BIOTECHNOLOGY AND BIOINFORMATICS

Subject Code:**18KP3Z09**

UNIT: I

Recombinant DNA technology

Recombinant DNA (rDNA) is a form of artificial DNA that is created by combining two or more sequences that would not normally occur together through the process of gene splicing.

Recombinant DNA technology is a technology which allows DNA to be produced via artificial means. The procedure has been used to change DNA in living organisms and may have even more practical uses in the future.

Recombinant DNA technology is one of the recent advances in biotechnology, which was developed by two scientists named Boyer and Cohen in 1973.

Recombinant technology begins with the isolation of a gene of interest (target gene). The target gene is then inserted into the plasmid or phage (vector) to form replicon. The replicon is then introduced into host cells to be cloned and either express the protein or not. The cloned replicon is referred to as recombinant DNA. The procedure is called recombinant DNA technology.

Cloning is necessary to produce numerous copies of the DNA since the initial supply is inadequate to insert into host cells. Cloning, In classical biology, a clone is a population of identical organisms derived from a single parental organism. For example, the members of a colony of bacterial cells that arise from a single cell on a petridish are clones. Molecular biology has borrowed the term to mean a collection of molecules or cells all identical to an original molecule or cell. Six steps of Recombinant DNA1. Isolating (vector and target gene) 2. Cutting (Cleavage) 3. Joining (Ligation) 4. Transforming 5. Cloning 6. Selecting (Screening)

Six basic steps are common to most recombinant DNA experiments

1. Isolation and purification of DNA:

Both vector and target DNA molecules can be prepared by a variety of routine methods. In some cases, the target DNA is synthesized in vitro.

2. Cleavage of DNA at particular sequences:

As we will see, cleaving DNA to generate fragments of defined length, or with specific endpoints, is crucial to recombinant DNA technology. The DNA fragment of interest is called insert DNA. In the laboratory, DNA is usually cleaved by treating it with commercially produced nucleases and restriction endonucleases.

3. Ligation of DNA fragments:

A recombinant DNA molecule is usually formed by cleaving the DNA of interest to yield insert DNA and then ligating the insert DNA to vector DNA (recombinant DNA or chimeric DNA). DNA fragments are typically joined using DNA ligase (also commercially produced). – T4 DNA Ligase

4. Introduction of recombinant DNA into compatible host cells:

In order to be propagated, the recombinant DNA molecule (insert DNA joined to vector DNA) must be introduced into a compatible host cell where it can replicate. The direct uptake of foreign DNA by a host cell is called genetic transformation (or transformation). Recombinant DNA can also be packaged into virus particles and transferred to host cells by transfection.

5. Replication and expression of recombinant DNA in host cells:

Cloning vectors allow insert DNA to be replicated and, in some cases, expressed in a host cell. The ability to clone and express DNA efficiently depends on the choice of appropriate vectors and hosts.

6. Identification of host cells that contain recombinant DNA of interest:

Vectors usually contain some genetic markers, or genes, that allow the selection of host cells that have taken up foreign DNA. The identification of a particular DNA fragment usually involves an additional step—screening a large number of recombinant DNA clones. This is almost always the most difficult step.

Applications of Recombinant DNA Technology:

The three important applications are: **(1) Applications in Crop Improvement (2) Applications in Medicines and (3) Industrial Applications.**

I. Applications in Crop Improvement:

Genetic engineering has several potential applications in crop improvement, such as given below:

1. Distant Hybridization:

With the advancement of genetic engineering, it is now possible to transfer genes between distantly related species. The barriers of gene transfer between species or even genera have been overcome. The desirable genes can be transferred even from lower organisms to higher organisms through recombinant DNA technology.

2. Development of Transgenic Plants:

Genetically transformed plants which contain foreign genes are called transgenic plants. Resistance to diseases, insects and pests, herbicides, drought; metal toxicity tolerance; induction of male sterility for plant breeding purpose; and improvement of quality can be achieved through this recombinant DNA technology. BT-cotton, resistant to bollworms is a glaring example.

3. Development of Root Nodules in Cereal Crops:

Leguminous plants have root-nodules which contain nitrogen fixing bacteria Rhizobium. This bacteria converts the free atmospheric nitrogen into nitrates in the root nodules. The bacterial genes responsible for this nitrogen fixation can be transferred now to cereal crops like wheat, rice, maize, barley etc. through the techniques of genetic engineering thus making these crops too capable of fixing atmospheric nitrogen.

4. Development of C_4 Plants:

Improvement in yield can be achieved by improving the photosynthetic efficiency of crop plants. The photosynthetic rate can be increased by conversion of C_3 plants into C_4 plants, which can be achieved either through protoplasm fusion or recombinant DNA technology C_4 plants have higher potential rate of biomass production than C_3 plants. Most C_4 plants (sorghum, sugarcane, maize, some grasses) are grown in tropical and subtropical zones.

II. Applications in Medicines:

Biotechnology, especially genetic engineering plays an important role in the production of antibiotics, hormones, vaccines and interferon in the field of medicines.

1. Production of Antibiotics:

Penicillium and Streptomyces fungi are used for mass production of famous antibiotics penicillin and streptomycin. Genetically efficient strains of these fungi have been developed to greatly increase the yield of these antibiotics.

2. Production of Hormone Insulin:

Insulin, a hormone, used by diabetics, is usually extracted from pancreas of cows and pigs. This insulin is slightly different in structure from human insulin. As a result, it leads to allergic reactions in about 5% patients. Human gene for insulin production has been incorporated into bacterial DNA and such genetically engineered bacteria are used for large scale production of insulin. This insulin does not cause allergy (see figure 1.6).

3. Production of Vaccines:

Vaccines are now produced by transfer of antigen coding genes to disease causing bacteria. Such antibodies provide protection against the infection by the same bacteria or virus.

4. Production of Interferon:

Interferon's are virus-induced proteins produced by virus-infected cells. Interferon are antiviral in action and act as first line of defense against viruses causing serious infections, including breast cancer and lymph nodes malignancy. Natural interferon is produced in very small quality from human blood cells. It is thus very costly also. It is now possible to produce interferon by recombinant DNA technology at much cheaper rate.

5. Production of Enzymes:

Some useful enzymes can also be produced by recombinant DNA technique. For instance, enzyme urikinase, which is used to dissolve blood clots, has been produced by genetically engineered microorganisms.

6. Gene Therapy:

Genetic engineering may one day enable the medical scientists to replace the defective genes responsible for hereditary diseases (e.g., haemophilia, phenylketonuria, alkaptonuria) with normal genes. This new system of therapy is called gene therapy.

7. Solution of Disputed Parentage:

Disputed cases of parentage can now be solved most accurately by recombinant technology than by blood tests.

8. Diagnosis of Disease:

Recombinant DNA technology has provided a broad range of tools to help physicians in the diagnosis of diseases. Most of these involve the construction of probes: short Segments of single stranded DNA attached to a radioactive or fluorescent marker. Such probes are now used for identification of infectious agents, for instance, food poisoning Salmonella, Pus forming Staphylococcus, hepatitis virus, HIV, etc. By testing the DNA of prospective genetic disorder carrier parents, their genotype can be determined and their chances of producing an afflicted child can be predicted.

9. Production of Transgenic Animals:

Animals which carry foreign genes are called transgenic animals.

Examples:

Cow, sheep, goat – therapeutic; human proteins in their milk. Fish like common carp, cat fish, salmon and gold fish contain human growth hormone (hGH).

III. Industrial Applications:

In industries, recombinant DNA technique will help in the production of chemical compounds of commercial importance, improvement of existing fermentation processes and production of proteins from wastes. This can be achieved by developing more efficient strains of microorganisms. Specially developed microorganisms may be used even to clean up the pollutants. Thus, biotechnology, especially recombinant DNA technology has many useful applications in crop improvement, medicines and industry.

Cloning

Cloning is the process of generating a genetically identical copy of a cell or an organism. Cloning happens all the time in nature. In biomedical research, cloning is broadly defined to mean the duplication of any kind of biological material for scientific study, such as a piece of DNA or an individual cell.

6

Gene Cloning Vectors: A General Account

The DNA that carries the desired gene to the host cell is called ***gene cloning vector.*** It is also known as ***cloning vector, vector, cloning vehicle*** *or* ***carrier DNA.***

Genetic engineering involves transfer of a desired gene to a host cell that has to be manipulated. The desired gene may be hydrolyzed by cellular enzymes immediately after introduction. Further, the chance for its expression is also very poor. In order to overcome these problems, the desired gene is inserted into a suitable vector for gene cloning.

Plasmids, ***viral DNAs*** and ***cosmids*** are used as gene cloning vectors.

Plasmids are circular, double-stranded DNAs usually present in prokaryotic cells. They can carry a foreign DNA of 5-15 kbp size to bacteria. Eg. *pBR322.* Some plasmids carry genes to plant cells. Eg. *Ti plasmid.*

Viral DNAs are linear or circular and single stranded or double stranded DNAs. DNAs of ***λ phage*** and ***M13*** phage are used to carry genes to bacteria. DNA of Simian Virus 40 (SV40) is used to carry genes to animal cells. Viral DNAs can carry DNAs of 10-25 kbp size.

Cosmids are a type of constructed plasmids containing complementary single stranded sites (cos-sites) of λ DNA. They can carry DNA segments of 25-45 kbp size to bacteria. Eg. *pHV 79.*

A suitable cloning vector is chosen and cut with a restriction enzyme. The desired gene is inserted into the cleaved vector by using DNA li[illegible]. The resulting recombinant vector (rDNA) is then introduced into a host cell for gene manipulation.

Plasmids

Plasmids *are small, circular, double stranded, extrachromosomal DNAs present in bacterial cells.* They are inherited sharply without the influence of chromosomal DNA. They replicate independently due to the presence of an origin of replication.

The plasmids are 1 kbp-200 kbp in size and have limited number of genes.

Most bacteria contain more than one copy of each plasmid. The number of copies of a plasmid present in a cell is called ***copy number***. The copy number of plasmids usually varies from 1 to 50. However, it can be further increased by treating the bacterial culture with ***chloramphenicol*** (an inhibitor of protein synthesis).

The genes for antibiotics resistance, nitrogen fixation, nodulation, environmental stresses, etc. occur in plasmid DNAs. The antibiotics-resistance in plasmids can be used as genetic marker to identify the strains containing the plasmids. Some plasmids code for some secondary metabolites.

Some plasmids, under certain conditions, integrate into the chromosomal DNA of the bacterium. Such plasmids are called ***episomes***. The integrated plasmid replicates along with the chromosomal DNA. Eg. *F- plasmid.*

The eukaryotes, except yeasts, do not have plasmids. The yeast *Saccharomyces cerevisiae* contains YEp (yeast episomal plasmid or 2- mi

7

Plasmids : Gene Cloning Vectors for Bacteria

***Plasmids** are small, circular, double-stranded, extrachromosomal DNAs present in bacterial cells.* They are inherited sharply without the influence of chromosomal DNA. They replicate independently due to the presence of an origin of replication.

The plasmids are 1 kbp- 200 kbp in size and have limited number of genes.

Most bacteria contain more than one copy of each plasmid. The number of copies of a plasmid present in a cell is called ***copy number***. The copy number of plasmids usually varies from 1 to 50. However, it can be further increased by treating the bacterial culture with ***chloramphenicol*** (an inhibitor of protein synthesis).

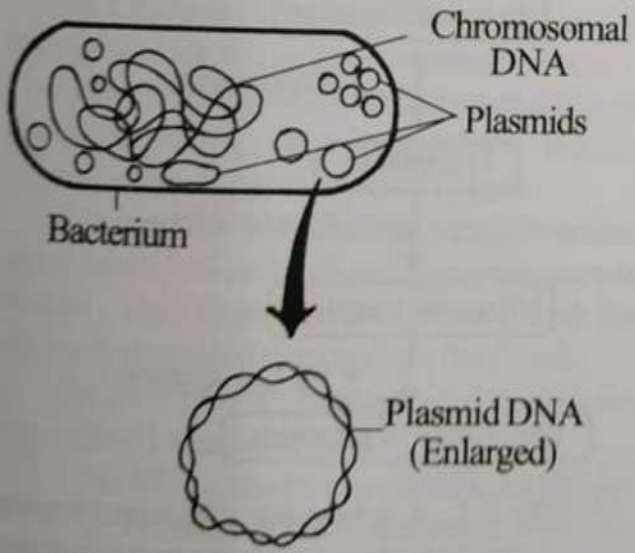


Fig.7.1: A bacterial cell showing chromosomal DNA and plasmids.

The genes for antibiotics resistance, nitrogen fixation, nodulation, environmental stresses, etc. are in plasmid DNA. The antibiotics resistance in plasmids can be used as ***genetic marker*** to identify those strains which contain the plasmids. Some plasmids code for some secondary metabolites.

Some plasmids, under certain conditions, integrate into the chromosomal DNA of the bacterium. Such plasmids are called ***episomes***. The integrated plasmid replicates along with the chromosomal DNA. Eg. *F- plasmid.*

The eukaryotes, except yeasts, do not have plasmids. The yeast *Saccharomyces cerevisiae* contains YEp (yeast episomal plasmid or 2-micron plasmid), YIp (yeast integrating plasmid) and ARS (Automatically replicating sequence) in the cells.

Classification of Plasmids

On the basis of conjugative transfer, plasmids are classified into two categories. They are-

i. Conjugative Plasmids : These plasmids are transferred from one bacterium to the other during conjugation. They contain ***'tra genes'*** for the conjugative transfer. Eg. *F-plasmid.*

ii. Non-conjugative Plasmids : These plasmids do not pass from one bacterium to another bacterium during conjugation. Eg. *Col E1 plasmid.*

On the basis of the functions, plasmids are classified into five types. They are-

i. F-Plasmid : F- plasmid or fertility plasmid contains some genes expressing the maleness in bacteria. The genes are known as ***'tra genes'.*** Eg. *F-plasmid of E.coli.*

If the supernatant has more than one type of plasmid, it is treated with ***ethydium bromide*** and ***centrifuged*** at high speed in ***cesium chloride*** density gradience. This results in separation of various plasmids according to their molecular weights. The plasmid is fractionated from appropriate layer and stored in TE buffer for cloning purposes. (Fig.7.3)

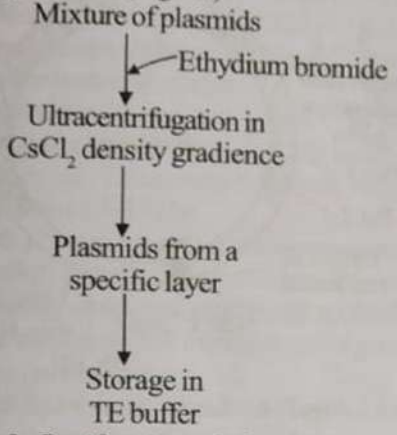


Fig.7.3: Purification of plasmid types.

Plasmid Cloning Vectors

The plasmid that carries the desired gene to the host cell is called a ***plasmid cloning vector.***

It is also known as ***cloning vector, cloning vechicle*** or ***carrier DNA***.

A. Chang and ***N. Cohen*** (1973) first proved the use of plasmids as gene cloning vectors.

They isolated plasmids from two different strains of bacteria. They are fused together using restriction enzyme and DNA ligase.

The joined DNA (chimeric plasmid) was then introduced into *E.coli* cells and its expression was studied. They found out that genes of both the plasmids expressed their traits in the *E.coli.*

Biotechnologists have adopted the following method for gene cloning:

1. A DNA segment is isolated from the genomic DNA using a restriction enzyme.

2. A suitable plasmid vector is isolated from a strain of bacterium.

3. The plasmid is cut with the same restriction enzyme used to isolate the desired DNA.

4. The source DNA is inserted into the cleaved plasmid DNA using DNA ligase. As a result, a ***chimeric DNA*** or ***rDNA*** is formed.

5. The ***rDNA*** is then introduced into a host bacterium that has to be manipulated. Here plasmid carries the inserted desired DNA to bacterium and being safe in it. So plasmid is known as a ***gene cloning vector*** or ***cloning vehicle.***

★ The plasmids ***pBR322, pBR324, pACYC177, pMB9, pRK646,*** etc. are useful to clone genes in *E.coli* cells.

★ ***pC194*** and ***pSAO501*** are gene cloning vectors for *Staphylococcus aureus.*

★ ***pBS161-I*** plasmid is suitable to clone genes in *Bacillus subtilis.*

★ pWWO plasmid is ideal to transfer genes to *Pseudomonas putida.*

Characteristics of Ideal Plasmid Vectors

The ideal plasmid vectors must have the following characteristic features:

1. Size : The plasmid must be small in size. The small size is helpful for easy uptake of chimeric DNA by host cells and for the isolation of plasmid without damage. The size of common plasmid vectors ranges from ***3kb*** to ***8.5 kb***. pRK646 plasmid is 3.4kb in size and pBR324 is about 8.3kb in size.

2. Copy number : The plasmid must be present in multiple copies.

3. Genetic markers : The plasmid must have one or a few genetic markers. These markers help us for the selection of organism that has recombinant DNA. ***Amp^r, Kan^r, Tet^r, Elimm, Ery^r, Str^r*** and ***LacZ*** genes are used as genetic markers in plasmid vectors.

4. Origin of Replication : The plasmid must have its own origin of replication and regulatory genes for self-replication.

5. Unique Restriction Sites : The plasmid must have unique restriction sites for common restriction enzymes in use. Eg. *EcoRI, SacI, KpnI, XmaI, SmaI, BamHI, XbaI, HindIII, BspMI, PstI and HindII.*

6. Insertional Inactivation : The plasmid must have unique sites for restriction en-

zymes in marker genes. This will help us for the selection of recombinants by insertional inactivation method. For example, ***LacZ gene*** marker has unique sites for KpnI, XmaI and Bam HI.

7. Pathogenicity : The plasmid should not have any pathogenic property.

Natural Plasmid Vectors

Some plasmids are isolated from bacteria and directly used for gene cloning without any modification. Such plasmids are known as ***natural plasmid vectors.*** Eg. *RP$_4$ plasmid of Pseudomonas, Col E1 of E.coli and YEp and YIp of yeasts.* Most of the natural plasmids cannot be used for gene cloning. This is because-

1. they are large in size.
2. they have no genetic markers.
3. they have no unique site for common restriction enzymes in the marker gene.
4. they confer pathogenicity to the host.

Artifical Plasmid Vector

The plasmid vectors created from wild plasmids are called ***artificial plasmid vectors.*** They are also called ***constructed vectors*** or ***based plasmid vectors*** or ***derived vectors.***

During construction some unwanted portions are removed from the wild type plasmid and desired sequences are insrted. Eg. *pBR322, RSF1010, pSC101, pUC8,* etc. The constructed plasmids are of much use in gene transfer experiments.

pBR322

pBR322 is an ***artificial plasmid.*** It was constructed by ***F.Bolivar*** and ***Rodriguez*** in 1977. It is a gene cloning vector for *E.coli.* It was constructed from two plasmids ***pSC101*** and ***Col EI*** and a transposon ***Tn3***. The copy number of this plasmid ranges from ***15-20*** per cell.

In the plasmid pBR322-

i. '***p***' indicates that it is a plasmid.

ii. ***BR*** indicates the names of workers ***F.Bolivar*** and ***Rodriguez,*** who created the plasmid.

iii. ***322*** is the specific number to distinguish the plasmid from others.

pBR322 is a ***circular, double-stranded plasmid DNA.*** It consists of ***4363 basepairs***. The plasmid has ***528 restriction sites*** for 66 restriction enzymes. Among these, 20 restriction enzymes cut it at unique restriction sites.

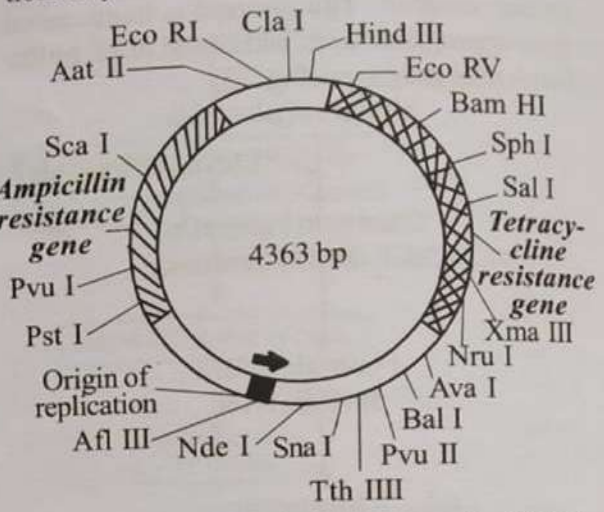


Fig.7.4: Restriction map of pBR322 showing unique restriction sites.

The pBR322 has two selectable gene markers-***tetracycline resistance gene*** (Tet^r) and ***ampicillin resistance gene*** (Amp^r). The ***Tet gene*** has unique sites for ***six restriction enzymes***. They are ***EcoR VI, BamHI, SphI, XmaI, SalI*** and ***NruI.*** If a gene is inserted into any of these restriction sites, the tetracycline resistance gene becomes inactive. Unique sites for Hind II and ClaI occur within the promoter of Tet^r gene.

The ***Amp^r gene*** has unique sites for ***three restriction enzymes PstI, PvuI*** and ***ScaI***. If a gene is inserted into any one of these unique sites, the ampicillin resistance gene becomes inactive.

The sequences other than Tet^r gene and Amp^r gene have ***unique sites*** for ***9 restriction enzymes***. There is ***no insertional inactivation*** when a gene is inserted into any one of these sites.

If a foreign DNA is cloned at Tet^r or Amp^r region of pBR322, the recombinant plasmid fails to confer resistance to the cell against antibiotic whose site was inactivated by that foreign DNA. This is the selectable property to screen for recombinant cells in a culture.

The ***ampicillin resistance*** determinant is the derivative of transposon ***Tn 3*** derived from ***RSF2124.*** The ***tetracycline resistance*** determinant is a derivative of ***pSC101*** derived from ***R-65 plasmid***. The remaining sequence is the derivative of a ***Col EI*** derivative ***pMBI.***

Advantages of pBR322 : 1. pBR322 is a ***small plasmid*** consisting of 4363 basepairs.

2. The copy number of pBR322 is ***15***. It can be increased upto 3000 by adding chloramphenicol to the bacterial culture.

3. Bacterial cells can uptake DNAs of 15 kbp size from the culture. But pBR322 is only 4.4 kbp in size. So it can carry relatively large DNA segments of ***5-10kbp***.

4. pBR322 has two selectable gene markers (Tet^r and Amp^r) for the selection of recombinants by insertional inactivation method.

5. The regulation and expression of a gene inserted into the plasmid is good.

6. pBR322 is used as a base plasmid for the in vitro construction of derived plasmid vectors such as pUC8, pUC9, pUC10, etc. and cosmids.

Uses : 1. pBR322 is being used to introduce desired genes into *E.coli* cells. Eg. *Somatostatin gene of man is introduced into E.coli through pBR322.*

2. Cloning vectors such as pBR327 and pBR325 are constructed from pBR322. Thus it serves as raw material for construction of other cloning vectors.

pUC Vectors

The plasmid vectors released from the ***University of California*** are termed ***pUC vectors;*** the letters ***UC*** stand for the name of that University. These vectors were constructed by ***Messings*** and his colleagues in 1983. They include a series of vectors such as pUC8, pUC9, pUC12, pUC13, pUC18, pUC19, etc. These are general purpose vectors that can clone foreign DNAs of less than ***10kb*** size in *E.coli*.

The pUC vectors consist of an ***origin of replication*** (ori), an ***ampicillin resistance sequence*** (Amp^r) and ***Lac Z′ gene.*** The ***ori sequence*** is taken from pBR322, which is a derivative of ***COLE1*** gene. ***Amp^r sequence*** is taken from ***pBR322*** plasmid. ***LacZ′ gene*** is the derivative of ***β-lactamase gene*** in the Lac operon of *E.coli*. The Lac Z′ gene is engineered in such a way that it should not be inactivated while inserting a ***multiple cloning site*** (MCS). The MCS is inserted at the coding region of β-lactamase gene. It provides ***multiple recognition sites*** for inserting foreign DNA. If anyone gene is inserted at the MCS, the β-lactamase gene becomes inactivated. Such recombinants will therefore form only ***white colonies***.

The general structure of all pUC vectors is the same, but they differ from each other in their multiple cloning sites (Box 7.2). The choice of restriction enzyme to be used to cut the plasmid is therefore determined by the MCS present in the vector.

pUC8

pUC8 is an ***artificial plasmid*** vector. It is a cloning vector constructed from ***pBR322***. *The letter 'p' stands for* ***plasmid***. *'UC' stands for* ***University of California***. It is ***2676 bp*** in size. It has an ***ampicillin resistance sequence*** (Amp^r), ***origin of replication*** from pBR322 and ***Lac Z′*** gene of *E.coli*. The Lac Z′ gene has a ***multiple cloning sequence*** (MCS) which has recognition sites for EcoRI, SmaI, Xma I, Bam HI, Sal I,Pst I, Acc I and Hind III.

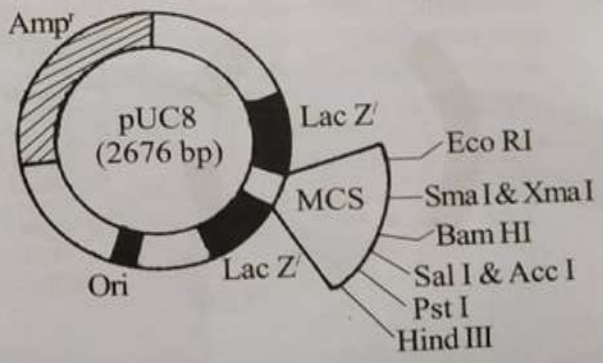


Fig.7.5: Structure of pUC8 vector.

Foreign gene is inserted into the MCS and it inactivates the Lac Z′ gene (See box 7.1). The ***recombinants*** are screened by growing the cells in a medium containing ***X-gal***. When a rDNA is introduced into a Lac Z⁻ mutant *E.coli*, the cell

Multiple Cloning Sites (MCS) Box.7.2

Multiple cloning site (MCS) is a *synthetic DNA segment that has a cluster of unique sites for restriction enzymes.* It is inserted into a gene cloning vector with a view to increasing the number of gene cloning sites.

The size of MCS usually varies from ***60 bp*** to ***84 bp***. The number and arrangement of restriction sites varies from MCS to MCS in different vectors. As MCS is a cluster of many restriction sites, it is also known as ***polylinker*** or ***polylinker sequence.***

Polylinkers are commercially available in the markets. The MCS is usually inserted into Lac Z′ gene in the vector. So, it is so designed that it does not affect the function of ***Lac Z′ gene.*** However, cloning of a gene at any one restriction site of the MCS inactivates the Lac Z′ gene. This property is useful for selection of the transformants by ***α-complementation.***

Uses : 1. Multiple cloning sequences are used to increase gene cloning sites in vector DNAs.

2. As they have unique sites for many restriction enzymes, DNA segments with different types of cut-ends can be inserted into the vector.

3. Restriction enzyme of choice can be used to insert a gene into the vector.

pUN121

pUN121 is an ***artificial plasmid*** vector.

Nilsson and colleagues constructed pUN121 plasmid vector from pBR322 in 1983. This plasmid is a ***positive selection vector*** to identify bacteria harbouring recombinant plasmid. It is used for constructing genomic libraries.

The pUN121 plasmid is ***4.4. kb*** in size. It consists of an ***ampicillin resistance gene*** (amp^r) and ***tetracycline resistance gene*** (tet^r) of pBR322 [illegible]. The original promoter of [illegible] replaced with ***p_L promoter*** of λ phage. The product of CI gene binds with P_L promoter and suppresses the expression of tet^r gene so that the pUN121 plasmid does not confer resistance to the cells against tetracycline. Bacteria harbouring pUN121 plasmid are, therefore, resistant to ampicillin but not to tetracycline in the medium.

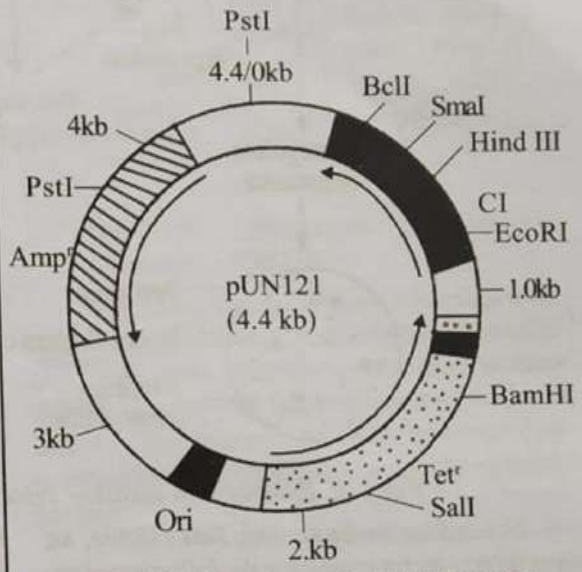


Fig.7.7: Structure of pUN121 vector.

The ***CI gene*** has unique sites for restriction enzymes such as ***EcoRI***, ***Hind III***, ***SmaI*** and ***Bcl I***. These sites are used to clone genes in the vector. If a foreign gene is inserted at ***CI gene***, the CI gene becomes inactivated and hence expression of suppressor of P_L promoter does not occur. Because of the absence of this suppressor, the P_L promoter does not work and the cells become resistant to tetracycline. Recombinant cells therefore grow into colonies in the medium containing tetracycline. Here, insertion of foreign gene enables tet^r gene to confer resistance to tetracycline; it can be used as a positive test for recombinants. So this vector is called ***positive selection vector*** (Fig.7.8).

Shuttle Vectors

The plasmid cloning vector that can exist and replicate in two different organisms is called ***shuttle vector.*** It has two origins of replication, each of which is specific to a host. Since

shuttle vectors replicate in two different hosts, they are often known as ***bifunctional vectors.*** Shuttle vectors can exist in two different hosts.

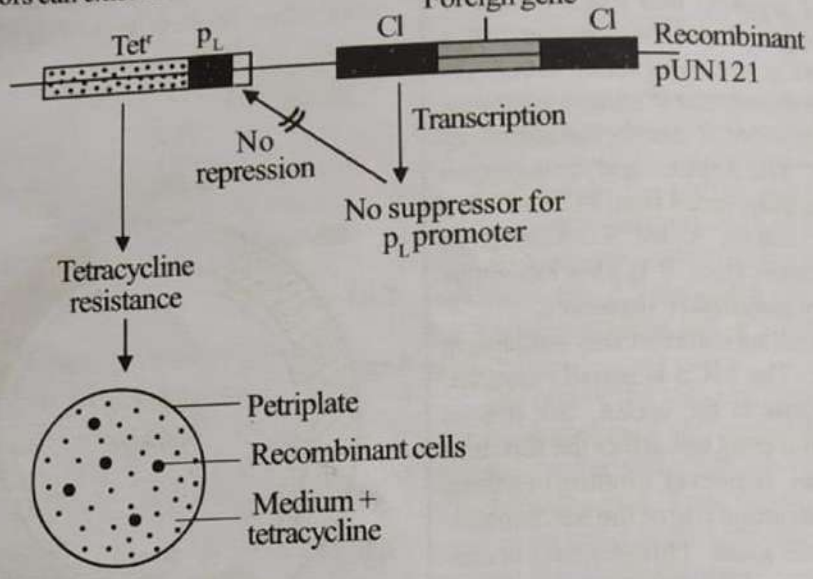


Fig.7.8: Principle of positive selection for recombinant pUN121.

According to ***Kado*** and ***Tait*** (1893), an ideal shuttle vector must have the following characteristics:

★ The vector must have origin of replication for two different hosts.

★ It must exist in the two specific hosts.

★ It must be small enough to carry long foreign DNAs.

★ It must have suitable genetic markers as selectable property.

★ It must be easily propagatable in vivo.

★ Expression of cloned gene must be easily detectable.

★ The vector must be non-pathogenic and should not induce any stress in the host.

The shuttle vectors pHV14, pEB10, pHP3, etc. replicate both in *Bacillus subtilis* and *E.coli.*

pJDB219 is another shuttle vector that can replicate in ***E.coli*** and ***yeast*** *(Saccharomyces cerevisiae).*

The initial cloning and in vivo amplification of rDNA (shuttle vector having foreign DNA) are carried out in the first host. To harvest the product of expression, the amplified rDNA is introduced into the second host.

pEB10

pEB10 is an ***artificial plasmid*** vector. It is a circular, double-stranded plasmid. It is a ***shuttle vector*** replicating in *B.subtilis* and *E.coli.* It is constructed by the combination of pBR 322 and pUB 110. It is ***8.9 kbp*** in size. It has two selectable markers-***ampicillin resistance gene*** (Amp^r) and ***kanamycin resistance gene*** (Kan^r).

The ***ampicillin resistance*** determinant is derived from ***pBR322***. It has an origin of replication for replication of the plasmid in *E.coli.*

The ***kanamycin resistance*** determinant is a derivative of ***pUB110***. It has yet another origin of replication that switches on replication in *Bacillus subtilis.*

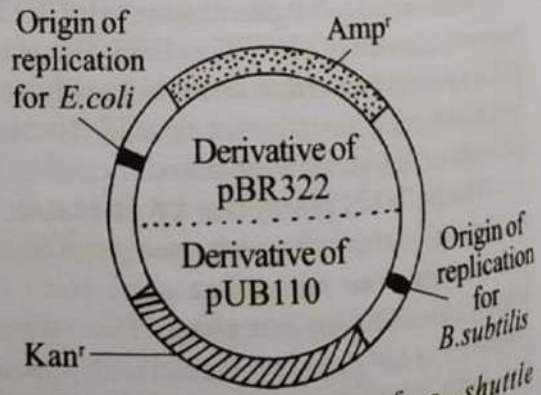


Fig.7.9: Structure of a shuttle vector, pEB10.

As the pEB10 has two different origins of replication, it can replicate both in *E.coli* and *B.subtilis*. The desired gene is first inserted into pEB10 to construct rDNA. The rDNA is introduced into *E.coli* cell by ***bacterial transformation***. *E.coli* cells are naturally resistant to chloramphenicol that inhibits protein synthesis. But *B. subtilis* is sensitive to chloramphenicol.

The recombinant *E.coli* culture is treated with chloramphenicol to increase the copy number of pEB10 in the *E.coli*. The amplified rDNA is isolated, purified and introduced into *Bacillus subtilis* by transformation. The transformants are screened and mass cultured for the expression of the cloned gene.

pJDB219

The pJDB219 is a ***shuttle vector***. It shuttles an inserted DNA between *E.coli* cell and yeast cell. It consists of the ***entire sequence*** of a ***pBR322,*** a ***2 μm plasmid*** and a selectable ***marker gene LEU2*** of yeast chromosome. It is an ***artificial plasmid*** vector.

The pBR322 derivative encodes for ***ampicillin resistance*** (Amp^r) and ***tetracycline resistance*** (Tet^r). It has an ***origin of replication*** essential for replication in *E.coli*.

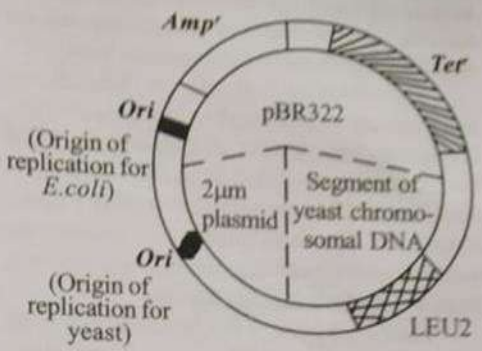


Fig.7.10: Structure of shuttle vector, pJDB219.

The ***2μm*** plasmid of yeast contributes an ***origin of replication*** of the plasmid in yeast cells. The LEU2 gene encodes for ***isopropyl malate dehydrogenase*** that converts pyruvic acid into leucine. The leucine can easily be assayed by growing the transformants in a medium lacking leucine.

The foreign DNA segment is inserted into pJDB219 to construct an rDNA. The rDNA is

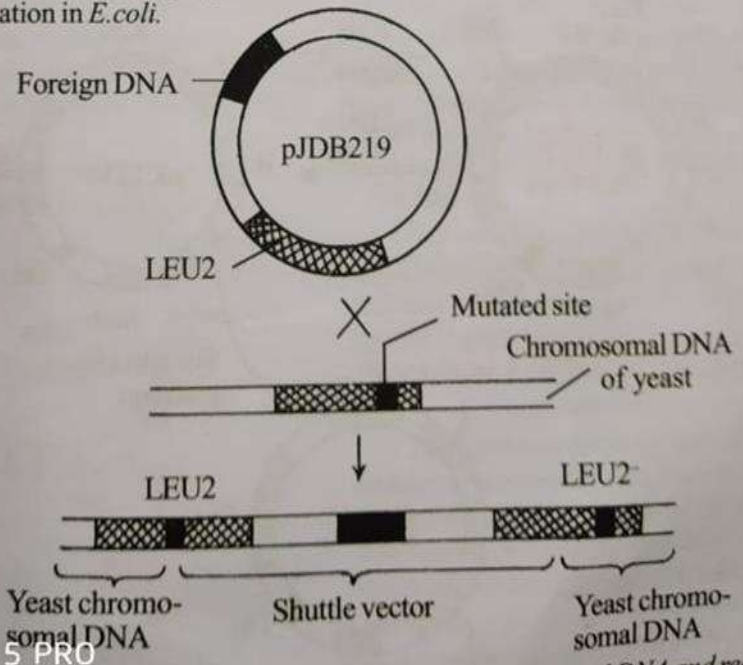


Fig.7.11: Homologous recombination between yeast chromosomal DNA and recombinant shuttle vector pJDB219.

8

Bacteriophage Vectors

The bacteriophage vector is a bacterial virus that carries a desired gene to a host cell.

Viral DNAs are linear or circular and single stranded or double stranded DNAs. DNAs of λ phage and M13 phage are used to carry genes to bacteria.

Vectors designed from ***M13 phage, λ phage, P1 phage*** and ***ϕ105 phage*** have been used for gene cloning in bacteria.

Natural Mechanism of Gene Transfer by Phages

Temperate phages are capable of integrating their genome with the host genome and are transmitted through cell divisions without causing host cell lysis. When they regain the ability of virus replication, the viral DNA picks up certain host genes along with it and integrates them in another host cell during reinfection. The ability of phages to transfer certain genes from one host to another host cell in subsequent infections, attracted scientists to develop phage-derived vectors.

I. M13 Bacteriophage Vectors

M13 is a filamentous bacteriophage that infects *E.coli*. It is a ***DNA virus***. It is closely allied to coliphages such as ***fd*** and ***f1***. It is 895 nm long and 6 nm in diameter.

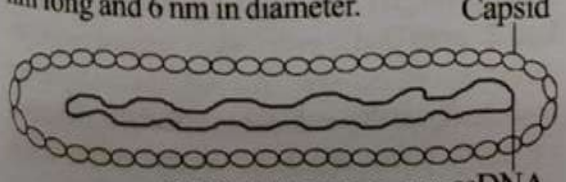


Fig.8.1: M13-phage: Structure of M13 phage.

It consists of an outer protein coat called ***capsid*** and a DNA residing in the capsid. The DNA is ***circular*** and ***single stranded***. It is made up of ***6408 bases***.

M13 Life Cycle

M13 phage is ***male-specific*** and infects *E.coli* cells bearing ***F pilli***. The phage first attaches to an F pilus of *E.coli* and injects its DNA into the cell. Inside the *E.coli*, the phage DNA strand (+) serves as a template for the synthesis of a complementary strand (-). It becomes a double stranded DNA (dsDNA). As the ds DNA is the product of replication, it is called ***replicative form*** (RF).

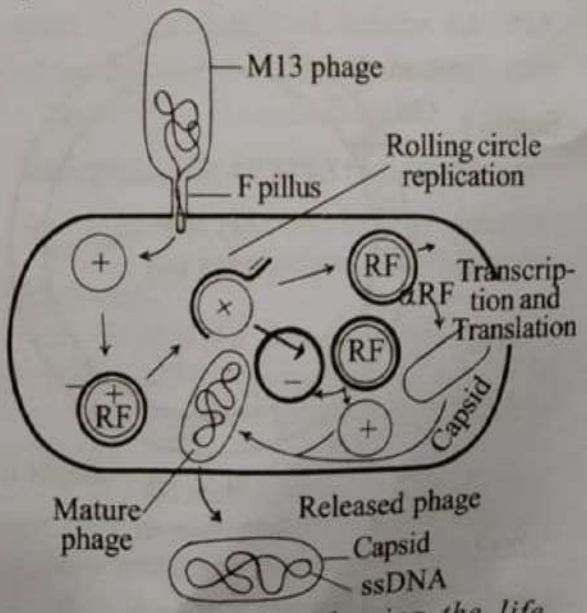


Fig.8.2:Diagram showing the life cycle of M13 phage.

II. Lambda Phage Vectors

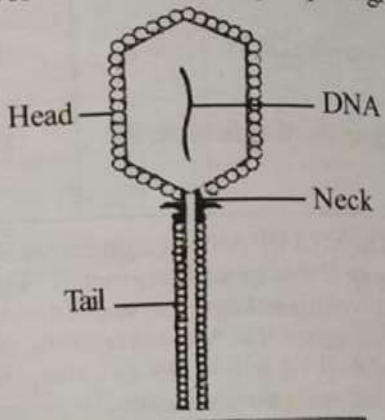
Lambda phage (λ phage) is a bacterial virus that infects the *E.coli*. It is 57 x 10^6 daltons in molecular weight. The phage consists of an icosahedral ***head*** and a ***flexible tail*** lacking contractile sheath. The phage DNA is packed inside the head (capsid). It is capable of integrating into genomic DNA of host cell and transmitted through cell division. Therefore, it is known as ***temperate phage***.

Lambda DNA

The DNA of λ phage is a linear duplex DNA with cohesive single-stranded extensions. It consists of 48,502 basepairs and the molecular weight is 32 x 10^6 daltons.

In the duplex region, more number of CG pairs are found at the left side than the right side. AT pairs are more at the right side. The λ DNA has 35-57% CG and 43-65% AT pairs.

The single stranded extensions of λ DNA are complementary to each other and consist of 12 nucleotides. They are known as ***complementary sites*** or ***cos ends***. The base sequence of one 5′ extension is GGGCGGCGACCT and of other 5′ extension is CCCGCCGCTGGA. The cos site is the significant feature of λ DNA that helps for circularization of the DNA and packaging of linear DNA inside the viral capsid. The free end of the cos site has a 5′-phosphate group.



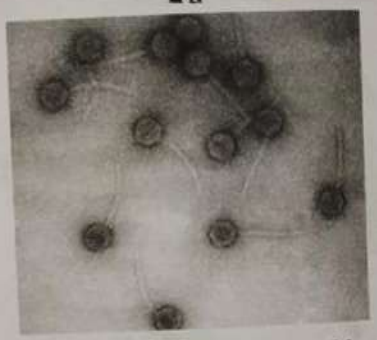


Fig.8.9: Electron micrograph of λ -phage.

Inside the host cell, the linear λ DNA becomes circular due to complementary basepairing between the cohesive ends.

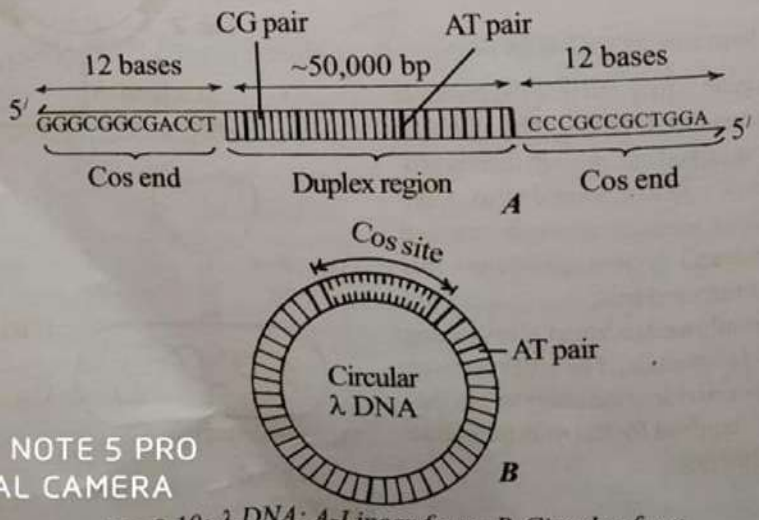


Fig.8.10: λ DNA: A-Linear form; B-Circular form.

Gene Map of λ DNA

The λ DNA is 48,502 basepairs in size. Nearly, one third of λ DNA is non-essential and does not involve in virus replication. ***Essential region*** has ***37 genes*** in definite clusters. It has a cluster of ***five genes*** (W, B, C, D and E) for ***head*** protein synthesis, a cluster of ***12 genes*** (F, Z, U, V, G, H, M, L, K, I and J) for ***tail*** synthesis and a ***gene*** (A) for assembly (maturation) of phages. All these 18 genes occur between ***1*** and ***19,400 basepairs***.

There is a ***non-essential sequence*** between ***19,400*** bp and ***27,975*** bp. This sequence is useful for gene cloning.

Following the non-essential sequence there are ***16 regulatory genes***. They involve in ***recombination, integration, DNA synthesis, regulation of DNA synthesis*** and ***regulation of lysis*** of host cells.

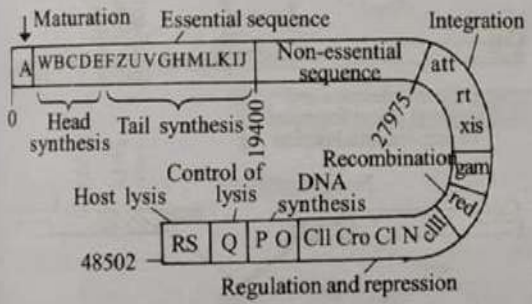


Fig.8.11: Gene map of λ DNA.

Life Cycle of λ Phage

The λ phage is a ***temperate*** bacteriophage that infects *E.coli*. It has both ***lytic*** and ***lysogenic*** life cycles, the latter is more common in λ phage. The frequency of lytic life cycle is increased by treating the culture with ***mitomycin*** or exposing it to ***X-rays.***

1. Lytic cycle

All the T series phages exhibit lytic cycle. These viruses infect bacteria and on completion of life cycle, they cause lysis or rupture of the bacterium. Hence the cycle is called ***lytic cycle***. Lytic cycle involves the following steps:-

1. Infection takes place by random collision between the phage and the bacterium.
2. The tail fibres select specific site on the surface of the bacterium.
3. The spikes anchor firmly.
4. A hole is made in the cell wall of bacterium by the enzyme lysozyme secreted by the tail.
5. The sheath contracts and this causes the tail core to penetrate the cell wall.
6. The DNA of the head is discharged into the cell through the tail core.
7. The protein coat does not enter in; it remains outside the cell.
8. The DNA of the phage takes over the protein synthesizing machinery of the bacterium.
9. The host DNA is degraded by the viral ***deoxyribonuclease.***

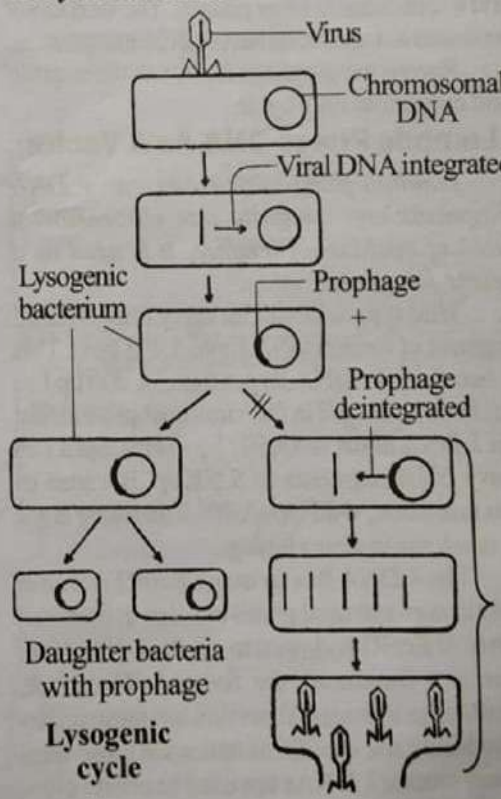


Fig.8.12:Life cycle of λ phage.

10. Viral DNA is transcribed into mRNA.
11. The mRNA translates proteins which are utilized for the synthesis of viral proteins.
12. Viral DNA also replicates to produce hundreds of copies.
13. Each DNA copy is enclosed by a protein coat producing new virus.

14. Soon the bacterial wall ruptures and the new viral particles are released.

15. A single phage can produce about 200 phages in twenty minutes after infection.

2. Lysogenic cycle

All the lambda series bacteriophages exhibit lysogenic cycle. In lysogenic cycle the bacterium is not lysed. As in the lytic cycle the viral DNA enters the bacterial cell. But the viral DNA does not take over the protein synthesizing machinery of the host. Instead it gets integrated with the bacterial chromosome. At this stage,the virus is called ***prophage.*** The bacterium with prophage is called ***lysogenic*** and it is resistant to other phages. The viral DNA replicates along with the bacterial chromosome.

Rarely, the prophage DNA can deintegrate and enters into lytic cycle.

Lambda Phage DNA As A Vector

Lambda phage DNA can carry DNA fragments and integrate into chromosomal DNA of host cell. Therefore, it is used as a ***vector*** *for gene cloning.*

Wild type λ DNA can carry only a small segment of foreign DNA (upto 5,498 bps). This is because, the DNA larger than 54,000 bp has not been packaged in the virus coat protein, but the λ DNA alone is 48,502 bp in size. So it can carry DNA segments of 5.5 Kbp. Because of this limitation, wild type λ DNA therefore is not in much use in gene cloning.

The λ DNA has so many EcoRI sites not only in non-essential genes but also in essential genes. If EcoRI is chosen to cut the DNA, more than two fragments are formed. Therefore, EcoRI sites in essential portion are removed by introducing site specific mutations at those sites. These mutant λ DNAs are used to create cloning vectors.

To make efficient λ cloning vectors, a segment of non-essential sequence is deleted from the wild type λ DNA and the two cut arms are joined together. The resulting $2\overline{21}$ DNAs can carry relatively large DNA fragments. Usually mutants of λ phages are used to construct cloning vehicles. The constructed λ DNA vectors are of two types:

i. Insertion vectors.

ii. Replacement vectors.

1. Insertion Vectors

Vector DNA constructed by removing a small segment of non-essential region and inserting desired DNA is called ***insertion vector.*** The λ insertion vector is designed so as to have a unique restriction site for gene cloning and the size of not less than 35kb. To construct a ***λ-insertion vector***, a small region is cut out from the wild type λ DNA and then restriction sites are removed through point mutations. For gene cloning, the unique site found in the non-essential region is used. A foreign DNA of about ***18 kbp*** can be cloned in λ-insertion vectors. Eg. λ *charon 16, λ gt 10,* etc.

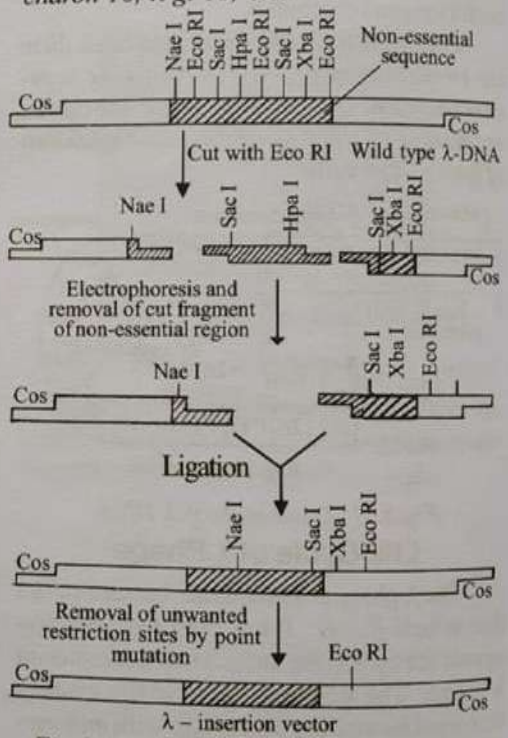


Fig.8.13: In vitro construction of λ-insertion vector.

The size of the insertion vector depends on the size of non-essential sequence deleted from the λ DNA. The restriction enzyme that cuts at its non-essential region is used to insert the foreign gene.

a. λgt 10 vector : This is a λ-insertional vector with the size of ***43kb***. It is constructed from wild type λ-DNA by removing a small EcoRI fragment from its non-essential sequence. The remaining portion of the non-essential se-

quence has a ***cI*** gene that encodes for a protein repressing ***cII*** gene, which takes part in lytic growth of λphage. The ***cII*** gene occurs in the right essential region of the λ-DNA. The ***cI*** gene has a unique site for EcoRI for gene cloning. Foreign DNA of 7kb size can be cloned in λgt 10 vector.

When a foreign DNA is inserted at EcoRI site of λgt 10, the cI gene becomes inactivated and hence the cI^+ phage becomes cI^- phage. To construct recombinant λgt 10, the DNA is cut with EcoRI and then the foreign DNA is inserted

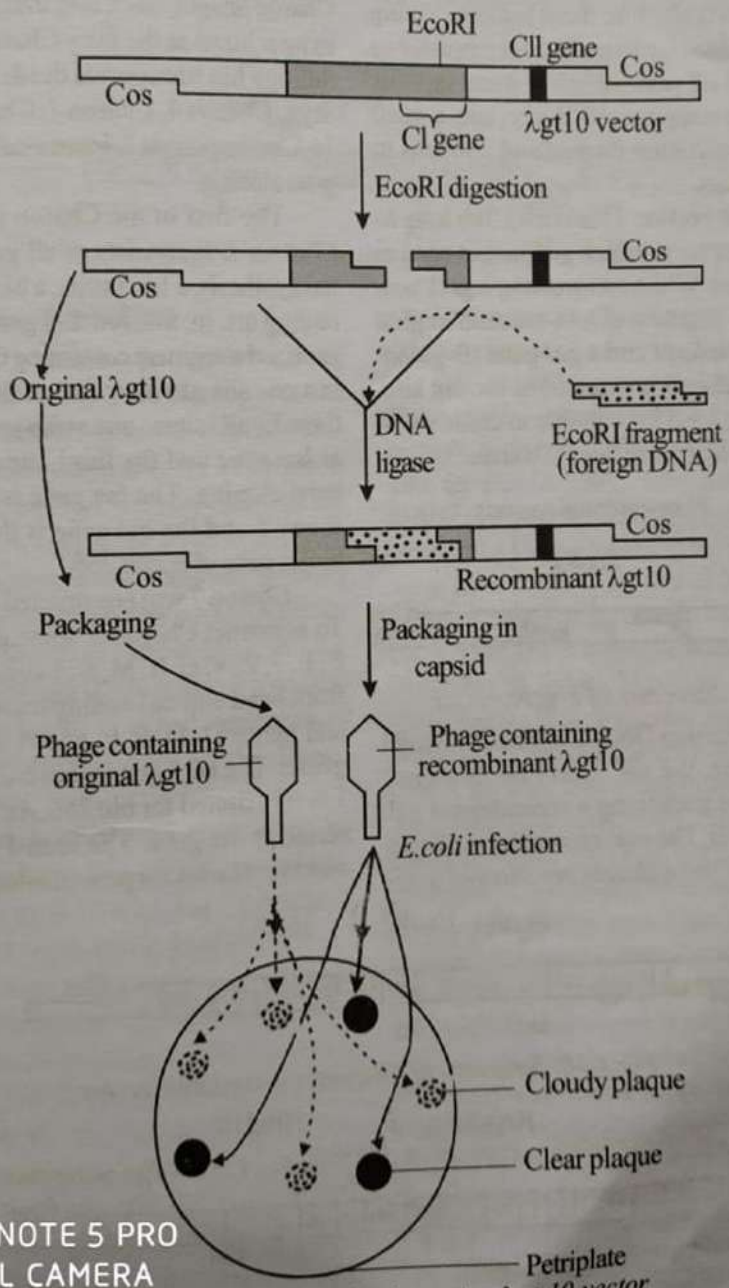


Fig.8.14: Insertional cloning using λgt 10 vector.

at that EcoRI cut site. The recombinant λgt10 is packaged in virus coat protein to form recombinant phages. These phages are allowed to infect *E.coli* in a petriplate. The phages infect *E.coli* cells, multiply in them and form definite plaques in the plate. Bacteria harbouring recombinant λgt10 form ***clear plaques*** due to lack of cl repressor protein. But, those bacteria having non-recombinant λgt10 produce cl repressor as usual so that ***cII*** gene which is necessary for lytic life cycle is suppressed. Hence, only a small number of phages are formed and it results in ***cloudy plaques.***

b. λgt 11 vector: This is a 43.7kb long λ-insertional vector in which ***gal*** gene is used as genetic marker. While constructing λ gt11 vector, a lengthy segment of non-essential region is cut out with ***SacI*** and a ***gal*** gene (β-galactosidase) of *E.coli* is inserted at the cut site. The resulting λ gt11 is suitable to clone small cDNAs which are less than 6kb size.

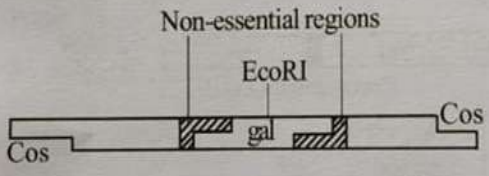


Fig.8.15: Structure of λ gt11 vector.

When a foreign DNA is inserted at EcoRI site of gal gene, the gal⁺ λ gt11 becomes gal⁻ λgt11; so after packaging a recombinant gal⁻ phage is formed. The non-recombinant phages remain gal⁺. If these phages are allowed to infect *E.coli* culture in a petriplate containing X-gal, the recombinant phages form ***white plaques*** while non-recombinant phages form ***blue plaques.***

c. Charon phages: In 1977, ***Blattner*** and colleagues constructed a series of λ-insertional vectors called ***charon phages.*** They are named Charon phages, since they transfer foreign genes to new hosts as the fairy Charon of Greek mythology has transported deads across the River Styx. Charon 4, Charon 7, Charon 11, Charon 16A are important λ-insertional vectors used for gene cloning.

The first of the Charon series vectors is ***Charon 4.*** It consists of all genes for head to tail synthesis, a lac 5 gene, a bio 256, a segment coding att, in, xis, red and gam genes, a nin 5 gene and a segment containing QSR gene. There is a cos site at either end of the vector. It has three EcoRI sites - one at ***bio*** gene, another one at ***lac*** gene and the third one at ***red*** gene for gene cloning. The lac gene is a derivative of *E.coli* 5 and the bio gene is the derivative of biotin gene of *E.coli* 256.

Charon 7 was constructed from Charon 4. To construct Charon 7, some genes (A, W, B, C, E, F, Z, V, G, H, M, K, I and J) are removed from head and tail coding region of Charon 4 and genes D, U, P, L, and N are brought together. The lac 5 gene is cut out from it and bio 1 is substituted for bio 256. An imm 434 is inserted at ***Xis*** gene. The imm 434 has Hind III and EcoRI sites for gene cloning.

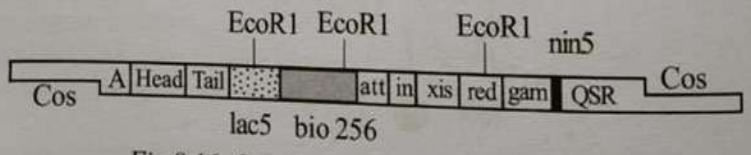


Fig.8.16: Structure of Charon 4 insertional vector.

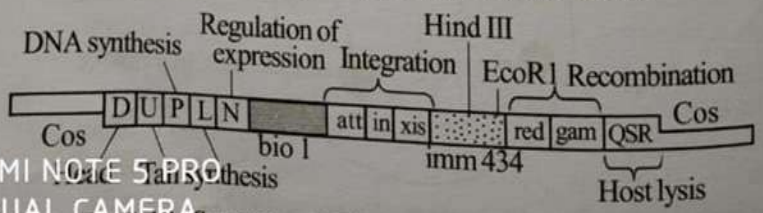


Fig.8.17: Structure of Charon 7 insertional vector.

Charon 12 consists of all genes for head and tail synthesis, LacZ gene of *E.coli*, a segment coding att, in, xis, red and gam genes, imm 80 gene of *E.coli*, nin 5 gene and a segment containing QSR genes. The LacZ gene has EcoRI site for gene cloning. So, the recombinants can be selected by insertional inactivation of lacZ gene.

DNAs, these vectors have been used to construct ***gene banks*** of eukaryotic organisms such as yeast, rat, mouse, *Drosophila* and man. These vectors have been created by suitable modifying restriction sites of λ DNA.

Wild type λ DNA has several restriction sites for many restriction enzymes in common

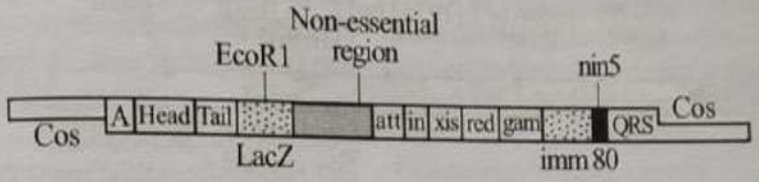


Fig.8.18: Structure of Charon 12.

2. Replacement Vectors

A vector in which one portion is replaced by a foreign gene during gene cloning is called ***replacement vector*** or ***substitution vector.*** This vector has two restriction sites for a restriction enzyme at both ends of the non-essential sequence. The DNA segment lying between the two restriction sites can be removed by using the restriction enzyme. A foreign DNA can be incorporated at the site of removed λ DNA fragment. Since a foreign DNA replaces a portion of λ DNA, the vector is called ***replacement vector.*** The replaced portion of λ DNA is known as ***stuffer region.***

Foreign DNA of 20 – 25 Kb size can be cloned in these vectors. Eg. *λ WES. λ3 can carry DNA segments upto 15 Kbp and λ pEMBL 4 can carry DNA segments of 23 Kbp size.*

There are some other λ-replacement vectors in which foreign DNA of 44 kb size can be cloned. Charon 34, Charon 35, Charon 40, λ DASH, λ L47, λ1059, etc. are examples for such vectors. In these vectors, the stuffer region constitutes a large segment of λDNA coding for head and tail synthesis and non-essential sequence. That is why they pick up large foreign DNAs. Here, head and tail proteins of λ phage [illegible] packaging of the recombinant λ DNA. Since they can pick up large

use. To construct a replacement vector, all restriction sites of an enzyme, except one on either side of stuffer region to be replaced, are inactivated by introducing point mutations. For example, wild type λ DNA has seven restriction sites for ***BanII.*** These sites occur at 581bp, 10,086 bp, 19,763 bp, 21,570 bp, 24,772 bp, 25,877 bp and 39,453 bp. Of these sites, 581bp site occur at head coding region and 25,877 bp site occurs at right end of non-essential sequence. Site at 39,453 bp is in the essential region. Hence 10,086 bp, 19,763 bp, 21,570 bp, 24,772 bp and 39453 bp recognition sites are inactivated by introducing point mutations. Now, 581 bp and 25,877 bp sites are available for BanII digestion. When this modified λ DNA is cut with BanII, a stuffer region of head, tail and non-essential sequences is formed and it can be replaced with a foreign gene. Therefore, it is known as a ***replacement vector.*** (Fig.8.19)

To insert a foreign gene into the λ-replacement vector, the vector is cut with BanII restriction enzyme and electrophoresed. The left arm of λ-vector is very small and the right arm is very long, but the stuffer region is in between the two. So, the λ DNAs found in the first and last bands are extracted from the gel and mixed with foreign gene obtained by BanII digestion. These DNA fragments are joined with DNA ligase to construct rDNAs. The rDNAs are

packaged in virus capsid to create recombinant phages which are then allowed to infect *E.coli* culture in a petriplate. After infection, the *E.coli* culture forms clear plaques of recombinant cells. If the two arms of λ-DNA are joined together, the size of the DNA construct is less than 3,700 bp, so it cannot be packed in virus capsid. So there is no need for selection of recombinant cells. The plaque forming colonies are directly used for further research and development.

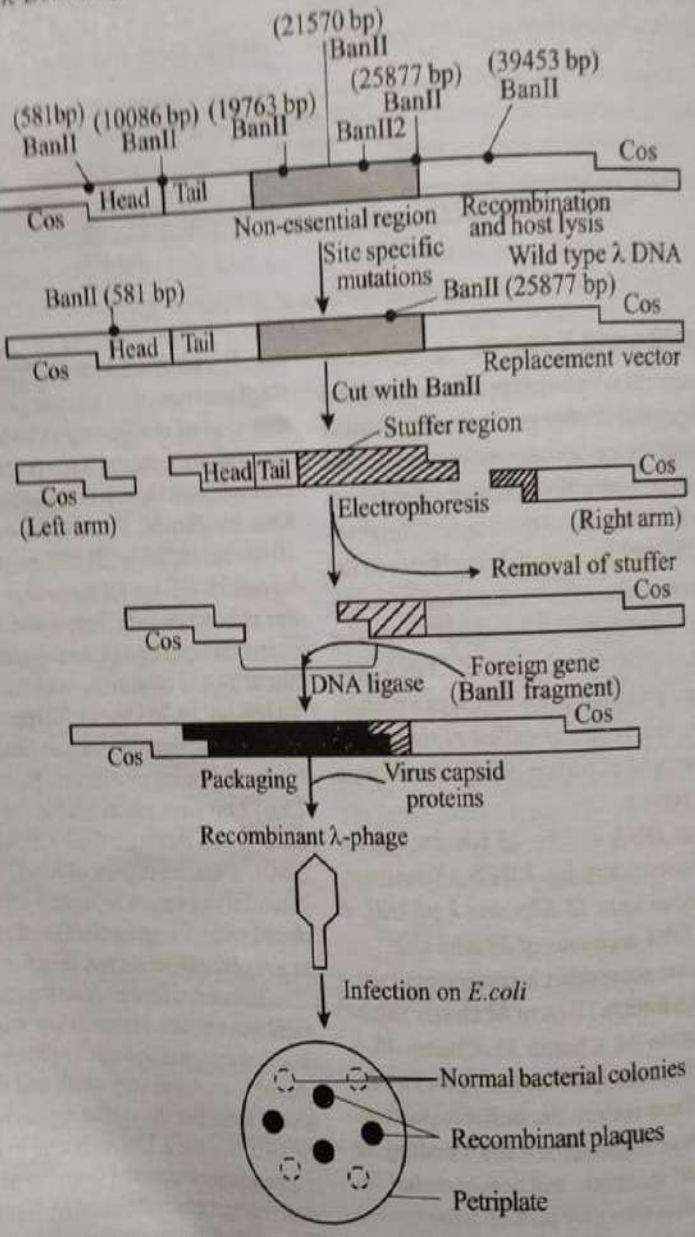


Fig.8.19: Gene cloning through λ-replacement vector.

Limitations of λ Vectors

1. For successful packaging of λ DNA in capsid, the two cos-sites should be separated by distance ranging from 37000 to 54,000 basepairs. If the distance is lower than 37,000 bp or higher than 54,000 bp, the packaging of DNA does not take place.

2. Sometimes, many λ DNAs join together by base pairing between their cos-sites and form relatively long DNA. Here, packaging of DNA does not take place successfully.

3. The λ DNA has no easily selectable marker gene. So transformants should be identified by DNA hybridization method.

4. The recombinant 2|21 DNA may enter the lytic life cycle.

5. The λ phage has narrow host-range.

Advantages : 1. The efficiency of gene transfer through λ phage is high.

2. Foreign DNAs upto 23 kbp can be packed in virus capsid and transduced to *E.coli.*

III. Cosmids

Cosmid** is an artificial plasmid containing cos-sites of λ DNA.* It is formed by joining ends of a linearized plasmid DNA with cos-sites of a λ DNA. It is a ***derived vector. The cosmid can be packaged in capsid of λ phage in vitro to form recombinant phage particles. It is ***linear*** inside the phage capsid. The cosmid gets circularized and behaves like a plasmid. Foreign DNA upto 45 kb size can be cloned in cosmid vectors.

Cosmid has an origin of replication, selectable markers and gene cloning sites of the plasmid DNA. They lack structural and regulatory genes of λ DNA. Hence there is no lysis and integration of cosmid DNA in the host cell. Eg. *Col EI cosmid, pHC 79, pJB8, pWE cosmid,* etc.

The fact that small segment of λ phage containing cos sites on a plasmid is enough for packaging of the DNA in virus capsid, was first discovered by a team of Japanese research workers in 1976. Cosmid was first constructed by ***Collins*** and ***Hohn*** in 1978.

Salient Features of Cosmids

1. Cosmid is a circular, double-stranded DNA.

2. It has two complementary single-stranded regions at both ends of a plasmid DNA. The two cos-ends form a duplex by base pairing.

3. At the cos-site, 3′ end and 5′ end of the same chain do not establish covalent bond during circularization. So a definite nick is present in each DNA strand.

4. The nicks are retained in the plasmid for a number of generations.

5. The cosmid DNA does not code for phage proteins and host cell lysis.

6. It does not involve in multiplication of phage particles.

7. It has an origin of replication from plasmid DNA for independent replication.

8. It has selectable marker genes and gene cloning sites of plasmid DNA.

9. The cosmid DNA is packaged within protein coat of bacteriophage to form infective phage particles. Cos-site is a prerequisite for invitro packaging of cosmid in phage protein coat.

10. After infection, the cosmid DNA does not integrate into chromosomal DNA of host cell. It exists as a definite extra chromosomal DNA and replicates independently.

Cosmid pLFR5

pLFR5 is the commonly used cloning vector suitable for cloning large DNA fragments upto 45 kbp. It is 6 kbp in size. It is constructed from *E.coli* plasmid pBR322 and two cos-ends of λ DNA.

The plasmid derived portion contributes an ***origin of replication*** (Ori) and ***tetracycline resistance gene*** (Tet^r). There is a ***multiple cloning site*** (MCS) between the Tet^r and cos-site.

A foreign DNA of upto 45 kbp is inserted into the MCS of pLFR5 and the rDNA is packaged into bacteriophage λ head in vitro. The phage thus formed delivers the DNA into *E.coli* while infecting the cell.

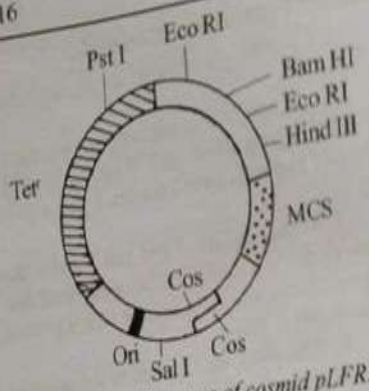


Fig.8.20: Structure of cosmid pLFR5.

Cosmid pJB8

pJB8 is constructed from the plasmid pBR322 and cos ends of λ DNA. It is 5.4 kbp in size. It has an ***origin of replication*** (Ori) and ***ampicillin resistance gene*** derived from pBR322 and two ***cos-ends*** from λ DNA.

A foreign DNA of about 45 kbp is inserted into Bam HI restriction site of the cosmid. The recombinant cosmid is packaged into λ phage head to form an infective phage particle. The phage delivers its DNA (rDNA) into *E.coli* while infecting the cell.

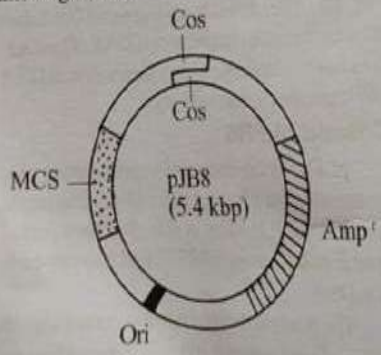


Fig.8.21: Structure of cosmid pJB8.

Cosmid pHC 79

pHC79 is constructed from pBR322 and cos-sites of λ DNA. It is 6.5 kbp in size. It can carry DNA fragments upto 40 kbp. The derivative of pBR322 has an ***origin of replication*** (Ori) and two marker genes- ***Ampicillin resistance gene*** (Amp^r) and ***Tetracycline resistance gene*** (Tet^r). The derivative of λ DNA has two cos-ends essential for in vitro packaging of the cosmid into λ phage head.

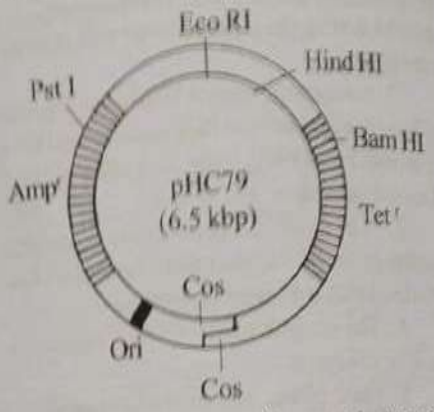


Fig.8.22: Structure of cosmid pHC79.

Gene Cloning Using Cosmids

The following are the essential steps in the construction of cosmid library of an organism, say for example *E.coli:*

1. After selecting a suitable cosmid vector, it is cut at MCS with a suitable restriction enzyme (eg. *BamHI*). Now cosmid is in linear form.

2. The cut cosmid is treated with phosphatase. This enzyme removes 5′ phosphate group from cut fragments so that polycosmid formation does not take place during ligation.

3. Foreign DNA is cut with BamH1 and electrophoresed to separate the DNA fragments based on their size.

4. DNA is reisolated from each of these DNA bands in the gel and mixed with linearized cosmid vector. Thus all DNA fragments of *E.coli* are mixed separately with cosmid.

5. The genomic DNA fragment joins with cosmid vector to form recombinant cosmid. It is in linear form. Thus concatemers of recombinant cosmids are created *in vitro*.

6. Each concatemer of recombinant cosmid is then treated with virus capsid for packaging

of recombinant cosmid in that capsid. As a result of packaging, infective phages containing recombinant cosmids are formed.

7. The phages are allowed to infect *E.coli* cells in a culture to deliver the recombinant DNA into *E.COLI* cells. Thus recombinant bacteria are formed.

Inside the bacterial cell, the linear recombinant DNA gets circularized by complementary basepairing between cos sites and remains as a plasmid. The set of bacterial colonies containing all segments of *E.coli* genome is called ***E.coli cosmid library.*** As cosmids can pick up 45kb long foreign DNAs, only 120 cosmid clones are enough to accommodate all DNA of *E.coli.*

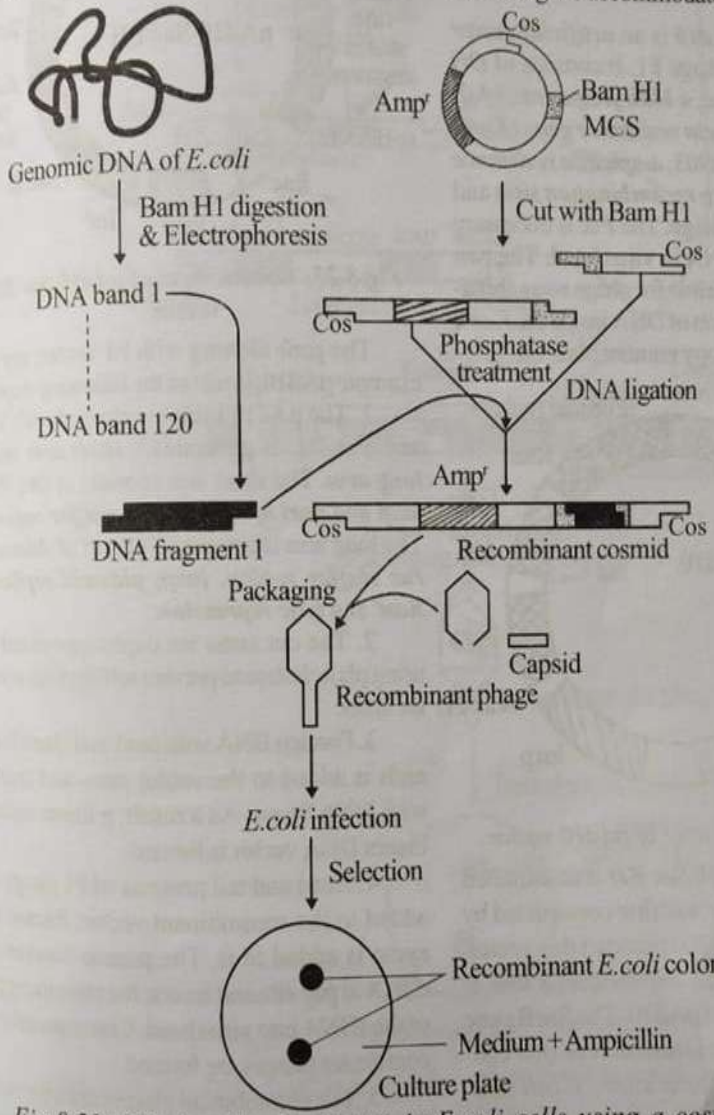


Fig.8.23: Cloning E.coli genome in E.coli cells using a cosmid vector. A clone contains a segment of E.coli genome. All 120 fragments preserved in the form of recombinant DNA in separate clone of bacteria constitute an E.coli cosmid library.

Restriction enzymes

A restriction enzyme is a protein that recognizes a specific, short nucleotide sequence and cuts the DNA only at that specific site, which is known as restriction site or target sequence.

- Most restriction enzymes are specific to a single restriction site
- Restriction sites are recognized no matter where the DNA came from
- The number of cuts in an organism's DNA made by a particular restriction enzyme is determined by the number of restriction sites specific to that enzyme in that organism's DNA.
- A fragment of DNA produced by a pair of adjacent cuts is called a RESTRICTION FRAGMENT.
- A particular restriction enzyme will typically cut an organism's DNA in to many pieces, from several thousand to more than a million!
- There is a great deal of variation in restriction sites even within a species.
- Although these variations do not have phenotypic expression beyond the base sequences themselves, the variants can be considered molecular "alleles," and they can be detected with sequencing techniques.
- As such, they can be used in mapping studies similar to the way true genes with known phenotypic effects can be used, but skipping the breeding steps and going straight to the molecules.
- These "molecular alleles" are a type of MOLECULAR MARKER, as they can be detected and located with labeled probes.

Types:

Naturally occurring restriction endonucleases are categorized into four groups (Types I, II III, and IV) based on their composition and enzyme cofactor requirements, the nature of their target sequence, and the position of their DNA cleavage site relative to the target sequence.DNA

sequence analyses of restriction enzymes however show great variations, indicating that there are more than four types. All types of enzymes recognize specific short DNA sequences and carry out the endonucleolytic cleavage of DNA to give specific fragments with terminal 5'-phosphates. They differ in their recognition sequence, subunit composition, cleavage position, and cofactor requirements,as summarised below:

Type I enzymes (EC 3.1.21.3) cleave at sites remote from a recognition site; require both ATP and S-adenosyl-L-methionine to function; multifunctional protein with both restriction digestion and methylase (EC 2.1.1.72) activities.

Type II enzymes (EC 3.1.21.4) cleave within or at short specific distances from a recognition site; most require magnesium; single function (restriction digestion) enzymes independent of methylase.

Type III enzymes (EC 3.1.21.5) cleave at sites a short distance from a recognition site; require ATP (but do not hydrolyse it); S-adenosyl-L-methionine stimulates the reaction but is not required; exist as part of a complex with a modification methylase (EC 2.1.1.72).

Type IV enzymes target modified DNA, e.g. methylated, hydroxymethylated and glucosyl-hydroxymethylated DNA

Type l

Type I restriction enzymes were the first to be identified and were first identified in two different strains (K-12 and B) of E. coli. These enzymes cut at a site that differs, and is a random distance (at least 1000 bp) away, from their recognition site. Cleavage at these random sites follows a process of DNA translocation, which shows that these enzymes are also molecular motors. The recognition site is asymmetrical and is composed of two specific portions—one containing 3–4 nucleotides, and another containing 4–5 nucleotides—separated by a non-specific spacer of about 6–8 nucleotides. These enzymes are multifunctional and are capable of both restriction digestion and modification activities, depending upon the methylation status of the target DNA. The cofactors S-Adenosyl methionine (AdoMet), hydrolyzed adenosine triphosphate (ATP), and magnesium (Mg^{2+}) ions, are required for their full activity. Type I restriction enzymes possess

three subunits called HsdR, HsdM, and HsdS; HsdR is required for restriction digestion; HsdM is necessary for adding methyl groups to host DNA (methyltransferase activity), and HsdS is important for specificity of the recognition (DNA-binding) site in addition to both restriction digestion (DNA cleavage) and modification (DNA methyltransferase) activity.

Type II

Typical type II restriction enzymes differ from type I restriction enzymes in several ways. They form homodimers, with recognition sites that are usually undivided and palindromic and 4–8 nucleotides in length. They recognize and cleave DNA at the same site, and they do not use ATP or AdoMet for their activity—they usually require only Mg^{2+} as a cofactor.These enzymes cleave the phosphodiester bond of double helix DNA. It can either cleave at the center of both strands to yield a blunt end, or at a staggered position leaving overhangs called sticky ends. These are the most commonly available and used restriction enzymes. In the 1990s and early 2000s, new enzymes from this family were discovered that did not follow all the classical criteria of this enzyme class, and new subfamily nomenclature was developed to divide this large family into subcategories based on deviations from typical characteristics of type II enzymes.These subgroups are defined using a letter suffix.

Type IIB restriction enzymes (e.g., BcgI and BplI) are multimers, containing more than one subunit.They cleave DNA on both sides of their recognition to cut out the recognition site. They require both AdoMet and Mg^{2+} cofactors. Type IIE restriction endonucleases (e.g., NaeI) cleave DNA following interaction with two copies of their recognition sequence. One recognition site acts as the target for cleavage, while the other acts as an allosteric effector that speeds up or improves the efficiency of enzyme cleavage. Similar to type IIE enzymes, type IIF restriction endonucleases (e.g. NgoMIV) interact with two copies of their recognition sequence but cleave both sequences at the same time.Type IIG restriction endonucleases (e.g., Eco57I) do have a single subunit, like classical Type II restriction enzymes, but require the cofactor AdoMet to be active. Type IIM restriction endonucleases, such as DpnI, are able to recognize and cut methylated DNA. Type IIS restriction endonucleases (e.g., FokI) cleave DNA at a defined distance from their non-

palindromic asymmetric recognition sites;this characteristic is widely used to perform in-vitro cloning techniques such as Golden Gate cloning. These enzymes may function as dimers. Similarly, Type IIT restriction enzymes (e.g., Bpu10I and BslI) are composed of two different subunits. Some recognize palindromic sequences while others have asymmetric recognition sites.

Type III

Type III restriction enzymes (e.g., EcoP15) recognize two separate non-palindromic sequences that are inversely oriented. They cut DNA about 20–30 base pairs after the recognition site. These enzymes contain more than one subunit and require AdoMet and ATP cofactors for their roles in DNA methylation and restriction digestion, respectively. They are components of prokaryotic DNA restriction-modification mechanisms that protect the organism against invading foreign DNA. Type III enzymes are hetero-oligomeric, multifunctional proteins composed of two subunits, Res (P08764) and Mod (P08763). The Mod subunit recognises the DNA sequence specific for the system and is a modification methyltransferase; as such, it is functionally equivalent to the M and S subunits of type I restriction endonuclease. Res is required for restriction digestion, although it has no enzymatic activity on its own. Type III enzymes recognise short 5–6 bp-long asymmetric DNA sequences and cleave 25–27 bp downstream to leave short, single-stranded 5' protrusions. They require the presence of two inversely oriented unmethylated recognition sites for restriction digestion to occur. These enzymes methylate only one strand of the DNA, at the N-6 position of adenosyl residues, so newly replicated DNA will have only one strand methylated, which is sufficient to protect against restriction digestion. Type III enzymes belong to the beta-subfamily of N6 adenine methyltransferases, containing the nine motifs that characterise this family, including motif I, the AdoMet binding pocket (FXGXG), and motif IV, the catalytic region (S/D/N (PP) Y/F).

Type IV

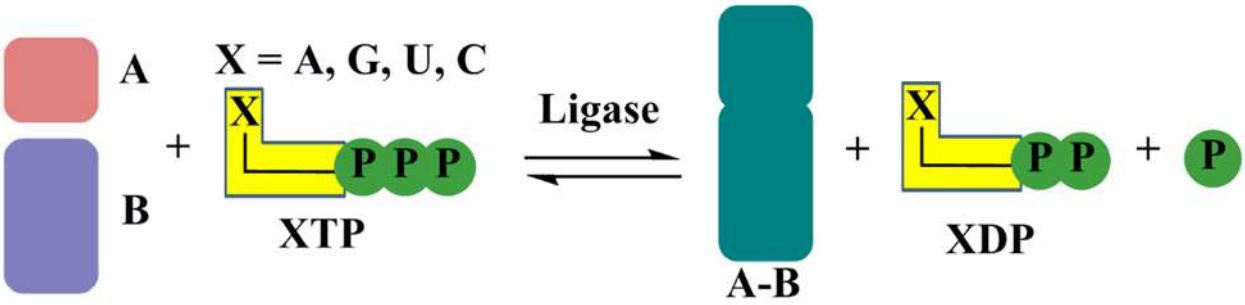
Type IV enzymes recognize modified, typically methylated DNA and are exemplified by the McrBC and Mrr systems of E. coli.

Type V

Type V restriction enzymes (e.g., the cas9-gRNA complex from CRISPRs utilize guide RNAs to target specific non-palindromic sequences found on invading organisms. They can cut DNA of variable length, provided that a suitable guide RNA is provided. The flexibility and ease of use of these enzymes make them promising for future genetic engineering applications.

Ligases

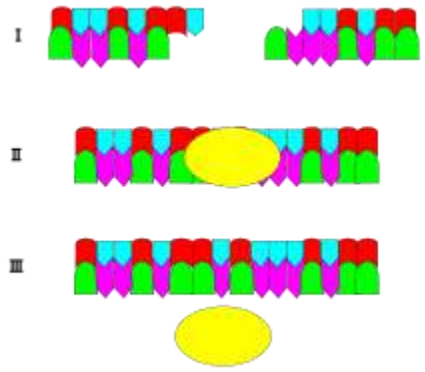
DNA ligase is an enzyme which can connect two strands of DNA together by forming a bond between the phosphate group of one strand and the deoxyribose group on another. It is used in cells to join together the Okazaki fragments which are formed on the lagging strand during DNA replication.



DNA ligase is a specific type of enzyme, a ligase, (EC 6.5.1.1) that facilitates the joining of DNA strands together by catalyzing the formation of a phosphodiester bond. It plays a role in repairing single-strand breaks in duplex DNA in living organisms, but some forms (such as DNA ligase IV) may specifically repair double-strand breaks (i.e. a break in both complementary strands of DNA). Single-strand breaks are repaired by DNA ligase using the complementary strand of the double helix as a template, with DNA ligase creating the final phosphodiester bond to fully repair the DNA.

DNA ligase is used in both DNA repair and DNA replication (see *Mammalian ligases*). In addition, DNA ligase has extensive use in molecular biology laboratories for recombinant DNA experiments . Purified DNA ligase is used in gene cloning to join DNA molecules together to form recombinant DNA.

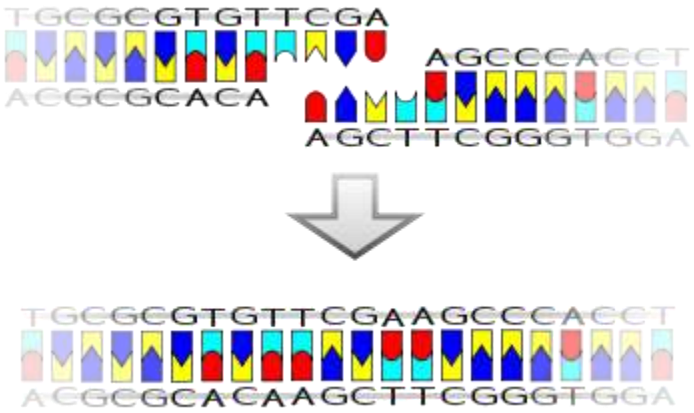
Enzymatic mechanism



The image demonstrates how ligase (yellow oval) catalyzes two DNA fragment strands. The ligase joins the two fragments of DNA to form a longer strand of DNA by "pasting" them together.

The mechanism of DNA ligase is to form two covalent phosphodiester bonds between 3' hydroxyl ends of one nucleotide ("acceptor"), with the 5' phosphate end of another ("donor"). Two ATP molecules are consumed for each phosphodiester bond formed. AMP is required for the ligase reaction, which proceeds in four steps:

1. Reorganization of activity site such as nicks in DNA segments or Okazaki fragments etc.
2. Adenylylation (addition of AMP) of a lysine residue in the active center of the enzyme, pyrophosphate is released;
3. Transfer of the AMP to the 5' phosphate of the so-called donor, formation of a pyrophosphate bond;
4. Formation of a phosphodiester bond between the 5' phosphate of the donor and the 3' hydroxyl of the acceptor.



A pictorial example of how a ligase works (with sticky ends)

Ligase will also work with blunt ends, although higher enzyme concentrations and different reaction conditions are required.

Types

E. coli

The *E. coli* DNA ligase is encoded by the *lig* gene. DNA ligase in *E. coli*, as well as most prokaryotes, uses energy gained by cleaving nicotinamide adenine dinucleotide (NAD) to create the phosphodiester bond. It does not ligate blunt-ended DNA except under conditions of molecular crowding with polyethylene glycol, and cannot join RNA to DNA efficiently.

The activity of E. coli DNA ligase can be enhanced by DNA polymerase at the right concentrations. Enhancement only works when the concentrations of the DNA polymerase 1 are much lower than the DNA fragments to be ligated. When the concentrations of Pol I DNA polymerases are higher, it has an adverse effect on E. coli DNA ligase

T4

The DNA ligase from bacteriophage T4 (a bacteriophage that infects *Escherichia coli* bacteria). The T4 ligase is the most-commonly used in laboratory research. It can ligate either cohesive or blunt ends of DNA, oligonucleotides, as well as RNA and RNA-DNA hybrids, but not single-stranded nucleic acids. It can also ligate blunt-ended DNA with much greater efficiency than *E. coli* DNA ligase. Unlike *E. coli* DNA ligase, T4 DNA ligase cannot utilize NAD and it has an absolute requirement for ATP as a cofactor. Some engineering has been done to improve the *in*

vitro activity of T4 DNA ligase; one successful approach, for example, tested T4 DNA ligase fused to several alternative DNA binding proteins and found that the constructs with either p50 or NF-kB as fusion partners were over 160% more active in blunt-end ligations for cloning purposes than wild type T4 DNA ligase. A typical reaction for inserting a fragment into a plasmid vector would use about 0.01 (sticky ends) to 1 (blunt ends) units of ligase. The optimal incubation temperature for T4 DNA ligase is 16 °C.

Bacteriophage T4 ligase mutants have increased sensitivity to both UV irradiation and the alkylating agent methyl methanesulfonate indicating that DNA ligase is employed in the repair of the DNA damages caused by these agents.

Mammalian

In mammals, there are four specific types of ligase.

1. DNA ligase I: ligates the nascent DNA of the lagging strand after the Ribonuclease H has removed the RNA primer from the Okazaki fragments.
2. DNA ligase III: complexes with DNA repair protein XRCC1 to aid in sealing DNA during the process of nucleotide excision repair and recombinant fragments. Of the all known mammalian DNA ligases, only Lig III has been found to be present in mitochondria.
3. DNA ligase IV: complexes with XRCC4. It catalyzes the final step in the non-homologous end joining DNA double-strand break repair pathway. It is also required for V(D)J recombination, the process that generates diversity in immunoglobulin and T-cell receptor loci during immune system development.

- DNA ligase II: appears to be used in repair. It is formed by alternative splicing of a proteolytic fragment of DNA ligase III and does not have its own gene, therefore it is often considered to be virtually identical to DNA ligase III.

DNA ligase from eukaryotes and some microbes uses adenosine triphosphate (ATP) rather than NAD.

Thermostable

Derived from a thermophilic bacterium, the enzyme is stable and active at much higher temperatures than conventional DNA ligases. Its half-life is 48 hours at 65 °C and greater than 1 hour at 95 °C. Ampligase DNA Ligase has been shown to be active for at least 500 thermal cycles (94 °C/80 °C) or 16 hours of cycling.This exceptional thermostability permits extremely high hybridization stringency and ligation specificity.

Measurement of activity

There are at least three different units used to measure the activity of DNA ligase:

- Weiss unit - the amount of ligase that catalyzes the exchange of 1 nmole of ^{32}P from inorganic pyrophosphate to ATP in 20 minutes at 37°C. This is the one most commonly used.
- Modrich-Lehman unit - this is rarely used, and one unit is defined as the amount of enzyme required to convert 100 nmoles of $d(A\text{-}T)_n$ to an exonuclease-III resistant form in 30 minutes under standard conditions.
- Many commercial suppliers of ligases use an arbitrary unit based on the ability of ligase to ligate cohesive ends. These units are often more subjective than quantitative and lack precision.

Research applications

DNA ligases have become indispensable tools in modern molecular biology research for generating recombinant DNA sequences. For example, DNA ligases are used with restriction enzymes to insert DNA fragments, often genes, into plasmids.

Controlling the optimal temperature is a vital aspect of performing efficient recombination experiments involving the ligation of cohesive-ended fragments. Most experiments use T4 DNA Ligase (isolated from bacteriophage T4), which is most active at 37 °C. However, for optimal ligation efficiency with cohesive-ended fragments ("sticky ends"), the optimal enzyme temperature needs to be balanced with the melting temperature T_m of the sticky ends being ligated,[13] the homologous pairing of the sticky ends will not be stable because the high temperature disrupts hydrogen bonding. A ligation reaction is most efficient when the sticky ends

are already stably annealed, and disruption of the annealing ends would therefore result in low ligation efficiency. The shorter the overhang, the lower the T_m.

Since blunt-ended DNA fragments have no cohesive ends to anneal, the melting temperature is not a factor to consider within the normal temperature range of the ligation reaction. The limiting factor in blunt end ligation is not the activity of the ligase but rather the number of alignments between DNA fragment ends that occur. The most efficient ligation temperature for blunt-ended DNA would therefore be the temperature at which the greatest number of alignments can occur. The majority of blunt-ended ligations are carried out at 14-25 °C overnight. The absence of stably annealed ends also means that the ligation efficiency is lowered, requiring a higher ligase concentration to be used.

A novel use of DNA ligase can be seen in the field of nano chemistry, specifically in DNA origami. DNA based self-assembly principles have proven useful for organizing nanoscale objects, such as biomolecules, nanomachines, nanoelectronic and photonic component. Assembly of such nano structure requires the creation of an intricate mesh of DNA molecules. Although DNA self-assembly is possible without any outside help using different substrates such as provision of catatonic surface of Aluminium foil, DNA ligase can provide the enzymatic assistance that is required to make DNA lattice structure from DNA over hangs.

Gene Cloning strategy

A set of techniques adopted for gene cloning for a particular purpose is said to be a gene cloning strategy. The basic 7 steps involved in gene cloning are:

1. Isolation of DNA [gene of interest] fragments to be cloned.
2. Insertion of isolated DNA into a suitable vector to form recombinant DNA.
3. Introduction of recombinant DNA into a suitable organism known as host.
4. Selection of transformed host cells and identification of the clone containing the gene of interest.
5. Multiplication/Expression of the introduced Gene in the host.
6. Isolation of multiple gene copies/Protein expressed by the gene.

7. Purification of the isolated gene copy/protein

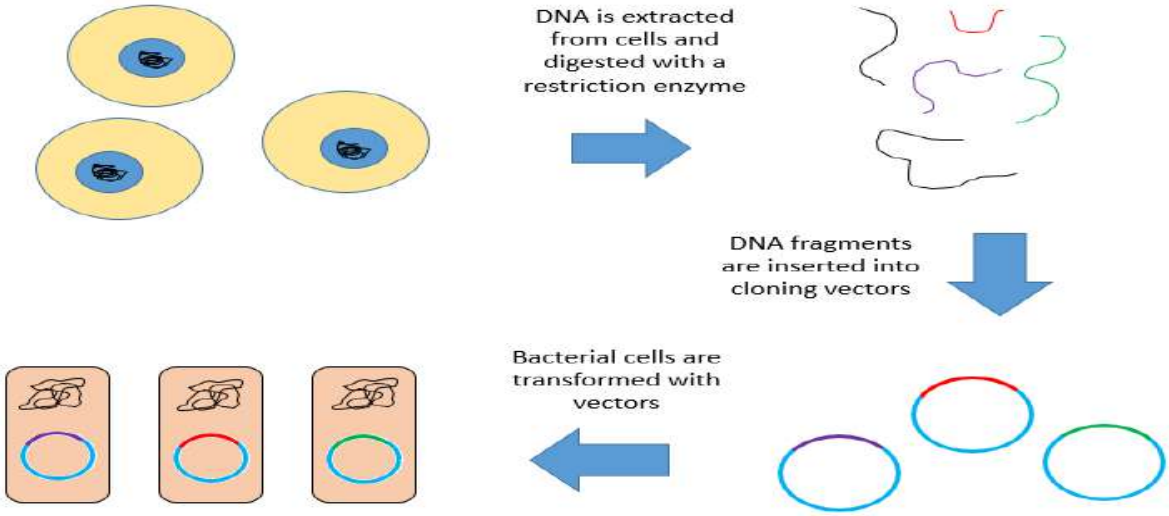
Genomic library construction

A collection of clones containing all DNA segments of the genome of an organism is called genomic DNA library.

Construction of a genomic library involves creating many recombinant DNA molecules. An organism's genomic DNA is extracted and then digested with a restriction enzyme. For organisms with very small genomes (~10 kb), the digested fragments can be separated by gel electrophoresis. The separated fragments can then be excised and cloned into the vector separately. However, when a large genome is digested with a restriction enzyme, there are far too many fragments to excise individually. The entire set of fragments must be cloned together with the vector, and separation of clones can occur after. In either case, the fragments are ligated into a vector that has been digested with the same restriction enzyme. The vector containing the inserted fragments of genomic DNA can then be introduced into a host organism.

Below are the steps for creating a genomic library from a large genome.

1. Extract and purify DNA.
2. Digest the DNA with a restriction enzyme. This creates fragments that are similar in size, each containing one or more genes.
3. Insert the fragments of DNA into vectors that were cut with the same restriction enzyme. Use the enzyme DNA ligase to seal the DNA fragments into the vector. This creates a large pool of recombinant molecules.
4. These recombinant molecules are taken up by a host bacterium by transformation, creating a DNA library.



cDNA library construction

cDNA is created from a mature mRNA from a eukaryotic cell with the use of reverse transcriptase. In eukaryotes, a poly-(A) tail (consisting of a long sequence of adenine nucleotides) distinguishes mRNA from tRNA and rRNA and can therefore be used as a primer site for reverse transcription. This has the problem that not all transcripts, such as those for the histone, encode a poly-A tail.

mRNA extraction

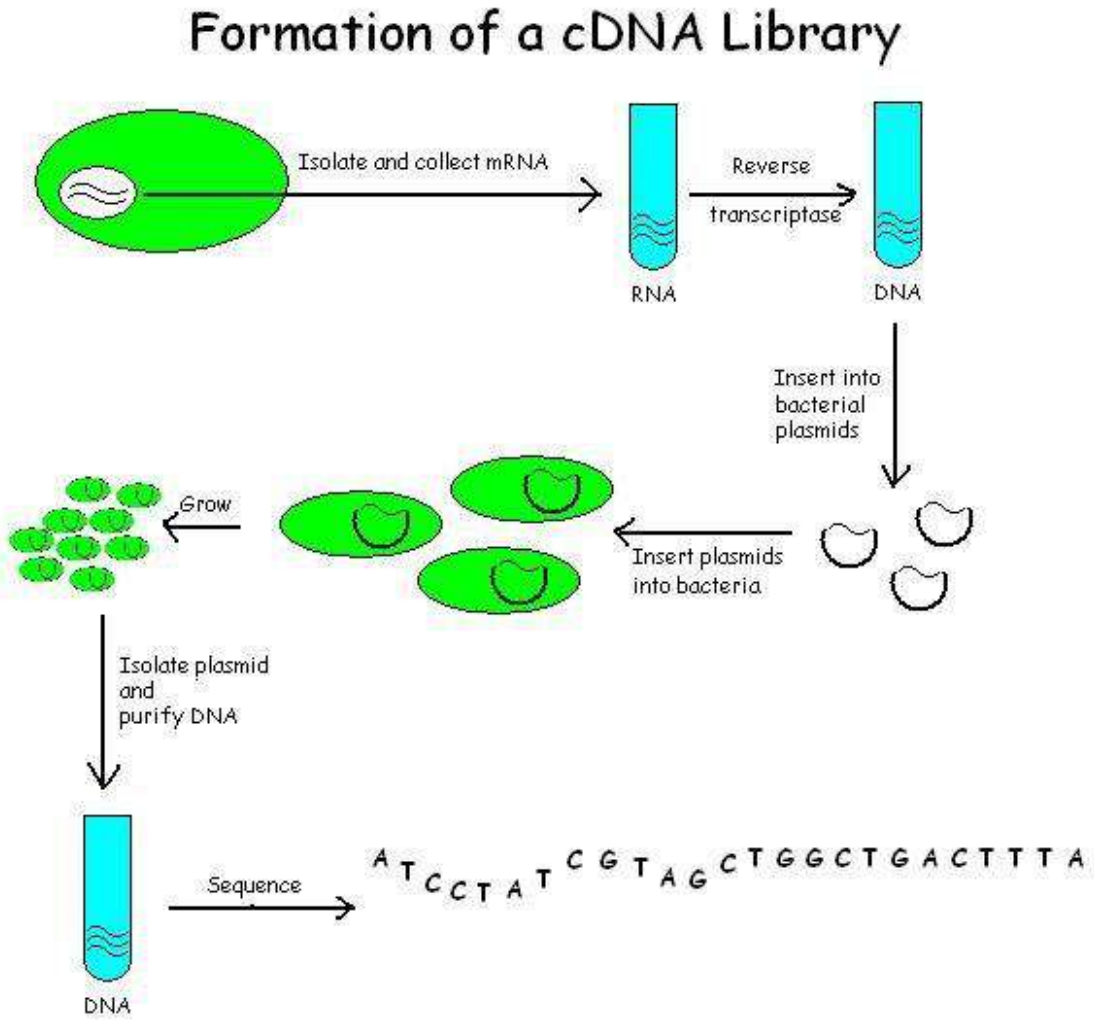
Firstly, the mRNA is obtained and purified from the rest of the RNAs. Several methods exist for purifying RNA such as trizol extraction and column purification. Column purification is done by using oligomeric dT nucleotide coated resins where only the mRNA having the poly-A tail will bind. The rest of the RNAs are eluted out. The mRNA is eluted by using eluting buffer and some heat to separate the mRNA strands from oligo-dT.

cDNA construction

Once mRNA is purified, oligo-dT (a short sequence of deoxy-thymidine nucleotides) is tagged as a complementary primer which binds to the poly-A tail providing a free 3'-OH end that can be extended by reverse transcriptase to create the complementary DNA strand. Now, the mRNA is removed by using a RNAse enzyme leaving a single stranded cDNA (sscDNA). This sscDNA is converted into a double stranded DNA with the help of DNA polymerase. However, for DNA polymerase to synthesize a complementary strand a free 3'-OH end is needed. This is provided by

the sscDNA itself by generating a hairpin loop at the 3' end by coiling on itself. The polymerase extends the 3'-OH end and later the loop at 3' end is opened by the scissoring action of S_1 nuclease. Restriction endonucleases and DNA ligase are then used to clone the sequences into bacterial plasmids.

The cloned bacteria are then selected, commonly through the use of antibiotic selection. Once selected, stocks of the bacteria are created which can later be grown and sequenced to compile the cDNA library.



Unit -II

Blotting techniques

A blot is a method of transferring proteins, DNA or RNA onto a carrier (for example, a nitrocellulose, polyvinylidene fluoride or nylon membrane). In many instances, this is done after a gel electrophoresis, transferring the molecules from the gel onto the blotting membrane, and other times adding the samples directly onto the membrane. After the blotting, the transferred proteins, DNA or RNA are then visualized by colorant staining (for example, silver staining of proteins), autoradiographic visualization of radiolabelled molecules (performed before the blot), or specific labelling of some proteins or nucleic acids. The latter is done with antibodies or hybridization probes that bind only to some molecules of the blot and have an enzyme joined to them. After proper washing, this enzymatic activity (and so, the molecules we search in the blot) is visualized by incubation with proper reactive, rendering either a colored deposit on the blot or a chemiluminescent reaction which is registered by photographic film.

Types of blotting techniques:

1. Southern blotting

Southern blotting techniques is the first nucleic acid blotting procedure developed in 1975 by Southern. Southern blotting is the techniques for the specific identification of DNA molecules.

2. Northern blotting

Northern blotting is the techniques for the specific identification of RNA molecules.

3. Western blotting

Western blotting involves the identification of proteins. Antigen + antibody

Southern Blotting

A Southern blot is a method used in molecular biology for detection of a specific DNA sequence in DNA samples. Southern blotting combines transfer of electrophoresis-separated DNA fragments to a filter membrane and subsequent fragment detection by probe hybridization.

The method is named after the British biologist Edwin Southern, who first published it in 1975.Other blotting methods (i.e., western blot,[2] northern blot, eastern blot, southwestern blot) that employ similar principles, but using RNA or protein, have later been named in reference to Edwin Southern's name. As the label is eponymous, Southern is capitalised, as is conventional of proper nouns. The names for other blotting methods may follow this convention, by analogy.

Method

1. Restriction endonucleases are used to cut high-molecular-weight DNA strands into smaller fragments.
2. The DNA fragments are then electrophoresed on an agarose gel to separate them by size.
3. If some of the DNA fragments are larger than 15 kb, then prior to blotting, the gel may be treated with an acid, such as dilute HCl. This depurinates the DNA fragments, breaking the DNA into smaller pieces, thereby allowing more efficient transfer from the gel to membrane.
4. If alkaline transfer methods are used, the DNA gel is placed into an alkaline solution (typically containing sodium hydroxide) to denature the double-stranded DNA. The denaturation in an alkaline environment may improve binding of the negatively charged thymine residues of DNA to a positively charged amino groups of membrane, separating it into single DNA strands for later hybridization to the probe (see below), and destroys any residual RNA that may still be present in the DNA. The choice of alkaline over neutral transfer methods, however, is often empirical and may result in equivalent results.
5. A sheet of nitrocellulose (or, alternatively, nylon) membrane is placed on top of (or below, depending on the direction of the transfer) the gel. Pressure is applied evenly to the gel (either using suction, or by placing a stack of paper towels and a weight on top of the membrane and gel), to ensure good and even contact between gel and membrane. If transferring by suction, 20X SSC buffer is used to ensure a seal and prevent drying of the gel. Buffer transfer by capillary action from a region of high water potential to a region of low water potential (usually filter paper and paper tissues) is then used to move the DNA from the gel onto the membrane; ion exchange interactions bind the DNA to the membrane due to the negative charge of the DNA and positive charge of the membrane.

6. The membrane is then baked in a vacuum or regular oven at 80 °C for 2 hours (standard conditions; nitrocellulose or nylon membrane) or exposed to ultraviolet radiation (nylon membrane) to permanently attach the transferred DNA to the membrane.
7. The membrane is then exposed to a hybridization probe—a single DNA fragment with a specific sequence whose presence in the target DNA is to be determined. The probe DNA is labelled so that it can be detected, usually by incorporating radioactivity or tagging the molecule with a fluorescent or chromogenic dye. In some cases, the hybridization probe may be made from RNA, rather than DNA. To ensure the specificity of the binding of the probe to the sample DNA, most common hybridization methods use salmon or herring sperm DNA for blocking of the membrane surface and target DNA, deionized formamide, and detergents such as SDS to reduce non-specific binding of the probe.
8. After hybridization, excess probe is washed from the membrane (typically using SSC buffer), and the pattern of hybridization is visualized on X-ray film by autoradiography in the case of a radioactive or fluorescent probe, or by development of colour on the membrane if a chromogenic detection method is used.

Hybridization of the probe to a specific DNA fragment on the filter membrane indicates that this fragment contains DNA sequence that is complementary to the probe. The transfer step of the DNA from the electrophoresis gel to a membrane permits easy binding of the labeled hybridization probe to the size-fractionated DNA. It also allows for the fixation of the target-probe hybrids, required for analysis by autoradiography or other detection methods. Southern blots performed with restriction enzyme-digested genomic DNA may be used to determine the number of sequences (e.g., gene copies) in a genome. A probe that hybridizes only to a single DNA segment that has not been cut by the restriction enzyme will produce a single band on a Southern blot, whereas multiple bands will likely be observed when the probe hybridizes to several highly similar sequences (e.g., those that may be the result of sequence duplication). Modification of the hybridization conditions (for example, increasing the hybridization temperature or decreasing salt concentration) may be used to increase specificity and decrease hybridization of the probe to sequences that are less than 100% similar.

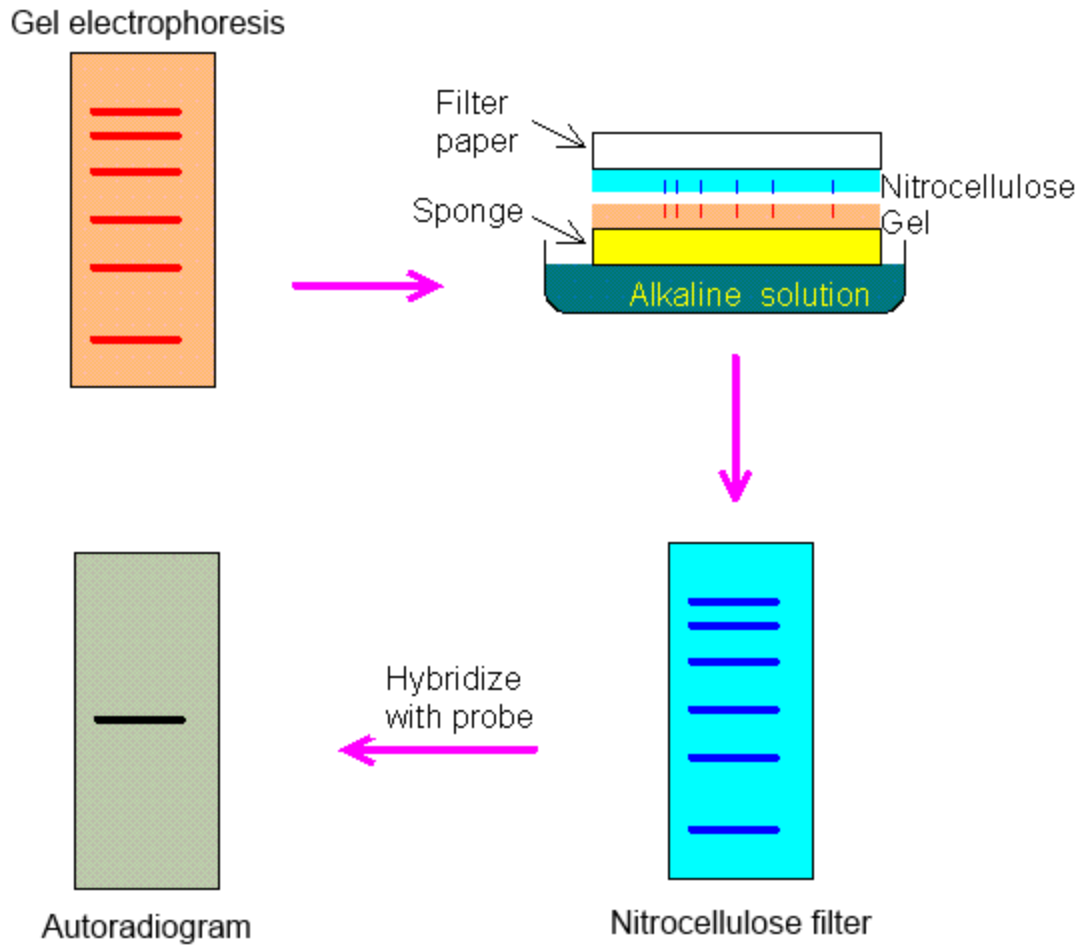


Figure 9-D-1. Southern blotting. (a) The DNA to be analyzed is digested with restriction enzymes and then separated by agarose gel electrophoresis. (b) The DNA fragments in the gel are denatured with alkaline solution and transferred onto a nitrocellulose filter or nylon membrane by blotting, preserving the distribution of the DNA fragments in the gel. (c) The nitrocellulose filter is incubated with a specific probe. The location of the DNA fragment that hybridizes with the probe can be displayed by autoradiography.

Applications

Southern blotting transfer may be used for homology-based cloning on the basis of amino acid sequence of the protein product of the target gene. Oligonucleotides are designed so that they are similar to the target sequence. The oligonucleotides are chemically synthesized, radiolabeled, and used to screen a DNA library, or other collections of cloned DNA fragments. Sequences that

hybridize with the hybridization probe are further analysed, for example, to obtain the full length sequence of the targeted gene.

Southern blotting can also be used to identify methylated sites in particular genes. Particularly useful are the restriction nucleases *MspI* and *HpaII*, both of which recognize and cleave within the same sequence. However, *HpaII* requires that a C within that site be methylated, whereas *MspI* cleaves only DNA unmethylated at that site. Therefore, any methylated sites within a sequence analyzed with a particular probe will be cleaved by the former, but not the latter, enzyme.

Northern Blotting

The northern blot, or RNA blot,is a technique used in molecular biology research to study gene expression by detection of RNA (or isolated mRNA) in a sample.

With northern blotting it is possible to observe cellular control over structure and function by determining the particular gene expression rates during differentiation and morphogenesis, as well as in abnormal or diseased conditions. Northern blotting involves the use of electrophoresis to separate RNA samples by size, and detection with a hybridization probe complementary to part of or the entire target sequence. The term 'northern blot' actually refers specifically to the capillary transfer of RNA from the electrophoresis gel to the blotting membrane. However, the entire process is commonly referred to as northern blotting. The northern blot technique was developed in 1977 by James Alwine, David Kemp, and George Stark at Stanford University,with contributions from Gerhard Heinrich. Northern blotting takes its name from its similarity to the first blotting technique, the Southern blot, named for biologist Edwin Southern. The major difference is that RNA, rather than DNA, is analyzed in the northern blot.

Procedure

A general blotting procedure starts with extraction of total RNA from a homogenized tissue sample or from cells. Eukaryotic mRNA can then be isolated through the use of oligo (dT) cellulose chromatography to isolate only those RNAs with a poly(A) tail. RNA samples are then separated by gel electrophoresis. Since the gels are fragile and the probes are unable to enter the

matrix, the RNA samples, now separated by size, are transferred to a nylon membrane through a capillary or vacuum blotting system.

Capillary blotting system setup for the transfer of RNA from an electrophoresis gel to a blotting membrane.

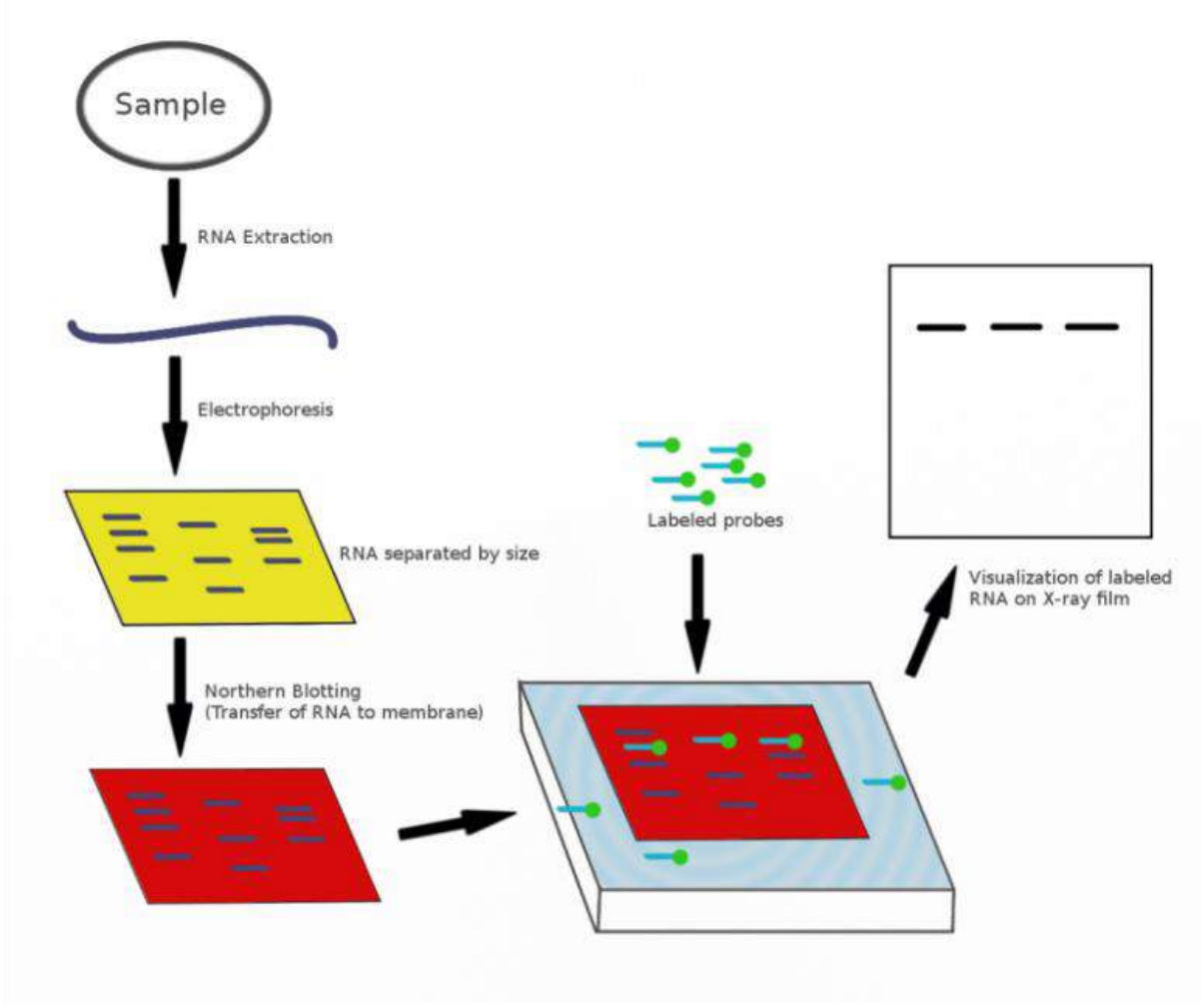
A nylon membrane with a positive charge is the most effective for use in northern blotting since the negatively charged nucleic acids have a high affinity for them. The transfer buffer used for the blotting usually contains formamide because it lowers the annealing temperature of the probe-RNA interaction, thus eliminating the need for high temperatures, which could cause RNA degradation. Once the RNA has been transferred to the membrane, it is immobilized through covalent linkage to the membrane by UV light or heat. After a probe has been labeled, it is hybridized to the RNA on the membrane. Experimental conditions that can affect the efficiency and specificity of hybridization include ionic strength, viscosity, duplex length, mismatched base pairs, and base composition.The membrane is washed to ensure that the probe has bound specifically and to prevent background signals from arising. The hybrid signals are then detected by X-ray film and can be quantified by densitometry. To create controls for comparison in a northern blot, samples not displaying the gene product of interest can be used after determination by microarrays or RT-PCR.

The RNA samples are most commonly separated on agarose gels containing formaldehyde as a denaturing agent for the RNA to limit secondary structure.[11][12] The gels can be stained with ethidium bromide (EtBr) and viewed under UV light to observe the quality and quantity of RNA before blotting. Polyacrylamide gel electrophoresis with urea can also be used in RNA separation but it is most commonly used for fragmented RNA or microRNAs. An RNA ladder is

often run alongside the samples on an electrophoresis gel to observe the size of fragments obtained but in total RNA samples the ribosomal subunits can act as size markers.Since the large ribosomal subunit is 28S (approximately 5kb) and the small ribosomal subunit is 18S (approximately 2kb) two prominent bands appear on the gel, the larger at close to twice the intensity of the smaller.

Probes

Probes for northern blotting are composed of nucleic acids with a complementary sequence to all or part of the RNA of interest, they can be DNA, RNA, or oligonucleotides with a minimum of 25 complementary bases to the target sequence. RNA probes (riboprobes) that are transcribed in vitro are able to withstand more rigorous washing steps preventing some of the background noise. Commonly cDNA is created with labelled primers for the RNA sequence of interest to act as the probe in the northern blot.The probes must be labelled either with radioactive isotopes (^{32}P) or with chemiluminescence in which alkaline phosphatase or horseradish peroxidase (HRP) break down chemiluminescent substrates producing a detectable emission of light.[16] The chemiluminescent labelling can occur in two ways: either the probe is attached to the enzyme, or the probe is labelled with a ligand (e.g. biotin) for which the ligand (e.g., avidin or streptavidin) is attached to the enzyme (e.g. HRP). X-ray film can detect both the radioactive and chemiluminescent signals and many researchers prefer the chemiluminescent signals because they are faster, more sensitive, and reduce the health hazards that go along with radioactive labels.The same membrane can be probed up to five times without a significant loss of the target RNA.



Applications

Northern blotting allows one to observe a particular gene's expression pattern between tissues, organs, developmental stages, environmental stress levels, pathogen infection, and over the course of treatment.The technique has been used to show overexpression of oncogenes and downregulation of tumor-suppressor genes in cancerous cells when compared to 'normal' tissue, as well as the gene expression in the rejection of transplanted organs. If an upregulated gene is observed by an abundance of mRNA on the northern blot the sample can then be sequenced to determine if the gene is known to researchers or if it is a novel finding. The expression patterns obtained under given conditions can provide insight into the function of that gene. Since the RNA is first separated by size, if only one probe type is used variance in the level of each band on the membrane can provide insight into the size of the product, suggesting alternative splice products

of the same gene or repetitive sequence motifs.The variance in size of a gene product can also indicate deletions or errors in transcript processing. By altering the probe target used along the known sequence it is possible to determine which region of the RNA is missing.

Advantages and disadvantages

Analysis of gene expression can be done by several different methods including RT-PCR, RNase protection assays, microarrays, RNA-Seq, serial analysis of gene expression (SAGE), as well as northern blotting. Microarrays are quite commonly used and are usually consistent with data obtained from northern blots; however, at times northern blotting is able to detect small changes in gene expression that microarrays cannot. The advantage that microarrays have over northern blots is that thousands of genes can be visualized at a time, while northern blotting is usually looking at one or a small number of genes.

A problem in northern blotting is often sample degradation by RNases (both endogenous to the sample and through environmental contamination), which can be avoided by proper sterilization of glassware and the use of RNase inhibitors such as DEPC (diethylpyrocarbonate). The chemicals used in most northern blots can be a risk to the researcher, since formaldehyde, radioactive material, ethidium bromide, DEPC, and UV light are all harmful under certain exposures. Compared to RT-PCR, northern blotting has a low sensitivity, but it also has a high specificity, which is important to reduce false positive results.

The advantages of using northern blotting include the detection of RNA size, the observation of alternate splice products, the use of probes with partial homology, the quality and quantity of RNA can be measured on the gel prior to blotting, and the membranes can be stored and reprobed for years after blotting.

For northern blotting for the detection of acetylcholinesterase mRNA the nonradioactive technique was compared to a radioactive technique and found as sensitive as the radioactive one, but requires no protection against radiation and is less time consuming.

Western Blotting

The western blot (sometimes called the protein immunoblot), or western blotting, is a widely used analytical technique in molecular biology and immunogenetics to detect specific proteins in a sample of tissue homogenate or extract.

In brief, the sample undergoes protein denaturation, followed by gel electrophoresis. A synthetic or animal-derived antibody (known as the primary antibody) is created that recognises and binds to a specific target protein. The electrophoresis membrane is washed in a solution containing the primary antibody, before excess antibody is washed off. A secondary antibody is added which recognises and binds to the primary antibody. The secondary antibody is visualised through various methods such as staining, immunofluorescence, and radioactivity, allowing indirect detection of the specific target protein.

Procedure

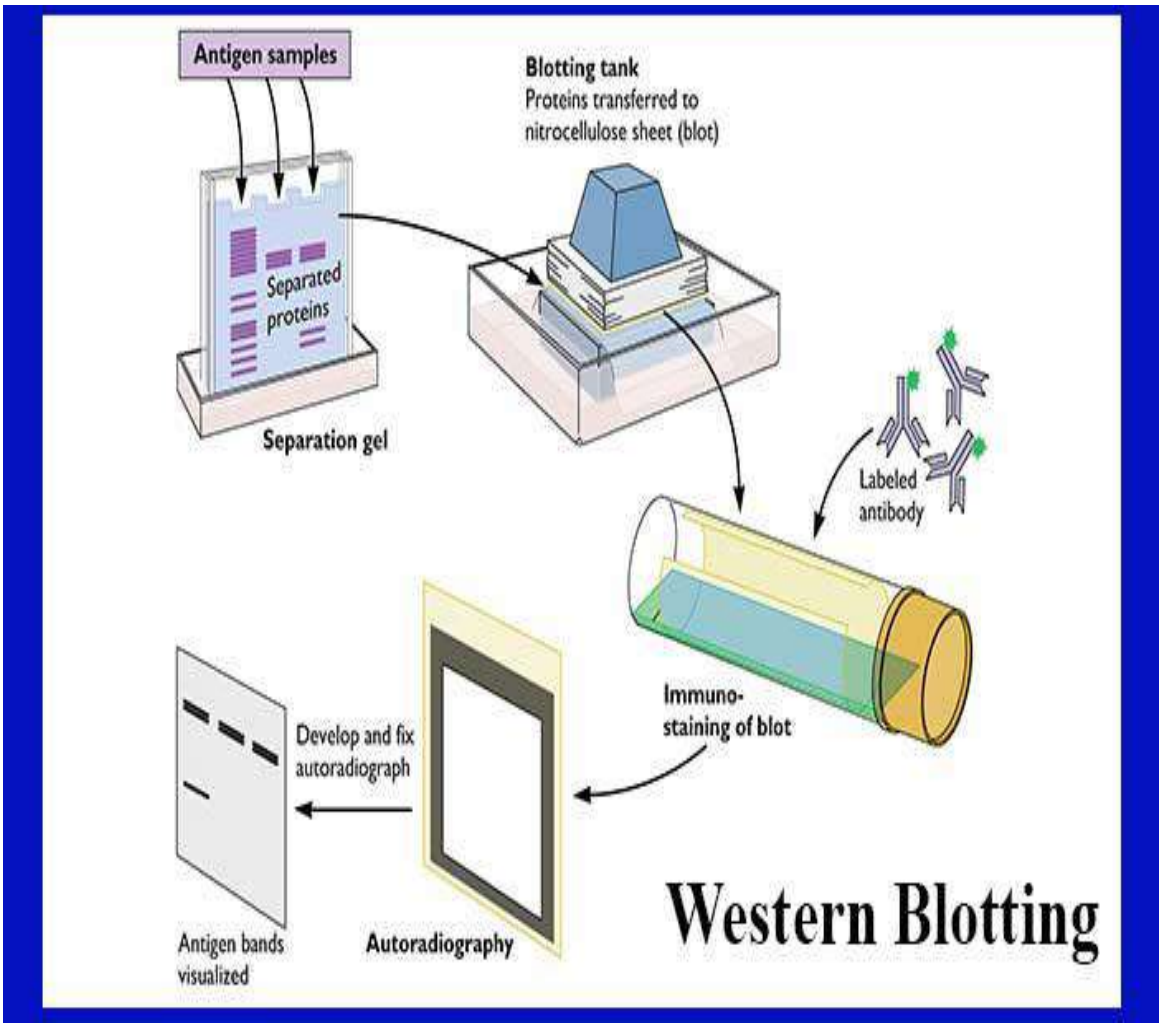
The western blot method is composed of a gel electrophoresis to separate native proteins by 3-D structure or denatured proteins by the length of the polypeptide, followed by an electrophoretic transfer onto a membrane (mostly PVDF or Nitrocellulose) and an immunostaining procedure to visualize a certain protein on the blot membrane. SDS-PAGE is generally used for the denaturing electrophoretic separation of proteins. SDS is generally used as a buffer (as well as in the gel) in order to give all proteins present a uniform negative charge, since proteins can be positively, negatively, or neutrally charged. This type of electrophoresis is known as SDS-PAGE (SDS-polyacrylamide gel electrophoresis). Prior to electrophoresis, protein samples are often boiled to denature the proteins present. This ensures that proteins are separated based on size and prevents proteases (enzymes that break down proteins) from degrading samples. Following electrophoretic separation, the proteins are transferred to a membrane (typically nitrocellulose or PVDF), where they are blocked with milk (or other blocking agents) to prevent non-specific antibody binding, and then stained with antibodies specific to the target protein. Lastly, the membrane will be stained with a secondary antibody that recognizes the first antibody staining, which can then be used for detection by a variety of methods. The gel electrophoresis step is included in western blot analysis to resolve the issue of the cross-reactivity of antibodies.

Gel electrophoresis

The proteins of the sample are separated using gel electrophoresis. Separation of proteins may be by isoelectric point (pI), molecular weight, electric charge, or a combination of these factors. The nature of the separation depends on the treatment of the sample and the nature of the gel.

By far the most common type of gel electrophoresis employs polyacrylamide gels and buffers loaded with sodium dodecyl sulfate (SDS). SDS-PAGE (SDS polyacrylamide gel electrophoresis) maintains polypeptides in a denatured state once they have been treated with strong reducing agents to remove secondary and tertiary structure (e.g. disulfide bonds [S-S] to sulfhydryl groups [SH and SH]) and thus allows separation of proteins by their molecular mass. Sampled proteins become covered in the negatively charged SDS, effectively becoming anionic, and migrate towards the positively charged (higher voltage) anode (usually having a red wire) through the acrylamide mesh of the gel. Smaller proteins migrate faster through this mesh, and the proteins are thus separated according to size (usually measured in kilodaltons, kDa). The concentration of acrylamide determines the resolution of the gel - the greater the acrylamide concentration, the better the resolution of lower molecular weight proteins. The lower the acrylamide concentration, the better the resolution of higher molecular weight proteins. Proteins travel only in one dimension along the gel for most blots.

Samples are loaded into *wells* in the gel. One lane is usually reserved for a *marker* or *ladder*, which is a commercially available mixture of proteins of known molecular weights, typically stained so as to form visible, coloured bands. When voltage is applied along the gel, proteins migrate through it at different speeds dependent on their size. These different rates of advancement (different *electrophoretic mobilities*) separate into *bands* within each *lane*. Protein bands can then be compared to the ladder bands, allowing estimation of the protein's molecular weight.



SDS-PAGE electrophoresis

It is also possible to use a two-dimensional gel which spreads the proteins from a single sample out in two dimensions. Proteins are separated according to isoelectric point (pH at which they have a neutral net charge) in the first dimension, and according to their molecular weight in the second dimension.

Transfer

To make the proteins accessible to antibody detection, they are moved from within the gel onto a membrane made of *nitrocellulose (NC) or polyvinylidene difluoride (PVDF*). The most commonly used method for transferring the proteins is called electroblotting. Electroblotting uses an electric

current to pull the negatively charged proteins from the gel towards the positively charged anode, and into the PVDF or NC membrane. The proteins move from within the gel onto the membrane while maintaining the organization they had within the gel. An older method of transfer involves placing a membrane on top of the gel, and a stack of filter papers on top of that. The entire stack is placed in a buffer solution which moves up the paper by capillary action, bringing the proteins with it. In practice this method is not commonly used due to the lengthy procedure time.

As a result of either transfer process, the proteins are exposed on a thin membrane layer for detection. Both varieties of membrane are chosen for their non-specific protein binding properties (i.e. binds all proteins equally well). Protein binding is based upon hydrophobic interactions, as well as charged interactions between the membrane and protein. Nitrocellulose membranes are cheaper than PVDF, but are far more fragile and cannot withstand repeated probings.

Total protein staining

Total protein staining allows the total protein that has been successfully transferred to the membrane to be visualised, allowing the user to check the uniformity of protein transfer and to perform subsequent normalization of the target protein with the actual protein amount per lane. Normalization with the so-called "loading control" was based on immunostaining of housekeeping proteins in the classical procedure, but is heading toward total protein staining recently, due to multiple benefits.[12] At least seven different approaches for total protein staining have been described for western blot normalization: Ponceau S, stain-free techniques, Sypro Ruby, Epicocconone, Coomassie R-350, Amido Black, and Cy5. In order to avoid noise of signal, total protein staining should be performed before blocking of the membrane. Nevertheless, post-antibody stainings have been described as well.[13]

Blocking

Since the membrane has been chosen for its ability to bind protein and as both antibodies and the target are proteins, steps must be taken to prevent the interactions between the membrane and the antibody used for detection of the target protein. Blocking of non-specific binding is achieved by placing the membrane in a dilute solution of protein – typically 3–5% bovine serum albumin (BSA) or non-fat dry milk (both are inexpensive) in tris-buffered saline (TBS) or I-Block,

with a minute percentage (0.1%) of detergent such as Tween 20 or Triton X-100. Although non-fat dry milk is preferred due to its availability, an appropriate blocking solution is needed as not all proteins in milk are compatible with all the detection bands.The protein in the dilute solution attaches to the membrane in all places where the target proteins have not attached. Thus, when the antibody is added, it cannot bind to the membrane, and therefore the only available binding site is the specific target protein. This reduces background in the final product of the western blot, leading to clearer results, and eliminates false positives.

Incubation

During the detection process the membrane is "probed" for the protein of interest with a modified antibody which is linked to a reporter enzyme; when exposed to an appropriate substrate, this enzyme drives a colorimetric reaction and produces a color. For a variety of reasons, this traditionally takes place in a two-step process, although there are now one-step detection methods available for certain applications.

Primary antibody

The primary antibodies are generated when a host species or immune cell culture is exposed to the protein of interest (or a part thereof). Normally, this is part of the immune response, whereas here they are harvested and used as sensitive and specific detection tools that bind the protein directly.

After blocking, a solution of primary antibody (generally between 0.5 and 5 micrograms/mL) diluted in either PBS or TBST wash buffer is incubated with the membrane under gentle agitation for typically an hour at room temperature, or overnight at 4°C. The antibody solution is incubated with the membrane for anywhere from 30 minutes to overnight. It can also be incubated at different temperatures, with lesser temperatures being associated with more binding, both specific (to the target protein, the "signal") and non-specific ("noise"). Following incubation, the membrane is washed several times in wash buffer to remove unbound primary antibody, and thereby minimize background. Typically, the wash buffer solution is composed of buffered saline solution with a small percentage of detergent, and sometimes with powdered milk or BSA.

Secondary antibody

After rinsing the membrane to remove unbound primary antibody, the membrane is exposed to another antibody known as the secondary antibody. Antibodies come from animal sources (or animal sourced hybridoma cultures). The secondary antibody recognises and binds to the species-specific portion of the primary antibody. Therefore, an anti-mouse secondary antibody will bind to almost any mouse-sourced primary antibody, and can be referred to as an 'anti-species' antibody (e.g. anti-mouse, anti-goat etc.). To allow detection of the target protein, the secondary antibody is commonly linked to biotin or a reporter enzyme such as alkaline phosphatase or horseradish peroxidase. This means that several secondary antibodies will bind to one primary antibody and enhance the signal, allowing the detection of proteins of a much lower concentration than would be visible by SDS-PAGE alone.

Horseradish peroxidase (HRP) is commonly linked to secondary antibodies to allow the detection of the target protein by chemiluminescence. The chemiluminscent substrate is cleaved by HRP, resulting in the production of luminescence. Therefore, the production of luminescence is proportional to the amount of HRP-conjugated secondary antibody, and therefore, indirectly measures the presence of the target protein. A sensitive sheet of photographic film is placed against the membrane, and exposure to the light from the reaction creates an image of the antibodies bound to the blot. A cheaper but less sensitive approach utilizes a 4-chloronaphthol stain with 1% hydrogen peroxide; the reaction of peroxide radicals with 4-chloronaphthol produces a dark purple stain that can be photographed without using specialized photographic film.

Western blot binding

As with the ELISPOT and ELISA procedures, the enzyme can be provided with a substrate molecule that will be converted by the enzyme to a colored reaction product that will be visible on the membrane (see the figure below with blue bands).

Another method of secondary antibody detection utilizes a near-infrared (NIR) fluorophore-linked antibody. The light produced from the excitation of a fluorescent dye is static, making fluorescent detection a more precise and accurate measure of the difference in the signal produced by labeled antibodies bound to proteins on a western blot. Proteins can be accurately quantified because the

signal generated by the different amounts of proteins on the membranes is measured in a static state, as compared to chemiluminescence, in which light is measured in a dynamic state.

A third alternative is to use a radioactive label rather than an enzyme coupled to the secondary antibody, such as labeling an antibody-binding protein like *Staphylococcus* Protein A or Streptavidin with a radioactive isotope of iodine. Since other methods are safer, quicker, and cheaper, this method is now rarely used; however, an advantage of this approach is the sensitivity of auto-radiography-based imaging, which enables highly accurate protein quantification when combined with optical software (e.g. Optiquant).

Detection and visualization

After the unbound probes are washed away, the western blot is ready for detection of the probes that are labeled and bound to the protein of interest. In practical terms, not all westerns reveal protein only at one band in a membrane. Size approximations are taken by comparing the stained bands to that of the marker or ladder loaded during electrophoresis. The process is commonly repeated for a structural protein, such as actin or tubulin, that should not change between samples. The amount of target protein is normalized to the structural protein to control between groups. A superior strategy is the normalization to the total protein visualized with trichloroethanol or epicocconone. This practice ensures correction for the amount of total protein on the membrane in case of errors or incomplete transfers.

Colony hybridization

Colony hybridization is a method of selecting bacterial colonies with desired genes.This method was discovered by Michael Grunstein and David S. Hogness.

Method

Colony hybridization begins with culturing sparsely populated bacterial colonies on a nutrient agar plate. These colonies are symmetrically replicated on a nitrocellulose filter by direct contact, after which the cells on the filter membrane are lysed and their DNA is denatured, allowing it to bind to the filter. These DNA clusters are then hybridized to a desired radioactively-labelled RNA or DNA probe (chosen specifically beforehand) and screened by autoradiography.

DNA clusters that exhibit a desired gene are then matched up to the corresponding (living) bacterial colonies, which can be isolated for further growth and experimentation.

Definition of Colony Hybridization:- Colony hybridization is the “Blot analysis technique” where the bacterial cells are transferred from the solid nutrient medium to the absorbent material. Colony hybridization can define as the method for the isolation of the specific DNA sequences or genes from the bacterial cells containing hybrid DNA, by the means of a nitrocellulose membrane filter. The transferring medium then goes through several chemical and physical treatment.

Transferring medium of Colony Hybridization:- The nitrocellulose filter paper is the transferring medium of the colony hybridization which forms replicas of the master plate. The nitrocellulose acts as a membrane which contains the exact copies of the gene to that of the master plate. Nitrocellulose filter paper acts as the “Blotting pad”.

Colony hybridization involves the following steps:

Preparation of Master plate:- First, inoculate the bacterial cell suspension on the solid agar medium to prepare the master plate. After the inoculation, the number of bacterial colonies will develop with different plasmids which refer as “Master or Reference plate”.

Formation of replicas over a nitrocellulose filter:- Then transfer the bacterial cells from the master plate on to the membrane or filter by the means of “Nitrocellulose filter”. Press the nitrocellulose filter paper over the surface of the master plate. This compression of the filter membrane will form replicas or copies of the bacterial cells as that of the master plate.

Treatment of filter medium with SDS:- After that treat the nitrocellulose filter paper with the detergent like SDS (Sodium dodecyl sulfate) to lyse the bacterial cells.

Treatment of filter medium with alkali:- Treat the filter medium with the alkali like sodium hydroxide in order to separate the DNA into single strands.

Fixation of DNA onto the filter medium:- To fix the DNA onto the nitrocellulose filter paper, either bake the filter paper at 80 degrees Celsius or expose it to the UV light.

Addition of radioactive probe:- Hybridize the nitrocellulose filter paper containing imprints of the plasmid DNA by the addition of radioactive RNA probe. This radioactive RNA probe will code the desired gene of sequence from the bacterial cells.

Washing and Autoradiography:- Wash the filter paper to remove unbound probe particles. After that, expose the nitrocellulose filter paper to the X-ray film by the method refer as "Autoradiography". The colony which will appear after autoradiography will refer as "Autoradiogram" which carry the genes of interest.

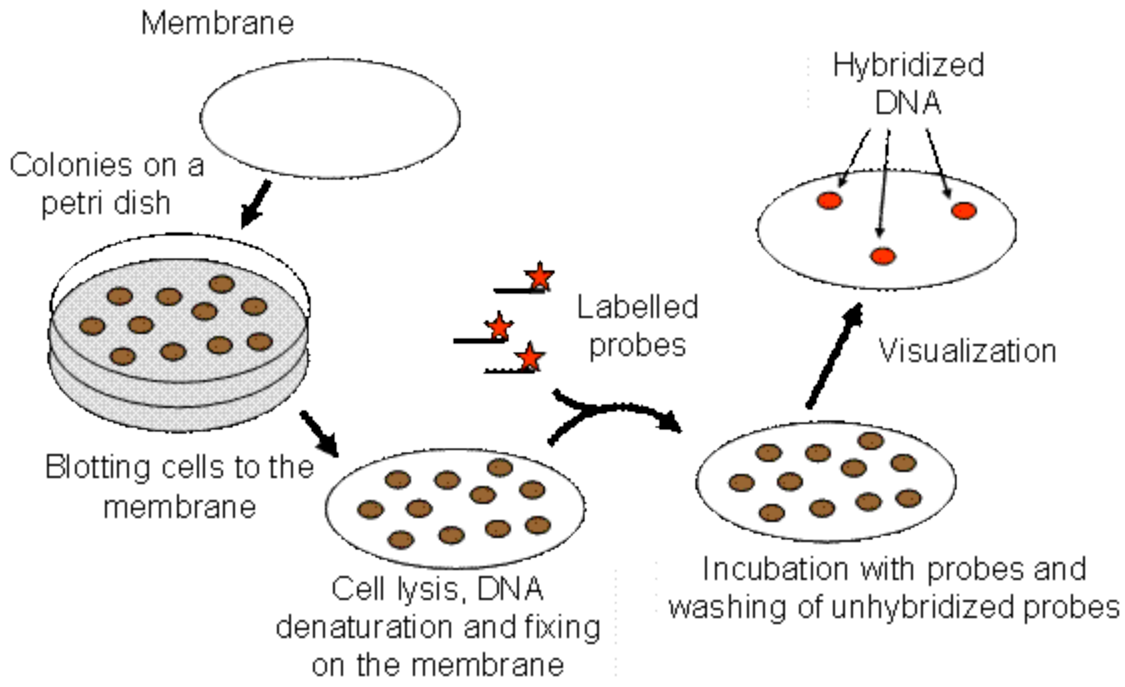
Identification of the desired gene:- Then compare the developed autoradiogram with the master plate to identify the colonies containing a gene of interest.

The cells which contain the desired gene can grow in the liquid medium and can further process for the isolation of recombinant plasmid DNA.

Conclusion:- Therefore we can conclude that the colony hybridization method is the "Screening technique" which makes the use of the radioactive probe. The radioactively labelled probe then screen or isolate the particular gene from the number of bacterial colonies.

Colony hybridization begins with culturing sparsely populated bacterial colonies on a nutrient agar plate. These colonies are symmetrically replicated on a nitrocellulose filter by direct contact, after which the cells on the filter membrane are lysed and their DNA is denatured, allowing it to bind

to the filter. These DNA clusters are then hybridized to a desired radioactively- labelled RNA or DNA probe (chosen specifically beforehand) and screened by autoradiography. DNA clusters that exhibit a desired gene are then matched up to the corresponding (living) bacterial colonies, which can be isolated for further growth and experimentation.



Polymerase Chain Reaction

Polymerase chain reaction (PCR) is a method widely used to rapidly make millions to billions of copies of a specific DNA sample, allowing scientists to take a very small sample of DNA and amplify it to a large enough amount to study in detail. PCR was invented in 1984 by the American biochemist Kary Mullis at Cetus Corporation. It is fundamental to much of genetic testing including analysis of ancient samples of DNA and identification of infectious agents. Using PCR, copies of very small amounts of DNA sequences are exponentially amplified in a series of cycles of temperature changes.

Procedure

The individual steps are as follows:

•Initialization: This step is only required for DNA polymerases that require heat activation by hot-start PCR.[12] It consists of heating the reaction chamber to a temperature of 94– 96 °C (201–205 °F), or 98 °C (208 °F) if extremely thermostable polymerases are used, which is then held for 1–10 minutes.

•Denaturation: This step is the first regular cycling event and consists of heating the reaction chamber to 94–98 °C (201–208 °F) for 20–30 seconds. This causes DNA melting, or

denaturation, of the double-stranded DNA template by breaking the hydrogen bonds between complementary bases, yielding two single-stranded DNA molecules.

•Annealing: In the next step, the reaction temperature is lowered to 50–65 °C (122–149 °F) for 20–40 seconds, allowing annealing of the primers to each of the single-stranded DNA templates. Two different primers are typically included in the reaction mixture: one for each of the two single-stranded complements containing the target region. The primers are single-stranded sequences themselves, but are much shorter than the length of the target region, complementing only very short sequences at the 3′ end of each strand.

It is critical to determine a proper temperature for the annealing step because efficiency and specificity are strongly affected by the annealing temperature. This temperature must be low enough to allow for hybridization of the primer to the strand, but high enough for the hybridization to be specific, i.e., the primer should bind only to a perfectly complementary part of the strand, and nowhere else. If the temperature is too low, the primer may bind imperfectly. If it is too high, the primer may not bind at all. A typical annealing temperature is about 3–5 °C below the Tm of the primers used. Stable hydrogen bonds between complementary bases are formed only when the primer sequence very closely matches the template sequence. During this step, the polymerase binds to the primer-template hybrid and begins DNA formation.

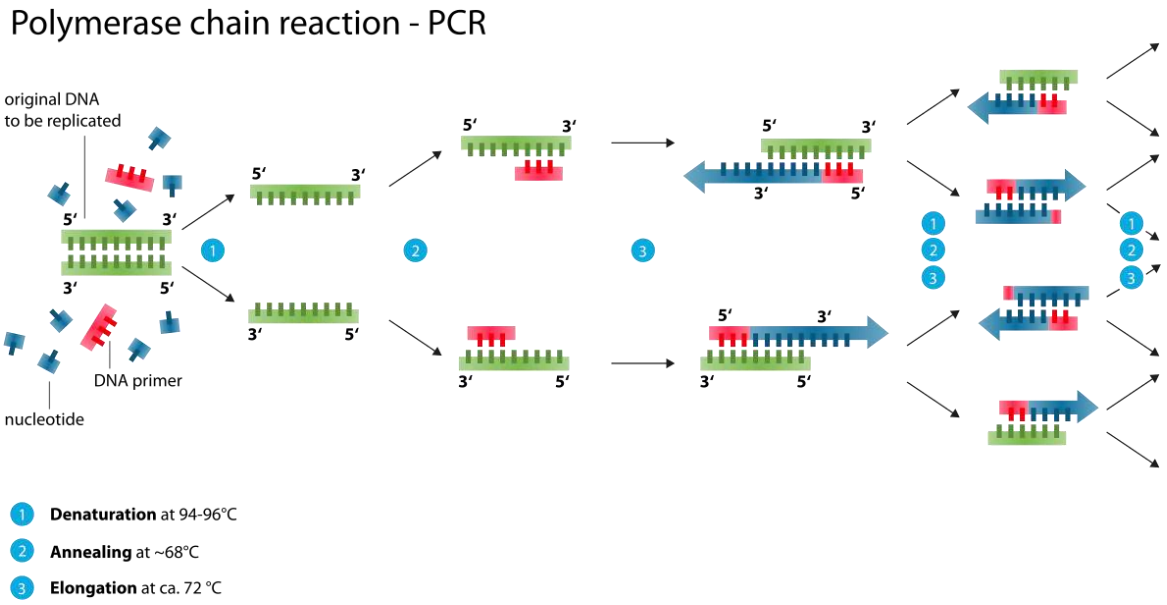
•Extension/elongation: The temperature at this step depends on the DNA polymerase used; the optimum activity temperature for the thermostable DNA polymerase of Taq polymerase is approximately 75–80 °C (167–176 °F),[13][14] though a temperature of 72 °C (162 °F) is commonly used with this enzyme. In this step, the DNA polymerase synthesizes a new DNA strand complementary to the DNA template strand by adding free dNTPs from the reaction mixture that is complementary to the template in the 5′-to-3′ direction, condensing the 5′-phosphate group of the dNTPs with the 3′-hydroxy group at the end of the nascent (elongating) DNA strand. The precise time required for elongation depends both on the DNA polymerase used and on the length of the DNA target region to amplify. As a rule of thumb, at their optimal temperature, most DNA polymerases polymerize a thousand bases per minute. Under optimal conditions (i.e., if there are no limitations due to limiting substrates or reagents), at each extension/elongation step, the number

of DNA target sequences is doubled. With each successive cycle, the original template strands plus all newly generated strands become template strands for the next round of elongation, leading to exponential (geometric) amplification of the specific DNA target region.

The processes of denaturation, annealing and elongation constitute a single cycle. Multiple cycles are required to amplify the DNA target to millions of copies. The formula used to calculate the number of DNA copies formed after a given number of cycles is 2n, where n is the number of cycles. Thus, a reaction set for 30 cycles results in 230, or 1,073,741,824, copies of the original double-stranded DNA target region.

•Final elongation: This single step is optional, but is performed at a temperature of 70–74 °C (158–165 °F) (the temperature range required for optimal activity of most polymerases used in PCR) for 5–15 minutes after the last PCR cycle to ensure that any remaining single- stranded DNA is fully elongated.

•Final hold: The final step cools the reaction chamber to 4–15 °C (39–59 °F) for an indefinite time, and maybe employed for short-term storage of the PCR products.



Gene amplification

Gene amplification refers to a number of natural and artificial processes by which the number of copies of a gene is increased "without a proportional increase in other genes".

Artificial DNA amplification research or diagnosis DNA amplification can be conducted

through methods such as:

•Polymerase chain reaction, an easy, cheap, and reliable way to repeatedly replicate a focused segment of DNA by polymerizing nucleotides, a concept which is applicable to numerous fields in modern biology and related sciences.

•Ligase chain reaction, a method that amplifies the nucleic acid used as the probe. For each of the two DNA strands, two partial probes are ligated to form the actual one; thus, LCR uses two enzymes: a DNA polymerase (used for initial template amplification and then inactivated) and a thermostable DNA ligase.

Transcription-mediated amplification, an isothermal, single-tube nucleic acid amplification system utilizing two enzymes, RNA polymerase and reverse transcriptase, to rapidly amplify the target RNA/DNA, enabling the simultaneous detection of multiple pathogenic organisms in a single tube.

Natural DNA amplification:

DNA replication is a natural form of copying DNA with the amount of genes remaining constant. However, the amount of DNA or the number of genes can also increase within an organism through gene duplication, a major mechanism through which new genetic material is generated during molecular evolution. Common sources of gene duplications include ectopic recombination, retrotransposition event, aneuploidy, polyploidy, and replication slippage.

A piece of DNA or RNA that is the source and/or product of either natural or artificial amplification or replication events is called an amplicon.

DNA Sequencing methods

DNA sequencing is the process of determining the nucleic acid sequence – the order of nucleotides in DNA. It includes any method or technology that is used to determine the order of the four bases: adenine, guanine, cytosine, and thymine.

Maxam-Gilbert sequencing method

.Maxam–Gilbert sequencing is a method of DNA sequencing developed by Allan Maxam and Walter Gilbert in 1976–1977. This method is based on nucleobase-specific partial chemical modification of DNA and subsequent cleavage of the DNA backbone at sites adjacent to the modified nucleotides.

An example Maxam–Gilbert sequencing reaction. Cleaving the same tagged segment of DNA at different points yields tagged fragments of different sizes. The fragments may then be separated by gel electrophoresis.

Maxam–Gilbert sequencing was the first widely adopted method for DNA sequencing, and, along with the Sanger dideoxy method, represents the first generation of DNA sequencing methods. Maxam–Gilbert sequencing is no longer in widespread use, having been supplanted by next-generation sequencing methods.

History

Although Maxam and Gilbert published their chemical sequencing method two years after Frederick Sanger and Alan Coulson published their work on plus-minus sequencing,Maxam–Gilbert sequencing rapidly became more popular, since purified DNA could be used directly, while the initial Sanger method required that each read start be cloned for production of single-stranded DNA. However, with the improvement of the chain-termination method (see below), Maxam–Gilbert sequencing has fallen out of favour due to its technical complexity prohibiting its use in standard molecular biology kits, extensive use of hazardous chemicals, and difficulties with scale-up.

Allan Maxam and Walter Gilbert's 1977 paper "A new method for sequencing DNA" was honored by a Citation for Chemical Breakthrough Award from the Division of History of Chemistry of the American Chemical Society for 2017. It was presented to the Department of Molecular & Cellular Biology, Harvard University.

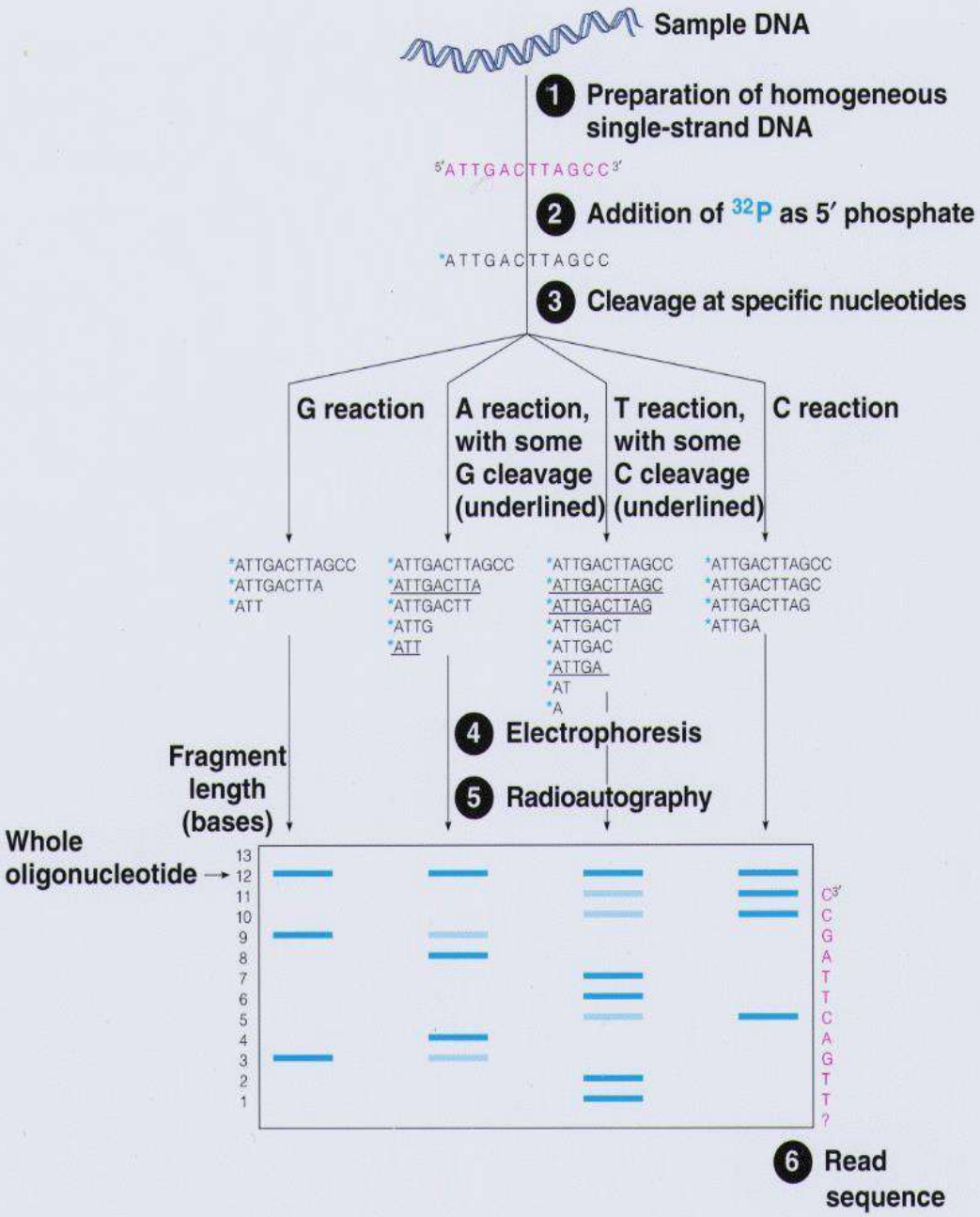
Procedure

Maxam–Gilbert sequencing requires radioactive labeling at one 5′ end of the DNA fragment to be sequenced (typically by a kinase reaction using gamma-^{32}P ATP) and purification of the DNA.

Chemical treatment generates breaks at a small proportion of one or two of the four nucleotide bases in each of four reactions (G, A+G, C, C+T). For example, the purines (A+G) are depurinated using formic acid, the guanines (and to some extent the adenines) are methylated by dimethyl sulfate, and the pyrimidines (C+T) are hydrolysed using hydrazine. The addition of salt (sodium chloride) to the hydrazine reaction inhibits the reaction of thymine for the C-only reaction. The modified DNAs may then be cleaved by hot piperidine; $(CH_2)_5NH$ at the position of the modified base. The concentration of the modifying chemicals is controlled to introduce on average one modification per DNA molecule. Thus a series of labeled fragments is generated, from the radiolabeled end to the first "cut" site in each molecule.

The fragments in the four reactions are electrophoresed side by side in denaturing acrylamide gels for size separation. To visualize the fragments, the gel is exposed to X-ray film for autoradiography, yielding a series of dark bands each showing the location of identical radiolabeled DNA molecules. From presence and absence of certain fragments the sequence may be inferred.

Figure 4A.4 Sequencing an oligonucleotide by the Maxam-Gilbert method



From Mathews and van Holde: *Biochemistry 2/e.* © The Benjamin/Cummings Publishing Co., Inc.

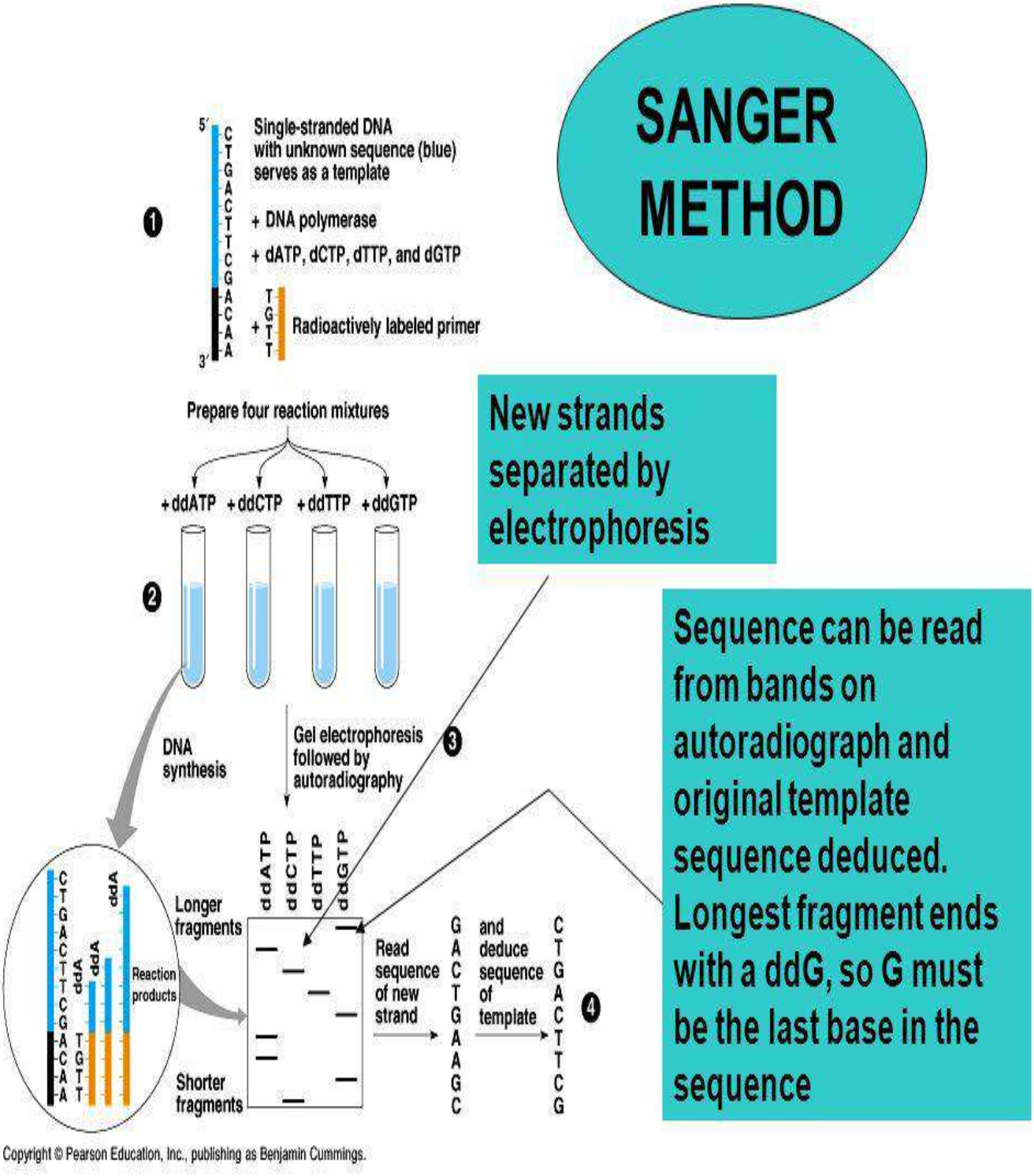
0

Sangers Sequencing method (chain-termination method)

The chain-termination method developed by Frederick Sanger and coworkers in 1977 soon became the method of choice, owing to its relative ease and reliability. When invented, the chain-termination method used fewer toxic chemicals and lower amounts of radioactivity than the Maxam and Gilbert method. Because of its comparative ease, the Sanger method was soon automated and was the method used in the first generation of DNA sequencers.

- Sanger's method of gene sequencing is also known as dideoxy chain termination method. It generates nested set of labelled fragments from a template strand of DNA to be sequenced by replicating that template strand and interrupting the replication process at one of the four bases.
- Four different reaction mixtures are produced that terminates in A. T. G or C respectively.

SANGER METHOD

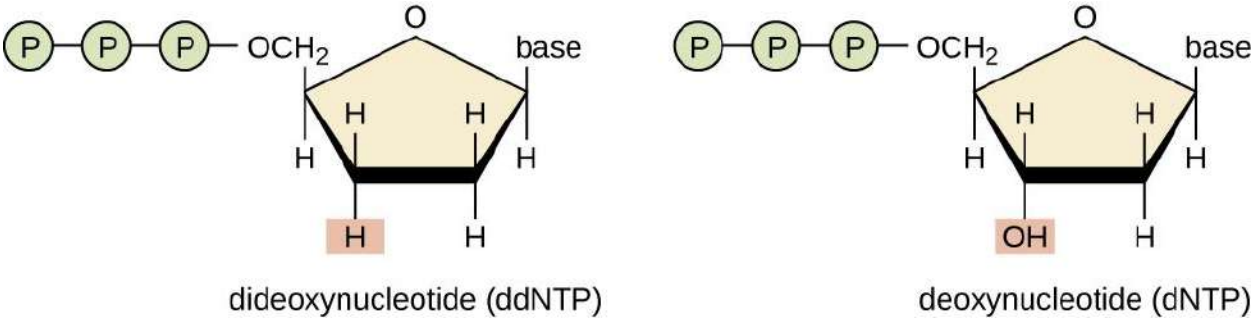


New strands separated by electrophoresis

Sequence can be read from bands on autoradiograph and original template sequence deduced. Longest fragment ends with a ddG, so G must be the last base in the sequence

Principle

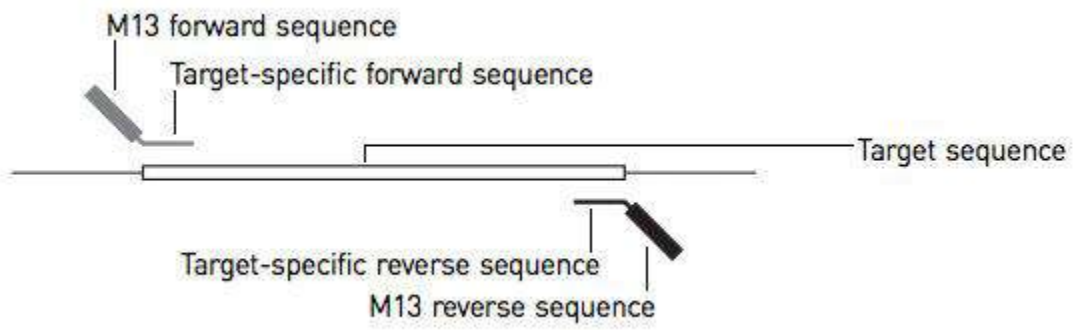
- A DNA primer is attached by hybridization to the template strand and deoxynucleosides triphosphates (dNTPPs) are sequentially added to the primer strand by DNA polymerase.
- The primer is designed for the known sequences at 3' end of the template strand.
- M13 sequences is generally attached to 3' end and the primer of this M13 is made.
- The reaction mixture also contains dideoxynucleoside triphosphate (ddNTPs) along with usual dNTPs.
- If during replication ddNTPs is incorporated instead of usual dNTPs in thegrowing DNA strand then the replication stops at that nucleotide.



- The ddNTPs are analogue of dNTPs
- ddNTPs lacks hydroxyl group (-OH) at c3 of ribose sugar, so it cannot make phosphodiester bond with nest nucleotide, thus terminates the nucleotide chain
- Respective ddNTPs of dNTPs terminates chain at their respective site. For example ddATP terminates at A site. Similarly ddCTP, ddGTP and ddTTP terminates at C, G and T site respectively.

Procedure

1. Template preparation:

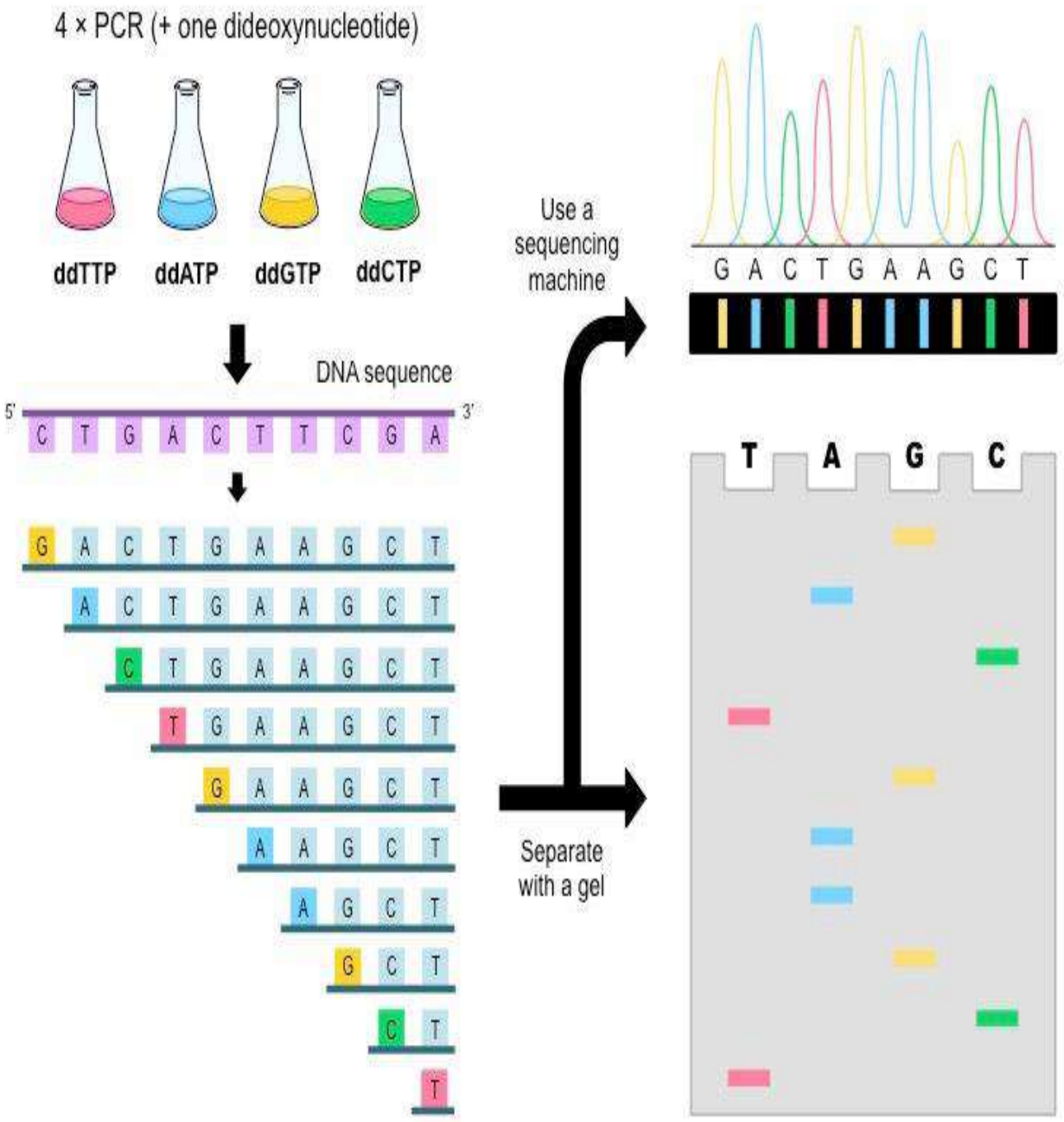


- Copies of template strand to be sequenced must be prepared with short known sequences at 3' end of the template strand.
- A DNA primere is essential to initiate replication of template , so primer preparation of known sequences at 3'end is always required.
- For this purpose a single stranded cloning vector M13 is flanked with template strand at 3'end which serves as binding site for primer.

2. Generation of nested set of labelled fragments:

- Copies of each template is divided into four batches and each batch is used for different replication reaction.
- Copies of standard primer and DNA polymerase I are used in all four batches.
- To synthesize fragments that terminates at A, ddATP is added to the reaction mixture on batch I along with dATP, dTTP,dCTP and dGTP, standard primer and DNA polymerase I.

Similarly, to generate, all fragments that terminates at C, G and T, the respective ddNTPs ie ddCTP,ddGTP and ddTTP are added respectively to different reaction mixture on different batch along with usual dNTPs.



Electrophoresis and gel reading:

- The reaction mixture from four batches are loaded into four different well on polyacrylamide gel and electrophoresed.
- The autoradiogramof the gel is read to determine the order of bases of complementary strand to that of template strand.

- The band of shortest fragments are at the bottom of autoradiogram so that the sequences of complementary strand is read from bottom to top.

DNA fingerprinting

DNA fingerprinting is a method used to identify living things based on samples of their DNA. Instead of looking at the whole sequence of a person's DNA, these techniques look at the presence or absence of common markers that can be quickly and easily identified.

DNA Fingerprinting Steps

Extracting DNA from Cells:

To perform DNA fingerprinting, you must first have a DNA sample! In order to procure this, a sample containing genetic material must be treated with different chemicals. Common sample types used today include blood and cheek swabs.

These samples must be treated with a series of chemicals to break open cell membranes, expose the DNA sample, and remove unwanted components – such as lipids and proteins – until relatively pure DNA emerges.

PCR Amplification (Optional):

If the amount of DNA in a sample is small, scientists may wish to perform PCR – Polymerase Chain Reaction – amplification of the sample.

PCR is an ingenious technology which essentially mimics the process of DNA replication carried out by cells. Nucleotides and DNA polymerase enzymes are added, along with "primer" pieces of DNA which will bind to the sample DNA and give the polymerases a starting point.

PCR "cycles" can be repeated until the sample DNA has been copied many times in the lab if necessary.

Treatment with Restriction Enzymes:

The best markers for use in quick and easy DNA profiling are those which can be reliably identified using common restriction enzymes, but which vary greatly between individuals.

For this purpose, scientists use repeat sequences – portions of DNA that have the same sequence so they can be identified by the same restriction enzymes, but which repeat a different number of times in different people. Types of repeats used in DNA profiling include Variable Number Tandem Repeats (VNTRs), especially short tandem repeats (STRs), which are also referred to by scientists as "microsatellites" or "minisatellites."

Once sufficient DNA has been isolated and amplified, if necessary, it must be cut with restriction enzymes to isolate the VNTRs. Restriction enzymes are enzymes that attach to specific DNA sequences and create breaks in the DNA strands.

In genetic engineering, DNA is cut up with restriction enzymes and then "sewn" back together by ligases to create new, recombinant DNA sequences. In DNA profiling, however, only the cutting part is needed. Once the DNA has been cut to isolate the VNTRs, it's time to run the resulting DNA fragments on a gel to see how long they are!

Gel Electrophoresis:

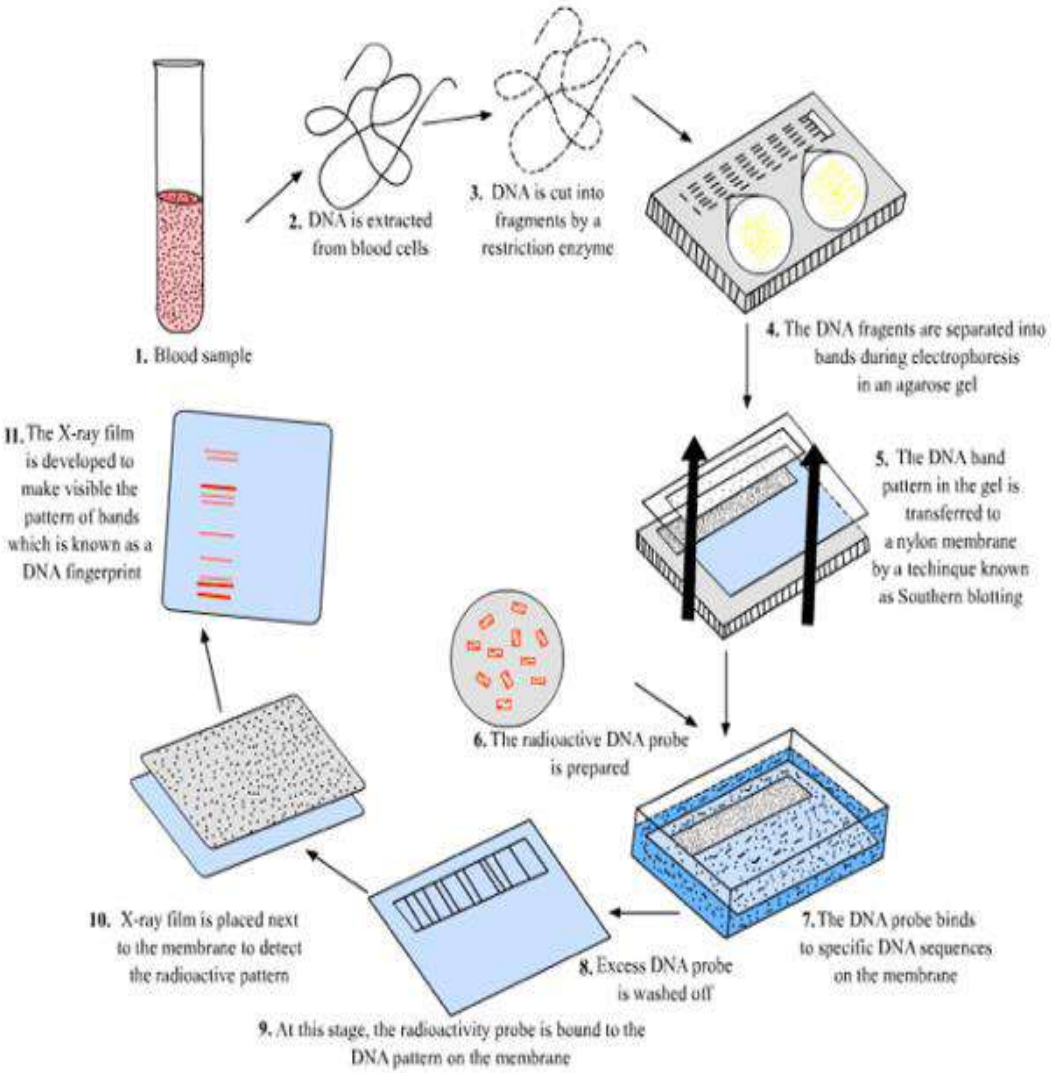
Gel electrophoresis is a brilliant technology that separates molecules by size. The "gel" in question is a material that molecules can pass through, but only at a slow speed.

Just as air resistance slows a big truck more than it does a motorcycle, the resistance offered by the electrophoresis gel slows large molecules down more than small ones. The effect of the gel is so precise that scientists can tell exactly how big a molecule is by seeing how far it moves within a given gel in a set amount of time.

In this case, measuring the size of the DNA fragments from the sample that has been treated with a restriction enzyme will tell scientists how many copies of each VTNR repeat the sample DNA contains.

It's called "electrophoresis" because, to make the molecules move through the gel, an electrical current is applied. Because the sugar-phosphate backbone of the DNA has a negative electrical charge, the electrical current tugs the DNA along with it through the gel.

By looking at how many DNA fragments the restriction enzymes produced and the sizes of these fragments, the scientists can "fingerprint" the DNA donor.



Steps in DNA fingerprinting techinque

Transfer onto Southern Blot:

Now that the DNA fragments have been separated by size, they must be transferred to a medium where scientists can “read” and record the results of the electrophoresis.

To do this, scientists treat the gel with a weak acid, which breaks up the DNA fragments into individual nucleic acids that will more easily rub off onto paper. They then “blot” the DNA fragments onto nitrocellulose paper, which fixes them in place.

Treatment with Radioactive Probe:

Now that the DNA is fixed onto the blotting paper, it is treated with a special probe chemical that sticks to the desired DNA fragments. This chemical is radioactive, which means that it will create a visible record when exposed to X-ray paper.

X-Ray Film Exposure:

The last step of the process is to turn the information from the DNA fragments into a visible record. This is done by exposing the blotting paper, with its radioactive DNA bands, to X-ray film.