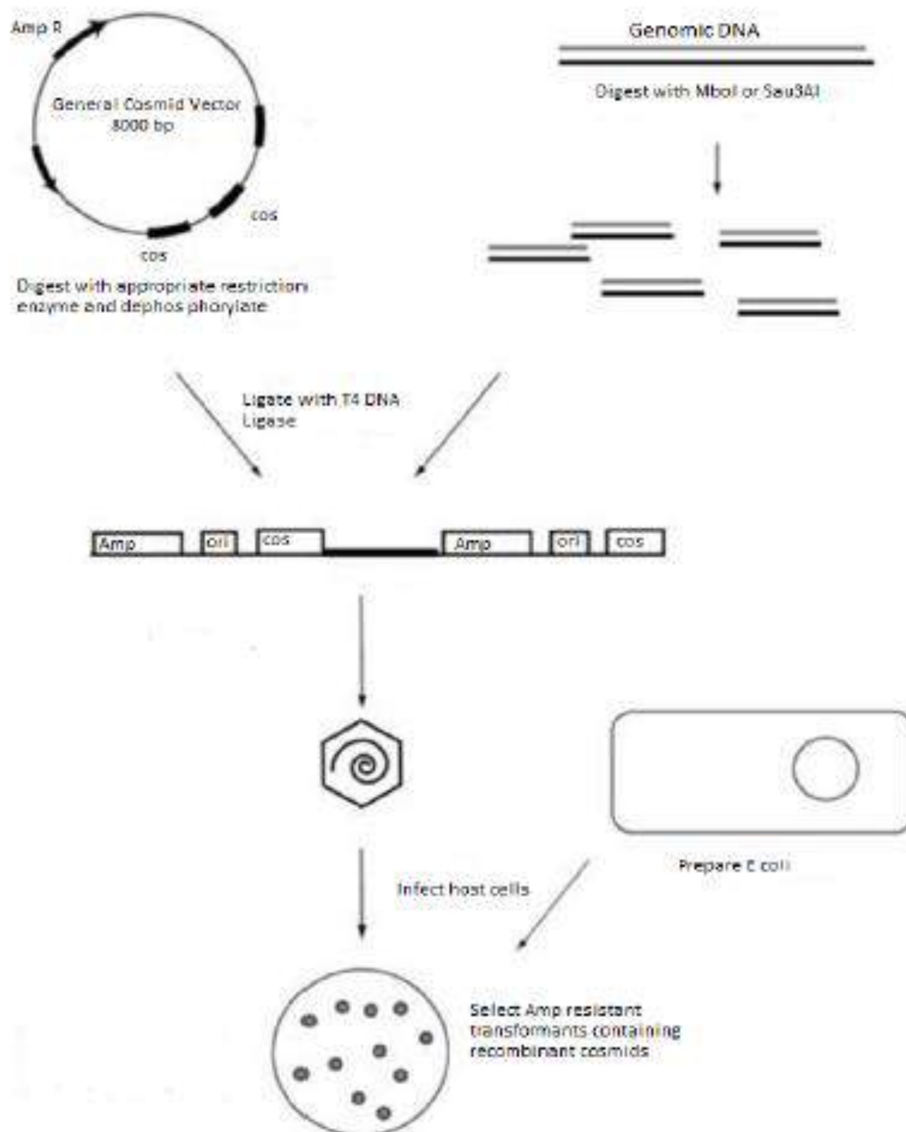
• Recombinant DNA is therefore obtained from colonies rather than plaques. They frequently also contain a gene for selection such as antibiotic resistance.

• They are able to load 37 to 52 kb of DNA, while normal plasmids are able to carry only 1–26 kb.

• Sometimes helper phage is used to assist in packaging of cosmid inside phage. Helper phage provides the essential proteins required for packaging which are lacked by cosmid vector.

• For packaging into a phage, concatemer formation is required (cosmid-insert-cosmid). This is generated by using two cos sites flanked by the insertion site for foreign DNA. Providing the inserted DNA in the right size, *in vitro* packaging cleaves the *cos* sites and replaces the recombinant cosmids in mature phage particles.

- Recombinant λ phages are used to infect an *E. coli* culture. Infected cells are plated on a selective medium and antibiotic-resistant colonies are grown. All colonies are recombinants, as non-recombinant linear cosmids are too small to be packaged into λ heads.
- Cosmids are widely exploited to build genomic libraries. The upper limit for the length of the cloned DNA is set by the space available within the λ phage particle. New DNA insert of size up to 44 kb can be inserted before the packaging limit of the λ phage particle is reached.



Schematics for Cosmid Library Construction

Artificial Chromosomes:

Artificial chromosomes are DNA molecules assembled *in vitro* from defined constituents that can function like natural chromosomes.

Types of artificial chromosomes:

i) BACs: Bacterial artificial chromosomes

ii) YACs: Yeast artificial chromosomes

iii) MACs: Mammalian artificial chromosomes

iv) HACs: Human artificial chromosomes

Bacterial Artificial Chromosomes: BAC:

Bacterial artificial chromosomes (BACs) are designed for the cloning of large DNA insert (typically 100 to 300 kb) in *E. coli* host. BAC vectors contain a single copy F-plasmid origin of replication (ori).

The F (fertility) plasmid is relatively large and vectors derived from it have a higher capacity than normal plasmid vectors. F-plasmid has F (fertility) factor which controls the replication and maintain low copy number. Also conjugation can take place between F+ bacteria (male) and F- bacteria (female) to transfer F-plasmid via pilus.

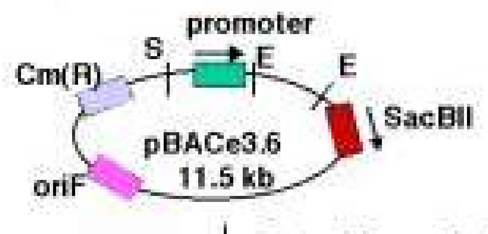
Common gene components of a bacterial artificial chromosome are:

1) **oriS, repE – F** for plasmid replication and regulation of copy number.

2) **parA** and **parB** for maintaining low copy number and avoiding two F plasmids in a single cell during cell division.

3) A selectable marker for antibiotic resistance; some BACs also have *lacZ* at the cloning site for blue/white selection.

4) T7 and Sp6 phage promoters for transcription of inserted genes.



The *par* genes, derived from F plasmid assist in the even distribution of plasmids to daughter cells during cell division and increase the likelihood of each daughter cell carrying one copy of the plasmid, even when few copies are present. The low number of copies is

useful in cloning large fragments of DNA because it limits the opportunities for unwanted recombination reactions that can unpredictably alter large cloned DNA over time.

The first BAC vector, **pBAC108L**, did not contain a selectable marker for recombinants. Thus, positive recombinants had to be identified by colony hybridization. Two widely used BAC vectors, **pBeloBAC11** and **pECBAC1**, are derivatives of pBAC108L in which the original cloning site is replaced with a *lac*Z gene carrying a multiple cloning site. **PBeloBAC11** has two *Eco*RI sites, one in the *lac*Z gene and one in the CMR gene, whereas pECBAC1 has only one *Eco*RI site in the *lac*Z gene. Further improvements to BACs have been made by replacing the *lac*Z gene with the *sac*B gene which is a negative selection marker. The product of *sacB* gene is levansucrase which can convert sucrose present in the media into levan, a toxin for the bacteria. Hence the colonies without insert would have intact *sacB* gene and thus cells die before forming colonies.

The F plasmid is relatively large and vectors constructed on it have a higher capacity for accepting inserted DNA. A similar cloning vector called a P1-derived artificial chromosome or PAC has also been produced from the bacterial P1 bacteriophage DNA. Both BACs and PACs can be used to clone fragments of 300kb and longer. They are often used to sequence the genome of organisms in genome projects.

Applications of BAC

● Contribution to models of disease: Inherited diseaseContribution to models of disease: Inherited disease

● BACs are now being utilized in modeling genetic diseases, often alongside transgenic mice.

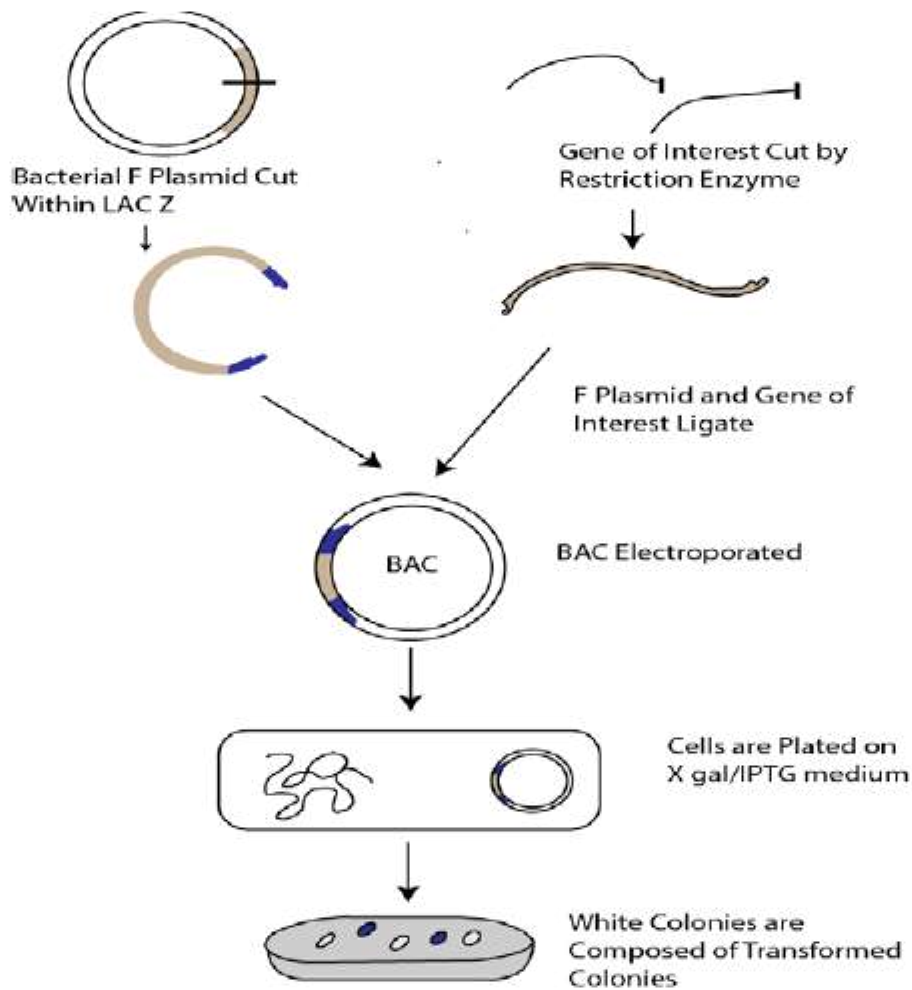
● BACs have been useful in this field as complex genes may have several regulatory sequences upstream of the encoding sequence, including various promoter sequences that will govern a gene's expression level.

● BACs have been used to study neurological diseases such as Alzheimer's disease or as in the case of aneuploidy associated with Down syndrome. There have also been instances when they have been used to study specific oncogenes associated with cancers.

● Contribution to models of disease: Infectious diseaseContribution to models of disease: Infectious disease

● The genomes of several large DNA viruses and RNA viruses have been cloned as BACs.

● These constructs are referred to as "infectious clones".



Transforming a Bacterium Using a BAC Vector

Yeast Artificial Chromosomes: YAC

• First described in 1983 by **Murray** and **Szostak**, a yeast artificial chromosome has sequences to exist inside *E. coli* as a circular plasmid and contains sequences to maintain as linear nuclear chromosome in yeast.

• As YAC vectors can accommodate 100-500 kb of insert DNA. The number of clones in a genomic library can be greatly reduced.

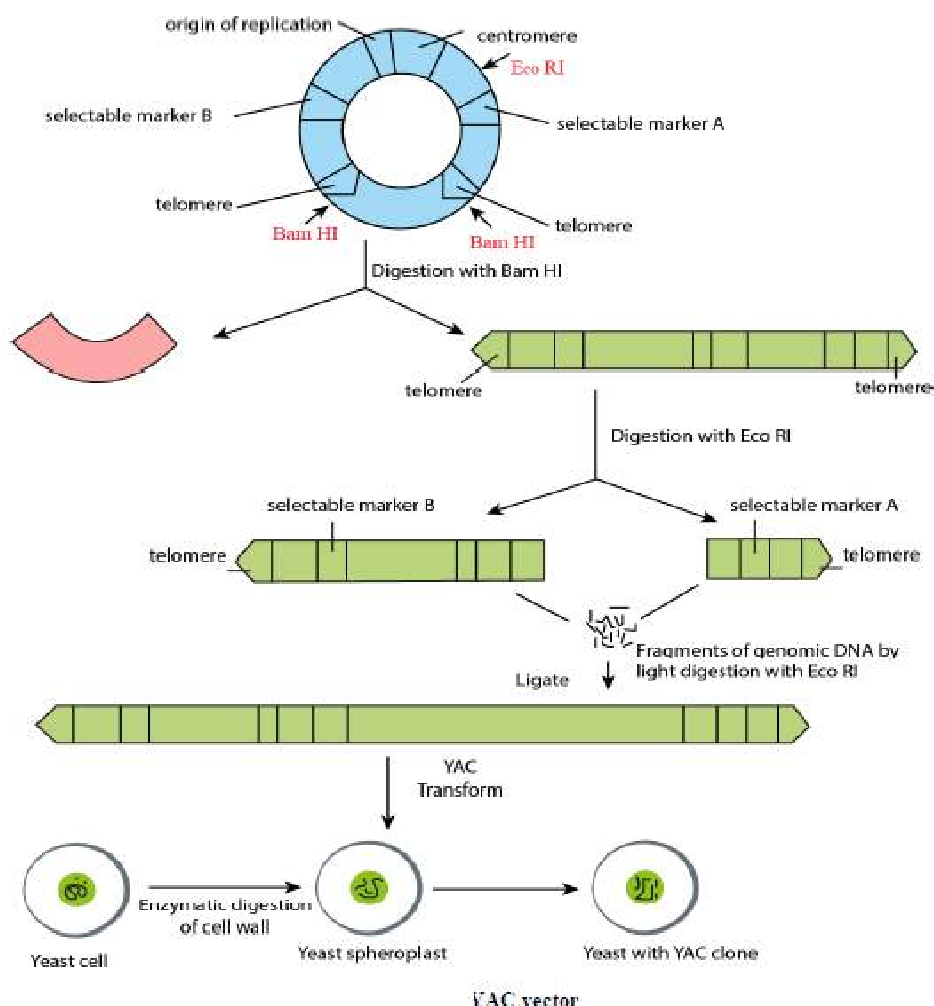
• YAC vectors have following elements:

- *E. coli* origin of replication
- Yeast origin of replication
- Elements of eukaryotic yeast chromosome (centromere and telomere region)
- Selection markers for both the host.

• YAC vector is initially propagated as circular plasmid inside bacterial host utilizing bacterial *ori* sequence. Circular plasmid is cut at specific site using restriction enzymes to generate a linear chromosome with two telomere sites at terminals. The linear chromosome is again digested at specific site with two arms with different selection marker. Genomic insert is then ligated into YAC vector using DNA ligase enzyme. The recombinant vectors are transformed into yeast cells and screened for the selection markers to obtain recombinant colonies.

• Yeast expression vectors, such as YACs, YIPs (yeast integrating plasmids), and YEPs (yeast episomal plasmids), have advantageous over bacterial artificial chromosomes (BACs). They can be used to express eukaryotic proteins that require post-translational modification. However, YACs have been found to be less stable than BACs.

Some recombinant plasmids have the ability to incorporate multiple replication origins and other elements that allow them to be used in more than one species (for example, yeast or *E. coli*).



YAC vector

Applications of YAC

● Yeast expression vectors, such as YACs, YIps (yeast integrating plasmids), and YEps (yeast episomal plasmids), have an advantage over bacterial artificial chromosomes (BACs) in that they can be used to express eukaryotic proteins that require post translational modification.

● By being able to insert large fragments of DNA, YACs can be utilized to clone and assemble the entire genomes of an organism. With the insertion of a YAC into yeast cells, they can be propagated as linear artificial chromosomes, cloning the inserted regions of DNA in the process. With this completed, two process can be used to obtain a sequenced genome, or region of interest: ➢ Physical Mapping ➢ Chromosome Walking

UNIT-IV

GENE TRANSFER TECHNIQUES

INTRODUCTION

• Gene transfer is to transfer a gene from one DNA molecule to another DNA molecule.

• The directed desirable gene transfer from one organism to another and the subsequent stable integration & expression of foreign gene into the genome is referred as genetic transformation.

• Transient transformation occur when DNA is not integreted into host genome

• Stable transformation occur when DNA is integrated into host genome and is inherited in subsequent generations.

• The transferred gene is known as transgene and the organism that develop after a successful gene transfer is known as transgenic.

METHODS OF GENE TRANSFER DNA

Transfer by natural methods

1. Conjugation
2. Bacterial transformation
3. Retroviral transduction
4. Agrobacterium mediated transfer

DNA TRANSFER BY ARTIFICIAL METHODS

Physical methods

1. Microinjection

2. Biolistics transformation

Chemical methods

1. DNA transfer by calcium phosphate method

2. Liposome mediated transfer

Electrical methods

1. Electroporation

Biolistics or microprojectiles

DNA transfer Biolistics or particle bombardment is a physical method that uses accelerated microprojectiles to deliver DNA or other molecules into intact tissues and cells. The method was developed initially to transfer genes into plants by Sanford. Biolistics transformation is relatively new and novel method amongst the physical methods for artificial transfer of exogenous DNA. This method avoids the need of protoplast and is better in efficiency. This technique can be used for any plant cells, root section, embryos, seeds and pollen.

The gene gun is a device that literally fires DNA into target cells. The DNA to be transformed into the cells is coated onto microscopic beads made of either gold or tungsten. Beads are carefully coated with DNA. The coated beads are then attached to the end of the plastic bullet and loaded into the firing chamber of the gene gun. An explosive force fires the bullet down the barrel of the gun towards the target cells that lie just beyond the end of the barrel. When the bullet reaches the end of the barrel it is caught and stopped, but the DNA coated beads continue on toward the target cells. Some of the beads pass through the cell wall into the cytoplasm of the target cells. Here the bead and the DNA dissociate and the cells become transformed. Once inside the target cells, the DNA is solubilised and may be expressed.

A gene gun or a biolistic particle delivery system is a device which can directly bombard small particles coated with the recombinant DNA on the nucleus of the target cell. This technique is often simply referred to as bio-ballistics or biolistics and has been successfully used in the transfection of both plant and animal cells.

In this technique the recombinant DNA is coated with microscopic tungsten particles known as micro-projectiles, which are then accelerated on a macro-projectile by firing a gunpowder charge or by using compressed gas to drive the macro-projectile.

At one end of the ‘gun’ there is a small aperture that stops the macro-projectile but allows the micro-projectiles to pass through. When directed at cells, these micro-projectiles carry the DNA into the cell and, in some cases, stable transformation will occur.

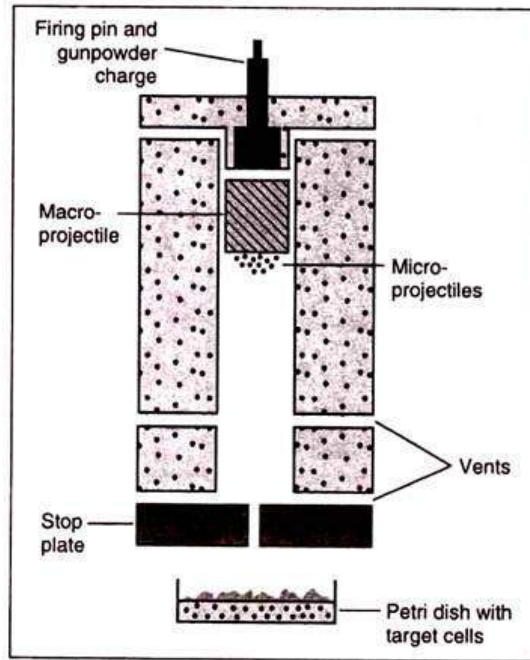
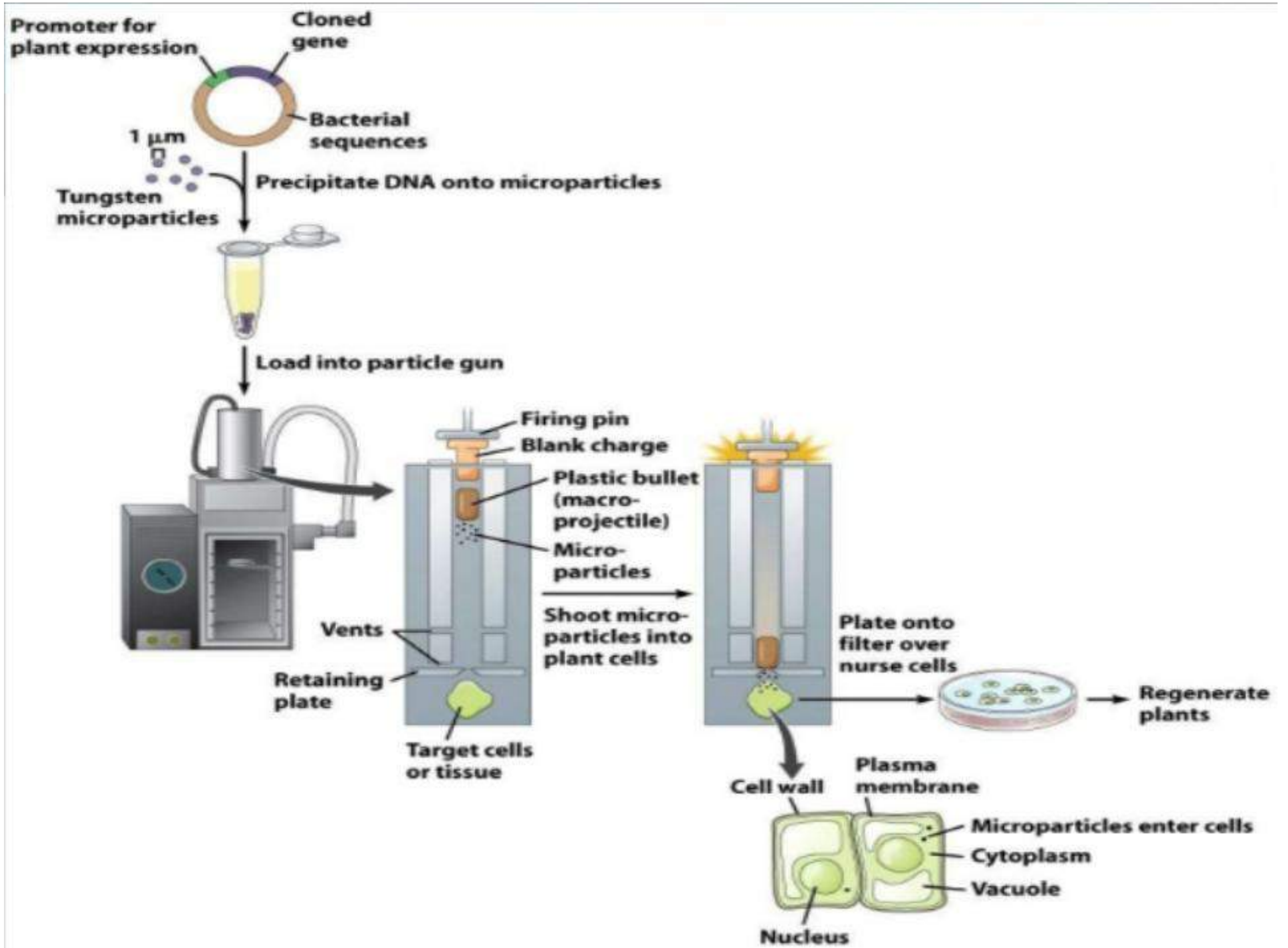


Fig. 5.9: Biolistic particle delivery system



• Biolistics or particle bombardment is a physical method that uses accelerated microprojectiles to deliver DNA or other molecules into intact tissues and cells.

• The gene gun is a device that literally fires DNA into target cells .

• The DNA to be transformed into the cells is coated onto microscopic beads made of either gold or tungsten.

• The coated beads are then attached to the end of the plastic bullet and loaded into the firing chamber of the gene gun.

• An explosive force fires the bullet with DNA coated beads towards the target cells that lie just beyond the end of the barrel.

• Some of the beads pass through the cell wall into the cytoplasm of the target cells

Advantages and limitations of biolistics

1. Requirement of protoplast can be avoided.
2. Walled intact cells can be penetrated.
3. Manipulation of genome of subcellular organelles can be achieved.

Limitations

1. Integration is random.
2. Requirement of equipments.

GENE TRANSFER TECHNIQUES: CHEMICAL METHODS

Introduction

Cell membrane is a sheet like assembly of amphipathic molecules that separate cells from their environment. These physical structures allow only the controlled exchange of materials among the different parts of a cell and with its immediate surroundings. DNA is an anionic polymer, larger molecular weight, hydrophilic and sensitive to nuclease degradation in biological matrices. They cannot easily cross the physical barrier of membrane and enter the cells unless assisted.

Various charged chemical compounds can be used to facilitate DNA transfer directly to the cell. These synthetic compounds are introduced near the vicinity of recipient cells thereby disturbing the cell membranes, widening the pore size and allowing the passage of the DNA into the cell.

An ideal chemical used for DNA transfer should have the ability to-

• Protect DNA against nuclease degradation.

• Transport DNA to the target cells.

• Facilitate transport of DNA across the plasma membrane.

• Promote the import of DNA into the nucleus. The commonly used methods of chemical transfection use the following,

1. Calcium Chloride
2. DEAE dextran

Calcium Chloride ($CaCl_2$) Mediated DNA Transfer:

This is used for the transformation of prokaryotic host cells.

Principle:

In the process of transformation all bacterial cells cannot uptake the exogenous DNA molecule. Those who are capable to take are called competent cells. So our aim in this step is to make bacterial cells more competent so that the possibility of transferring of the recombinant DNA into the host cell increases to a higher fold. $CaCl_2$ makes the cell wall of the bacteria more permeable to the exogenous DNA and thus increases the competence of the host cell.

Procedure:

Growing E. Coli cells are isolated and suspended in 50 mM $CaCl_2$ at a concentration of 108-1010 cells/ml. The cells may be incubated for 12- 24 hr. to increase the frequency of transformation. The recombinant DNA is then added.

Efficient transformation takes only a few minutes and the cells are plated on a suitable medium for the selection of transformed clones. The frequency of transformed cells is 106-107 per mg of plasmid DNA; this is about one transformation per 10,000 plasmid molecules.

The transformed cells are suitably diluted and spread thinly on a suitable medium so that each cell is well separated and produces a separate colony. Generally, the medium is so designed that it permits only the transformed cells to divide and produce colonies. This frequency can be further improved by using special E. Coli strains, e.g., SK1590, SK1592, X1766, etc.

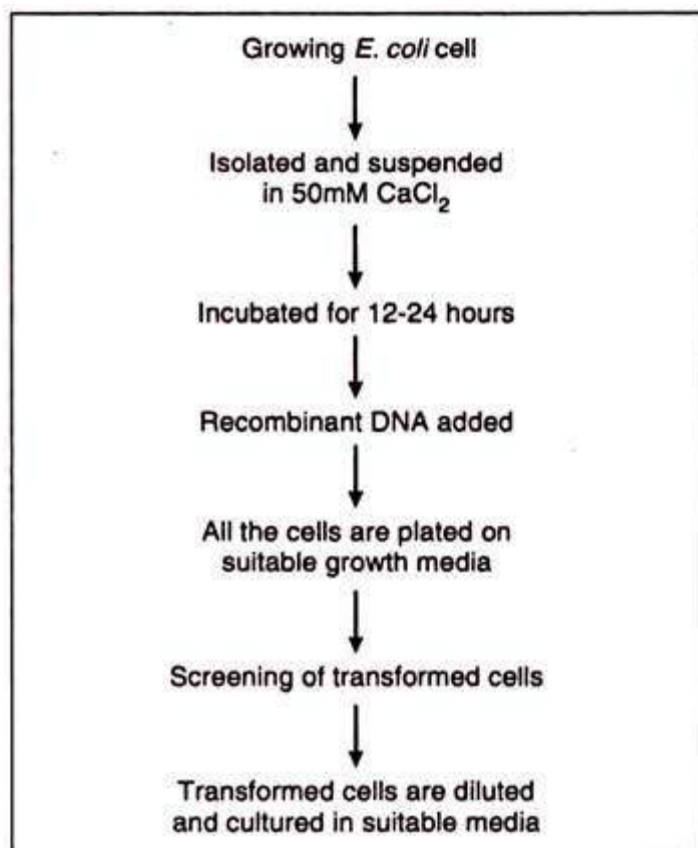


Fig. 5.5: Calcium chloride ($CaCl_2$) mediated DNA transfer

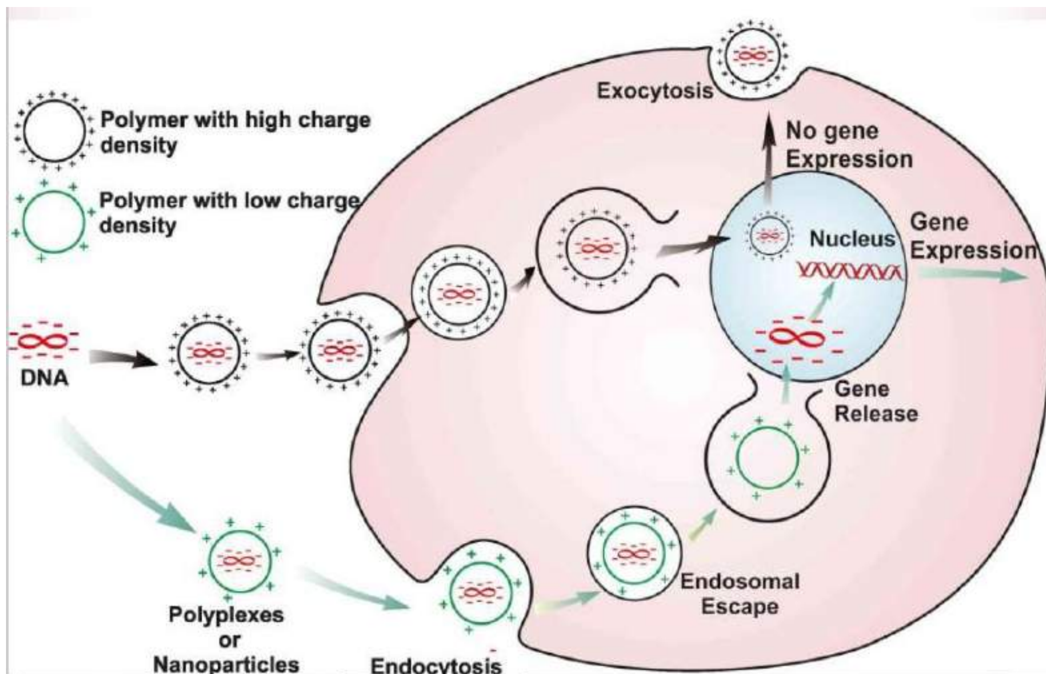
DEAE-Dextran (Diethylaminoethyl Dextran) mediated DNA transfer

• This method was initially reported by Vaheri and Pagano in 1965 for enhancing the viral infectivity of cell but later adapted as a method for plasmid DNA transfer.

• Diethylaminoethyl dextran (DEAE-dextran) is a soluble polycationic carbohydrate that promotes interactions between DNA and endocytotic machinery of the cell.

• In this method, the negatively charged DNA and positively charged DEAE – dextran form aggregates through electrostatic interaction and form apolyplex. A slight excess of DEAE – dextran in mixture results in net positive charge in the DEAE – dextran/ DNA complex formed. These complexes, when added to the cells, bind to the negatively charged plasma membrane and get internalized through endocytosis. Complexed DNA delivery with DEAE-dextran can be improved by osmotic shock using DMSO or glycerol.

• Several parameters such as number of cells, polymer concentration, transfected DNA concentration and duration of transfection should be optimized for a given cell line.



Advantages

• Simple and inexpensive

• More sensitive

• Can be applied to a wide range of cell types

• Can be used for transient transfection.

Disadvantages

• Toxic to cells at high concentrations

• Transfection efficiency varies with cell type

• Can only be used for transient transfection but not for stable transfection

• Typically produces less than 10% delivery in primary cells. Another polycationic chemical, the detergent Polybrene, has been used for the transfection of Chinese hamster ovary (CHO) cells, which are not amenable to calcium phosphate transfection.

GENE TRANSFER TECHNIQUES: BIOLOGICAL METHODS

Introduction

The main gene transfer methods using biological means are as follows:

• Bacterial gene delivery i.e. bactofection.

• Delivery using a viral vector i.e. transduction

Bactofection

It is a method of direct gene transfer using bacteria into the target cell, tissue, organ or organism. Various bacterial strains that can be used as vectors in gene therapy. The genes located on the plasmids of the transformed bacterial strains are delivered and expressed into the cells. The gene delivery may be intracellular or extracellular. It has a potential to express various plasmid-encoded heterologous proteins (antigens, toxins, hormones, enzymes etc.) in different cell types. Strains that are invasive and having better cell to cell spread are more efficient.

Uses

• Bactofection can be used for DNA vaccination against various microbial agents such as viruses, fungi, protozoans and other bacteria.

• It can be used in the treatment of several tumours like melanoma, lung carcinoma and colon carcinoma in mice.

Advantages

• Simple, selective and efficient transfection.

• Low synthesis cost and can be administered easily.

Disadvantages

• Unwanted side effects associated with host-bacteria interaction. This can be reduced by using genetically modified bacteria which contain suicide genes that ease the bacterial destruction and thus reduces the risk of clinical infections.

Transduction

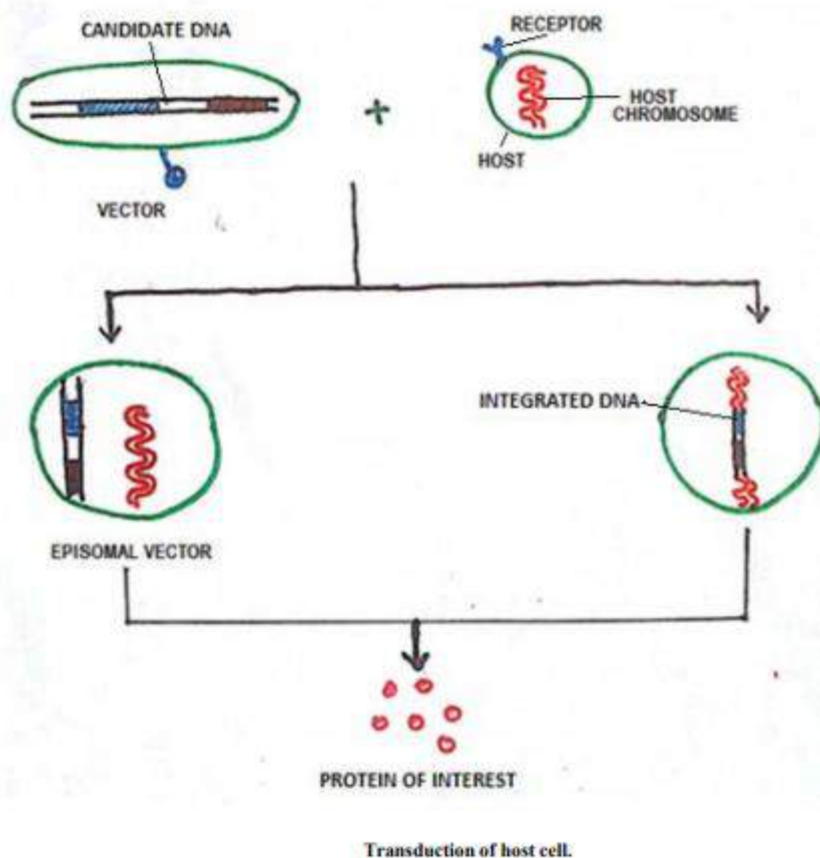
This method involves the introduction of genes into host cell’s genome using viruses as carriers. The viruses are used in gene transfer due to following features-

• Efficiency of viruses to deliver their nucleic acid into cells

• High level of replication and gene expression.

The foreign gene is packaged into the virus particles to enter the host cell. The entry of virus particle containing the candidate gene sequences into the cell and then to the nuclear genome is a receptor- mediated process. The vector genome undergoes complex processes ending up

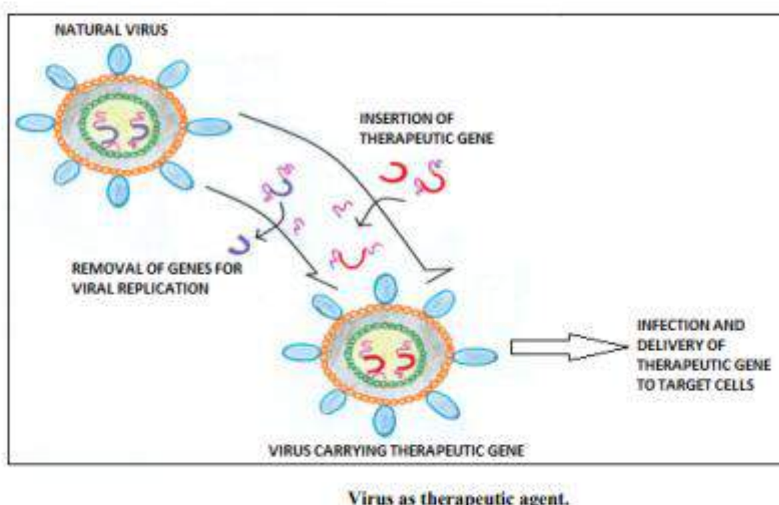
with ds-DNA depending on the vector that can persist as an episome or integrate into the host genome followed by the expression of the candidate gene.



Transduction of host cell.

Viral vectors as therapeutic agents

Viruses have paved a way into clinical field in order to treat cancer, inherited and infectious diseases. They can be used as vectors to deliver a therapeutic gene into the infected cells. They can be genetically engineered to carry therapeutic gene without having the ability to replicate or cause disease.



Virus as therapeutic agent.

Selection of host cells/ organism containing vector sequences:

Selection of the transformed cells from the non-transformed population is done by using selectable marker genes that confers resistance to antibiotics. Hence, cells only having the vector with the resistance gene for the antibiotic would grow in the selection media containing the antibiotic (ampicillin, tetracycline etc.); while the non-transformed cells would die.

SELECTION OF RECOMBINANT CLONES

SELECTION

– After the introduction of recombinant DNA into the host cells, it is essential to identify those cells which received rDNA molecule - screening (or) selection.

– The vector or foreign DNA present in the recombinant cells expresses certain characters or traits, while non-recombinants do not expess the traits.

– Following the methods for screening or selection of recombinant clones.

DIRECT SELECTION OF RECOMBINANTS

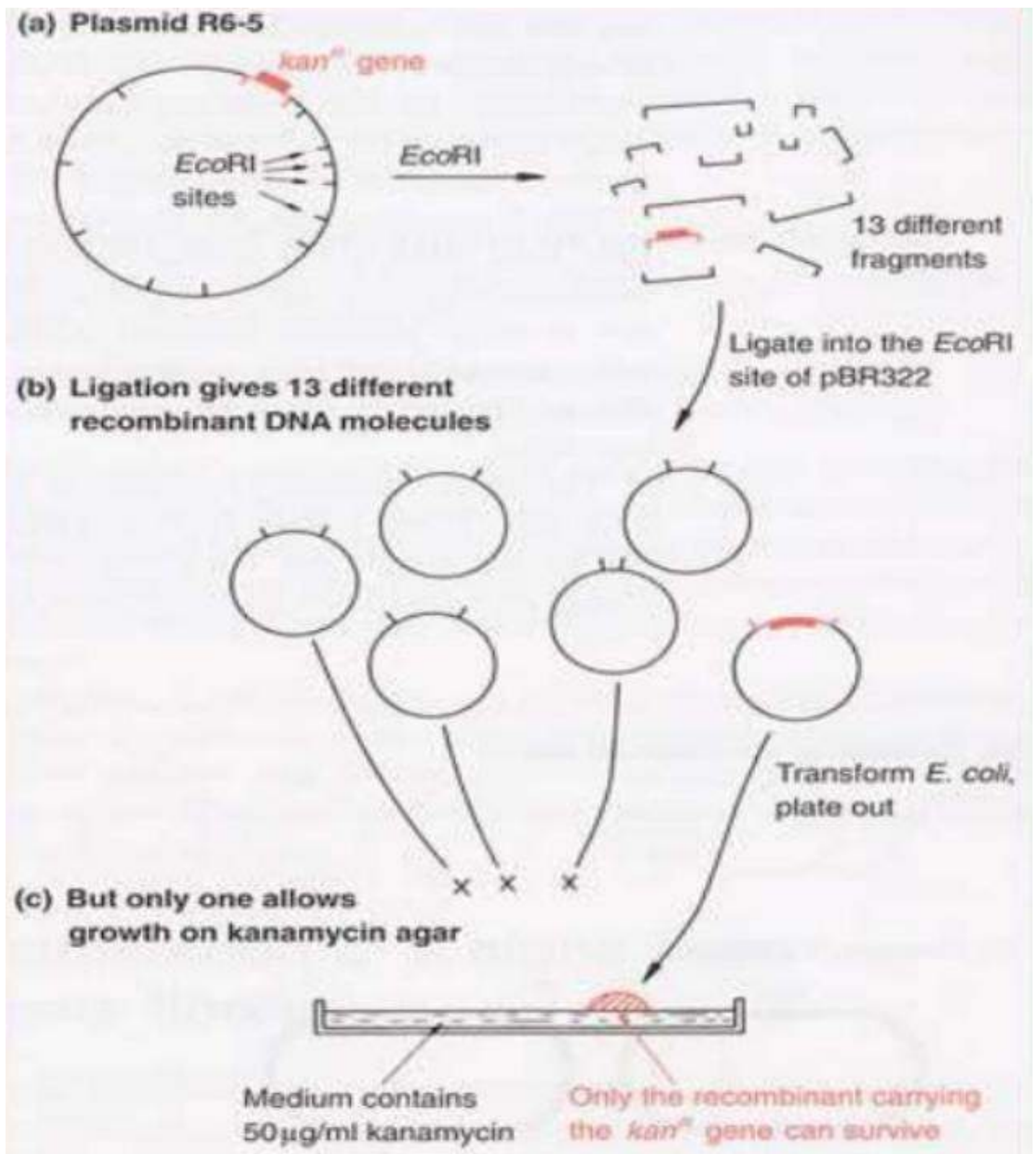
The simplest example of Direct Selection occurs when the desired gene specifies resistance to an antibiotic. As an example will will consider an experiment to clone the gene for kanamcycin resistance from plasmid R6-5.

– This plasmid carries genes for resistances to four antibiotics: kanamycin, chloramphenicol, streptomycin and sulphonamide. The kanamycin resistance gene lies within one of the 13 EcoRI fragments (a).

– To clone this gene the EcoRI fragments of R6-5 would be inserted into the EcoRI site of a vector such as pBR322. The ligated mix will comprise many copies of 13 different recombinant DNA molecules, one set of which carries the gene for kanamycin resistance.

– Insertional inactivation cannot be used to select recombinants when the EcoRI site of pBR322 is used. This is becuase this site does not lie in either the ampicillin or the tetracycline resistance of this plasmid.

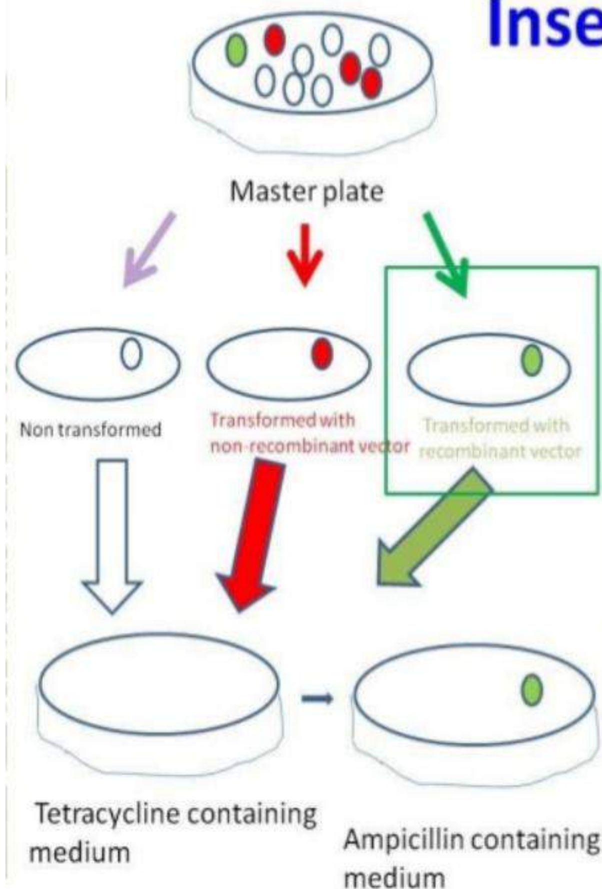
– But this is immaterial for cloning the kanamycin resistance gene because in this case the coloned gene can be used as the selectable marker. Transformants are plated onto kanamycin agar, on which the only cells able to survice and produce colonies are those recombinants that contain the cloned kanamycin resistance gene.



INSERTIONA L INACTIVATIO N SELECTION METHOD

Insertional inactivation is the inactivation of a gene upon insertion of another gene in its place or within its coding sequence.

– This helps in selection of recombinant colonies in rDNA technology.



BLUE-WHITE SCREENING

– It is a powerful method for screening recombinants.

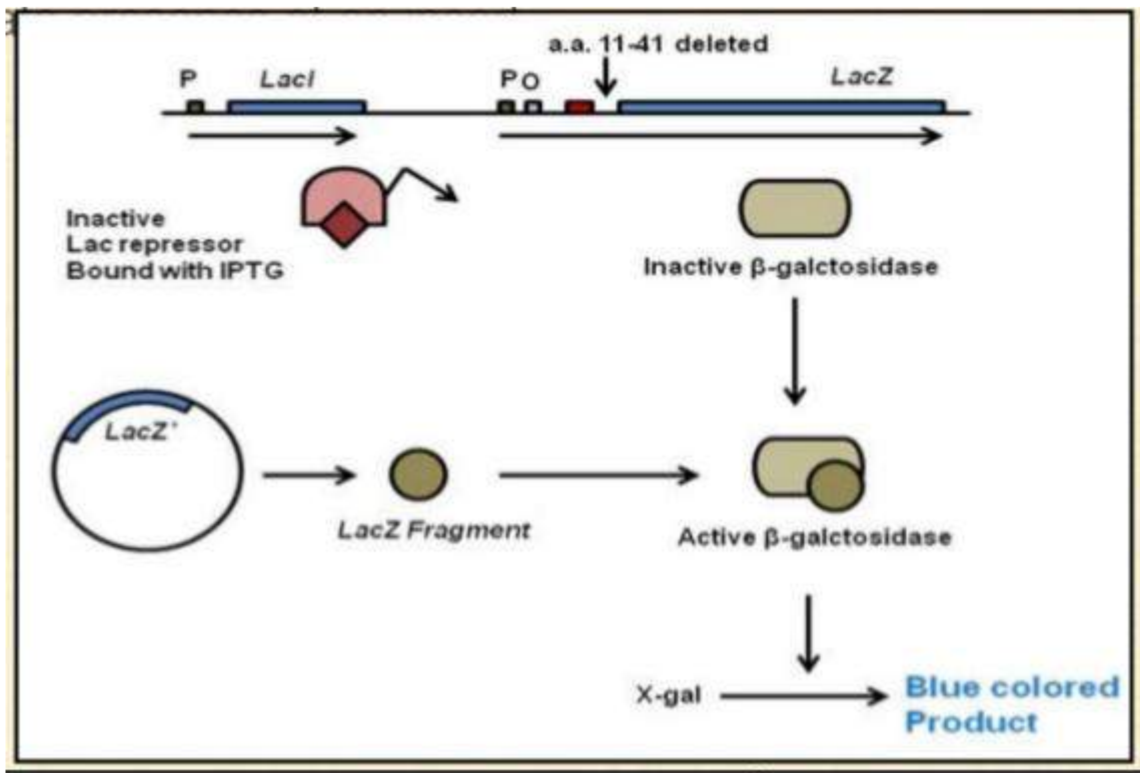
– In this method a reporter gene lacZ is inserted in the vector (encodes β-galactosidase).

– β-galactosidase breaks a synthetic substrate, X-gal into an insoluble blue colured product.

– If a foreign gene is inserted into lacZ, this gene will be inactivated; therefore no blue colour will develop.

– The host cells containing recombinant will form white coloured substrate on the medium containing X-gal.

– The host cells containing non recombinants wil turn blue in colour. – On the basis of colony colour the recombinants can be selected.



Screening by hybridization

• Nucleic acid hybridization is the most commonly used method of library screening first developed by Grunstein and Hogness in1975 to detect DNA sequences in transformed colonies using radioactive RNA probes.

• It relies on the fact that a single-stranded DNA molecule, used as a probe can hybridize to its complementary sequence and identify the specific sequences.

• This method is quick, can handle a very large number of clones and used in the identification of cDNA clones which are not full-length (and therefore cannot be expressed). The commonly used methods of hybridization are,

a) Colony hybridization

b) Plaque hybridization

COLONY HYBRIDIZATION TECNIQUE

– Colony blot hybridization is applied to DNA or RNA released from blotted microbial colonies.

– The microbial colonies are transferred (blotted) to a membrane.

– The cells are lysed in place to release the nucleic acids.

– The RNA or DNA (after denaturation) is fixed to the filter and hybridized with a labelled probe.

– Blocking reagent may be added prior to the probe to prevent unspecific binding.

– Excess probe is washed away and the membrane is visualized by UV or autoradiography.

– Colony blot hybridization can be used for screening clones or bacterial isolates.

Plaque hybridization

Plaque hybridization, also known as Plaque lift, was developed by Benton and Davis in 1977 and employs a filter lift method applied to phage plaques. This procedure is successfully applied to the isolation of recombinant phage by nucleic acid hybridization and probably is the most widely applied method of library screening. The method of screening library by plaque hybridization is described below-

• The nitrocellulose filter is applied to the upper surface of agar plates, making a direct contact between plaques and filter.

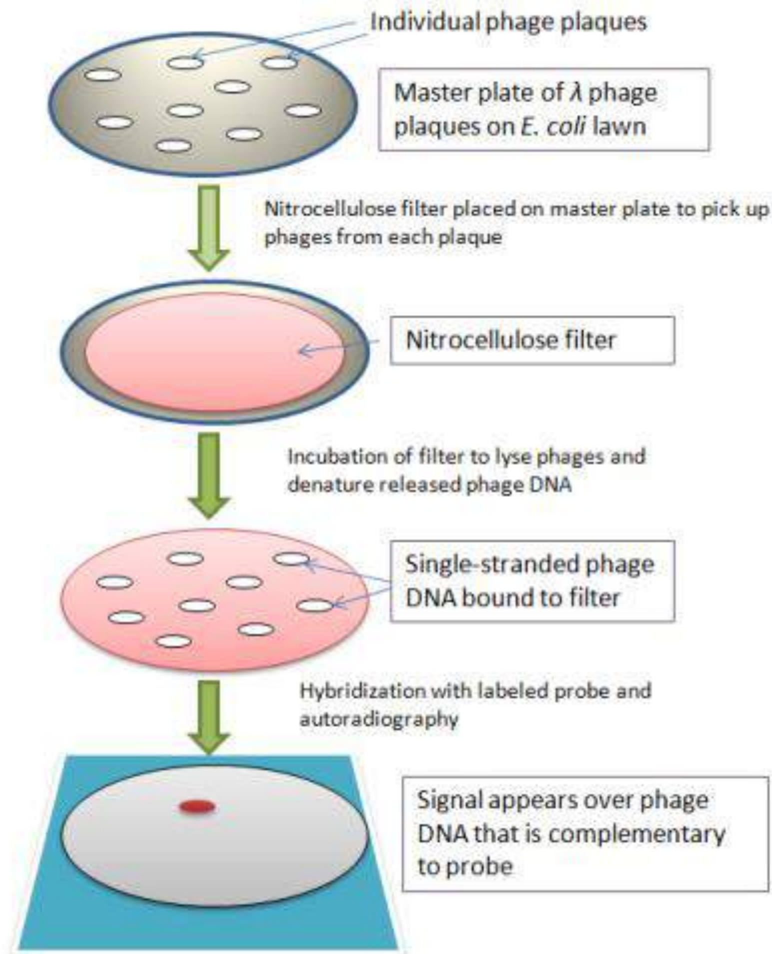
• The plaques contain phage particles, as well as a considerable amount of unpackaged recombinant DNA which bind to the filter.

• The DNA is denatured, fixed to the filter, hybridized with radioactive probes and assayed by autoradiography.

Advantages

• This method results in a 'cleaner' background and distinct signal (less background probe hybridization) for λ plaque screening due to less DNA transfer from the bacterial host to the nitrocellulose membrane while lifting plaques rather than bacterial colonies.

• Multiple screens can be performed from the same plate as plaques can be lifted several times.



Schematic process for screening libraries by Plaque hybridization.

Immunological screening

This involves the use of antibodies that specifically recognize antigenic determinants on the polypeptide. It does not rely upon any particular function of the expressed foreign protein, but requires an antibody specific to the protein. Earlier immunoscreening methods employed radio-labeled primary antibodies to detect antibody binding to the nitrocellulose sheet . It is now superseded by antibody sandwiches resulting in highly amplified signals. The secondary antibody recognizes the constant region of the primary antibody and is, additionally, conjugated to an easily assayable enzyme (e.g. horseradish peroxidase or alkaline phosphatase) which can be assayed using colorimetric change or emission of light using X-ray film.

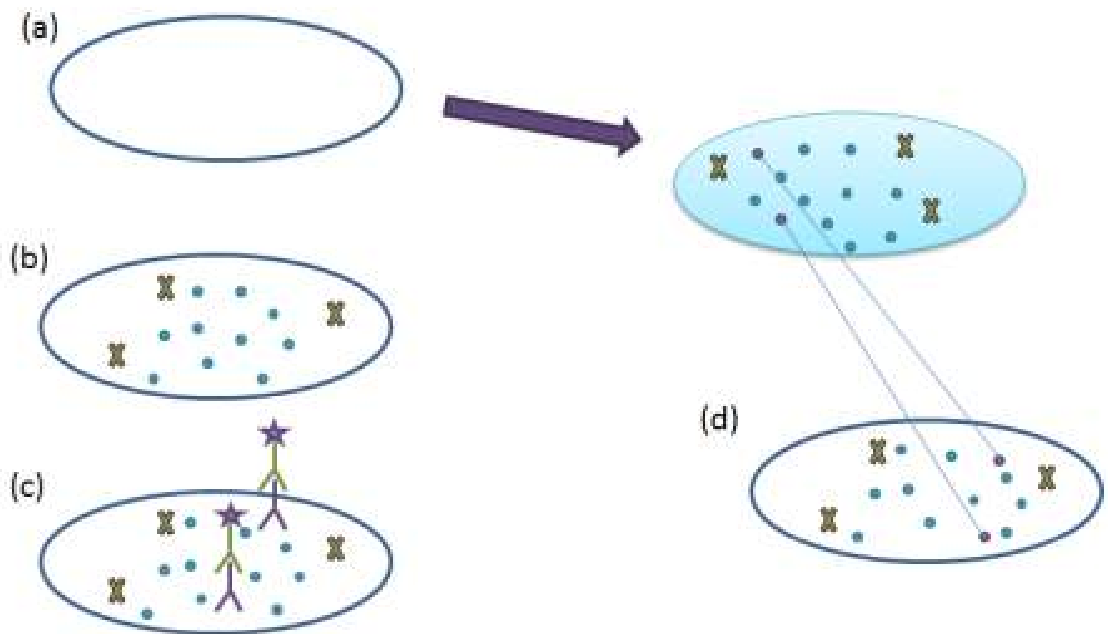
• In this technique, the cells are grown as colonies on master plates and transferred to a solid matrix.

• These colonies are subjected to lysis releasing the proteins which bind to the matrix.

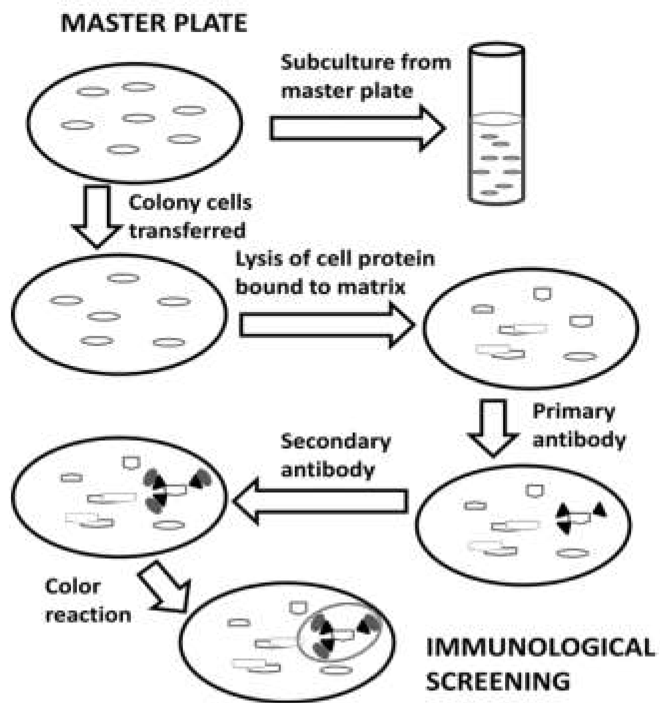
• These proteins are treated with a primary antibody which specifically binds to the protein (acts as antigen), encoded by the target DNA. The unbound antibodies are removed by washing.

• A secondary antibody is added which specifically binds to the primary antibody removing the unbound antibodies by washing.

• The secondary antibody carries an enzyme label (e.g., horse radishperoxidase or alkaline phosphatase) bound to it which converts colorless substrate to colored product. The colonies with positive results (i.e. colored spots) are identified and subcultured from the master plate.



(a). Schematic process of immunological screening (a) a nitrocellulose disk is placed onto the surface of an agar plate containing the phage library. Both agar plate and disk are marked so as to realign them later. (b) When the nitrocellulose disk is lifted off again, proteins released from the bacteria by phage lysis bind to the disk. (c) These proteins bind to specific antibody. (d) Plaques formed by bacteriophage that express the protein bound to the antibody will be detected by emission of light. The positive clones can be identified by realignment.



Schematic process of immunological screening using antibody sandwich.

Blotting Techniques:

Blotting:

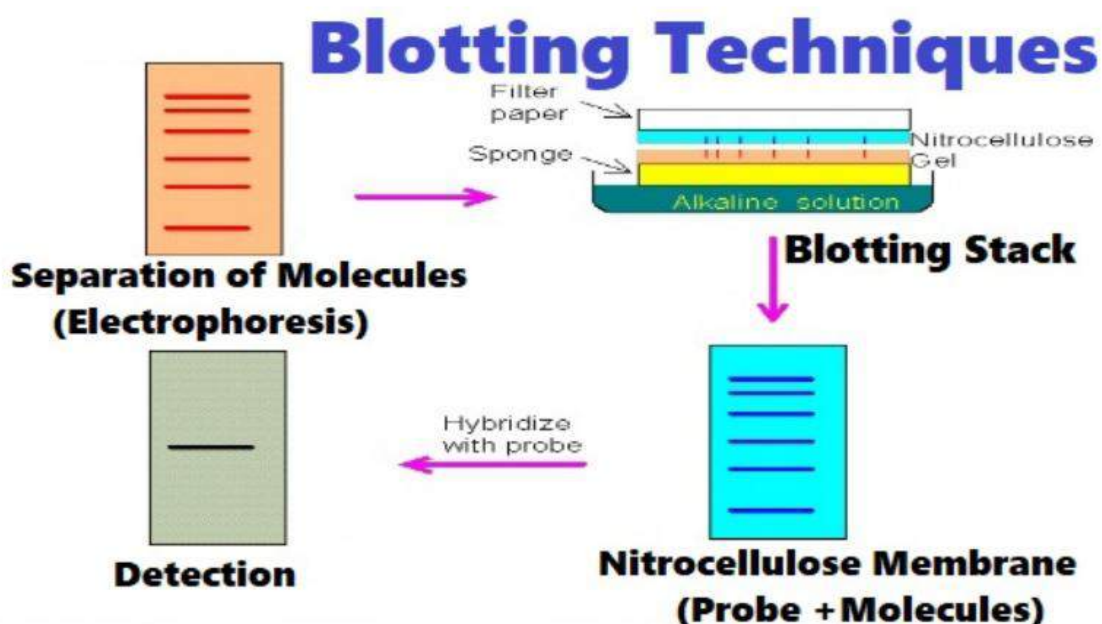
Blots are techniques for transferring DNA , RNA and proteins onto a carrier so they can be separated, and often follows the use of a gel electrophoresis. The Southern blot is used for transferring DNA, the Northern blot for RNA and the western blot for PROTEIN.

Blotting is a method of transferring proteins, DNA, RNA onto a carrier (special membrane). This membrane may be nitrocellulose, PVDF or nylon membrane. This process can be done after gel electrophoresis, by transferring the molecules from gel onto the surface of blotting membrane but sometime samples are directly added onto the membrane.

The process of transferring of the denatured fragments out of the gel and onto a membrane made from nylon (or sometimes nitrocellulose) where they become accessible for analyzing using a probe is blotting. They are analytical techniques used for identification of a specific DNA, RNA or a protein and are collectively referred to as blotting techniques.

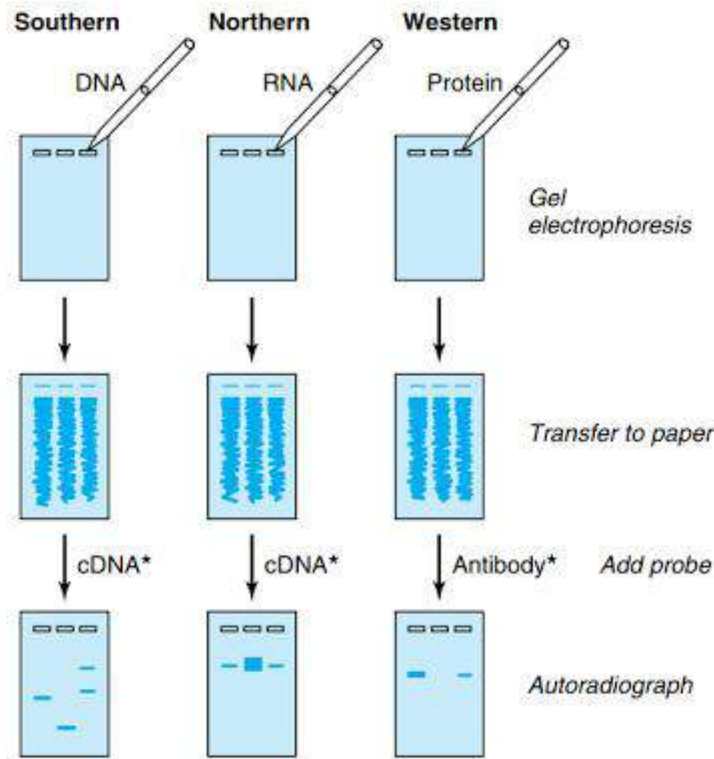
The key to this method is **hybridization** which is a process of forming a double-stranded molecule between a single-stranded probe and a single-stranded target segment. It is a phenomenon in which single-stranded deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) molecules anneal to complementary DNA or RNA.

Blotting of nucleic acid is the central technique for hybridization studies. Nucleic acid labeling and hybridization on membranes have formed the basis for a range of experimental techniques involving understanding of gene expression and its organization, Identifying and measuring specific proteins in complex biological mixtures etc.



Principle of blotting

The mixture of molecules such as proteins, DNA, RNA etc. (Fragmented if necessary) is separated by electrophoresis which are immobilized on a matrix. The probe is added to the matrix to bind to the molecules. Any unbound probes are then removed by washing. The place where probe is bound corresponds to the location of immobilized target molecule.



Southern Blotting:

Southern blot hybridization detects target DNA fragments that have been size-fractionated by gel electrophoresis. This technique was invented in 1975 by E.M. southern.

Principle:

In this technique we exploit the property of a radio-labelled probe with the single stranded DNA. If we want to detect the presence of a specific sequence in our mixed DNA sample then we will accordingly design the probe which will have complementary sequence to our target sequence.

Procedure:

In this procedure, called a Southern blot, DNA from the sample is cleaved into restriction fragments with a restriction endonuclease, and the fragments are spread apart by gel electrophoresis. The double-stranded helix of each DNA fragment is then denatured into single strands by making the pH of gel basic, and the gel is **"blotted"** with a sheet of nitrocellulose, transferring some of the DNA strands to the sheet.

Next, a probe consisting of purified, single-stranded DNA corresponding to a specific gene (or mRNA transcribed from that gene) is poured over the sheet.

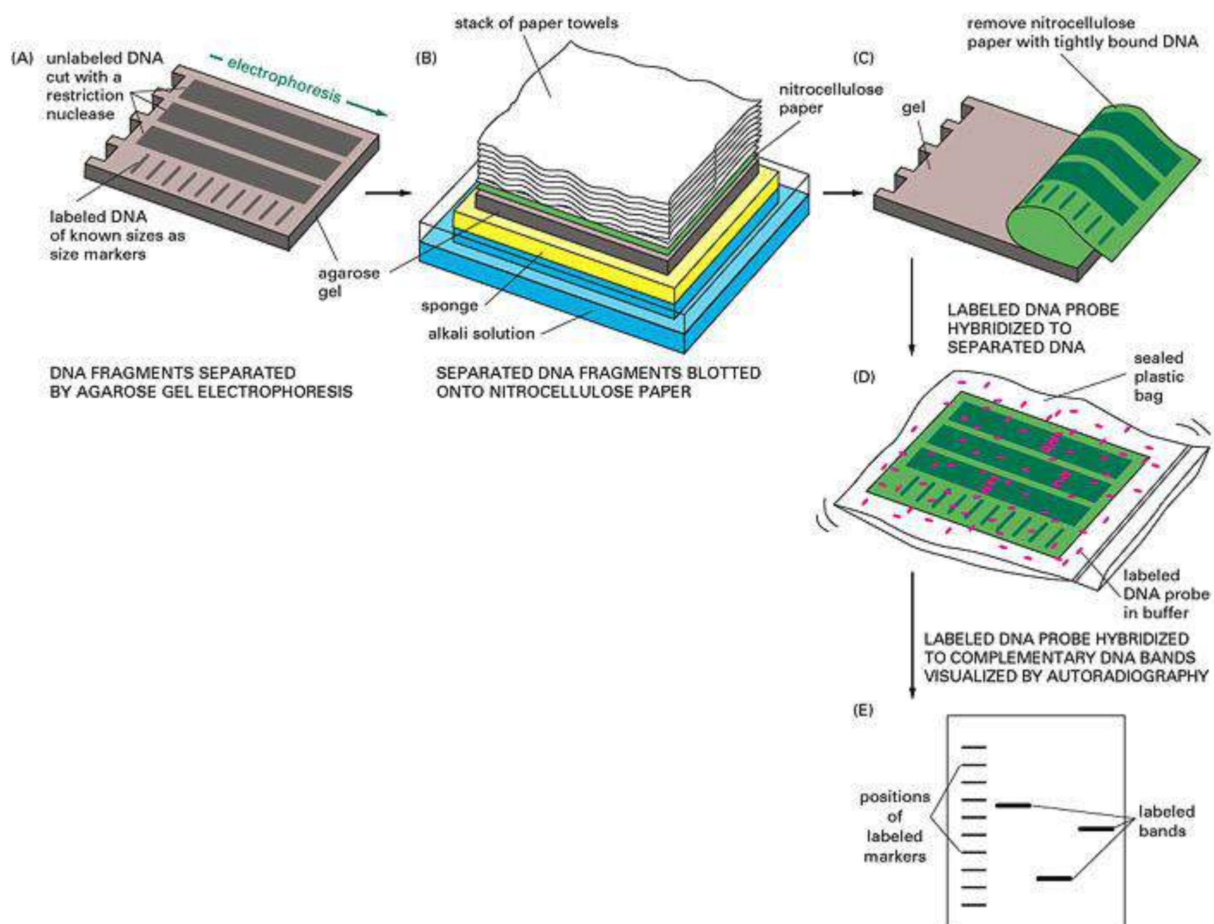
Any fragment that has a nucleotide sequence complementary to the probe's sequence will hybridize (base pair) with the probe. If the probe has been labelled with ^{32}P, it will be radioactive, and the sheet will show a band of radioactivity where the probe is hybridized with the complementary fragment.

1. Whether a particular gene is present and how many copies are present in the genome of an organism.

2. The degree of similarity between the chromosomal gene and the probe sequence.

3. Whether recognition sites for particular restriction endonucleases are present in the gene. By performing the digestion with different endonucleases, or with combinations of endonucleases, it is possible to obtain a restriction map of the gene, i.e., an idea of restriction enzyme sites in and around the gene which will assist in attempts to clone the gene.

4. Whether re-arrangements have occurred during the cloning process.



Uses of Southern Blotting:

1. To identify a single gene among thousands of fragments of DNA and to detect sequences of DNA in an organism's genome.

2. Used in gene discovery and gene mapping.

3. To analyse the genetic patterns in an organism's DNA.

4. To identify gene mutation, deletion, duplication, and gene rearrangement involved in diseases.

5. To determine the number of copies of a particular DNA sequence presented in the genome of an organism.

6. To identify related DNA sequence in the genome and to determine if there is a gene family (a group of similar genes).

7. To detect certain cancers and genetic diseases

Northern Blotting:

Northern blotting is a simple extension of Southern blotting, and derives its name from the earlier technique. It is one of the key techniques in molecular biology, its principal aim being the measurement of RNA (in particular mRNA).

Principle:

RNA molecules are separated by size and detected on a membrane using a hybridization probe with a base sequence complementary to all, or a part, of the sequence of target RNA.

Procedure:

RNA is extracted from the cells of interest, but precautions must be taken to avoid degradation of single-stranded RNA by ribonuclease (RNase), which is found on the skin and on glassware. Wear gloves use specially treated plastics and glassware to avoid accidental introducing ribonuclease to extraction prep.

Addition of di-ethyl-pyro-carbonate (DEPC) inhibits ribonuclease activity and also baking at high temperature destroys ribonuclease activity (only useful for treating heat resistant equipment like glassware).

It is performed in following steps:

Step 1:

RNA is isolated from several biological samples (e.g., various tissues, various developmental stages of same tissue, etc.)

Step 2:

The RNA samples are separated according to their size on an agarose gel.

Step 3:

The gel is then blotted on a nylon membrane or a nitrocellulose filter.

Step 4:

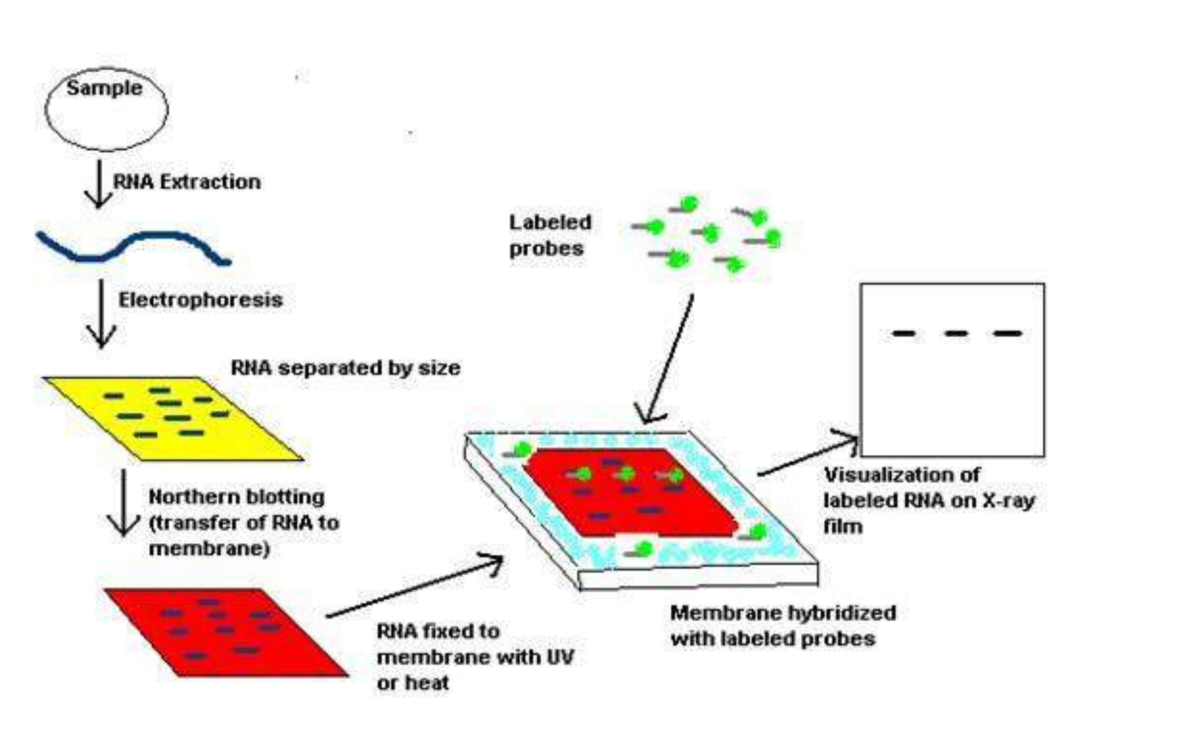
The membrane is placed in a dish containing hybridization buffer with a labelled probe. RNA blots are most usually probed with cDNA fragments.

Step 5:

The membrane is washed to remove unbound probe.

Step 6:

The labelled probe is detected via autoradiography (if a radioactive probe is used) or via a chemiluminescence reaction (if a chemically labelled probe is used). In both cases this results in the formation of a dark band on an X-ray film.



Uses of Northern Blotting:

1. Northern blotting allows researchers to determine gene expression patterns. This indicates a myriad of practical applications, allowing researchers to compare patterns of gene expression in cells of tissues, cells of patients undergoing treatment and cells of different developmental stages.

2. Northern blot analysis can also be used to detect cancerous pancreatic cells and tissues. In one study the researchers review that the pancreatic cancers exhibited 3- fold, 10-fold and 15-fold increase in mRNA of a certain receptor, indicating for the first time that this receptor was involved in carcinogenesis of pancreatic cancer. This information has been obtained by using Northern blotting technique.

3. Northern blotting can also enable the scientists to know the function of unknown proteins.

4. This technique enables the scientists to detect the size of RNA.

5. It also allows them to observe the alternate splice products, using the probes with partial homology.

Western Blotting:

The Western Blot is an analytical technique used to detect specific proteins in a given sample of tissue homogenate or extract. Sometime referred to as immune blotting, this technique uses gel electrophoresis to separate native or denatured proteins by the length of the polypeptide (denaturing conditions) or by the 3-D structure of the protein (native/ non-denaturing conditions).

Other related techniques include using antibodies to detect proteins in tissues and cells by immune staining and enzyme-linked immune sorbent assay (ELISA). This method originated from the laboratory of George Stark at Stanford. The name Western blot was given to the technique by W. Neal Burnette and is a play on the name Southern Blot, a technique for DNA detection developed earlier by Edwin Southern.

Principle:

It is an analytical method wherein a protein sample is electrophoresed on an SDS-PAGE and electro transferred onto nitrocellulose membrane. The transferred protein is detected using specific primary antibody and secondary enzyme labelled antibody and substrate.

Procedure:

In this technique first of all the sample of proteins is separated on the basis of their molecular mass using SDS-PAGE or two-dimensional electrophoresis. Electrophoresis moves the proteins from the gel and onto the nitrocellulose where proteins adhere.

To detect a specific protein, an antibody to that protein must be available. The nitrocellulose membrane itself has many non-specific sites that can bind proteins, including antibodies which must be blocked with a non-specific protein solution, such as re-hydrated powdered milk.

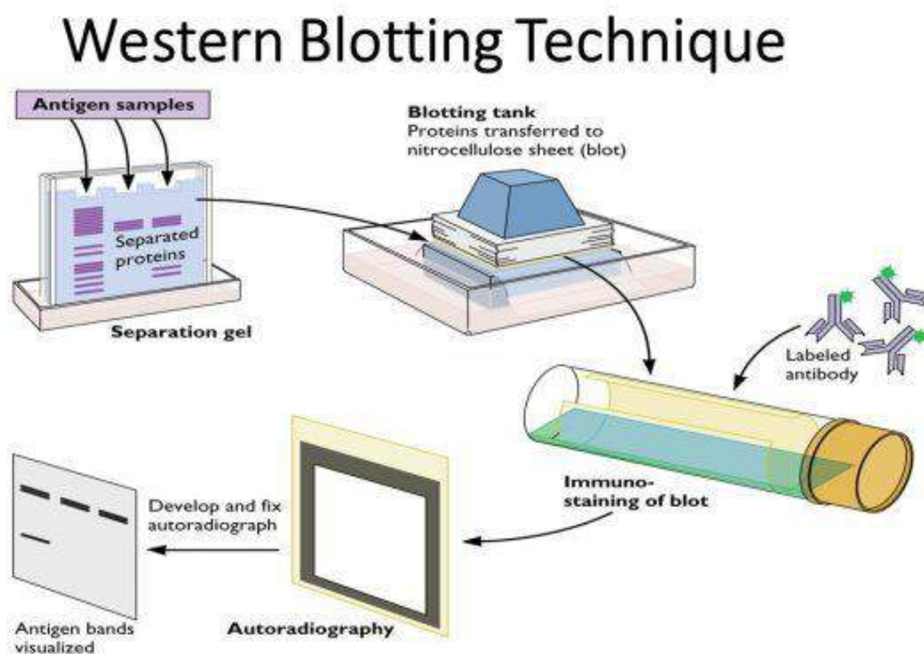
The primary antibody is added in the milk solution and binds to the protein of interest. The antibody protein complex is detected using a secondary antibody that has a label attached to it

(Fig. 3.27). Often a reporter enzyme such as alkaline phosphatase is linked to the secondary antibody, and the addition of lumiphos or X-phos to the blot allows detection of the protein band.

Uses of Western Blotting:

Western blotting is mostly used as a medical diagnostic technique. A positive Western blot can usually confirm an HIV infection. The confirmatory HIV test employs a Western blot to detect anti-HIV antibody in a human serum sample.

A Western blot is also used as the definitive test for Bovine spongiform encephalopathy (BSE, commonly referred to as 'mad cow disease'). Western blotting is also helpful in the diagnosis of some forms of Lyme disease.



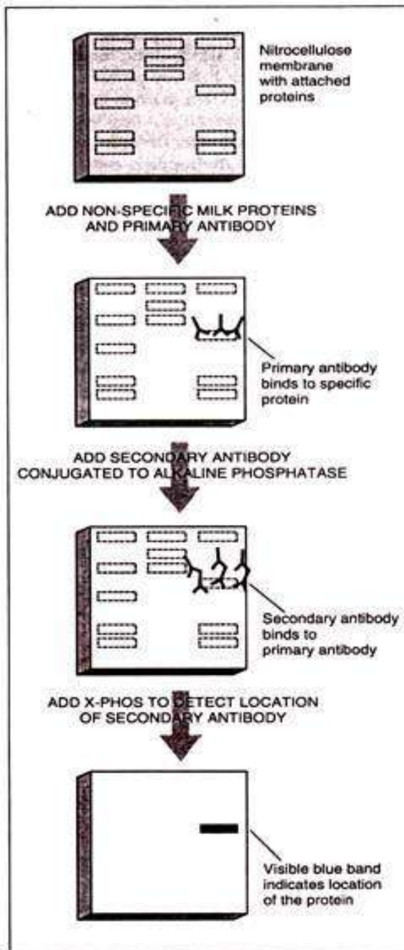


Fig. 3.28: Western Blotting technique

The **eastern blot**, or **eastern blotting**, is a biochemical technique used to analyze protein post-translational modifications (PTM) including the addition of lipids, phosphates, and glycoconjugates. It is most often used to detect carbohydrate epitopes. Thus, eastern blot can be considered an extension of the biochemical technique of western blot. Multiple techniques have been described by the term "eastern blot(ting)", most use prospo protein blotted from SDS-PAGE gel on to a PVDF or nitrocellulose membrane. Transferred proteins are analyzed for post-translational modifications using probes that may detect lipids, carbohydrate, phosphorylation or any other protein modification. Eastern blotting should be used to refer to methods that detect their targets through specific interaction of the PTM and the probe, distinguishing them from a standard far-western blot. In principle, eastern blotting is similar to lectin blotting (i.e. detection of carbohydrate epitopes on proteins or lipids).

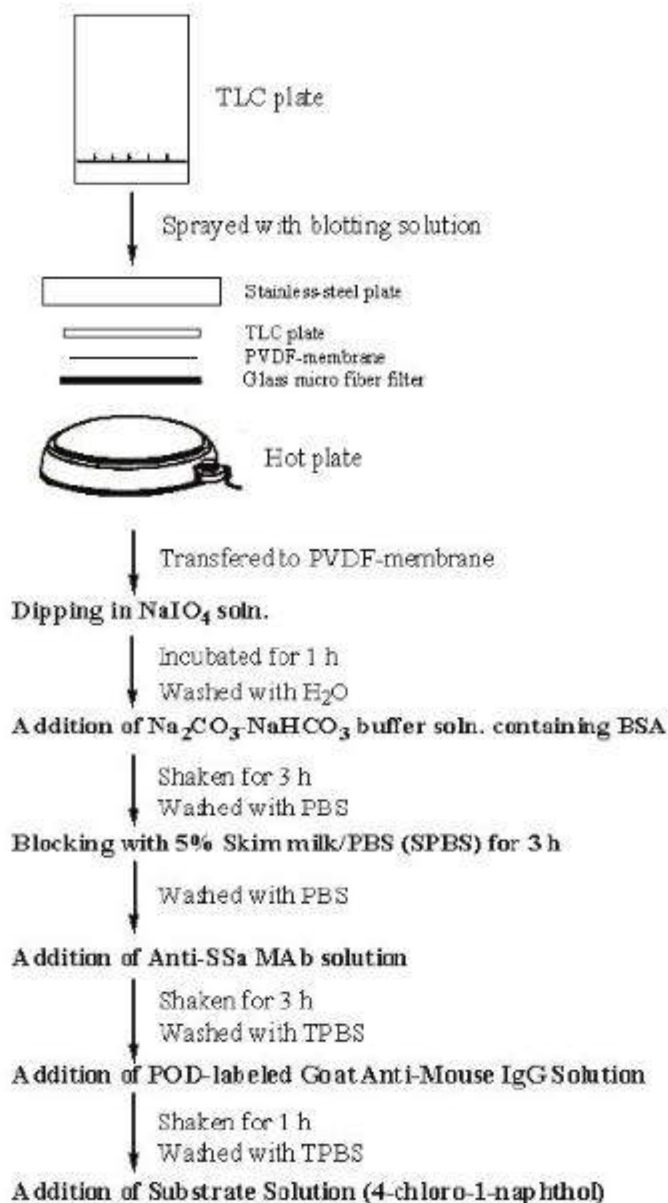
Application

- detection of protein modifications in bacterial species Ehrlichia- E. muris and IOE.
- The technique showed that the antigenic proteins of the non-virulent E. muris is more post-translationally modified than the highly virulent IOE

• PMT play an important role in translocation across biological membranes. • Expression of post-translated proteins is important in several diseases Eastern Blotting.

Far-eastern blotting is a technique developed in 1994 by Taki and colleagues at the Tokyo Medical and Dental University, Japan for the analysis of lipids separated by high-performance thin layer chromatography (HPTLC).

• The lipids are transferred from the HPTLC plate to a polyvinyledene difluoride (PVDF) membrane for further analysis, for example by enzymatic or ligand binding assays or mass spectrometry.



POLYMERASE CHAIN REACTION (PCR) AND ITS APPLICATIONS

Introduction:

Polymerase chain reaction (PCR) is a widely employed technique in molecular biology to amplify single or a few copies of DNA, generating millions of copies of a particular DNA sequence. The polymerase chain reaction results in the selective amplification of a target region of a DNA or RNA molecule. PCR has been extensively exploited in cloning, target detection, sequencing etc. The method consists of thermal cycles of repeated heating followed by cooling of the reaction mixture to achieve melting and primer hybridization to enable enzymatic replication of the DNA.

A basic PCR set up requires the following essential components and reagents :

1. Template DNA containing the DNA region (target) to be amplified.

2. Primers that are complementary to the 5' ends of each of the sense (Forward primer) and anti-sense strand of the DNA target (Reverse primer).

3. Taq polymerase or other thermostable, high fidelity DNA polymerase (Pfu polymerase isolated from Pyrococcus furiosus).

4. Deoxyribonucleotide triphosphates (dNTPs), which are the building-blocks for a newly synthesized DNA strand.

5. Buffer solutions to provide a suitable chemical condition for optimum activity and stability of the DNA polymerases.

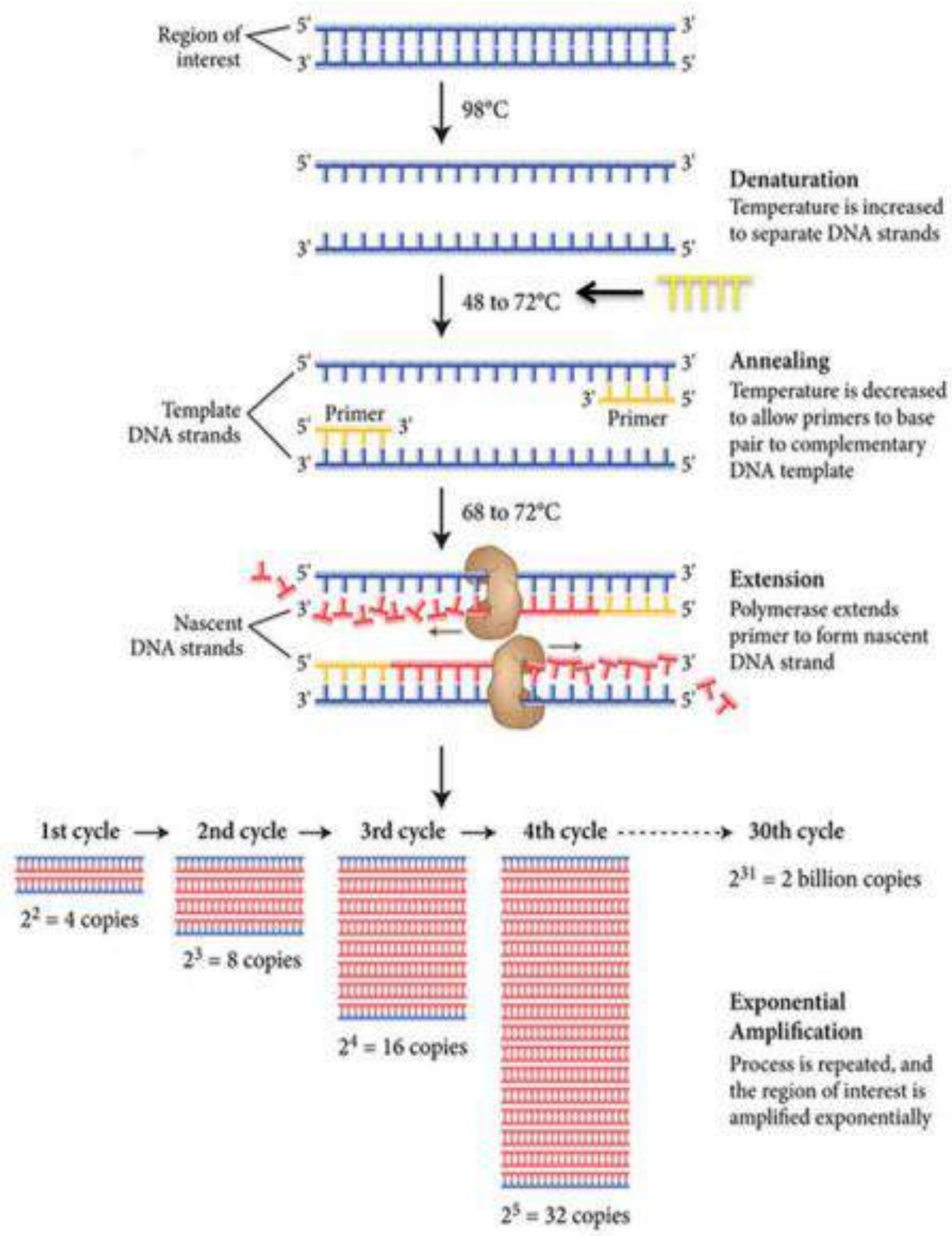
6. Divalent cations (eg. magnesium or manganese ions). They act as a co-factor for Taq polymerase which increases its polymerase activity

Procedure:

Typically, PCR is designed of 20-40 repeated thermal cycles, with each cycle consisting of 3 discrete temperature steps: denaturation, annealing and extension. The thermal cycles are often proceeded by a temperature at a high range (>90°C), and followed by final product extension or brief storage at 4 degree celsius. In PCR cycles, the temperatures and the duration of each cycle is determined based on various parameters like the type of DNA polymerase used, the melting temperature (Tm) of the primers, concentration of divalent ions and dNTPs in the reaction etc.

The various steps involved are:-

a) Initial Denaturation
b) Denaturation
c) Annealing
d) Extension
e) Final extension.



: The sequential steps of PCR

Restriction Fragment Length Polymorphism (RFLP):

Restriction fragment length polymorphism, commonly pronounced "rif-lip", special types of SNPs, are gene markers which result in a restriction site being changed. When digested with a restriction endonuclease the loss of the site is revealed because two fragments remain joined together.

Analysis Technique:

The basic technique for detecting RFLPs involves fragmenting a sample of DNA by a restriction enzyme, which can recognize and cut DNA wherever a specific short sequence occurs, in a process known as a restriction digest. The resulting DNA fragments are then separated by length by agarose gel electrophoresis, and transferred to a membrane via the Southern blot procedure.

Hybridization of the membrane to a labelled DNA probe then determines the length of the fragments which are complementary to the probe. A RFLP occurs when the length of a detected fragment varies between individuals. Each fragment length is considered an allele, and can be used in genetic analysis.

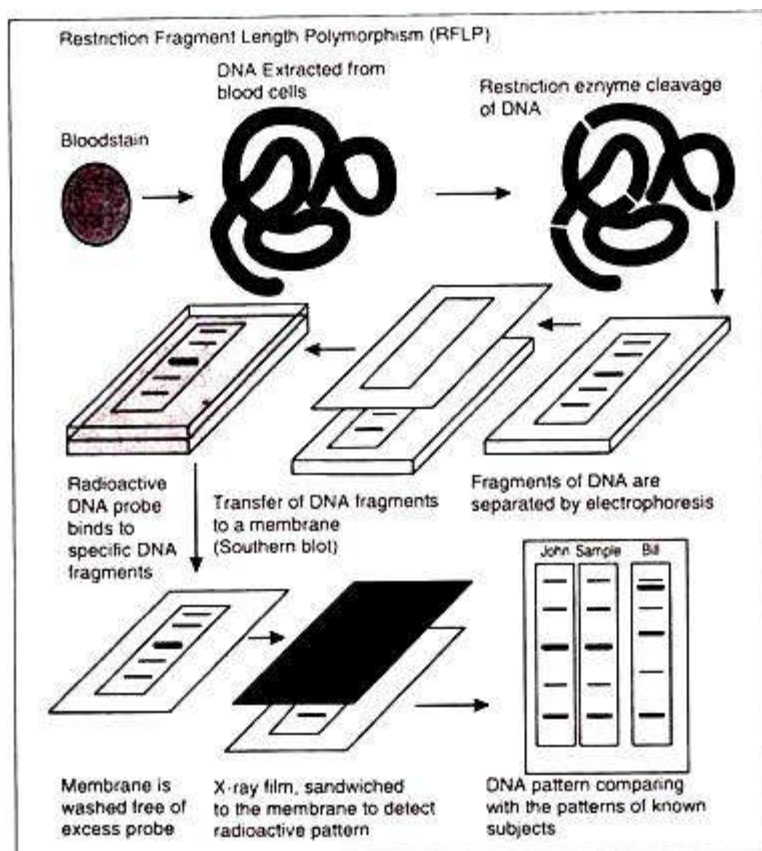


Fig. 8.2: RFLP analysis by taking DNA sample from blood stain

Applications:

Analysis of RFLP variation in genomes was a vital tool in genome mapping and genetic disease analysis. If researchers were trying to initially determine the chromosomal location of a

particular disease gene, they would analyse the DNA of members of a family afflicted by the disease, and look for RFLP alleles that show a similar pattern of inheritance as that of the disease.

Once a disease gene was localized, RFLP analysis of other families could reveal who was at risk for the disease, or who was likely to be carriers of the mutant genes.

Advantages:

RFLP markers are co-dominant. Hence identifications of patterns of genomes in homozygous and heterozygous individuals are easier.

Disadvantages:

May require large amounts of DNA samples since distinct RFLP markers require different restriction enzymes.

Random Amplification of Polymorphic DNA (RAPD):

The RAPD, pronounced **"rapid"**, utilizes short synthetic oligonucleotides of random sequences as primers to amplify Nano gram amounts of total genomic DNA under low annealing temperatures by PCR. Amplification products are generally separated on agarose gels and stained with ethidium bromide.

At an appropriate annealing temperature during the thermal cycle, oligonucleotide primers of random sequence bind several priming sites on the complementary sequences in the template genomic DNA and produce discrete DNA products if these priming sites are within an amplifiable distance of each other.

The profile of amplified DNA primarily depends on nucleotide sequence homology between the template DNA and oligonucleotide primer at the end of each amplified product. Nucleotide variation between different sets of template DNAs will result in the presence or absence of bands because of changes in the priming sites.

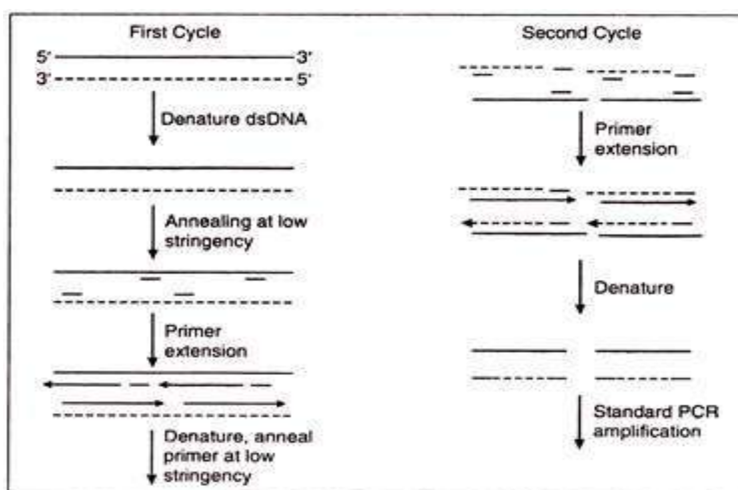


Fig. 8.5: Schematic diagram of RAPD reaction for 2 loci (derived from Welsh and McClelland, 1994)

Applications:

1. Used in genetic mapping and DNA finger printing.

2. Sex determination.

3. Generation of specific PCR primers for anonymous Genomes.

4. Quantitative analysis of mixed bio-samples.

5. Determination of paternity and kinship relationships.

6. Analyses of interspecific gene flow and hybrid speciation.

7. Determination of taxonomic identity.

Advantage:

Can use random primer sequences.

Disadvantages:

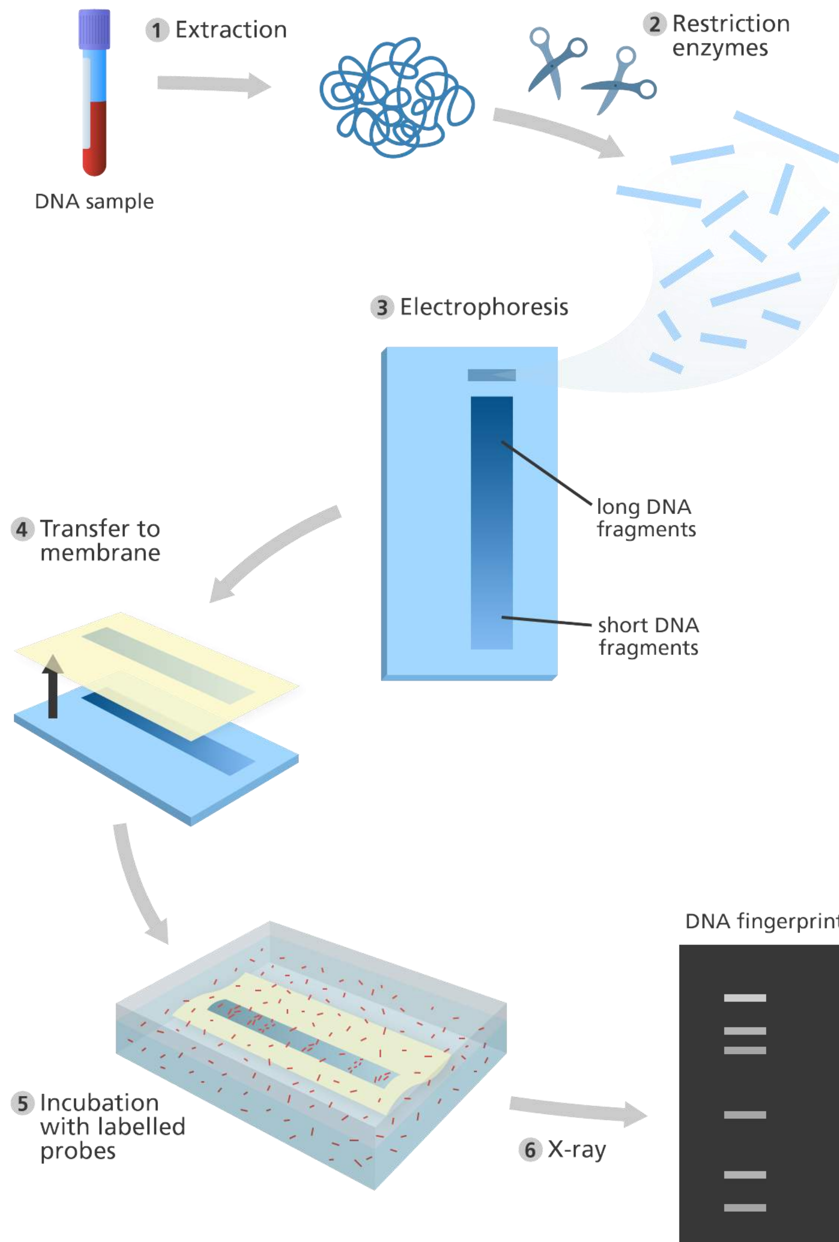
(1) RAPD markers are dominant. Hence, distinguishing homozygous and heterozygous individuals is not possible.

(2) Reproducibility of the experiment is very low since low annealing temperatures are used.

<u>DNA Fingerprinting-</u>

- DNA fingerprinting or DNA profiling is a process used to determine the nucleotide sequence at a certain part of the **DNA** that is unique in all human beings.
- The process of DNA fingerprinting was invented by Sir Alec Jeffrey at the University of Leicester in 1985.
- The DNA of every human being on the planet is 99.9% same. However, about 0.1% or 3 x 10^6base pairs (out of 3 x 10^9 bp) of DNA is unique in every individual.
- Human genome possesses numerous small non-coding but inheritable sequences of bases which are repeated many times. They do not code for proteins but make-up 95% of our genetic DNA and therefore called the ―junk DNA.
- They can be separated as satellite from the bulk DNA during density gradient centrifugation and hence called satellite DNA.
- In satellite DNA, repetition of bases is in tandem. Depending upon length, base composition and numbers of tandemly repetitive units, satellite DNAs have subcategories like microsatellites and mini-satellites.

- Satellite DNAs show polymorphism. The term polymorphism is used when a variant at a locus is present with a frequency of more than 0.01 population.



- Variations occur due to mutations. These mutations in the non-coding sequences have piled up with time and form the basis of DNA polymorphism (variation at genetic level arises due to mutations).
- The junk DNA regions are thus made-up of length polymorphisms, which show variations in the physical length of the DNA molecule.

- At specific loci on the chromosome the number of tandem repeats varies between individuals. There will be a certain number of repeats for any specific loci on the chromosome.
- Depending on the size of the repeat, the repeat regions are classified into two groups. **Short tandem repeats (STRs)** contain 2-5 base pair repeats and **variable number of tandem repeats (VNTRs)** have repeats of 9-80 base pairs.
- Since a child receive 50% of the DNA from its father and the other 50% from his mother, so the number VNTRs at a particular area of the DNA of the child will be different may be due to insertion, deletion or mutation in the base pairs.
- As a result, every individual has a distinct composition of VNTRs and this is the main principle of DNA fingerprinting.
- As single change in nucleotide may make a few more cleavage site of a given nucleotide or might abolish some existing cleavage site.
- Thus, if DNA of any individual is digested with a restriction enzyme, fragments pattern (sizes) will be produced and will be different in cleavage site position. This is the basics of DNA fingerprinting.

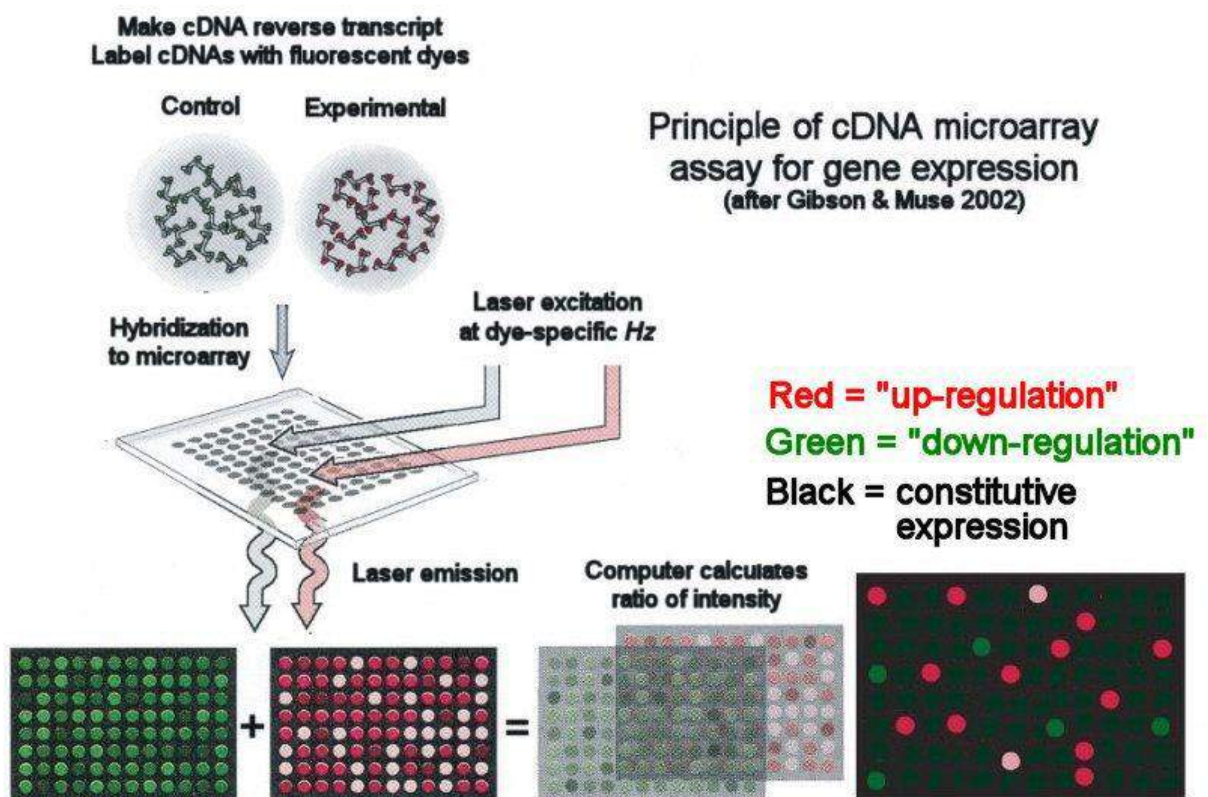
DNA Microarray

- **DNA** microarrays are solid supports, usually of glass or silicon, upon which DNA is attached in an organized pre-determined grid fashion.
- Each spot of DNA, called a probe, represents a single gene.
- DNA microarrays can analyze the expression of tens of thousands of genes simultaneously.
- There are several synonyms of DNA microarrays such as DNA chips, gene chips, DNA arrays, gene arrays and biochips.

Principle of DNA microarrays

- The principle of DNA microarrays lies on the hybridization between the **nucleic acid** strands.
- The property of complementary nucleic acid sequences is to specifically pair with each other by forming hydrogen bonds between complementary nucleotide base pairs.

- For this, samples are labeled using fluorescent dyes.
- At least two samples are hybridized to chip.
- Complementary nucleic acid sequences between the sample and the probe attached on the chip get paired via hydrogen bonds.
- The non-specific bonding sequences while remain unattached and washed out during the washing step of the process.
- Fluorescently labeled target sequences that bind to a probe sequence generate a signal.
- The signal depends on the hybridization conditions (ex: temperature), washing after hybridization etc while the total strength of the signal, depends upon the amount of target sample present.
- Using this technology the presence of one genomic or cDNA sequence in 1,00,000 or more sequences can be screened in a single hybridization.



The reaction procedure of DNA microarray takes places in several steps:

1. **Collection of samples**

- The sample may be a cell/tissue of the organism that we wish to conduct the study on.

- Two types of samples are collected: healthy cells and infected cells, for comparison and to obtain the results.

2. **Isolation of mRNA**

- RNA is extracted from the sample using a column or solvent like phenol-chloroform.
- From the extracted RNA, mRNA is separated leaving behind rRNA and tRNA.
- As mRNA has a poly-A tail, column beads with poly-T-tails are used to bind mRNA.
- After the extraction, the column is rinsed with buffer to isolate mRNA from the beads.

3. **Creation of labeled cDNA**

- To create cDNA (complementary DNA strand), reverse transcription of the mRNA is done.
- Both the samples are then incorporated with different fluorescent dyes for producing fluorescent cDNA strands. This helps in distinguishing the sample category of the cDNAs.

4. **Hybridization**

- The labeled cDNAs from both the samples are placed in the DNA microarray so that each cDNA gets hybridized to its complementary strand; they are also thoroughly washed to remove unbounded sequences.

5. **Collection and analysis**

- The collection of data is done by using a microarray scanner.
- This scanner consists of a laser, a computer, and a camera. The laser excites fluorescence of the cDNA, generating signals.
- When the laser scans the array, the camera records the images produced.
- Then the computer stores the data and provides the results immediately. The data thus produced are then analyzed.
- The difference in the intensity of the colors for each spot determines the character of the gene in that particular spot.

Application of recombinant DNA technology

1. **Production of Transgenic Plants**:By utilizing the tools and techniques of genetic engineering it is possible to produce transgenic plants or the genetically modified plants. Many transgenic plants have been developed with better qualities like resistance to herbicides, insects or viruses or with expression of male sterility etc.

2. **Production of Transgenic Animals**:By the use of rec DNA technology, desired genes can be inserted into the animal so as to produce the transgenic animal. The method of rec DNA technology aids the animal breeders to increase the speed and range of selective breeding in case of animals. It helps for the production of better farm animals so as to ensure more commercial benefits. Another commercially important use of transgenic animals is the production of certain proteins and pharmaceutical compounds. Transgenic animals also contribute for studying the gene functions in different animal species. Biotechnologists have successfully produced transgenic pigs, sheep, rats and cattle.

3. **Production of Hormones**:By the advent of techniques of rec DNA technology, bacterial cells like E.coli are utilized for the production of different fine chemicals like insulin, somatostatin, somatotropin and pendorphin. Human Insulin Hormone i.e., Humulin is the first therapeutic product which was produced by the application of rec DNA technology.

4. **Production of Vaccines:**Vaccines are the chemical preparations containing a pathogen in attenuated (or weakened) or inactive state that may be given to human beings or animals to confer immunity to infection. A number of vaccines have been synthesized biologically through rec DNA technology, these vaccines are effective against numerous serious diseases caused by bacteria, viruses or protozoa. These include vaccines for polio, malaria, cholera, hepatitis, rabies, smallpox, etc. The generation of DNA vaccines has revolutionized the approach of treatment of infectious diseases. DNA-vaccine is the preparation that contains a gene encoding an immunogenic protein from the concerned pathogen.

5. **Biosynthesis of Interferon**:Interferon's are the glycoproteins which are produced in very minute amounts by the virus-infected cells. Interferon's have antiviral and even anti-cancerous properties. By recDNA technology method, the gene of human fibroblasts (which produce interferon's in human beings) is inserted into the bacterial plasmid.These genetically engineered bacteria are cloned and cultured so that the gene is expressed and the interferon's are produced in fairly high quantities. This interferon, so produced, is then extracted and purified.

6. **Production of Antibiotics**:Antibiotics produced by microorganisms are very effective against different viral, bacterial or protozoan diseases. Some important antibiotics are tetracyclin, penicillin, streptomycin, novobiocin, bacitracin, etc.recDNA technology helps in increasing the production of antibiotics by improving the microbial strains through modification of genetic characteristics.

7. **Production of Commercially Important Chemicals**:Various commercially important chemicals can be produced more efficiently by utilizing the methods of rec DNA technology.

A few of them are the alcohols and alcoholic beverages obtained through fermentation; organic acids like citric acid, acetic acid, etc. and vitamins produced by microorganisms.

8. **Application in Enzyme Engineering**:As we know that the enzymes are encoded by genes, so if there are changes in a gene then definitely the enzyme structure also changes. Enzyme engineering utilizes the same fact and can be explained as the modification of an enzyme structure by inducing alterations in the genes which encode for that particular enzyme.

9. **Prevention and Diagnosis of Diseases:**Genetic engineering methods and techniques have greatly solved the problem of conventional methods for diagnosis of diseases. It also provides methods for the. prevention of a number of diseases like AIDS, cholera, etc. Monoclonal antibodies are useful tools for disease diagnosis. Monoclonal antibodies are produced by using the technique called hybridoma technology.

10. **Gene Therapy:**Gene therapy is undoubtedly the most beneficial area of genetic engineering for human beings. It involves delivery of specific genes into human body to correct the diseases. Thus, it is the treatment of diseases by transfer and expression of a gene into the patients' cells so as to ensure the restoration of a normal cellular activity.

11. **Practical Applications** of Genetic Engineering:recDNA technology has an immense scope in Research and Experimental studies. It is applied for:

a. Localizing specific genes.
b. Sequencing of DNA or genes.
c. Study of mechanism of gene regulation.
d. Molecular analysis of various diseases.
e. Study of" mutations in DNA, etc.

12. **Applications in forensic science:** The applications of rec DNA technology (or genetic engineering) in forensic sciences largely depend on the technique called DNA profiling or DNA fingerprinting. It enables us to identify any person by analysing his hair roots Wood stains, serum, etc. DNA fingerprinting also helps to solve the problems of parentage and to identify the criminals.

13. **Biofuel Production**: Biofuels are derived from biomass and these are renewable and cost effective. Genetic engineering plays an essentially important role in a beneficial and largescale production of biofuels like biogas. bio hydrogen biodiesel bio-ethanol., etc. Genetic engineering helps to improve organisms for obtaining higher product yields and product tolerance.

14. **Genetically stable high producing microorganisms** are being developed by using modern recDNA techniques, which aid in an efficient production of bioenergy.

15. **The energy crop plants** are those plants which use solar energy in a better way for production of biomass. Genetic improvements of these energy crop plants greatly help for quick and high Product on of biomass which in turn reduces the biofuel production cost. The fermenting microbes which are utilized for biogas production are improved at the genetic level for achieving better result.

16. **Environment Protection:**Genetic engineering makes its contributions to the environment protection in various ways. Most important to mention are the new approaches utilized for waste treatments and bioremediation Environment protection means the conservation of resources and hence to limit the degradation of environment.