College of Engineering, Pune

(An Autonomous Institute of Government of Maharashtra)

Applied Science Department

Tel: 020-25507034 Fax: 020-25507219

CT16002 – Biology for Engineers

UNIT IV: Expression and transmission of Genetic Information

DNA replication, Enzyme driven process of DNA cloning, Protein synthesis- Transcription & translation Techniques for optimization: a. At molecular level: Recombinant DNA Technology, DNA hybridization, PCR, DNA microarray

---

DNA Replication

DNA is capable of self reproduction. Parent DNA produces two daughter DNA molecules which are exact copies / replicas of parent DNA in N2 base sequence. Hence DNA duplication is called as DNA replication.

Definition-- It is making exact copies of parent DNA. It is bidirectional.

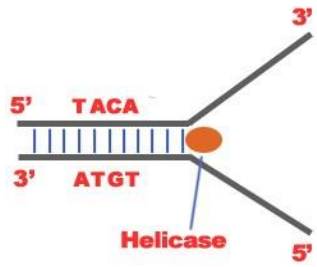
Steps of DNA Replication --

1) The first major step for the **DNA Replication** to take place is the breaking of hydrogen bonds between bases of the two antiparallel strands. The unwounding of the two strands is the starting point.

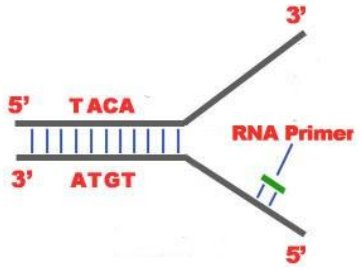
2) Activation of deoxyribonucleotides by energy & enzyme **Phosphorylase.**

3) Enzyme **Endonuclease** makes cut to one of the strands of DNA. The splitting happens in places of the chains which are rich in A-T. That is because there are only two bonds between Adenine and Thymine (there are three hydrogen bonds between Cytosine and Guanine). **Helicase** is the enzyme that splits the two strands. The initiation point where the splitting starts is called "origin of replication". The structure that is created is known as "**Replication Fork".**

4) Topoisomerase, helix destabilizing protein stabilize replication fork.



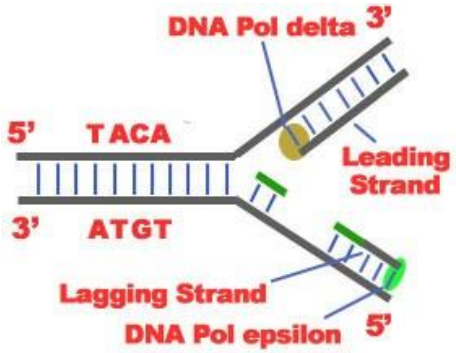
5) Primase synthesizes RNA primer that attaches at 3' end. RNA primer functions as 5' end of new strand.



6) Two separated strands function as templates. Initiation of replication occurs at 3' end.

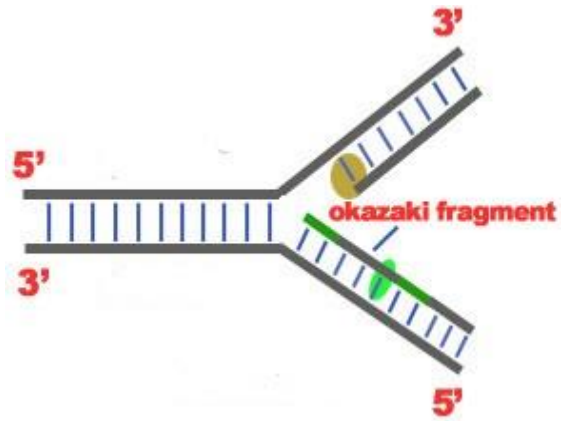
7) DNA polymerase adds nucleotides in 5' → 3'direction continuously complementary to the nucleotides of the template, e.g. A—T. This strand is called **leading strand.**

8) On the other strand DNA polymerase adds nucleotides in 5' → 3'direction in short fragments. Hence it is called as **lagging strand**. The newly synthesized DNA fragments are called as **Okazaki fragments. These fragments are joined by DNA ligase & become continuous.**



9) Once replication is completed, RNA primer is removed& DNA nucleotides are synthesized by **DNA polymerase.**

10) Proof reading – Mismatched N2 bases are removed by **endonuclease** & appropriate N2 bases are introduced by **DNA polymerase**.
11) In daughter DNA, 2 strands coil around each other to form helix**.**



12) Thus in daughter DNA one strand is old i.e. conserved & another strand is new. Hence it is known as **"semiconservative replication".**

Protein Synthesis:

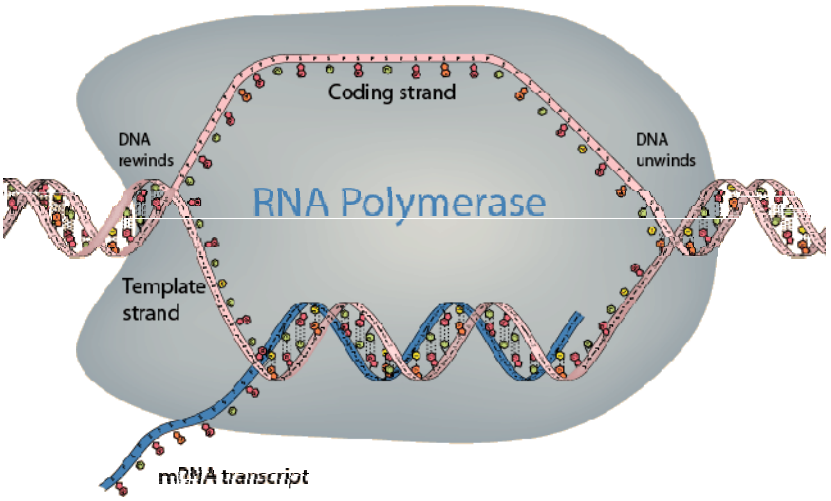
Protein synthesis is the main mechanism in body with respect to growth and changes in the cell. It results in the production of amino acid chains which are for proteins (important component in body). But for short term reactions importance of protein synthesis is the production of variety enzymes for different reactions as needed by the body for that moment. Since we cannot exist without enzymes, protein synthesis is needed for our existence. The process of protein synthesis translates the codons (nucleotide triplets) of the messenger RNA (mRNA) into the 20-symbol code of amino acids that build the polypeptide chain of the proteins.

DNA ⟹ mRNA ⟹ PROTEIN is known as CENTRAL DOGMA in the process of protein synthesis. It has 2 steps ⇒ – transcription (DNA ~~mRNA~~) and translation (mRNA PROTEIN).

TRANSCRIPTION

The first step in protein synthesis is the transcription of mRNA from a DNA gene in the nucleus. In this phase, one strand of DNA double helix acts as a template to synthesize its complimentary strand i.e. mRNA. Transcription starts with an enzyme called **polymerase** copying the DNA sequence to a similar molecule called messenger RNA (mRNA). This synthesis takes place in a specific portion of DNA which is known as transcription bubble. In the bubble, DNA strands are separated or unzipped

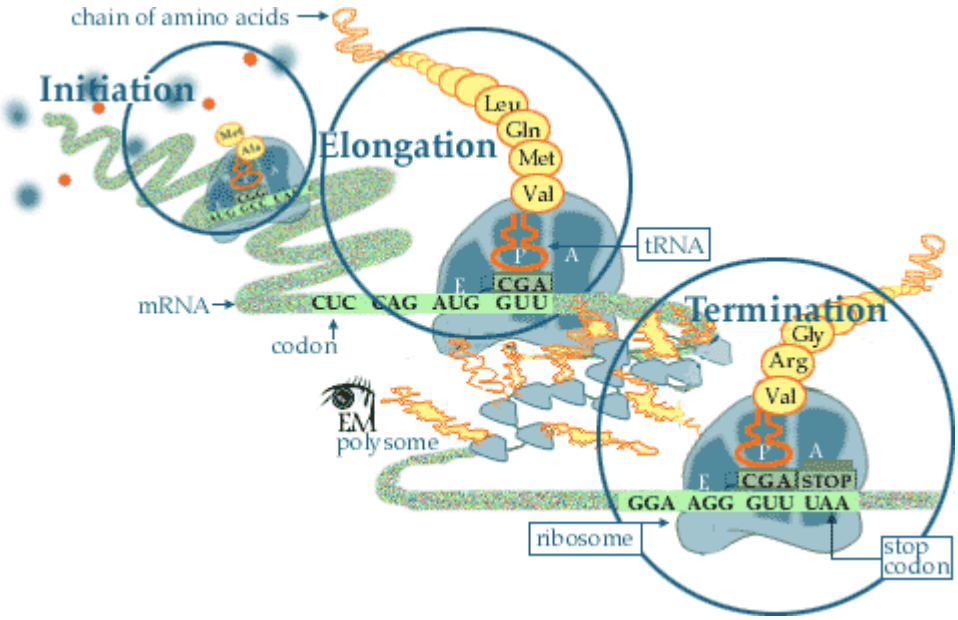
The process of mRNA translation begins from its 5′-end towards its 3′-end as the polypeptide chain is synthesized from its amino-terminal (N-end) to its carboxyl-terminal (C-end). It replaces T with U (Uracil), a helper base, making it clear that the mRNA is a copy. The bases (A, T, G, C) on one strand of the DNA specify the order of bases on the new strand of mRNA (A, U, G, C). At the end of transcription, DNA stays inside the nucleus while the RNAs migrate from the nucleus into the cytoplasm.



TRANSLATION

This is the second phase of protein synthesis where the ribosomes in the cytoplasm use transfer RNA (tRNA) to attach to the mRNA and translate the bases into amino acids. tRNA molecules bring the specified amino acids that the ribosome links together to make a protein. Translation has four steps viz. Activation and charging, Initiation, Elongation and Termination

Before the actual translation, **activation**, which is not part of translation in technical sense, occurs. In this step, the correct amino acid (AA) is joined to the correct t RNA. It is required for translation to proceed. When the activated tRNA has an amino acid linked to it, it is termed as "**charged**". The AA is joined by its carboxyl group to the 3' OH of the tRNA by an ester bond.



Initiation: In the cytoplasm, protein synthesis is actually initiated by the AUG codon on mRNA. The AUG codon signals both the interaction of the ribosome with m-RNA and also the tRNA with the anticodons (UAC). The tRNA which initiates the protein synthesis has N-formyl-methionine attached. The formyl group is really formic acid converted to an amide using the $-NH_2$ group on methionine (left most graphic)

The next step is for a second tRNA to approach the mRNA (codon -CCG). This is the code for proline. The anticodon of the proline tRNA which reads this is GGC. The final process is to start growing peptide chain by having amine of proline to bond to the carboxyl acid group of methinone (met) in order to elongate the peptide.

Elongation: Elongation of the peptide begins as various tRNA's read the next codon. In the example on the left the next tRNA to read the mRNA is tyrosine. When the correct match with the anticodons of a tRNA has been found, the tyrosine forms a peptide bond with the growing peptide chain. The proline is now hydrolyzed from the tRNA. The proline tRNA now moves away from the ribosome and back into the cytoplasm to reattach another proline amino acid.

Elongation and Termination: When the stop signal on mRNA is reached, the protein synthesis is terminated. The last amino acid is hydrolyzed from its t-RNA. The peptide chain leaves the ribosome. The N-formyl-methionine that was used to initiate the protein synthesis is also hydrolyzed from the completed peptide at this time.

The ribosome is now ready to repeat the synthesis several more times. The events following translation include post-translational modification and protein folding. During

and after synthesis, polypeptide chains often fold to attain secondary, tertiary and quaternary structures. This process is known as **maturation of protein** and it takes place in Golgi Complex.

GENETIC ENGINEERING

Recombinant DNA Technology, **Genetic modification/manipulation (GM)** and **Gene splicing** are terms that apply to the direct manipulation of an organism's genes. Genetic engineering is different from traditional breeding, where the organism's genes are manipulated indirectly.

Definition --Genetic engineering uses the techniques of molecular cloning and transformation to alter the structure and characteristics of genes directly.

Genetic engineering techniques have found some successes in numerous applications. Some examples are in improving crop technology, the manufacture of synthetic human insulin through the use of modified bacteria, the manufacture of erythropoietin in hamster ovary cells, and the production of new types of experimental mice such as the oncomouse (cancer mouse) for research.

Definition Clone: A clone is a group of identical copies. A clone of a cell is a group of cells of a single type isolated and allowed to reproduce to create a population of identical cells. A clone of a DNA molecule is an isolated DNA molecule which is multiplied number of times to produce a large amount of identical copies.

Definition Cloning : It is a method of producing identical copies of cells / molecules/organisms.

There are a number of ways through which genetic engineering and cloning together is accomplished. Essentially, the process has following steps :-

1. Isolation of the gene/ DNA fragment of interest (known function) from an organism **(A)** . It is known as an **insert.**
2. Enzymatic cleavage **(B)** and joining **(C)** of insert DNA to another DNA molecule (cloning vector) to form recombinant DNA (rDNA) i.e. vector + insert DNA **(D)**
3. Transformation of a host cell that includes transfer and maintenance of rDNA molecule in the host organism **(E)**
4. Identification of transformed cells (with rDNA) and their selection from non transformants
5. Amplification of rDNA to get multiple copies in a cell **(F)**
6. Cell multiplication **(G)** to get clones (population of genetically identical individuals carrying multiple copies of foreign DNA)

Isolation of gene of interest:

- Isolation is achieved by identifying the gene of interest that the user wishes to insert into the organism, usually on the basis of existing knowledge about various functions of genes. This segment of DNA is the molecule is used for the cloning process.
- Isolation is carried out by using variety of Restriction Endonucleases or Restriction Enzymes (e.g. *Eco*R1) that recognize the site of cleavage (cut) on DNA at specific palindromic sequences.

- Broadly these enzymes are categorized in three categories Type I, II and III. Recognition site and cleavage site is same in Type II while it is different in Type I and III. Thus, type II enzymes are widely used in genetic engineering especially for gene manipulation.

The details of these groups are as follows

Type I *Eco*B, *Eco*K	Recognition site for type I enzyme is 15 bp long and cleavage site 1000 bp away; the enzyme has restrictive subunit, modification subunit and specificity subunit; requires Mg^{++}, S-adenosyl methionine and ATP as cofactors
Type II *Eco*R1/R2 *Hae* III *BamH*1	The length of recognition site for type II enzyme varies from 4/5/6/8 or more bp long and cleavage site 1000 bp away; the enzyme has restrictive subunit, modification subunit and specificity subunit; more stable and requires Mg^{++} as cofactors
Type III *Mbo*II *Fok*I	Made up of 2 subunits for recognition and cleavage; requires ATP for energy and Mg as co factor; recognition site is non palliondromic and cleavage site 25-27 bp away

- As a result of these enzymes, the DNA from donor cell is cut in such a way that it has sticky/staggered cuts or sometimes blunt ends.

Insertion of the above mentioned gene of interest into a cloning / transfer vector

- This can be done with the opening of vector DNA molecule with a cut by the same restriction enzyme (i.e. producing similar kind of the sticky or blunt ends).
- Thus the foreign DNA and the vector DNA (plasmid) now share the complementarities at these ends.
- Then the addition of DNA ligase joins the two DNA molecules by ligating their nucleotides with each other. This prepares the recombinant DNA vector i.e. the recombinant plasmid.
- A vector is a cloning vehicle i.e. the agent used to multiply the isolated gene. It itself is a synthesized DNA molecule that carries the isolated gene into a host where it can replicate producing many of its copies.
- Hence the same number of the copies the gene of interest are generated. The product thus generated is called the **Recombinant DNA** (rDNA)and the technique is called gene cloning.
- A variety of vectors which are acutally the cloning vehicles have been developed which allow the multiplication of the inserted gene.
- The most common vectors are plasmids. Other vectors can also be used, such as viral vectors, and non-prokaryotic ones such as liposomes, or even direct insertion using DNA guns.

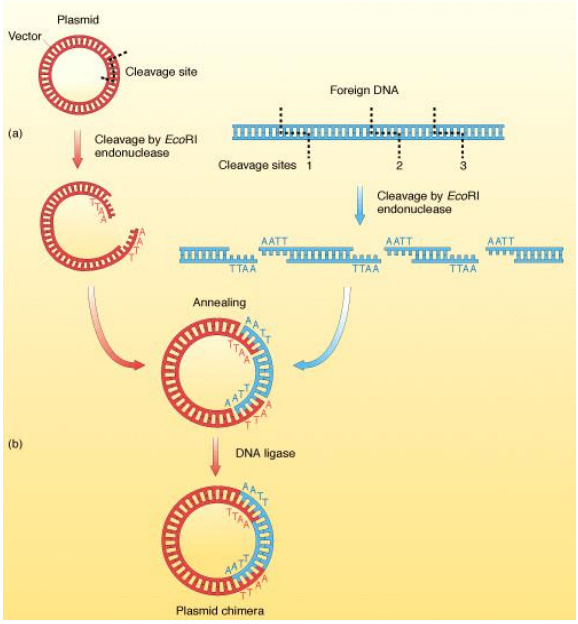
Commonly used vectors:

- **Plasmids**: Extra-chromosomal, closed, circular, double stranded DNA molecules present in bacterial cells. May present independently or integrated in the bacterial chromosome. Capable of independent replication and

transmission. Many times

specify number of properties of the host. e.g. antibiotic resistance, heavy metal resistance, nitrogen fixation. Plasmids are used with *E. coli.* cells. They are introduced in the bacterial cells by the process of transformation. The *E. coli* cells and plasmid DNA are incubated together at 0°C in $CaCl_2$ solution, then subjected to the heat shock by rapidly changing the temperature to 43°C. Few cells of the bacteria take up the plasmid DNA. Plasmid DNA can also be introduced in *E. coli* cells by electroporation during which the mixture of the cells and the plasmid is subjected to a high voltage pulse. The cell membrane becomes permeable and the plasmid can enter the cells.

- **Bacteriophages**: Also called phages which are the viruses that infect bacteria. They have the ability to attach the host (bacterium) and transfer their own DNA to the host. Most frequently used phage vectors are lambda (λ) phage and M13. The genome of λ phage is a linear double stranded with 12 bases long single stranded strech at 5' end which can be used as cohesive/sticky ends. Also there is a large non-essential region in its DNA which can be replaced with the foreign DNA. Such λ vectors can accommodate large DNA inserts around 40,000 – 53,000 b p (base pair) long. These vectors can then be packaged into the infectious phage particles. The infectious phages bring about the lysis of the host cell.
- BACs: Bacterial artificial chromosomes, A **bacterial artificial chromosome (BAC)** is a DNA construct, from plasmids. They can contain very long DNA inserts (100,000-300,000) as a foreign DNA.
- YACs: **Yeast artificial chromosome** (short **YAC**) is a vector used to clone large DNA fragments (larger than 100 kb and up to 3000 kb). It is an artificially constructed chromosome and contains the sequences needed for replication and preservation in yeast cells.
- **Shuttle vectors**: these are the plasmid DNAs that can be propagated in the cells of 2 or more different species.



Transfer of the vector by transformation of the host cell is the next step. The commonest type of host organism used for introducing the vectors (plasmids and viral vectors) is *Escherichia coli*. Its DNA metabolism is well-understood and plasmids are the naturally occurring circular DNA molecules in the *E. coli*.

Transformation is the process by which the *E. coli* are made receptive by some heat shock or electric shock and the recombinant plasmid DNA is kept in the surrounding medium of such competent *E. coli* cells. The recombinant plasmid DNA enters the *E. coli* cells. Such organisms are called **GMO or Genetically Modified Organisms**.

Selection of the transformed organisms is required to find out only those organisms which have received the recombinant from those which have not. The selection is carried out by various ways –

a) Direct Selection:

- If the cloned DNA itself codes for resistance to the antibiotic ampicillin (amp^R), the recombinants can be allowed to grow on the minimal medium containing ampicillin.
- Thus only such recombinants will grow and develop colonies on medium, which contain amp^R gene on its plasmid vector.
- But with this method one can not segregate the transformed cells with re-ligated plasmid (i.e. cut by endonuclease & joined again)vector from the recombinant plasmid.

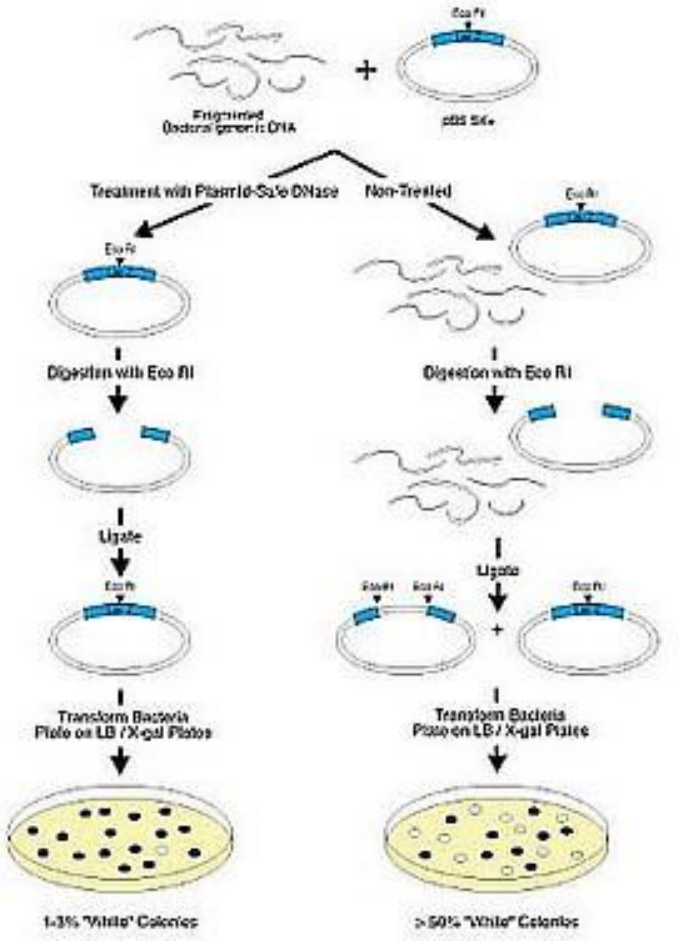
b) Insertional Selection Inactivation:

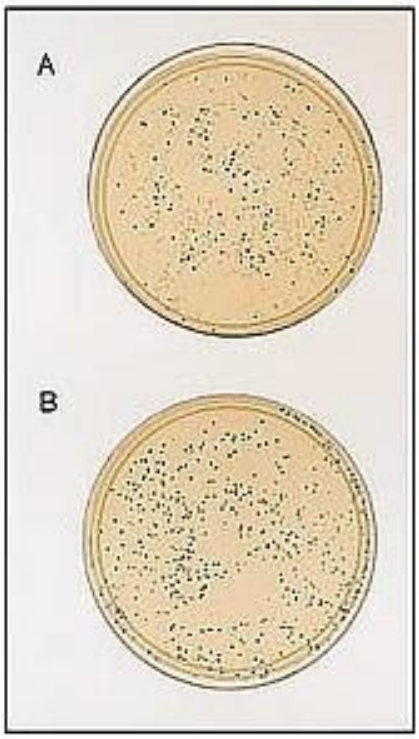
- In this method, as discussed earlier, the vector plasmid is constructed with some specific DNA sequences which confer the antibiotic resistance to the host cell to two different antibiotics (e.g. ampicillin and tetracyclin).
- It is designed in such a way that the foreign gene gets inserted in this part of the plasmid disrupting one of these gene. As a result the resistance for one antibiotic is lost by the recombinant plasmid whereas the plasmid DNA without the insert (foreign DNA) has the intact resistance for both the antibiotics.
- Thus the host cells that have not taken the plasmid cannot grow on the medium containing both the antibiotics.
- Those cells which that get transformed by the non-recombinant DNA have intact resistance for both the antibiotics.
- The *E. coli* cells that get transformed by the recombinant plasmid lose resistance for one of the antibiotics and retain for the other.
- Hence the cells are first grown on medium with ampicillin and then on medium with tetracyclin
- This strategy can be exploited for identifying the host cells that have taken the foreign gene. These cells are then grown in cultures to produce multiple copies of the foreign gene (as explained in the diagrammatic representation).

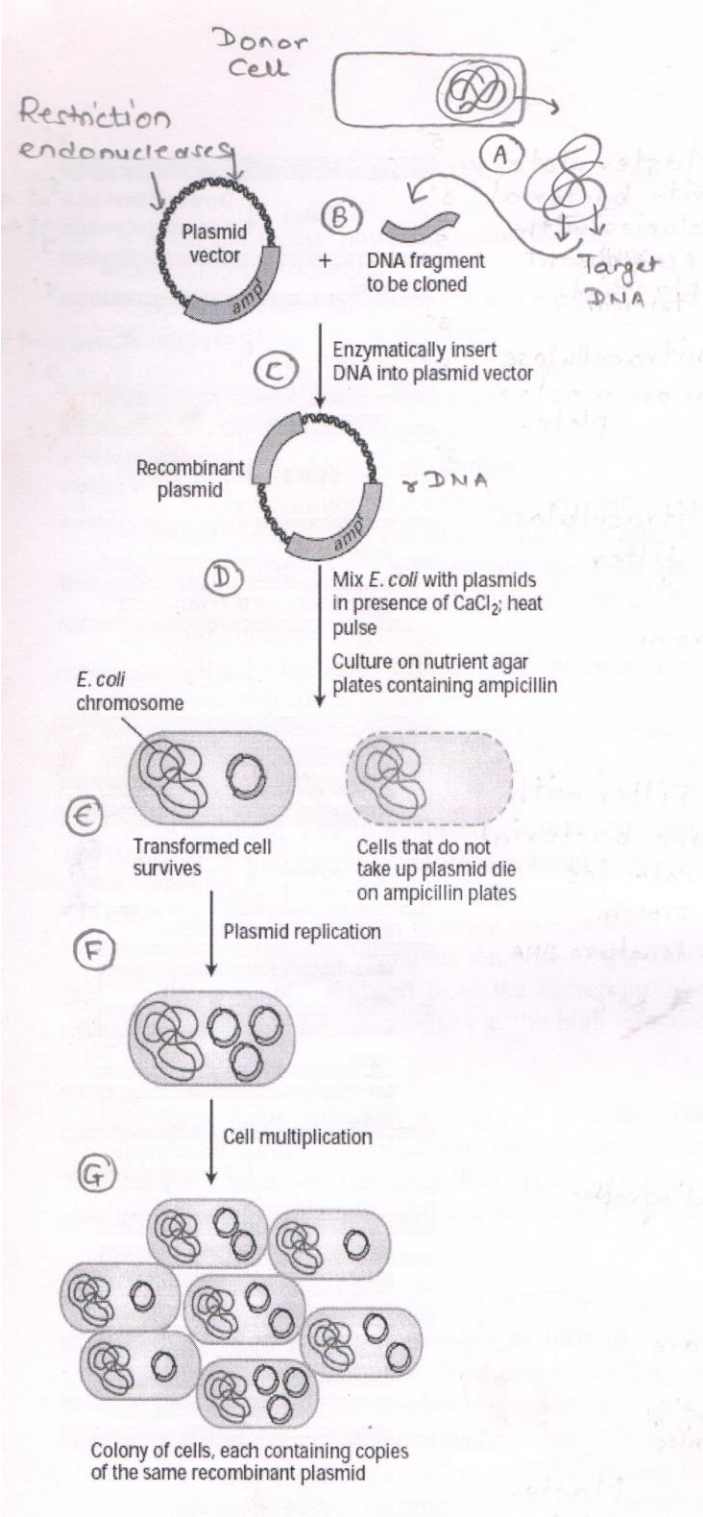
c) Blue White Colony

- In this method instead of antibiotic resistance gene, a lacZ gene is constructed which is responsible for the synthesis of beta-galactosidase enzyme.

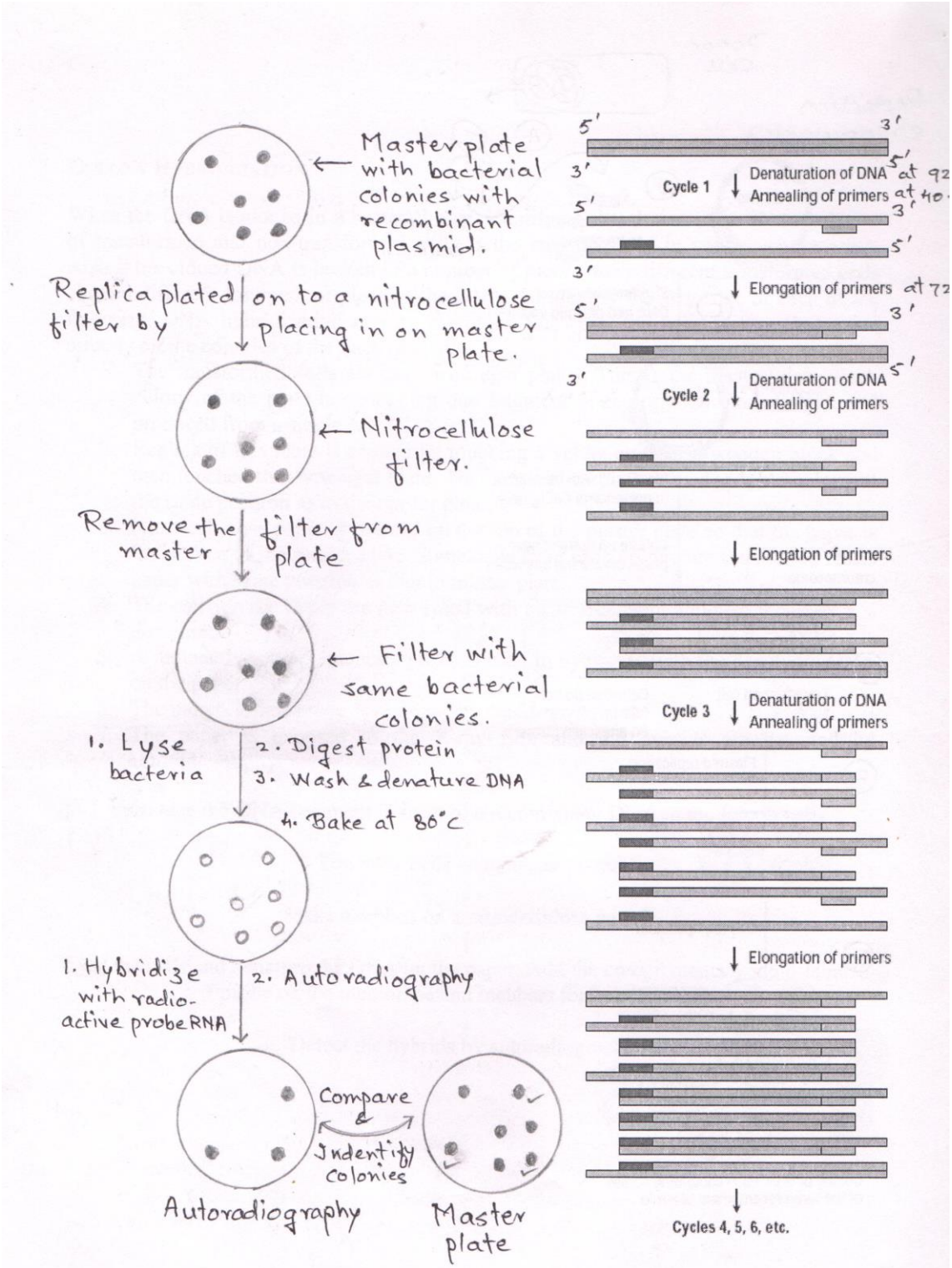
- The enzyme reacts with chemical B galactose and develops blue colour.
- In the recombinant plasmid, the introduced gene disrupts(alter function) lacZ gene and thus prevents further reaction if grown on the medium containing B galactose.
- Against this, the non recombinant plasmid certainly develops blue colour on the growing plate and can easily be differentiated from white colonies of recombinant plasmids which use glucose from the medium.





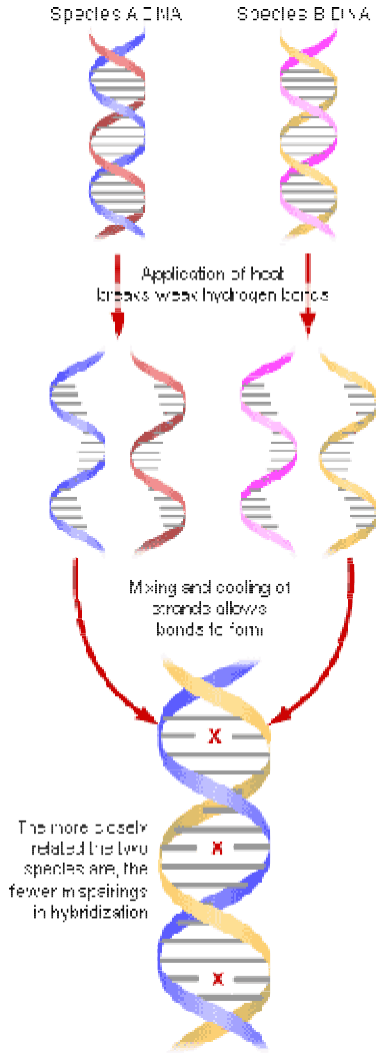


Gene Cloning



Colony Hybridization Polymerase Chain reaction

DNA HYBRIDIZATION:



- Hybridization technique makes use of the ability of DNA molecule to renature or reanneal when it is converted to single stranded structure.
- It is the method that detects DNA by its hydrogen bonding capability - we call this "probing" the DNA, and we say that the probe "hybridizes" to the target sequence.
- Also if any two DNA molecules share any complementarity, they can bind each other to form a double stranded DNA molecule.
- When any two different DNA samples are to be compared they are first denatured. They are then mixed and the mixture is cooled i.e. allowed to renature.
- During renaturation, the two DNA strands of the original molecule anneal and also the strands of the two different molecules anneal if they have significant similarity. Depending upon the extent of the similarity between them, they will form DNA duplexes of various lengths. Thus they form hybrids (if the renatured DNA has formed from the two different sources present in the mixture).
- These hybrids are the partial duplexes.
- This technique is highly useful in detecting similarity between two genes or DNA molecules form different organisms or species etc.
- Similarly, a small complementary probe also can hybridize with the small region of the DNA to be detected or studied. This probe is usually labeled with radioactive or fluorescent label.

The level of the radioactivity or fluorescence is detected after the two DNA solutions are mixed and allowed to renature. The intensity of the label is directly proportional to the complementarity of the two DNA molecules. This method can be implemented to detect a gene sequence.

Like DNA – DNA hybridization RNA –DNA hybridization can also be carried out.

The overall process can be summarize as follows,

Denatured DNA sample 1 + Denatured DNA sample 2 → Cool for renaturation and/or hybridization → Hybrid DNA duplexes → Screen the label by its radioactivity or fluorescence.

COLONY HYBRIDIZATION

When the DNA is cloned in a bacterial host, identifying the cloned DNA in the mixture of transformed and non-transformed cells is the step involved in genetic engineering. Also if the cloned DNA is broken into number of pieces, then different transformed cells contain different fragment or piece of the DNA. To identify the sequence of each of the fragment, DNA hybridization is very useful. In this case, the hybridization is carried directly on the colonies of the bacterial cells.

1. The transformed cells are grown on agar plates. This is the master plate. Each colony of the plate is containing one fragment of the original cloned DNA and produced from a single cell.
2. Replica of this plate is created by touching a velvet cloth or a wooden block and then touched to a new agar plate. The new replica plate is having same colony at the same position as in the master plate.
3. Nitro-cellulose paper is pressed on the top of the master plate so that the paper is the replica of the master plate. Some cells from each colony are transferred to the paper with same position as that in master plate.
4. The cells on the paper are now lysed with alkali treatment and also their DNA is denatured.
5. A radioactive complementary probe is used to hybridize with the denatured DNA on the paper.
6. The unhybridized probe is removed by washing.
7. The paper is exposed to the X-ray film and the hybrids are detected by autoradiography.

Introduce the DNA fragment in form of a recombinant DNA in the E. coli cells

↓

Grow the cells on a master plate.

↓

Make a replica on a nitrocellulose paper.

↓

Lyse the cells and denature the DNA on the paper. Add the complementary radio-labeled probe on the membrane and incubate for hybridization

↓

Detect the hybrids by autoradiography.

POLYMERASE CHAIN REACTION (PCR)

- It is an extremely powerful technique that allows a million fold amplification of selected DNA sequence.
- It is used to clone the given DNA sequence *in vitro* without using living cells during cloning process.
- It makes the use of the ability of the enzyme DNA polymerase to carry out the semi- conservative replication of DNA.
- PCR results in the amplification of a chosen region of DNA molecule when the sequences of the borders are known.
- Two synthetic oligonucleotides that are complementary to end parts of the chosen sequence can anneal to the ends of the DNA segment.
- The amplification of this segment between the two defined ends can be then continued by synthesizing or replicating the DNA with the help of a special enzyme DNA polymerase.
- The oligonucleotides act as the primers for the enzyme to complete the synthesis/replication.
- This amplification is achieved by a repetitive series of cycles involving three steps.
  1. Denaturation: The DNA sample to be amplified (template DNA) is denatured by heating at 92^{o}C.
  2. Annealing: The oligonucleotide primers added to the separated template DNA strands and the temperature is reduced to 40-60 ^{o}C.
  3. Synthesis by extending the primers: At 72^{o}C DNA polymerase that has been added in the reaction mixture extends the 3' ends of the oligonucleotide primers complementarily using the template DNA. The DNA polymerase used in PCR is a thermostable, isolated from the bacterium *Thermus aquaticus*. It is called *Taq* polymerase which survives at the high temperature required during denaturation step.

These three steps represent a single PCR cycle. The products of the first cycle are replicated for further amplification. This reaction can be performed many times to supply unlimited copies of the amplified DNA. During repetition of the cycle regular denaturation of the freshly synthesized double stranded DNA molecules is carried out.

Applications of PCR

- After 25 cycles the target DNA is amplified about 10^6 fold.
- PCR can detect and amplify as little as 1 DNA molecule. It allows successful cloning of DNA even from samples which are more than 40,000 years old or mummified human bodies, extinct animals like wooly mammoth or dinosaurs.
- Highly useful in new fields like molecular archaeology, molecular palaeontology
- To trace the evolution of pathogenic viruses, forensic medicine, detection of viral infections before causing the disease or showing the symptoms, prenatal diagnosis of genetic disorders

☐ It was used for Human Genome Project.

DNA Microarray

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

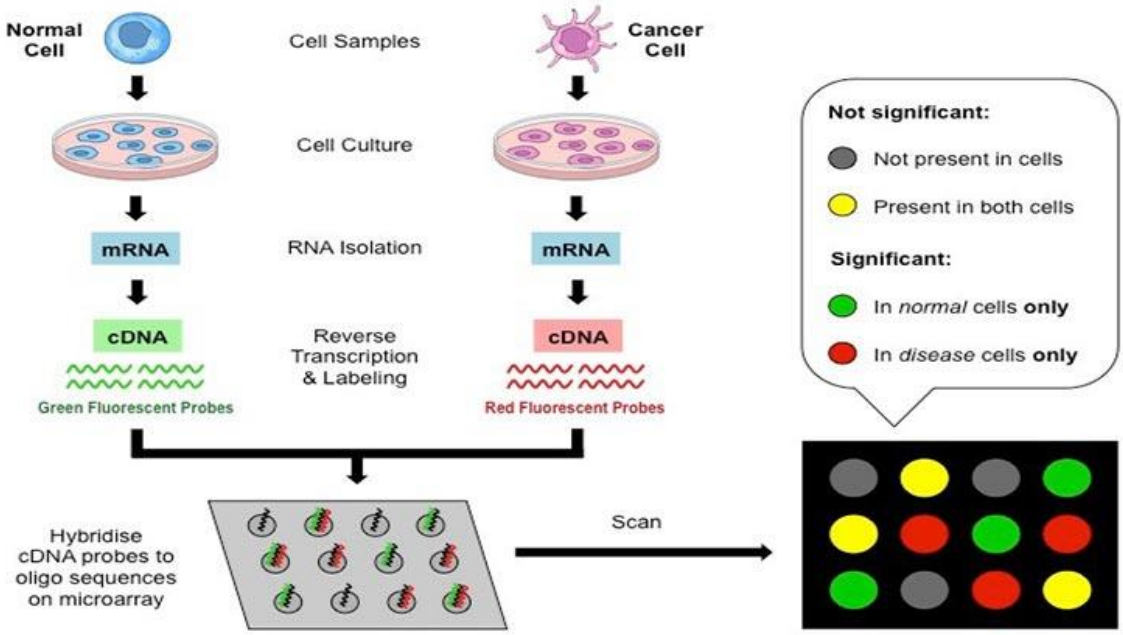


Image Source: BioNinja

Principle of DNA Microarray Technique

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

Steps Involved in cDNA based Microarray

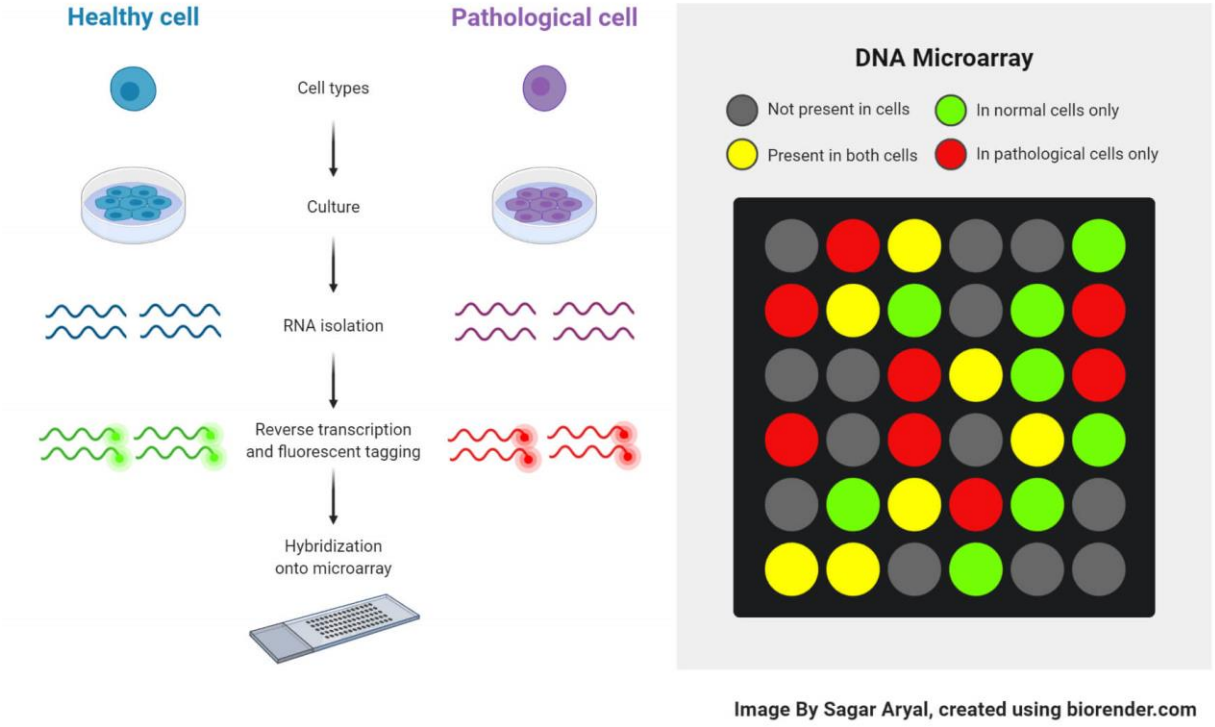


Image By Sagar Aryal, created using biorender.com

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

1. **Collection of samples**
- The sample may be a cell/tissue of the organism that we wish to conduct the study on.
- Two types of samples are collected: healthy cells and infected cells, for comparison and to obtain the results.

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

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

4. **Hybridization**

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

5. **Collection and analysis**

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

Applications of DNA Microarray

- In humans, they can be used to determine how particular diseases affect the pattern of gene expression (the expression profile) in various tissues, or the identity (from the expression profile) of the infecting organism. Thus, in clinical medicine alone, DNA microarrays have huge potential for diagnosis.

Besides, it has applications in many fields such as:

- Discovery of drugs
- Diagnostics and genetic engineering
- Alternative splicing detection
- Proteomics
- Functional genomics
- DNA sequencing
- Gene expression profiling
- Toxicological research (Toxicogenomics)

Advantages of DNA Microarray

- Provides data for thousands of genes in real time.
- Single experiment generates many results easily.
- Fast and easy to obtain results.
- Promising for discovering cures to diseases and cancer.
- Different parts of DNA can be used to study gene expression.

Disadvantages of DNA Microarray

- Expensive to create.
- The production of too many results at a time requires long time for analysis, which is quite complex in nature.
- The DNA chips do not have very long shelf life.