BIOTECHNOLOGY

Gene cloning

Gene cloning involves separation of specific gene or DNA fragments from a donor cell, attaching it to small carrier molecule called vector and then replicating this recombinant vector into a host cell.

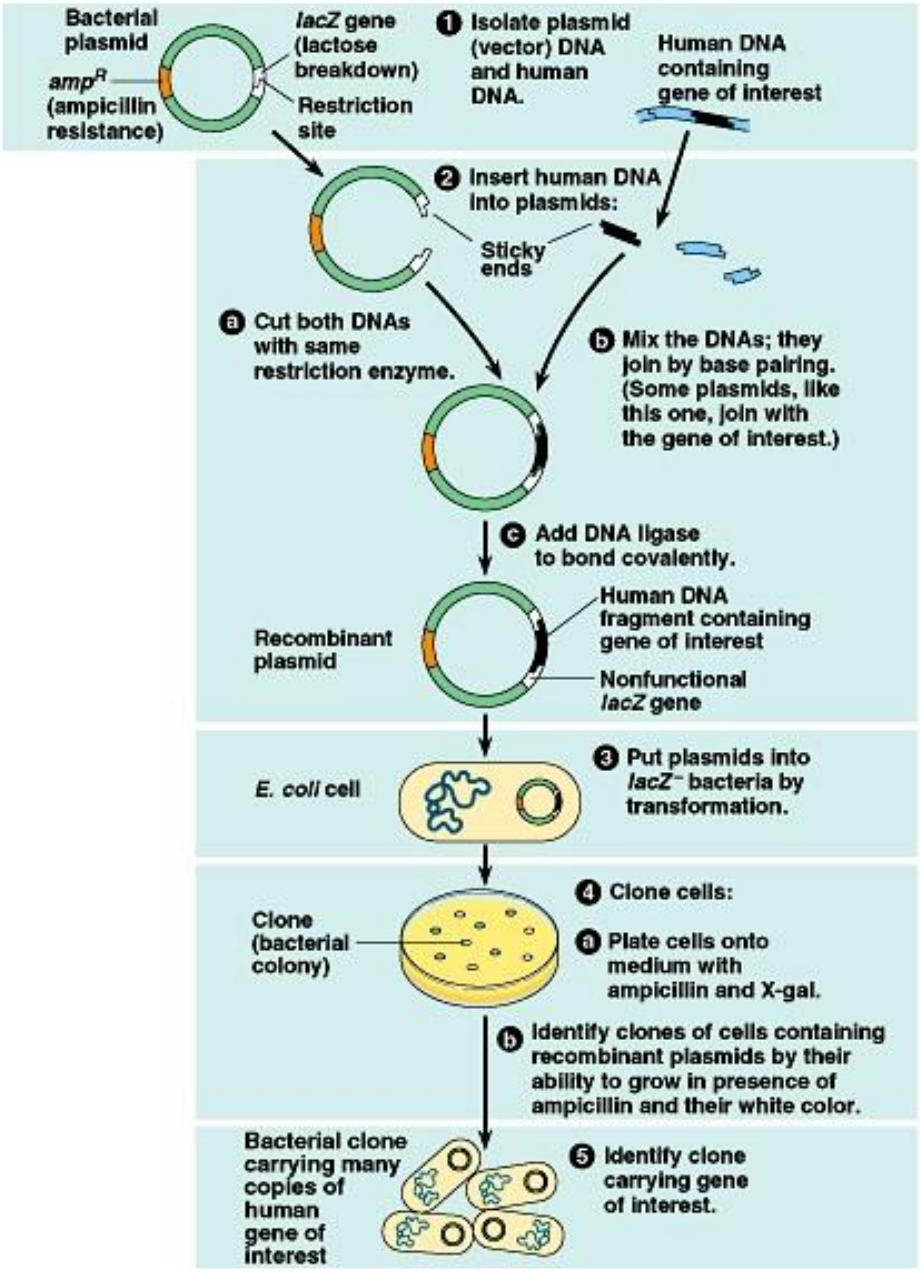
Major steps in cloning

Steps involved in gene cloning:

1. **Isolation of donor DNA fragment or gene**
2. **Selection of suitable vector**
3. **Incorporation of donor DNA fragment into the vector**
4. **Transformation of recombinant vector into a suitable host cell**
5. **Isolation of recombinant host cell**

1. Isolation of donor DNA fragment or gene

- At first a donor DNA fragment should be isolated. There are two method for isolation of desired gene or DNA fragment.
- Using restriction endonuclease enzyme: the enzyme restriction endonuclease is a key enzyme in molecular gene cloning. It has specific restriction site for its action. The enzyme RE generates a DNA fragment either with blunt end or with sticky end
- Using reverse transcriptase enzyme: reverse transcriptase enzyme Synthesizes complementary DNA strand of the desired gene using its mRNA.



---

2. Selection of suitable cloning vector:

- When donor DNA fragment is incorporated into a host cell, it will not replicates because the isolated gene do not have the capacity to replicated itself. So before introduction of donor fragment into host, a suitable vector should be selected.

- Cloning vector is the DNA molecule capable of self-replication inside the host cell. the main function of cloning vector is to replicates the inserted DNA fragment inside the host cell.
- Examples of cloning vectors: Plasmid, BAC, YAC, Λ-bacteriophase, expression vectors etc.
- Characteristics of a cloning vectors
    - It must be self-replicating inside host cell
    - It must possess restriction site for RE enzymes
    - Introduction of donor DNA fragment must not interfere with replication property of the vector
    - It must possess some marker gene such that it can be used for later identification of recombinant cell.

3. Incorporation of donor DNA fragment with Plasmid vector:

- The plasmid vector is cut open by the same RE enzyme used for isolation of donor DNA fragment
- The mixture of donor DNA fragment and plasmid vector are mixed together.
- In the presence of DNA ligase, base pairing of donor DNA fragment and plasmid vector occurs forming recombinant vector in the mixture.

4. Transformation of recombinant vector into suitable host:

- The recombinant vector is transformed into suitable host cell. ie bacterial cell
- Some bacteria are naturally transformable, they take up the recombinant vector automatically. For examples: Bacillus, Haemophillus, Helicobacter pylori, are naturally competent
- Some other bacteria are not naturally competent, in those bacteria recombinant vector are incorporated by artificial method such as Ca++ ion treatment, electroporation etc.

5. Isolation of recombinant cell:

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

Vectors- properties of an ideal vector, different types

A vector is a DNA molecule that has the abil-ity to replicate autonomously in an appropri-ate host cell and into which the gene of inter-est (a foreign genetic sequence) is integrated. When we insert a foreign genetic sequence into the vector the aim is either to obtain numer-ous copies of the gene of interest or to obtain the product of that.

Accordingly the design and features of the vectors used for the assigned activity vary. Due to this there are two types of vectors - the expression vectors and clon-ing vectors.

Characteristic Features of an Ideal Vec-tor:

1. It should be able to replicate autono-mously. When the objective of cloning is to obtain a large number of copies of the DNA insert, the vector replication must be under relaxed control so that it can gener-ate multiple copies of itself in a single host cell.

2. It should be easy to isolate and purify.
3. It should be easily introduced into the host cells, i.e., transformation of the host with the vector should be easy.
4. The vector should have suitable marker genes that allow easy detection and/or se-lection of the transformed host cells.
5. When the objective is gene transfer, it should have the ability to integrate either itself or the DNA insert it carries into the genome of the host cell.
6. The cells transformed with the vector con-taining the DNA insert (recombinant DNA) should be identifiable and selectable from those transformed by the unaltered vector.
7. A vector should contain unique target sites for as many restriction enzymes as pos-sible into which the DNA insert can be in-tegrated.
8. When expression of the DNA insert is de-sired, the vector should contain at least suitable control elements, e.g., promoter, operator and ribosome binding sites.

Types of vector

Vectors are of two types:

- Cloning vector
- Expression vector

Cloning vectors

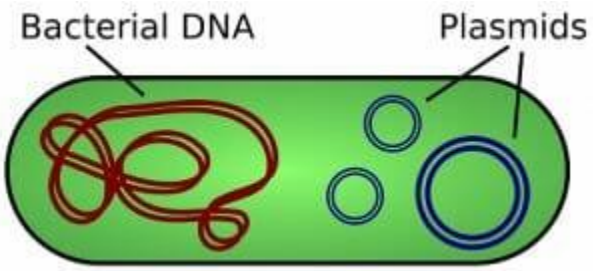
- Cloning vectors are small piece of DNA which have the ability and used to introduce foreign gene of interest into the host cell.
- They can be stably maintained insides the host cell.
- Cloning vector are generally used to obtain multiple copies of desired foreign gene.
- **Example-** Plasmid, Cosmid and Phages, BACs, YACs.
- These type of vectors generally contains selectable marker, origin of replication and a restriction site.

Expression vector

- Expression vector is a type of vector which not only introduces a gene of interest into the host cell but also aids in the analysis of the foreign gene via relevant protein product expression.
- It is type of vector which is used to obtain or analyses the gene product, which may be RNA or protein of the inserted desired gene.
- **Example-** Only plasmid vector.
- Expression vector contains enhancer, promoter region, start/stop codon, transcription initiation, selectable marker, ori sites, and restriction site.

Plasmids

- A plasmid is a small, circular piece of DNA that is different than the chromosomal DNA, which is all the genetic material found in an organism's chromosomes.
- It replicates independently of chromosomal DNA.
- Plasmids are mainly found in bacteria, but they can also be found in archaea and multicellular organisms.
- Plasmids usually carry at least one gene, and many of the genes that plasmids carry are beneficial to their host organisms.
- Although they have separate genes from their hosts, they are not considered to be independent life.



Functions of Plasmids

- Plasmids have many different functions.

- They may contain genes that enhance the survival of an organism, either by killing other organisms or by defending the host cell by producing toxins.
- Some plasmids facilitate the process of replication in bacteria.
- Since plasmids are so small, they usually only contain a few genes with a specific function (as opposed to a large amount of noncoding DNA).
- Multiple plasmids can coexist in the same cell, each with different functions.

General Types of Plasmids

- **Conjugative and Non-Conjugative**
  - There are many ways to classify plasmids from general to specific.
  - One way is by grouping them as either conjugative or non-conjugative.
  - Bacteria reproduce by sexual conjugation, which is the transfer of genetic material from one bacterial cell to another, either through direct contact or a bridge between the two cells.
  - Some plasmids contain genes called transfer genes that facilitate the beginning of conjugation.
  - Non-conjugative plasmids cannot start the conjugation process, and they can only be transferred through sexual conjugation with the help of conjugative plasmids.

Incompatibility

- Another plasmid classification is by incompatibility group.
- In a bacterium, different plasmids can only co-occur if they are compatible with each other.
- An incompatible plasmid will be expelled from the bacterial cell.

- Plasmids are incompatible if they have the same reproduction strategy in the cell; this allows the plasmids to inhabit a certain territory within it without other plasmids interfering.

Specific Types of Plasmids

There are five main types of plasmids:

- **Fertility F-plasmids**
- **Resistance plasmids**
- **Virulence plasmids**
- **Degradative plasmids**
- **Col plasmids**

- **Fertility F-plasmids**
  - Fertility plasmids, also known as F-plasmids, contain transfer genes that allow genes to be transferred from one bacteria to another through conjugation.
  - These make up the broad category of conjugative plasmids.
  - F-plasmids are episomes, which are plasmids that can be inserted into chromosomal DNA.
  - Bacteria that have the F-plasmid are known as F positive (F+), and bacteria without it are F negative (F–).
  - When an F+ bacterium conjugates with an F– bacterium, two F+ bacterium result.
  - There can only be one F-plasmid in each bacterium.

- **Resistance Plasmids**
  - Resistance or R plasmids contain genes that help a bacterial cell defend against environmental factors such as poisons or antibiotics.
  - Some resistance plasmids can transfer themselves through conjugation.

- When this happens, a strain of bacteria can become resistant to antibiotics.
- Recently, the type bacterium that causes the sexually transmitted infection gonorrhea has become so resistant to a class of antibiotics called quinolones that a new class of antibiotics, called **cephalosporins**, has started to be recommended by the World Health Organization instead.
- The bacteria may even become resistant to these antibiotics within five years.
- According to NPR, overuse of antibiotics to treat other infections, like urinary tract infections, may lead to the proliferation of drug-resistant strains.

- **Virulence Plasmids**
  - When a virulence plasmid is inside a bacterium, it turns that bacterium into a pathogen, which is an agent of disease.
  - Bacteria that cause disease can be easily spread and replicated among affected individuals.
  - The bacterium Escherichia coli (E. coli) has several virulence plasmids. E. coli is found naturally in the human gut and in other animals, but certain strains of E. coli can cause severe diarrhea and vomiting.
  - Salmonella enterica is another bacterium that contains virulence plasmids.

- **Degradative Plasmids**
  - Degradative plasmids help the host bacterium to digest compounds that are not commonly found in nature, such as camphor, xylene, toluene, and salicylic acid.
  - These plasmids contain genes for special enzymes that break down specific compounds.

    - Degradative plasmids are conjugative.

- **Col Plasmids**
    - Col plasmids contain genes that make bacteriocins (also known as colicins), which are proteins that kill other bacteria and thus defend the host bacterium.
    - Bacteriocins are found in many types of bacteria including E. coli, which gets them from the plasmid ColE1.

Applications of Plasmids

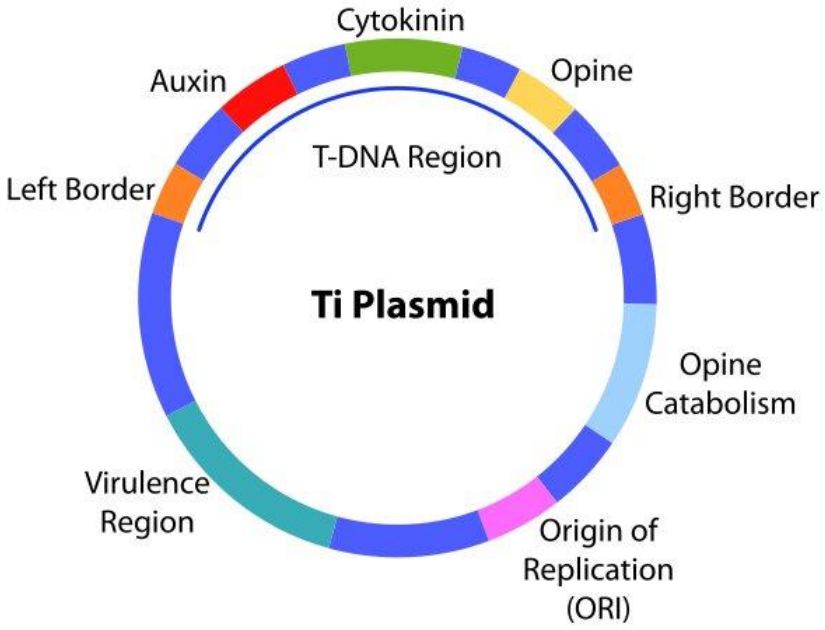
- Humans have developed many uses for plasmids and have created software to record the DNA sequences of plasmids for use in many different techniques.
- Plasmids are used in genetic engineering to amplify, or produce many copies of, certain genes.
- In molecular cloning, a plasmid is a type of vector.
- A vector is a DNA sequence that can transport foreign genetic material from one cell to another cell, where the genes can be further expressed and replicated.
- Plasmids are useful in cloning short segments of DNA. Also, plasmids can be used to replicate proteins, such as the protein that codes for insulin, in large amounts.
- Additionally, plasmids are being investigated as a way to transfer genes into human cells as part of gene therapy.
- Cells may lack a specific protein if the patient has a hereditary disorder involving a gene mutation.
- Inserting a plasmid into DNA would allow cells to express a protein that they are lacking.

Ti plasmid

- The **Ti-plasmid** in the bacteria is known to induce crown gall disease in plants by transferring crucial regions from the plasmid.
- These crucial regions were seen to modify the plant cells into a tumour to produce synthetic plant hormones and cause crown gall.
- This led the scientists to believe that there is a scope for bioengineering techniques to modify the plants using Ti-plasmid for our own use.

Features of Ti Plasmid

- **Virulence Region:** The virulence region codes for virulence genes that are responsible for the transfer of T-DNA to the plant cells and also recruiting various effector proteins for infecting the plant cells.



- **T-DNA:** The T-DNA region is the crucial region that gets transferred to the plant cell for infection. It is approximately 15-20 kbp in length and is transferred to the plant cell via means of genetic recombination.
- **Opine Catabolism:** The opine catabolism region is the region from where the bacteria sources its nutrients for the whole process. Opines are derivatives of amino acid or sugar phosphates that can be catabolized

to use in the form of nutrients. The types of opines found in Ti-plasmid are nopaline and octopine types.

- **Origin of Replication:** The origin of replication is the region where replication of the plasmid is initiated.

Uses in Bioengineering

- The ability of the Ti-plasmid to modify the plant cells has been taken advantage of for the production of transgenic plants.
- The plasmids have been modified into a cloning vector now which are no more pathogenic to plants.
- They are being used to transfer genes of our interest into the plant and produce plants with better quality and quantity.
- It is therefore known as 'nature's genetic engineer'.

Bacteriophage

- A bacteriophage is a virus that infects a bacterial cell and reproduces inside it.
- They vary a lot in their shape and genetic material.
- A bacteriophage may contain DNA or RNA.
- The genes range from four to several thousand.
- Their capsid can be isohedral, filamentous, or head-tail in shape.

Bacteriophage Structure

The bacteriophage consists of a polyhedral head, a short collar and a helical tail.

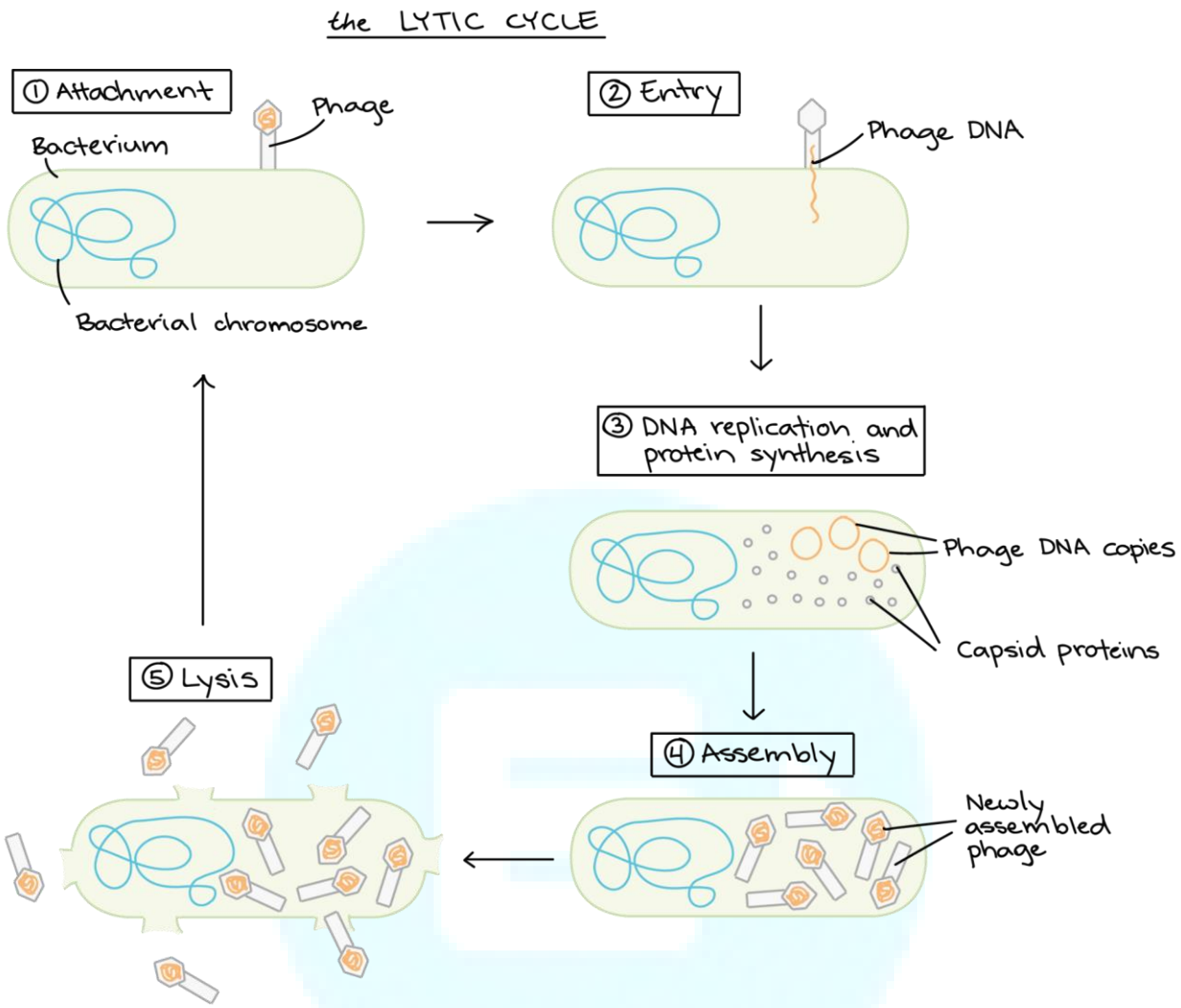
- **Head-** The head consists of 2000 capsomeres with double-stranded DNA enclosed within.
- **Tail-** The tail consists of an inner hollow tube which is surrounded by a contractile sheath with 24 annular rings. The distal end consists of a basal plate with tail fibres at each corner. The bacteriophage attaches to the bacteria with the help of these tail fibres.

Bacteriophage Life Cycle

Bacteriophage exhibits two major types of life cycles:

- **Lytic Cycle or Virulent Cycle**
- **Lysogenic Cycle or Temperate Cycle**

- **Lytic cycle**
  - In the lytic cycle, a phage acts like a typical virus: it hijacks its host cell and uses the cell's resources to make lots of new phages, causing the cell to lyse (burst) and die in the process.

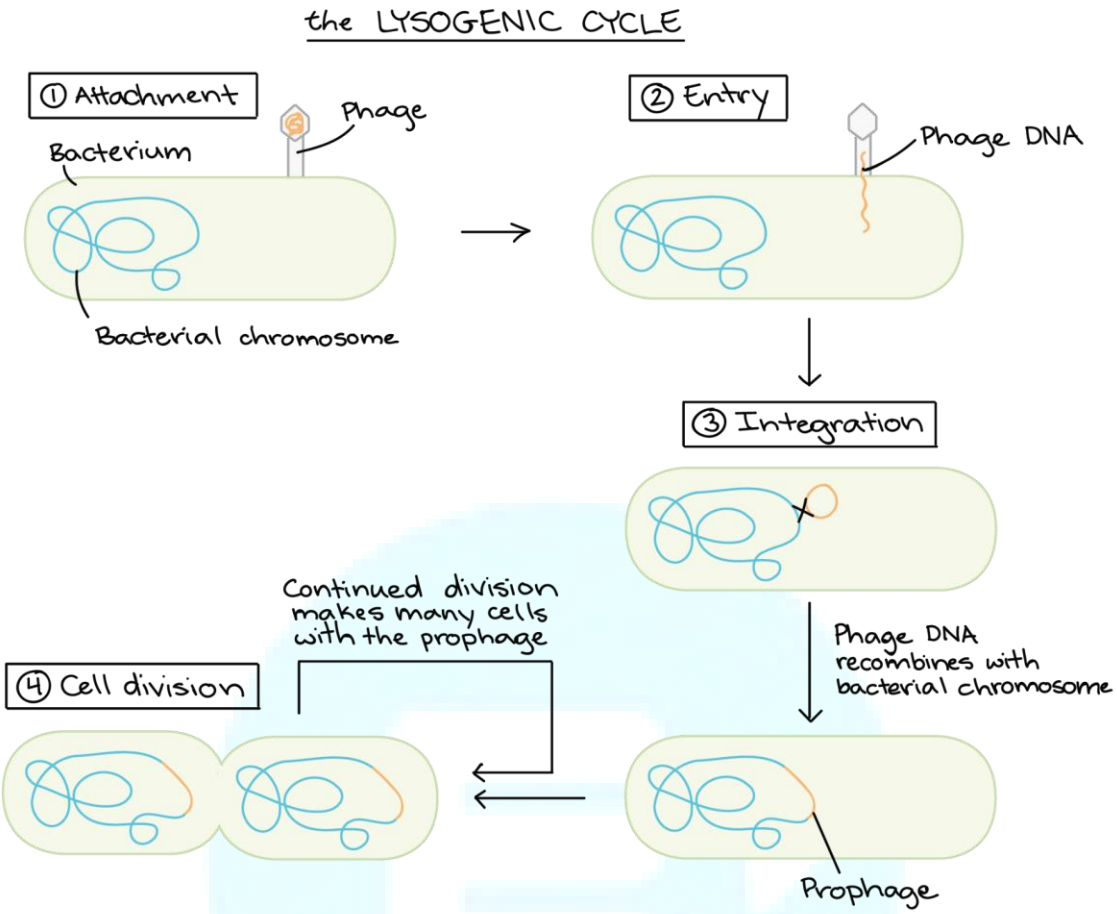


- <u>The stages of the lytic cycle are:</u>
  - **Attachment:** Proteins in the "tail" of the phage bind to a specific receptor (in this case, a sugar transporter) on the surface of the bacterial cell.
  - **Entry:** The phage injects its double-stranded DNA genome into the cytoplasm of the bacterium.
  - **DNA copying and protein synthesis:** Phage DNA is copied, and phage genes are expressed to make proteins, such as capsid proteins.

    - **Assembly of new phage:** Capsids assemble from the capsid proteins and are stuffed with DNA to make lots of new phage particles.
    - **Lysis:** Late in the lytic cycle, the phage expresses genes for proteins that poke holes in the plasma membrane and cell wall. The holes let water flow in, making the cell expand and burst like an overfilled water balloon.

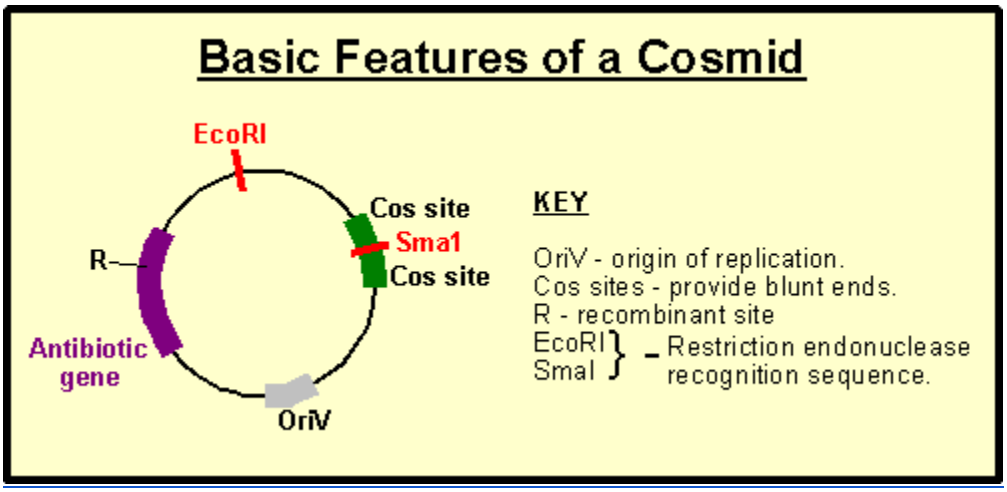
Cell bursting, or lysis, releases hundreds of new phages, which can find and infect other host cells nearby. In this way, a few cycles of lytic infection can let the phage spread like wildfire through a bacterial population.

- **Lysogenic cycle**
    - The lysogenic cycle allows a phage to reproduce without killing its host.
    - Some phages can only use the lytic cycle, but the phage we are following, lambda (\lambdaλlambda), can switch between the two cycles.
    - In the lysogenic cycle, the first two steps (attachment and DNA injection) occur just as they do for the lytic cycle.
    - However, once the phage DNA is inside the cell, it is not immediately copied or expressed to make proteins.
    - Instead, it recombines with a particular region of the bacterial chromosome.
    - This causes the phage DNA to be integrated into the chromosome.



- The integrated phage DNA, called a **prophage**, is not active: its genes aren't expressed, and it doesn't drive production of new phages.
- However, each time a host cell divides, the prophage is copied along with the host DNA, getting a free ride.
- The lysogenic cycle is less flashy (and less gory) than the lytic cycle, but at the end of the day, it's just another way for the phage to reproduce.
- Under the right conditions, the prophage can become active and come back out of the bacterial chromosome, triggering the remaining steps of the lytic cycle (DNA copying and protein synthesis, phage assembly, and lysis).

Cosmids



- A cosmid, first described by Collins and Hohn in 1978, is a type of hybrid plasmid with a bacterial "ori" sequence and a "cos" sequences derived from the lambda phage.
- It is formed by joining ends of a linearized plasmid DNA with cos-site of lambda DNA.
- It is a derived vector.
- The cosmid DNA can be packed in a capsid of lambda phage in vitro to form recombinant phage particles.
- It is linear inside the phage capsid.
- The cosmid gets circularized and behaves like a plasmid.
- Cosmid has an origin of replication, selectable markers, and gene cloning sites of plasmid DNA.
- They lack structural and regulatory genes of lambda DNA.
- Hence there is no lysis and integration of cosmid DNA in the host cell.
- Examples: **Col EI cosmid, pHC 79, pJB8, pWE cosmid, etc.**

Salient features of Cosmids

- Cosmid is a circular ds DNA
- It has two complementary single-stranded regions at both ends of a plasmid DNA. The two cos-ends form a duplex by base pairing.
- The cosmid DNA does not code for phage proteins and host cell lysis.
- It does not involve in ,multiplication of phage particles.

- It has an origin of replication from plasmid DNA for independent replication.
- It has selectable marker genes and gene cloning sites of plasmid DNA
- The cosmid DNA is packed within protein coat of bacteriophage to form inactive phage particles. Cos-site is a prerequsites for invitro packaging of cosmid in phage protein coat.
- After infection, the cosmid DNA does not integrate into host chromosomal DNA. It exits 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 lambda DNA.
- The plasmid derived portion contributes an origin of replication (Ori) and tetracycline resistance gene.
- There is a MCS between the orogin of replication and cos-site.
- A foreign DNA of upto 45 kbp is inserted into the MCS of pLR5 and the rDNA is packed into bacteriophage head in vitro.
- The phage thus formed delivers the DNA into E.coli while infecting the cell.

Cosmid pJB8

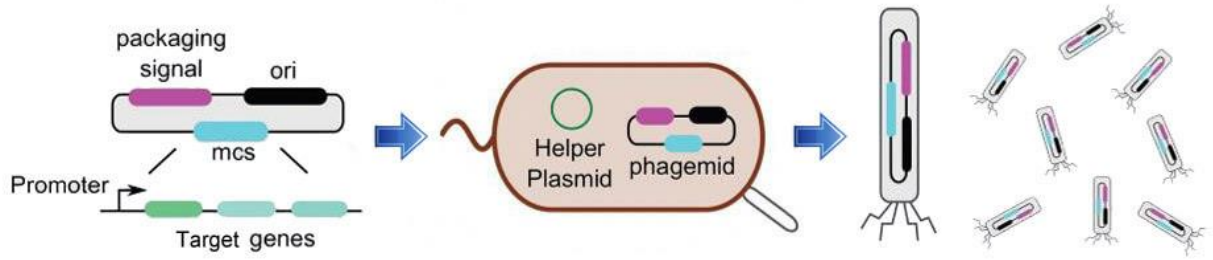
- pJB8 is constructed from the plasmid pBR322 and cos sites of lambda DNA.
- It is 5.4 kbp in size.
- It has an origin of replication and ampicillin resistance gene derived from pBR322 and two cos-ends from lambda DNA.
- A foreign DNA of about 45 kbp is inserted into BamHI or RcoRI or HindIII restriction site of the cosmid.

- The recombinant cosmid is packaged into lambda phage head to form an infective phage particle.
- The phage delivers its rDNA into E.coli while infecting the cell.

Cosmid pHC79

- pHC79 is constructed from pBR322 and cos-sites of lambda 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 and two marker genes- ampicillin resistance gene and tetracycline resistance gene.
- The derivative of the cosmid into lambda phage head.

Phagemid



- A phagemid is a normal plasmid which derived from a coliphage and a plasmid.
- It carries both a plasmid replication origin allowing isolation of double stranded plasmid from the cytoplasm of cells, and a replication origin usually from a single stranded phage such as f1, fd or M13 which allows the plasmid to enter a single strand replication mode in which only one of the strands is packaged into the virus particle when helper phage is added to the cell.
- Besides, the double stranded circular DNA of the plasmid form of the vector contains a variety of multiple cloning sites which make it

convenient for DNA recombination, gene manipulation and protein purification.

- Creative Biogene can offer you several phagemid vectors which can be used in the study of high-throughput screening of protein-protein interactions, selection of proteins with specific binding properties (high affinity binders), enzyme inhibitor screening, and ligand screening etc.

Enzymes in gene cloning

The seven types of enzymes are:

1. **Alkaline Phosphatase**
2. **Terminal Transferase**
3. **Thermo-stable DNA Polymerases**
4. **Bacteriophage RNA Polymerases**
5. **Nucleases: DNase and RNase**
6. **Polynucleotide Kinase and**
7. **DNA Ligase**

1. **Alkaline Phosphatase:**
    - Alkaline phosphatase removes 5′ phosphate groups from DNA and RNA.
    - It will also remove phosphates from nucleotides and proteins.
    - These enzymes are most active at alkaline pH therefore known as **alkaline phosphatase**.

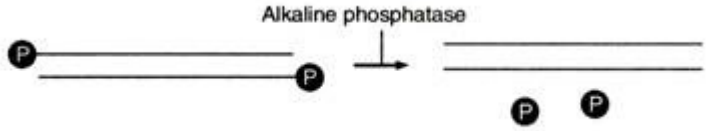


Fig. 13.4. Removal of phosphate by alkaline phosphatase.

There are several sources of alkaline phosphatase that differ in how easily they can in-activated:

- Bacterial alkaline phosphatase (BAP) is the most active of the enzymes, but also the most difficult to destroy at the end of the dephosphorylation reaction.
- Calf intestinal alkaline phosphatase (CIP) is purified from bovine intestine. This is phosphatase most widely used in molecular biology labs because, although less active than BAP, it can be effectively destroyed by protease digestion or heat (75 °C for 10 minutes in the presence of 5 mM EDTA).
- Shrimp alkaline phosphatase is derived from a cold-water shrimp and is promoted for being readily destroyed by heat (65°C for 15 minutes).

There are two primary uses for alkaline phosphatase in DNA manipulations:

- Removing 5′ phosphates from plasmid and bacteriophage vectors that have been cut with a restriction enzyme. In subsequent ligation reactions, this treatment prevents self-ligation of the vector and thereby greatly facilitates ligation of other DNA fragments into the vector (e.g. sub-cloning).
- Removing 5′ phosphates from fragments of DNA prior to labeling with radioactive phosphate. Polynucleotide kinase is much more effective in phosphorylating DNA if the 5′ phosphate has previously been removed.

It is usually recommended that dephosphorylation of DNAs with blunt or 5′-recessed ends be conducted using a higher concentration alkaline phosphatase or at higher temperatures than for DNAs with 5′ overhangs.

2. **Terminal Transferase:**
    - Terminal transferase catalyzes the addition of nucleotides to the 3′ terminus of DNA.

- Interestingly, it works on single-stranded DNA, including 3′ overhangs of double-stranded DNA, and is thus an example of a DNA polymerase that does not require a primer.
- It can also add homo-polymers of ribo-nucleotides to the 3′ end of DNA.
- The much preferred substrate for this enzyme is protruding 3′ ends, but it will also, less efficiently, add nucleotides to blunt and 3′-recessed ends of DNA fragments.
- Cobalt is a necessary cofactor for activity of this enzyme.

Some of its uses are:

- Labeling the 3′ ends of DNA:

  Most commonly, the substrate for this reaction is a fragment of DNA generated by digestion with a restriction enzyme that leaves a 3′ overhang, but oligodeoxynucleotides can also be used. When such DNA is incubated with tagged nucleotides and terminal transferase, a string of the tagged nucleotides will be added to the 3′ overhang or to the 3′ end of the oligonucleotide.

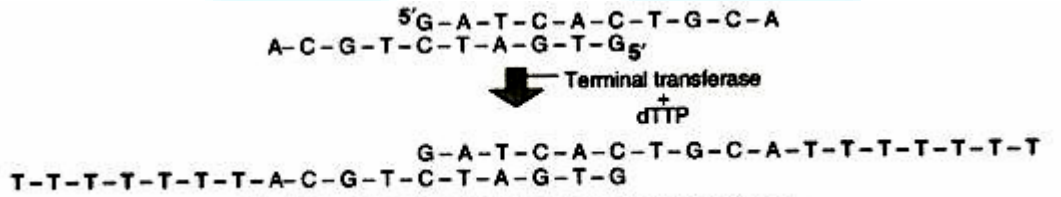


Fig. 13.5. Role of terminal transferase in adding poly (T).

- Adding Complementary Homopolymeric Tails to DNA:

  This clever procedure was commonly used in the past to clone cDNAs into plasmid vectors, but has largely been replaced by other, much more efficient techniques.

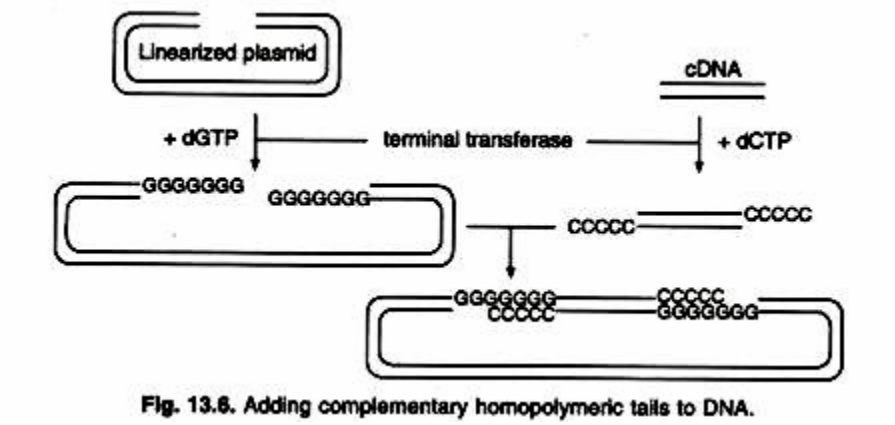


Fig. 13.6. Adding complementary homopolymeric tails to DNA.

Basically, terminal transferase is used to tail a linearized plasmid vector with G's and the cDNA with C's. When incubated together, the complementary G's and C's anneal to "insert" the cDNA into the vector, which is then transformed into E. coli.

Terminal transferase is a mammalian enzyme, expressed in lymphocytes. The enzyme purchased commercially is usually produced by expression of the bovine gene in E. coli.

3. Thermo-stable DNA Polymerases (TAQ Polymerase):

- It is interesting how some unimportant discoveries become something of immense practical importance after some time.
- Such is the history of Taq DNA polymerase.
- The original report of this enzyme, purified from the hot springs bacterium Thermus aquaticus, was published in 1976.
- Roughly 10 years later, the polymerase chain reaction was developed and shortly thereafter **"Taq"** became a household word in molecular biology circles.
- Currently, the world market for Taq polymerase is in the hundreds of millions of dollars each year.
- The thermophilic DNA polymerases, like other DNA polymerases, catalyze template-directed synthesis of DNA from nucleotide

triphosphates. A primer having a free 3′ hydroxyl is required to initiate synthesis and magnesium ion is necessary.

- In general, they have maximal catalytic activity at 75° to 80°C, and substantially reduced activates at lower temperatures.
- At 37°C, Taq polymerase has only about 10% of its maximal activity.
- In addition to Taq DNA polymerase, several other thermostable DNA polymerases have been isolated and expressed from cloned genes.

Table 13.2. Properties of commonly used polymerases.

Polymerase	3′ → 5′ Exonuclease	Source and Properties
Taq	No	From Thermus aquaticus. Half-life at 95°C is 1.6 hours.
Pfu	Yes	From Pyrococcus furiosus. Appears to have the lowest error rate of known thermophilic DNA polymerases.
Vent	Yes	From Thermococcus litoralis; also known as Tli polymerase. Half-life at 95°C is approximately 7 hours.

4. Bacteriophage RNA Polymerases:

- Phage-encoded DNA-dependent RNA polymerases are used for in vitro transcription to generate defined RNAs. Most commonly, the reaction utilizes ribo-nucleotides that are labeled with radio-nuclides or some other tag, and the resulting labeled RNA is used as a probe for hybridization.
- Other applications of in vitro transcription including making RNAs for in vitro translation or to study RNA struction and function.
- Several bacteriophage RNA polymerases are commercially available.
- They are named after the phage that encodes them, and either purified from phage-infected bacteria or produced as recombinant proteins.

- Many of the plasmids used for carrying cloned DNA incorporate promoters for bacteriophage RNA polymerases adjacent to the cloning site.
- This allows one to readily obtain either mRNA sense or antisense transcripts from the inserted DNA.
- The process is often called **run-off transcription**, because the plasmid is cut with a restriction enzyme downstream of the inserted DNA, which causes the polymerase to fall off the template when it reaches that spot.
- If we assume that the RNA transcribed has the polarity of a mRNA (e.g. sense), it is easy to modify the construct to express an antisense RNA – simply reverse the orientation of the transcribed region.
- Indeed, most plasmids used for in vitro transcription have two different phage polymerase promoters flanking the insertion site, which allows transcription of sense RNA with one polymerase and antisense with the other.

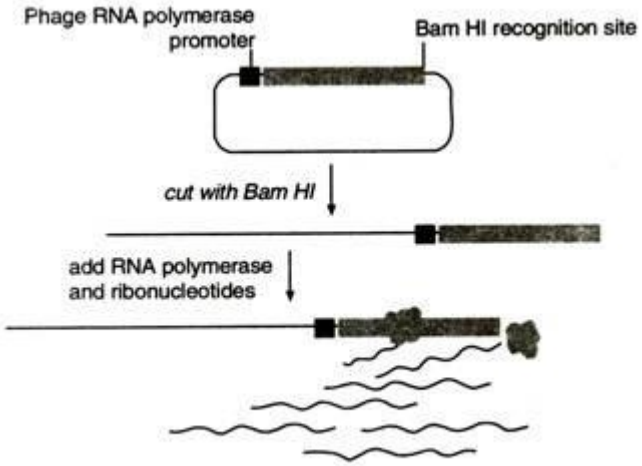


Fig. 13.8. RNA formation from inserted DNA.

5. **Nucleases: DNase and RNase:**
   - Most of the time nucleases are the enemy of the molecular biologist who is trying to preserve the integrity of RNA or DNA samples.

- However, deoxyribonucleases (DNases) and ribonucleases (RNases) have certain indispensible roles in molecular biology laboratories.
- Numerous types of DNase and RNase have been isolated and characterized.
- They differ among other things in substrate specificity, cofactor requirements, and whether they cleave nucleic acids internally (endonucleases), chew in from the ends (exonucleases) or attack in both of these modes.
- In many cases, the substrate specificity of a nuclease depends upon the concentration of enzyme used in the reaction, with high concentrations promoting less specific cleavages.
- The most widely used nucleases are DNase I and RNase A, both of which are purified from bovine pancreas: Deoxyribonuclease I cleaves double-stranded or single stranded DNA.
- Cleavage preferentially occurs adjacent to pyrimidine (C or T) residues, and the enzyme is therefore an endonuclease.
- Major products are 5′-phosphorylated di, tri and tetra nucleotides.
- In the presence of magnesium ions, DNase I hydrolyzes each strand of duplex DNA independently, generating random cleavages.
- In the presence of manganese ions, the enzyme cleaves both strands of DNA at approximately the same site, producing blunt ends or fragments with 1-2 base overhangs.
- DNase I does not cleave RNA, but crude preparations of the enzyme are contaminated with RNase A; RNase-free DNase I is readily available.

Applications:

- Eliminating DNA (e.g. plasmid) from preparations of RNA.
- Analyzing DNA-protein interactions via DNase foot printing.

- Nicking DNA prior to radio-labeling by nick translation.

Ribonuclease A is an endoribonuclease that cleaves single-stranded RNA at the 3′ end of pyrimidine residues. It degrades the RNA into 3′-phosphorylated mononucleotides and oligonucleotides. The major use of RNase A is eliminating or reducing RNA contamination in preparations of plasmid DNA.

6. **Polynucleotide Kinase:**
   - Polynucleotide kinase (PNK) is an enzyme that catalyzes the transfer of a phosphate from ATP to the 5′ end of either DNA or RNA.
   - It is a product of the T4 bacteriophage, and commercial preparations are usually products of the cloned phage gene expressed in E. coli.

The enzymatic activity of PNK is utilized in two types of reactions:

- In the “forward reaction”, PNK transfers the gamma phosphate from ATP to the 5′ end of a polynucleotide (DNA or RNA).
- The target nucleotide is lacking a 5′ phosphate either because it has been dephosphorylated or has been synthesized chemically.
- In the “exchange reaction”, target DNA or RNA that has a 5′ phosphate is incubated with an excess of ADP, PNK will first transfer the phosphate from the nucleic acid onto an ADP, forming ATP and leaving a dephosphorylated target.
- PNK will then perform a forward reaction and transfer a phosphate from ATP onto the target nucleic acid.

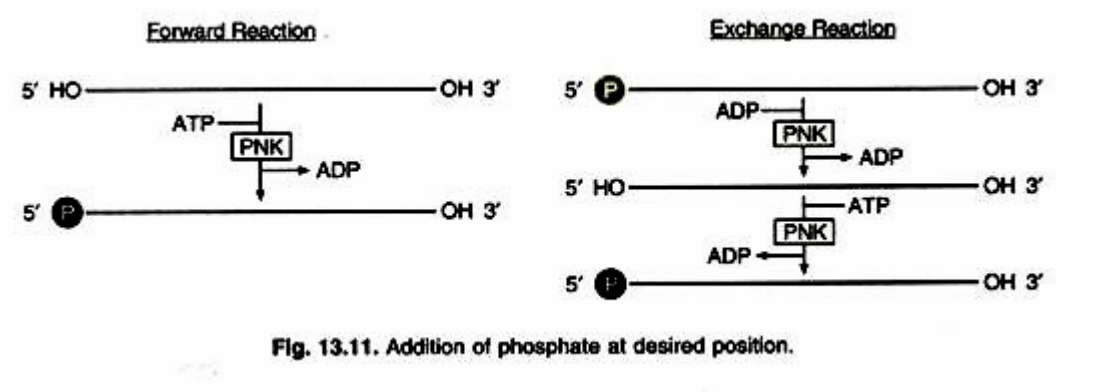


Fig. 13.11. Addition of phosphate at desired position.

- As you might expect, the efficiency of phosphorylating via the exchange reaction is considerably less than for the forward reaction.
- In addition to its phosphorylating activity, PNK also has 3′ phosphatase activity, although this has little significance to molecular technologists.

There are two major indications for phosphorylating nucleic acids and hence uses of PNK are:

- Phosphorylating linkers and adaptors (fragments of DNA ready for ligation) which require a 5′ phosphate. This includes products of polymerase chain reaction, which are typically generated using non-phosphorylated primers.
- Radiolabelling oligonucleotides, usually with 32P, for use as hybridization probes. PNK is inhibited by small amounts of ammonium ions, so ammonium acetate should not be used to precipitate nucleic acids prior to phosphorylation.
- Low concentrations of phosphate ions, or NaCl concentrations greater than about 50 mM, also inhibit this enzyme.

7. **DNA Ligase:**

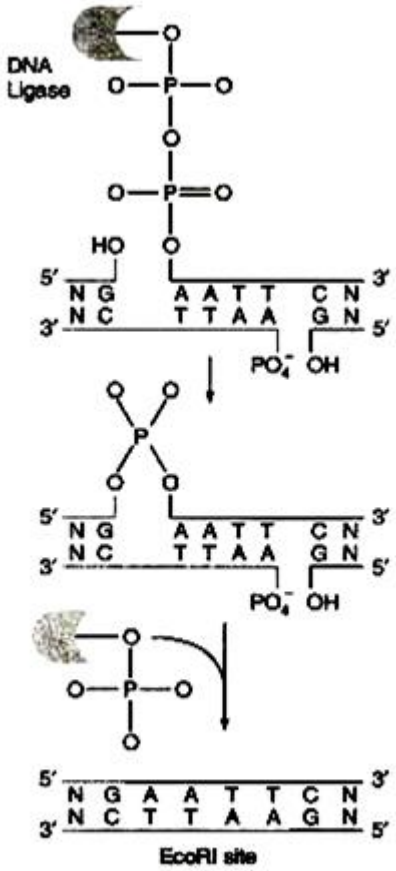


Fig. 13.12. Sealing of nick by ligase

- The term recombinant DNA includes the concept of recombining fragments of DNA from different sources into a new and useful DNA molecule.
- Joining linear DNA fragments together with covalent bonds is called ligation.
- More specifically, DNA ligation involves creating a phosphodiester bond between the 3′ hydroxyl of one nucleotide and the 5′ phosphate of another.
- The enzyme used to ligate DNA fragments is T4 DNA ligase, which originates from the T4 bacteriophage.
- This enzyme will ligate DNA fragments having overhanging, cohesive ends that are annealed together, as in the EcoRI.
- This is equivalent to repairing “nicks” in duplex DNA.

- T4 DNA ligase will also ligate fragments with blunt ends, although higher concentrations of the enzyme are usually recommended for this purpose.

Probes

The following points highlight the three types of nucleic acid probes. The probes are:

1. **Oligonucleotide Probes**
2. **DNA Probes**
3. **RNA Probes**

1. **Oligonucleotide Probes:**
   - These are synthesized chemically as oligonucleotides based on the information available on the amino acid sequence of the protein of interest.
   - These oligo nucleotides can be used as a probe the identification of gene which encode for that particular protein.
   - However, due to degeneracy of the genetic code, construction of oligonucleotide is car-ried out with those that are rich in methionine or tryptophan residue or with only two codons.
   - Generally oligonucleotide probe is used to screen cDNA libraries.

2. **DNA Probes:**
   - These are longer than the oligonucleotides.
   - Thus, clones of longer DNA sequence is used as a DNA probes.
   - The sequence obtained from cDNA library (cDNA clones) can be used to probe genomic library to identify genomic clones.
   - The same probe can also be used to reprobe the same cDNA library to identify more cDNA clones.
   - The genomic DNA clones are used to screen cDNA library or a genomic library.

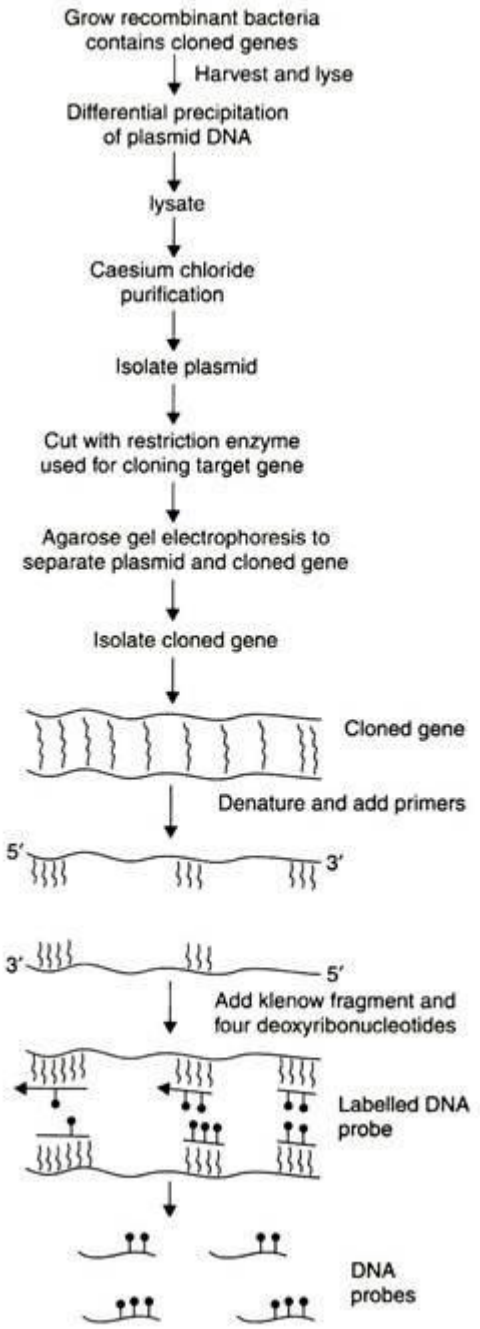


Fig. 13.8 Preparation of labelled probes

Preparation of DNA Probe:

The DNA probe is prepared by random primer method as follows:

- In double stranded DNA containing the sequence that is to act as the probe is dena-tured and an oligonucleotide sample containing all possible sequences of six nucleotides

is added (it is statistical certainty that some of the molecules of the oligonucleotide mixture will hybridize to the unlabelled, denatured probe DNA).

- In the presence of klenow fragment and four deoxyribunucleotides, one of the four deoxyribonucleotides is labelled.
- The bound oligonucleotides act as primers for DNA synthesis.
- The synthesised DNA is labelled and can be used as a probe to detect the presence of a complementary DNA sequence in a source DNA sample.

- There are at least **two possible sources of probes**: one is from cloned DNA, second is the nucleotide sequence of a synthetic probe, based on the probable nucleotide sequence that is deduced from the known aminoacid sequence of the protein encoded by the target gene.

3. **RNA Probes:**
   - RNA probes used only under certain circumstances.
   - Purification of particular RNA is generally used as a specific probe for the corresponding DNA specifies.
   - By employing positive- negative screening or differential hybridized approach, often possible to identify clones for RNA in one population but not so in other population.

   - The colony is first probed with labelled RNA from one population.
   - After the location of hybridized colonies, labelled RNA is then washed off from the membrane and the membrane is then probed with labelled RNA from the second population and hybridized colonies are identified.

- The first RNA probes were mRNAs of a gene that are abundantly expressed in a cell, labelled with 32po4.
- One of the efficient ways of preparing RNA probes involves transcription from a target gene cloned in a plasmid.
- Transcription of this gene is initiated from a promoter that is specifically recognised by an RNA polymerase.
- Due to unspecificity of RNA pol recogni-tion, specific promoter can be selected.
- The promoter of a bacteriophage is very specific for the phage polymerase.
- Some researchers have utilized phage (SP6) T7 promoter and the corre-sponding enzyme to transcribe a DNA (target gene) cloned in a plasmid vector.

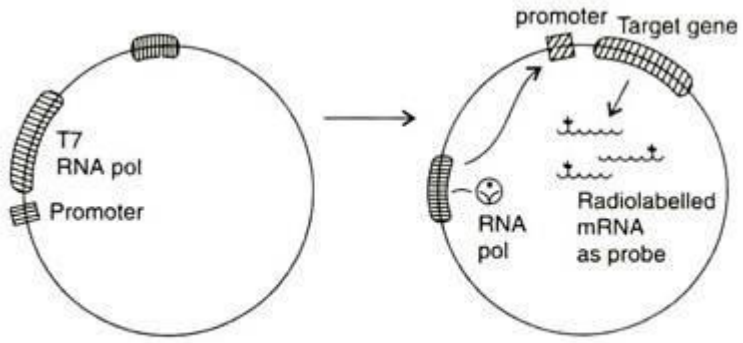


Fig 13.9 Preparation of RNA probe

- The gene for the phage polymerase and the phage promoter were both cloned in a vector, and a target gene sequence for transcription inserted downstream of the phage promoter.
- Cells transformed with recombinant DNA are supplied with RNA precursor nucleotides, of which one nucleotide is radio labelled.
- Transcription of the inserted gene is therefore labelled and may be used as RNA probes.

ENTRI

Molecular marker

- Molecular marker is identified as genetic marker.
- Molecular marker is a DNA or gene sequence within a recognized location on a chromosome which is used as identification tool.
- In the pool of unknown DNA or in a whole chromosome, these molecular markers helps in identification of particular sequence of DNA at particular location.

- **Applications:**
    - It plays a crucial role in gene mapping by identifying the position of linked genes in the chromosome which inherited together.
    - It also detect any alteration in a sequence of DNA or any genetic oddity. It ascertains genes involved in genetic disorders.
    - It is used to determine different characters in a DNA sequence which is used to distinguish between individuals, populations or species.
- Different types of genetic polymorphism can be used as Genetic markers. On the basis of polymorphisms detected in the genetic makeup of individuals that may vary in the length of a DNA sequence or in the identity of nucleotides located at specific position in chromosome, some of the common genetic markers are- RFLP, SSLP, etc
- Genetic markers can be classified as PCR based and hybridization based.
    - PCR based genetic markers: RAPD, ISSR, EST-SSR, microsatellite, CAPS etc.
    - Hybridization based genetic markers: RFLP, VNTRs, in which targeted gene is digested with restriction enzymes and then hybridized with RFLP probe.

Quality for a good genetic marker:

- Genetic markers should be largely polymorphic in nature

- They should be selectively neutral
- Assay for detecting markers should be simple and rapid
- Genetic markers should occur frequently within genome
- The genetic marker (gene) should show codominant inheritance pattern.
- They should be highly reproducible
- They should not interact with other markers while using multiple markers at a same time

Types of genetic markers:

1. **Random Amplified Polymorphic DNA (RAPD):**

- RAPD was developed by Welsh and McClelland along with Williams in 1990.
- It is pronounced as 'rapid'.
- It is based on PCR assay and it doesn't need require any prior sequencing of DNA.
- This procedure uses short arbitrary primer of 8-12 bp that randomly amplifies the region of DNA.
- This primer serves as both forward and reverse primer.
- This reaction proceeds when a single primer anneals to the genomic DNA at two distinct sites on the complementary strand of DNA template.
- The amplification of segment of DNA depends on the positions complimentary to the primers' sequence.
- The fragments obtained from RAPD are between 0.2 to 5.0kb and can be viewed by using agarose gel electrophoresis stained by ethidium bromide or with the help of polyacrylamide gel electrophoresis.
- If any mutation occurs in the primer binding region then no any PCR product will be produced, yielding a distinct pattern of amplified DNA segments on the gel.
- **Application:**

- Distinct pattern of amplification is seen in different samples. This is why RAPD can be used for studying polymorphism.
- RAPD is applicable for the mapping of genome, analyzing linkage, and individual specific genotyping.
- RAPD markers are dominant in nature so it has restrictions for mapping purpose.
- RAPD is strictly laboratory dependent so it requires sensitivity.

- **Demerits:**
  - It has demerits as poor reproducibility, yields faint products, problems occur in band scoring.

2. Restriction Fragment Length Polymorphism (RFLP):

- It was one of the first methods used for the analysis of DNA in various fields such as forensic science.
- It is a hybridization based technique.
- It was invented by Alec Jeffreys, an English scientist in 1984 during his research in genetic diseases.
- RFLP uses particular restriction endonuclease enzymes that cuts at its specific site yielding fragments of various lengths along with the fragment of interest.
- The length of the distinct fragments is determined by using blotting, now replaced with sequencing.
- RFLP markers are largely locus-specific and are co-dominant in nature due to the nature of restriction endonuclease used.
- **Steps for RFLP are as follows:**
  - DNA extraction is done from saliva, blood or other samples and is purified.
  - Restriction endonucleases digests the purified DNA resulting restriction fragments.
  - Now the restriction fragments are examined using gel electrophoresis.

- The gel is now treated with luminescent dyes for the visibility of DNA bands.

- **Applications:**
  - RFLP was one of the first techniques applied for genetic fingerprinting/profiling.
  - It is used for identification of inherited diseases, carrier of that diseases, genetic mapping, and heterozygous detection.
  - The molecular basis of the RFLP is that any point mutations as such deletions, substitutions and insertions or alterations like duplications, inversions within the genome can eliminate or form new restriction sites. These alterations in genome can be detected by analyzing fragments of variable length, digested with restriction endonuclease enzyme

- **Demerits:**
  - requires relatively large DNA sample
  - laborious and tedious process
  - sensitivity and more precautions for contamination required

3. Amplified Fragment Length Polymorphism (AFLP):

- Zabeau and Vos invented the AFLP technique in 1993.
- AFLP was originally developed by the KeyGene in 1990.
- It is a PCR based technique for fingerprinting. It includes both PCR and RFLP.
- The basis of AFLP is the amplification of selected fragments followed by restriction digestion of whole genomic DNA of specific organism.
- **The steps for the AFLP are as follows:**
  - DNA extraction and its restriction digestion followed by ligation with the short adaptor sequences.
  - Amplification of restricted fragments by PCR

    - Analysis of results in gel electrophoresis or PAGE followed by autoradiography.
- **Applications:**
    - AFLP has its ability for rapid generation of marker fragments for any organism without prior sequencing of DNA is required.
    - Also, it needs only small fragments of starting template DNA relative to RAPD and ISSR (inter-simple sequence repeats) and has much higher reproducibility.
    - AFLP is largely used for crop improvement programs, parentage and genomic interpretation of various crop species.
- **Demerits:**
    - AFLP require large DNA samples and require purification

Homoploymer tailing

- A homopolymer is a polymer where all the subunits are same.
- Tailing is the adding a series of nucleotides on to the 3′-OH termini of a double stranded DNA molecule using enzyme terminal deoxynucleotidyl transferase of calf thymus.
- In homopolymer tailing, gene of interest is tailed with one nucleotide and vector is tailed with an complementary base.
- It is the method for joining DNA molecules by annealing of complementary homopolymer sequence.
- Oligo(dA) sequences is added to the 3′ ends of one molecule and oligo(dT) blocks to the 3′ end of another molecule.
- This method is catalyzed by enzyme terminal deoxynucleotidyl transferase by addition of nucleotides to the 3′ end of a DNA fragment

Process

- They are then allowed to incubate together toallow hydrogen bonding to bring them together.

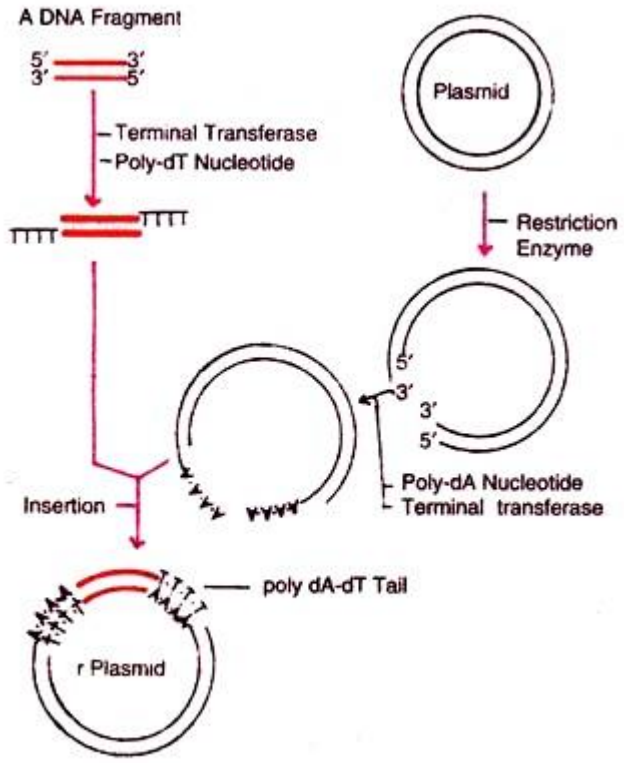
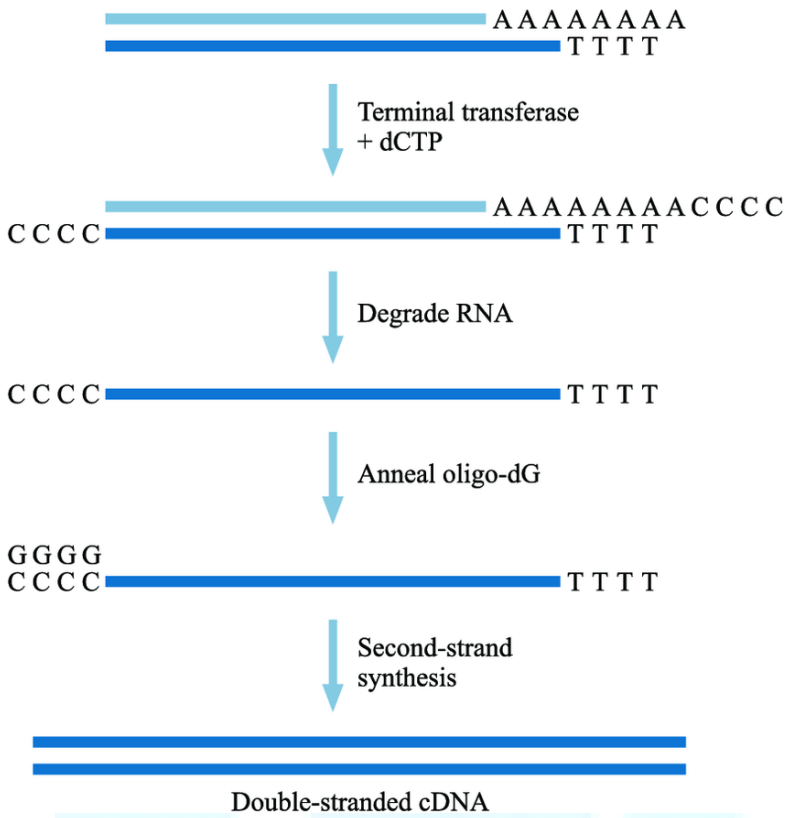


Fig. 11.3 : Homopolymer tails.

- Homopolymer tailing is achieved by annealing.
- It is a technique by which sticky ends can be produced on an blunt-ended DNA molecule.
- The reaction when carried out in the presence of one deoxynucleotide a homopolymer tail is produced.
- For ligation of two tailed molecules the hompolymers must be complementary.
- Poly dC tails are attached to the vector and poly dG to the DNA to be cloned.

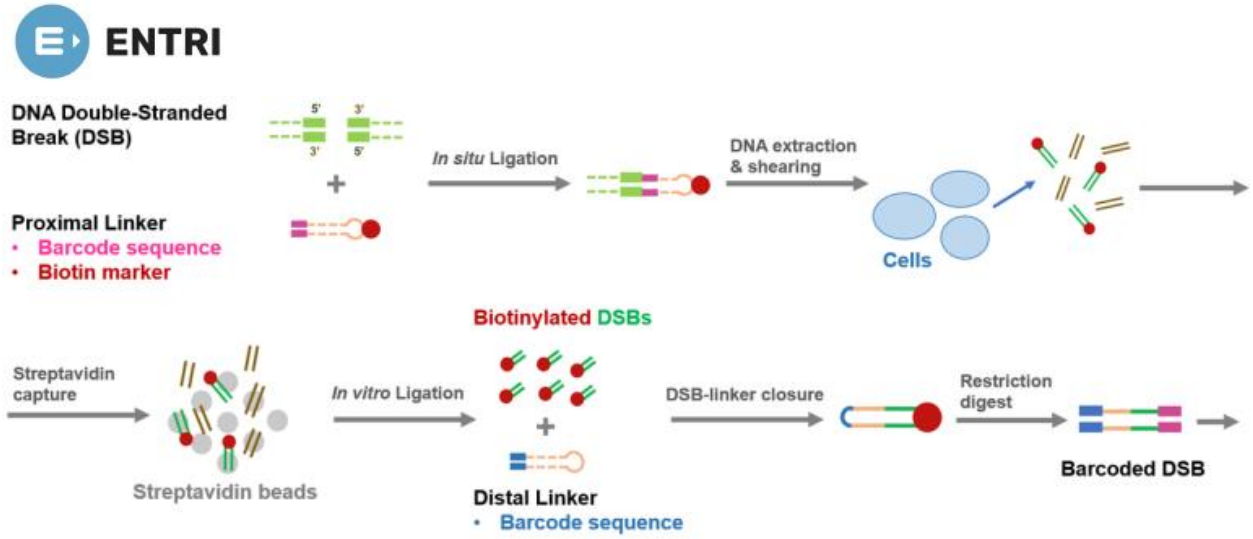
Uses of Homoploymer tailing



- This is more widely used approach for making recombinant DNA.
- Segments of 50-100 poly A and poly G are added to one fragment.
- Equal length segments of poly-T and poly-C are added to the other fragment.
- Homopolymer tailing to provide a priming site for second-strand synthesis.

Linkers

- The linkers are short double-stranded sequences of DNA. Both the ends of a linker molecule are blunt in nature. They are chemically synthesized oligonucleotides.
- It contains restriction sites for the identification of restriction enzymes. The restriction enzymes cleave the ligated linker and the DNA fragment to produce cohesive ends.
-
-

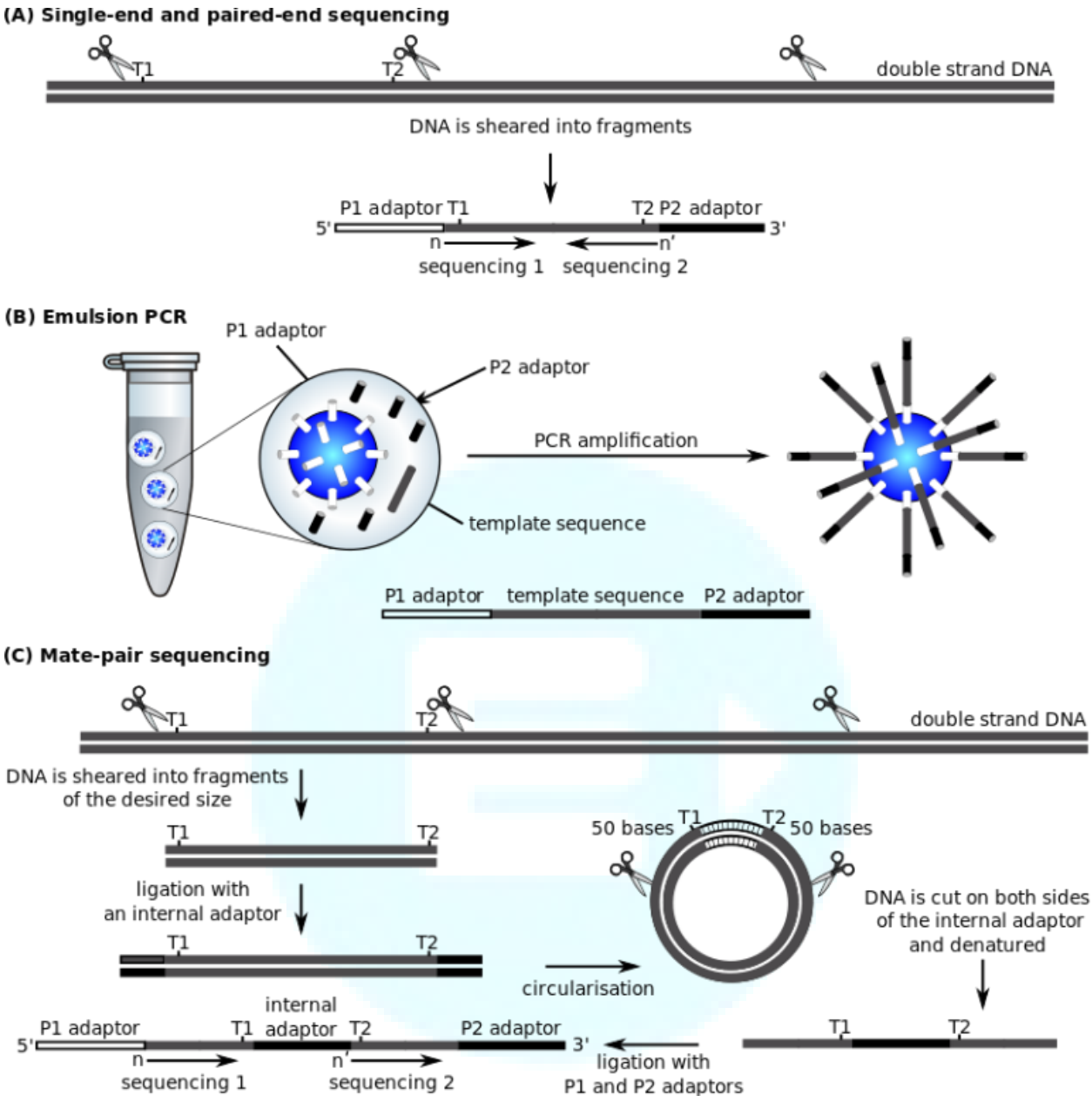


→ PCR → Next-Generation Sequencing

- One drawback of linkers is that the DNA fragment sometimes already possesses the restriction sites for producing cohesive ends. This limits the use of linker molecules.
- **Example:** Eco-RI linkers and Sal-I linkers

Adaptors

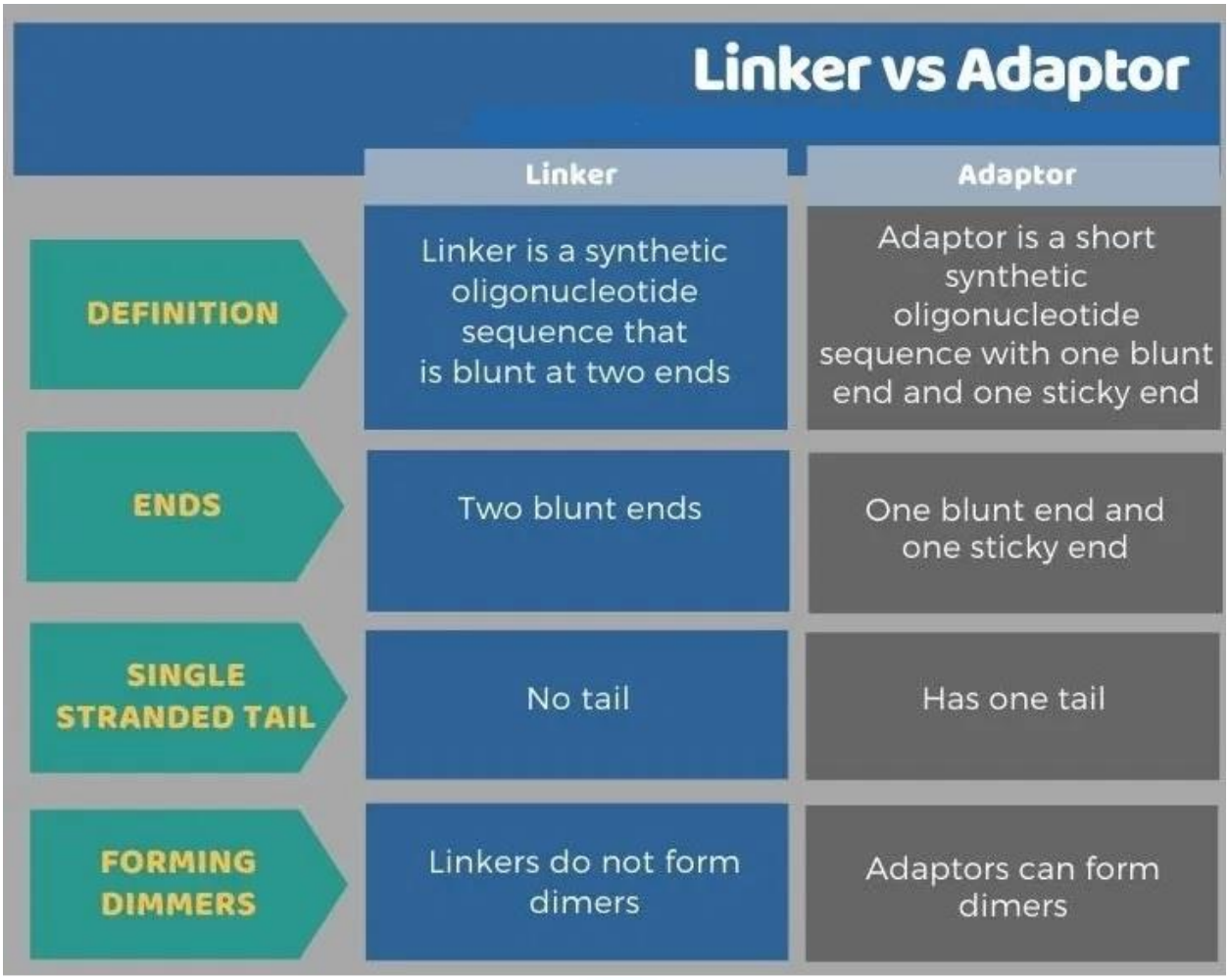
- An adaptor molecule is a single or double-stranded, chemically synthesized oligonucleotide that is used for the ligation of DNA and RNA strands.
- The adaptors can have both blunt ends or one sticky and one blunt end. The sticky end helps in easy ligation of the DNA fragments. The adaptors also have restriction sites that can be used to create new protruding terminuses by the action of restriction enzymes.
- One disadvantage of adaptor molecules is that two sticky ends can join to form dimers. This can however be prevented by treating the molecules with alkaline phosphatases.
- Uses of adaptor molecules include the addition of a sticky end to cDNA for easy ligation into the plasmids and similar vector ligations.



What are the Similarities Between Linker and Adaptor?

- Both linker and adaptor are double-stranded short oligonucleotide sequences.
- They carry internal restriction sites.
- Moreover, they are chemically synthesized DNA molecules and are synthetic molecules.

- They can link two DNA molecules together.
- After ligation of linkers and adaptors, the DNA is again restricted with restriction enzymes in order to produce sticky ends.



Genetic engineering techniques

Polymerase chain reaction

- **Polymerase chain reaction (PCR)** is a common laboratory technique used to make many copies (millions or billions!) of a particular region of DNA.
- This DNA region can be anything the experimenter is interested in.
- For example, it might be a gene whose function a researcher wants to understand, or a genetic marker used by forensic scientists to match crime scene DNA with suspects.
- Typically, the goal of PCR is to make enough of the target DNA region that it can be analyzed or used in some other way. For instance, DNA amplified by PCR may be sent for sequencing, visualized by gel electrophoresis, or cloned into a plasmid for further experiments.
- PCR is used in many areas of biology and medicine, including molecular biology research, medical diagnostics, and even some branches of ecology.

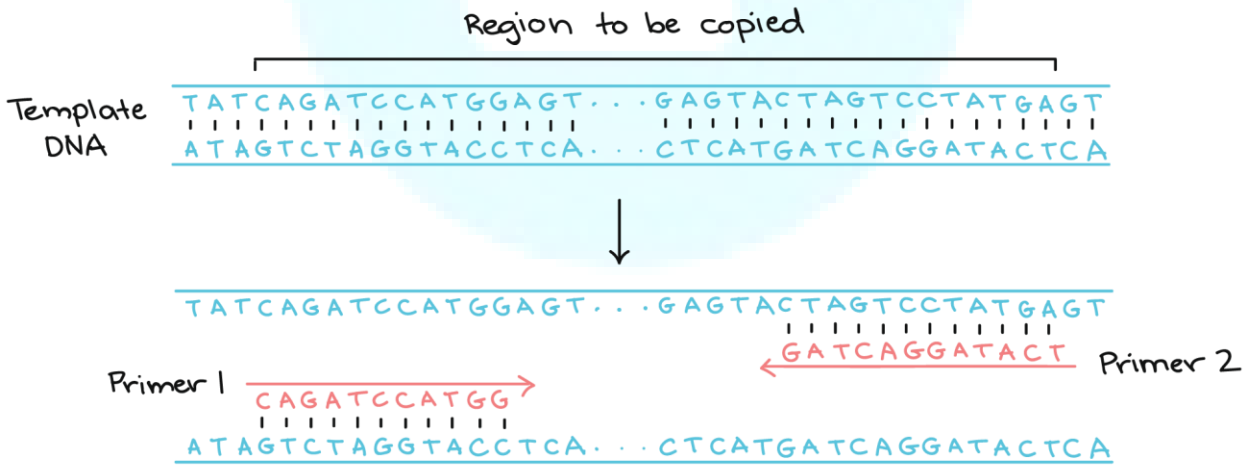
Taq polymerase

- Like DNA replication in an organism, PCR requires a DNA polymerase enzyme that makes new strands of DNA, using existing strands as templates.
- The DNA polymerase typically used in PCR is called Taq polymerase, after the heat-tolerant bacterium from which it was isolated (Thermus aquaticus).
- T. aquaticus lives in hot springs and hydrothermal vents.
- Its DNA polymerase is very heat-stable and is most active around 70 °\text C70°C70, °, start text, C, end text (a temperature at which a human or E. coli DNA polymerase would be nonfunctional).

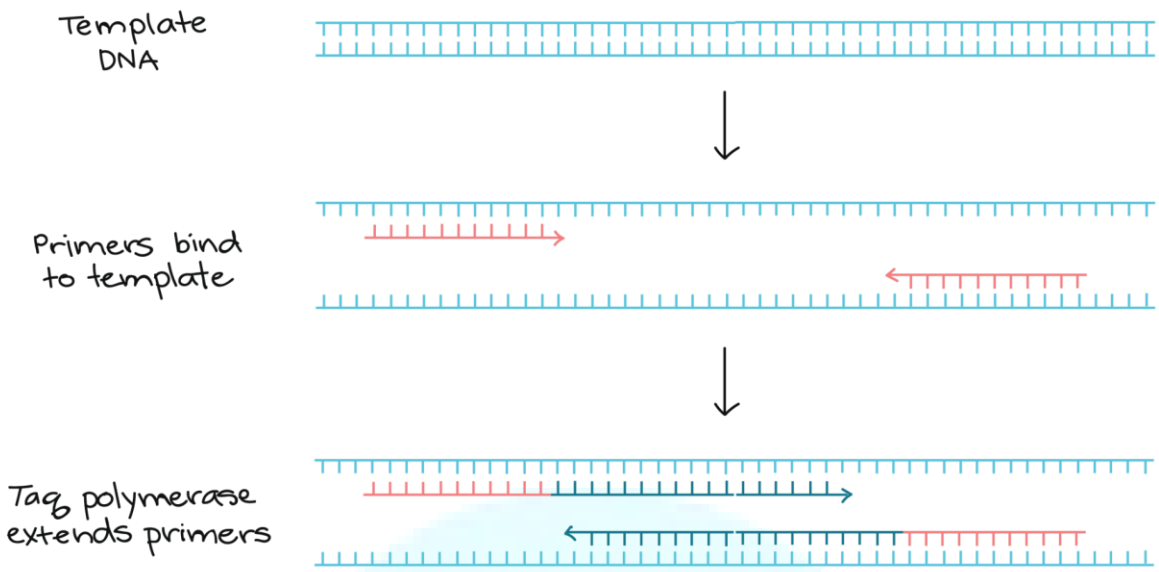
- This heat-stability makes Taq polymerase ideal for PCR. As we'll see, high temperature is used repeatedly in PCR to denature the template DNA, or separate its strands.

PCR primers

- Like other DNA polymerases, Taq polymerase can only make DNA if it's given a primer, a short sequence of nucleotides that provides a starting point for DNA synthesis.
- In a PCR reaction, the experimenter determines the region of DNA that will be copied, or amplified, by the primers she or he chooses.
- PCR primers are short pieces of single-stranded DNA, usually around 202020 nucleotides in length.
- Two primers are used in each PCR reaction, and they are designed so that they flank the target region (region that should be copied).
- That is, they are given sequences that will make them bind to opposite strands of the template DNA, just at the edges of the region to be copied.
- The primers bind to the template by complementary base pairing.



When the primers are bound to the template, they can be extended by the polymerase, and the region that lies between them will get copied

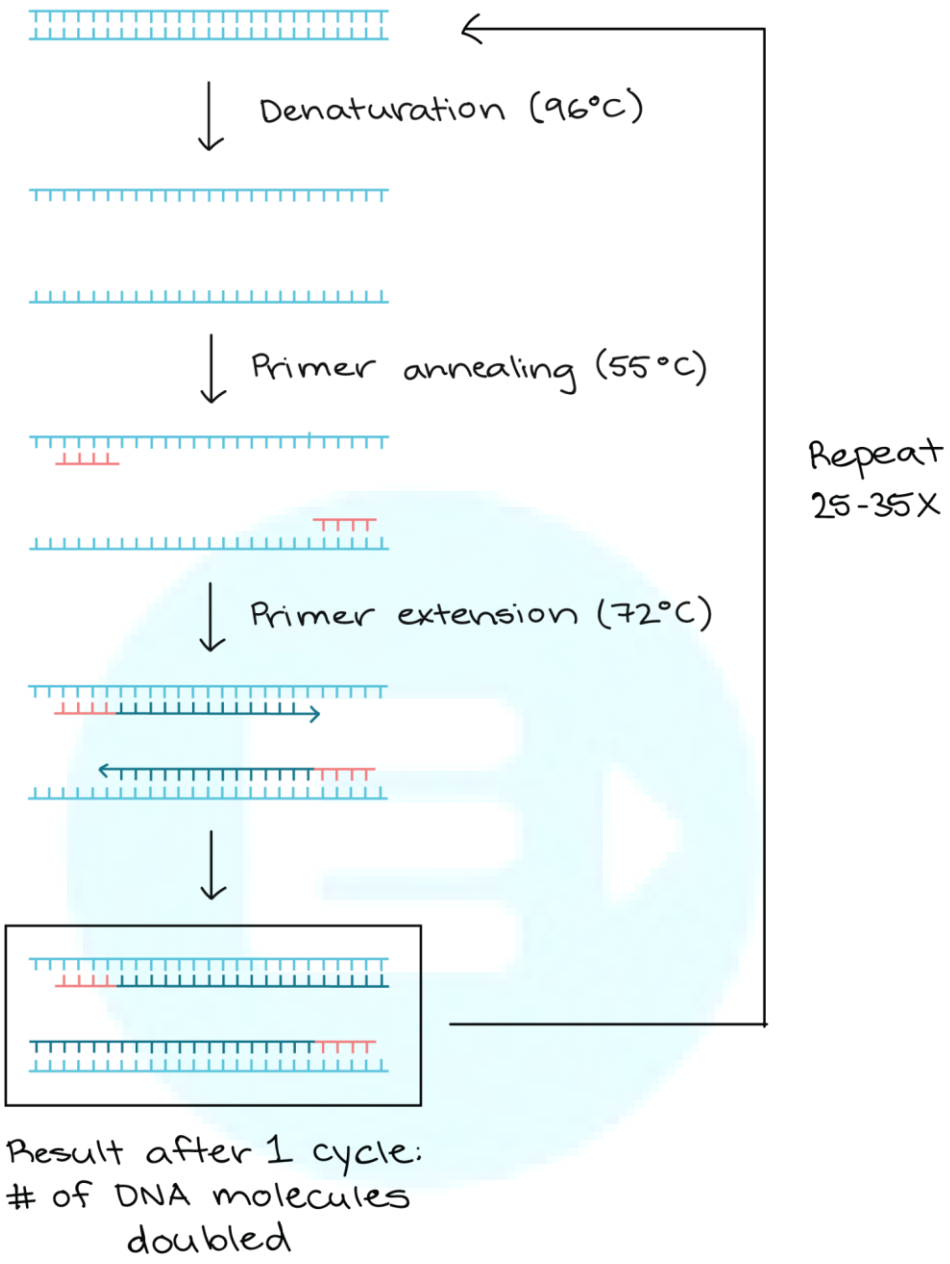


The steps of PCR

The key ingredients of a PCR reaction are Taq polymerase, primers, template DNA, and nucleotides (DNA building blocks). The ingredients are assembled in a tube, along with cofactors needed by the enzyme, and are put through repeated cycles of heating and cooling that allow DNA to be synthesized.

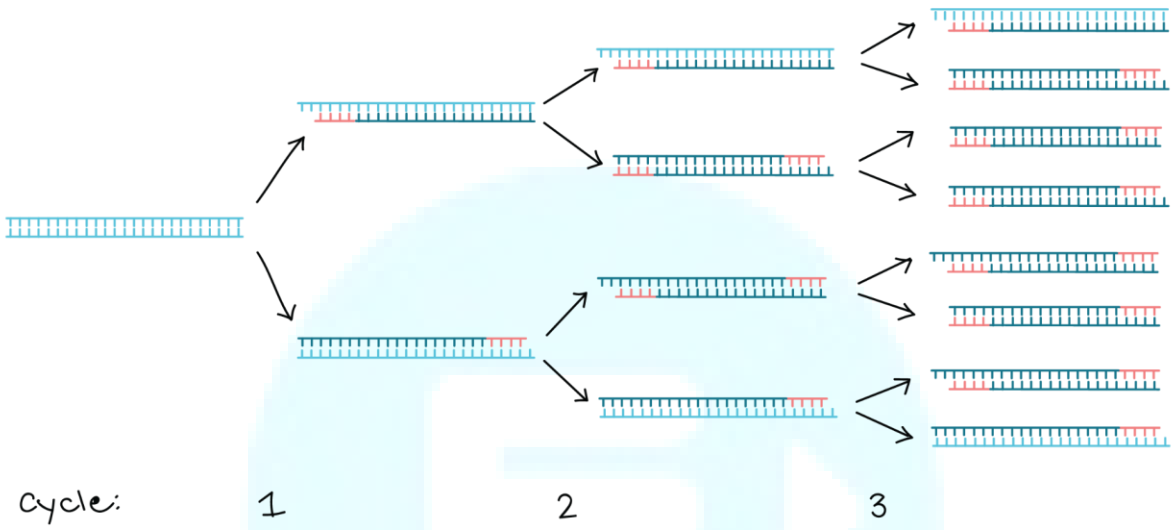
The basic steps are:

- **Denaturation (96°C):** Heat the reaction strongly to separate, or denature, the DNA strands. This provides single-stranded template for the next step.
- **Annealing (55-65°C):** Cool the reaction so the primers can bind to their complementary sequences on the single-stranded template DNA.
- **Extension (72°C):** Raise the reaction temperatures so Taq polymerase extends the primers, synthesizing new strands of DNA.



- This cycle repeats 25-35 times in a typical PCR reaction, which generally takes 2-4 hours, depending on the length of the DNA region being copied.
- If the reaction is efficient (works well), the target region can go from just one or a few copies to billions.
- That's because it's not just the original DNA that's used as a template each time.

- Instead, the new DNA that's made in one round can serve as a template in the next round of DNA synthesis.
- There are many copies of the primers and many molecules of Taq polymerase floating around in the reaction, so the number of DNA molecules can roughly double in each round of cycling.



Applications of PCR

- Using PCR, a DNA sequence can be amplified millions or billions of times, producing enough DNA copies to be analyzed using other techniques.
- For instance, the DNA may be visualized by gel electrophoresis, sent for sequencing, or digested with restriction enzymes and cloned into a plasmid.
- PCR is used in many research labs, and it also has practical applications in forensics, genetic testing, and diagnostics.
- For instance, PCR is used to amplify genes associated with genetic disorders from the DNA of patients (or from fetal DNA, in the case of prenatal testing).
- PCR can also be used to test for a bacterium or DNA virus in a patient's body: if the pathogen is present, it may be possible to amplify regions of its DNA from a blood or tissue sample.

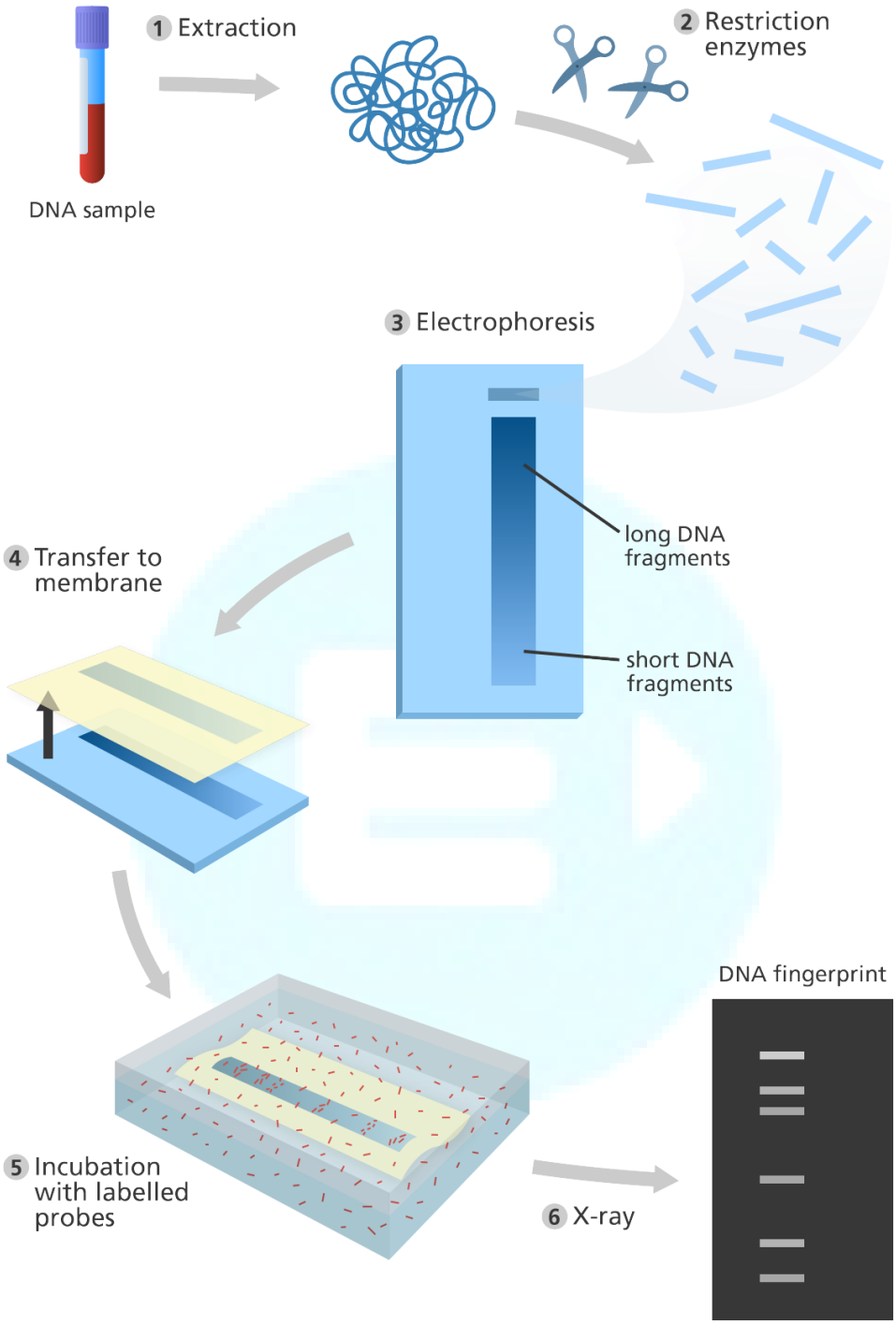
DNA fingerprinting

- "DNA fingerprinting is a procedure that shows the hereditary cosmetics of living things. It is a strategy for finding the distinction between the satellite DNA areas in the genome."
- DNA profiling, DNA testing, DNA examination, Genetic profile, DNA distinguishing proof, genetic fingerprinting, and genetic investigation are a portion of the mainstream names utilized for DNA fingerprinting. This technique was invented by Alec Jeffreys in 1984.

Principle of DNA fingerprinting

- The human genome consists of innumerable small noncoding sequences which are inheritable and repeatedly present.
- They can be separated from the bulk DNA as satellite upon performing density gradient centrifugation and thus known as satellite DNA.
- They can be categorized into either microsatellites or microsatellites depending on the length, base composition and tandemly repetitive units.
- These satellite DNAs show polymorphism and this polymorphism is the basis of DNA fingerprinting.
- The repeat regions can be divided into two groups based on the size of the repeat - variable number tandem repeats (VNTRs) and short tandem repeats.
- These repeats act as genetic markers and every individual inherits these repeats from their parents.
- Thus, every individual has a particular composition of VNTRs and this is the main principle of the DNA fingerprinting technique.

DNA Fingerprinting Steps



1. Collection of organic example blood, spit, buccal swab, semen, or solid tissue.

2. DNA extraction.

3. Restriction absorption or PCR intensification.
4. Agarose gel electrophoresis, slim electrophoresis or DNA sequencing.
5. Interpreting outcomes.

The Process of DNA Fingerprinting

Sample collection, DNA extraction, absorption or intensification and investigation results are significant advances.

- **Stage 1: Sample Collection**
  - DNA can be acquired from any bodily sample or liquid. Buccal smear, salivation, blood, amniotic liquid, chorionic villi, skin, hair, body liquid, and different tissues are significant kinds of samples utilized.

- **Stage 2: DNA Extraction**
  - We need to initially get DNA. To play out any genetic applications, DNA extraction is one of the most significant advances. Great quality and amount of DNA expands the conceivable outcomes of getting better outcomes.
  - Following strategies can utilize:
    - Phenol-chloroform DNA extraction strategy
    - CTAB DNA extraction strategy
    - Proteinase K DNA extraction strategy

- **Stage 3: Restriction Absorption, Enhancement or DNA Sequencing**
  - Three regular strategies are utilized:
    - RFLP based STR investigation
    - PCR based investigation
    - Real-time PCR investigation
- **Stage 4: Analysis of Results**

- Utilizing the southern blotting, agarose gel electrophoresis, narrow electrophoresis, ongoing intensification, and DNA sequencing, the outcomes for different DNA profiling can be gotten in which rt-PCR and sequencing are much of the use in forensic science.

- **Stage 5: Interpreting Results**
  - By looking at DNA profiles of different examples, varieties and likenesses between people can be distinguished.
  - Outstandingly, the whole procedure is presently nearly automatic. We don't need to do anything, the computer gives us conclusive outcomes.

Applications of DNA Fingerprinting

- Utilizing the DNA fingerprinting strategy, the natural personality of an individual can be uncovered. For approving one's character, there is no other preferable alternative over DNA fingerprinting.
- Gravely harmed dead bodies can be distinguished.
- It is utilized to detect maternal cell contamination.
- One of the significant downsides of pre-birth determination is maternal cell tainting. The amniotic liquid or CVS test contains the maternal DNA or maternal tissue, once in a while. Contamination expands the opportunity of false-positive outcomes, particularly on account of carrier recognition. Utilizing VNTRs and STRs markers with PCR-gel electrophoresis, maternal cell tainting can be recognized during pregnancy hereditary testing.
- One of the most significant uses of the current strategy is in the crime scene examination and criminal check. The example is gathered from the crime site which could be salivation, blood, hair follicle, or semen. DNA is removed and investigated against the suspect, utilizing the two

markers we clarified previously. By coordinating DNA band designs criminal's connected to wrongdoing can be built up.

Utilizing Blood-Typing in Paternity Tests

- The procedure of DNA fingerprinting was discovered by Alec Jeffreys in 1984, and it originally opened up for paternity testing in 1988.
- Before this kind of DNA investigation was accessible, blood classifications were the most widely recognized calculation considered human paternity testing. Blood bunches are a mainstream case of Mendelian hereditary qualities at work.
- All things considered, there are various human blood bunches with numerous alleles, and these alleles display a scope of predominance designs.

DNA Fingerprinting and Farming

- A few DNA minisatellite tests have yielded piece profiles that show up valuable for plant reproducing work.
- These part profiles show no variety when vegetative spread material is broken down.
- So also, examples obtained through self-inbreeding species show indistinguishable profiles.
- Interestingly, hereditary recombination in cross-pollinating species brings about exceptionally factor, normally singular, explicit piece profiles.
- Along these lines various cultivars can be recognized, as additionally can genotypes of wild species in characteristic populaces.
- These piece profiles can likewise be used in parentage examination, as has just been led in rice and apples, in this way empowering us to explain the source of deficiently recorded cultivars.

- Also, evaluations of hereditary variety dependent on similitude lists determined from section profiles show a nearby relationship with known degrees of hereditary relatedness.

Blotting techniques

- Blotting is the technique in which nucleic acids or proteins are immobilized onto a solid support generally nylon or nitrocellulose membranes.
- Blotting of nucleic acid is the central technique for hybridization studies.
- Nucleic acid labeling and hybridization on membranes have formed the basis for a range of experimental techniques involving understanding of gene expression, organization, etc.
- Identifying and measuring specific proteins in complex biological mixtures, such as blood, have long been important goals in scientific and diagnostic practice.
- More recently the identification of abnormal genes in genomic DNA has become increasingly important in clinical research and genetic counseling.
- Blotting techniques are used to identify unique proteins and nucleic acid sequences.
- They have been developed to be highly specific and sensitive and have become important tools in both molecular biology and clinical research.

General principle

- The blotting methods are fairly simple and usually consist of four separate steps: electrophoretic separation of protein or of nucleic acid fragments in the sample; transfer to and immobilization on paper support; binding of analytical probe to target molecule on paper; and visualization of bound probe. Molecules in a sample are first separated by electrophoresis and then transferred on to an easily handled support medium or membrane.

- This immobilizes the protein or DNA fragments, provides a faithful replica of the original separation, and facilitates subsequent biochemical analysis.
- After being transferred to the support medium the immobilized protein or nucleic acid fragment is localized by the use of probes, such as antibodies or DNA, that specifically bind to the molecule of interest.
- Finally, the position of the probe that is bound to the immobilized target molecule is visualized usually by autoradiography.
- Three main blotting techniques have been developed and are commonly called
    - **Southern blotting**
    - **Northern blotting**
    - **Western blotting**

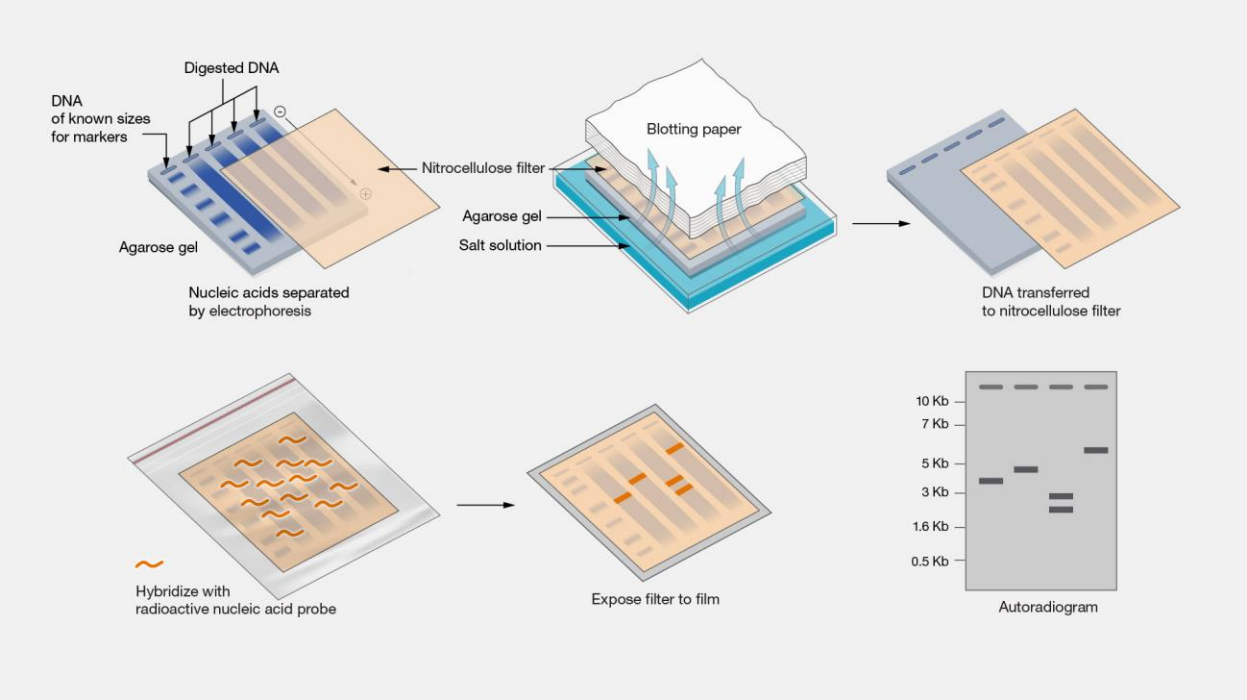
Southern blot

Southern blot is a method used to check for the presence of a DNA sequence in a DNA sample.

The method is named after its inventor, the British biologist Edwin Southern.

The procedure for Southern blot technique is as detailed below:

- Restriction endonucleases are used to cut high-molecular-weight DNA strands into smaller fragments, which are then electrophoresed on an agarose gel to separate them by size.
- If the DNA fragments are larger than 15 kb, then prior to blotting, the gel may be treated with an acid, such as dilute HCl, which depurinates the DNA fragments, breaking the DNA into smaller pieces, thus allowing more efficient transfer from the gel to membrane.
- If alkaline transfer methods are used, the DNA gel is placed into an alkaline solution (containing NaOH) to denature the double-stranded DNA.



- The denaturation in an alkaline environment may improve binding of the negatively charged DNA to a positively charged membrane, separating it into single DNA strands for later hybridization to the probe and destroys any residual RNA that may still be present in the DNA.
- A sheet of nitrocellulose (or 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.
- 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 used to move the DNA from the gel on to 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.
- The membrane is then baked in a vacuum or regular oven at 80 °C for 2 hours or exposed to ultraviolet radiation (nylon membrane) to permanently attach the transferred DNA to the membrane.

- 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.
- After hybridization, excess probe is washed from the membrane 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. 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 (ie, 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.

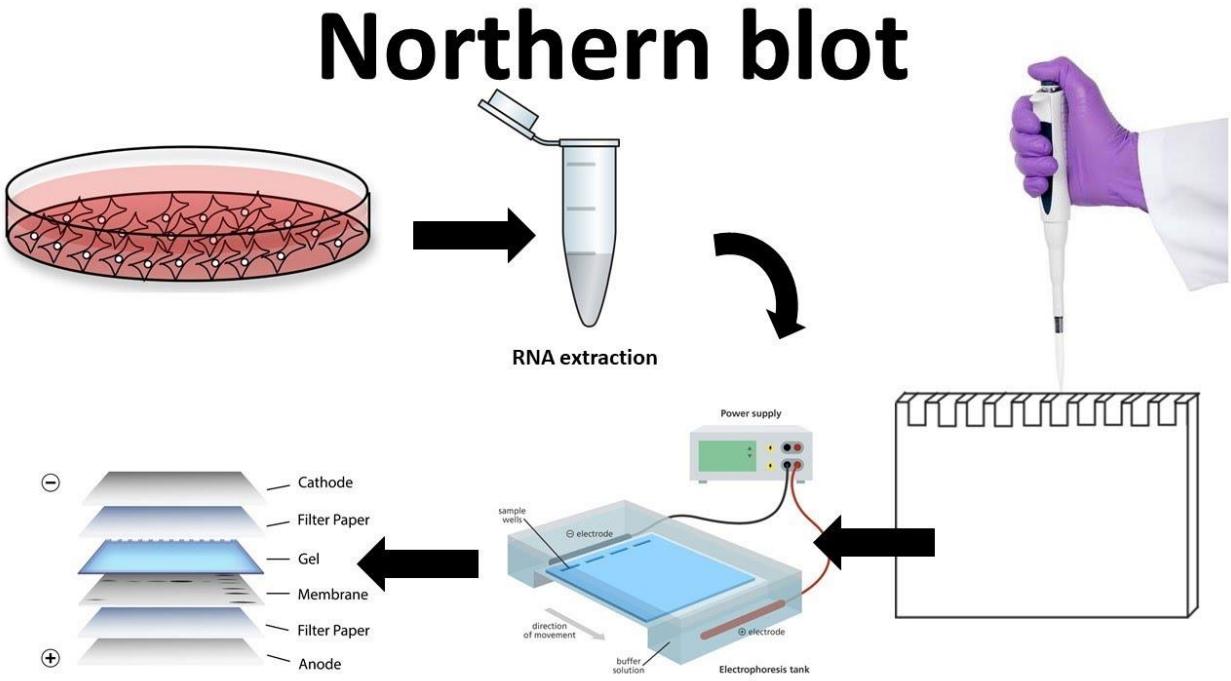
Northern blot

- The northern blot technique is used 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 levels during differentiation, morphogenesis, as well as abnormal or diseased conditions.
- This technique was developed in 1977 by James Alwine, David Kemp and George Stark at Stanford University.
- Northern blotting takes its name from its similarity to the first blotting technique, the Southern blot.
- The major difference is that RNA, rather than DNA, is analyzed in the northern blot.

Procedure

- The blotting procedure starts with extraction of total RNA from a homogenized tissue sample.
- The mRNA can then be isolated through the use of oligo (dT) cellulose chromatography to maintain only those RNAs with a poly(A) tail.
- RNA samples are then separated by gel electrophoresis.
- 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 preventing RNA degradation by high temperatures.
- 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.
- The membrane is washed to ensure that the probe has bound specifically.

- The hybrid signals are then detected by X-ray film and can be quantified by densitometry.



Applications

- Northern blotting allows in observing a particular gene's expression pattern between tissues, organs, developmental stages, environmental stress levels, pathogen infection.
- The technique has been used to show over expression of oncogenes and down regulation 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 up regulated 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 & disadvantages

- Analysis of gene expression can be done by several different methods including RT-PCR, RNase protection assays, microarrays, 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.

Western blot

- A western blot, sometimes called a protein immunoblot, is an antibody-based technique used to detect the presence, size and abundance of specific proteins within a sample.
- The technique was developed in 19791 by Harry Towbin and colleagues and later named the "western blot" due to the technique's similarity to Southern blotting.
- Briefly, proteins in an aqueous sample are separated by electrophoresis.
- Following transfer to an appropriate membrane, the samples are probed using target-specific antibodies.
- These antibodies can be detected, and the size and abundance of the bound proteins evaluated in comparison to known standards or controls.

Western blot protocol

- **Western blot gel**
    - Before a western blot can be performed, the proteins in the sample must be separated.
    - This is typically achieved by protein electrophoresis, such as sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE) or native PAGE, which separates proteins based on their molecular weight or charge.
    - The specific separation method chosen will depend on the aim of the analysis.
    - For a clean image, samples are centrifuged to remove any solids, in order to load only the soluble fraction.
    - If your protein of interest is in the insoluble fraction (e.g., cell membrane-bound proteins) investigate pretreatment methods to liberate and solubilize it first.
    - Solids will impair the running of the gel and it is likely your protein of interest will remain in the stacking gel.

- It is also important to load appropriate control samples and size marker ladders to enable interpretation of the final blot.

- **Loading control for western blot**
  - It is essential, especially when trying to compare protein expression between different samples, to know how much sample has been loaded as this may not be apparent from the blot alone.
  - For example, when assessing a blot, the band from one sample may appear twice as bright as another sample.
  - This could mean that there is twice as much of the target protein in that sample, or it could mean that more sample or a more concentrated sample has been loaded in one lane than the other.
  - Running a duplicate protein gel and developing with Coomassie stain5 can help to remove this uncertainty as it will show the amount of total protein6 in each sample lane and can reveal any loading inconsistencies.
  - Detecting expression of a ubiquitous protein that should be even between all of your samples, such as actin in whole cell and cytoplasmic samples, can also be used as a loading control and helps to ensure consistent transfer of protein samples to the membrane.
  - However, this type of control can be problematic when comparing models in which “control” proteins are differently expressed, such as degeneration models.

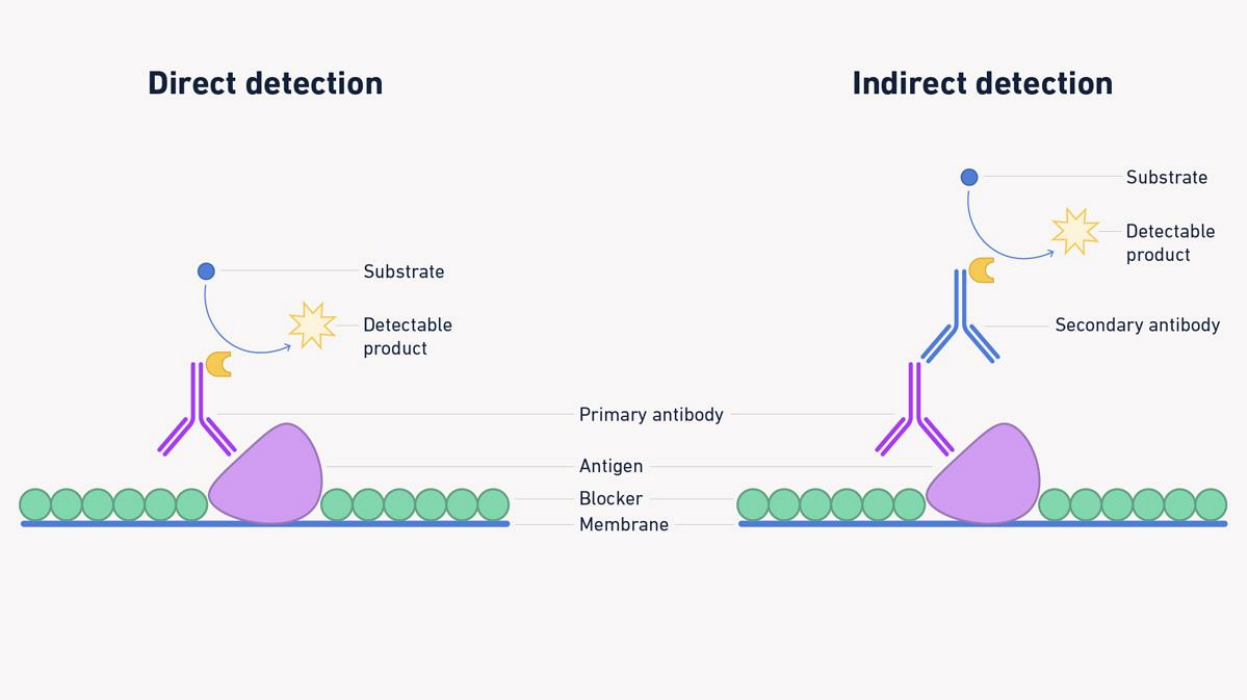
- **Protein transfer (blotting)**
  - Proteins must be transferred from the protein gel to an appropriate membrane (typically nitrocellulose or polyvinylidene difluoride (PVDF)) to facilitate antibody probing.

- A number of techniques can be used for transfer, including capillary transfer, diffusion transfer and **vacuum blotting**, but by far the most common due to its speed and efficiency is **electroblotting** (also called electroelution or electrophoretic transfer).
- Here, the protein gel is sandwiched against the transfer membrane and an electrical current is applied.
- Proteins from the gel are carried across and attach to the membrane tightly.
- Within electroblotting, there are also multiple strategies for transfer, known as wet, semi-dry and dry transfer.
- Wet transfer is efficient and offers flexible buffering but is time-consuming.
- Semi-dry transfer is quicker and still offers flexibility but is less efficient than wet transfer for large proteins.
- Dry transfer is efficient and quick but offers less flexibility than the other methods.
- Transfer efficiency can be assessed prior to probing using a removable stain such as Ponceau S.

- **Western blot blocking**
  - Due to the high affinity of blotting membranes for proteins, after transfer it is important to block any remaining binding sites to prevent subsequent non-specific binding of the assay detection antibodies.
  - This is achieved by incubating the membrane with a proteinaceous liquid such as milk or serum.
- **Western blot washing**
  - Following blocking, it is important to wash the membrane between each step to remove excess or unbound reagents.

- Insufficient or uneven washing can lead to poor quality/patchy blots and high background.
- However, over washing can diminish the target signal so it is important to optimize the number and duration of wash steps.
- Ensure the membrane is well covered with an appropriate buffer and apply gentle agitation to wash the membrane evenly without damaging it.
- Commonly used buffers include tris-buffered saline (TBS) and phosphate-buffered saline (PBS), often with the inclusion of Tween 20 (TBST and PBST).

- **Western blot antibody**

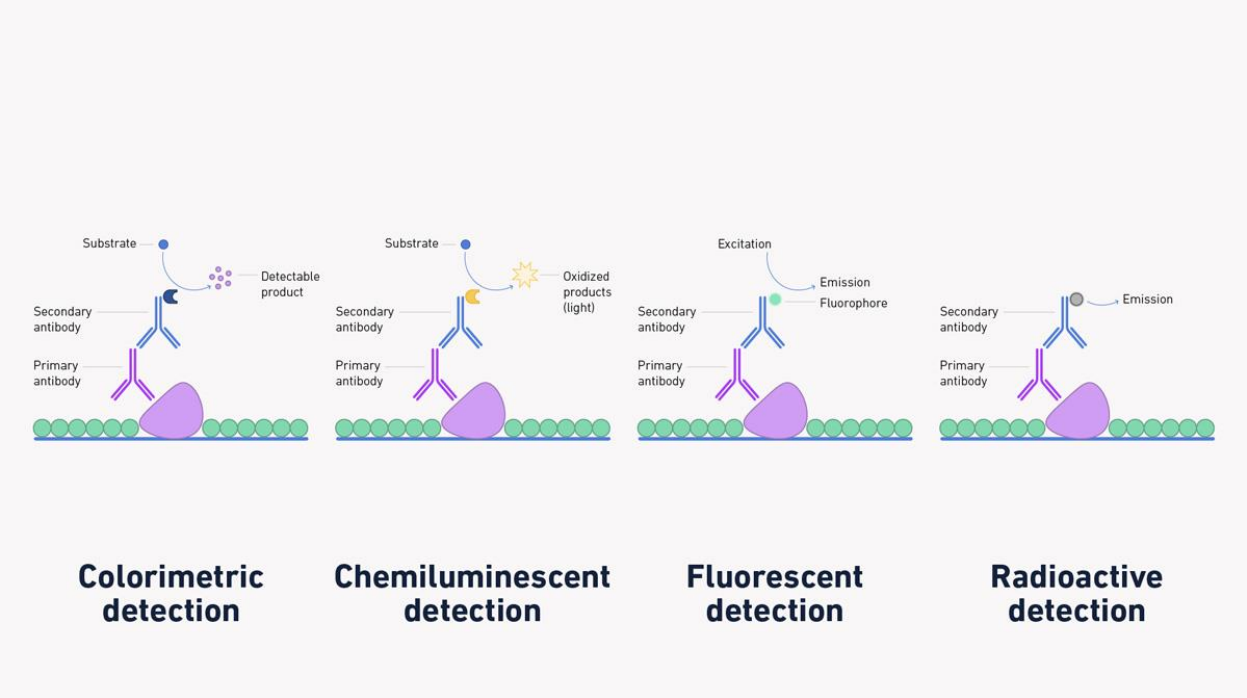


- While it is possible to use direct detection (a single antibody that recognizes the target and is detectable) for a western blot, more often an indirect method is applied.
- Here, a primary antibody is used to probe the membrane and bind any target protein present.

- Then, a secondary antibody is used that binds the primary antibody and is detectable.
- As with all steps, optimization, in this case choosing the "right" antibody and determining the optimal concentration, is key to a good blot.

- **Western blot secondary antibody**
  - When an indirect detection assay is used, a secondary antibody will need to be applied after washing excess unbound primary antibody off the membrane.
  - The secondary antibody should be specific to the species of the primary antibody (e.g., mouse anti-rabbit if the primary antibody was derived from a rabbit) and possess the necessary conjugate for the chosen detection method.

- **Western blot analysis**



- There are multiple methods for detection and subsequent visualization of western blots including **colorimetric, chemiluminescent, fluorescent and radioactive detection**.
- Both **colorimetric and chemiluminescent detection** require conjugation of an enzyme to the detection antibody and are considered very sensitive techniques.
- Horseradish peroxidase (HRP) and alkaline phosphatase (AP) are the most commonly used enzymes, with HRP generally favored due to its stability, amenability to most conjugations and low cost.
- During detection, a substrate is added to the membrane, which is acted on by the conjugated enzyme, bringing about a chemical change.
- If performing colorimetric detection, a chromogenic substrate is chosen that will produce a change that can be visualized and imaged directly.
- However, prompt imaging of the blot is important as colors will fade as the blot dries.
- In chemiluminescent detection, the signal produced only lasts as long as the reaction between the enzyme and substrate is occurring (typically 1–24 hours).
- During this time, the signal can be recorded by exposing X-ray film or using digital imaging to make a permanent record.
- In fluorescent detection, the detection antibody is conjugated with a fluorophore rather than an enzyme.
- When light of a specific wavelength is shone on them, they become excited and emit light of a different specific wavelength.
- This can then be captured visually using digital imagers, such as an avalanche photodiode (APD), photomultiplier tube (PMT) or charge-coupled device (CCD) camera.

- While specialist equipment is required to undertake the excitation and detection steps, there is no substrate step in fluorescent detection, shortening the protocol.
- It is also possible to multiplex fluorophores11 within a western blot assay.
- Radioactive detection, where a radioisotope is conjugated to the detection antibody and the emitted radiation is detected on X-ray film, was used extensively in the past.
- However, the technique requires special handling to protect personnel from the radiation, is expensive and has a limited shelf life due to radioactive decay.
- Therefore, the technique has mostly been replaced in favor of other available detection methods.

Dot Blotting Techniques

- The drawbacks of blotting techniques have lead to the development of dot blotting technique which is more advanced, less time consuming, accurate and applicable to a wide variety of gene/source simultaneously.
- The dot or slot blotting technique is the most widely used of all techniques for analysing.
- None of the blot methods require electrophoresis prior to blotting and hybridization.
- Hybridization of cloned DNA without electrophoretic separation is called as **dot blotting**.

DNA sequencing

- DNA sequencing refers to methods for determining the order of the nucleotides bases adenine,guanine,cytosine and thymine in a molecule of DNA.

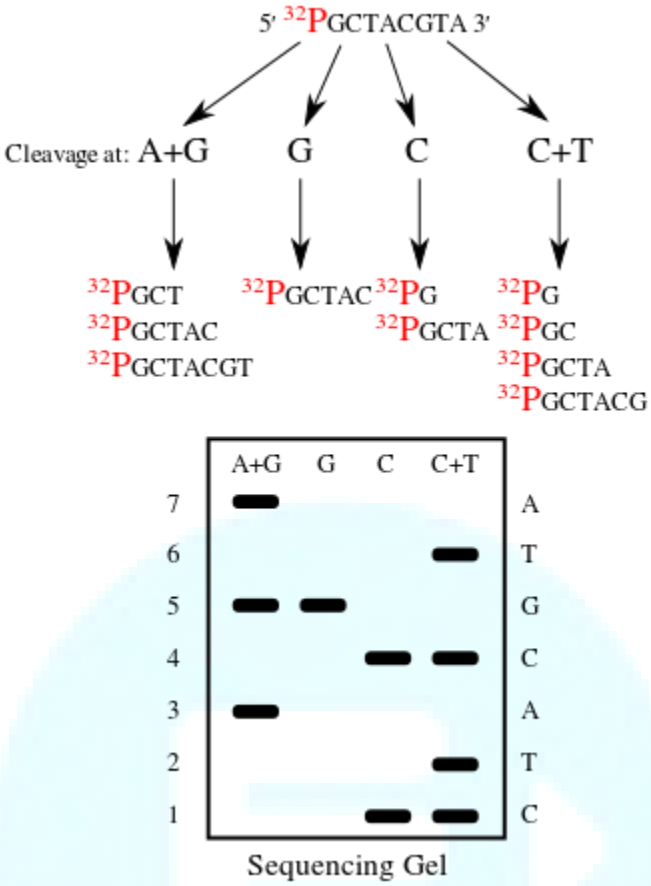
- The first DNA sequence was obtained by academic researchers, using laboratories methods based on 2- dimensional chromatography in the early 1970s.
- By the development of dye based sequencing method with automated analysis, DNA sequencing has become easier and faster.

Two main methods are widely known to be used to sequence DNA:

  - **The Chemical Method** (also called the Maxam–Gilbert method after its inventors).
  - **The Chain Termination** Method (also known as the Sanger dideoxy method after its inventor).
- Maxam–Gilbert technique depends on the relative chemical liability of different nucleotide bonds, whereas the Sanger method interrupts elongation of DNA sequences by incorporating dideoxynucleotides into the sequences.
- The chain termination method is the method more usually used because of its speed and simplicity.

Maxam-Gilbert method

- In 1976-1977, Allan Maxam and Walter Gilbert developed a DNA sequencing method based on chemical modification of DNA and subsequent cleavage at specific bases.
- The method requires radioactive labelling at one end and purification of the DNA fragment to be sequenced.
- Chemical treatment generates breaks at a small proportions of one or two of the four nucleotide based in each of four reactions (G,A+G, C, C+T).
- Thus a series of labelled fragments is generated, from the radiolabelled end to the first ‘cut’ site in each molecule.
- The fragments in the four reactions are arranged side by side in gel electrophoresis for size separation.



- To visualize the fragments, the gel is exposed to X-ray film for autoradiography, yielding a series of dark bands each corresponding to a radiolabelled DNA fragment, from which the sequence may be inferred.

Key Features

- Base-specific cleavage of DNA by certain chemicals
- Four different chemicals, one for each base
- A set of DNA fragments of different sizes
- DNA fragments contain up to 500 nucleotides

Advantages

- Purified DNA can be read directly
- Homopolymeric DNA runs are sequenced as efficiently as heterogeneous DNA sequences
- Can be used to analyze DNA protein interactions (i.e. footprinting)

- Can be used to analyze nucleic acid structure and epigenetic modifications to DNA

Disadvantages

- It requires extensive use of hazardous chemicals.
- It has a relatively complex set up / technical complexity.
- It is difficult to "scale up" and cannot be used to analyze more than 500 base pairs.
- The read length decreases from incomplete cleavage reactions.
- It is difficult to make Maxam-Gilbert sequencing based DNA kits.

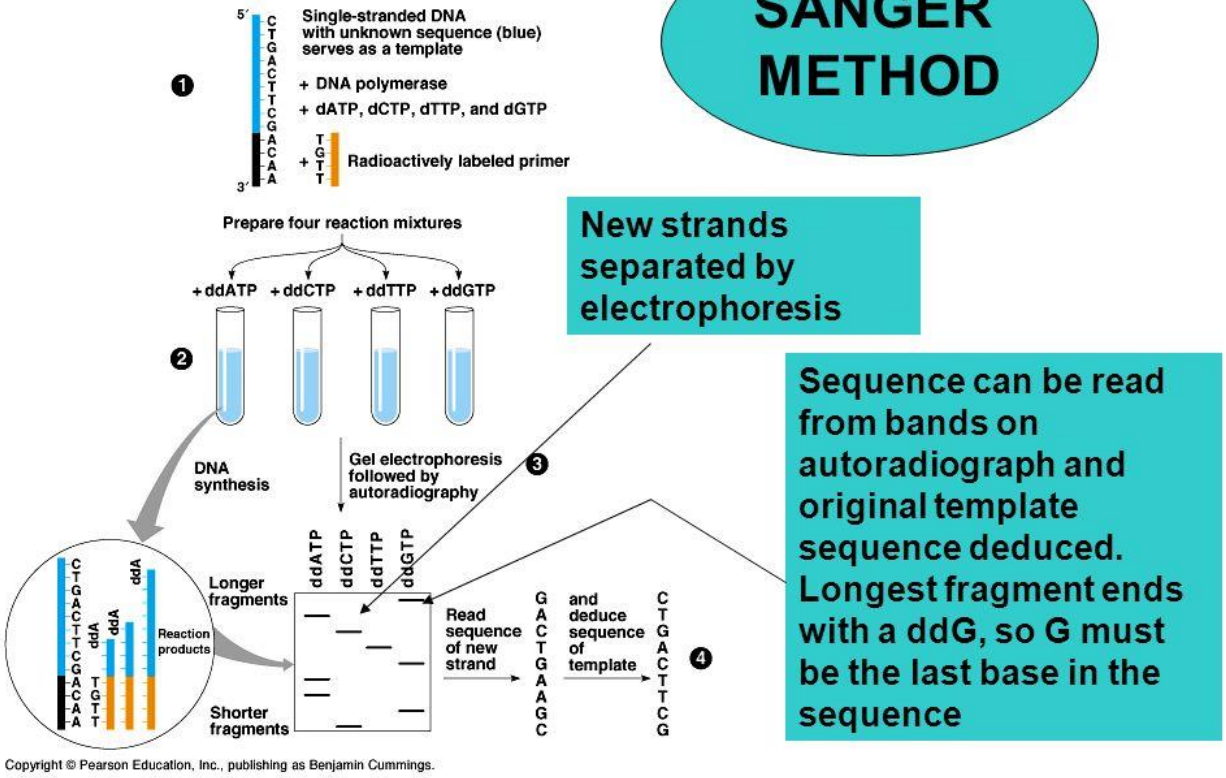
Sanger-Coulson method

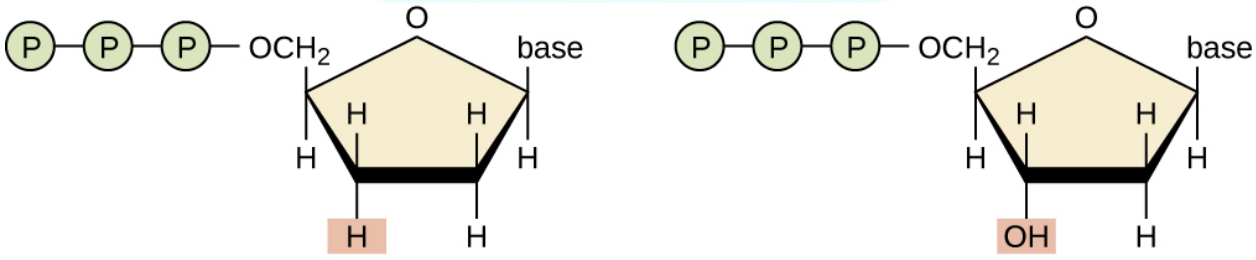
- Sanger's method of gene sequencing is also known as dideoxy chain termination method. It generates nested set of labelled fragments from a template strand of DNA to be sequenced by replicating that template strand and interrupting the replication process at one of the four bases.
- Four different reaction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

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 the growing DNA strand then the replication stops at that nucleotide.

SANGER METHOD



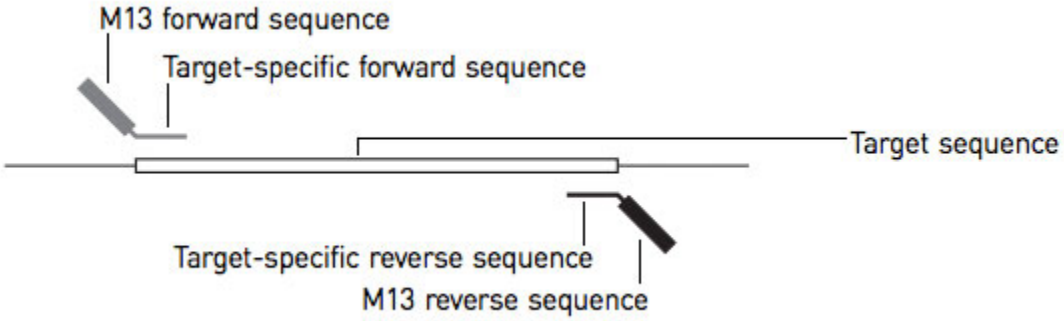


- The ddNTPs are analogue of dNTPs.
- ddNTPs lacks hydroxyl group (-OH) at c3 of ribose sugar, so it cannot make phosphodiester bond with nest nucleotide, thus terminates the nucleotide chain.

- 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:**

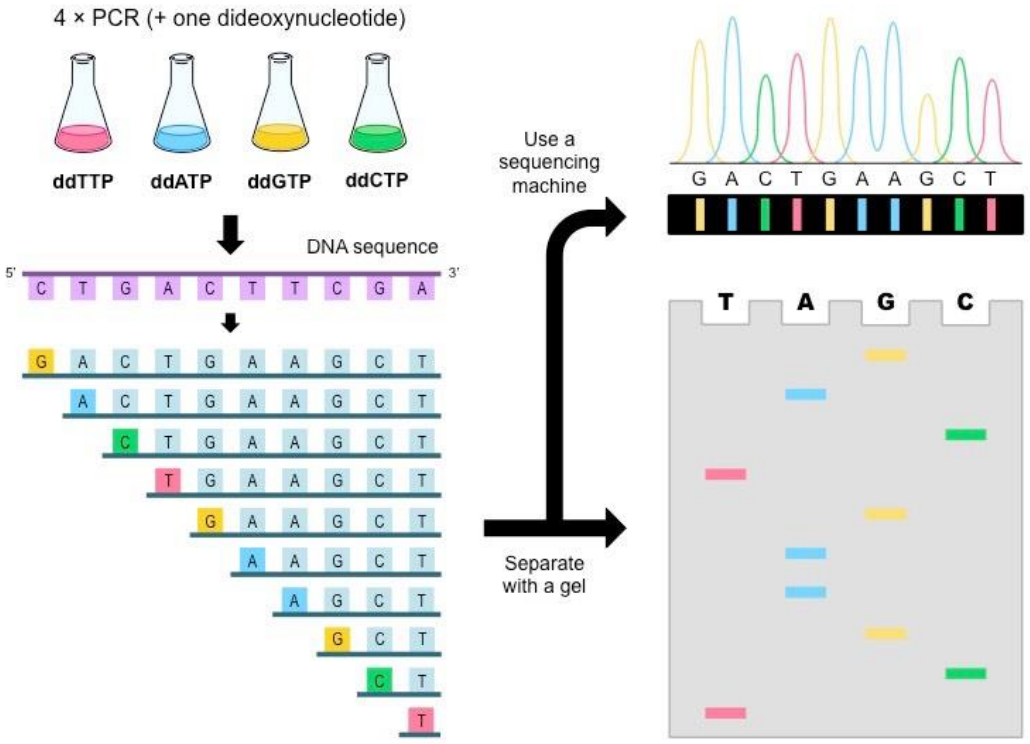


- Copies of template strand to be sequenced must be prepared with short known sequences at 3' end of the template strand.
- A DNA primere is essential to initiate replication of template , so primer preparation of known sequences at 3'end is always required.
- For this purpose a single stranded cloning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



3. Electrophoresis and gel reading:

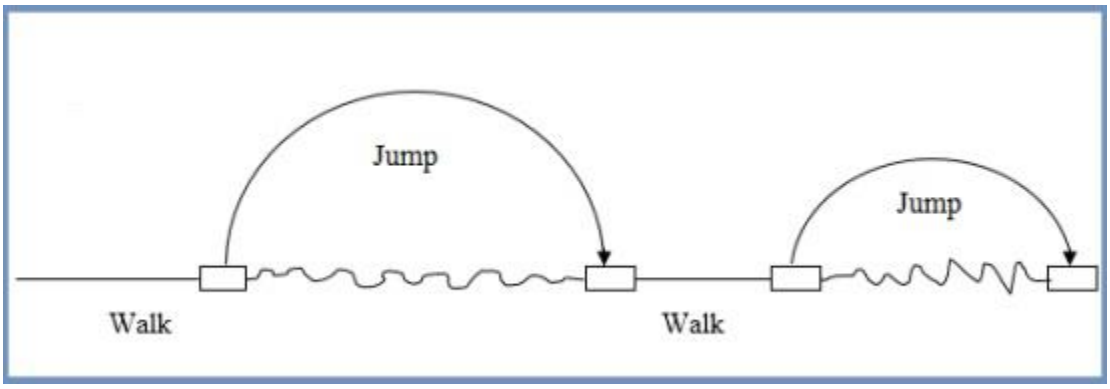
- The reaction mixture from four batches are loaded into four different well on polyacrylamide gel and electrophoresed.
- The autoradiogramof the gel is read to determine the order of bases of complementary strand to that of template strand.
- The band of shortest fragments are at the bottom of autoradiogram so that the sequences of complementary strand is read from bottom to top.

Significance of DNA Sequencing

- Information obtained by DNA sequencing makes it possible to understand or alter the function of genes.
- DNA sequence analysis demonstrates regulatory regions that control gene expression and genetic “hot spots” particularly susceptible to mutation.
- Comparison of DNA sequences shows evolutionary relationships that provide a framework for definite classification of microorganisms including viruses.
- Comparison of DNA sequences facilitates identification of conserved regions, which are useful for development of specific hybridization probes to detect microorganisms including viruses in clinical samples.
- DNA sequencing has become sufficiently fast and inexpensive to allow laboratory determination of microbial sequences for identification of microbes. Sequencing of the 16S ribosomal subunit can be used to identify specific bacteria. Sequencing of viruses can be used to identify the virus and distinguish different strains.
- DNA sequencing shows gene structure that helps research workers to find out the structure of gene products.

Chromosome jumping

- Chromosomal jumping is a technique used in molecular biology for physical mapping of genomes of the organisms.
- This technique was introduced to overcome a barrier of the chromosomal walking which arose upon finding the repetitive DNA regions during the cloning process. Therefore, chromosome jumping technique can be considered as a special version of chromosomal walking. It is a rapid method compared to chromosomal walking and enables bypassing of the repetitive DNA sequences which are not prone to be cloned during chromosomal walking.



- Chromosomal jumping narrows the gap between the target gene and the available known markers for genome mapping.
- Chromosome jumping tool starts with the cutting of a specific DNA with special restriction endonucleases and ligation of the fragments into circularized loops.
- Then a primer designed from a known sequence is used to sequence the circularized loops.
- This primer enables jumping and sequencing in an alternative manner.
- Hence, it can bypass the repetitive DNA sequences and rapidly walk through the chromosome for the search of the target gene.
- The discovery of the gene encodes for cystic fibrosis disease was done using the chromosomal jumping tool.
- Combined together, chromosomal jumping and walking can enhance the genome mapping process.

Genomic library

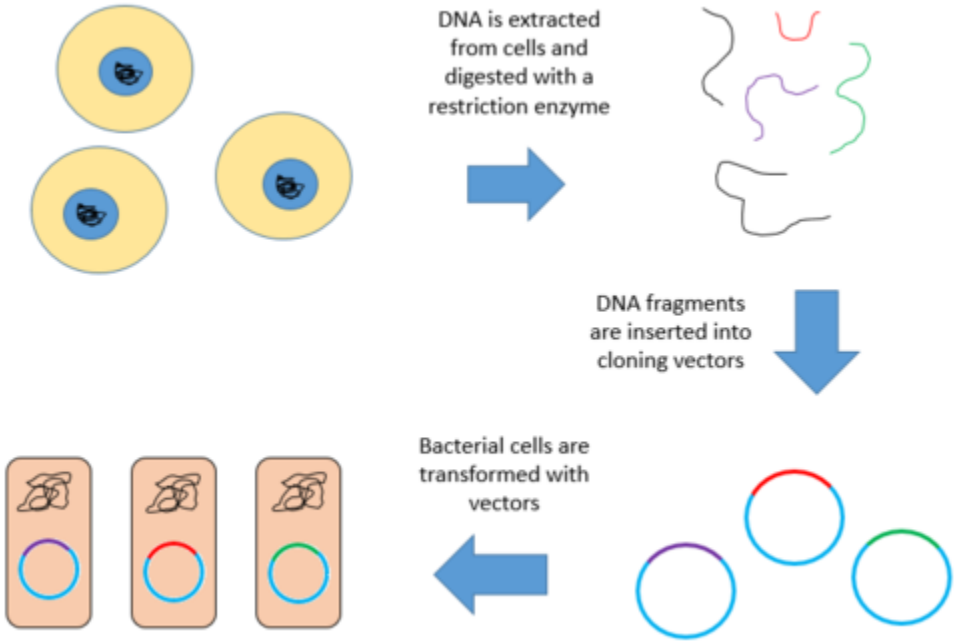
- A genomic library is a collection of the total genomic DNA from a single organism.
- The DNA is stored in a population of identical vectors, each containing a different insert of DNA.
- In order to construct a genomic library, the organism's DNA is extracted from cells and then digested with a restriction enzyme to cut the DNA into fragments of a specific size.
- The fragments are then inserted into the vector using DNA ligase.

- Next, the vector DNA can be taken up by a host organism - commonly a population of Escherichia coli or yeast - with each cell containing only one vector molecule. Using a host cell to carry the vector allows for easy amplification and retrieval of specific clones from the library for analysis.
- There are several kinds of vectors available with various insert capacities.
- Generally, libraries made from organisms with larger genomes require vectors featuring larger inserts, thereby fewer vector molecules are needed to make the library.
- Researchers can choose a vector also considering the ideal insert size to find the desired number of clones necessary for full genome coverage.
- Genomic libraries are commonly used for sequencing applications.
- They have played an important role in the whole genome sequencing of several organisms, including the human genome and several model organisms

Genomic library construction

- 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.

- Extract and purify DNA.
- Digest the DNA with a restriction enzyme. This creates fragments that are similar in size, each containing one or more genes.
- 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.
- These recombinant molecules are taken up by a host bacterium by transformation, creating a DNA library.

cDNA library

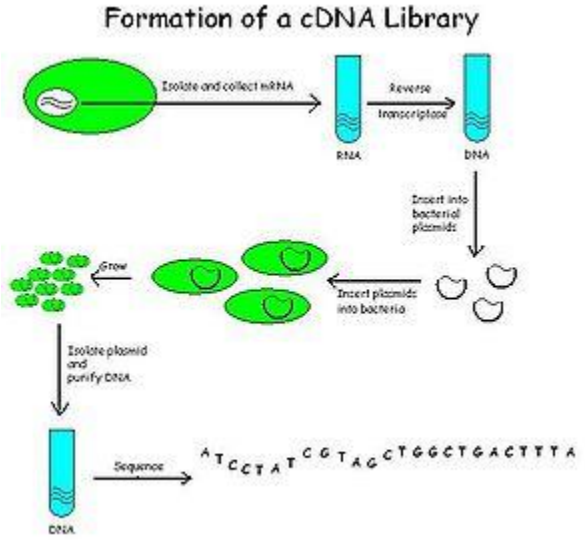
- A cDNA library is a combination of cloned cDNA (complementary DNA) fragments inserted into a collection of host cells, which constitute some portion of the transcriptome of the organism and are stored as a "library".
- cDNA is produced from fully transcribed mRNA found in the nucleus and therefore contains only the expressed genes of an organism.
- Similarly, tissue-specific cDNA libraries can be produced.
- In eukaryotic cells the mature mRNA is already spliced, hence the cDNA produced lacks introns and can be readily expressed in a bacterial cell.
- While information in cDNA libraries is a powerful and useful tool since gene products are easily identified, the libraries lack information about enhancers, introns, and other regulatory elements found in a genomic 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.
  - The mRNA is removed by using an RNAse enzyme, leaving single-stranded cDNA (sscDNA).
  - The sscDNA is converted into 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 coiling on itself at the 3' end, generating a hairpin loop.
  - The polymerase extends the 3'-OH end, and later the loop at 3' end is opened by the scissoring action of S1 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.

cDNA Library uses

- cDNA libraries are commonly used when reproducing eukaryotic genomes, as the amount of information is reduced to remove the large numbers of non-coding regions from the library.
- cDNA libraries are used to express eukaryotic genes in prokaryotes.
- Prokaryotes do not have introns in their DNA and therefore do not possess any enzymes that can cut it out during transcription process.
- cDNA does not have introns and therefore can be expressed in prokaryotic cells.
- cDNA libraries are most useful in reverse genetics where the additional genomic information is of less use.
- Additionally, cDNA libraries are frequently used in functional cloning to identify genes based on the encoded protein's function.
- When studying eukaryotic DNA, expression libraries are constructed using complementary DNA (cDNA) to help ensure the insert is truly a gene.

Site specific mutagenesis/Site directed mutagenesis

- **Site-directed mutagenesis** is a molecular biology method that is used to make specific and intentional mutating changes to the DNA sequence of a gene and any gene products.
- Also called **site-specific mutagenesis or oligonucleotide-directed mutagenesis**, it is used for investigating the structure and biological activity of DNA, RNA, and protein molecules, and for protein engineering.

- Site-directed mutagenesis is one of the most important laboratory techniques for creating DNA libraries by introducing mutations into DNA sequences.
- There are numerous methods for achieving site-directed mutagenesis, but with decreasing costs of oligonucleotide synthesis, artificial gene synthesis is now occasionally used as an alternative to site-directed mutagenesis.
- Since 2013, the development of the CRISPR/Cas9 technology, based on a prokaryotic viral defense system, has also allowed for the editing of the genome, and mutagenesis may be performed in vivo with relative ease.

Basic mechanism

- The basic procedure requires the synthesis of a short DNA primer.
- This synthetic primer contains the desired mutation and is complementary to the template DNA around the mutation site so it can hybridize with the DNA in the gene of interest.
- The mutation may be a single base change (a point mutation), multiple base changes, deletion, or insertion. The single-strand primer is then extended using a DNA polymerase, which copies the rest of the gene.
- The gene thus copied contains the mutated site, and is then introduced into a host cell in a vector and cloned. Finally, mutants are selected by DNA sequencing to check that they contain the desired mutation.

Approaches

- The original method using single-primer extension was inefficient due to a low yield of mutants.
- This resulting mixture contains both the original unmutated template as well as the mutant strand, producing a mixed population of mutant and non-mutant progenies.
- Furthermore, the template used is methylated while the mutant strand is unmethylated, and the mutants may be counter-selected due to

presence of mismatch repair system that favors the methylated template DNA, resulting in fewer mutants. Many approaches have since been developed to improve the efficiency of mutagenesis.

- A large number of methods are available to effect site-directed mutagenesis, although most of them have rarely been used in laboratories since the early 2000s, as newer techniques allow for simpler and easier ways of introducing site-specific mutation into genes.

1. Kunkel's method

- In 1985, Thomas Kunkel introduced a technique that reduces the need to select for the mutants.
- The DNA fragment to be mutated is inserted into a phagemid such as M13mp18/19 and is then transformed into an E. coli strain deficient in two enzymes, dUTPase (dut) and uracil deglycosidase (udg).
- Both enzymes are part of a DNA repair pathway that protects the bacterial chromosome from mutations by the spontaneous deamination of dCTP to dUTP.
- The dUTPase deficiency prevents the breakdown of dUTP, resulting in a high level of dUTP in the cell.
- The uracil deglycosidase deficiency prevents the removal of uracil from newly synthesized DNA. As the double-mutant E. coli replicates the phage DNA, its enzymatic machinery may, therefore, misincorporate dUTP instead of dTTP, resulting in single-strand DNA that contains some uracils (ssUDNA).
- The ssUDNA is extracted from the bacteriophage that is released into the medium, and then used as template for mutagenesis.
- An oligonucleotide containing the desired mutation is used for primer extension.
- The heteroduplex DNA, that forms, consists of one parental non-mutated strand containing dUTP and a mutated strand containing dTTP.

- The DNA is then transformed into an E. coli strain carrying the wildtype dut and udg genes.
- Here, the uracil-containing parental DNA strand is degraded, so that nearly all of the resulting DNA consists of the mutated strand.

2. Cassette mutagenesis

- Unlike other methods, cassette mutagenesis need not involve primer extension using DNA polymerase.
- In this method, a fragment of DNA is synthesized, and then inserted into a plasmid.
- It involves the cleavage by a restriction enzyme at a site in the plasmid and subsequent ligation of a pair of complementary oligonucleotides containing the mutation in the gene of interest to the plasmid.
- Usually, the restriction enzymes that cut at the plasmid and the oligonucleotide are the same, permitting sticky ends of the plasmid and insert to ligate to one another.
- This method can generate mutants at close to 100% efficiency, but is limited by the availability of suitable restriction sites flanking the site that is to be mutated

3. PCR site-directed mutagenesis

- Depiction of one common way to clone a site-directed mutagenesis library (i.e., using degenerate oligos).
- The gene of interest is PCRed with oligos that contain a region that is perfectly complementary to the template (blue), and one that differs from the template by one or more nucleotides (red).
- Many such primers containing degeneracy in the non-complementary region are pooled into the same PCR, resulting in many different PCR products with different mutations in that region.
- The limitation of restriction sites in cassette mutagenesis may be overcome using polymerase chain reaction with oligonucleotide

"primers", such that a larger fragment may be generated, covering two convenient restriction sites.

- The exponential amplification in PCR produces a fragment containing the desired mutation in sufficient quantity to be separated from the original, unmutated plasmid by gel electrophoresis, which may then be inserted in the original context using standard recombinant molecular biology techniques.
- There are many variations of the same technique. The simplest method places the mutation site toward one of the ends of the fragment whereby one of two oligonucleotides used for generating the fragment contains the mutation.
- This involves a single step of PCR, but still has the inherent problem of requiring a suitable restriction site near the mutation site unless a very long primer is used.
- Other variations, therefore, employ three or four oligonucleotides, two of which may be non-mutagenic oligonucleotides that cover two convenient restriction sites and generate a fragment that can be digested and ligated into a plasmid, whereas the mutagenic oligonucleotide may be complementary to a location within that fragment well away from any convenient restriction site.
- These methods require multiple steps of PCR so that the final fragment to be ligated can contain the desired mutation.
- The design process for generating a fragment with the desired mutation and relevant restriction sites can be cumbersome. Software tools like SDM-Assist can simplify the process.

4. Whole plasmid mutagenesis

- For plasmid manipulations, other site-directed mutagenesis techniques have been supplanted largely by techniques that are highly efficient but relatively simple, easy to use, and commercially available as a kit.

- An example of these techniques is the Quikchange method,[16] wherein a pair of complementary mutagenic primers are used to amplify the entire plasmid in a thermocycling reaction using a high-fidelity non-strand-displacing DNA polymerase such as pfu polymerase.
- The reaction generates a nicked, circular DNA. The template DNA must be eliminated by enzymatic digestion with a restriction enzyme such as DpnI, which is specific for methylated DNA.
- All DNA produced from most Escherichia coli strains would be methylated; the template plasmid that is biosynthesized in E. coli will, therefore, be digested, while the mutated plasmid, which is generated in vitro and is therefore unmethylated, would be left undigested.
- Note that, in these double-strand plasmid mutagenesis methods, while the thermocycling reaction may be used, the DNA need not be exponentially amplified as in a PCR. Instead, the amplification is linear, and it is therefore inaccurate to describe them as a PCR, since there is no chain reaction.

5. In vivo site-directed mutagenesis methods

- Delitto perfetto
- Transplacement "pop-in pop-out"
- Direct gene deletion and site-specific mutagenesis with PCR and one recyclable marker
- Direct gene deletion and site-specific mutagenesis with PCR and one recyclable marker using long homologous regions
- In vivo site-directed mutagenesis with synthetic oligonucleotides

Applications

- Site-directed mutagenesis is used to generate mutations that may produce a rationally designed protein that has improved or special properties (i.e.protein engineering).

- Investigative tools – specific mutations in DNA allow the function and properties of a DNA sequence or a protein to be investigated in a rational approach.
- Furthermore, single amino-acid changes by site-directed mutagenesis in proteins can help understand the importance of post-translational modifications.
- For instance changing a particular serine (phosphoacceptor) to an alanine (phospho-non-acceptor) in a substrate protein blocks the attachment of a phosphate group, thereby allows the phosphorylation to be investigated.
- This approach has been used to uncover the phosphorylation of the protein CBP by the kinase HIPK2
- Another comprehensive approach is site saturation mutagenesis where one codon or a set of codons may be substituted with all possible amino acids at the specific positions.
- Commercial applications – Proteins may be engineered to produce mutant forms that are tailored for a specific application. For example, commonly used laundry detergents may contain subtilisin, whose wild-type form has a methionine that can be oxidized by bleach, significantly reducing the activity the protein in the process.
- This methionine may be replaced by alanine or other residues, making it resistant to oxidation thereby keeping the protein active in the presence of bleach.

Gene targeting

- Gene targeting is a method for modifying the structure of a specific gene without removing it from its natural environment in the chromosome in a living cell.
- This process involves the construction of a piece of DNA, known as a gene targeting vector , which is then introduced into the cell where it replaces or modifies the normal chromosomal gene through the process of homologous recombination.

The Homologous Recombination Process

- Homologous recombination is a process that occurs within the chromosome and which allows one piece of DNA to be exchanged for another piece.
- It is a cellular mechanism that is probably part of the normal process cells use to repair breaks in their chromosomes.
- Homologous recombination requires that the pieces of DNA undergoing recombination be almost identical (homologous) in sequence.
- In addition, sequences on either side of the target should be identical, to promote more efficient targeting and recombination.
- By constructing a sequence that is homologous to a target sequence (such as a gene), laboratory researchers can replace one of the cell's own copies of a particular gene with a copy that has been altered in some way.
- It is also possible to replace only a part of a gene, such as one portion of its protein coding region.
- This permits the introduction of a mutation into specific cellular genes, which can either stop the gene functioning altogether (called a **"knock out "**) or can mimic changes to genes that have been implicated in human diseases.
- The ability to target DNA constructs to particular locations in chromosomes is a very powerful tool because it allows the modification of more or less any gene of interest, in more or less any way desired.
- Homologous recombination of a DNA vector into a gene of interest can be done in almost any cell type but occurs at a very low frequency, and it is therefore important to detect the few cells that have taken up the gene.
- Gene targeting vectors are designed with this in mind.

- The simplest strategy is to include an antibiotic resistance gene on the vector, which interrupts the sequence homologous to the gene of interest and thus makes the inserted gene nonfunctional.
- This **"selectable marker"** gene makes the cells that possess it resistant to antibiotics, and can then be used to eliminate cells that are not genetically modified.
- An example of a selectable marker that is commonly used for this purpose is the puromycin-N-acetyl-transferase (pac) gene, which confers resistance to the antibiotic puromycin, a drug that inhibits the function of ribosomes.
- After the introduction of the DNA construct, the cells are cultured with puromycin in the medium .
- This allows the selection of single cells that have incorporated the DNA construct into their own chromosomes.
- Cells lacking the pac gene will die in a culture medium containing puromycin.
- Once the puromycin resistant cells have been expanded into cell lines, the DNA of these cells can then be analyzed to select out a subset of the cells in which the introduced construct has integrated into the correct (target) gene.
- For reasons that are not yet fully understood, the rate of homologous recombination in mouse embryonic stem (ES) cells is substantially higher than that of most other cells.
- Once a clone of ES cells with the correct targeting event has been identified, these cells can be used to introduced into the mouse via the process of blastocyst injection, which allows the study of gene function in the bodies of living, intact animals. Until very recently mice were the only organisms in which it has been possible to introduce targeted mutations into the germ line.

- The development of nuclear transfer (moving the nucleus from one cell to another), however, has allowed gene targeting to be done in other mammalian species, such as sheep and pigs.

Adding or Deleting Genetic Material

- As well as mutating or knocking out specific genes, gene targeting allows the introduction of novel pieces of DNA into a specific chromosomal location (this is often termed a "knock-in").
- This allows researchers to examine the function of a gene in a variety of ways.
- For example, it is possible to examine where in the animal the gene is normally expressed by insertion (knock-in) of a fluorescent protein (such as green fluorescent protein, GFP) into the gene so that the cells expressing the gene begin to glow.
- In addition to changing single genes it is also possible to remove or alter large pieces of chromosomes.
- Technologies also now exist that allow genes to be removed not just in a whole animal, as described above, but in a subset of cells or in a particular tissue.
- This can be achieved by modifying the vector to include target sites (termed loxP sites) for an enzyme called Cre recombinase.
- When the Cre enzyme is present in a mouse cell in which the target gene is surrounded by loxP sites, it will cut this gene out of the chromosome.
- This allows the function of this gene, which may be required for the mouse to normally develop, to be analyzed in a particular cell type or tissue where only the Cre recombinase is expressed.

Therapeutic Potential of Gene Targeting

- It is hoped that gene targeting may eventually become useful in treating some human genetic disorders such as hemophilia and Duchenne muscular dystrophy.

- Treating human disease by the types of genetic approaches mentioned above is termed **"gene therapy."**
- This could, in principle, be achieved by replacing the defective gene with a normal copy of the gene in the affected cells of an individual undergoing treatment.
- In order to make this potential treatment effective it will be necessary to develop technologies that increase the frequency with which targeting occurs. This is currently the subject of much research.
- The development of nuclear transfer technology also has opened up the possible alternative method of using homologous recombination for gene therapy based on cell transfer.
- Gene targeting would be used to replace the defective genes in selected somatic cells in culture, and their nuclei could then be transferred into stem cells.
- The stem cells can then be differentiated into the affected cell type (for example, into bone marrow cells for hemophilia) and these cells could then be transplanted to patients.

Human Genome Project

Human genome project (HGP) was an international scientific research project which got successfully completed in the year 2003 by sequencing the entire human genome of 3.3 billion base pairs. The HGP led to the growth of bioinformatics which is a vast field of research. The successful sequencing of the human genome could solve the mystery of many disorders in humans and gave us a way to cope up with them.

Goals of the human genome project

Goals of the human genome project include:

- Optimization of the data analysis.

- Sequencing the entire genome.
- Identification of the complete human genome.
- Creating genome sequence databases to store the data.
- Taking care of the legal, ethical and social issues that the project may pose.

Methods of the human genome project

In this project, two different and significant methods are typically used.

- Expressed sequence tags wherein the genes were differentiated into the ones forming a part of the genome and the others which expressed RNAs.
- Sequence Annotation wherein the entire genome was first sequenced and the functional tags were assigned later.

The process of the human genome project

- The complete gene set was isolated from a cell.
- It was then split into small fragments.
- This DNA structure was then amplified with the help of a vector which mostly was BAC (Bacterial artificial chromosomes) and YAC (Yeast artificial chromosomes).
- The smaller fragments were then sequenced using DNA sequencers.
- On the basis of overlapping regions, the sequences were then arranged.
- All the information of this genome sequence was then stored in a computer-based program.
- This way the entire genome was sequenced and stored as genome database in computers. Genome mapping was the next goal which was achieved with the help of microsatellites (repetitive DNA sequences).

Features

Features of the Human genome project include:

- Our entire genome is made up of 3164.7 million base pairs.

- On average, a gene is made up of 3000 nucleotides.
- The function of more than 50 percent of the genes is yet to be discovered.
- Proteins are coded by less than 2 percent of the genome.
- Most of the genome is made up of repetitive sequences which have no coding purposes specifically but such redundant codes can help us better understand of genetic development of humanity through the ages.

Applications of HGP

- As the goals of the human genome project were achieved, it led to great advancement in research.
- Today, if any disease arises due to some alteration in a certain gene, then it could be traced and compared to the genome database that we already have.
- In this way, a more rational step could be taken to deal with the problem and can be fixed with more ease.

Human gene therapy

With the advancement of medical science, several treatments have been discovered to treat deadly diseases. However, there are a few diseases that are genetic and as a result, are not easily curable. In a few such cases, the application of gene therapy has been found to be useful.

It is an experimental technique through which healthy genes are inserted into an individual or embryo to treat disease. Gene therapy paves ways to replace faulty or mutated genes with new ones.

How does Gene Therapy Works?

There are several ways through which gene therapy works.

- Replace a mutated gene with a healthy version of that gene.
- Introduce a new functioning gene to fight disease.
- Inactivate a faulty gene that is causing disease.

This gene therapy diagram shows that first, the defective genes are spotted. Then, medical experts use healthy genes to replace faulty ones. Finally, the new gene restores the functionality of the existing cells. Some portions of DNA containing useful proteins enter the cells through the vectors. Inside the cell, DNA/genes start making useful proteins. After some time, the damaged cells heal and remove the source of the disease.

Different Kinds of Gene Therapy

Primarily, there are **two types** of gene therapy.

- **Somatic Gene Therapy**

  The human body mainly consists of somatic or stem cells. This process uses healthy genes to replace damaged ones. The therapy targets the defective cells of an individual who is suffering from a disease. Somatic cells are mainly non-reproductive. That means the effects of this therapy will not transfer to the future generation. Hence, it is considered to be one of the safest applications of gene therapy.

- **Germline Gene Therapy**

  This therapy targets the germ cells of the body that produce eggs or sperms. Germline gene therapy process includes the infusion of functional DNA into cells. However, the effect of this therapy can affect future generations. Therefore, the usage of this therapy is restricted in many places. For example, the European Union does not allow this process.

Gene Therapy Application

With time, the popularity of this therapy is increasing. The application of gene therapy includes the following:

- Effectively cures several genetic disorders.
- Treats diseases like brain tumours, Alzheimer's, Parkinson's, Haemophilia, and several others.
- Useful for the diseases that traditional medicine cannot cure.
- Solely destroys disease-causing cells without affecting other cells.
- Can be used on individuals, as well as embryos.

However, this treatment has some temporal or permanent side effects too.

Challenges of Gene Therapy

- The new genes have to reach the right place.
- On reaching the exact location, this gene has to start becoming functional.
- The genes can cause harm if they reach the wrong cells.
- Sometimes targeted cells stop the new genes from entering. The immune system of a body often also tries to kill the inserted gene.
- It has to make sure that the new cells are affecting the functions of other cells.

ENTRI

Tansgenic Animals

- **Trangenic animals** are the animals with the modified genome.
- A foreign gene is inserted into the genome of the animal to alter its DNA.
- This method is done to improve the genetic traits of the target animal.
- Initially, the improvement of genetic traits was done by selective breeding methods.
- In this, the animals with desired genetic characteristics were mated to produce an individual with improved genetic characteristics.
- Since this technique was time-consuming and expensive, it was later replaced by recombinant DNA technology.
- **Transgenesis** is the phenomenon in which a foreign gene with desired characteristics is introduced into the genome of the target animal.
- The foreign gene that is introduced is known as the transgene, and the animal whose genome is altered is known as transgenic.
- These genes are passed on to the successive generations.
- The transgenic animals are genetically engineered and are also known as genetically modified organisms.
- The first genetically modified organism was engineered in the year 1980.
- Let us have a detailed look at the process, importance and applications of transgenic animals.

Methods for Creating Transgenic Animals

The transgenic animals are created by the following methods:

- **Physical Transfection**
    - In this method, the gene of interest is directly injected into the pronucleus of a fertilized ovum. It is the very first method that proved to be effective in mammals. This method was applicable to a wide variety of species. Other methods of physical transfection include particle bombardment, ultrasound and electroporation.

- **Chemical Transfection**
    - One of the chemical methods of gene transfection includes transformation. In this method, the target DNA is taken up in the presence of calcium phosphate. The DNA and calcium phosphate co-precipitates, which facilitates DNA uptake. The mammalian cells possess the ability to take up foreign DNA from the culture medium.

- **Retrovirus-Mediated Gene Transfer**
    - To increase the chances of expression, the gene is transferred by means of a vector. Since retroviruses have the ability to infect the host cell, they are used as vectors to transfect the gene of interest into the target genome.

- **Viral Vectors**
    - Viruses are used to transfect rDNA into the animal cell. The viruses possess the ability to infect the host cell, express well and replicate efficiently.

- **Bactofection**
    - It is the process by which the gene of interest is transferred into the target gene with the help of bacteria.

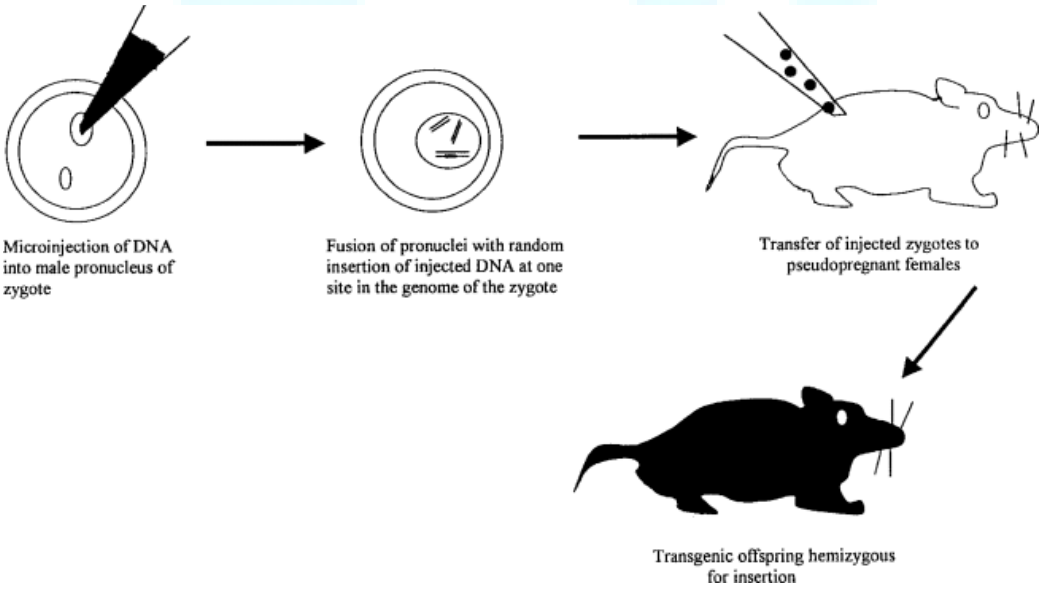
Examples of Transgenic Animals

Following are the examples of transgenic animals:

- **Dolly Sheep**
    - Dolly the sheep was the first mammal to be cloned from an adult cell. In this, the udder cells from a 6-year-old Finn Dorset white sheep were injected into an unfertilized egg from a Scottish Blackface ewe, which had its nucleus removed. The cell was made

to fuse by electrical pulses. After the fusion of the nucleus of the cell with the egg, the resultant embryo was cultured for six to seven days. It was then implanted into another Scottish Blackface ewe which gave birth to the transgenic sheep, Dolly.

- **Transgenic Mice**
  - Transgenic mice are developed by injecting DNA into the oocytes or 1-2 celled embryos taken from female mice. After injecting the DNA, the embryo is implanted into the uterus of receptive females.



Applications Of Transgenic Animals

The transgenic animals are created because of the benefits they provide to the man. Let us discuss a few of them here.

- **Normal Physiology and Development**
    - In transgenic animals, a foreign gene is introduced due to which the growth factor is altered. Hence, these animals facilitate the study of gene regulation and their effect on the everyday functions of the body.

- **Study of Diseases**
    - Transgenic animals are specially designed to study the role of genes in the development of certain diseases. Moreover, in order to devise a cure for these diseases, the transgenic animals are used as model organisms. These transgenic models are used in research for the development of medicines. For example, we have transgenic models for diseases such as Alzheimer's and cancer.

- **Biological Products**
    - A number of biological products such as medicines and nutritional supplements are obtained from transgenic animals. Research for the manufacture of medicines to treat diseases such as phenylketonuria (PKU) and hereditary emphysema is going on. The first transgenic cow, Rosie (1997), produced milk containing human protein (2.4 grams per litre). This milk contains the human gene alpha-lactalbumin and could be given to babies as an alternative to natural cow milk.

- **Vaccine Safety**
    - Transgenic animals are used as model organisms for testing the safety of vaccines before they are injected into humans. This was conventionally done on monkeys.