Uttar Pradesh Rajarshi Tandon
Open University

PGBCH - 118 N
Genetic Engineering

Block- 1

Material and Tools- I

Course Design Committee

Dr. (Prof.) Ashutosh Gupta, School of Science, UPRTOU, Prayagraj	Chairman
Prof. Prof. Umesh Nath Tripathi Department of chemistry Deen Dayal Upadhyaya Gorakhpur University, Gorakhpur	Member
Prof. S.I. Rizvi Department of Biochemistry University of Allahabad, Prayagraj	Member
Prof. Dinesh Yadav Department of Biotechnology Deen Dayal Upadhyaya Gorakhpur University, Gorakhpur	Member
Prof. Sharad Kumar Mishra Department of Biotechnology Deen Dayal Upadhyaya Gorakhpur University, Gorakhpur	Member
Dr. Ravindra Pratap Singh Assistant Professor (Biochemistry) School of Science, UPRTOU, Prayagraj	Member
Dr. Dharmveer Singh Assistant Professor (Biochemistry) School of Science, UPRTOU, Prayagraj	Course Coordinator

Course Preparation Committee

Dr. Gopal Dixit Assistant Professor Department of Botany, Upadhi Mahavidyalaya, Pilibhit, U.P.	**Author**	**Block-1-2**	Unit: 1-5
Dr. Arun Kumar Pandey Assistant Professor Department of Botany, PSMPG College, Maharajganj. U.P.	**Author**	**Block-2-3**	Unit: 6-9
Dr. Sadhana Singh Assistant Professor- Biochemistry School of Sciences, UPRTOU, Prayagraj	**Author**	**Block-4**	Unit: 10, 12
Dr. Anuradha Singh Assistant Professor- Botany School of Sciences, UPRTOU, Prayagraj	**Author**	**Block-4**	Unit: 11
Dr. Mohd. Khalid Masroor Retd. Associate Professor-Botany, University of Allahabad, U.P.	**Editor**	(Block- 01, 02, 03&04, Unit: 1, 2,3,4,5, 10, 11, &12)	
Dr. Rajiv Ranjan Associate Professor, MLKPG College, Balrampur, U.P.	**Editor**	(Block- 02& 03)	Unit: (6-9)
Dr. Dharmveer Singh Assistant Professor (Biochemistry) School of Sciences, UPRTOU, Prayagraj			**(SLM & Course Coordinator)**

PGBCH – 118, Genetic Engineering
©UPRTOU , 2024
ISBN :

Printed By: K.C.Printing & Allied Works, Panchwati, Mathura -281003.

COURSE INTRODUCTION

In this course, you will learn about genetic engineering. Where genetic engineering fits into molecular biology is examined. Gene expression in prokaryotes and eukaryotes, DNA and RNA isolation, the principle of nucleic acid hybridization, PCR, DNA sequencing methods, enzymes in nucleic acids and DNA sequencing, and immune technology are all briefly discussed. However, genetic engineering is a sort of biotechnology that involves directly changing an organism's DNA to alter its characteristics. Cloning from genomic DNA entails separating and inserting DNA pieces into vectors for analysis. PCR amplifies specific DNA sequences, allowing for comprehensive examination and cloning. DNA fingerprinting identifies people based on their distinct genetic characteristics. Screening and assessing recombinants entails choosing and validating effective genetic changes. Immuno technology, such as ELISA and Western blotting, detects specific proteins or antigens. Transgenic animals, which are made by introducing foreign genes into their genomes, are utilized in science, agriculture, and medicine as models for researching gene function and creating therapeutic proteins will cover in this course. Genetic engineering offers huge prospects for furthering science and improving people's lives, but it must be approached with caution due to ethical, social, and environmental considerations. The course is organized in the following blocks:

Block 1 *covers the Material and Tools- I*

Block 2 *deals the Genetic Engineering-I*

Block 3 *describes the Genetic Engineering-II*

Block 4 *this block covers the Genetic Engineering-III*

Introduction

The following three units are included in the first block of genetic engineering are as:

Unit-1: This unit discusses basic molecular biology and its history, as well as gene expression in prokaryotes and eukaryotes, which entails turning DNA into functional proteins. Gene expression in prokaryotes, such as bacteria, is uncomplicated. Gene expression in eukaryotic cells is regulated by chromatin structure, transcription factors, and post-translational modifications, which provide more complex control over protein synthesis.

Unit-2: This unit covers the isolation of nucleic acids. The isolation of DNA and RNA, radiolabelling of nucleic acids, and end labeling are explained using nick translation and primer extension; nonetheless, the principle of nucleic acid hybridization and DNA sequencing methods such as Maxam-Gilbert and Sanger-Coulson sequencing are presented.

Unit-3: This unit covers enzymes that interact with nucleic acids. Restriction enzymes (Type II) and restriction mapping are explored. DNA modifying enzymes such as nucleases, polymerases, end-modifying enzymes, and DNA ligases are briefly described. DNA ligases join DNA fragments, essential for creating recombinant DNA. Together, they enable precise genetic manipulation and analysis.

Unit-1: History of genetics

Contents

1.1. Introduction

The history of genetics form classical era with contribution by Pythagorus, Hippocrates, Aristotle, Epicurus and others. In ancient green, Hippocrates suggested that all organs of body of a present gave off invisible "seed" miniaturised components were transmitted during sexual intercourse and combine with mother's womb to form a baby. The modern genetics began with the work Modern genetics began with the work of Augustinian Friar Gregor Johann Mendel. His works on pea plants published in 1866, provided initial evidence that, on its rediscovery in 1900,s helped to established the theory of Mendelian inheritance. In 1900 rediscovery of Vries, Carlcorrens and Erich Von. Tschermak lead to rapid advances in genetics. By 1995 the basic principles of mendelian genetics had been studied in wide variety of organisms most notably the fruit fly Drosophila melanogaster.

Thomas Hunt Morgan's research with fruit flies in the early 1900s established chromosomes as gene carriers, so confirming the chromosomal hypothesis of heredity. The discovery of the DNA double helix structure by James Watson and Francis Crick in 1953, based on Rosalind Franklin's critical X-ray diffraction photographs, transformed genetics. This finding laid the groundwork for molecular genetics, emphasizing DNA's function in inheritance and protein synthesis. After the rediscovery of mendalian work there was a feud between Willium Bateson and Pearson over the hereditary mechanism, solved by Ronold fisher in his work "The correlation between Relative on the

supposition of mendalain inheritance" .In 1910 the Thomos Hunt Morgan showed that genes reside on specific chromosomes. He later showed that genes occupy specific locations on the chromosomes.

The second half of the twentieth century saw considerable improvements, notably the invention of recombinant DNA technology in the 1970s, which enabled genetic engineering. The Human Genome Project, finished in 2003, mapped the whole human genome, revealing information about hereditary illnesses and human evolution. Genetics is now an essential component of professions such as medicine, biotechnology, and forensic science, helping to advance our understanding of life and heredity.

Geneticists investigate hereditary processes, including the inheritance of traits, characteristics, and diseases. Genetics focuses on the molecular instructions that transmit information from generation to generation. Geneticists have shown that certain genetic variations are linked to disease and that varying genes enhances a species' ability to adapt to environmental changes. Although significant advances in genetics research, such as deciphering the genetic code, isolating disease-causing genes, and successfully cloning plants and animals, has occurred since the mid-twentieth century, the field's history spans approximately 150 years. Genomics studies genes, genetics, inheritance, molecular biology, biochemistry, and biological statistics while also incorporating advanced technology, computer science, and mathematics.

Genetics began with the emergence of evolutionary theories. The origin of species and how species variety evolved as a result of Charles Darwin and Wallace's studies in 1858. They described how new species emerged through evolution and how natural selection worked to create new forms. However, they were unaware of the function that genes played in these phenomena.

Within the time, Mendelian Principles and the Chromosomal Theory of Inheritance were developed. Mendel's work remained mostly unknown. It wasn't until 1900 that Mendelian ideas were rediscovered, and periodicals began to cite his work.

In the 1870s, the substance in the nucleus was identified as a nucleic acid. DNA was discovered to be the genetic material between the 1920s and the mid-1950s. Griffith's research with a bacterial strain proved the theory.

In 1928 Fredrick Griffith showed that genes could be transferred in what is now known as Griffith's experiments, injection into a mouse of a deadly strain of bacteria that had been heat killed

transferred genetics information to a safe strain of the some bacteria, killing the mouse. In 1941, George Wells Beadle and Edward Lawrie Tatum showed that mutation in genetics caused in specific steps of metabolic pathways. This showed gene code for specific proteins leading to the "one gene on enzyme" hypothesis. Oswald Avery, Colin Munro Macleod and Maclyn McCarty showed in 1944 that DNA holds gene information.

In 1952, Rosalind Franklin and Raymond Gosling produced a striking clear X-ray diffraction pattern indicating a helical form.

In 1960, Jacob and Collaborators discovered the operon which consists of sequences of gene whose expression is coordinated by operator DNA. Richard J. Robert and Philip sharp discovered in 1977 that genes can be split into segments. The word genetics was introduced in 1905 by English biologist William Bateson, who was one of the discoverers of Mendel's work and who become a champion of Mendel's principle of inheritance.

Avery, MacLeod, and McCarty also demonstrated that DNA, rather than protein or RNA, was responsible for the genetic inheritance and evolution of the bacterial strains investigated by Griffith. In their seminal work, Watson and Crick identified the structure of DNA, and others proposed that DNA carried a genetic code. The code was first discovered in the 1960s. Crick discovered transcription and translation, which resulted in the establishment of the "central dogma of molecular biology".

In 1947, Salvador Iuria discovered the reactivation of irradiated phage leading to many further studies on the fundamental process of repair of DNA damage. In 1958, Meselson and Stahl demonstrated that DNA replicates semi conservatively leading the understanding that each of the individual strands in double stranded DNA serves as a template for new stranded synthesis. Mid-late twentieth century and early twenty-first century. We saw the emergence of molecular biology and molecular genetics. Various advanced technologies entered the knowledge base around this time. This encompassed molecular biology, recombinant DNA technologies, and biotechnology methods. During this time, technologies for radiolabelling DNA with radioactive or fluorescent tags were found in order to provide diagnostic and therapeutic approaches as well as research tools. Restriction enzymes were identified and utilized to create recombinant DNA molecules containing foreign DNA that could be readily produced in bacterial strains. Then came proced ures like PCR (Polymerase chain reaction) and a slew of other biotechnology techniques, and new uses emerged in medicine, pharmacotherapy, and research.

Objectives

After reading this units the learner will be able to

- To know basic molecular biology
- To know the gene expression and its regulations
- To know gene expression in prokaryotes and in eukaryotes

1.2. Basic molecular biology

The term "molecular biology" was first used in 1945 by English Physicist William Astbury, who discover it as an approach focussed on discerning the under pinings of biological phenomenon i.e. undiscovering the physical and chemical structures and properties of biological molecules as well as there interaction with other individual. These interactions explain observation of so called classical biology, which instead studies biological processes act large scale and higher levels of organization. However, molecular biology is the study of the molecular systems that operate within cells. It focuses on how DNA, RNA, and proteins work together to govern biological processes. DNA (deoxyribonucleic acid) stores genetic information that is translated into RNA (ribonucleic acid). RNA is then translated into proteins, which are the cell's workhorses, performing duties such as catalyzing reactions (enzymes), providing structural support, and controlling gene expression. Replication (copying DNA), transcription (creating RNA from DNA), and translation (converting RNA into proteins) are all important activities. Techniques such as PCR (polymerase chain reaction) and sequencing are critical for investigating these molecular interactions.

Warren Weaver, an American scientist, invented the term "Molecular Biology" in the year 1938. According to records, the discovery of molecular biology began in the early 1940s, and its basic development occurred in 1953, when two molecular scientists named James Watson and Francis Crick invented the double-helical structure of the DNA molecules. In 1950, Francis Crick, James Watson were the first to describe the double helix model for the chemical structure of DNA. This is often considered a laud mark event for the nascent field because it provided a physic-chemical basis by which to understand the idea of nucleic acids as the primary substances of biological inheritance.

Molecular biology is the study of biology at a molecular level. The field combines biology and chemistry, particularly genetics and biochemistry. Molecular biology is primarily concerned with understanding the relationships between a cell's many systems, particularly the interactions between

DNA (deoxyribonucleic acid), RNA (ribonucleic acid), and protein production, as well as how these interactions are regulated.

Molecular biology entails the study of the chemical organization of the cell. Molecules analysis organization comprises the smallest component capable of performing all the activities (structural and catalytic) of a substance. One or more atoms constitute each molecule.

Many modules comprise the various and cellular and sub-cellular components of an organism. Molecules form hat only the physical structure of the organism but communicate information between the various compartments of the cell. This communication can be the transfer of information from DNA to RNA and finally to protein or the subtle regulation of cell internal homeostatic processes. This communication relies on the interaction of various molecules to insure the fidelity of the message or cellular regulation. Molecular biology is a branch of biology that is also closely related to other sub-disciplines like biochemistry, cell biology, genetics and genomics. Molecular biology is the branch of biology that studies the molecular bases of biological activity. Application of molecular Biology:

1. Restriction Enzymes
2. DNA hybridization
3. Rapid DNA sequencing
4. Genetic Engineering and Expression cloning
5. The Polymerase Chain Reaction (PCR)

i. DNA (Deoxyribonucleic Acid)

DNA (Deoxyribonucleic Acid) is a molecule that contains the genetic information needed for the growth, development, function, and reproduction of all living things, including many viruses. DNA is a long polymer made up of repeating units called nucleotides, each of which contains a phosphate group, a sugar molecule (deoxyribose), and a nitrogenous base. There are four types of nitrogenous bases in DNA.

- Adenine (A)
- Thymine (T)
- Cytosine (C)
- Guanine (G)

The sequence of these nucleotides contains genetic information. In the double helix structure, bases pair precisely (A with T and C with G) via hydrogen bonding to ensure proper replication.

Structure of DNA

1. **Double Helix:**

DNA has a double helix structure that resembles a twisted ladder. The sides of the ladder are composed of alternating sugar and phosphate groups, while the rungs are made up of pairs of nitrogenous bases. James Watson and Francis Crick discovered the double helix structure in 1953, with input from Rosalind Franklin and Maurice Wilkins.

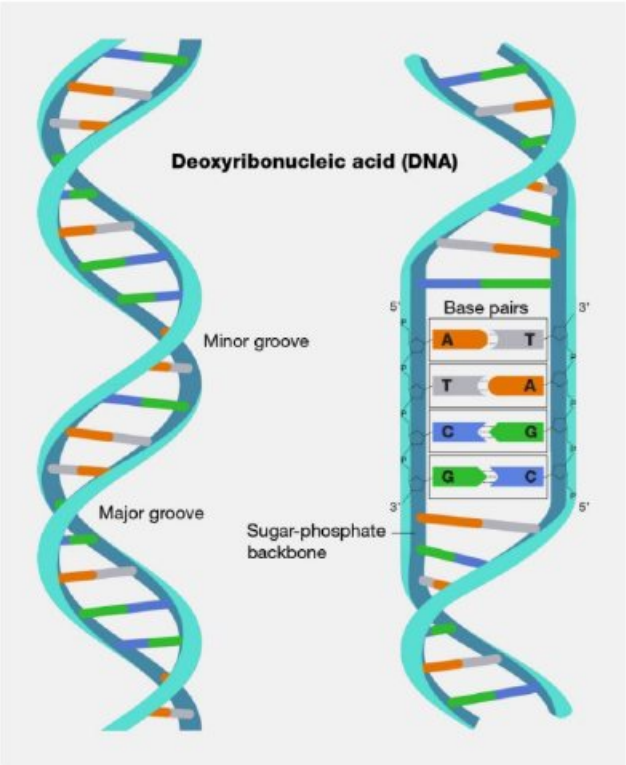


Fig.1.1: Structure of double strand DNA

2. **Nucleotides**: the nucleotides in DNA molecules has following components such as

- **Phosphate Group**: Connects the sugar molecules of adjacent nucleotides, forming the backbone of the DNA strand.

- **Deoxyribose Sugar**: A five-carbon sugar molecule that forms part of the backbone.
- **Nitrogenous Bases**: Four types of bases that pair specifically with each other: Adenine (A), Thymine (T), Cytosine (C), and Guanine (G).

3. **Base Pairing**:

   - **Complementary Base Pairing**:
     - **Adenine (A)** pairs with **Thymine (T)**.
     - **Cytosine (C)** pairs with **Guanine (G)**.
   - **Hydrogen Bonds**: Base pairs are held together by hydrogen bonds (A-T pairs have two hydrogen bonds; C-G pairs have three).

4. **Antiparallel Strands**:

   - **Description**: The two strands of DNA run in opposite directions. One strand runs 5' to 3', and the other runs 3' to 5'.
   - **Significance**: The antiparallel orientation is crucial for DNA replication and function.

Functions of DNA:

1. **Genetic Information Storage**:

DNA segments containing instructions for protein synthesis. Each gene contains a unique sequence of nucleotides that defines the amino acid sequence of a protein. The sequence of nucleotides in DNA is read in triplets (codons), with each specifying a certain amino acid.

2. **Replication**:

DNA replicates during cell division, guaranteeing that each new cell obtains an exact duplicate of the genetic material. The key enzymes involved are DNA polymerase, which adds new nucleotides, and helicase, which unwinds the DNA.

3. **Transcription**:

DNA is transcribed into messenger RNA (mRNA), which transports genetic information from the DNA to the ribosomes, where proteins are produced. RNA polymerase is the enzyme that synthesizes RNA from the DNA template.

4. **Translation**:

At the ribosome, mRNA is translated into protein by tRNA molecules, which add amino acids to the growing polypeptide chain in accordance with the mRNA's codon sequence.

5. **Mutation and Variation**:

Changes in the DNA sequence can cause genetic variation and potentially alter an organism's features. Mutations can result from environmental influences or faults in DNA replication. Variations in DNA sequences provide variety within populations and can drive evolution.

DNA in Cells:

1. **Chromosomes**:

DNA is divided into structures known as chromosomes. Humans have 23 pairs of chromosomes in each cell, which hold all of their genetic information. DNA is bundled with proteins (histones) to form a complex termed chromatin, which aids in organizing and regulating access to the DNA.

2. **Nucleus**:

In eukaryotes (plants, mammals, fungi, and protists), DNA is stored in the nucleus, a membrane-bound organelle. In prokaryotes (bacteria and archaea), DNA is found in the nucleoid area, which is not membrane bound.

DNA Technologies:

1. **Sequencing**: Determining the precise order of nucleotides in a DNA molecule, used in genetic research, diagnostics, and forensic science.

2. **Cloning**: Creating genetically identical copies of DNA fragments, cells, or organisms.

3. **Genetic Engineering**: Modifying DNA to alter an organism's traits or produce new products.

4. **PCR (Polymerase Chain Reaction)**: Amplifying specific DNA sequences to analyze or manipulate them.

ii. RNA (Ribonucleic Acid)

RNA is a single-stranded molecule involved in various roles, including coding, decoding, regulation, and expression of genes. RNA is a single-stranded molecule composed of nucleotides. Each nucleotide consists of a ribose sugar, a phosphate group, and a nitrogenous base (adenine [A], cytosine [C], guanine [G], or uracil [U]).There are several types of RNA, each with a specific function:

Types of RNA:

- **mRNA (messenger RNA)**: Messenger RNA (mRNA) is a single-stranded RNA molecule that transports genetic information from the nucleus to the cytoplasm's ribosomes, which build proteins. The process begins with transcription, which creates an mRNA strand from a DNA template. This mRNA strand contains codons, which are three-nucleotide sequences that correspond to specific amino acids or signal when protein synthesis begins or ends. Once transcribed, mRNA is processed in eukaryotes, which includes the addition of a 5' cap and a poly-A tail, as well as splicing to eliminate introns. The mature mRNA then departs the nucleus and moves into the cytoplasm, where it connects with ribosomes.Ribosomes read mRNA codons and assemble the correct amino acids to form a polypeptide chain, which eventually folds into a functioning protein.The significance of mRNA in gene expression is crucial because it acts as an intermediary between the genetic information in DNA and the creation of proteins, which perform a variety of tasks within cells. Recent breakthroughs in biotechnology have used mRNA for medicinal uses, such as mRNA vaccines, which instruct cells to create antigens that elicit an immune response, as demonstrated by the COVID-19 vaccines.
- **tRNA (transfer RNA)**: Transfer RNA (tRNA) is a tiny RNA molecule that converts the genetic code delivered by messenger RNA (mRNA) into proteins. Each tRNA molecule has a distinct shape that enables it to act as an adapter, matching a specific amino acid to the relevant

codon on the mRNA strand during protein synthesis.A sequence of three nucleotides that complement an mRNA codon. Located at the 3' end of the tRNA, a specific amino acid is covalently bonded by an enzyme known as aminoacyl-tRNA synthetase. During translation, tRNA molecules deliver amino acids to the ribosome, which is the cellular machinery that produces proteins.

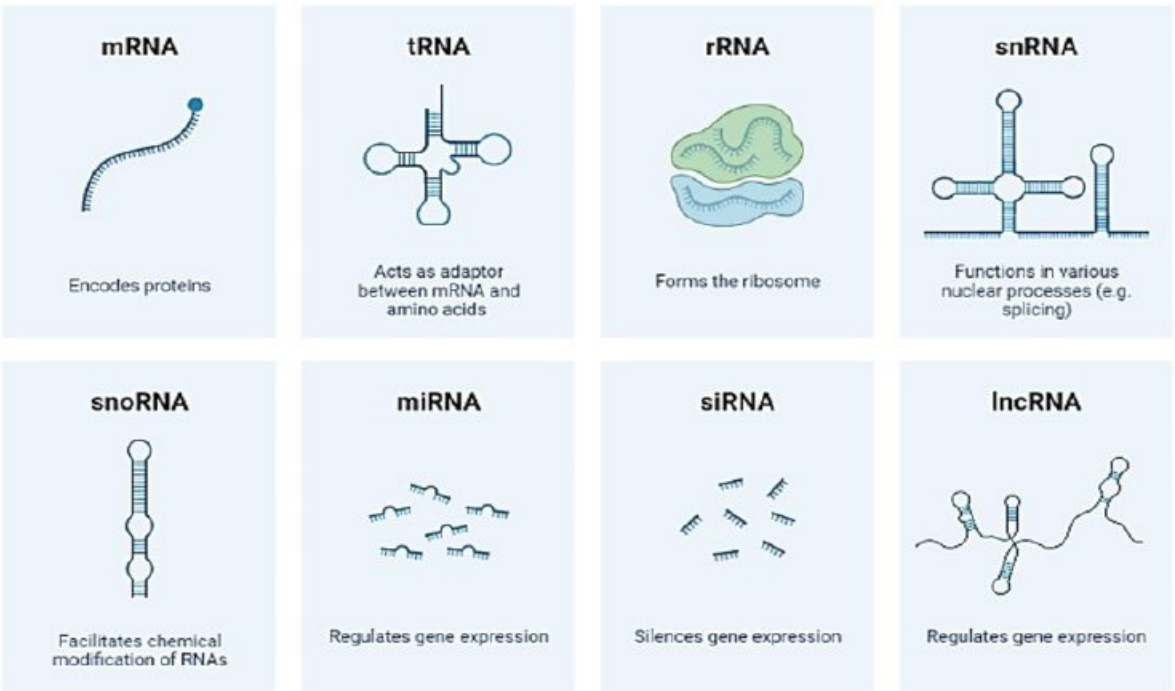


Fig.1.2: Different types of RNA

- **rRNA (ribosomal RNA)**: Ribosomal RNA (rRNA) is a type of RNA that, along with proteins, makes up the structural and functional core of ribosomes, the cellular organelles responsible for protein synthesis. rRNA is the most abundant kind of RNA in cells, and it is essential for translating mRNA into proteins.Ribosomes are made up of two subunits, each containing rRNA and ribosomal proteins.In prokaryotes, the ribosome is made up of two subunits: 30S small and 50S big, resulting in a 70S ribosome.In eukaryotes, the ribosome is composed of a 40S small subunit and a 60S big subunit, resulting in an 80S ribosome.The "S" represents Svedberg units, a measure of sedimentation rate during centrifugation that shows size and form. The ribosome's structure is built around rRNA, which serves as a scaffold for ribosomal proteins.

- **snRNA (small nuclear RNA)**: tiny nuclear RNA (snRNA) is a type of tiny RNA found in the nucleus of eukaryotic cells. These RNAs are typically 150-300 nucleotides long and play an

important role in RNA processing, particularly in the splicing of pre-messenger RNA. SnRNAs are typically found in complex with proteins, creating tiny nuclear ribonucleoproteins (snRNPs, pronounced "snurps"). The most well-known snRNAs are U1, U2, U4, U5, and U6, each with a unique sequence and function. SnRNAs have a fundamental role in pre-mRNA splicing, which involves the removal of non-coding portions (introns) and the joining of coding regions (exons) to form mature mRNA. SnRNAs, as part of snRNPs, are the fundamental components of the spliceosome, a vast ribonucleoprotein complex that performs splicing.Certain snRNAs have a direct role in the catalytic steps of splicing, helping to align the pre-mRNA and catalyzing the chemical events that eliminate introns and ligate exons.

- **miRNA (microRNA) and siRNA (small interfering RNA):** MicroRNA (miRNA) and small interfering RNA (siRNA) are two types of small, non-coding RNA molecules that regulate gene expression using RNA interference (RNAi) processes. miRNAs are generally 20-24 nucleotides long. They are produced from endogenous genes as primary miRNAs (pri-miRNAs), which are then processed in the nucleus by the enzyme Drosha to become precursor miRNAs. Pre-miRNAs are transported to the cytoplasm, where Dicer, an RNase III enzyme, converts them into mature miRNA duplexes. The guide strand of the duplex is integrated into the RNA-induced silencing complex (RISC), whereas the other strand is often destroyed. miRNAs regulate gene expression post-transcriptionally by base-pairing with complementary sequences in the 3' untranslated regions (3' UTRs) of target mRNAs. This interaction typically leads to mRNA degradation or translational repression, thereby reducing protein synthesis from the target mRNA. miRNAs play roles in various cellular processes, including development, differentiation, proliferation, and apoptosis.

- **siRNA (small interfering RNA):** siRNAs are approximately 20-25 nucleotides long.They are frequently derived from foreign sources, such as viruses, or are deliberately injected into cells.Dicer processes lengthy double-stranded RNA (dsRNA) precursors into siRNA duplexes.Like miRNA, one strand of the siRNA duplex (the guide strand) is loaded into the RISC.siRNAs mediate mRNA degradation by forming perfect or near-perfect base pairs with their target mRNAs.As a result, the target mRNA is cleaved and then degraded, thereby suppressing its expression. siRNAs are commonly employed in research and therapeutic applications to reduce the expression of certain genes.

iii. Proteins

Proteins are vital macromolecules that serve a wide range of roles in living organisms. They are made up of amino acids joined together by peptide bonds, and their function is determined by their unique sequence and three-dimensional structure. Proteins are involved in nearly every biological process and are essential for tissue and organ structure, function, and regulation.

Structure of Proteins:

1. **Amino Acids**:

- **Building Blocks**: Proteins are made up of 20 different amino acids. Each amino acid has a basic structure consisting of an amino group ($–NH_2$), a carboxyl group (–COOH), a hydrogen atom, and a variable side chain (R group).

- **Classification**: Amino acids are classified based on their side chains as nonpolar, polar, acidic, or basic.

2. **Levels of Protein Structure**:

- **Primary Structure**: The linear sequence of amino acids in the polypeptide chain. This sequence is determined by the gene encoding the protein.

- **Secondary Structure**: The local folding of the polypeptide chain into structures such as alpha-helices and beta-sheets, stabilized by hydrogen bonds.

- **Tertiary Structure**: The overall three-dimensional shape of a single polypeptide chain, formed by interactions among the secondary structural elements and side chains. This structure is stabilized by various interactions, including hydrogen bonds, disulfide bridges, hydrophobic interactions, and ionic bonds.

- **Quaternary Structure**: The assembly of multiple polypeptide chains (subunits) into a functional protein complex. Not all proteins have quaternary structure.

Functions of Proteins:

Enzymes are proteins that catalyze biological reactions, increasing their rate without being consumed. Amylase breaks down starches into sugars, while DNA polymerase creates new DNA

strands. Structural proteins help cells and tissues maintain their shape and structure. Collagen (in connective tissues), keratin (hair and nails), and actin and tubulin (cytoskeleton).Transport proteins help chemicals move across cell membranes or throughout the body. Hemoglobin (which transports oxygen in the blood) and membrane transporters (which let molecules move across cellular membranes).Regulatory proteins govern a variety of biological activities, including gene expression and cell signaling. Transcription factors govern gene expression, while hormones such as insulin regulate blood sugar levels. Defence proteins defend the body against infections and other dangerous chemicals. Antibodies bind and neutralize foreign substances.

Protein Synthesis:

1. **Transcription**:

- **Process**: The DNA sequence of a gene is transcribed into messenger RNA (mRNA) in the nucleus.

2. **Translation**:

- **Process**: mRNA is translated into a polypeptide chain at the ribosome. Transfer RNA (tRNA) molecules bring amino acids to the ribosome, where the mRNA sequence is read in sets of three nucleotides (codons) to determine the sequence of amino acids.

3. **Post-Translational Modifications**:

- **Process**: After synthesis, proteins may undergo various modifications, such as phosphorylation, glycosylation, or cleavage, which are essential for their final functional state and activity.

iv. Central Dogma of Molecular Biology

The Central Dogma of Molecular Biology is a framework for understanding the flow of genetic information within a biological system. Formulated by Francis Crick in 1958, it describes the process by which genetic information is transferred from DNA to RNA to protein. The central dogma outlines three key processes: replication, transcription, and translation.

1. **Replication**:

- **Description**: The process by which DNA makes a copy of itself.

- **Function**: Ensures that genetic information is accurately passed on to daughter cells during cell division.

- **Mechanism**: DNA polymerase enzymes synthesize a new complementary strand for each original DNA strand, resulting in two identical DNA molecules.

2. **Transcription**:

- **Description**: The process by which a segment of DNA is copied into RNA by the enzyme RNA polymerase.

- **Function**: Converts genetic information stored in DNA into a form (mRNA) that can be used to produce proteins.

- **Mechanism**: RNA polymerase binds to the promoter region of a gene, unwinds the DNA, and synthesizes a complementary RNA strand using one of the DNA strands as a template. The resulting RNA molecule is called messenger RNA (mRNA).

3. **Translation**:

- **Description**: The process by which the genetic code carried by mRNA is decoded to produce a specific protein.

- **Function**: Synthesizes proteins, which perform a vast array of functions within the organism.

- **Mechanism**: Ribosomes, composed of rRNA and proteins, facilitate the assembly of amino acids into polypeptide chains. Transfer RNA (tRNA) molecules bring amino acids to the ribosome, where they are added to the growing polypeptide chain according to the sequence of codons in the mRNA.

Flow of Information:

- **DNA → RNA → Protein**:

- **DNA to RNA** (Transcription): Genetic information in DNA is transcribed into mRNA.

- **RNA to Protein** (Translation): The information in mRNA is translated into a specific sequence of amino acids to form a protein.

Significance:

- The central dogma is fundamental to understanding molecular biology and genetics.

- It explains how genetic information is maintained and expressed within cells.

- It provides the basis for many biotechnological applications, such as genetic engineering, gene therapy, and the development of mRNA vaccines.

v. Replication

DNA replication is the process of copying a single DNA molecule to create two identical DNA molecules. This procedure is necessary for cell division. The steps in DNA replication are:

Initiation:

Replication begins at certain DNA molecule regions known as replication origins. Prokaryotes normally have one origin, whereas eukaryotes have several origins. Helicase enzymes unwind the double helix by disrupting the hydrogen bonds connecting the complementary bases, resulting in two single-stranded DNA templates. The unwinding produces Y-shaped structures known as replication forks at each replication origin.

Primer Synthesis:

An RNA polymerase enzyme called primase synthesizes a short RNA primer complementary to the DNA template. This primer provides a starting point for DNA synthesis because DNA polymerases can only add nucleotides to an existing strand of nucleic acid.

Elongation:

DNA polymerase III (prokaryotes) or DNA polymerase δ and ε (eukaryotes) expands the RNA primer by adding DNA nucleotides that are complementary to the template strand. The leading strand is continually produced in the 5' to 3' direction, towards the replication fork. The lagging strand is generated discontinuously in tiny segments known as Okazaki fragments, away from the replication fork. These parts are eventually put together. DNA polymerases proofread any mismatched nucleotides, ensuring high fidelity in DNA replication.

Primer Removal and Replacement:

Exonuclease activity degrades RNA primers. In prokaryotes, DNA polymerase I eliminates RNA primers and replaces the gaps with DNA. In eukaryotes, RNase H eliminates RNA primers while DNA polymerase δ fills in the gaps.

Ligation:

The enzyme DNA ligase seals the nicks between the Okazaki fragments on the lagging strand, resulting in a single DNA strand.

Termination:

Replication continues until the entire DNA molecule is copied. In circular prokaryotic DNA, this entails separating the two connected DNA molecules. Replication in linear eukaryotic DNA terminates when replication forks converge and telomeres replicate.

vi. Transcription

Transcription is the process of copying genetic information from DNA into messenger RNA (mRNA), which is then used to synthesize proteins. This is the initial step of gene expression, where the information encoded in DNA is used by the cell to generate proteins. A full description of the transcription process:

1. **Initiation:**

Transcription begins when RNA polymerase binds to a specific section of DNA known as the promoter, which is located upstream of the gene being transcribed. In eukaryotes, transcription factors

help RNA polymerase recognize and bind to the promoter. Once bound, RNA polymerase unwinds a little portion of the DNA double helix, revealing the template strand for RNA synthesis.

2. **Elongation**:

RNA polymerase works along the DNA template strand, forming a complementary RNA strand by adding ribonucleotides (ATP, UTP, GTP, and CTP) in the 5' to 3' direction. Each ribonucleotide is inserted in accordance with the base-pairing regulations (adenine with uracil, cytosine with guanine). As it proceeds along the DNA, RNA polymerase creates a transcription bubble in which the DNA is unraveled, an RNA-DNA hybrid is produced, and newly synthesized RNA is displaced.

3. **Termination**:

In prokaryotes, termination occurs when RNA polymerase encounters a terminator sequence, prompting the enzyme to disengage from the DNA and release the freshly created RNA transcript. Termination in eukaryotes includes more complex mechanisms and frequently requires additional components. For mRNA, transcription extends beyond the coding sequence, and the pre-mRNA is cleaved at a precise location downstream of the coding section.

RNA Processing (Eukaryotes):

- **5' Capping**: A modified guanine nucleotide (7-methylguanosine cap) is added to the 5' end of the pre-mRNA, protecting it from degradation and aiding in ribosome binding during translation.
- **Polyadenylation**: A poly-A tail (a series of adenine nucleotides) is added to the 3' end of the pre-mRNA, enhancing stability and facilitating export from the nucleus.
- **Splicing**: Non-coding regions called introns are removed from the pre-mRNA, and the coding regions called exons are joined together by the spliceosome, forming mature mRNA. This mature mRNA is then exported from the nucleus to the cytoplasm for translation.

vii. Genetic Code

The genetic code is a set of rules that defines how the information encoded in DNA is translated into proteins. It consists of sequences of nucleotides in DNA that are decoded into amino acid sequences during protein synthesis. The genetic code is universal, meaning it is nearly identical in all living organisms, which underscores its fundamental role in biology. The Features of the Genetic Code are

Codons:

The genetic code is read in groups of three nucleotides known as codons. Each codon represents a specific amino acid or a stop signal. For example, the codon "AUG" defines the amino acid methionine while also serving as the start codon for translation.

Degeneracy:

The genetic code is degenerate, which means that most amino acids are encoded by several codons. This redundancy protects against mutations. For example, the amino acid leucine can be transcribed using six distinct codons.(UUA, UUG, CUU, CUC, CUA, and CUG).

Start and Stop Codons:

AUG is the universal start codon, which indicates the start of translation and encodes methionine. Three stop codons (UAA, UAG, and UGA) indicate the conclusion of the translation process, allowing the ribosome to release the freshly created protein.

Universal Code:

The genetic code is nearly universal across all organisms, from bacteria to humans. This universality suggests that all life forms share a common evolutionary ancestor.

Examples and Functions:

1. **Protein Synthesis:**
    - During translation, the ribosome reads the mRNA sequence in sets of three nucleotides (codons). Each codon is matched with a corresponding tRNA molecule that carries the appropriate amino acid, facilitating the synthesis of a polypeptide chain.

2. **Mutations and Genetic Variation:**

- Changes in the nucleotide sequence of DNA can lead to mutations, which may alter the codons and, consequently, the amino acids in a protein. Some mutations are silent (do not change the amino acid), while others can affect protein function or lead to diseases.

3. **Evolutionary Conservation:**

- The presence of similar gene, portion of genes, or chromosome segments in different species, reflecting both the common origin of species, and an important functional property of the conserved element in known as evolutionary conservation.
- The conservation of the genetic code across different species highlights its evolutionary significance and the fundamental nature of the processes involved in protein synthesis.

The genetic code is central to the understanding of molecular biology and genetics, providing the blueprint for translating genetic information into functional proteins that drive cellular activities and biological processes. A gene that has remain unchanged throughout evolution. Conservation of a gene indicates that it is unique and essential. There is not an extra copy of that gene with which evolution can thicker and changes in the gene are likely to be lethal.

viii. Gene Regulation

Gene regulation is key to the ability of an organization to respond to environmental changes. Gene regulation is one of the fundamental process that a cell carries out in order to produce the transcript that will lead to the proteins and it is essential function for which a lots of cells energy is devoted to. It is also something that probably a lot f cells genome is devoted to regulate so a cell will respond in the connect of stress or the development of the organism. Gene regulation is the process by which cells control gene expression, controlling when and to what extent a gene's output (typically a protein) is generated. This regulation is essential for cells' normal function, allowing them to adapt to environmental changes, differentiate into different cell types, and maintain homeostasis. Gene regulation can take place at several stages, including transcription, post-transcription, translation, and post-translational. So gene regulation is the process used to control the timing, location and amount in which gene are expressed. The process in complicated and is carried out by variety of un-changes, including through regulatory proteins and chemical modification of DNA.

1. **Transcriptional Regulation:**

Promoters and enhancers are examples of regulatory sequences that influence the commencement of transcription. Promoters are positioned near the start of genes and are recognized by

RNA polymerase, whereas enhancers are distant regulatory elements that interact with promoters to raise transcription levels. Proteins that attach to specific DNA sequences around the gene, promoting or inhibiting transcription. Activators (which increase transcription) and repressors (which reduce transcription) are two examples. The packing of DNA into chromatin influences its accessibility. Histone alterations and chromatin remodelling can condense (make chromatin less accessible) or relax (make it more accessible). Transcriptional regulation is the means by which a cell regulates the conversion of DNA to RNA (transcription) thereby archest rating gene activity. A single form altering the number of copies of RNA that are transcribed to the temporal control of when the gene is transcript. This control allows the cell or organisms to respond to a variety of intra and extracellular signals and thus mount a response.

2. **Post-Transcriptional Regulation**:

In eukaryotes, pre-mRNA is spliced to eliminate introns and link exons. Alternative splicing enables the creation of several mRNA isoforms from a single gene, resulting in distinct protein variations.The nucleotide sequence of RNA molecules can change after transcription, potentially affecting the protein output.The stability and decay of mRNA molecules affect the amount of protein generated. The 5' cap, 3' UTR, and poly-A tail are all regulatory factors in mRNA that regulate its stability. So post- transcriptional regulation occurs once the RNA polymerases has been attached to the genes promoter and is synthesizing the nucleotide sequence.

3. **Translational Regulation**:

The presence of translation initiation factors, as well as certain sequences in the mRNA (such as the 5' UTR), might influence translation initiation. mRNA localization to certain cell areas can have an impact on protein production and function. Proteins that bind to mRNA and impact translation by either inhibiting or facilitating ribosome binding or movement. So translational regulation refers to the control of the levels of proteins important to the cellular response to stressors, growth cues, and differentiation.

4. **Post-Translational Regulation**:

Proteins can be changed in a variety of ways, such as phosphorylation, acetylation, ubiquitination, and glycosylation, all of which affect their function, localization, stability, and interactions with other molecules. The ubiquitin-proteasome system and associated pathways control protein degradation, dictating levels and functional activity inside the cell. Post-translational regulation is the process of controlling the amounts of active proteins. There are various varieties. It is either

achieved by reversible processes (post-translational changes, such as phosphorylation or sequestration) or by irreversible actions (proteolysis).

Mechanisms of Gene Regulation:

Gene regulator is the process used to control the timing, location and amount in which gene are expressed. The process is carried out by a variety of changes including through regulatory proteins and chemical modification of DNA.

1. Operons (in Prokaryotes):

Operons are regulatory complexes that control and coordinate protein production in response to cellular requirements. The operon genes appear constantly on the RNA and are regulated by a single promoter. The promoter is then regulated by regulatory elements based on the cell's metabolic requirements. In bacteria, operons are collections of genes controlled by a single promoter and regulatory elements. The expression of genes within an operon can be coordinated in response to the cell's needs. For example, E. coli's lac operon regulates lactose metabolism.

1. Operons are regulatory complexes that control and coordinate protein synthesis based on requirements.
2. The operon genes occur continuously on the DNA and are controlled by a single promoter;.
3. The promoter is regulated regulatory elements based on the cells metabolic need
4. The main constituents continuously of operon are promoter.
5. They participate in gene regulation by inducing or repressing genes.
6. The regulatory proteins are either repressor or activators that bind to the operator gene and determine the fate of operon.
7. Operons are continuously found in bacteria's, virus and bactriophages.
8. Lac operon and trp operon are operon found in prokaryotes.

2. Epigenetic Regulation:

Epigenetic changes modify gene expression without altering the DNA sequence. These changes include DNA methylation and histone modifications, which can be inherited and affect gene expression patterns. Epigenetics is the study of inheritable features or persistent changes in cell function that occur without altering the DNA sequence. Epigenetic changes typically entail a change that is not erased by cell division and impacts gene expression regulation; such effects on cellular and physiological phenotypic features may be caused by environmental influences or be a natural aspect of development. They can lead to cancer.

3. **Small RNAs:**

MicroRNAs control gene expression by binding to complementary mRNA regions, causing mRNA breakdown or translation inhibition. Small interfering RNAs (siRNAs) are engaged in RNA interference (RNAi), which targets certain mRNAs for destruction and thereby silences gene expression.

4. **Signal Transduction Pathways:**

External inputs, like as hormones or growth factors, can activate signal transduction pathways, resulting in changes in gene expression. These routes frequently entail cascades of protein changes and interactions, which eventually influence transcription factors and other regulatory proteins.

Significance of Gene Regulation:

- Gene regulation ensures the correct timing and spatial expression of genes during development, leading to proper cell differentiation and organ formation.
- Cells can adapt to changes in their environment by regulating gene expression, allowing them to respond to stress, nutrient availability, and other factors.
- Dysregulation of gene expression can lead to various diseases, including cancer, genetic disorders, and autoimmune diseases.

ix. Transcriptional Regulation

Transcriptional regulation is the process by which cells control the transcription of specific genes, regulating when and to what extent a gene's output is produced. This control is necessary to ensure that genes are expressed at the appropriate periods and levels to suit the needs of the cell. Transcriptional regulation is made up of several different mechanisms and components, such as regulatory sequences, transcription factors, and chromatin changes. For instances In E. coli, the lac operon is controlled by lactose availability. Lactose deactivates a repressor protein, allowing genes required for lactose metabolism to be expressed. Transcriptional regulation is important in eukaryotic development because it controls whether specific genes are switched on or off in different tissues or stages. Components of Transcriptional Regulation are:

1. **Promoters:**

Promoters are DNA sequences positioned upstream of a gene that serve as a binding site for RNA polymerase and general transcription factors to commence transcription. In eukaryotes, the transcription machinery recognizes certain motifs in promoters, such as the TATA box.

2. **Enhancers:**

Enhancers are DNA sequences that can be found distant from the gene that they control. They improve the chance of transcription by binding transcription factors and interacting with the promoter. Enhancers can loop around to bring distant regulatory components closer to the promoter, so promoting transcription start.

3. **Silencers:**

Silencers are DNA sequences that inhibit gene transcription. They bind repressor proteins, which prevent the promoter from interacting with the RNA polymerase. Silencers help to control the time and location of gene expression by inhibiting transcription when it is not required

4. **Transcription Factors:**

Transcription factors attach to certain DNA regions (enhancers or promoters) and boost transcription initiation by enabling transcription machinery assembly. Transcription factors bind to silencers or other regulatory areas, inhibiting transcription by preventing RNA polymerase or other activators from binding. Proteins essential for the basal transcription machinery, such as RNA polymerase and other promoter-initiated transcription factors.

5. **Chromatin Modifications:**

Chemical modifications to histone proteins (such as acetylation, methylation, and phosphorylation) can modify chromatin structure and affect gene expression. For example, histone acetylation stimulates transcription by increasing chromatin accessibility. The addition of methyl groups to DNA, most commonly at cytosine residues in CpG dinucleotides, can suppress transcription by blocking transcription factor binding or attracting repressive proteins

6. **Coactivators and Corepressors:**

Proteins that interact with activators to boost transcription by altering chromatin shape or recruiting more transcription machinery. Proteins that bind with repressors and suppress transcription by increasing chromatin condensation or preventing transcription machinery assembly.

Mechanisms of Transcriptional Regulation:

1. **Positive Regulation:**

Positive regulation occurs when activators bind to enhancers or promoters, increasing the chance of transcription.In the presence of a specific signal, an activator protein may attach to an enhancer, boosting the recruitment of RNA polymerase to the promoter and increasing transcription.

2. **Negative Regulation:**

Negative regulation occurs when repressors bind to silencers or other regulatory areas, decreasing the chance of transcription. A repressor protein may bind to a silencer region, preventing RNA polymerase from connecting to the promoter and so suppressing transcription.

3. **Gene Regulatory Networks:**

Complex networks of interactions between transcription factors, enhancers, silencers, and other regulatory components regulate the expression of many genes in response to developmental signals, environmental conditions, or cellular requirements.

4. **Feedback Loops:**

Feedback loops regulate gene expression through the gene product or downstream signaling pathways, allowing cells to fine-tune their responses and maintain homeostasis.

x. Molecular Techniques

Molecular biology relies on various techniques to study and manipulate nucleic acids and proteins:

a) Polymerase Chain Reaction (PCR)

The Polymerase Chain Reaction (PCR) is a popular molecular biology technique for amplifying specific DNA sequences, allowing researchers to create millions of copies of a single section of DNA from a small starting material. PCR is used for a variety of purposes, including genetic research, medical diagnostics, forensic analysis, and more.

1. **DNA Template**: Sample DNA with the target sequence to be amplified. This could come from any biological source, including blood, tissue, or a microbiological sample.

2. **Primers**: Short, single-stranded DNA sequences that complement the target's flanking regions. Two primers are used: a forward primer and a reverse primer, which anneal to the opposite strands of DNA.

3. **DNA Polymerase**: Enzymes that add complimentary nucleotides to template DNA to create new strands. Taq polymerase, produced from the bacterium Thermus aquaticus, is the most often used polymerase in PCR because it is heat stable and can resist the high temperatures employed throughout the process.

4. **Deoxynucleotide Triphosphates (dNTPs)**: DNA building blocks, such as dATP, dTTP, dCTP, and dGTP, are added to the developing strand during amplification.

5. **Buffer and Salt**: A buffer solution and salts (such as magnesium chloride) are added to maintain the optimal pH and ionic environment for the enzyme activity and DNA polymerase function.

Basic PCR Process:

1. **Denaturation**: The reaction mixture is heated to approximately 94-98°C (201-208°F). This high temperature denatures, or separates, the double-stranded DNA into two single strands by breaking the hydrogen bonds between complementary bases.
2. **Annealing**: The temperature is reduced to roughly 50-65°C (122-149°F), depending on the primers' melting temperature. This enables the primers to bind (anneal) to the complementary sequences of the single-stranded DNA template. The annealing temperature is carefully calibrated to ensure that the primers bind specifically.
3. **Extension**: The temperature is raised to around 72°C (162°F), which is the ideal temperature for Taq polymerase activity. The DNA polymerase extends primers by adding nucleotides to the 3' end, resulting in the synthesis of new DNA strands that are complementary to the template strand.

The denaturation, annealing, and extension steps are repeated in multiple cycles (typically 20-40 cycles). Each cycle doubles the amount of DNA, leading to exponential amplification of the target sequence.

b) Gel Electrophoresis

Gel electrophoresis is a laboratory technique for separating and analyzing macromolecules such as DNA, RNA, and proteins according to their size and charge. The method entails combining these molecules in a gel matrix, which is commonly comprised of agarose or polyacrylamide. This gel behaves like a molecular sieve.The basic steps of gel electrophoresis are as follows:

- **Preparation of the Gel**: The gel is created by dissolving agarose or polyacrylamide in a buffer solution and then cooling and solidifying in a mold. This produces a gel slab with wells in which samples can be loaded.
- **Loading the Samples**: The macromolecule samples, combined with a loading dye, are pipetted into the gel's wells. The loading dye helps to visualize the samples and ensures they sink into the wells.
- **Applying an Electric Field**: The gel is placed in an electrophoresis chamber containing an electrically conductive buffer solution. An electric field is applied across the gel, forcing charged molecules to move through the matrix. DNA and RNA are negatively charged and travel

towards the positive electrode (anode), however proteins can go to either electrode depending on their net charge at a given pH.

- **Separation of Molecules**: The gel matrix functions as a sieve, allowing smaller molecules to go quicker and further than bigger ones. This leads to the separation of molecules based on size.
- **Visualization**: Following electrophoresis, the separated molecules are frequently dyed to make them visible. Ethidium bromide or SYBR Green is typically used for DNA, whereas Coomassie Brilliant Blue or silver stain is utilized for proteins. The stained gel is then inspected using UV light or other imaging methods to determine the results

c) DNA Sequencing

DNA sequencing is the process of determining the exact sequence of nucleotides (adenine, cytosine, guanine, and thymine) in a DNA molecule. This technique is crucial for understanding genetic information, identifying genetic variations, and advancing fields such as genomics, medicine, and evolutionary biology. DNA sequencing allows scientists to analyze the genetic makeup of organisms, study mutations, and explore genetic diversity. There are several methods are used for DNA Sequencing:

1. **Sanger Sequencing:**

This approach, also called chain-termination sequencing, was developed by Frederick Sanger. It is the selective inclusion of chain-terminating dideoxynucleotides (ddNTPs) during DNA replication. The DNA to be sequenced is first amplified via PCR.The DNA is combined with a DNA polymerase, standard deoxynucleotides (dNTPs), and a small amount of fluorescently tagged dNTPs. Each ddNTP ends the DNA chain at a distinct length.The resultant DNA fragments are size-separated using capillary electrophoresis. The fluorescent labels on the ddNTPs are detected, and the sequence is read using the color emitted by each ddNTP. Suitable for smaller-scale sequencing initiatives, such as sequencing individual genes or specific areas of the genome.

2. **Next-Generation Sequencing (NGS):**

NGS technology, also known as high-throughput sequencing, enable the simultaneous sequencing of millions of DNA fragments, resulting in a full view of the genome.DNA is split into smaller fragments, with adapters attached to both ends of each fragment.DNA fragments are sequenced using a variety of techniques, including sequencing-by-synthesis (Illumina), sequencing-by-ligation (SOLiD), and ion semiconductor sequencing (Ion Torrent).The sequence readings are then aligned and combined to recover the original DNA sequence. Ideal for large-scale applications like whole-genome

sequencing, exome sequencing, and transcriptome analysis.Next-generation sequencing refers to genome sequencing, resequencing, transcriptome profiling, and protein-DNA interactions. Resequencing is required to investigate genome variants, which cannot be accomplished by studying a single genome. There is an increasing need for low-cost sequencing, which has led to the development of next-generation sequencing or high-throughput sequencing technology capable of producing dozens or millions of sequences concurrently.

3. **Third-Generation Sequencing**:

These approaches outperform NGS in terms of read length and sequencing speed. Pacific Biosciences developed this approach to sequence individual DNA molecules in real time utilizing zero-mode waveguide (ZMW) technology. Oxford Nanopore Technologies developed this technology for sequencing DNA by passing it through a nanopore and monitoring electrical current changes. Useful for sequencing complicated genome areas, finding structural variants, and real-time sequencing applications.

xi. Cloning

Cloning is the process of creating genetically identical copies of an organism, cell, or DNA sequence. Cloning can be performed at various levels, from molecular cloning of DNA fragments to whole-organism cloning. The primary methods of cloning include molecular cloning, cellular cloning, and reproductive cloning. Each type of cloning has different applications and implications. There are several types of Cloning:

1. **Molecular Cloning**:

Cloning is a set of experimental methods in molecular biology that are used to assemble recombinant DNA molecules and to direct their replication within host organism. It is the process of producing several copies of a given DNA segment. Molecular cloning allows researchers to reproduce and examine certain genes or DNA sequences. The DNA segment of interest is extracted from the source organism. The DNA fragment is put into a vector (plasmid or viral DNA) that can reproduce within the host cell. The vector is inserted into a host cell (such as bacteria, yeast, or mammalian cells).Host cells that have successfully taken up the vector are chosen and tested for the presence of cloned DNA. Gene expression studies, protein production, functional gene analysis, and genetic engineering.

2. **Cellular Cloning**:

Unicellular organisms such as bacteria and yeast naturally produce clones of themselves when they replicate asexually by binary fission this is known as cellular cloning. The nuclear DNA replicates

by the process of mitosis which creates an exact replica of the genetic material. The technique of creating several identical cells from a single cell. Cellular cloning is commonly employed in research to generate cell lines with specific properties. A single cell is isolated and grown in a laboratory. The cell is allowed to divide, resulting in a population of genetically identical cells. Cellular functions are investigated, drugs are tested, and cell lines are created for research and biotechnology. The therapeutic cloning involves creating a cloned embryo for the sole purpose of producing embryonic stem cells with same DNA as the donor cell. These stem cell can be used in experiments aimed at understanding diseases and developing new treatment or diseases.

3. **Reproductive Cloning:**

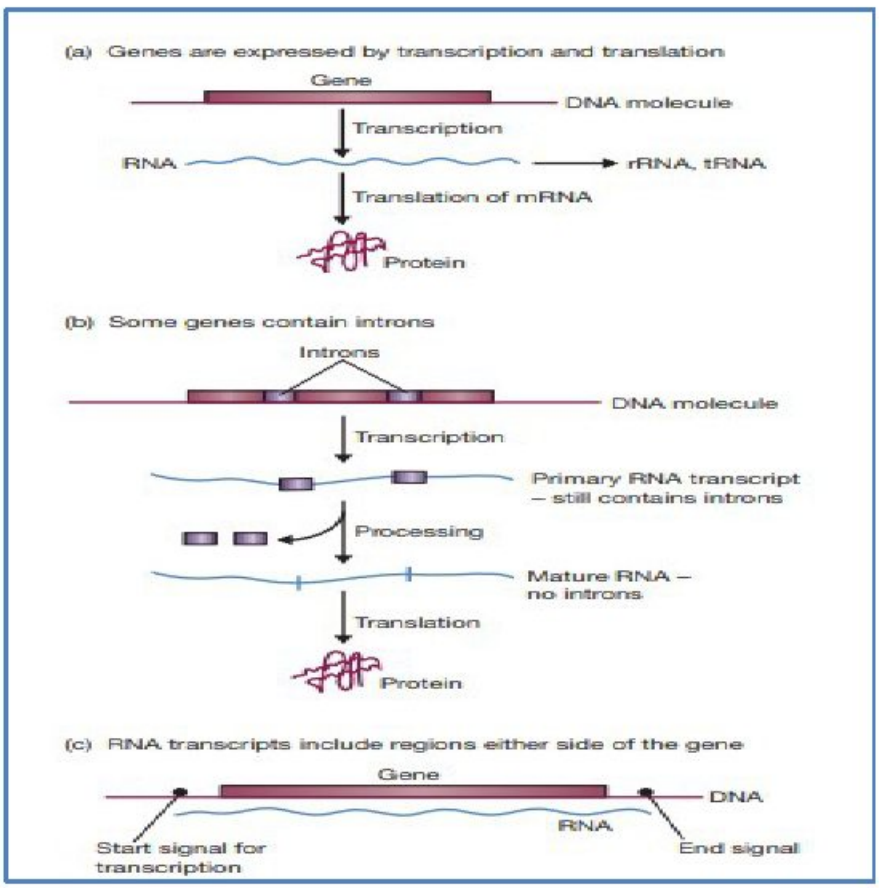
Reproducing cloning is defined as the deliberate promotion of genetically identical individual. Each newly produced individual is a clone of the original monozygotic (identical) twins is natural clones. Clones contain identical set of genetic material in the nucleus, the compartment that contains chromosomes of every cell in their bodies. Thus the cell form two clones have the some DNA and the some gene in their nucleus. The process of developing a genetically similar organism from a somatic cell. Reproductive cloning tries to create a new organism that is genetically identical to an existing one. A somatic cell (any cell other than a sperm or egg) has its nucleus removed and transferred to an enucleated egg cell. The egg cell, which now contains the somatic nucleus, is stimulated to mature into an embryo. The embryo is put into a surrogate mother and grows into a complete organism. Research into developmental biology and its potential for agricultural and medical applications, yet ethical and technical barriers prevent practical implementation. Cloning allows researchers to study the functions of specific genes and their roles in health and disease. Cloning genes into expression systems can produce proteins for research, therapeutics, and vaccines. Molecular biology has numerous applications in medicine, agriculture, and biotechnology.

1.3. Gene expression

All genes have to be expressed in order to function. Gene expression is the process by which information encoded in a gene is translated into a function. This mostly occurs by via the transcription of RNA molecules, that code for proteins or non coding RNA molecules that serve other functions. Gene expression begins with transcription into a complementary RNA strand. For some genes, such as those that encode transfer RNA (tRNA) and ribosomal RNA (rRNA), the transcript is the functionally significant molecule.

For other genes, the transcript is converted into a protein molecule.This is mostly accomplished by the transcription of RNA molecules that code for proteins or non-coding RNA molecules that perform other tasks. Gene expression can be thought of as a "on/off switch" that controls when and where RNA molecules and proteins are produced, as well as a "volume control" that determines how much of each product is produced.

Gene expression is closely regulated, with significant variations depending on the environment and cell type. Many genes' RNA and protein products control the expression of other genes. This process is essential to all forms of life and is strictly regulated to ensure that genes are expressed at the proper time, in the correct location, and in the appropriate amounts. Generally, we can assess gene expression by examining a phenotype or trait. Examples include measuring protein activity. If a protein's activity can be measured, the gene that encodes that protein is most likely turned on, or we can characterize it as such. We can also seek for trends and characteristics. For example, if a beautiful butterfly wing with numerous colors is the consequence of distinct genes being turned on in different parts of the butterfly wing, we may evaluate that gene expression simply by examining the butterfly's wing and identifying the locations of the colors.The primary phases of gene expression include transcription, RNA processing, translation, and post-translational modifications.



1. Transcription

The first stage in gene expression is called transcription. During transcription, the DNA of a gene is converted into an RNA transcript known as messenger RNA (mRNA). RNA is chemically similar to DNA, but it has only one strand of bases rather than two. And, while DNA has the base thymine (T), RNA has the base uracil (U). RNA polymerase, an enzyme found inside the nucleus, is responsible for transcription. After the gene is transcribed, the mRNA leaves the nucleus and transports the information to the ribosomes, the cell's protein factories.

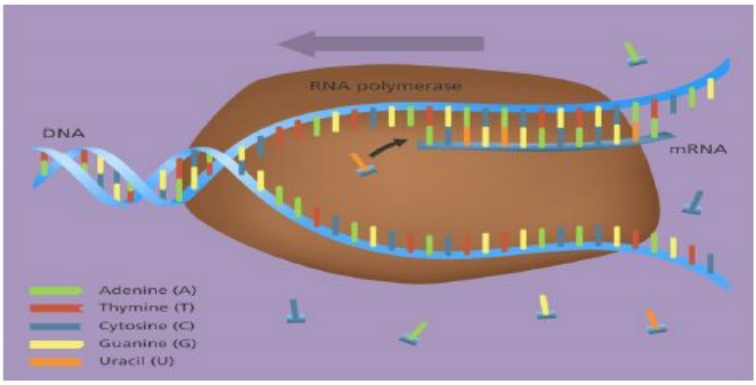


Fig.1.3: An illustration showing the process of transcription – where the DNA of a gene is transcribed into messenger RNA (mRNA).Image credit: Laura Olivares Boldú / Wellcome Connecting Science

Source: What is gene expression? (yourgenome.org)

- **Initiation**:Transcription begins at a promoter, which is a specific DNA sequence positioned upstream of the gene. RNA polymerase and transcription factors work together to bind to the promoter. The DNA double helix unwinds at the promoter, creating a transcription bubble and exposing the template strand.
- **Elongation**:RNA polymerase proceeds along the DNA template strand, introducing complementary RNA nucleotides (adenine pairs with uracil, cytosine with guanine). RNA is generated from the 5' to 3' direction, resulting in a developing RNA strand.
- **Termination**:Transcription continues until the RNA polymerase finds a termination signal in the DNA sequence. The RNA transcript is liberated from the DNA template, while the RNA polymerase separates from the DNA.

2. RNA Processing

In eukaryotes, the initial RNA transcript (pre-mRNA) undergoes several modifications before becoming mature mRNA that can be translated into a protein.

- **Capping**: A modified guanine nucleotide is added to the 5' end of the pre-mRNA. This cap protects the RNA from degradation and helps in ribosome binding during translation.
- **Splicing**: Non-coding regions (introns) are removed from the pre-mRNA by the spliceosome. The coding regions (exons) are joined together to form a continuous coding sequence.
- **Polyadenylation**: A series of adenine nucleotides (poly-A tail) is added to the 3' end of the pre-mRNA. This tail enhances the stability of the RNA and facilitates its export from the nucleus.

1. Translation

The second step in gene expression is called translation. During translation, mRNA is transformed into a string of amino acids, which are the basic building blocks of proteins. When mRNA reaches the ribosomes, it is read by a carrier molecule known as transfer RNA (tRNA). The tRNA reads mRNA in three-base chunks called codons, each of which encodes one amino acid. For example, the codon 'GGU' encodes the amino acid glycine. The tRNA briefly attaches to the mRNA codon and transports the matching amino acid to the ribosome. The ribosome binds the amino acids together, forming a lengthy chain known as a polypeptide. This process continues until the ribosome reaches a'stop' codon, which is a collection of three nucleotides that indicates the end of the protein. There are following involved in Translation

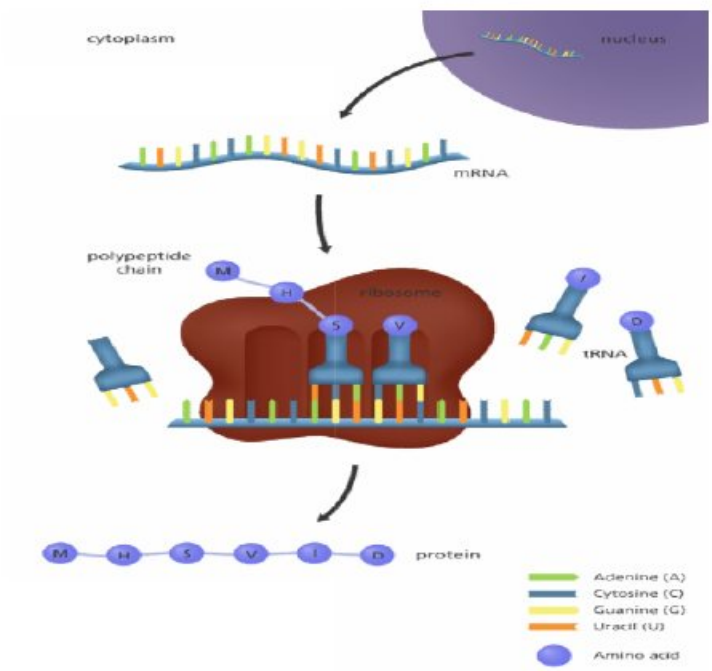


Fig.1.4: An illustration showing the process of translation – where the mRNA is translated into amino acids, creating a protein.Image credit: Laura Olivares Boldú / Wellcome Connecting Science

Source: Source: What is gene expression? (yourgenome.org)

1. **Initiation**:

- **mRNA Binding**: The small ribosomal subunit attaches to mRNA around the 5' cap and looks for the start codon (AUG).
- **tRNA Binding**: A tRNA molecule bearing the complementary anticodon (UAC) and methionine (the amino acid predicted by AUG) attaches to the start codon.
- **Ribosome Assembly:** The big ribosomal subunit joins the complex, resulting in a functioning ribosome.

2. **Elongation**:

- **Codon Recognition**: The ribosome proceeds along the mRNA, and each codon is recognized by a corresponding tRNA that contains the proper amino acid.
- **Peptide Bond Formation**: The ribosome catalyzes the synthesis of peptide bonds between adjacent amino acids, which extends the polypeptide chain.
- **Translocation**: The ribosome proceeds along the mRNA, moving the tRNA from the A site to the P site, then to the E site, where the empty tRNA is released.

3. **Termination**:

- **Stop Codon Recognition**: When the ribosome meets a stop codon (UAA, UAG, or UGA), translation stops.
- **Release Factor Binding**: Release factors bind to the stop codon, facilitating polypeptide release and ribosome disassembly.

4. **Post-Translational Modifications**

Polypeptides are frequently subjected to further changes after translation in order to function completely. These changes can influence the protein's stability, activity, localization, and interactions with other molecules.

Regulation of Gene Expression

Gene expression is tightly regulated at multiple levels to ensure precise control over cellular functions. Regulation can occur at the transcriptional, post-transcriptional, translational, and post-translational levels.

Transcriptional Regulation

1. **Transcription Factors**:

In molecular biology and genetics transcriptional regulation is means by which a cell regulates the conversion of DNA to RNA, thereby orchestrating gene activity. Proteins that bind to specific DNA sequences (enhancers and silencers) to increase or decrease transcription.

2. **Epigenetic Modifications**:

- **DNA Methylation**: Addition of methyl groups to cytosine residues, typically repressing gene expression.
- **Histone Modifications**: Chemical modifications of histones (e.g., acetylation, methylation) affecting chromatin structure and gene accessibility.
- **Transcriptions**: the process of making RNA from a DNA template by RNA polymerase.
- **Transcriptional factors**: a substance such as a protein that contributes to the cause of specific biochemical reactions or bodiles process.
- **Promoters**: a region of DNA that initiate transcription of particular gene.
- **Sigma factors:** Specialize bacterial cofactors that complete with RNA polymerase and encode sequence specifically.
- **Coactivator:** a protein that works with transcription factors to increase the rate of gene transcription.

3. **Chromatin Remodeling**:

- **Nucleosome Positioning**: ATP-dependent complexes reposition nucleosomes, altering access to DNA.

Post-Transcriptional Regulation

1. **Alternative Splicing**:

- **Generating Diversity**: Different combinations of exons can be spliced together, producing multiple mRNA variants from a single gene.

2. **RNA Editing**:

- **Base Modification**: Enzymes alter nucleotide sequences in RNA, potentially changing the encoded protein.

3. **mRNA Stability**:

- **Regulatory Elements**: Sequences in the mRNA (e.g., AU-rich elements) influence its degradation rate.

4. **RNA Interference (RNAi):**

- **siRNAs and miRNAs**: Small RNAs that bind to mRNA, leading to its degradation or translational repression.

Translational Regulation

1. **Ribosome Binding:**

- **Regulatory Proteins**: Proteins that bind to the mRNA and affect its translation efficiency.

2. **Internal Ribosome Entry Sites (IRES):**

- **Alternative Initiation**: Some mRNAs contain IRES elements that allow translation initiation independent of the 5' cap.

Post-Translational Regulation

1. **Protein Modifications:**

- **Chemical Modifications**: Affect protein function, localization, and interactions.

2. **Protein Degradation:**

- **Ubiquitin-Proteasome System**: Targeted degradation of proteins by the proteasome.

Techniques for Studying Gene Expression

Gene expression research focuses on how genes are transcribed into RNA and translated into proteins. To measure and study gene expression, a variety of approaches are utilized, each with their own set of advantages and uses. Here are some common techniques:

Quantitative PCR (qPCR)

It is also called real time PCR or quantitative real time PCR. It measures the amount of a given RNA molecule by first converting it to complementary DNA (cDNA) and then amplifying it with PCR It is PCR based techniques that couple amplification of a target DNA sequence with quantification of the concentration of that DNA species in reaction. . Used to accurately quantify gene expression levels, validate microarray data, and detect low-abundance transcripts.

2. RNA Sequencing (RNA-Seq)

It is techniques that use next generation sequencing to reveal the presence of quantity of RNA molecules in biological sample providing a snapshot of gene expression in the samples, salso known as transcriptome. RNA-Seq includes sequencing the entire transcriptome, which provides a complete picture of all expressed genes. It counts RNA levels and detects new transcripts, splice variants, and fusion genes. Used to identify new genes, comprehend complex transcriptional landscapes, and compare gene expression across circumstances or tissues.

3. Microarrays

It is multiplex lab-on -a-chip. Its purpose is to simultaneously detect the expression of thousands of biological interaction, It is a two dimensional array on a solid substance-usually a glass slid or silicon thin film cell- that assays (test) large amount of biological material using high throughout screening miniaturized, multiplexed and parallel processing and detection methods. The concept and methodology of microarrays was first introduced and illustrated in antibody microarrays by Tse Wen change in 1983 in scientific publication. Microarrays employ a grid of DNA probes on a solid surface to hybridize with labeled cDNA from samples. The level of hybridization reflects gene expression levels. Used to profile gene expression under various situations, discover differentially expressed genes, and investigate gene function and regulation.

4. Northern Blotting

It refers to laboratory procedures for detecting a specific RNA sequence in blood or tissue samples. This approach separates RNA molecules using gel electrophoresis, transfers them to a membrane, and then detects specific RNAs with tagged probes. Used to determine the size and amount of individual mRNA molecules, as well as to study alternative splicing and RNA processing.

5. In Situ Hybridization (ISH)

It is a type of hybridization that uses a labelled complementary DNA,RNA or modified nucleic acid strand (i.e. a probe) to localise specific DNA or RNA sequence is a portion or section of tissue (in-situ) or if the tissue is small enough (e.g. plants seeds drosophila embryos) in the entire tissue. ISH detects the spatial and temporal expression of genes within tissue slices or whole organisms by using tagged RNA or DNA probes.Used to investigate gene expression patterns during development, in various tissues, and in response to environmental changes.

6. Western Blotting

Western blotting is a technique that uses antibodies to detect certain proteins in samples. The method involves using gel electrophoresis to separate the samples proteins. The separate proteins are transferred out of the gel to the surface of a membrane. This is some time used to diagnose disease. While it detects protein levels rather than mRNA, it offers information on gene expression at the protein level. Modifications and analysis of protein-protein interactions.

6. Reporter Gene Assays

Reporter Gene Assays are typically used to measure the regulatory ability of an known DNA-sequence. This is done by thinking the unknown promoter sequence to an easily detectable reporter gene whose product can be easily detected and quantifiably measured. Reporter genes, like luciferase or GFP, are fused to the gene of interest's regulatory regions. Their expression is used to determine the

activity of the gene's promoter and other regulatory elements.Used to investigate gene regulation, identify regulatory elements, and test for transcription factor activity.

8. Chromatin Immunoprecipitation (ChIP)

This is type of experimental techniques used to investigate the interaction between protein and DNA in the cell. It aims to determine whether specific proteins are associated with specific genomic regions such as transcription factor on proteins or other DNA binding sites and possibly defined cistromes. ChIP is the process of cross-linking proteins to DNA, immunoprecipitating DNA-protein complexes with particular antibodies, and analyzing the associated DNA to find transcription factor binding sites and histone changes. It is used to investigate protein-DNA interactions, epigenetic changes, and transcriptional regulation.

9. Single-Cell RNA Sequencing (scRNA-Seq)

It has become a powerful tool for describing cell sub population classification and cell heterogeneity by achieving high throughput and multidimensional analysis of individual cells and circumventing short coming of traditional sequencing. scRNA-Seq measures gene expression at the single-cell level, providing information about cellular heterogeneity and unusual cell populations. Used to investigate developmental processes, tumor heterogeneity, and immunological responses.These techniques provide a comprehensive toolkit for studying gene expression, each offering unique insights into the complex regulation and function of genes in various biological contexts.

Applications of Gene Expression Studies

- Identifying biomarkers for cancer, cardiovascular disease, and genetic abnormalities.
- Determining illness stages and progression.
- Identifying pharmacological targets and understanding their mechanisms of action.
- Personalized treatment by predicting patient responses to medications.
- Understanding gene function and relationships.
- Identify regulatory elements and processes.
- Investigating gene expression patterns throughout development and differentiation.
- Understanding the role of certain genes in embryogenesis and organ development.
- Profiling gene expression in cancer types to find oncogenes and tumor suppressors.
- Develop tailored medicines and track treatment outcomes.
- Understanding gene expression in brain development, function, and diseases.

- Identifying genes associated with neurodegenerative illnesses and mental health disorders.
- Understanding gene expression in immune cells and immunological responses.
- Identifying targets for vaccines and immunotherapy.
- Gene expression research aim to improve crop attributes like yield, disease resistance, and stress tolerance.
- Developing GMOs with desired properties.
- Investigating the effect of environmental changes and contaminants on gene expression in organisms.
- Monitoring ecosystem health and biodiversity using gene expression profiles.
- Comparing gene expression patterns across species to better understand evolutionary links.
- Investigating how gene expression impacts adaptation and speciation.
- Adapting medical therapies based on unique gene expression profiles.
- Developing targeted medicines for complicated disorders.
- Identifying target genes for therapeutic interventions.
- Tracking the impact of gene therapy on gene expression in treated cells and tissues.

1.4. Gene expression in prokaryotes

The control of gene expression in prokaryotes can be positive or negative. In general, negative regulation involves RNA polymerases and the repressor, which presents RNA polymerases from transcribing prokaryotic genes. Gene expression is simpler in prokaryotes than in eukaryotes due to their simpler cellular structure and organization. Bacteria and other prokaryotic cells lack a nucleus, allowing transcription and translation to take place concurrently in the cytoplasm. Prokaryotic gene expression is a well-coordinated process that requires accurate transcription and translation regulation. The ability to rapidly respond to environmental changes through operons, sigma factors, and regulatory proteins is crucial for the survival and adaptability of prokaryotic organisms. Understanding these mechanisms not only provides insights into fundamental biological processes but also has practical applications in biotechnology and medicine.Gene expression in eukaryotes is a highly regulated and complex process, involving multiple steps from the DNA level to the production of functional proteins. This process occurs within a more compartmentalized cellular environment compared to prokaryotes. The main stages of gene expression in eukaryotes include transcription, RNA processing, translation, and post-translational modifications.

Additionally, gene expression in eukaryotes is subject to sophisticated regulatory mechanisms at each stage.

Types of Gene Expression:

- **Constitutive Gene Expression:** It is a gene that is transcribed continually as opposed to a facultative gene, which is only transcribed when needed. It also refers to genes that are constitutively expressed in most cells, such as those that encode housekeeping function proteins such as ribosomal proteins, tRNA, and rRNA. Other genes produce proteins that are only essential for cell proliferation under specific environmental conditions. Thus, constitutive expression of these genes is a waste of energy.
- **Inducible gene expression:** It is a gene whose expression is either responsive to environmental changes or dependent on the position in the cell cycle. Induction is an increase in gene expression due to the presence of an inducer. While our gene provide all the instruction for the proteins. We make our individual traits are influenced by the regulation of gene expression. Such genes are expressed only when their product is required. Inducible gene expression is the process of activating gene expression in response to a chemical in the environment, and these genes are referred to as inducible genes. Catabolic pathway enzymes, such as lactose and galactose, are examples of inducible expression.
- **Repressible gene expression:** Repressible gene are those in which the presence of a substance (a corepressor) in the environment turns off the expression of those gene (structural genes) involved in metabolism of that substance e.g. tryptophan represses the expression of trap gene. If a biosynthetic enzyme is present in an environment where bacteria proliferate, its ongoing synthesis is a waste. Thus, in the event of an externally available nutrition molecule, the enzymes required for its manufacture must be turned off. Such expression is referred to as repressible gene expression.

Transcription in Prokaryotes

RNA polymerase

In prokaryotes, a single RNA polymerase synthesizes all forms of RNA (mRNA, tRNA, and rRNA). The RNA polymerase holoenzyme consists of the core enzyme (composed of α, β, β′, and ω subunits) and a sigma factor (σ) that promotes transcription start by allowing the RNA polymerase to recognize and bind to certain promoter regions on the DNA.

Promoters

Promoters are DNA sequences located upstream of the coding region of a gene that signal the RNA polymerase where to begin transcription. Prokaryotic promoters typically have two key regions:

- **-10 region (Pribnow box)**: A consensus sequence (TATAAT) located about 10 bases upstream of the transcription start site.
- **-35 region**: Another consensus sequence (TTGACA) located about 35 bases upstream of the transcription start site.

Steps of Transcription

1. **Initiation**:

- The RNA polymerase holoenzyme binds to the promoter, forming a closed complex.
- The DNA is unwound to form an open complex, allowing RNA polymerase to access the template strand.
- Transcription begins at the +1 site, where the first ribonucleotide is added.

2. **Elongation**:

- The sigma factor dissociates from the RNA polymerase after initiation.
- RNA polymerase synthesizes the RNA strand by adding ribonucleotides complementary to the DNA template strand, moving in the 5′ to 3′ direction.

3. **Termination**:

- Transcription continues until RNA polymerase encounters a termination signal, which can be either rho-dependent or rho-independent.
- **Rho-dependent termination**: The rho protein binds to the nascent RNA and moves along it until it catches up with the RNA polymerase, causing termination.
- **Rho-independent termination**: A GC-rich hairpin loop followed by a series of uracil residues in the RNA causes the RNA polymerase to pause and dissociate from the DNA template.

Translation in Prokaryotes

In prokaryotes, translation begins as transcription continues, a process known as linked transcription-translation. The ribosome binds to the mRNA as soon as the ribosome binding site (Shine-Dalgarno sequence) becomes available.

Shine-Dalgarno Sequence

This ribosomal binging site in bacterial messenger RNA become known as the shine-Dalgarno (SD) sequence. It enables initiation of proteins synthesis by aligning the ribosome with the start codon. Simply put genes are read in groups of the three letters, but you need to let the ribosome known where

to start, this sequence is common is bacteria, but rarer in archaea. It is also present in some chloroplast and mitochondrial transcripts. This sequence was proposed by Australian scientist John Shine and Iynn Dalgarno in 1973. The Shine-Dalgarno sequence (AGGAGG) is a purine-rich region in mRNA that occurs upstream of the start codon (AUG). It complements a region of the 16S rRNA in the small ribosomal subunit, allowing the ribosome to bind to the mRNA.

Steps of Translation

1. **Initiation:**

- The small ribosomal subunit binds to the Shine-Dalgarno sequence on the mRNA.
- The initiator tRNA carrying formylmethionine (fMet) binds to the start codon (AUG) on the mRNA.
- The large ribosomal subunit joins the complex, forming a functional ribosome.

2. **Elongation:**

- Aminoacyl-tRNAs, carrying specific amino acids, enter the A site of the ribosome.
- Peptidyl transferase, an enzymatic activity of the ribosome, forms peptide bonds between the amino acids.
- The ribosome translocates along the mRNA, moving the tRNA from the A site to the P site, and the empty tRNA exits from the E site.

3. **Termination:**

- When a stop codon (UAA, UAG, or UGA) is encountered, release factors bind to the ribosome, promoting the release of the newly synthesized polypeptide.
- The ribosomal subunits dissociate, ready to initiate another round of translation.

Regulation of Gene Expression

The gene expression can be broadly regulated at five different levels in prokaryotes

Gene expression in prokaryotes is primarily regulated at the transcriptional level, allowing cells to respond quickly to environmental changes. Key mechanisms of regulation include operons, sigma factors, and regulatory proteins.

Operons

An operon is a cluster of genes transcribed as a single mRNA molecule, regulated by a common promoter and regulatory sequences. Operons allow the coordinated expression of genes that function in the same pathway. The most well-known example is the lac operon in E. coli, which controls the metabolism of lactose.

- **lac Operon**: The lac operon consists of three structural genes (lacZ, lacY, and lacA) involved in lactose metabolism. It is regulated by the lac repressor (LacI) and the catabolite activator protein (CAP).
- **In the absence of lactose**: The LacI repressor binds to the operator, blocking transcription.
- **In the presence of lactose**: Lactose (or allolactose) binds to LacI, causing a conformational change that releases LacI from the operator, allowing transcription.
- **In the absence of glucose**: cAMP levels increase, and cAMP binds to CAP. The CAP-cAMP complex binds to the promoter, enhancing RNA polymerase binding and transcription.

Sigma Factors

Sigma factors are proteins that bind to RNA polymerase and direct it to specific promoters. Different sigma factors recognize different sets of promoters, allowing the cell to regulate groups of genes in response to environmental conditions. For example:

- **σ70**: The primary sigma factor in E. coli, responsible for the transcription of housekeeping genes.
- **σ32**: Activated under heat shock conditions, inducing the expression of heat shock proteins.

Regulatory Proteins

Regulatory proteins can act as activators or repressors, influencing the binding of RNA polymerase to the promoter.

- **Activators**: Bind to enhancer sequences or the promoter region, increasing transcription.
- **Repressors**: Bind to operator sequences or other regulatory regions, blocking transcription.

1.5. Gene expression in Eukaryotes

A eukaryote is defined as any cell or creature with a clearly defined nucleus. The eukaryotic cell has a distinct nuclear membrane that surrounds the nucleus, where chromosome pairs are encountered. Organelles found in eukaryotic cells include mitochondria, the Golgi apparatus, the endoplasmic reticulum, and lysosomes.

Eukaryotic transcription

Eukaryotic transcription is defined as the process by which eukaryotic cells repeat data in a strand of DNA and copy genetic information into a new molecule of RNA. Eukaryotic transcription is

the initial stage in gene expression, in which certain portions of DNA are converted into RNA by a unique enzyme called RNA polymerase. This produces an antiparallel RNA fiber known as a main transcript.Eukaryotic transcription is carried out in the nucleus of the cell by one of three RNA polymerases, depending on the RNA being transcribed, and occurs in three successive stages:

1. Initiation
2. Elongation
3. Termination

Unlike prokaryotic RNA polymerase, which can connect to a DNA template on its own, eukaryotes require several additional proteins, known as transcription factors, to first bind to the promoter region and then recruit the proper polymerase. A transcription pre-initiation complex (PIC) is formed after all of the transcription factors and RNA polymerase have been assembled and bound to the promoter.

The most well studied core promoter element in eukaryotes is a short DNA sequence known as a TATA box, which is located 25-30 base pairs upstream of the transcription start point. Only 10-15% of mammalian genes have TATA boxes, with the remainder containing alternative core promoter elements; however, the methods by which transcription is begun at TATA box promoters are well understood.

The TATA box, as a key promoter region, serves as the binding site for TATA-binding protein (TBP), a subunit of Transcription Factor II D (TFIID). After TFIID attaches to the TATA box via the TBP, five other transcription factors and RNA polymerase create a pre-initiation complex in a sequence of phases. Transcription Factor II H (TFIIH) is responsible for separating opposing strands of double-stranded DNA so that the RNA polymerase can access a single-stranded DNA template. However, the pre-initiation complex alone drives a modest, or basal, transcription rate.

Other proteins, known as activators and repressors, as well as any associated coactivators or corepressors, control transcription rate. Activator proteins promote transcription, whereas repressor proteins inhibit transcription.

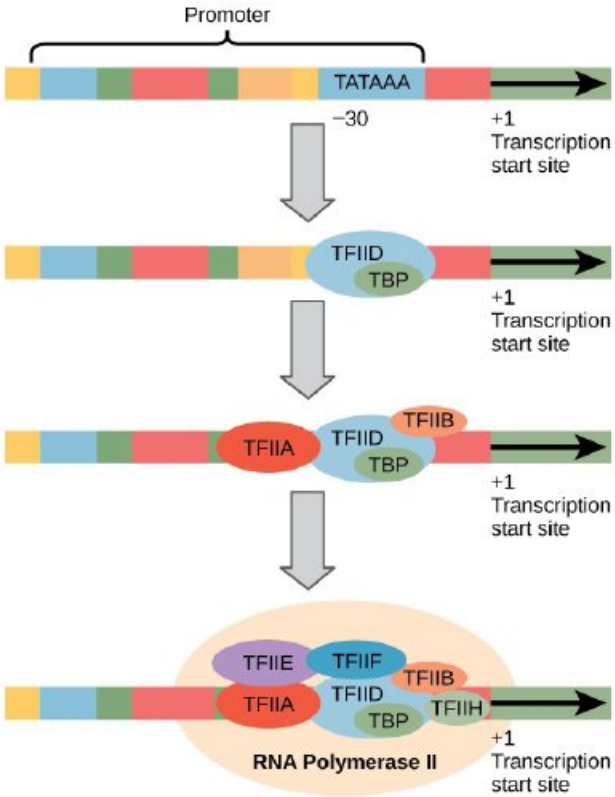


Fig.1.5: Eukaryotic Transcription Initiation: A generalized promoter of a gene transcribed by RNA polymerase II is shown. Transcription factors recognize the promoter, RNA polymerase II then binds and forms the transcription initiation complex.

Source:15.6: Eukaryotic Transcription - Initiation of Transcription in Eukaryotes - Biology LibreTexts

RNA Polymerases

Eukaryotes have three main types of RNA polymerases, each responsible for transcribing different classes of genes:

- **RNA Polymerase I:** Transcribes ribosomal RNA (rRNA) genes. Itis located in the nucleolus and transcribes the 28S, 18S, and 5.8S rRNA genes.
- **RNA Polymerase II:** Transcribes messenger RNA (mRNA) and some small nuclear RNAs (snRNAs). It is located in the nucleoplasm and transcribes protein-coding genes, to yield pre-mRNA, and also the genes encoding small nucleolar RNAs (snoRNAs) involved in rRNA processing and small nuclear RNAs (snRNAs) involved in mRNA processing, except for U6 snRNA.
- **RNA Polymerase III:** Transcribes transfer RNA (tRNA) genes and other small RNAs. I is also located in the nucleoplasm. It transcribes the genes for tRNA, 5S rRNA, U6 snRNA, and

the 7S RNA associated with the signal recognition particle (SRP) involved in the translocation of proteins across the endoplasmic reticulum membrane.

- Three eukaryotic RNA polymerases are huge and complicated, with 12 or more subunits each.
- Some eukaryotic enzyme subunits share DNA sequence similarities with the core enzyme of E. coli RNA polymerase.
- Four to seven subunits of each eukaryotic RNA polymerase are distinct, with no similarities to bacterial or other eukaryotic RNA polymerase subunits.

Promoters and Transcription Factors

Eukaryotic promoters are more complex than prokaryotic promoters, containing several elements:

- **Core Promoter**: Includes the TATA box (located around -25 to -30 nucleotides upstream of the transcription start site) and the initiator (Inr) sequence.
- **Proximal Promoter Elements**: Include the CAAT box and GC-rich regions.
- **Enhancers and Silencers**: Distal regulatory elements that can significantly enhance or repress transcription from a distance.

Transcription factors are proteins that bind to specific DNA sequences to regulate transcription. They can be divided into general transcription factors, which are required for the assembly of the transcriptional machinery, and specific transcription factors, which regulate gene expression in response to various signals.

Steps of Transcription

1. **Initiation**:

- **Pre-Initiation Complex Formation**: General transcription factors (such as TFIID, which includes the TATA-binding protein) and RNA polymerase II assemble at the core promoter to form the pre-initiation complex.
- **DNA Unwinding**: Helicase activity unwinds the DNA to form the transcription bubble.

2. **Elongation**:

- **RNA Synthesis**: RNA polymerase II moves along the DNA template strand, synthesizing the RNA transcript in the 5' to 3' direction.

3. **Termination**:

- **Polyadenylation Signal**: Transcription continues until RNA polymerase II transcribes a polyadenylation signal (AAUAAA) in the RNA. This signal is recognized, and the RNA is cleaved and polyadenylated.

RNA Processing

In eukaryotes, the primary RNA transcript (pre-mRNA) undergoes several processing steps before becoming mature mRNA capable of translation.

Capping

- **5' Cap Addition**: A 7-methylguanosine cap is added to the 5' end of the pre-mRNA. This cap protects the RNA from degradation and is important for ribosome binding during translation.

Splicing

- **Intron Removal**: Non-coding regions (introns) are removed from the pre-mRNA by the spliceosome, a complex of small nuclear RNAs (snRNAs) and proteins.
- **Exon Joining**: The coding regions (exons) are joined together to form a continuous coding sequence.

Polyadenylation

- **Poly-A Tail Addition**: A string of adenine nucleotides (poly-A tail) is added to the 3' end of the pre-mRNA. This tail enhances the stability of the RNA and facilitates its export from the nucleus.

Translation in Eukaryotes

Translation is the process by which the mRNA sequence is decoded to synthesize a polypeptide. This process occurs in the cytoplasm, where ribosomes, transfer RNAs (tRNAs), and various translation factors work together.

Steps of Translation

1. **Initiation**:

- **mRNA Binding**: The small ribosomal subunit, along with initiation factors, binds to the 5' cap of the mRNA and scans for the start codon (AUG).
- **tRNA Binding**: The initiator tRNA, carrying methionine, binds to the start codon.
- **Ribosome Assembly**: The large ribosomal subunit joins the complex, forming a functional ribosome.

2. **Elongation**:

- **Codon Recognition**: tRNAs with complementary anticodons bring the corresponding amino acids to the ribosome.
- **Peptide Bond Formation**: Peptidyl transferase activity of the ribosome forms peptide bonds between amino acids, elongating the polypeptide chain.
- **Translocation**: The ribosome moves along the mRNA, shifting the tRNA from the A site to the P site, and the empty tRNA exits from the E site.

3. **Termination**:

- **Stop Codon Recognition**: When the ribosome encounters a stop codon (UAA, UAG, or UGA), release factors bind to the ribosome, promoting the release of the completed polypeptide.
- **Ribosomal Disassembly**: The ribosomal subunits dissociate, ready to initiate another round of translation.

Post-Translational Modifications

Newly synthesized polypeptides often undergo various modifications to become fully functional proteins. These modifications can affect protein folding, stability, localization, and activity.

Types of Post-Translational Modifications

1. **Phosphorylation**:

- **Addition of Phosphate Groups**: Kinases add phosphate groups to specific amino acids (serine, threonine, or tyrosine), regulating protein activity and signaling pathways.

2. **Glycosylation**:

- **Addition of Carbohydrates**: Carbohydrates are added to proteins, affecting their stability, localization, and interactions.

3. **Acetylation**:

- **Addition of Acetyl Groups**: Acetyl groups are added to lysine residues, often regulating gene expression by modifying histones.

4. **Ubiquitination**:

- **Addition of Ubiquitin**: Ubiquitin molecules are attached to proteins, targeting them for degradation by the proteasome.

5. **Proteolytic Cleavage**:

- **Protein Processing**: Some proteins are synthesized as inactive precursors and are activated by cleavage of specific peptide bonds.

Regulation of Gene Expression

Gene expression in eukaryotes is tightly regulated at multiple levels to ensure precise control over cellular functions. Regulation can occur at the transcriptional, post-transcriptional, translational, and post-translational levels.

Transcriptional Regulation

1. **Transcription Factors**:

- **Activators and Repressors**: Proteins that bind to specific DNA sequences (enhancers and silencers) to increase or decrease transcription.

2. **Epigenetic Modifications**:

- **DNA Methylation**: Addition of methyl groups to cytosine residues, typically repressing gene expression.
- **Histone Modifications**: Chemical modifications of histones (e.g., acetylation, methylation) affecting chromatin structure and gene accessibility.

3. **Chromatin Remodeling**:

- **Nucleosome Positioning**: ATP-dependent complexes reposition nucleosomes, altering access to DNA.

Post-Transcriptional Regulation

1. **Alternative Splicing**:

- **Generating Diversity**: Different combinations of exons can be spliced together, producing multiple mRNA variants from a single gene.

2. **RNA Editing**:

- **Base Modification**: Enzymes alter nucleotide sequences in RNA, potentially changing the encoded protein.

3. **mRNA Stability**:

- **Regulatory Elements**: Sequences in the mRNA (e.g., AU-rich elements) influence its degradation rate.

4. **RNA Interference (RNAi)**:

- **siRNAs and miRNAs**: Small RNAs that bind to mRNA, leading to its degradation or translational repression.

Translational Regulation

1. **Ribosome Binding**:

- **Regulatory Proteins**: Proteins that bind to the mRNA and affect its translation efficiency.

2. **Internal Ribosome Entry Sites (IRES)**:

- **Alternative Initiation**: Some mRNAs contain IRES elements that allow translation initiation independent of the 5' cap.

Post-Translational Regulation

1. **Protein Modifications**:

- **Chemical Modifications**: Affect protein function, localization, and interactions.

2. **Protein Degradation**:

- **Ubiquitin-Proteasome System**: Targeted degradation of proteins by the proteasome.

1.6. Summary

Molecular biology is a dynamic and rapidly expanding field that gives critical insights into the workings of life. Understanding the molecular basis of biological processes enables scientists to develop new solutions to complex problems in health, agriculture, and industry. The incorporation of molecular biology into other domains continues to increase our knowledge and capabilities, resulting in improvements that improve people's quality of life. All biological activities are dependent on DNA, which acts as the blueprint for life and permits genetic information to be handed down through generations. Its discovery and comprehension changed biology and medicine, propelling discoveries in genetics, biotechnology, and a range of other fields. The Central Dogma of Molecular Biology emphasizes the directing flow of genetic information, focusing on the fundamental processes that allow genes to be expressed as functional proteins. Gene regulation is a complicated and dynamic process that allows cells to accurately control the expression of their genetic material, ensuring proper responses to internal and external stimuli while also preserving cellular function and stability. Transcriptional regulation regulates gene expression, allowing cells to adapt, differentiate, and control their growth and development. Lactose availability regulates the lac operon in E. coli. Lactose deactivates a repressor protein, allowing genes involved in lactose metabolism to be expressed. In eukaryotes, transcriptional regulation is vital in development because it determines whether specific genes are turned on or off in different tissues or stages. Gel electrophoresis is a fundamental technique in molecular biology, genetics, and biochemistry. Its uses include DNA fingerprinting, RNA analysis, protein purification, and genetic issue detection. Its ability to distinguish complex macromolecule combinations makes it a vital tool for scientific research and clinical diagnostics.

1.7. Terminal questions

Q.1. Discuss the history and scope of basic molecular biology.

Answer: ----------

Q.2. What are genetics? Discuss the structure and functions of gene

Answer: ----------

Q.3. What do you understand by gene expression and regulation?

Answer: ----------

Q.4. Discuss the gene expression in prokaryotes.

Answer: ----------

Q.5. Discuss the process and mechanism of Gene expression in eukaryotes.

Answer: ----------

Q.6. Write the short notes on

a) Translation

b) transcription

c) DNA sequencing

1.8. Further suggested readings

1. Robert Schlei, Genetics and Molecular Biology,2nd Edition
2. McGraw-Hill, Cell and Molecular Biology, Human Genetics: Concepts and Application, 9th Edition.
3. Desmond S. T. Nicholl, An Introduction to Genetic Engineering Third Edition, University of the West of Scotland, Paisley, UK

4. Robert Schleif, Genetics and Molecular Biology, second edition, The Johns Hopkins University Baltimore, Maryland

5. NPTEL – Bio Technology – Genetic Engineering & Applications

6. T.A. Brown, Gene Cloning And Dna Analysis An Introduction, Sixth Edition, University Of Manchester Manchester.

Unit-2: Working with nucleic acids

Contents

2.1. Introduction

Nucleic acids are large biomolecules that are crucial in all cells and Viruses. They are composed of nucleotide, which are the monomer components: a 5 carbon sugar, a phosphate, group and a nitrogenous base. The two main classes of nucleic acids are deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). If the sugar is ribose, the polymer is RNA, if the sugar is deoxyribose; a variant of ribose, the polymer is DNA. Nucleic acids are chemical compounds that are found in nature. They carry information in the cells and makeup genetic materials. These acids are very common in all living things. Where they create, encode, and store information in every living cell of every life form on earth. In term they send and express that information inside and outside the cell nucleolus.

String of nucleotides are bonded to form spiralling, backbones and assembled into chains of bases or base pairs selected from the five primary or canonical nucleobases.

The isolation of DNA and RNA is critical in molecular biology because it allows for the study of genetic material for a variety of applications including cloning, sequencing, gene expression studies, and diagnostics. Despite similarities in their isolation methods, distinct techniques are used to address

the unique features of each form of nucleic acid. This handbook describes the key procedures and methods for separating DNA and RNA. Cell lysis, protein and other contamination removal, and DNA precipitation are all steps in the DNA isolation process. The goal is to produce pure, high-molecular-weight DNA suitable for subsequent uses. DNA isolation techniques include phenol-chloroform extraction, silica column-based approaches, and magnetic bead-based procedures.To avoid RNase degradation, RNA isolation must be done carefully. Cells are lysed, RNases are inactivated, proteins and other impurities are removed, and RNA is precipitated.The separation of DNA and RNA is critical for many molecular biology and genetic research applications. While both processes involve common phases such as cell lysis and precipitation, the methodologies differ to account for the unique characteristics of DNA and RNA. Technological advancements have streamlined these operations, making them more efficient and accessible. The proper execution of these protocols ensures the extraction of high-quality nucleic acids appropriate for a variety of scientific studies.The procedures for cell lysis and precipitation differ to accommodate for the unique characteristics of DNA and RNA. Technological advancements have streamlined these operations, making them more efficient and accessible. The proper execution of these protocols ensures the extraction of high-quality nucleic acids appropriate for a variety of scientific studies.End labeling is a technique for attaching radioactive isotopes to the ends of DNA or RNA molecules, making them easier to identify and analyze in a variety of molecular biology applications such as hybridization tests, sequencing, and probing.Nucleic acid hybridization is a fundamental molecular biology technique that relies on the fact that complementary strands of nucleic acids (DNA or RNA) can establish specific base-pair interactions. This method is used in a variety of applications, including nucleic acid sequence detection, identification, and characterization in research and diagnostics.Nucleic acid hybridization is the process of annealing single-stranded nucleic acids (probes) to their complementary sequences (targets). The sequence of bases in RNA determines hybridization selectivity. A bonds with T (or U in RNA), and C bonds with G via hydrogen bonding.

Objectives

After reading this unit, the learner will able to know

- The Isolation of DNA and RNA and its
- the radiolabelling of nucleic acids through end labelling, nick translation, and by primer extension
- the nucleic acid hybridization and its significance

- the DNA sequencing methods through Maxam-Gilbert and Sanger-Coulson sequencing

2.2. Isolation of DNA and RNA

The separation of DNA and RNA is an important step in molecular biology and genetic studies. The nucleic acids are extracted from cells or tissues to investigate their structure, function, and expression. Because DNA and RNA have different chemical and structural properties, the methods for isolating them differ. Friedrich Miescher isolated an acidic substance known as nuclein, now known as nucleic acid from the nuclear material of salmon sperm in 1989.

The ability of organisms to pass genetic information from one generation to the next is made possible by nucleic acids, which are molecules that store information for cellular growth and reproduction. Jones established in 1920 that there are two different kinds of nucleic acids: ribonucleic acid (often referred to as RNA) and deoxyribonucleic acid (also known as DNA).

2.2.1. DNA isolation

DNA isolation is routine procedure to collect DNA for subsequent molecular analysis. Isolation of genomic DNA involves extraction of proteins by organic solvents followed by alcohol precipitation. Isolation of DNA is a crucial method in molecular biology that allows researchers to remove and purify DNA from cells or tissues for use in downstream applications such as cloning, sequencing, and PCR. Cell lysis, elimination of proteins and other impurities, and DNA precipitation are all common steps in the process. Here are the specific procedures and typical methods for DNA isolation:The first and most important stage in swotting a specific DNA chain is DNA isolation. This vital technique is used to break down a complicated DNA population and genome structure. In a snippet, the process entails extracting a component from the cortex and purifying the nucleic acid. Surprisingly, the crude metric has several origins. As previously said, DNA isolation can be performed on any organism, including hair, blood, nails, sperm, bloodstains, urine, saliva, bones, nails, and so on.

DNA isolation processes and purposes differ. The technique utilized to isolate DNA will be determined by numerous criteria, including sample size, age, and source. However, despite the differences in approach, the isolation process stays consistent. In the medical field, biotechnology is used to diagnose important concerns with life-threatening diseases such as cancer. Scientists are significant DNA hunters, looking to identify diverse species and understand how they are managed.

The DNA isolation steps involve five major steps. For extraction of DNA there are various steps and methods involved.

- ❖ Preparation of a cell extract.
- ❖ DNA purification from cell extract.
- ❖ The concentration of the DNA sample.
- ❖ Measuring the DNA purity.
- ❖ Monitoring the DNA concentration.

1. Preparation of a cell extract:

To extract DNA from a tissue/cells of interest, the cells must be separated and the cell membranes broken. The "Extraction buffer" assists in carrying out these procedures.The extraction buffer contains chemicals such as EDTA (Ethylene Diamine Tetra Acetate), which removes Mg^{2+} ions that are necessary for maintaining the overall structure of the cell membrane, and SDS (Sodium Dodecyl Sulfate), which aids in disrupting the cell membranes by removing the lipids. After lysing the cells, the final step in preparing a cell extract is to remove insoluble cell debris. Centrifugation can be used to pellet cell debris and partially digested organelles, resulting in a reasonably clear supernatant.

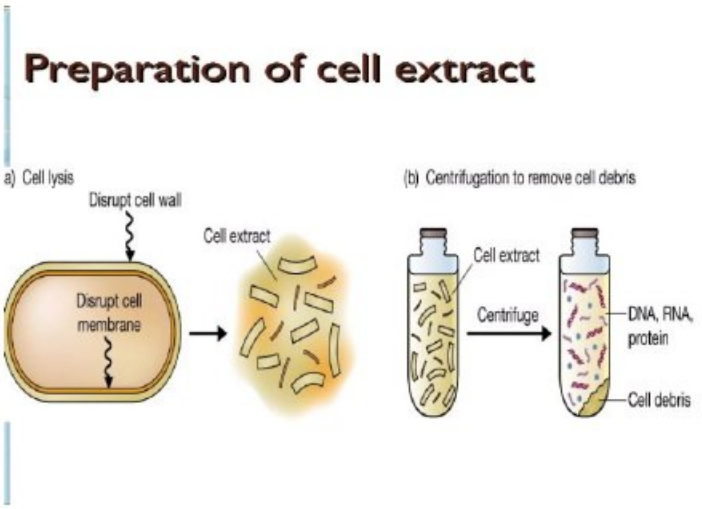
There are two types of techniques for breaking open bacterial cells: physical methods, which rupture the cells with mechanical pressures, and chemical approaches, which cause cell lysis by exposing them to chemical substances that damage the integrity of the cell walls. Chemical approaches are most typically utilized with bacterial cells when the purpose is to prepare DNA.

LYSIS:

In the lysis refers to the shattering of the cell wall and cellular membranes (in particular, the plasma and nuclear membranes).

- Mechanical force disrupts the cell wall.
- Chemical lysis often consists of one chemical attacking the cell wall and another damaging the cell membrane.

- For cell wall lysis: lysozyme, ethylenediaminetetraacetate (EDTA), or a mixture of both.
- For cell membrane lysis: sodium dodecyl sulfate (SDS).



1. Lysozyme: digests the polymeric compounds that give the cell wall its rigidity

2. EDTA: remove magnesium, that are essential for preserving the overall structure of the cell envelope, and inhibits cellular DNase

3. SDS: removes lipid molecules and thereby cause disruption of the cell membranes

After lysing the cells, the final stage in preparing a cell extract is the removal of insoluble cell debris. Centrifugation can be used to pellet components like partially digested cell wall fractions, leaving the cell extract in a reasonably clear supernatant.

2. DNA purification from cell extract:

Purification procedures can eliminate impurities and yield pure DNA. Solvent extraction is a regularly used method for removing impurities from nucleic acids. A mixture of phenol and chloroform is commonly used to remove proteins. Precipitation with ethanol or isopropanol is commonly used to concentrate nucleic acids. To improve precipitation efficiency in a product mixture, an inert carrier, such as glycogen, can be introduced when the target acid nuclei amount is low. A variety of procedures can be used to remove these contaminants, leaving the DNA in a pure form. The most commonly used procedures are.

- Ethanol precipitation:
- Phenol-chloroform extraction
- Minicolumn purification.

a. Ethanol precipitation:

- Ethanol precipitation usually is done by ice-cold ethanol or isopropanol.
- Since DNA is insoluble in these alcohols, it will aggregate together, giving a pellet upon centrifugation
- Precipitation of DNA is improved by increasing the ionic strength, usually by adding sodium acetate

b. Phenol-chloroform extraction:

Phenol-chloroform extraction denatures proteins in the sample. • After centrifugation, denaturated proteins remain in the organic phase. Aqueous phase with nucleic acid is combined with chloroform to eliminate phenol residues from solution.

c. Minicolumn purification.

Minicolumn purification is based on the fact that nucleic acids can adsorb to the solid phase salt concentration of the buffer (ca or other) dependent on the pH and salt concentration of the buffer.

3. The concentration of DNA samples:

The most commonly utilized method of concentration is ethanol precipitation. Absolute ethanol will efficiently precipitate polymeric nucleic acids in the presence of salt at -20°C or lower temperatures. With a concentrated solution of DNA, one can use a glass rod to draw out the adhering DNA strands, but for dilute solutions precipitated DNA can be collected by centrifugation and redissolving in an adequate proportion of water.

4. Measuring the DNA purity:

To determine DNA purity, measure absorbance between 230nm and 320nm to look for additional impurities. The most common purity calculation is the ratio of absorbance at 260nm to reading at 280nm. Good-quality DNA will have an A260/A280 ratio of 1.7-2.0. A reading of 1.6 does not make the DNA inappropriate for any purpose, but lower ratios indicate that additional contaminants are present. The ratio can be determined after adjusting for turbidity (absorbance at 320nm).

DNA purity (A260/A280) = (A260 reading – A320 reading) ÷ (A280 reading – A320 reading)

Strong absorbance at 230nm may indicate the presence of chemical components or chaotropic salts in pure DNA. A 260nm to 230nm ratio can assist determine the amount of salt (A chaotropic agent is a substance which disrupts the structure of and denature, macromolecules such as proteins and nucleic acids e.g. DNA and RNA) carryover in pure DNA. For example, as the ratio decreases, the amount of thiocyanate salt increases.

5. Monitoring the DNA concentration:

Many molecular biology applications rely on accurate DNA concentration measurements. Spectrophotometry and fluorescence are routinely employed to determine genomic and plasmid DNA concentrations. The DNA sample is measured with a fluorometer, and nucleic acid concentrations are determined by comparing the sample's fluorescence emission to a fluorescence curve constructed using established nucleic acid concentration standards.

DNA concentration can be determined by measuring the absorbance at 260 nm (A260) in a spectrophotometer using a quartz cuvette. (A cuvett for spectrophotometric measurements is small, clear rectangular vessel). For greatest accuracy, readings should be between 0.1 and 1.0. An absorbance of 1 unit at 260 nm corresponds to 50 μg genomic DNA per ml (A260 =1 for 50 μg/ml; based on a standard 1 cm path length.

2.2.2. Isolation of RNA

Total RNA isolation is the method that helps to separate pure RNA form tissues and the mixtures of DNA or proteins. Isolating RNA is a fundamental method in molecular biology and industry, allowing researchers to study gene expression, sequence, and perform other RNA-based investigations. RNA extraction is a process to isolate various or a specific type of RNA molecules for gene expression studies." RNA extraction or isolation is the process of separating any sort of RNA from a cell. It is useful for studying gene expression, as well as identifying and classifying pathogens. It is accurate, reproducible, and adaptable since it uses nucleic acids for detection. The procedure includes lysing cells, preserving the RNA from degradation, eliminating proteins and other impurities, and purifying the RNA. Given RNA's sensitivity to destruction by ubiquitous RNases, additional efforts must be taken to maintain its integrity.Obtaining high-quality RNA is the first, and often most important, step in many molecular techniques, including reverse transcription real-time PCR (RT-

qPCR), transcriptome (transcriptomics is the study of all RNA molecules in one cell or population of cells. Hence, while genomics studies the changes in the DNA, which can be seen as the causative mechanism, transcriptomics usually aims at observing the consequences that these changes have at an RNA level. Analysis with next-generation sequencing, array analysis, digital PCR, northern analysis, and cDNA library construction. To provide the most sensitive and biologically meaningful results, the RNA isolation technique must comprise several critical phases preceding, during, and following the actual RNA purification.

RNA extraction-based RT-PCR detection methods were not very popular. The RNA extraction procedure is time-consuming and prone to contamination. Chomczynski and Sachhi (1987) proposed the first single-step RNA extraction method. There have been thousands of published procedures for plant RNA extraction. However, the majority of the described techniques rely on three methods:

Phenol-based method: Phenol-based protocols are most commonly utilized for plants with high secondary metabolite concentrations. Phenol is commonly used with chloroform to eliminate polysaccharide impurities.

Trizol-based method: Trizol has been used to extract RNA from some tissues, such as Arabidopsis seeds. Trizol is a commercially available reagent that mixes phenol and guanidine isothiocyanate in a single solution to produce high yields of RNA during a fast isolation technique. This approach removes pollutants by combining trizol with chloroform, isopropyl alcohol, and ethanol in RNase-free water.

CTAB-based method: The cetyltrimethylammonium bromide, or CTAB-based approach, has been employed in plant samples with high polysaccharide, secondary product, or RNase levels. The CTAB extraction buffer is made up of 2% cetyltrimethylammonium bromide, 1% polyvinylpyrrolidone (PVP), 100mM Tris-HCl, 1.4M NaCl, and 20mM EDTA. This process begins by eliminating carbohydrates, next proteins, and ultimately secondary metabolites. This procedure can also contain sodium dodecyl sulfate (SDS) at a concentration of around 0.5%. SDS is a detergent that forms compounds with protein, aiding in the removal of protein impurities during extraction.

The Main Steps in plant RNA extractioncan be split into four step

Homogenization: Tissue can be purchased fresh or frozen. Furthermore, when suitable lab conditions are not available (for example, when harvesting from the field), materials can be placed in a

mortar and pounded into a fine powder with a pestle before adding to the buffer extraction solution. In other cases, researchers have access to a bead-based homogenizer, which can help pulverize the tissue into powder.Isolation: Once the tissue has been finely powdered, the buffer extraction is added. In this instance, any phenol, trizol, or CTAB buffer extraction solution may be utilized. The goal of this step is to lyse the cells and release RNA.

However, this is a vital stage for contamination, because when the membranes are ruptured, all of the cell contents are discharged. However, this is a vital phase in the contamination process because when the membranes are ruptured, all of the cell contents are discharged at once. This indicates that the RNA has been mixed with all other cell components, allowing metabolites such as polysaccharides and polyphenols to have closer contact with the RNA. Simultaneously, RNAses begin to work, increasing the danger of RNA breakdown. Several of the extraction solutions operate as RNAse inhibitors, minimizing the risk of RNA damage.

Clearing: In this stage, several complimentary reagents are utilized to eliminate impurities such as polysaccharides, proteins, secondary metabolites, and other potential pollutants. Polysaccharides/secondary metabolites and proteins are removed using reagents such as chloroform and isopropyl alcohol, respectively. Furthermore, contaminating genomic DNA can be eliminated using column-based or enzyme-based techniques.

Precipitation: In order to prepare the RNA for use in subsequent processes, this step involves further removing any extra impurities from the final solution and purifying and cleaning it. Here, the extracted RNA can be preserved using 70% ethanol, water treated with 0.1% (v/v) diethyl pyrocarbonate (DEPC), or water free of RNase and DNase. Finally, RNA needs to be stored at -80°C after it has been dissolved in one of these chemicals.

For plant RNA extraction there are mainly solutions are used

1. Cell Lysis

The initial step involves breaking open the cell membrane to release the cellular contents. This is typically achieved using a lysis buffer containing:

- **Guanidiniumthiocyanate** or other chaotropic agents to denature proteins and inactivate RNases.

- **Detergents** (e.g., SDS or Triton X-100) to dissolve cell membranes.
- **Buffering agents** (e.g., Tris-HCl) to maintain pH stability.

2. Inhibition of RNases

RNases, which can rapidly degrade RNA, must be inhibited throughout the process. This can be achieved by:

- Using RNase inhibitors.
- Ensuring all solutions, tubes, and equipment are RNase-free (e.g., treated with DEPC).

3. Removal of Proteins and Other Contaminants

After cell lysis, proteins and other cellular debris must be removed. This can be done through enzymatic digestion or chemical extraction.

4. RNA Precipitation

RNA is separated from the solution by precipitation, commonly using alcohol (ethanol or isopropanol) in the presence of salts.

Quantification of a plant RNA isolation

One can calculate the μg RNA/mg tissue in order to estimate the RNA yield. The primary form of RNA found in the extraction, ribosomal RNA content, is used to assess the quality. As a result, rRNA estimate on a gel is used to indirectly measure mRNA quality.

An equipment used to confirm the quantity and quality (or purity) of the extracted RNA is a NanoDrop spectrophotometer. The ratio of absorbance at 230 nm, 260 nm, and 280 nm is used by researchers to evaluate the presence and purity of the extracted RNA.

If the RNA is pure, the 260/280 ratio should be approximately 2.

However, a significantly lower ratio indicates the presence of phenol, protein, or other pollutants that absorb heavily at or near 280 nm.

On the other hand, the 260/230 ratio is predicted to fall between 2 and 2.2. If this value is significantly lower, it may suggest the presence of pollutants (for example, carbohydrates, EDTA, guanidine isothiocyanate, and phenol absorb at 230 nm). Ratios lower than intended may necessitate extra cleaning.

The bioanalyzer is another tool for determining RNA amount and quality. In this technology, RNA samples are separated via electrophoretic separation on microfabricated chips

1. Cell Lysis

The initial step involves breaking open the cell membrane to release the cellular contents. This is typically achieved using a lysis buffer containing:

- **Guanidiniumthiocyanate** or other chaotropic agents to denature proteins and inactivate RNases.
- **Detergents** (e.g., SDS or Triton X-100) to dissolve cell membranes.
- **Buffering agents** (e.g., Tris-HCl) to maintain pH stability.

2. Inhibition of RNases

RNases, which can rapidly degrade RNA, must be inhibited throughout the process. This can be achieved by:

- Using RNase inhibitors.
- Ensuring all solutions, tubes, and equipment are RNase-free (e.g., treated with DEPC).

3. Removal of Proteins and Other Contaminants

After cell lysis, proteins and other cellular debris must be removed. This can be done through enzymatic digestion or chemical extraction.

4. RNA Precipitation

RNA is separated from the solution by precipitation, commonly using alcohol (ethanol or isopropanol) in the presence of salts.

2.3. Radiolabelling of nucleic acids- end labelling

Radioactive end labelling is useful for visualizing and allowing the detection of nucleic acids at trace concentration. Radioactive end-labelling can be carried out RNA, DNA, or other modified nucleic acids. Radio labelling typically detect specific radio labelled nucleotides enzymatically, DNA and RNA sequences are detected and authorized. It depending on the type of application, DNA and RNA sequences are labelled at 5' or 3' end. A few of the typical reasons for radiolabeling include:

1. To generate information on gene integrily and copy number (biot).
2. To diagnose specific sequences and chromosomal aberrations (*in-situ* hybridization)
3. To simultaneously measure the relative expression of RNAs (microarrays analysis)
4. To discover proteins nucleic acid interactions (electrophoretic mobility shift assays or FRET)

Keeping track of small amounts of nucleic acid is a common challenge in cloning methods. This issue is exacerbated at each stage of the process, as losses mean that the amount of material available decreases after each step. Labeling the nucleic acid with a marker can aid in tracing and identifying the material throughout the procedure. So what can be utilized for the label?

Types of label – radioactive

Radioactive tracers have been widely employed in biochemistry and molecular biology for many years, with well-established protocols. The most widely used isotopes are tritium (3H), carbon-14 (14C), sulfur-35 (35S), and phosphorus-32 (32P). Tritium and 14C are low-energy emitters, while 35S is a'medium' emitter and 32P is a high-energy one. Thus, 32P poses a greater risk than radioactive isotopes are commonly employed to label nucleic acids, however they pose more risks than non-radioactive labeling approaches.

The other isotopes require special attention when used. Radioactive waste is subject to severe legislative storage and disposal restrictions. To mitigate the risks associated with working with high-energy isotopes, other methods like fluorescent dyes and enzyme-linked labels have gained popularity in recent years. While both approaches have advantages for specific applications (e.g., DNA sequencing), radioactive labels remain the favored choice for routine tracing experiments. Radiolabelling is a common phrase used to describe this process.

To trace DNA and RNA samples, mark the nucleic acid with a radioactive molecule, such as a deoxynucleoside triphosphate (dNTP) labeled with 3H or 32P. The amount of nucleic acid present can then be determined using a scintillation counter. This is often done by calculation, taking into account the amount of radioactivity in the sample.Radiolabelling can also produce extremely radioactive nucleic acids for hybridization investigations. Such molecules are referred to as radioactive probes, and they are extremely useful for recognizing certain DNA or RNA sequence. Have a range of applications (see Sections 3.5 and 8.2). The primary distinction between labelling for tracing and labelling for probes is specific activity, which is a measure of how radioactive the molecule is. A low specific activity is sufficient for tracing, whereas a high specific activity is required for probes. In probe preparation, the radioactive label is often the high-energy ß-emitter 32P. The following section discusses common ways for labeling nucleic acid molecules.

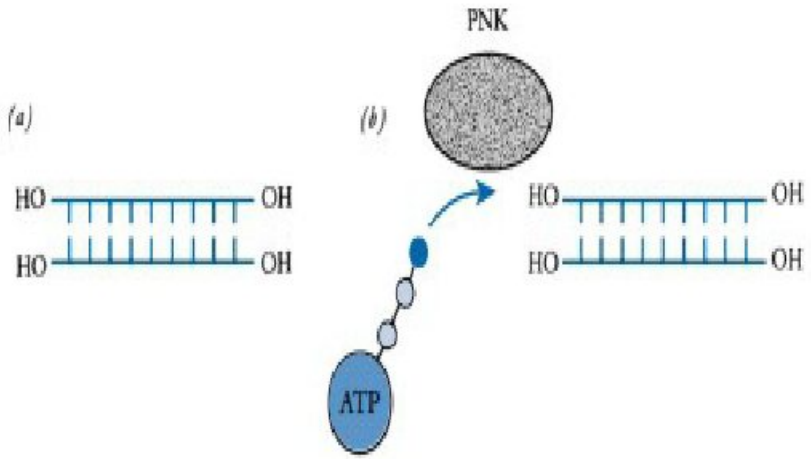


Fig 2.1. : End labelling DNA using polynucleotide kinase (PNK). (a) DNA is dephosphorylated using phosphatase to generate 5' -OH groups. (b) The terminal phosphate of [γ-32P]ATP (solid circle) is then transferred to the 5' terminus by PNK. The reaction can also occur as an exchange reaction with 5'-phosphate termini

Source: AnIntroductiontoGeneticEngineering.pdf

2.3.1. End labelling:

The end labeling procedure involves using polynucleotide kinase to deposit the terminal phosphate group of ATP onto the 5-hydroxyl termini of nucleic acid molecules. Radioactive labeling

of an ATP donor results in a nucleic acid with poor specific activity, as only the ends of the molecules become radioactive (Fig. 3.2).

Methods

5' End Labeling

5' end labeling involves attaching a label to the 5' end of the nucleic acid. This is typically accomplished using T4 polynucleotide kinase (PNK) and γ-32P ATP.

Steps:

1. **Dephosphorylation (if necessary)**: The 5' end of the nucleic acid may first be dephosphorylated using alkaline phosphatase to remove any existing phosphate groups.

2. **Phosphorylation**: T4 PNK transfers the γ-phosphate from γ-32P ATP to the 5' hydroxyl group of the nucleic acid.

3' End Labeling

3' end labeling involves adding a labeled nucleotide to the 3' end of the nucleic acid using terminal deoxynucleotidyltransferase (TdT).

Steps:

1. **Polymerization**: TdT adds α-32P-labeled deoxynucleotide triphosphates (dNTPs) to the 3' hydroxyl group of the nucleic acid.

Applications

- **Detection**: Radiolabeled nucleic acids can be detected using autoradiography or phosphorimaging, while fluorescently labeled nucleic acids can be detected using fluorescence microscopy or other fluorescence-based techniques.

- **Quantification**: The extent of hybridization or binding in assays can be quantified by measuring the intensity of the label.

- **Probing**: Labeled probes can hybridize to complementary sequences for gene mapping, expression studies, and mutation analysis.

Advantages

- **High Sensitivity**: Radioactive labels, particularly 32P, provide high sensitivity, enabling the detection of minute amounts of nucleic acids.

- **Specificity**: The labeled ends specifically hybridize to complementary sequences, allowing precise targeting in various applications.

Considerations

- **Safety**: Handling radioactive materials requires stringent safety protocols to protect researchers and the environment.

- **Stability**: Labeled nucleic acids should be used promptly as some labels, especially radioactive ones, can degrade over time.

End labeling is a powerful technique for tracking, detecting, and quantifying nucleic acids. Whether using radioactive or fluorescent labels, the method provides high sensitivity and specificity, making it essential for a wide range of molecular biology applications.

2.3.2.Nick translation:

The enzyme DNA polymerase I is responsible for translating nicks created in the phosphodiester backbone of the DNA double helix. Nicks can arise spontaneously or due to low concentrations of DNase I in the reaction mixture. DNA polymerase I catalyzes strand-replacement reactions that add new dNTPs to the DNA chain. Using radioactive dNTPs can result in a strongly labeled DNA molecule (Fig. 2.2).

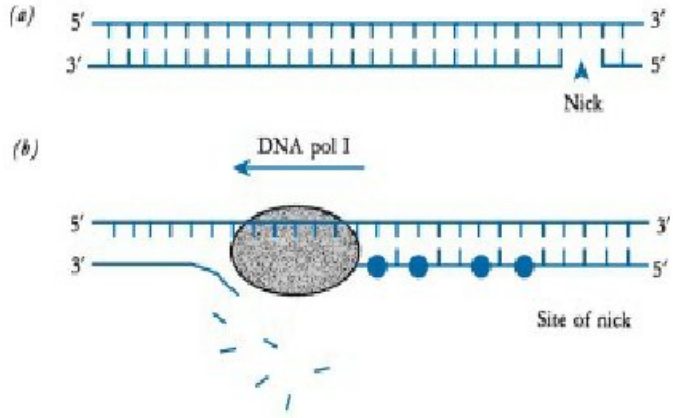
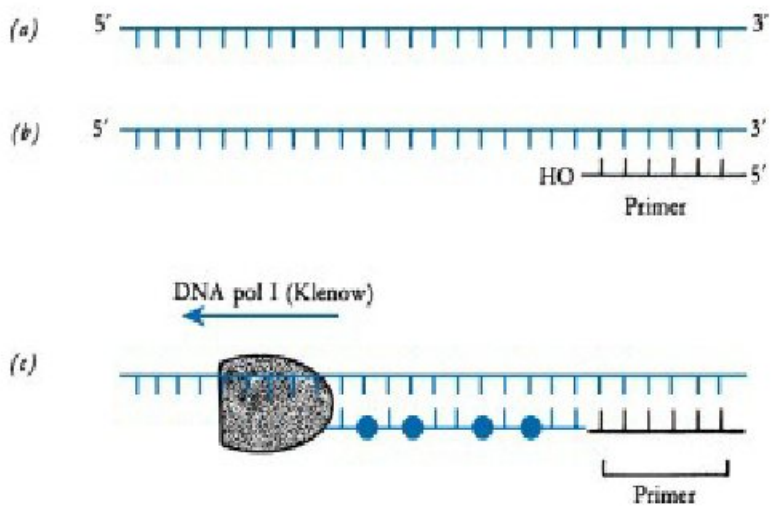


Fig.2.2: Labelling DNA by nick translation. (a) A single-strand nick is introduced into the phosphodiester backbone of a DNA fragment using DNase I. (b) DNA polymerase I then synthesises a copy of the template strand, degrading the non-template strand with its 5——————→3——————exomuclease activity. If [α32P]dNTP is supplied this will be incorporated into the newly synthesised strand

Source: AnIntroductiontoGeneticEngineering.pdf



Fg.2.3: Labelling DNA by primer extension (oligolabelling). (a) DNA is denatured to give single-stranded molecules. (b) An oligonucleotide primer is then added to give a short double-stranded region with a free 3-OH group. (c) The Klenow fragment of DNA polymerase I can then synthesise a copy of the template strand from the primer, incorporating [α32P]dNTP (solid circles) to produce a labelled molecule with a very high specific activity

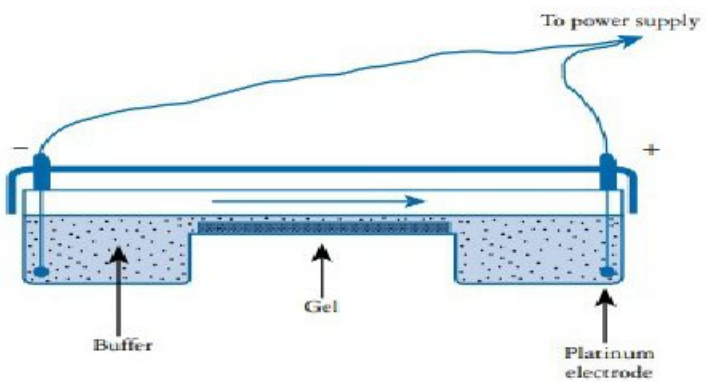
2.3.3. Labelling by primer extension

Labeling by primer extension is a technique that uses random oligonucleotides (often hexadeoxyribonucleotide sequences) to initiate DNA strand synthesis. In most labeling procedures, not all radioactive dNTP is incorporated into the target sequence. Non-incorporated isotopes are typically removed before employing the probe. By DNA polymerase. To label the DNA, heat it and anneal the oligonucleotide primers to the single-stranded DNA. The Klenow fragment of DNA polymerase (see Section 4.2.2) synthesizes a copy of the template, primed by the oligonucleotide's 3-hydroxyl group. Incorporating a labeleddNTP produces DNA with high specific activity (Fig. 3.4).

Radiolabelling reactions generally require separating labeled DNA from unincorporated nucleotides in the reaction mixture. A easy solution is to perform a small-scale gel filtration with a suitable media. Using a Pasteur pipette, the method involves first passing labeled DNA through the column, then unbound nucleotides. Radioactivity can be measured in fractions and used to compute total DNA activity, specific activity, and isotope incorporation percentage.

2.4. Nucleic acid hybridisation

Nucleic acid hybridization can provide insights into sequence complexity, as detailed in Section 2.6.1. The base-pairing link between complementary sequences has significant implications for both cell function and scientific exploration. This approach is highly sensitive and can detect individual DNA sequences even in complicated combinations. Typically, a single pure sequence is labeled with 32P and utilized as a probe.



The DNA to be probed is typically denatured and attached to a nitrocellulose or nylon membrane. Hybridization involves forming duplexes in a sealed plastic bag or tube at 65-68°C over

several hours. After washing off the extra probe, the degree of hybridization can be measured using a scintillation spectrometer or an autoradiograph (exposing the sample to X-rays).

2.5. DNA sequencing methods

In the mid-1970s, when molecular cloning techniques advanced, simple methods for determining DNA nucleotide sequence were created. These advancements paved the way for extensive investigation of the structure and function of several genes. To determine a DNA sequence, bases must be identified sequentially, allowing for the processing of each base individually. There are three major requirements to obtain this

- DNA fragments need to be prepared in a form suitable for sequencing
- The technique used must achieve the aim of presenting each base in turn in a form suitable for identification.
- The detection method must permit rapid and accurate identification of the bases

Creating and preparing DNA fragments is a basic technical process. The fragments are frequently cloned sequences provided for sequencing in an appropriate vector; with careful attention to detail, this can be accomplished very cheaply. The 'informatic problem' of determining the location of a segment within the genome is more challenging. Section 3.7.2 will discuss two techniques to addressing this issue.

The sequencing protocol is technical, not experimental. There are various variations on the basic procedure, however the most commonly used techniques involve the enzymatic method. The goal is to create overlapping segments that end at distinct bases and vary in length by one nucleotide, regardless of the strategy used. This is known as a collection of nested fragments.There are two primary ways for sequencing DNA. Allan Maxam and Walter Gilbert invented a method that uses chemicals to split DNA at certain places, resulting in fragments that differ by one nucleotide each. The second approach, discovered by Fred Sanger and Alan Coulson, involves enzymatic production of DNA strands that end in a changed nucleotide.

The enzyme method has largely superseded the chemical method as the preferred technique. However, chemical sequencing can still give useful data to confirm enzymatic results.

Fluorographic detection technologies can replace radioactive isotopes. This is especially essential in DNA sequencing because it speeds up the process and allows for automation. In Chapter 10, we shall discuss genome sequencing in more detail.

The difficulty of preparing a segment for sequencing depends on the project size. Sequencing a previously extracted and identified gene fragment requires an appropriate length and form for the sequencing technique.Sequencing technique is crucial in this scenario, with two options to consider. The first is an orderly sequencing strategy. The technique includes tracking fragments and noting their relative order throughout the project. The sequence follows the order of the fragments. The second option is shotgun sequencing, which generates and processes fragments randomly. The sequence is then assembled by searching for sequence overlaps on a computer. Figure 2.5 shows the two strategies.

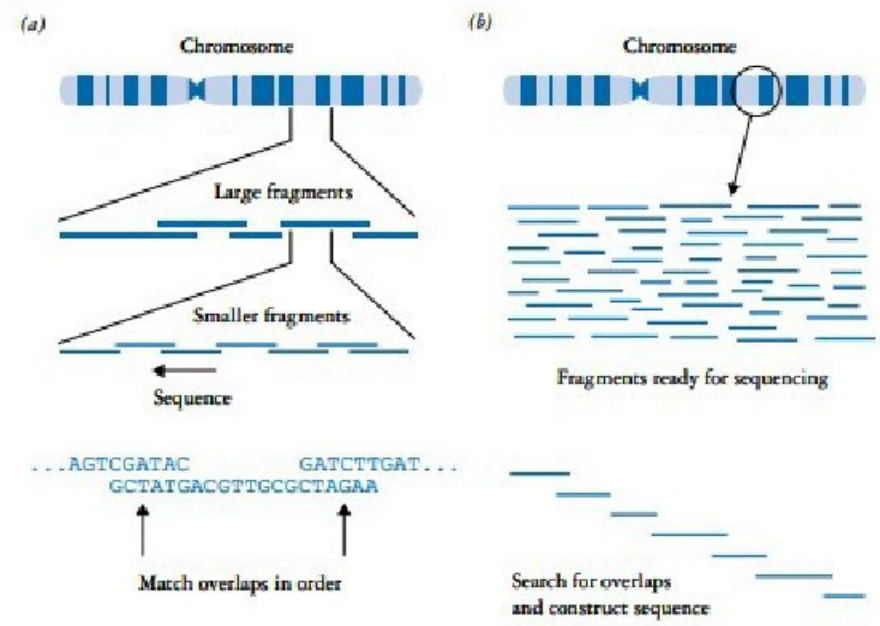


Fig.2.5: Two strategies for sequencing large stretches of DNA

2.6. Maxam-Gilbert sequencing:

The sequencing by chemicals degration also known as chemical sequencing was published in 1977 by Maxam and Gilbert requiring chemical modifications of the DNA and further cleavage and electrophoresis. The Maxam and Gilbert sequencing approach was a significant breakthrough and become the methodology of choice since it allowed the direct sequencing of purified DNA, without requiring previous in-vivo cloning ssDNA preparation steps.

Maxam-Gilbert sequencing is also known as chemical sequencing, in this sequencing the use radioactive labelingis use at one 5' end of the DNA fragment which is to be sequenced. This approach was developed after studying the interaction between Lac repressor and lac operator in vitro. This approach involves radiolabeling one end of DNA and partially cleaving it using five particular chemical reactions for each nucleotide. This results in five populations of radiolabelled molecules extending from a common point to the site of terminal cleavage. DNA bases control the length of each population, which is made up of a variety of molecules. Electrophoresis through polyacrylamide gels resolves these populations, and autoradiography detects the end-labelled molecules. This procedure remains unchanged from when it was developed.

Procedure:

The procedure involved the radioactive labelling of the 5' –p end of the double stranded DNA (dsDNA) with 32p-dATP using polynucleotide kinase. Then the DNA is denatured with DiMethyl sulfoxide (DMSO) at 90 °C and resulting ssDNA molecule are segregated via electrophoresis. Nitrogenous base-specific reations are carried out to modify the adenosine (A), Cytosine (C) Guanine (G) and Thymine (T) , residues, allowing the chemical cleavage of the ssDNA at the 5-p side of such positions. The A and T reactions also generate some minor G and C cleavage respectively, which should be taken into account (showing as weaker signal later on). Subsequent polyacrylamide gel electrophoresis and autoradiography allows the ssDNA fragment separation by size and the detection of the radiolabelled. DNA band pattern on a X-ray film that encodes the DNA sequence, from which the sequence may be referred.

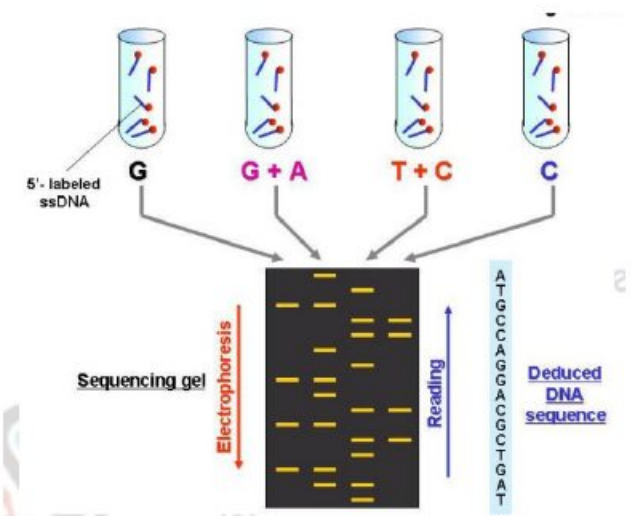


Fig.2.6: Maxam Gilbert Sequencing Method

Source: 1501755049Sequencing.pdf

The chemical degradation approach has one advantage over chain termination sequencing: the DNA sequence is retrieved directly from the original molecule, rather than an enzymatic copy. This technology allows for the sequencing of synthetic nucleotides, analysis of DNA changes like methylation, and research of secondary structure and protein-DNA interactions through chemical protection or interference tests.

2.7. Sanger's sequencing:

Maxam-Gilbert methodology was very popular at that time, it use hazardous chemicals and technically challenging, being also difficult to scale up. Thus it was eventually replaced by cleaner and more convenient alternative, namely the Sanger approach. Sanger sequencing is a method of DNA sequencing that involves electrophoresis is and is based. On the random incorporation of chain terminating dideoxynucleotides by DNA polymerase during *in-vitro* DNA replication. After first being developed by Fredrick Sanger and colleagues in 1977, it become the most widely used sequencing method for approximately 40 years. It was first commercialized by applied bio systems in 1986.

Frederick Sanger and his associates created the chain termination method of sequencing. The Sanger-Coulson process differs significantly from Maxam and Gilbert's, although producing similar results. In this situation, DNA polymerase's Klenow fragment creates a copy of the DNA to be sequenced. In this reaction, single-stranded DNA serves as the template. A primer is required to give the 3' terminal for DNA polymerase to synthesize the copy (see Fig. 3.9). To produce nested fragments, a modified dNTP is added to each reaction. These dNTPs lack the required hydroxyl group at the 3' position of deoxyribose for chain elongation. These modified dNTPs are called dideoxynucleoside triphosphates (ddNTPs).

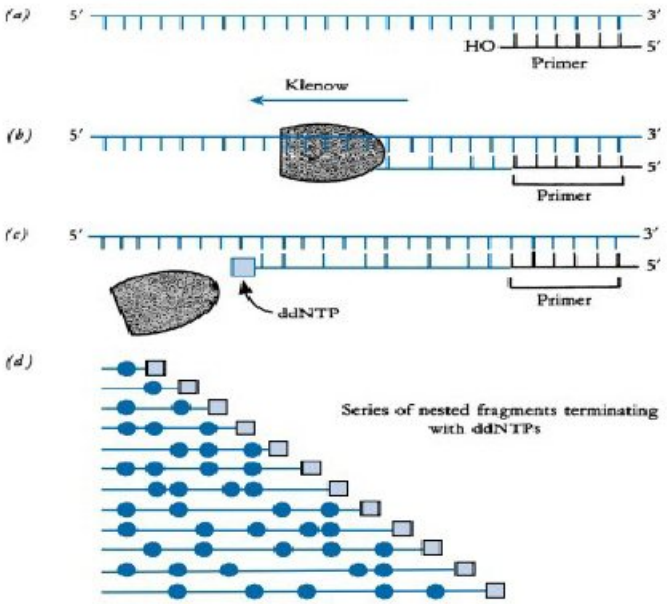


Fig.2.7: DNA sequencing using the dideoxy chain termination (Sanger–Coulson) method.

The four ddNTPs (A, G, T, and C forms) are used in four reactions, each with the four regular dNTPs. The dideoxy form has a low concentration and is rarely absorbed into the developing DNA chain. Each reaction generates fragments that end at a given nucleotide, resulting in nested fragments from all four reactions. To identify the DNA chain, include a radioactive dNTP in the reaction mixture. Using [α-35S]dATP allows for greater sequence reading from a single gel compared to 32P-labeled dNTPs previously used.Dideoxy sequencing requires more subcloning into multiple vectors to generate fragments, compared to chemical sequencing. Several plasmid vectors are now available (see to Section 5.2), and some are suitable for DNA sequencing investigations. Cloning the DNA into a vector, such as the bacteriophage M13 (see Section 5.3.3), can produce single-stranded DNA during infection. This provides an appropriate substrate for sequencing processes.

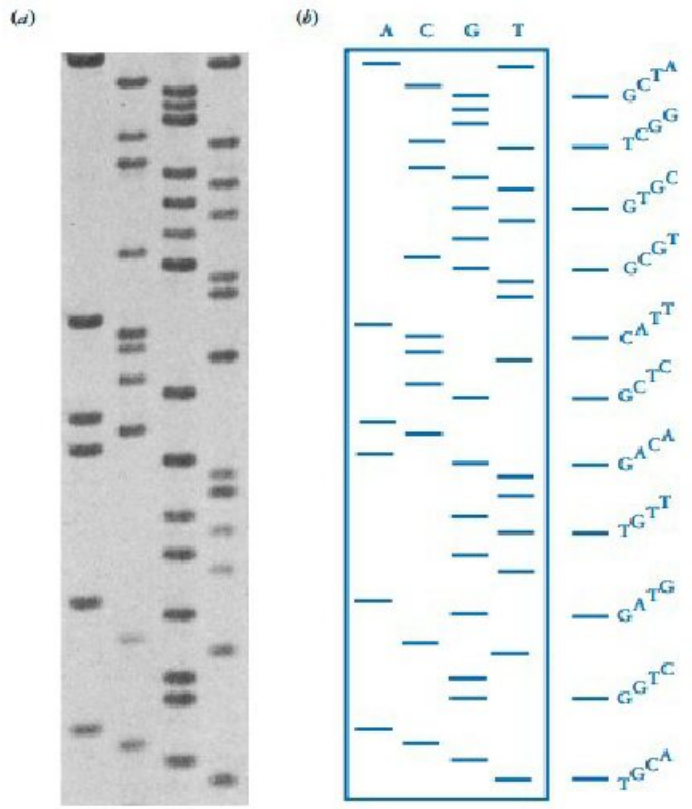


Fig. 2.8: Reading a DNA sequence. (a) An autoradiograph of part of a sequencing gel, and (b) a tracing of the autoradiograph.

2.8. Summary

The isolation of DNA is an important step in molecular biology research. The method chosen is determined by the sample type, purity requirements, and downstream applications. Phenol-chloroform extraction, silica column-based procedures, and magnetic bead-based methods are all commonly employed, each with merits and disadvantages. Proper execution of these protocols results in the extraction of high-quality DNA suitable for a variety of molecular biology techniques.Isolating RNA is an important step in molecular biology research because it allows us to analyze gene expression and other RNA-related processes. The method chosen is determined by the sample type, purity requirements, and downstream applications. Phenol-chloroform extraction, silica column-based procedures, and magnetic bead-based methods are all commonly utilized, with each having its own set of advantages and disadvantages. By adhering to established methods and taking steps to prevent RNA degradation, researchers can get high-quality RNA suitable for numerous analysis and experiments.

End labelling is a technique for attaching radioactive isotopes to the ends of DNA or RNA molecules, making them easier to identify and analyze in a variety of molecular biology applications such as hybridization tests, sequencing, and probing.Radiolabeled nucleic acids can be detected via autoradiography or phosphor imaging. Labeled probes can hybridize with complementary sequences to perform gene mapping, expression investigations, and mutation research. End labeling is an effective approach for tracking and quantifying nucleic acids in research and diagnostic applications, with excellent sensitivity and specificity due to the radioactive signal.

DNA sequencing identifies the nucleotide sequence of a DNA sample. There are two traditional ways of sequencing: Sanger sequencing or chain termination, and Maxam-Gilbert sequencing via chemical cleavage.Sanger sequencing is still considered the "gold standard" because to its high accuracy. Automated Sanger sequencing still generates the longest reads. However, it is slow and costly compared to second and third generation technologies, which are faster and more cost-effective. Accuracy and quick readings are major issues with these technologies, making final assembly and alignment laborious and computationally challenging. The method chosen depends on the application and advantages/disadvantages of the sequencing methodology.

2.9. Terminal questions

Q.1. What is the nucleic acids? Discuss the structure and functions of nucleic acids.

Answer: ----------

Q.2. Define DNA and how it isolate form plant cell.

Answer: ----------

Q.3. Discuss the process and purification of RNA isolation.

Answer: ----------

Q.4. Discuss the Radiolabelling of nucleic acids- end labelling.

Answer: ----------

Q.5. Write the principle of nucleic acid hybridization.

Answer: ----------

Q.6. What do you understand about DNA sequencing? Different methods of DNA sequencing.

Answer: ----------

2.10. Furthers suggested readings

1. Desmond S. T. Nicholl, An Introduction to Genetic Engineering Third Edition, University of the West of Scotland, Paisley, UK
2. Robert Schleif, Genetics and Molecular Biology, second edition, The Johns Hopkins University Baltimore, Maryland
3. NPTEL – Bio Technology – Genetic Engineering & Applications
4. T.A. Brown, Gene Cloning and DNA Analysis, an Introduction, Sixth Edition, University Of Manchester.
5. Robert Schlei, Genetics and Molecular Biology,2nd Edition
6. McGraw-Hill, Cell and Molecular Biology, Human Genetics: Concepts and Application, 9th Edition.

Contents

3.1. Introduction

Much like specialized tools designated for specific task, enzymes catalyse unique biochemical reactions, facilitating transformation in biological systems without being consumed in the process. By choosing the right enzyme, i.e. molecular tool for the application, it brings natures processes from *in-vivo* to *in-vitro*, supporting discoveries in molecular biology and enabling precise diagnosis of disease. In living things, enzymes function as biological catalysts to quicken chemical reactions. They are necessary for several biological functions, such as respiration of cells, DNA replication, and metabolism. Enzymes are useful instruments in molecular biology and biotechnology that enable a broad range of commercial and experimental uses. This introduction examines the many applications of enzymes as instruments, highlighting their adaptability, specificity, and effectiveness. Enzymes have the ability to speed up chemical reactions dramatically often by millions of times enabling processes that would not otherwise be feasible to complete in a reasonable amount of time. Their capacity to reduce the activation energy needed for reactions accounts for their efficiency. Enzymes also show a surprising degree of selectivity. Usually, an enzyme operates on a single type of substrate

or catalyzes only one kind of reaction. Because of their specificity, enzymes are perfect for focused applications in both industry and research, as they guarantee accurate and regulated biological reactions.

In molecular biology, enzymes are essential for working with DNA. For example, restriction enzymes can isolate and analyze genes by cutting DNA at particular sequences. After that, DNA ligases unite DNA fragments, enabling the creation of recombinant DNA molecules. DNA polymerases are necessary for the polymerase chain reaction (PCR), which amplifies particular DNA sequences. For the purposes of cloning, sequencing, and genetic disease diagnosis, this technology is indispensable.Additionally essential to DNA sequencing technology are zymes. DNA polymerases that include chain-terminating nucleotides produce fragments of different lengths that can be examined to ascertain the DNA sequence in Sanger sequencing. Enzymes are also used in more recent techniques like next-generation sequencing to manufacture, segment, and prepare DNA for sequencing.

Specialized proteins called restriction enzymes sometimes referred to as restriction endonucleases; cleave DNA at certain nucleotide sequences called recognition sites. These enzymes, which were first identified in the late 1960s, are normally present in bacteria and operate as a defense against viral DNA. A specific sequence of four to eight base pairs is recognized by each restriction enzyme, which then cuts the DNA inside or close to this spot to produce either blunt or sticky ends.

Restrictions enzymes are essential tools in molecular biology and genetic engineering because they can precisely cut DNA at specific spots. They are employed in the process of gene cloning, in which recombinant DNA molecules are made by inserting DNA pieces into vectors. Genetically modified organisms (GMOs), gene therapy, and the creation of biopharmaceuticals are just a few examples of the tremendous strides in genetic research, science, and medicine that have been made possible by this technology. In forensic science, genetic mapping, and paternity testing, restriction fragment length polymorphism (RFLP) analysis is one of the methods made possible by restriction enzymes. Because of their specificity and accuracy, scientists are able to precisely change DNA sequences, which open up new avenues for scientific discovery in a variety of domains.

Objectives:

After reading this unit, the learn will able to

- Know about restriction enzymes, its types and functions

- Also know about uses and restriction mapping of restrictions enzymes type –II
- Able to understand DNA modifying and end modifying enzyme's

3.2. Enzymes

Enzymes are proteins that catalyse a substrate's reaction. Enzymes used in genetic engineering specifically target DNA strands. Let us look at the enzymes that are utilized as instruments for genetic engineering and gene. Enzymes aid in particular tasks that are critical to the operation and overall health of the organism. They help to accelerate chemical reactions in the human body. They are necessary for respiration, digestion, muscle and nerve function, and other processes. The specificity and sensitivity of enzyme make them a valuable analytical tool for quantifying various compounds that act as enzyme substrate, activator or inhibitor.

Enzymes are biological catalysts that are virtually usually proteins. It accelerates the rate of a particular chemical reaction in the cell. The enzyme is not damaged during the reaction and can be utilized repeatedly. A cell includes thousands of distinct types of enzyme molecules, each tailored to a specific chemical reaction. The enzyme is used to make and improve nearly 400 every day consumer and commercial products. They are used in foods and beverages processing, animal nutrition, textiles, house hold cleaning and fuel of cars and energy generation.

An enzyme is a biological catalyst that is often a protein, but can also be RNA. A catalyst's purpose is to accelerate a reaction. And the genome encodes a large number of enzymes that produce proteins or RNAs that speed up numerous chemical reactions and perform thousands of different functions within cells. All enzymes were considered to be proteins until the 1980s, when the catalytic activity of some nucleic acids known as ribozymes (or catalytic RNAs) was demonstrated, disproving this premise. Because so little is known about the enzymatic action of RNA, this discussion will concentrate exclusively on protein enzymes. Recombinant DNA technology involves using enzymes and various laboratory technologies to manipulate and isolate DNA segments of interest. These method can be used to combine (or splice) DNA from different species or to create gene with new functions. The resulting copies are often referred to as recombinant DNA.

One or more polypeptide chains, or chains of amino acids, make up a big protein enzyme molecule. The distinctive protein folding patterns, which are crucial for enzyme specificity, are determined by the amino acid sequence. The protein structure may get denatured and lose its

enzymatic activity if the enzyme is exposed to perturbations, such as variations in pH or temperature. Denaturation can occasionally be reversed, but not always.

A cofactor is an extra chemical that is bound to certain enzymes and has a direct role in the catalytic action, making it necessary for enzymatic activity. Certain enzymes require both an inorganic metal ion and an organic molecule, such as a vitamin, to function as a cofactor. An enzyme might have a cofactor attached to it either tightly or loosely. When the cofactor is closely linked, it is called a prosthetic group. Prosthetic groups are non-peptide (non proteins) compounds that mostly attached to proteins and assist them in different way. They can be inorganic (like metal) or organic (carbon containing) and bind tightly to their targets. Prosthetic groups can bind via covalent (electron sharing) or non-covalent bonds. They also can help proteins binds other cellular components or a set as carrier of electrons or molecules (protons, H^+) and oxygen) to assist a cell in moving electrons or molecules form one place to another. By attaching to a specific group of proteins, called enzymes, prosthetic groups can make enzymes active or increase their activity prosthetic groups that attach to enzymes are after called cofactors or coenzyme because they help the enzyme to functions. A restriction enzyme, restrict endonuclease, REase, ENase or restrictase is an enzyme that claves DNA into fragments at or near specific recognition sites within molecules known as is restriction sites, Restriction enzymes are one class of the broader endonuclease group of enzymes.

One of the most significant functions of enzymes is to facilitate digestion. Digestion is the process of converting the food we ingest into energy. Our saliva, pancreas, intestines, and stomach all contain enzymes. They degrade lipids, proteins, and carbs. Enzymes employ these nutrients to promote cell development and repair. An enzyme also helps in, breathing, developing muscle, Nerves function and detoxifying our body.

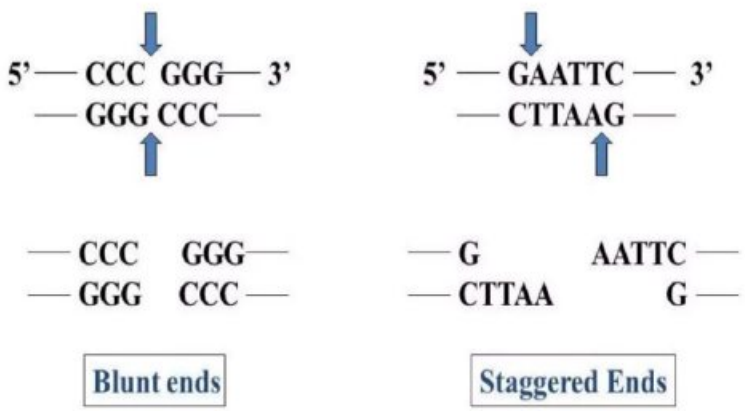
Enzymes have numerous industrial and medical applications. Wine fermentation, bread leavening, cheese curdling, and beer brewing have all been done since antiquity, but it wasn't until the nineteenth century that these events were recognized as the result of enzyme catalytic activity. Enzymes have since become more important in industrial operations involving organic chemical reactions. In medicine, enzymes are used to eliminate disease-causing microbes, promote wound healing, and diagnose certain disorders.

The body has thousands of individual enzymes. Each enzyme performs a single function. For example, sucrase is an enzyme that degrades a sugar known as sucrose. Lactase degrades lactose, a type of sugar present in milk. The most often used digestive enzymes are:

- **Carbohydrase**: Carbohydrase degrades carbs into sugars.
- **Lipase:** Lipase converts lipids into fatty acids.
- **Protease:** Protease is responsible for the breakdown of proteins into amino acids.

3.3. Restriction Enzymes

Enzymes produced by certain bacteria that have the ability to cleave DNA molecules at or near specific base sequences are known as restriction enzymes.One of the most crucial instruments in the recombinant DNA technology is the restriction enzyme.These are DNA-cutting enzymes found in prokaryotes, or bacteria that identify particular DNA locations known as restriction sites and cause a double strand break in the DNA at or close to these locations. The ends of the newly formed DNA will either be blunt or sticky (staggered) after cutting.



In other words the bacteria create a protein called a restriction enzyme, which cleaves DNA at particular locations. These proteins are found in microorganisms and are required for viral and other foreign DNA defence mechanisms. The names given to restriction enzymes are derived from the organism in which they were isolated, such as genus and strain. EcoRI is one of the most significant and commonly utilized restriction enzymes.

Bacteriophages are prevented from harming live bacteria by the restriction enzymes. They identify the bacteriophage, cut at its restriction points, and eliminate its DNA.

Restriction enzymes, also known as restriction endonucleases, are specialized proteins that cut DNA at specific sequences. Restriction enzymes, which originated in bacteria, function as a defence mechanism against viral DNA. Restriction enzymes identify and cleave distinct nucleotide sequences that typically range from four to eight bases long. Nucleases have been employed as key tools in different recombinant DNA techniques, particularly the class of nucleases known as restriction endonucleases, which is widely used in the production of recombinant DNA molecules. Restriction endo nucleases are DNases that detect certain nucleotide sequences and cleave double-stranded DNA at specified sites. They were identified in the 1950s as part of the restriction modification (R-M) system, which bacteria use to protect themselves against invading bacteriophages and genetic elements. Bacteria mostly remove undesired foreign DNAs via this R-M mechanism, and they are considered the equivalent of an immune system.More than 2000 restriction enzymes have been documented to far, and they are classified as type I, type II, type III, and type IV (Table 3.1). The approved nomenclature for R-M enzymes is based on Smith and Nathans' (1973) criteria, which were modified by Szybalski et al. (1988). The rules are as follows.

Table 3.1: Features of classified as type I, type II, type III, and type IV restriction enzymes

Features	Type-I	Type-II	Type-II	Type-IV
R-Mactive structure	Single enzyme three subunits (RMS) complex Mg	Separate enzymes. R. dimer	Single enzyme with two subunit complex	Separate monomeric enzymes
Cofactors	Restriction activity require ATP hydrolysis	-	-	Restriction activity is stimulated by AdoMet
Recognition site	Asymmetric bipartite	Palindromic	Asymmetric	Asymmetric
Cleavage	Variable distance either side	Same site	25-27 bp to 3' side	14 bp to 3' side

Genetic engineering benefits greatly from the use of restriction enzymes. They can be utilized in labs after being separated from the bacteria. The DNA contains brief, distinct nucleotide sequences called recognition sequences that are recognized by restriction enzymes. The precise and diversified cleavage capabilities of restriction enzymes allows for genome editing in a wide range of applications.

We provide a selection of high-quality, performance-tested restriction endonucleases for restriction digest and cloning applications. These enzymes cover a wide range of overhang needs, including a subset that allows for fast DNA digestion in 15 minutes or less. Promega's MULTI-CORE™ Buffer is a universal restriction enzyme buffer that facilitates multiple enzyme digestions. Bovine Serum Albumin (BSA) is also available to improve enzyme stability or to serve as a carrier protein.

The restriction enzyme breaks through the DNA molecule by hydrolyzing the link between neighbouring nucleotides when it identifies a sequence in DNA. With the aid of enzyme methylases, the bacteria adds a methyl group to the adenine or cytosine bases in the recognition sequence, preventing the destruction of its own DNA sequences.

In microscopic organisms like bacteria, there are about 400 known restriction enzymes that can recognize and cut over 100 different DNA patterns. The development of recombinant DNA technology led to the discovery of restriction enzymes as well as the elucidation of their specific mechanism and mode of action.

3.4. Types of Restriction Enzymes

Restrictions enzymes typically have two subunits: a restriction endonuclease that cleaves intruding double-stranded DNA and a methyl-transferase that modifies the host cell genome to protect it from cleavage. Restriction enzymes are widespread. They are categorized into four kinds based on their recognition sequence, subunit composition, cleavage position, and cofactor requirements.

The different classes (Types I, II, III, and IV) of restriction enzymes are distinguished by their structural makeup, cleavage selectivity, and enzymatic activity requirements. Each variety has unique characteristics and uses, especially in genetic engineering and research. Type II enzymes are especially preferred in molecular biology due to their ease of use and accuracy when cutting close to or at their recognition sites. For more information and specifics on the different types of enzymes, go to the table below.

Type I Restriction Enzymes

These enzymes that limit DNA cut it at a distance from the sequences that recognize it. DNA is cleaved at random locations distant from their recognition sequences by type I restriction enzymes, which are intricate, multipurpose proteins. As cofactors, they need ATP, S-adenosylmethionine

(SAM), and Mg^2+. Three subunits make up these enzymes: one is involved in recognition, one in restriction, and one in modification (methylation). For Type I enzymes, the recognition site normally consists of two unique sequences separated by a nonspecific spacer. After attaching itself to the recognition site, the enzyme starts a process known as translocation, which involves moving along the DNA and cleaving it at a non-specific, remote place. Because of their intricacy, Type I enzymes are less helpful for precise genetic engineering. Nevertheless, they are not very useful because they do not generate discrete restriction segments. These are intricate enzymes that modify and restrict many subunits. Originally believed to be rare, they are now shown to be common and of great biochemical interest thanks to genetic study.

Type II Restriction Enzymes

Type II restriction enzymes are the most commonly utilized in molecular biology because to their simplicity and precision. They cut DNA at predetermined locations inside or near their recognition sequences, which are typically palindromic and 4-8 base pairs long. Unlike Type I enzymes, Type II enzymes require only Mg^{2+} ions, not ATP or SAM.

These enzymes make precise cuts at locations that are nearer or inside the restriction sites. We see patterns of gel bands and discrete restriction pieces. In labs, they are only utilized for gene cloning and DNA analysis. These proteins are not connected to one another. They have the name of the species of bacterium from which they were isolated. For instance, EcoRI has been isolated from the E. Coli bacterium. There are two sorts of cuts produced by the restriction enzymes. Sticky ends result in an overhang, and blunt ends cut the DNA in the middle of the recognition sequence.Type II enzymes are classified further according to their recognition and cutting patterns:

Type IIP (Prototypic): It identifies palindromic sequences and cut within them. Consider EcoRI (GAATTC).

Type IIS (Shifted): It identifies non-palindromic sequences and cuts at a predetermined distance from the recognition site. As an example, consider FokI.

Type IIB (Bifunctional): It recognize palindromic sequences and cut along both sides of the recognition site. Example: BcgI.

Type IIG: It requires both a recognition site and a distinct binding domain for action, as do Types I and III. Example: BplI.

Type III Restriction Enzymes

These proteins have two subunits, Res and Mod, and are multifunctional. It's a methyltransferase modification. The Mod subunit identifies the unique DNA sequence for the system. Type III restriction enzymes typically cleave DNA at 25–27 base pairs downstream from their recognition sites. In addition to needing Mg^{2+} and SAM for activity, these enzymes require ATP for cleavage but do not hydrolyze it. Two subunits make up type III enzymes: one for recognising DNA and the other for cleaving and modifying it. DNA looping and cleavage are the outcome of the complex they create with their recognition site and another identical enzyme attached to a second site in a head-to-head configuration. One such is EcoP15I, which can identify the CAGCAG sequence.

Type IV Restriction Enzymes

Type IV restriction enzymes target and cleave modified DNA, particularly methylated or hydroxymethylated bases. Unlike the other types, they do not recognize specific sequences but rather the presence of certain modifications. These enzymes play a crucial role in host defense mechanisms against foreign DNA that might be modified to evade other restriction enzymes. Examples of Type IV enzymes include McrBC, which targets methylated cytosines.

In DNA mapping, molecular cloning, and genetic engineering, restriction enzymes are essential instruments. They make it possible for researchers to precisely cut DNA, which makes it easier to add or remove genes. For the purpose of producing recombinant DNA molecules, which may be inserted into host cells to create proteins or investigate gene activity, this precision is crucial. Additionally, restriction enzymes are employed in diagnostic procedures to find genetic mutations and variations, such as restriction fragment length polymorphism (RFLP) studies.

3.5. Uses and Restriction mapping of restriction enzyme

Restriction mapping is a technique for mapping an unknown section of DNA by breaking it into pieces and then detecting the positions of the breakpoints. This procedure depends on the employment of proteins called restriction enzymes, which can cut or digest DNA molecules at short, particular sequences called restriction sites. After a DNA segment has been digested with a restriction

enzyme, the resultant fragments can be studied using gel electrophoresis, a laboratory procedure for separating DNA bits based on their size. These are categorized into four types: I, II, III, and IV. Type II enzymes are most commonly used in molecular methods. Restriction mapping is the process of determining the relative locations of restriction sites on a DNA segment. Some enzymes detect four bases as restriction sites and make frequent cuts, resulting in relatively smaller fragments, whereas others recognize six or eight bases and make uncommon cuts to generate bigger fragments.

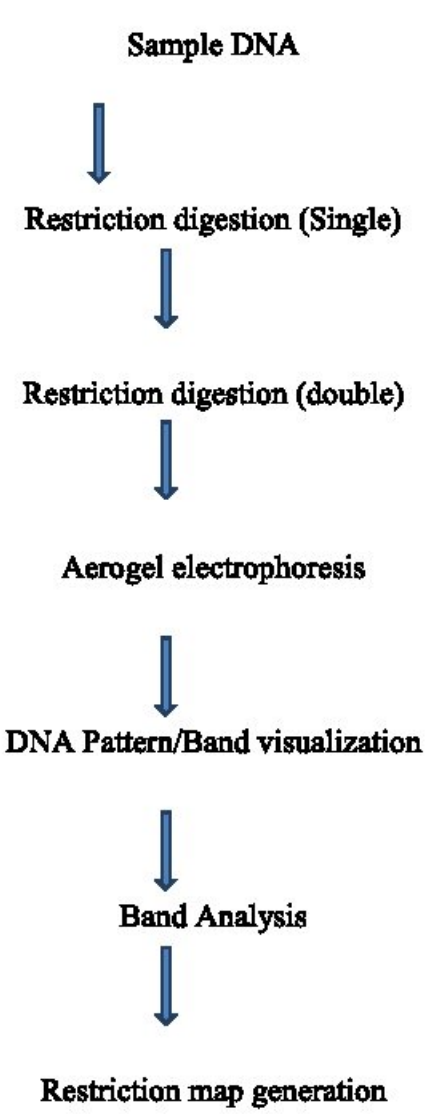
One popular method for creating a restriction map is to digest an unknown DNA sample in three different ways. Two pieces of the DNA sample are digested separately with distinct restriction enzymes, and a third portion of the DNA sample is double digested with both restriction enzymes at the same time.

Each digestion sample is then separated using gel electrophoresis, and the DNA fragment sizes are reported. The total length of pieces in each digestion will be equal. However, the length of each individual DNA fragment relies on the placements of its restriction sites. Each restriction site can be mapped according to the length of fragments.

The information from the double digestion is particularly useful for correctly mapping the sites. The final drawing of the DNA segment that shows the positions of the restriction sites is called a restriction map.

The position of restriction sites on a DNA fragment can be determined by digesting it with several restriction enzymes, both singly and in combination, and evaluating the resulting fragment sizes. Restriction mapping was once an important method for defining a cloned piece of DNA, but DNA sequencing has made this task easier. Nonetheless, analysing restriction sites (or fragment sizes) can be beneficial in comparing the chromosomal organization of different strains.

Process of restriction mapping



A restriction map is created by digesting a DNA molecule with restriction enzymes (preferably uncommon cutters) and then analyzing the resulting products for size using agarose gel electrophoresis. The size of the DNA fragment produced by restriction endonuclease digestion can be determined using standards run concurrently. Double digestion of DNA entails utilizing two independent restriction enzymes to position different restriction enzyme sites in relation to one another. The different sized DNA fragments obtained after restriction digestion can be used to calculate the distance between restriction sites. Data from many digestions can be combined to produce a more detailed restriction map.

Applications of Restriction Mapping

In a molecular biology, restriction maps are used as a reference to engineer plasmid or other relatively short pieces of DNA and some time for longer genomic DNA. Generating a restriction map

is typically the initial step in defining an unknown DNA and identifying different overlapping clones for contig mapping. Restriction maps are also used to identify the prerequisites for manipulating data for various purposes. Only a restriction map allows for the identification of sub cloning sites and the selection of appropriate restriction endonucleases for cleavage. The pattern of created restriction fragments can also be utilized to determine molecular relationships between two different species. Another use of restriction maps is to compare the mitochondrial DNA (mtDNA) of eukaryotic species.

Aside from being lower in size than the nuclear genome, an additional benefit of mtDNA is that it undergoes changes through mutation. It is approximately ten times faster than the nuclear genome, allowing for the determination of phylogenetic relationships between extremely closely related species, or even between different populations of the same species. Restriction enzymes can also be used to identify specific single nucleotide polymorphisms (SNPs). If a restriction enzyme can be discovered that cuts only one potential allele of a piece of DNA, it can be used to genotype the sample without fully sequencing it. The sample is initially run through a restriction digest to cleave the DNA, followed by gel electrophoresis.

Generation of Restriction Maps

The map's resolution is determined by the frequency of the restriction site in the DNA fragment, which takes into account both its size and base composition. To construct restriction maps of tiny DNA molecules such as plasmids, PCR fragments, and inserts, relatively affordable restriction enzymes that break DNA infrequently are commonly utilized. Rare cutters can be useful for larger vectors like cosmids and artificial chromosomes. This is called long-range restriction mapping. Such enzymes can be used to create restriction maps of complete chromosomes, but the resulting DNA fragments must be separated using pulsed field gel electrophoresis (PFGE) or comparable procedures (see Chapter 6 for further information). However, because a cell's genome produces so many fragments, restriction mapping is more practicable for comparing smaller pieces of DNA, which are typically a few thousand nucleotides in length.

The DNA to be restriction mapped is usually housed within a well-characterized plasmid or viral vector whose sequence is known. In reality, there are generally numerous known restriction sites immediately flanking the uncharacterized pure DNA, which simplifies generating the map. The following description assumes that the unknown DNA has been placed into a plasmid vector; however,

the principles can easily be adapted to other scenarios. Thus, two commonly used approaches for generating restriction maps are discussed below:

Restriction Map Generated by Digestion of Unknown DNA with Multiple Restriction Enzymes:

However, because a cell's genome generates so many fragments, restriction mapping is more practical for comparing smaller regions of DNA, which are typically a few thousand nucleotides long.

Consider a 6,000-bp plasmid vector with a 3,000-bp segment of unknown DNA (Figure 3.16). The Kpn I and Bam HI enzymes have unique recognition sites inside the vector, immediately surrounding the unknown DNA. Assume that separate digestion with Kpa I produces two fragments with sizes of 900 and 8,100 bp. There is a single Kpn I site in the vector, and the existence of a 900 bp fragment indicates that the unknown DNA also has a single Kpn I site, which is 900 bp away from the Kpn I site in the vector. Assume that digestion with Bam H I produces three fragments of sizes 600, 2,200, and 6,200 bp.

The existence of 600 and 2,200 bp pieces indicates that the unknown DNA contains two Bam H I sites. As a result, we can conclude that one Bam H I site is 2,800 bp (600 + 2,200 bp) away from the Bam H I site in the sample vector. The second Bam Hl site can be 600 or 2,200 base pairs away from the vector's Bam HI site. At this point, it is impossible to determine which of these various perspectives is right. Thus, digestion with individual enzymes determines the presence or lack of recognition sites, as well as the number of sites for a certain enzyme in the unknown DNA.

To determine the precise location of the second Bam Hl site, the plasmid is digested with Kpn 1 and Bam HI simultaneously. Assume this double digest produces pieces of 600, 900, 1,300, and 6,200 bp. The 600 bp fragment is identical to that produced after digestion with Bam H I alone. The 900 and 1,300 bp fragments indicate that Kpn I cuts within the 2,200-bp Bam HI segment. If the technique described above is repeated with a larger set of enzymes, a far more complete pulsed map will be produced. In essence, double digests are utilized to establish the proper order and orientation of the pieces.

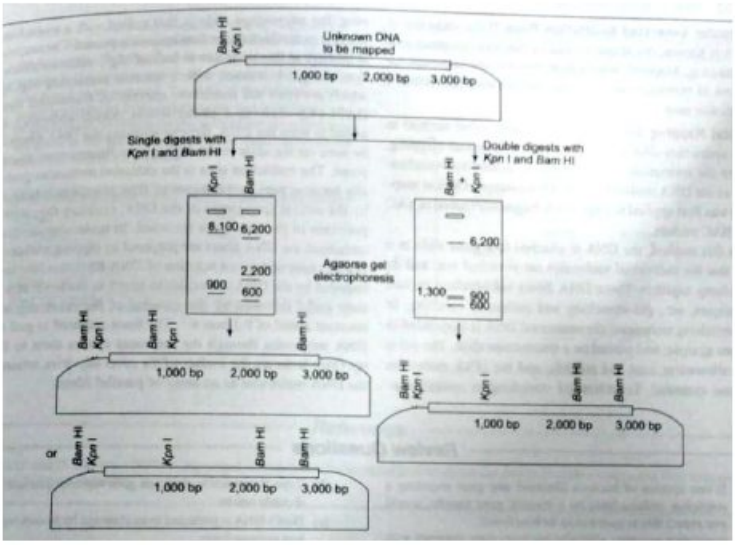


Fig. 3.1: Restriction mapping by multiple restriction enzyme catalysed digestion of unknown DNA

The effectiveness of this procedure is dependent on achieving complete digestion of the DNA with each of the enzymes used, as partial digestion will result in fragments that are a major cause of confusion. One technique to prevent this issue is to total up the predicted sizes of all the pieces in each lane, which should be nearly equivalent to the complete DNA. Another source of confusion is the formation of two fragments of about equal size that appear as a single fragment on agarose gels. This condition is frequently suspected by detecting an unusually bright fragment on the gel or a fragment that is larger than expected.

Computer Generated Restriction Maps

Computer Generated Restriction Maps If the DNA sequence is known, it can be fed into computer programs (such as Mapper), which will search for dozens of restriction enzyme recognition sites and provide a restriction map.

Optical Mapping

In this approach, the DNA is glued to a glass slide so that the individual molecules spread out and do not cluster together. These DNA filaments are formed using two techniques: gel stretching and molecular combing. The gel stretching technique suspends chromosomal DNA in molten agarose and places it on a microscope slide. The gel is then allowed to cool and solidify, causing the DNA molecules to expand. To use gel stretching in optical mapping, the microscope slide is first treated with

a restriction enzyme, after which molten agarose is poured. The enzyme is currently inactive due to a shortage of Mg^2. map.

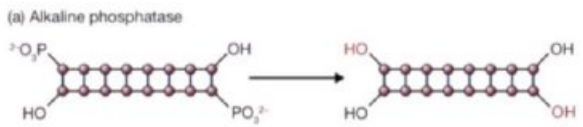
Following solidification, the gel is rinsed with aMgCl solution, which activates the restriction enzyme. A fluorescent dye, DAPI (4,6-diamino-2-phenylindole dihydrochloride), is used to stain the DNA, allowing the DNA fibers to be viewed on the slide with a high-power fluorescence microscope. The restriction sites in the stretched molecules eventually become gaps as the degree of fiber extension decreases due to the natural springiness of the DNA, allowing the relative positions of the cuts to be recorded.

3.6. DNA modifying enzymes

DNA modifying enzymes are a group of specialized proteins that play an essential role in the maintenance and expression of genetic information. These enzymes catalyze chemical reactions that modifying the structure or sequence of DNA including replication, repair, recombination and transcription. Restriction enzymes and DNA ligases perform cutting and joining roles in DNA manipulation. DNA modifying enzymes encompass all other enzymes used in genetic engineering. These enzymes are involved in degradation, synthesis, and modification of nucleic acids.

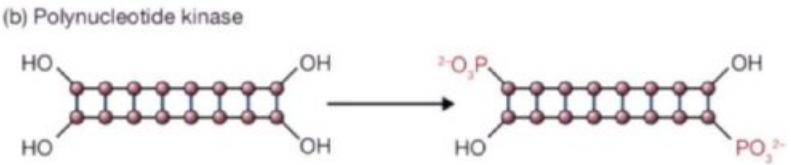
A). Alkaline Phosphatase

This enzyme removes phosphate group at the 5' terminus of a DNA molecule.



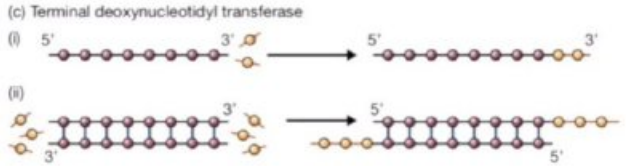
B). Polynucleotide kinase

This enzyme add phosphate group at the 5' terminus of a DNA molecule.



C). Terminal deoxyneucleotidyl transferase

This enzyme adds one or more deoxyribose neucleotides at the 3' terminus of a DNA molecule.



DNA ligase is a crucial biological enzyme that repairs damaged phosphodiester links, which can occur at random or during DNA replication or recombination. In genetic engineering, gaps in sugar-phosphate chains are sealed to create recombinant DNA by combining molecules from diverse sources. It functions as molecular glue, connecting DNA fragments. DNA ligase is an important enzyme in genetic engineering as it plays a significant role in ensuring experiment success.T4 DNA ligase, isolated from E. coli cells infected with bacteriophage T4, is the most commonly utilized enzyme in research. The enzyme is more effective at sealing gaps in fragments with cohesive ends, but it can also glue blunt-ended DNA strands together under certain conditions.T4 DNA Ligase forms a phosphodiester link between the 5'-phosphate and 3'-hydroxyl termini in duplex DNA or RNA. The enzyme can repair single-strand nicks in duplex DNA, RNA, or DNA/RNA hybrids. It can link DNA fragments with cohesive or blunt termini, but not single-stranded nucleic acids. The T4 DNA Ligase needs ATP as a cofactor.

Types of modifying enzymes

Nuclease enzymes disrupt the phosphodiester link between nucleotides to breakdown nucleic acids. Endonucleases, such as restriction enzymes, make cuts inside DNA strands.Exonucleases are a second type of nuclease that degrades DNA at its ends. Apart from restriction enzymes, four helpful nucleases are commonly utilized in genetic engineering.

- **Nuclease Bal 31:** Nuclease Bal 31 is a complicated enzyme. Its major activity is a fast-acting 3' exonuclease combined with a slow-acting endonuclease. When Bal 31 is present in large concentrations, these actions efficiently shorten DNA molecules from both ends.
- **Exonuclease III (exonucleases):** Exonuclease III is a 3' exonuclease that generates molecules with protruding 5' termini.

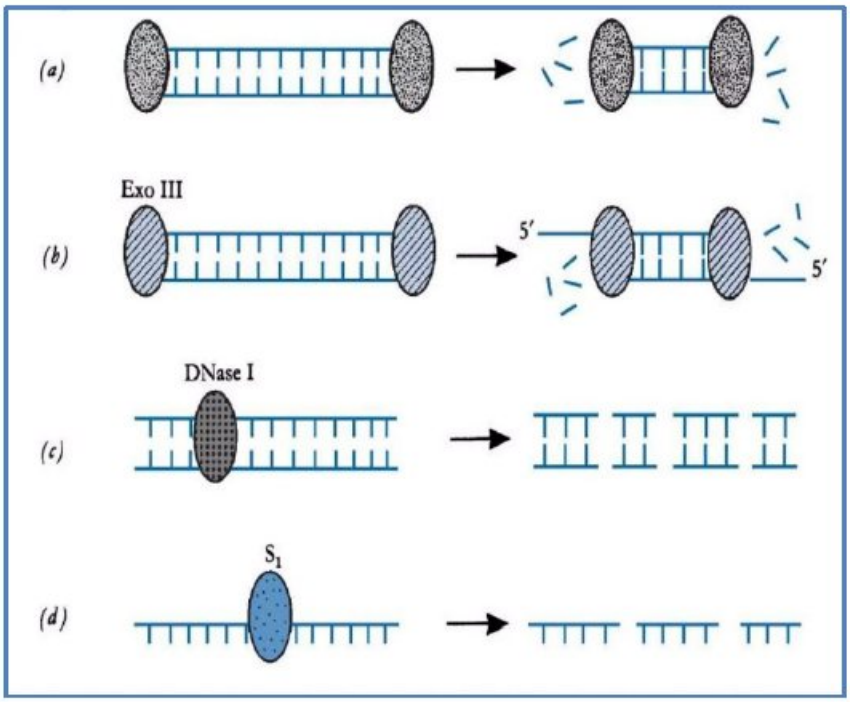
- **Deoxyribonuclease I (DNase I):** DNase I cuts either single-stranded or double-stranded DNA at essentially random sites.
- **S1-nuclease (endonucleases):** Nuclease S1 is specific for single-stranded RNA or DNA.

3.6.1. Nucleases

A nuclease (also archaically, known as nucleodepolymerase or polynucleotidage) is an enzyme capable of slowing the phosphodiester bonds between nucleotide of nucleic acids. Nuclease variously effect single and double standard breaks in their target molecules. In living organism they are essential machinery for many as peels of DNA repair. Defects in certain nucleases can cause genetic instability or immunodeficiency. Nucleases are also expensively used in molecular cloning.

There are two primary classification based on the locus of an activity. Exonuclease digests nucleic acids for the ends. Endonucleases act on regions in the middle of target molecules. They are further subcategorized as deoxyribonucleases and ribonucleases. Any enzyme that cleaves nucleic acids is called a nuclease. Nuclease enzymes degrade nucleic acids by breaking the phosphodiester link that connects the nucleotides. As members of the hydrolase enzyme class, nucleases are often selective enzymes that operate only on specific substrates, such as ribonucleases on RNA and deoxyribonucleases on DNA.

Restriction enzymes are good examples of endonucleases since they cut within the DNA strand. Restriction enzymes are nucleases that only split DNA molecules containing specific subunits. Some break the target DNA molecule at random locations (Type I), while others divide it only at the recognition site (Type II) or at a fixed distance from it (Type III). Type II and III restriction enzymes are effective instruments for determining the sequence of bases in DNA molecules. They serve a critical part in recombinant DNA technology, also known as genetic engineering.



Nucleases are a diverse collection of enzymes that hydrolyze DNA and RNA phosphodiester linkages. They perform critical functions in genetic quality control in nature, including as DNA proofreading, base, nucleotide, mismatch, and double-strand repairs, homologous recombination, and turnover. Nucleases are also commonly utilized in molecular and cell biology applications requiring precise nucleic acid manipulation, such as restriction digestion, degradation, and trimming. Nucleases are classified into groups depending on their functions and selectivity for various forms of nucleic acid substrate.

Nuclease Types:

Endonucleases: The enzymes that cause internal cuts in DNA molecules are called endonucleases. However, the endonucleases, with varying degrees of site recognition, cut both DNA and RNA from the middle of the chain. Sequence-specific cloning and gene analysis tools are restriction endonucleases, also referred to as restriction enzymes. These endonucleases are known as restriction endonucleases, and the sequences or recognition sites that they identify are known as recognition sites. Known as "molecular scissors," restriction endonucleases are crucial to the field of recombinant DNA technology.Restriction endonucleases (also known as restriction enzymes) are sequence-specific cloning and gene analysis tools. Each Restriction Endonuclease acts by determining the length of a DNA sequence. When it finds its unique recognition sequence, it binds to the DNA and cuts it by catalyzing the hydrolysis of the phosphodiester bond between neighbouring nucleotides. All

restriction enzymes cut DNA twice, once in each of the two strands of the double helix at specified locations in their sugar-phosphate backbones. A palindrome in DNA is a sequence of nucleotide base pairs that reads the same on both strands while the reading orientation remains constant. The recognition sequences of a restriction enzyme EcoR1 are provided below to understand the palindromic nucleotide sequence:

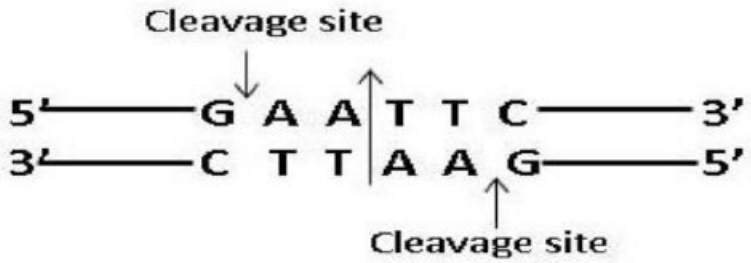


Fig: 3.2: EcoR1 recognition sequence

In genetic engineering, restriction enzymes are utilized to create recombinant DNA molecules. Figure 2 illustrates the procedures for creating recombinant DNA with restriction enzymes. This comprises the same restriction enzyme cleaving the desired DNA and the vector DNA, producing the same kind of sticky ends. An enzyme known as DNA ligase connects the desired DNA fragment with the sticky ends of the vector DNA.

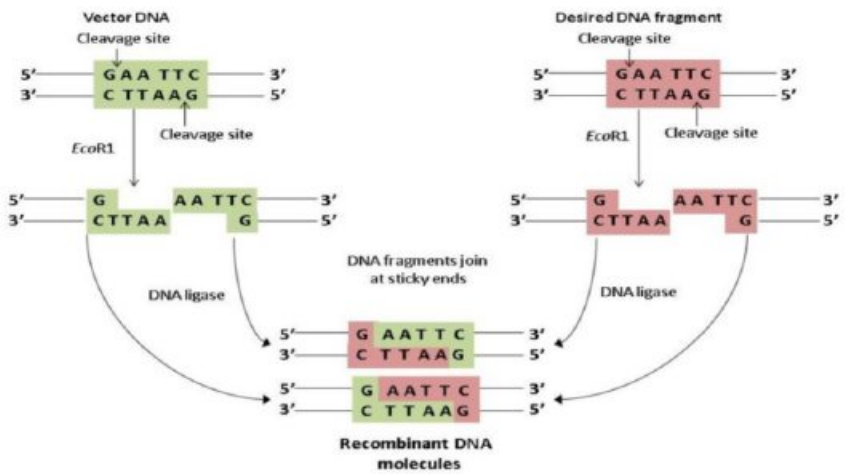


Fig.3.3: Steps in the formation of Recombinant DNA

Source: How does restriction endonuclease function? (vedantu.com)

There are three types of restriction endonucleases: Type I, Type II, and Type III. Type I restriction endonucleases cut DNA at random places more than 1,000 base pairs away from the recognition sequence. Type III restriction endonucleases break the DNA approximately 25 base pairs from the recognition sequence. Both Types I and II require energy in the form of ATP to function properly. Type II restriction endonucleases are simple and do not require ATP to cleave DNA within the recognition sequence. Thus, only Type II Restriction Endonucleases are used in recombinant DNA technology.

Exonucleases: Endonuclease is an enzyme that cleaves the nucleotide sequence at a specific point in the DNA. In contrast to endonucleases, exonucleases cleave nucleotides at the 3' or 5' ends of DNA and RNA chains one at a time. Exonucleases preferentially break down single-stranded DNA and remove overhangs. Endonuclease activity begins with an evaluation of the length of the DNA sequence; once a specific recognition sequence is identified, endonuclease binds to it and chops each of the DNA strands at certain locations.Endonucleases, often known as molecular scissors, are utilized in genetic engineering, assaying, cloning, and producing recombinant proteins.

In prokaryotes and eukaryotes, exonucleases are of 3 types:

- Decapping 5' to 3' exonuclease
- Independent 5' to 3' exonuclease
- Poly A specific 3' to 5' exonucleases

The majority of exonucleases remove mononucleotides in the 3'-5' direction of the DNA/RNA strand by biochemical catalysis. Exonucleases remove nucleotides from the terminal locations of nucleic acids one at a time, resulting in overhangs known as sticky ends. Exonucleases have important roles in genetic quality control, DNA proofreading during replication, homologous recombination, and DNA repair. As a result, they are believed to be the defenders of genome stability, hence exonucleases are crucial to numerous DNA metabolism activities.

Exonuclease is an important enzyme in the maintenance of cellular metabolism. DNA transcription and translation generate all biological components, particularly proteins. So, any DNA misformation will result in protein loss or dysfunction, disrupting cellular pathways and signaling. To prevent such errors, these specific exonucleases proofread DNA. These exonucleases use catalysis to

generate error-free DNA, ensuring that biological functions are not affected. Exonucleases inhibitors can be used to regulate the activity of exonucleases.Exonucleases modify DNA and check it for mistakes in order to preserve genetic stability and integrity. This is an extremely precise technique that calls for a high level of specificity. Exonucleases proofread DNA during DNA polymerization, searching for any odd structures that could interfere with DNA replication and mending damaged DNA in the process. Their application has earned them the moniker "DNA protectors."

Some examples of exonucleases are:

Exonuclease I: It degrades extra ssDNA primers in vitro from a combination of double-stranded products. Holds applications in PCR.

Exonuclease III: It performs unidirectional nested deletion.

Lambda exonuclease: It removes single nucleotides from the 5' ends of double-stranded DNA.

DNases: DNases, or deoxyribonucleases, degrade DNA rather than RNA, and different varieties can cut from both inside and outside. DNase I is frequently used in research to remove contaminated DNA from RNA and protein samples, clean cell cultures, and perform DNA fragmentation studies. The DNase enzyme functions in cells by breaking down extracellular DNA (ecDNA) ejected by apoptosis, necrosis, and neutrophil extracellular traps (NET) to assist minimize inflammatory responses that might otherwise be generated. There are numerous deoxyribonucleases that belong to one of two families (DNase I or DNase II), each with unique substrate specificities, chemical processes, and biological activities. DNase's laboratory applications include the purification of proteins recovered from bacterial organisms. Additionally, DNase has been utilized to treat disorders caused by ecDNA in blood plasma. DNase assays are also becoming more used in studies.

RNases: In contrast, ribonucleases prefer to digest RNA rather than DNA. RNases, also known as RNase A and RNase H, are commonly used to remove tainted RNA from samples and RNA testing. RNases are frequently more selective for single-stranded RNA than hybridized RNA (with the exception of RNase H).RNases are enzymes that help degrade RNA. RNases have multiple cysteine residues that create several intramoleculardisulfide connections; the robust nature of these enzymes renders them resistant to many decontamination methods, and powerful chemical treatments are frequently required to remove RNases from surfaces and solutions.RNases aid in the breakdown of

RNA by catalyzing cleavage on the molecule. By severing the connections between nucleotides, they disassemble RNA into smaller pieces and give other enzymes access. In labs where researchers work with RNA, RNases are typically regarded as pollutants because they reduce the quality of samples and RNA integrity.

Table: Properties of ribonucleases (RNases)

Ribonuclease	Applications	Reaction catalyzed	Substrate	Reaction products
RNase A, DNase and protease-free	Removal of RNA from plasmid or genomic DNA preparations Removal of RNA from recombinant protein preparations Ribonuclease protection assays	Cleavage of phosphodiester bonds between 3'-C or 3'-U residues and the 5'-OH residue of an adjacent nucleotide	ssRNA	3'-CMP 3'-UMP oligonucleotides with terminal 3'-CMP 3'-UMP
RNase I	Removal of RNA from DNA solutions Removal of RNA from recombinant protein preparations Ribonuclease protection assays	Cleavage of all phosphodiester bonds in an RNA temp	ssRNA	3'-NMP
RNase T1	Removal of RNA from DNA solutions RNA sequencing Ribonuclease protection assays	Cleavage of phosphodiester bonds between 3'-G residues and the 5'-OH residue of an adjacent nucleotide	ssRNA	3'-GMP oligonucleotides with terminal 3'-GMP

Strand-specific Nucleases: Several nucleases have been produced due to their selectivity for single- or double-stranded nucleic acids. Micrococcal nuclease, S1 nuclease, and Mung Bean nuclease are used to selectively digest and remove overhangs in single-stranded DNA and RNA (including DNA-RNA hybridization). In contrast, duplex-specific nucleases prefer double-stranded DNA or RNA over single-stranded DNA or RNA.

Other nucleases: Streptococcus pyogenes produces the nuclease known as Cas9 (streptococcus pyogenes associated protein 9), which is used in the CRISPR/Cas9 genome editing

technique. Combining the Cas9 nuclease with a sequence-specific guide RNA allows for the modification, insertion, or deletion of a particular region of the genome.

3.6.2. Polymerases

Polymerase is an enzyme that synthesizes long chain of polymers or nucleic acids. DNA polymerase and RNA polymerase are used to assembles DNA and RNA molecules, respectively, by coping a DNA template strand using a base-pairing interactions or RNA by halt ladder replication. DNA polymerases for the thermopiles bacterium, thermos aquaticus (Taq) is used in the polymerases chain reactions important techniques of molecular biology. Polymerases may be template dependent or template independent.

DNA polymers are a group that catalyzes the production of phates (NTP performing the most fund DNA replication and replication Different polymerase types exist in both prokaryotes and rukaryotes. All DNA polymerases for synthesis of DNA mporeisainion to dNTPs, an initiating oligomer known as a primer, carrying and hydroxyl gresep that is employed as a beginning point of the growth DNA polyannotintrute synthesis de novo from ties. The majority of polymerases are uplate-dependent. By aligning with the complementary regass in the DNA template, the priser supplies the DNA polymerase with a double-sounded structure.The enzyme subsequently proceeds along the DNA template, extending the perimer in five directions (Fig. 3.4), following the Watson-Crick base pairing rule, which states that A pairs with T and C pairs with G. In addition to major 5'-3' polymerase activity, most polymerases have 3'-5' proofreading activity, which is responsible for reducing errors in newly produced DNA.

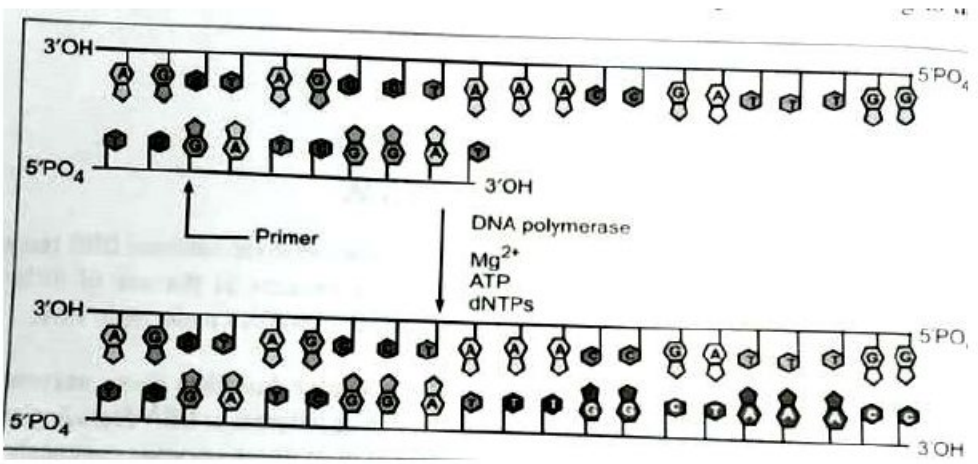


Fig.3.4: DNA polymerase has 5,-3, polymerization activity using DNA template

Source: Jayant K. Pal and Saroj S. Ghaskadbiby , Fundaments of nuclear biology, oxford publication

The most widely used polymerase is DNA polymerase I from E. coli. This enzyme, in addition to 5'-3' polymerase and 3'-5' exonuclease activities, exhibits 5'-3' exonuclease activity. Its 5'-3' exonuclease activity is very useful in the process of nick translation. This enzyme is employed in the nick translation procedure to prepare labeled DNA probes for hybridization. Klenow, a modified variant of this enzyme, is more often utilized in several DNA labeling procedures, including random primer labelling and 5' end labelling. Klenow is a large piece of DNA polymerase 1 produced from enzyme cleavage by the protease subtilin. It keeps the subunit. 5'-3' activity is represented by 5'-3' polymerase and 3'-5' exonuclease (see Figure 3.2). The tiny fragment contains 5'-3' e modest exonuclease activity.

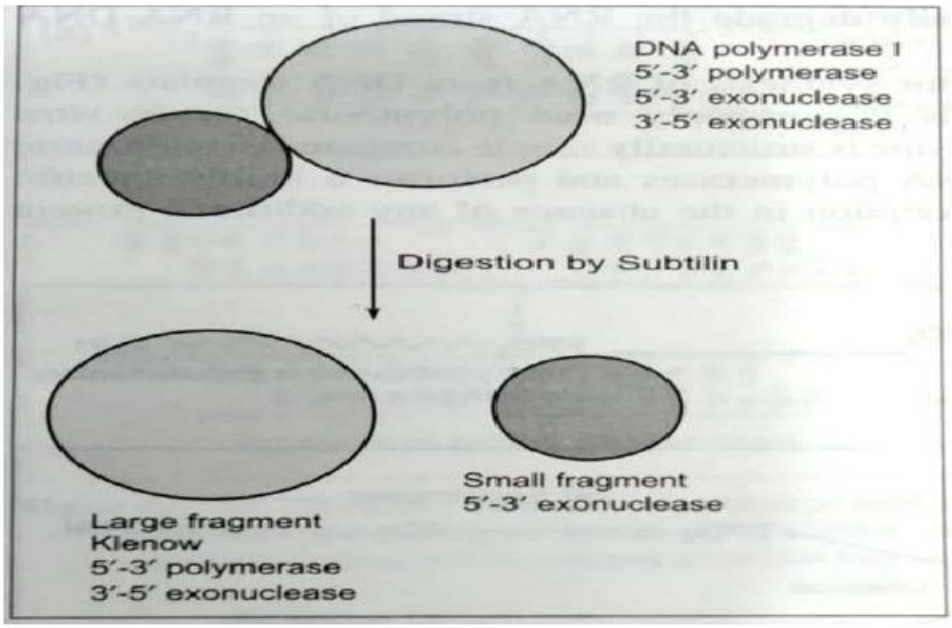


Fig. 3.5: DNA polymerase has one small and one large sub- unit. These can be separated on cleavage of holoenzyme by protease subtilin giving one large and one small subunit. The large subunit is called as Klenow and retains 5' - 3' polymerase and 3' - 5' exonuclease activity, while the small subunit has 5' - 3' exonuclease activity.

Source: Jayant K. Pal and Saroj S. Ghaskadbiby , Fundaments of nuclear biology, oxford publication

Arthur Kornberg was the first to purify and characterize E.coli DNA polymerase. It is a single-chain polypeptide currently referred to as DNA polymerase-I. Scientists have discovered five DNA polymerases in E. coli. DNA polymerases are primarily responsible for duplicating a cell's DNA content during division. They accomplish this by adding nucleotides to the 3'-OH group of the developing DNA strand.DNA polymerases are enzymes that catalyze the creation of new DNA strands from existing ones, which is essential for DNA replication. DNA polymerase 1 (Pol 1) and DNA polymerase 3 (Pol 3) are the two main polymerases involved in this process. Pol 1 is largely responsible for eliminating RNA primers and filling in gaps with DNA nucleotides during DNA replication, whereas Pol 3 is the primary polymerase responsible for synthesizing new DNA strands. Polymerase enzymes, which synthesise copies of nucleic acid molecules, are utilized in numerous genetic engineering methods.When describing a polymerase enzyme, the terms 'DNA-dependent' or 'RNA-dependent' might be used to identify the kind of nucleic acid template employed by the enzyme.Thus, a DNA-dependent DNA polymerase copies DNA into DNA, an RNA-dependent DNA polymerase copies RNA into DNA, and a DNA-dependent RNA polymerase converts DNA to RNA.

DNA Polymerases

In E. Coli, five DNA polymerases have been found. The structures, roles, rates of polymerization, and processivities of each DNA polymerase vary.

DNA Polymerase I

DNA polymerase I is encoded by the polA gene. It is a single polypeptide that plays a role in recombination and repair. It exhibits both 5'→3' and 3'→5' exonuclease activity. DNA polymerase removes the RNA primer from the lagging strand via 5'→3' exonuclease activity and fills the gap. The DNA polymerase I incorporates in modification of nucleotides. It is also active in multiple buffers, including restriction enzyme, PCR, and RT buffers. The main Applications of DNA polymerase I is in DNA labelling and second-strand synthesis of cDNA in conjunction with RNaseH.

DNA Polymerase II: The polB gene codes for DNA Polymerase II. There are seven components in it. Its primary function is to support and repair DNA polymerase III. Its exonuclease activity is 3'→5'.

DNA polymerase III: The primary enzyme for replication in E. Coli is DNA Polymerase III. Coding for it is the polC gene. DNA polymerase III has the highest rate of polymerization and processivity. It also contains 3'→5' exonuclease activity for proofreading.

DNA polymerase IV: A gene called dinB codes for DNA Polymerase IV. When DNA replication stalls at the replication fork during the SOS reaction, its primary function is to repair the damaged DNA. Translation (translation synthesis (TLS) is one of the pathway to overcome stalled replication in which specific polymerases (TLS) performed by pass synthesis across DNA damage) polymerases are DNA polymerase II, IV, and V.

DNA Polymerase V: Additionally, translation synthesis during the SOS response and DNA repair are facilitated by DNA Polymerase V. UmuC monomer and UmuD dimer comprise its composition.

Similar to prokaryotic cells, eukaryotic cells also include a large number of DNA polymerases that serve a variety of purposes, such as nuclear and mitochondrial DNA replication. DNA polymerase γ and α are primarily responsible for nuclear DNA replication. In humans, at least 15 different DNA polymerases have been found.

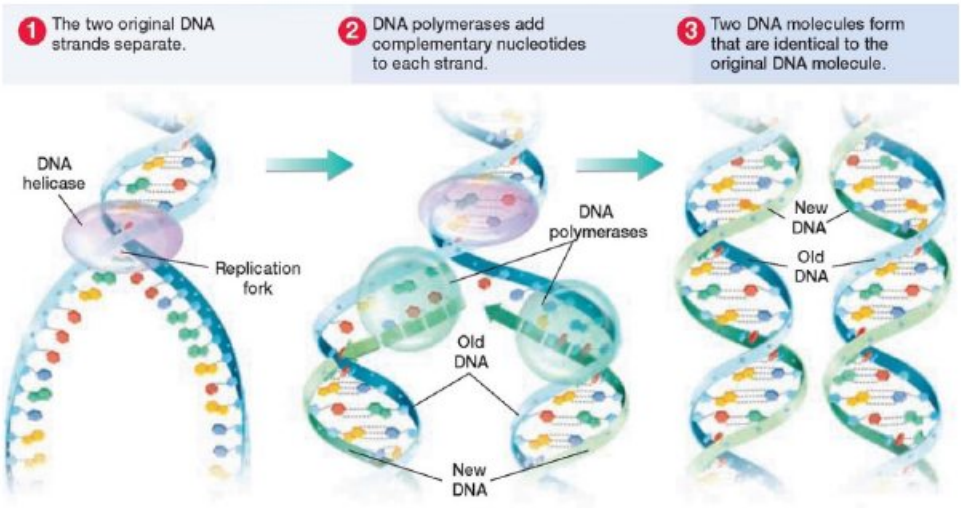
- DNA polymerase δ – In eukaryotes, this enzyme is the primary replication agent. For proofreading, it also contains a 3'→5' exonuclease activity.
- DNA polymerase α – Primers are created by DNA polymerase α, which is its primary job. The activity of the smaller component is primase. There is polymerization activity in the biggest subunit. It functions as a primer for Okazaki fragments, which DNA polymerase α
- DNA polymerase ϵ – Repair of DNA is the primary function. Primers for Okazaki fragments are extracted from the lagging strand.
- DNA polymerase γ – It is the main replicative enzyme for mitochondrial DNA.

DNA polymerase Function of DNA polymerase

Replication

The replication process is the primary way DNA polymerase synthesizes new DNA. Maintaining and transmitting genetic information from one generation to the next is a vital function. When DNA polymerase functions in pairs, it replicates two DNA strands at the same time.

Deoxyribonucleotides are introduced into the growing DNA strand's 3'-OH group. Polymerization activity causes DNA strands to expand in a 5'→3' direction. Guanine pairs with cytosine, and adenine pairs with thymine. To initiate replication, DNA polymerases require the addition of a primer to the nucleotides. In prokaryotes, DNA polymerase III is the major enzyme responsible for replication. DNA polymerase *r* is the major enzyme involved in replication in eukaryotes. DNA polymerase I use 5'→3' exonuclease activity to remove the RNA primer and replace it with polymerase activity on the trailing strand.



Repair

Maintaining genome integrity is critical, and the replication process is a massive operation. DNA repair is the continual process of correcting any faults in the genome produced by DNA damage, other than replication errors. There are various ways DNA can be repaired.

3.6.3. DNA ligase

DNA ligase is the type of enzyme that facilitates the joining of DNA strand together by catalyzing the formation of phosphodiester bond. It play a role in replacing single stand break is duplex DNA in living organism, but some forms (such as DNA ligase IV) may specifically repair double strand breaks (i.e. a break in both complementary strand of DNA). Single strand break an repaired by DNA ligase using the complementary strand of double helix as a template, with DNA ligase creating the final phosphodiester bond to fully repair the DNA.

DNA ligase is an important cellular enzyme, as its function is to repair broken phosphodiester bonds that may occur at random or as a consequence of DNA replication or recombination.It can therefore be thought of as molecular glue, which is used to stick pieces of DNA together.DNA ligases catalyze the creation of a phosphodiester link between the 5' phosphate and the 3' hydroxyl end, separated by a nick (Fig. 3.6a and b). Bacterial DNA ligases.For example, E. coli and B. subtilis employ NAD hydrolysis as an energy source, whereas ATP serves as a cofactor for DNA ligases from bacteriophages (e.g., T4 and T7) and eukaryotic cells. In vivo, the enzyme is absolutely required for the joining of Okazaki fragments during DNA replication. The enzyme is present in all living creatures. The most often utilized ligases in molecular biology research are E. coli DNA ligase and T4 DNA ligase.

The ligation process is carried out in three steps using two relatively stable intermediates. The first step involves the covalent transfer of the S-adenyl group of NAD or ATP to an epsilon-amino group of a lysine residue in the enzyme, resulting in the formation of a ligase adenylate intermediate and the subsequent release of NMN or ppi. This reaction takes place regardless of the presence or absence of DNA substrates. The AMP transfer activates the 5' phosphate group by establishing a new pyrophosphate linkage in the second intermediate, DNA-adenylate.The DNA-AMP intermediate releases a stoichiometric amount of AMP when it forms a phosphodiester bond with the acceptor molecule's 3' OH group. The penultimate step involves the nucleophilic displacement of the AMP via the attack of 3' OH at the activated 5' PO group. Each stage of the ligase reaction is reversible. The rate of ligation and the nature of the end products are affected by a variety of parameters, including temperature, enzyme concentration, substrate DNA concentration, and the shape and flexibility of DNA molecules.

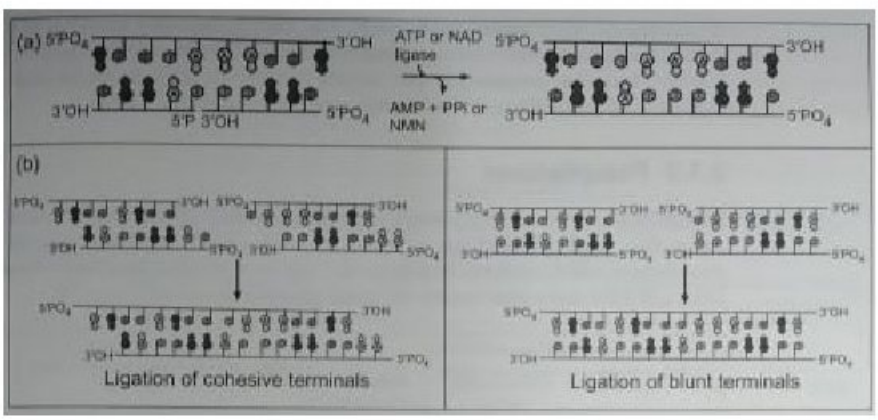


Fig.3.6: Ligase joins the two DNA fragments by synthesizing a phosphodiester bond between 5'PO, and 3'0H termini separated by a nick. Ligase requires ATP (T4 DNA ligase) or NAD (E coll DNA ligase) for making a bond (b) Ligase can join both cohesive and blunt termini of the DNA fragments

T4 DNA Ligase

- The enzyme repairs single-strand nicks in duplex DNA, RNA, or DNA/RNA hybrids.
- It also joins DNA fragments with either cohesive or blunt termini, but has no activity on single-stranded nucleic acids.
- The T4 DNA Ligase requires ATP as a cofactor.

Applications

- ❖ Cloning of restriction enzyme generated DNA fragments
- ❖ Cloning of PCR products
- ❖ Joining of double-stranded oligonucleotide linkers or adaptors to DNA
- ❖ Site-directed mutagenesis

T4 RNA Ligase

T4 RNA Ligase catalyzes the ATP-dependent intra-and intermolecular formation of phosphodiester bonds between 5'-phosphate and 3'-hydroxyl termini of oligonucleotides, single-stranded RNA and DNA.

Applications

- ❖ Joining RNA to RNA
- ❖ Specific modifications of tRNAS
- ❖ Site-specific generation of composite primers for PCR

3.6.4. Phosphatases:

A phosphate se is enzyme that uses water to cleave a phosphoric acid monoester into a phosphate ion and an alcohol. Because a phosphate enzyme catalyzes the hydrolysis of its substrate, it

is subcategories of hydrolases. Phosphatase enzymes are essential to many biological functions, because phosphorylation (e.g. by protein kinase) and dephosphorylation (by phosphatases) serve diverse roles in cellular regulation and signalling. Whereas phosphatases remove phosphate groups from molecules, kinases catalyze the transfer of phosphate groups to molecules from ATP. Together, Kinases and phosphatases direct a form of post translational modification that is essential to coll's regulatory network.

Phosphatases are enzymes that remove terminal phosphates. The most often utilized phosphatase is E. coli's bacterial alkaline phosphatire, which catalyzes the hydrolysis of 3' and/or 5' terminal phosphates from DNA and RNA, as well as a wide range of organic phospho-monoesters (Fig 3.7).

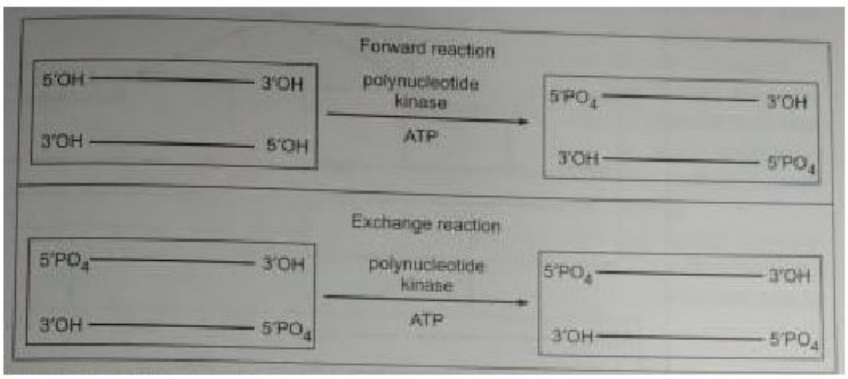


Figure 3.7: Alkaline phosphatase dephosphorylases the termini by removing the terminal phosphate. This process is referred to as baptisation

The rate and efficiency of phosphatase reaction are determined by pH, ionic strength, and temperature. The optimal pH range is 8-9.5, and the process has no observable activity below pH 6. The ideal temperature for the reaction is around 65°C. The activity increases by up to tenfold as the concentration of the salts, such as KCI or NaCl, increases from 0.1 to 1 M.

Za is essential for both catalytic activity and enzyme stability. The enzyme is mostly utilized for dephosphorylating the termini of DNA or RNA. During molecular cloning, dephosphorylation of 5' PO of the vector termini reduces self-ligation of vector, thus boosting the yield of intermolecular ligation products.

3.6.5. Polynucleotide Kinases

Kinases transport the phosphate group. They are a hugely diversified collection of enzymes that catalyze the transfer of gamma phosphoryl groups from NTP to an acceptor molecule. Kinases are broadly classified as protein kinases, carbohydrate kinases, and polynucleotide kinases based on the type of phosphate acceptor.The most prevalent polynucleotide kinase is T4 polynucleotide kinase. This transfers the terminal phosphate group of ATP to the 5' OH group of DNA, RNA, or oligonucleotides. This process is completely reversible, therefore under the right conditions, polynucleotide kinase can catalyze both phosphorylation of dephosphorylated ends, known as the forward reaction, and phosphate exchange with phosphorylated ends, known as the exchange reaction (Fig 3.8). This enzyme is commonly used to mark the 5 termini of nucleic acids.

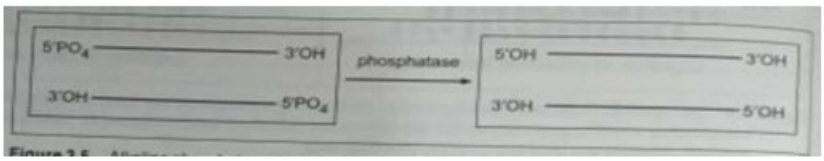


Fig. 3.8: Polynucleotide kinase phosphorylates the termini by transferring the phosphate group from ATP either in exchange reaction or forward reaction. Forward reaction is more efficient

3.7. End modifying enzymes

A class of specialized proteins known as DNA modifying enzymes is crucial to the expression and preservation of genetic information. These enzymes catalyse chemical processes such as transcription, recombination, replication, and repair that alter the sequence or structure of DNA. The significance of DNA-modifying enzymes and their possible uses in molecular biology will be discussed in this article.

Double-stranded DNA termini frequently need to be modified in order to make the molecule ready for cloning. A 5' monophosphate is needed by DNA ligases to adenylate the donor end, whereas a 3' hydroxyl group is needed for the acceptor end. The sequences that are to be attached must also be compatible; that is, they must be cohesive ends with complementary overhangs to other cohesive ends or blunt ends joined to other blunt ends. End alterations are carried out to enhance cloning efficiency, guarantee compatibility of the ends to be united, and optimize the location of translated and regulatory sequences.

Among the most well-known enzymes that change DNA is DNA polymerase. During replication, it catalyzes the insertion of nucleotides into an expanding DNA chain by utilizing an

already-existing template strand. Since DNA polymerase makes sure that every new cell obtains an exact copy of the genetic material, it is necessary for the proper duplication of genetic material during cell division. DNA polymerases are also essential for replacing and removing erroneous nucleotides from damaged DNA.

The genetic information that organisms need to survive and reproduce is stored in their DNA. Cellular function depends critically on the stability of DNA molecules. Thus, a variety of endogenous (such as free radicals generated by metabolic intermediates) and exogenous (such as UV light, electric radiation, chemical poisons, Figure 1) factors continuously harm the DNA of human organism cells. Damage to cells can occur 104 times in a single day. Damage that cannot be fixed in a timely manner will result in malignant tumors, unchecked cell proliferation, death, and genetic variety. For instance, prolonged exposure to intense sunlight on human skin can readily lead to skin cancer. Direct repair (DR), mismatch repair (MMR), base excision repair (BER), nucleotide excision repair (NER), and double strand break repair (DSBR), which includes non-homologous end joining (NHEJ) and homologous recombination (HR), are among the techniques used to repair DNA in mammals.

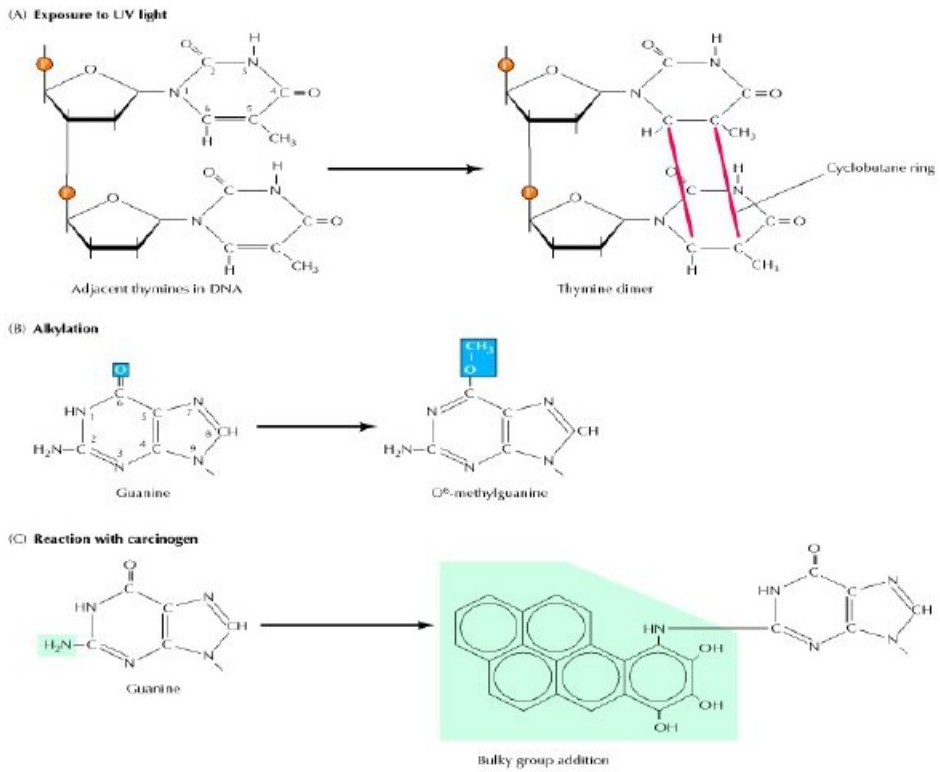


Fig.3.9. Examples of DNA damage induced by radiation and chemicals.

UV light causes the creation of pyrimidine dimers, which are made up of two neighbouring pyrimidines (for example, thymines) connected by a cyclobutane ring structure. (B) Alkylation is the process of adding methyl or ethyl groups to specific places on the DNA bases. In this example, alkylation of guanine's O6 position yields O6-methylguanine. (C) Many carcinogens, such as benzo-(a) pyrene, react with DNA bases, resulting in the addition of large bulky chemical groups to the DNA molecule.

Direct repair (DR)

DNA damage can be easily repaired using direct repair, which recovers the damaged base in a single step. O6-alkylguanine DNA alkyltransferase (AGT), which is widely distributed in the body, is a significant repair factor in DR. Additionally, practically every kind of cancer cell, including lung, pancreatic, and colon cancer cells, expresses this enzyme. The repair of AGT, of which O6-methylguanine DNA methyltransferase, MGMT, is the most significant DR enzyme, is strongly linked to the resistance of tumor cells to the methylation medicines (temozolomide, TMZ) and chlorinated medications (carmustine, lomustine). Guanine methylation and alkylation damage happen at the O6 location when the alkylating substance interacts with the body's DNA. By moving the methyl group from the O6 position of O6-methylguanine to the cysteine residue at position 145 on the MGMT, it is able to irreversibly inactivate itself and restore the guanine on the DNA strand (Figure 2). Tumor cells become resistant to chemotherapeutic medicines and fight against alkylating agents as a result of MGMT, which also shields cells from tumors and prevents DNA damage from alkylating groups.

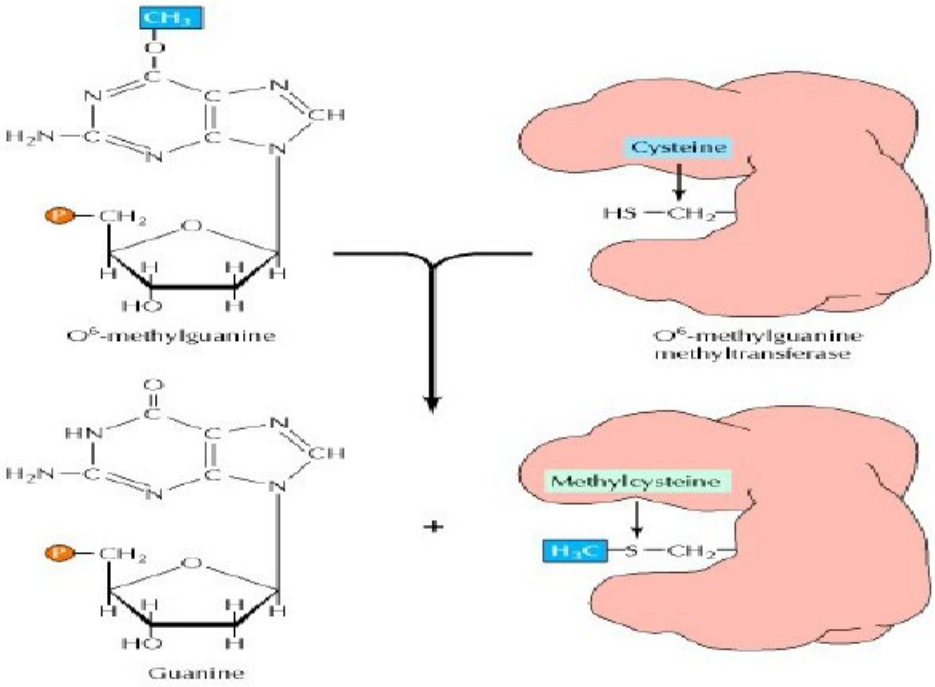


Fig.3.9: Repair of O^6-methylguanine.

DNA ligase is an additional important enzyme that modifies DNA and is necessary for both DNA replication and DNA repair. DNA ligase creates a continuous strand during replication by bridging the gaps between freshly formed DNA fragments. Additionally, single-strand breaks in DNA are repaired by DNA ligase, which also joins the broken ends of double-strand breaks. Apart from DNA polymerase, DNA ligase, and topoisomerases, numerous other enzymes that change DNA are essential for preserving the genetic stability of an organism. Helicases, for instance, are enzymes involved in transcription, repair, and replication of DNA that unwind the double helix. In order to expose the template strand and provide other enzymes access to the DNA, helicases are necessary.

DNA modifying enzymes play an important role in the integrity and expression of genetic information. The advent of novel DNA-modifying enzymes and technologies, such as CRISPR-Cas, has opened up new paths for genetic study and holds great promise for medical and agricultural applications. However, the ethical and safety concerns of gene editing must be carefully balanced against the potential benefits of the technique.

3.8. Summary

Nucleases are either sequence-independent or sequence dependent. Endonucleases cleave DNA down its length, whereas exonucleases digest DNA from either end. On the other hand, sequence

dependent restriction endonucleases will only cleave the DNA if a specific sequence is present together with an anticipated set of DNA fragments, the sizes of which can be identified. Single-stranded DNA or RNA is not cleaved by restriction endonucleases. RNases would selectively destroy RNA, whether endonucleolytically or exonucleolytically.Polymerases, with the exception of terminal deoxynucleotidyltransferases, create polynucleotides using a template, whereas ligases combine two DNA molecules separated by a nick. Phosphatases remove terminal phosphates from DNA, RNA, and oligonucleotides, whereas polynucleotide kinase adds phosphates to these molecules. The capacity of these enzymes to produce specific alterations to DNA molecules in the test tube is the foundation of recombinant DNA technology.

In biotechnology, restriction enzymes are invaluable instruments, with several varieties possessing distinct characteristics and uses. Type I, III, and IV enzymes offer insights into intricate DNA interactions and alterations, but Type II enzymes are most frequently utilized due to their straightforward and accurate cutting process. For the advancement of genetic research and biotechnology, it is essential to comprehend the diversity and activity of these enzymes. DNA modifying enzymes that alter DNA are crucial for preserving the accuracy and expression of genetic information. With great promise for use in medicine and agriculture, the discovery of novel DNA-modifying enzymes and technologies, including CRISPR-Cas, has opened up new directions in genetic research. However, the potential advantages of the technology must be carefully balanced against the ethical and safety concerns associated with gene editing.

3.9. Terminal questions

Q.1. What are enzymes? Discuss the types of enzymes use in prokaryotic cell division.

Answer:----------

Q.2. What are the restriction enzymes? Discuss the role of restriction enzymes in DNA repair.

Answer:----------

Q.3. Discuss about different types of restriction enzymes

Answer:----------

Q.4. Discuss about Uses and Restriction mapping of restriction enzyme type –II.

Answer:----------

Q.5. What are DNA Modifying enzymes? Discuss its role in DNA and RNA.

Answer:----------

Q.6. Briefly discuss about Nuclease and its types,

Answer:----------

Q.7. Polymerase and its role in DNA medication and repair.

Answer:----------

Q.8. What are end modifying enzymes discuss it function and significance.

Answer:----------

3.10. Further suggested readings

1. Robert Schlei, Genetics and Molecular Biology,2nd Edition.
2. McGraw-Hill, Cell and Molecular Biology, Human Genetics: Concepts and Application, 9th Edition.
3. Desmond S. T. Nicholl, An Introduction to Genetic Engineering Third Edition, University of the West of Scotland, Paisley, UK.
4. Robert Schleif, Genetics and Molecular Biology, second edition, The Johns Hopkins University Baltimore, Maryland.
5. NPTEL – Bio Technology – Genetic Engineering & Applications.
6. T.A. Brown, Gene Cloning and DNA Analysis an Introduction, Sixth Edition, University of Manchester Manchester.

Uttar Pradesh Rajarshi Tandon
Open University

PGBCH - 118 N
Genetic Engineering

Block- 2

Genetic Engineering- I

Course Design Committee

Dr. (Prof.) Ashutosh Gupta, School of Science, UPRTOU, Prayagraj	Chairman
Prof. Prof. Umesh Nath Tripathi Department of chemistry Deen Dayal Upadhyaya Gorakhpur University, Gorakhpur	Member
Prof. S.I. Rizvi Department of Biochemistry University of Allahabad, Prayagraj	Member
Prof. Dinesh Yadav Department of Biotechnology Deen Dayal Upadhyaya Gorakhpur University, Gorakhpur	Member
Prof. Sharad Kumar Mishra Department of Biotechnology Deen Dayal Upadhyaya Gorakhpur University, Gorakhpur	Member
Dr. Ravindra Pratap Singh Assistant Professor (Biochemistry) School of Science, UPRTOU, Prayagraj	Member
Dr. Dharmveer Singh Assistant Professor (Biochemistry) School of Science, UPRTOU, Prayagraj	Course Coordinator

Course Preparation Committee

Dr. Gopal Dixit Assistant Professor Department of Botany, Upadhi Mahavidyalaya, Pilibhit, U.P.	**Author**	**Block-1-2**	Unit: 1-5
Dr. Arun Kumar Pandey Assistant Professor Department of Botany, PSMPG College, Maharajganj. U.P.	**Author**	**Block-2-3**	Unit: 6-9
Dr. Sadhana Singh Assistant Professor- Biochemistry School of Sciences, UPRTOU, Prayagraj	**Author**	**Block-4**	Unit: 10, 12
Dr. Anuradha Singh Assistant Professor- Botany School of Sciences, UPRTOU, Prayagraj	**Author**	**Block-4**	Unit: 11
Dr. Mohd. Khalid Masroor Retd. Associate Professor-Botany, University of Allahabad, U.P.	**Editor**	(Block- 01, 02, 03&04, Unit: 1, 2,3,4,5, 10, 11, &12)	
Dr. Rajiv Ranjan Associate Professor, MLKPG College, Balrampur, U.P.	**Editor**	(Block- 02& 03)	Unit: (6-9)
Dr. Dharmveer Singh Assistant Professor (Biochemistry) School of Sciences, UPRTOU, Prayagraj			**(SLM & Course Coordinator)**

PGBCH – 118, Genetic Engineering
©UPRTOU , 2024
ISBN :

Printed By: K.C.Printing & Allied Works, Panchwati, Mathura -281003.

Introduction

The following three units are included in the first block of genetic engineering are as:

Unit-4: This unit covers the Cloning vectors include pBR322, pUC8, Lambda (λ)-phage vector, M13 phage, cosmids, phasmids, shuttle vectors, and artificial chromosomes for bacteria, yeast, and mammals. BACs and YACs are employed for large DNA fragments, whereas MACs aid in mammalian cell cloning. These vectors allow for effective gene cloning, modification, and expression in a variety of organisms, as explained.

Unit-5:

This unit covers the applications of cDNA include synthesis from mRNA. Cloning cDNA in plasmid or bactriophages vectors. Expression of cloned cDNA molecules, cDNA library. cDNA is synthesized from mRNA using reverse transcription. The cDNA is then cloned into plasmid or bacteriophage vectors for propagation. These vectors enable the expression of cloned cDNA in host cells. A cDNA library, consisting of diverse cDNA clones, represents the mRNA population of a cell, facilitating gene discovery and study.

Unit-6: This unit cover the cloning form genetic DNA. Cloning is a technique scientists use to make exact genetic copies of living things. Genes, cells, tissues, and even whole animals can all be cloned. Genomic DNA libraries consist of DNA fragments that have been cloned into vectors. To create DNA fragments for cloning, genomic DNA is fragmented and placed into vectors. Ligation links these fragments to vectors. The libraries are subsequently packaged into phages or plasmids, which are amplified in host cells for further examination, are briefly discussed in this unit.

4.1. Introduction

In molecular biology, cloning vectors are crucial instruments for replicating and modifying DNA fragments inside host cells. By acting as carriers, these vectors enable scientists to clone, examine, and modify genes and genomic sequences. A cloning vector is a small piece of DNA that can be stably maintained in an organism and into which a foreign DNA fragment can be inserted for cloning purpose. The cloning vector may be DNA taken form a virus, the cell of a higher organisation, or it may be the plasmid of a bacterium. The vector contains features that allow for the convenient insertion of a DNA fragment into the vector or its removal from the vectors. Cloning vectors are available in a variety of forms, each tailored to fit a particular need and host organism. They are essential for genetic research, biotechnology, and medicine because of peculiarities in their design that make DNA insertion, replication, and selection easier. The cloning vectors have many characteristics, such as element expression, numerous cloning sites, selectable markers, and replication origin. A DNA

sequence known as the replication origin gives the vector the ability to reproduce on its own inside a host cell. This function makes sure that every time the host cell divides, the vector and any added DNA are replicated. Replication origins of various vectors are tailored to certain hosts, like yeast or E. Coli. Genes known as selectable markers are those included into the vector that give resistance to particular drugs or permit growth under particular circumstances. Researchers can identify cells that have successfully incorporated the vector by using these markers. Auxotrophic indicators in yeast vectors or antibiotic resistance genes like ampicillin resistance in plasmid vectors are examples of common selectable markers. DNA fragments within host cells. The Multiple Cloning Site (MCS), which is often referred to as a polylinker, is a section of the vector that has several distinct sites for restriction enzyme recognition. This makes it possible for foreign DNA segments to be easily inserted into the vector. Numerous different DNA sequences can be cloned more easily thanks to the MCS. In order to induce the transcription of the inserted gene, promoters, enhancers, and terminators are added to vectors used in gene expression experiments. These components guarantee that the gene is adequately controlled and expressed at the desired level in the host cell. There are various types of cloning vectors available, including plasmid vectors, cosmids and phagemids vectors, bacterial artificial chromosomes (BACs), yeast artificial chromosomes (YACs), and mammalian artificial chromosomes (MACs). Small, circular DNA molecules known as plasmid vectors replicate on their own in bacterial cells. Because of their effectiveness and simplicity, they are among the most often used cloning vectors. Plasmids are perfect for cloning small to medium-sized DNA fragments since they can hold inserts up to 15 kb in length. Vectors such as pBR322 and pUC are examples.

In molecular cloning, vector is any practices E.g. plasmids, cosmids, Lambda phase) used as a vehicle to artificially carry a foreign nucleic sequence usually DNA into another cell, when it can be replicated and /or expressed. A vector containing foreign DNA is termed recombinant DNA. Most commonly used vectors are plasmids. The vectors itself generally carries a DNA sequence that consist of an inserted and a large sequence that served as the "backbone" of vectors. The purpose of the vectors which transfer the genetic information to another cell is typically to isolate multiply or express the insert in the target cell. All vectors may be used for cloning and are therefore cloning vectors, but they are also vectors designed especially for cloning, while others may be designed specifically for others purpose, such as transcription and protein expression.

Vectors designed specifically for the expression of the transgene in the target cell are called expression vectors, and gradually have a promoter sequence that drives expression of the transgene. Simple vectors only transcribe not translate.

Bacterial Artificial Chromosomes (BACs) are designed to clone huge DNA segments ranging in size from 100 kb to 300 kb. They include elements from the F plasmid, such as a low-copy number replication origin, centromere, and telomere. Because of their ability to keep massive inserts stable, BACs are utilized in genome mapping and sequencing programs like the Human Genome Project. Yeast Artificial Chromosomes (YACs) are vectors that replicate and operate like real chromosomes inside yeast cells. They can hold much larger DNA pieces, ranging from 200 kb to 2 Mb. YACs have centromeres, telomeres, and a replication origin, just like yeast chromosomes. They're employed to clone big genes and build genomic libraries. Similar to YACs, Mammalian Artificial Chromosomes (MACs) are made specifically for usage in mammalian cells. Their replication origins, telomeres, and centromeres enable them to replicate and segregate similarly to the normal chromosomes seen in mammalian cells. Large DNA segments can be carried by MACs, which are useful for gene therapy, functional genomics, and gene cloning. Bacteriophages, which are viruses that infect bacteria, are the source of phage vectors. It is possible to package and introduce DNA pieces into bacterial cells using these vectors. M13 phage vectors and λ-phage vectors are two examples. Phage vectors can be used to produce single-stranded DNA for sequencing and for high-efficiency cloning. DNA pieces up to 45 kb can be cloned using cosmids, hybrid vectors that combine plasmid and λ-phage components. Applications for phage display and plasmid replication are made possible by phagemids, which are plasmid vectors containing characteristics of bacteriophages.

The functions of the cloning vectors utilized are functional genomics, gene expression, and gene cloning. Cloning vectors are essential to contemporary molecular biology because they make it possible to work with DNA and investigate genes in a variety of biological systems. Their adaptability and unique qualities allow scientists to thoroughly examine genetic material, resulting in advances in biotechnology, medicine, and genetic research.

Objectives:

After reading this unit, the learner will able to know

- the plasmids vectors and its application in genetic engineering
- the about the pbr322 and puc8vectrs
- about lambda (λ)-phage vector and m13 phage
- about cosmids, phasmids and shuttle vector

4.2. Plasmid:

A plasmid is a small extra chromosomal DNA molecule within cell that is physically separated forms chromosomal DNA and can replicate. They can most commonly found as small circular double-stranded DNA molecules in and eukaryotic organism. Plasmid often carry useful gene such as for antibiotic resistance. Plasmids are usually very small and contain additional gene for special circumstance.

A plasmid is a little circular piece of double-stranded DNA that is found inside the cell but does not bind to or separate from the DNA of the chromosome. The plasmid has an independent replication origin that it uses to determine what constitutes real replication. Nevertheless, chromosomal DNA has no influence over plasmid replication. Working on the conjugation process in the 1950s, it was discovered that a transmissible genetic factor determines a bacteria's maleness. Every female bacterium that conjugates with another female becomes a male. The F (fertility) factor is a hereditary characteristic of males that is passed down through cell-to-cell interaction. F is a distinct genetic element as a result. This element's genetic name, plasmid, was first used in 1952 by J. Lederberg. Therefore, in addition to a bacterial cell's chromosome, plasmids are small circular DNA molecules that replicate themselves and are double stranded. During cell division, it multiplies on its own and is inherited by both daughter cells.

Therefore, the bacterial chromosome does not control its operation. The word "episome" was originally used in 1960 by Jacob Schaeffer and Wollman to refer to the extrachromosomal genetic material that integrated the bacterial chromosome during replication.

Per bacterial cell, there can be one, hundreds, or even more plasmids. Five to one hundred genes that control various biological processes are found on a plasmid. Extra chromosomal genetic material called plasmids is present in bacteria but is not required for growth and does not exist in an extracellular form. Round, little, and capable of self-replication are plasmids. These circular DNA fragments are autonomous, free-floating, and able to replicate within the host cell. Natural plasmids can be present in some eukaryotic cells as well as bacteria. They are distinct from chromosomal DNA in eukaryotes. Plasmids are autonomous DNA circles that replicate

on their own, separate from the chromosomal genome of bacteria. Genes that are typically not necessary for cell development or survival are found on plasmids.

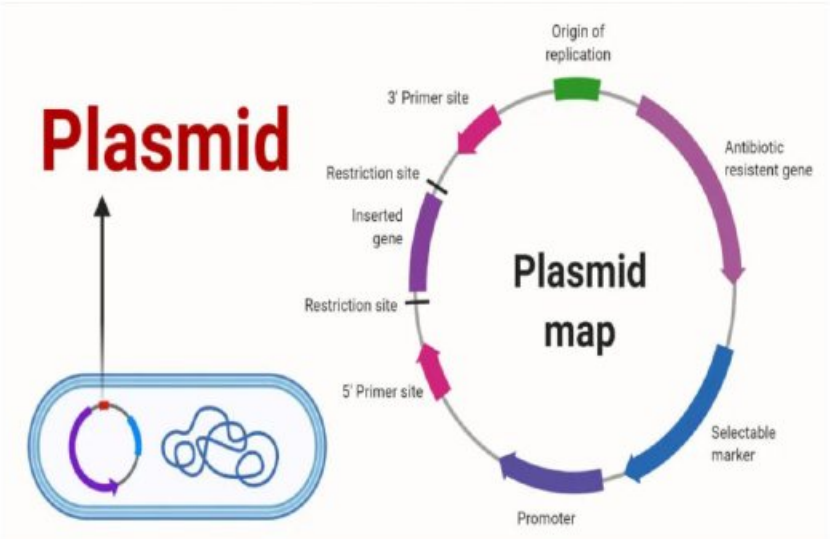
Certain plasmids can be intentionally created in the lab, integrate into the host DNA, and act as cloning vectors. Plasmids are genetic units that can multiply and function steadily without being necessary for the basic operations of the host cell, which are primarily bacteria. They are physically distinct from the host cell's chromosome. A plasmid is a DNA structure found in cells that can exist and multiply without the assistance of chromosomes. It seems that plasmids impede gene activation in a variety of investigated organisms. Genes that offer genetic advantages, including antibiotic resistance, can be found on some plasmids. They could have between 1,000 and several hundred thousand base pairs of DNA. The majority of them are present in bacteria. Plants and yeast are examples of eukaryotes that have plasmids. Plasmids are utilized as cloning vectors in recombinant DNA technology for gene transfer and manipulation because of their capacity for autonomous replication. Plasmids contain a large number of bacteria's antibiotic-resistant genes.

Plasmids are found in many different prokaryotes, ranging in size from 1 to more than 200 kbp, and are typically not necessary for cell division and growth. The E. coli F plasmid has an average size, about 1% smaller than the E. coli at-chromosome. Additionally, plasmids are moved from one bacterial cell to another by conjugation. When bacterial cells are given plasmids encoding a particular gene, they proliferate quickly and generate more of the required DNA fragment.

Although they are less frequent than in bacteria, plasmids are also present in some higher organisms. For instance, the 2-μ circle or 2-μm plasmid, a circular ds DNA plasmid, is present in the majority of yeast strains. Chromosome DNA duplicates during bacterial cell division, and the cell's replication mechanism also replicates plasmid DNA. The plasmid copies are distributed equally across the daughter cells and must be able to reproduce independently of the host chromosome. The cell is not harmed by the plasmid's multiplication. The plasmid is regarded as a real replicon since it bears its own origin of replication (ori). A unique DNA segment linked to cis-acting regulatory regions is the source of replication for a 1s plasmid. Plasmids depend on the host cell for various enzymatic functions, raw materials, and energy needs. Therefore, the smallest segment of plasmid DNA that may replicate independently and retain a normal copy number is known as a plasmid replicon.

Based on their replication, transferability, and function, plasmids can be categorized. For instance, conjugative plasmids have genes that facilitate the development of a pilus, a structure that

facilitates the conjugative transfer of plasmids from one bacterial cell to another. The quick dissemination of genetic features both within and between bacterial populations is made possible by this horizontal gene transfer mechanism. If a conjugative plasmid is also co-infected with the host cell, non-conjugative plasmids, which are devoid of these genes, can still be transferred.



Source: Plasmids- Definition, Properties, Structure, Types, Functions, Examples (microbenotes.com)

Plasmids are used in genetic engineering as vectors to introduce new genes into cells. To create proteins or study gene activities, researchers can insert a desired gene into a plasmid, which can then be injected into bacteria or other cells. There are two forms of plasmid replication, rigorous and relaxed, and both are tightly controlled.

There is a limited quantity of stringent plasmid copies per cell because they multiply in tandem with the host cell's genome. On the other hand, several copies of relaxed plasmids can replicate within a host cell without being influenced by the chromosomal replication cycle. This copy number can impact plasmid-borne gene expression, which in turn can impact the phenotypic traits of the host. In genetic engineering and biotechnology, plasmids are very helpful cloning vectors.

Researchers have the ability to clone foreign DNA sequences into plasmids, which are then inserted into host cells mostly bacteria. Using this method, vast amounts of proteins, enzymes, and other biomolecules that are encoded by the added genes can be produced. Additionally, unique regulatory elements can be added to plasmids to enable targeted gene expression in the host organism. Additionally, plasmids are critical to the study of gene regulation and function as well as the advancement of recombinant DNA technology. Plasmid-based systems, for instance, have been

applied to synthetic biology, gene therapy, and vaccination development. Plasmids, for instance, have the ability to transfer therapeutic genes to target cells during gene therapy, possibly correcting genetic abnormalities.

Structure of plasmid

Circular, double-stranded DNA molecules that are relatively tiny and exist independently of chromosomal DNA in cells are called plasmids. Although they are mostly found in bacteria, they can also be found in eukaryotes and archaea. Independently replicating from the chromosome, plasmids frequently harbor genes that bestow benefits onto the host cell, such resistance to antibiotics or the ability to metabolize unusual compounds. A typical plasmid consists of the following essential elements:

Origin of Replication (ori)

This particular DNA sequence is the starting point for replication. The copy number (the number of copies of the plasmid within a cell) and replication process of the plasmid are determined by the ori. It's necessary for the plasmid to multiply without relying on the chromosomal DNA of the host.

Selectable Marker Genes

These genes allow cells carrying the plasmid to be recognized and chosen. Antibiotic resistance is provided by common selectable markers. Antibiotic resistance to tetracycline, kanamycin, or ampicillin is a common sign. When the matching antibiotic is present, cells that have picked up the plasmid can survive; cells that do not contain the plasmid cannot.

Cloning Site

A site where foreign DNA pieces can be introduced. Scientists use this location to insert additional genes into plasmids for experiments. The MCS, also known as a polylinker, is a short DNA tract that contains numerous distinct restriction enzyme recognition sites. This area enables the simple insertion of foreign DNA segments into the plasmid using particular enzymes, allowing cloning and genetic manipulation.

Promoter Region

This is a sequence that triggers the transcription of a downstream gene. In designed plasmids, the promoter is frequently chosen based on the intended expression system and can be inducible or constitutive, depending on the requirement for controlled gene production.

Other genes: Plasmids can contain a variety of genes that help the host cell, such as those that encode enzymes for toxin degradation, antibiotic synthesis, or survival in hostile environments, as well as reporter genes.

The reporter genes express easily visible features, such as fluorescence (e.g., GFP green fluorescent protein) or color change (e.g., LacZ, which turns blue in the presence of X-gal). Reporter genes assist in monitoring the expression and location of plasmid-encoded genes. Based on the size and function of the plasmid, the number of genes carried might vary from a few thousand to hundreds of thousands of base pairs.

Additionally, due to damage or enzymatic action, plasmids can exist in the following physical forms: supercoiled, relaxed, or linear. The most efficient and compact form for expression and replication is the supercoiled one. Plasmids are useful tools in genetic engineering, synthetic biology, and molecular biology because of their structural flexibility and simplicity. Their ability to transport particular genes or regulatory elements can be designed, enabling precise control over the expression of genes in a variety of host organisms.

Types of plasmids:

On the basis of function plasmids are divided into several types.

Sex factor or fertility (F) factor: Sex factor, or F factor, refers to the plasmids of male cells that give their host cells the ability to transfer chromosomal markers but not the other characteristics. Conjugative plasmids with tra (for transfer) genes are called F plasmids. The bacterial DNA sequence that enables a bacterium to make the pilus required for conjugation is known as the fertility factor, commonly referred to as the F factor or sex factor. It has several other genetic sequences that are in charge of replication, incompatibility, and other processes, in addition to clusters of tra genes. The F factor is an episome, meaning that it has the potential to integrate into the bacterial genome and persist in this form through multiple cell divisions, but eventually

R (resistance) plasmids: In 1946, Shigella caused multiple outbreaks of diarrhea in Japan, leading to the development of antibiotic treatment resistance. A strain of S. dysenteriae that causes dysentery was identified in Japan in 1955 and proved to be resistant to four medications: tetracycline, chloramphenicol, streptomycin, and sulfanilamide. These plasmids contain genes that provide resistance against disinfectants, antibiotics, and other potentially dangerous compounds. In terms of public health, the proliferation of resistant plasmids is a serious concern.

Virulence Plasmids: These plasmids carry genes that make bacteria more harmful. They might encode proteins that aid in the adherence of bacteria to host cells or toxins. Genes for different traits are carried by virulence plasmids, which facilitate the infection of higher species by bacteria, including people and plants. There are several ways in which bacteria might infect the target organism. Virulence plasmid-carrying bacteria can infect animal cells by adhering to their membranes and secreting poisons (Figure 8.19). The genes that encode protein filaments, or adhesins, are found on the plasmids. Adhesins resemble pili in general but differ in length and thickness. They connect to target cells via identifying their cell-surface receptors. The bacteria release toxins after adhesion, which have the ability to enter the target cell and kill it. As an illustration, harmful E. Coli strains typically depend on virulence factors carried by plasmids. Different pathogenic strains of Escherichia coli have a wide variety of toxins, such as heat-labile enterotoxin (similar to choleratoxin), heat-stable enterotoxin, hemolysin (lysed red blood cells), and Shiga-like toxin (similar to the toxin of Shigella, which causes dysentery). On the other hand, virulence plasmids produce comparable adhesins, often known as "colonization factors," that help germs adhere to mammalian cells. Tumor-inducing plasmid (pTi) of A. tumefaciens is another well-known example of a pathogenic plasmid that creates crown gall tumors on injured plant cells.

Degradative Plasmids: These plasmids contain genes that generate enzymes that enable bacteria to break down odd substances like pesticides and pollution. A lot of research has been conducted on P demonas' degradative plasmid. Through their unique metabolic pathways, the pseudomonads have been found to catalyze a variety of rare complex orga chemicals. Scientist Anand Mohan Chakrabarty, who was born in India and raised in the United States, has extracted plasmids from several Pseudomonas putida cultures that have the ability to use a variety of complex organic compounds, including 3-chlorobenz biphenyls and 2,4-D salicylate.

Col Plasmids: Numerous bacterial strains are capable of producing bacteriocin, a proteinaceous toxin that is fatal to other strains in the same genus. The sensitive cells are killed by an

Ecoltare tosinserted. Coficin synthesis is carried out by the Col factor-containing plasmids found in & café cells. Many Col plasmide cha E. coli exist, but they are all based on bacterial chromosomes. Col plasmids are carriers of genes that produce bacteriocins, which are proteins that confer a competitive advantage to the host bacterium by killing competing bacteria.

Heavy-metal resistance plasmids: Numerous bacterial strains possess genetic resistance to heavy metals, such as Hg++, Ag+, Cd++, Co++, CrO4, Cu++, Ni++, Pb+++, and Zn++, among others.

Cryptic plasmids: Every bacterium was found to contain a low molecular weight DNA plasmid during the process of isolating plasmid DNA from a large number of bacteria. They are non-functional since they lack any genes. The term "crypos plasmid" refers to these plasmids. Plasmids appear to be the norm rather than the exception. These plasmid types most likely carry genes whose properties are currently unknown. The yeast 2-μm plasmid, for instance, does not include any genes that may give any phenotypic features. The mauriceville plasmid, a unique retroelement discovered in the mitochondria of some Neurospora strains, is another instance of a cryptic plasmid. This 3.6 kbp closed circular ds DNA plasmid, known as a mitochondrial retroplasmid (mRP), contains the reverse transcriptase gene.

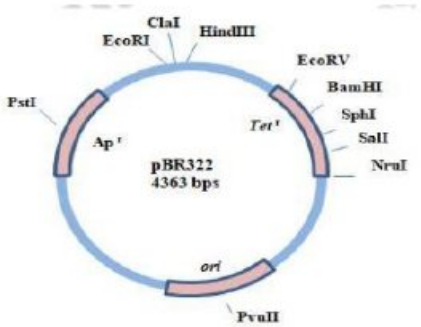
4.3. pBR322

It is plasmid and wasone of the first widely used E. coli cloning rectors Created in 1977 m the laboratory of Herbert Boyer. IT was named after Francisco Bolivar Zapata. The P Stands for "plasmid" and BR for Bolivar and Rodriguez. It has two antibiotic resistance genes, the gene between encoding the ampicillin resistance (AmpR) proteins and the gene tetA encoding the tetracycline resistance (Tet^R) protein.

One of the most used plasmids in genetic engineering and molecular biology is pBR322. The acronym pBR322 stands for Rodriguez, Bolivar, and Bolivar; 322 is the number identifier. With so many uses in cloning, pBR322 is a widely used cloning vector in E. coli. Bolivar and Rodriguez created it in 1977, and it's a prime example of a cloning vector. With its length of roughly 4,361 base pairs and other essential characteristics, the plasmid pBR322 is an adaptable instrument for DNA manipulation.A fundamental instrument in the realm of molecular biology is pBR322. Since it includes recognition sites for restriction enzyme activity and antibiotic resistance genes as a selectable marker, pBR322 is a helpful vector. It was among the first plasmids to be widely employed in cloning, which

paved the way for the creation of more sophisticated vectors. Numerous other plasmids and vectors used in research and biotechnology today were created using its design ideas and features. pBR322 plasmid vector has the following elements:

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



The structure of pBR322 contain

- **Origin of Replication (ori):**

The ColE1 plasmid is the source of replication in pBR322. This particular sequence facilitates the independent replication of the plasmid in a bacterial host, usually Escherichia coli (E. coli). The plasmid can be present in the cell in a moderate number of copies thanks to the ColE1 origin. Plasmid pBR322's replication origin is also referred to as pMB1. This plasmid has a copy number of 15-20.

- **Restriction enzyme Sites:**

The sequence of pBR322 includes multiple distinct restriction enzyme sites, such as those for EcoRI, PstI, BamHI, SalI, and HindIII. There are about forty distinct restriction sites in the pBR322 genome. These locations allow foreign DNA pieces to be inserted; they are found inside the antibiotic resistance genes. The tetracycline resistance region has around 11 distinct restriction sites. When foreign DNA is inserted into a restriction site in an antibiotic resistance gene, insertional inactivation may result. For instance, the ampicillin resistance gene's activity can be interfered with by introducing

DNA into the PstI site, rendering the cell ampicillin-sensitive. This offers a method for finding viable recombinants.

- **Selectable marker sites:**

This plasmid's genome contains two genes that are either antibiotic resistance genes or selectable marker sites.

- Ampicillin resistance site – the ampicillin gene codes for β-lactamase, which can be used for screening microorganisms when a foreign DNA is being inserted in the plasmid.
- Tetracycline resistance site – this gene degrades the antibiotic tetracycline and can be used for screening microorganisms.

These antibiotic resistance genes are useful in screening organisms after cloning.

Foreign genes have been inserted and expressed in E. coli through the extensive usage of pBR322 as a cloning vector. It's an easy tool for gene cloning research because of its well-characterized structure and abundance of cloning sites. pBR322 is used by researchers to map genes and investigate gene function through the introduction and analysis of different DNA fragments. The study of gene function and regulation is made easier by the plasmid's distinctive restriction sites and markers for antibiotic resistance, which enable the production of mutations within particular genes.

4.4. pUC8

PUC 8 is a derivative of plasmid pBR322 and filamentous phase M13. Plasmid puc 8:15 bears a 2.3 kb Vitreoscill a chromosomal fragment (Hind III Hind III) containing the Vitreoscilla hemoglobin gene (vgb).

pUC18 plasmids are 2686 bp in size, high copy number plasmids. In molecular biology, the pUC8 plasmid vector is frequently utilized for cloning and expressing genes in Escherichia coli (E. coli). It is a member of the high-copy-number pUC plasmid family, which is designed to carry a range of genetic components beneficial for expression and cloning research. The acronym pUC8 represents plasmid, University of California, the location of its construction, and the order of discovery within the same laboratory. It is developed from pBR322, with only minor modifications. The ampicillin resistance gene and replication origin in pUC8 and pBR322 are identical. The ampicillin resistance

gene's unique recognition sequence has been lost due to changes in only a few nucleotide sequences. Structures have been changed to include a cluster of recognition sites in the LacZ' gene rather than the tetracycline resistance gene. The University of California's Messing and associates created this set of cloning vectors. The pUC series of plasmids was created by Vieira and Messing (1982) and is derived from pBR322. This vector has the name of the University of California, where they conducted their initial preparation. These plasmids include the amp' gene, have a high copy number, and are around 2,700 bp length. The E. coli chromosome is the source of the lac Z' gene, and the replication origin is taken from pBR322 (Fig. 4.1).

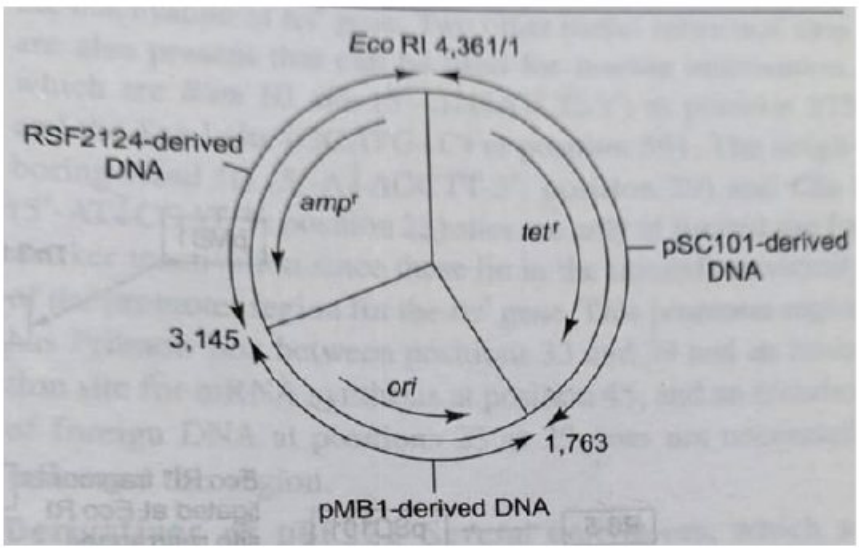


Fig. 4.1: pUC18 plasmids

The lac operator region (Olac) and the lac promoter (Plac), an inducible promoter, are also included in this family of vectors. In addition to these, the pUC series of vectors has a few more components that give this vector better attributes. The following are these components: The lac repressor protein is encoded by the lac I gene. Lac repressor is produced by the plasmid to augment host levels. Since the plasmid has a large copy number, it's possible that the host lac autorepressor levels won't be enough to effectively repress the lac operator on the plasmids. The lac Z' gene, which results in the production of the N-terminal portion of the B-galactosidase protein, is transcribed from the lac promoter. The tetr gene section of pBR322 has been substituted with a gene segment that codes for a portion of the κ-galactosidase enzyme, into which a number of specially created restriction sites, known as MCS or polylinkers, have been inserted. The pUC series, which includes pUC8, is still an important tool in labs all over the world and contributed significantly to the development of molecular cloning methods. These characteristics apply to the pUC8 vector.

- **High copy number**: A fortuitous mutation within the replication origin during pUC8 synthesis led to a copy number of 500–700 even prior to replication.
- **Size:** Its length is 2750 bp, and the cloning vector should not be larger than 10 kb. The plasmid's modest size makes it easy to insert huge amounts of foreign DNA into the vector.
- **Restriction site**: presence of a cluster of RS in the LacZ gene, which gives the vector more versatility in the kinds of DNA fragments it may insert.
- **One step identification of recombinant:** We may obtain our recombinant colony by putting ampicillin and x-gal into medium and letting the culture grow on it.

Because of its large copy number and effective transformation, pUC8 is frequently employed for the cloning and subcloning of DNA fragments. The insertion of different DNA sequences is made easier by the MCS's many restriction sites. Recombinant clones can be easily identified with the lacZ gene and the blue/white screening approach. Recombinant DNA technology and genetic engineering both make extensive use of this technique. While pUC8 is not an expression vector by itself, it can be utilized in conjunction with other components to investigate the expression of specific genes. It is frequently employed to multiply DNA fragments that will subsequently be introduced into expression vectors.

4.5. Lambda (λ)-phage vectors

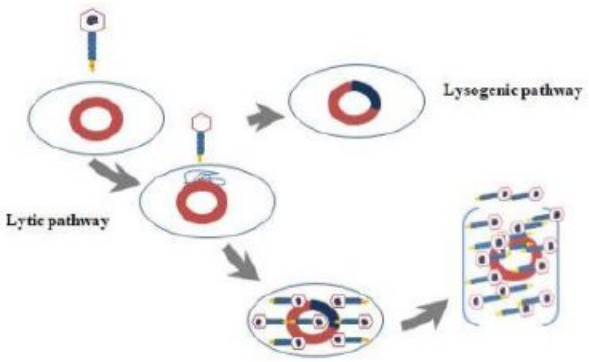
It is a type of vector commonly used in molecular biology for constrictive cDNA libraries and seeing by differential hybridization. It differs high transformation efficiencies and reproducibility, contains 48490 bp genome that contains two 5 overhangs to circularize the phage genome once it has entered a bacterial cytoplasm. The lambda phage genome encodes many proteins essential for viral replication, but also recombination.

The proteins involved in the latter process, the annealase and exonuclease, have also key tools in molecular cloning. The lambda phage exomuclease digests dsDNA in a5' to3' direction with a catalytic rate that has been studied using a variety of single molecule approaches.

Bacteriophage-based cloning vectors called lambda (λ)-phage vectors are developed from the bacteriophage lambda virus, which infects Escherichia coli (E. coli). Lambda phages are approximately 50 kilobase double-stranded DNA phages that may replicate in both the lytic and lysogenic modes in E. Coli. In molecular biology, these vectors are frequently employed for the

cloning of large DNA segments, the creation of genomic libraries, and other genetic manipulation methods. The advantage of lambda phage vectors is their high transformation efficiency since they use high-quality, readily available packing extracts;

A bacteriophase labda simply attaches itself to the host bacterium's cell membrane and injects its DNA into it. The linear double-stranded DNA molecule entered the host cell and cycled through the cos sites at both ends to generate a circular DNA molecule that resembled a plasmid. Following circularization, it chooses between its two life cycles. It integrates with the bacterial genome throughout the isogenic cycle, multiplies, and is transferred with the bacterial genome. Using the machinery of the host cell, it replicates while in the lytic cycle and continues until the bacteria dies from an excess of virons.



Structure of Lambda Phage

1. **Head (Capsid)**: The double-stranded DNA genome of the phage is enclosed in an icosahedral protein shell called the head. The genome of the wild-type λ-phage is roughly 48.5 kilobases (kb) long.
2. **Tail:** The flexible tail makes it easier for the phage DNA to be injected into the host cell by adhering to the surface of the bacterial host.
3. **Cos Sites:** The short, cohesive ends of the λ DNA are called cos sites, and they enable the circularization of linear DNA when it enters the host cell. They play a vital role in encasing the DNA within phage particles

Features of Lambda Phage Vectors

- **Large DNA Insert Capacity**

Lambda vectors are capable of holding huge foreign DNA inserts, usually up to 20 kb. Their ability to clone big genes and build genomic libraries makes them perfect.

- **Replacement Vectors**

Replacement vectors, or insertionional vectors, are modified λ-phages that permit the insertion of foreign DNA. They include non-essential sections that can have the desired DNA substituted for them, like the stuffer fragment. To create room for the insert, the center region between two cos sites is frequently deleted. Additionally, only DNA molecules falling into a particular size range—typically 38–51 kb—can be packaged into phage particles.

- **Cosmid Vectors**
- **Infection and Cloning Efficiency**

Cosmids are hybrids of plasmids and λ-phage vectors that incorporate the characteristics of both. They have cos sites for phage particle packaging, antibiotic resistance genes, and a plasmid origin of replication. Greater DNA fragment sizes (up to 45 kb) can be carried by cosmids than by conventional λ vectors.

- **Screening and Selection**

There are a number of techniques for screening recombinant phages, including plaque assays, in which a lawn of bacteria is covered in plaques, or clear zones. It is possible to differentiate recombinant plaques from non-recombinants using phenotypic markers or by looking for specific genetic elements missing.

The whole genome of an organism is broken up and cloned into a vector using lambda vectors, which are then used to construct genomic libraries for further screening and examination. Their purpose is to clone huge genes or several gene fragments, a task that is sometimes challenging with plasmid vectors because of their limited capacity. Using lambda vectors to clone and express genes in bacterial hosts, researchers may examine the function and control of genes. as cloning applications require a large number of clones, λ-phage vectors are helpful due to their high efficiency of phage infection as compared to plasmid transformation. Cosmids and lambda phage vectors have been crucial in furthering genetic research, especially in the areas of functional studies and genomics.

4.6. M13 Phage

The bactriophages known as M13 forms the basis of cloning systems designed to easily introduce. It also has been a filamentus bacteriophage which infects E.coli host. The M13 genome has the following characteristics:

Circular single stranded DNA

6400bp long

The genome codes for a total of 10 genes.

Gene VIII codes for the major structural protein of the bacteriophage particles

Gene III codes for the minor coat proteins.

The gene VIII proteins form a tubular array of approx 2700 identical subunits surrounding the viral genome.

Approximately five to eight copies of the gene III proteins are located at the ends of the filaments phage (genome plus gene III assembly)

Allows binding to bacterial "sex" pilus (pilus is bacterial surface structure of E.coli which harbor the F factors, extra chromosomal elements)

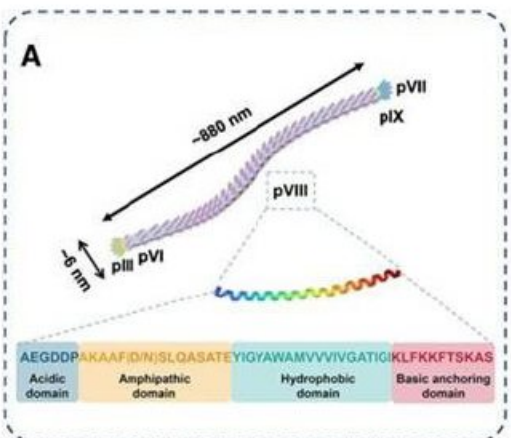
In molecular biology, the filamentous bacteriophage known as M13 phage is frequently utilized as a vector for cloning and expression when it comes to Escherichia coli (E. coli). The genome consists of approximately 6.4 kb of single-stranded circular DNA encased in a proteinaceous sheath. The plus (+) strand refers to the DNA strand found in phages. Once inside the E. Coli host, it uses the bacterial machinery to transform into a double-stranded DNA molecule known as the replicative form (RF). It is possible to obtain single-stranded or double-stranded M13 phage to use as a cloning vector. Similar to plasmid vectors, replicative form double stranded vectors are altered and replicated inside E. coli hosts. M13 phage can be gathered in order to isolate single-stranded vectors.

Phagemid vectors are plasmids that include a little fragment of the M-13, fd, or F1 filamentous phage, which may transport up to 10 kb of passenger DNA. Examples include the plasmids in the pEMBL series and the pBluescript family. When the bacteria is concurrently infected with an M13

helper phage, the M13 origin of replication permits the packaging of the plasmid into an M-13 phage. M13 has a non-lytic life cycle, which permits the continued production of offspring phage without causing host cell death, in contrast to other bacteriophages that lyse their host cells. Because of this characteristic, M13 is a desirable tool for a number of uses, such as the synthesis of single-stranded DNA and phage display. The rolling circle process replicates DNA in the M13 bacteriophage. Through this process, the DNA polymerase stretches free 3'OH after cleaving one strand. As the rising point circles the loop template, the circle's 3' end is enlarged. Single-stranded DNA generates a tail when the 5'end is misplaced. Cloning systems that are intended to quickly introduce mutations into genes incorporated into the phage genome are based on the bacteriophage referred to as "M13". Additionally, it has been employed in a number of "combinatorial" DNA and peptide libraries as well as "phage display" techniques.

Structure of M13 Phage

M13 bacteriophage is thin, having a diameter of around 6 nm and a length of around 880 nm. The M13 phage genome is packaged in five capsid proteins: pIII, pVI, pVII, pVIII, and pIX. These proteins help determine the length of the phage. Among these capsid proteins, pIII and pVI are at one end of the phage and are capped, while pVII and pIX are at the other end. Each minor capsid protein normally contains 3 to 5 copies. The most abundant capsid protein is pVIII, which accounts for 98% of the phage's total mass. It is composed of three domains, with around 2700 copies: an N-terminal domain (1–20 amino acids), an intermediate hydrophobic domain (21–39 amino acids), and a positively charged region (40–50 amino acids) that interacts electrostatically with the phage's genomic DNA. The electrostatic contact between the positively charged domain of pVIII and the phage DNA allows these pVIII proteins to arrange themselves in a helical pattern around the phage DNA, creating a negatively charged surface on the M13 phage.



Applications

1. **Phage Display**

Phage display technology is one of the M13 phage's most important uses. The process of phage display entails the surface expression of proteins or peptides on the phage particle. Researchers can display a protein on the phage surface by fusing a gene expressing a peptide or protein of interest with the gene encoding one of the phage's coat proteins (usually pIII). This method is used to select high-affinity binding molecules from huge libraries and to examine interactions between proteins, peptides, and DNA.

2. **Single-Stranded DNA Production**

The M13 phage is a naturally occurring source of single-stranded DNA that can be used for a number of molecular biology procedures, such as DNA sequencing, site-directed mutagenesis, and complementary DNA (cDNA) synthesis. Unlike most other cloning vectors, which produce double-stranded DNA, M13 produces single-stranded DNA.

3. **Non-Lytic Infection**

The M13 phage does not cause host cell lysis. Rather, it continually buds from the bacterial membrane, which keeps the host culture alive and permits the continued production of phage particles. This characteristic renders M13 advantageous in scenarios necessitating extended synthesis of recombinant proteins or peptides.

4. **M13 Derivatives**

For particular uses, multiple altered forms of M13, also referred to as M13 vectors or M13 phagemids, have been created. Combining components of phage vectors and plasmids, phagemids enable bacterial cell reproduction as plasmids and package as single-stranded DNA within phage particles. They are flexible instruments for expression and cloning because of their dual functionality.

5. **Library Screening**

A large number of variants (such as peptide libraries) may be included in M13 phage display libraries. These variants can be screened for binding to certain targets, making it possible to identify

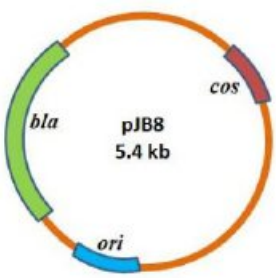
molecules with high affinity and specificity. The development of therapeutic antibodies and drug discovery both heavily rely on this technique.

4.7. Cosmids

A cosmid is a type of hybrid plasmid that contains a lambda phage cos sequence. Often used as cloning vectors in genetic engineering, cosmid can be used to build genomic libraries. They were first described by Collins and Hohn in 1978. They contained 37 to 52 kb of DNA, Limits based on the bateriophage size. They can replicate as plasmids if they have a suitable origin of replication e.g. SV40 ori in mammalian cell.

Cosmids are circular double stranded DNA (ds DNA) molecule like plasmids. However they have where the ds-DNA backbone is necked on each strand. The necking sites on the two strands are situated at 12 base pair apart making the circular molecules liner with 12 base single stranded DNA overhangs.

Cosmids are a type of hybrid plasmid-bacteriophage vector that is used to clone large DNA segments, usually between 30 and 45 kilobases (kb). They blend characteristics of bacteriophage lambda (λ) and plasmids. The word "cosmid" comes from the sequences known as cos sites found in bacteriophage λ that are necessary for encasing DNA in phage particles. Because cosmids replicate bacterially, they offer an extra advantage over bacteriophage λ-based cloning vectors. They are chimeric cloning vectors made of bacteriophage λ and bacterial plasmid segments. They have a nearby COS site that aids in circularization once they enter the host cell. Cosmids can be maintained in bacteria since they are the source of bacterial replication. Tetracycline, an antibiotic resistance gene, is also present, and it permits the selection of altered host cells. Additionally, DNA segments can be ligated into the distinct Multiple Cloning Site (MCS) region found on cosmids. Recombinant DNA can be directly utilized to infect E. Coli cells by packing it into freshly manufactured λ particles or virons.



pJB-8 is an illustration of a cosmid vector. Cosmids are plasmid vectors that have a selectable marker and a replication origin. Additionally, DNA fragments can be added to the unique restriction enzyme recognition site found in cosmids.

The key Features of Cosmids are such as:

- **Cos Sites**: The λ-phage genome contains particular DNA sequences called cos sites, which enable the packaging of DNA into phage particles. These regions are required for the λ-phage capsid proteins to identify and package the DNA during the in vitro packaging process. Large DNA segments can be efficiently packaged and transferred into host bacteria thanks to the cos sites present in cosmids.
- **Plasmid Elements**: Because cosmids have a plasmid origin of replication, they can multiply on their own inside the bacterial host, usually Escherichia coli (E. coli). Cosmids can proliferate and remain stable in bacterial cells because to this characteristic. Cosmids include antibiotic resistance genes that enable the selection of altered bacterial cells, such as those resistant to kanamycin or ampicillin.
- **Large Insert Capacity**: Cosmids can carry DNA pieces up to 45 kb in size, which is substantially more than what plasmids can manage. They are perfect for cloning big genes, gene clusters, or entire genomic regions because of their ability.
- **In Vitro Packaging and Transduction**: Cos sites aid in the packaging of DNA segments that have been cloned into cosmids into λ-phage particles in vitro. After infecting E. coli cells, these phage particles can introduce the cosmid DNA into the host. Particularly for big DNA inserts, this transduction approach is more effective than conventional transformation techniques.
- **Cloning and Expression**: Although cosmids are mainly utilized as cloning vectors for large DNA fragments, cloned genes can also be expressed using cosmids. Cosmids, on the other hand, usually lack potent promoters or other regulatory components for regulated gene expression, in contrast to expression vectors.

Genomic libraries, which are collections of DNA fragments that comprise an organism's whole genome, are commonly created using cosmids. Cosmids' enormous insert capacity makes it possible to clone extensive genomic sections, which facilitates the study of intricate genes and regulatory elements. Large-scale genome mapping and sequencing initiatives, including the Human Genome Project, have made use of cosmids. Their capacity to transport sizable DNA fragments makes it easier

to identify genetic loci and put together consecutive sequences. By cloning and examining sizable genomic areas, including genes and their regulatory components, cosmids are utilized by researchers to investigate the function and control of genes. Positional cloning, which involves cloning the necessary genomic material to determine the genetic cause of a characteristic or illness, can be facilitated using cosmids region.

4.8. Phasmids

The phagemid or phasmids is a hybrid between a phase and a plasmid, and it produces a cloning vector that grow as a plasmid in E. coli and is packaged single strand DNA is Viral particles delivery to M. tuberculosis replication as well as an F1 ori that enable a single stranded replication and packaging into phage particle. These plasmid introduced gene into the mycobacterium chromosome by site-specific recombination between phase and bacterial attachment sites. The recombination can be introduced into mycobacterium where the shuttle plasmid that occurs the genes of interest can lysogenize and be stably maintained as prophase. In 1987, Jacobs et al., contracted the first recombinant shuttle phagemid as chimeras containing mycobacteriaophage. DNA inserted into an E. coli cosmid. Those shuttle vectors permitted the first introduction of foreign DNA into m. smegmatis and mycobacterium Bovis BCG but not to M. tuberculosis.

Phasmids, sometimes referred to as phagemid vectors, are hybrid cloning vectors that incorporate bactriophages and plasmid characteristics. With a plasmid replication origin for double-stranded DNA isolation and a phage replication origin for single-stranded replication, phagemid vectors are adaptable instruments that are produced from both plasmids and coliphages. In order to infect and transducer, phagemid can be packaged into bactriophages particles or replicate as plasmids within a bacterial host. A Phagemid is a hybrid vector containing the A Phage DNA and the λ origin of replication from the plasmid. A linearized Plasmid DNA is inserted into cleaved A DNA to create it.

This is referred to as plasmid lifting. Homologous recombination between a DNA and the chromosomal DNA of E. coli occurs at this location. It facilitates the recombinant phagemid-containing phage particles' in vivo proliferation. The plasmid component in E. coli is what allows phage mid to exist independently as a plasmid. In E. coll., it might be freely liberated. For instance, AZAP

In phage display technology, peptides or proteins are shown on the surface of phage particles by the usage of phagamids. Researchers can display a protein on the phage surface by fusing a gene encoding a peptide or protein of interest to the coat protein gene of the phagemid. High-affinity ligands, antibodies, and other binding molecules are chosen from sizable libraries using this technique. Phasmids' capacity to synthesize single-stranded DNA is very helpful for procedures like DNA sequencing and site-directed mutagenesis. The ssDNA can be used in hybridization experiments or as a template for introducing particular mutations.

Genes can be cloned and expressed in bacterial hosts using plasmids. They facilitate the creation of recombinant proteins and the simple insertion and preservation of foreign DNA. Libraries of genetic variations can be produced using phagemid vectors and subsequently tested for desirable characteristics like binding affinity or enzymatic activity.

4.9. Shuttle vector

Shuttle vectors are the vectors that can replicate in prokaryotes as well as eukaryotes. The vector is initially identified and replicated in E. coli and is later transformed into yeast cell. For this purpose, the vector should contain the ORi sited of both the organism. There are different shuttle vector available in the case of yeasts. These includes yeast episomal plasmid (YEps), yeast integrative plasmids (γIPS) and yest centromeres plasmid (γ C.Ps) and yeast replecative plasmids (γ Rps). γIPs integrate genes of interest in yeast chromosome through homologues recombination.

Specialized plasmid vectors known as shuttle vectors are made to replicate in a variety of host species, usually prokaryotic (bacteria) and eukaryotic (yeast, mammalian cells, etc.). Their ability to perform two tasks, namely the cloning and modification of genes in one organism and their subsequent functional study in a different, makes them indispensable instruments for molecular biology research. In genetic engineering, gene cloning, and gene expression research, shuttle vectors are frequently employed. Shuttle vectors have two or more replication origins, each unique to a distinct host species. For instance, the Escherichia coli plasmid pUC (ColE1 origin) is a typical source of bacterial replication, but the 2μ plasmid may be the source of replication in yeast. The shuttle vector can reproduce in both bacterial and eukaryotic cells thanks to its dual origin. The shuttle vector's dual origin allows it to proliferate in both bacterial and eukaryotic cells. Shuttle vectors contain selectable marker genes for each host system. Common markers impart resistance to medicines like ampicillin and kanamycin in bacteria. Markers for eukaryotic cells include genes conferring antibiotic resistance,

such as G418 or hygromycin, as well as auxotrophic indicators, such as the yeast URA3 gene, which allow selection based on the ability to thrive under minimal conditions. A shuttle vector often includes a multiple cloning site (MCS) with numerous restriction enzyme sites for foreign DNA insertion. This feature makes it easier to insert and remove genetic material. Shuttle vectors can be used to manufacture proteins in a variety of host systems, allowing researchers to select the best system for creating a protein with certain features, such as precise folding and post-translational modifications. Shuttle vectors can be employed in functional genomics to investigate gene function by expressing or knocking down genes in different host systems. They are also utilized in gene complementation experiments to restore the function of a deficient gene in a mutant organism. Shuttle vectors, on the other hand, are valuable tools in genetic engineering for introducing novel features or functionalities into organisms, and they are widely utilized in biotechnology to produce therapeutic proteins, vaccines, and other physiologically active molecules.

4.10. Bacterial Artificial Chromosomes (BACs)

A bacterial artificial chromosome (BAC) in an engineered DNA molecule used to clone DNA sequence in bacterial cells (e.g. E.coli) BAC are often used in connection with DNA sequencing. Segments of an organisms DNA ranging from 100,000 to about 300,000 bp, can be inserted into BACs.

The BAC with then inserted DNA are taken up bacterial cells. As the bacterial DNA grow and drive, they amplify the BAC. DNA, while can then be isolated and used in sequencing DNA. A large piece of DNA can be engineered in a faison that allows into be propagated as a circular artificial chromosomes or BAC. Each BAC is a DNA clone containing roughly 100 to 300 thousands bp of cloned DNA. Because BAC is much smaller than the endogenous bacterial chromosome, it is straight forward to purify the BAC, DNA away from the rest of the bacterial cell's DNA and thus have the cloned DNA is a purified form.

Large cloning vectors called Bacterial Artificial Chromosomes (BACs) are used to clone and preserve large DNA pieces in E. coli hosts. These fragments typically range in size from 100 to 300 kilobases (kb). The replication origin (ori) of the F-plasmid is present in BAC vectors in a single copy. Compared to regular plasmid vectors, the F (fertility) plasmid vectors have a higher capacity due to its relatively large size. The F (fertility) factor in the F-plasmid regulates replication and keeps the number of copies low. Additionally, F+ bacteria (male) and F- bacteria (female) can conjugate in order

to transfer F-plasmid via pilus. Large-scale DNA sequence analysis is needed in functional genomics, genome sequencing initiatives, and other fields that make considerable use of BACs. A bacterial artificial chromosome's gene components are:

1. oriS, repE – F responsible for plasmid replication and regulation of copy number.
2. parA and parB for maintaining low copy number and avoiding two F plasmids in a single cell during cell division.
3. A selectable marker for antibiotic resistance; fue BACs also have blue/white selection having lacZ gene at the cloning site.
4. T7 and Sp6 phage based promoters examining transcription of genes of interest

BACs are capable of holding massive DNA inserts up to 300 kb in size. Because BACs have a far larger capacity than cosmids or conventional plasmid vectors, they are the perfect choice for cloning big genes or genomic areas. Low-copy number BACs originate from the F plasmid (fertility plasmid), which is frequently their source of replication. This origin minimizes the possibility of recombination or instability of big insertion by guaranteeing that only one to two copies of the BAC are retained per bacterial cell. Genes for antibiotic resistance, such as those for ampicillin or kanamycin resistance, are commonly found in BAC vectors. The bacterial colonies that have successfully taken up the BAC can be chosen thanks to these indicators. Multiple cloning sites (MCS) are present in BACs to allow the insertion of foreign DNA.

Large DNA fragment insertion is made easier by the MCS's variety of restriction enzyme sites. The purpose of BACs is to keep big DNA inserts in bacterial cells steadily. The big insert is kept intact without frequent recombination or loss thanks to low-copy replication.

Large-scale genome mapping initiatives, like the Human Genome Project, have relied heavily on BACs. Large genomic sections can be cloned thanks to them, and the cloned segments can then be sequenced and assembled into continuous sequences (contigs). By fragmenting and cloning an organism's whole genome into BACs, genomic libraries are created using BACs. This method enables the thorough investigation of an organism's genetic makeup.

4.11. Yeast artificial chromosome (YAC)

Yeast artificial chromosomes (YACs) are a human engineered DNA used to clone DNA sequences in yeast cells. YAC's are often used in connection with the mapping and sequencing of

genomes segments of a organisms DNA up to one million base pairs in length, can be inserted into YAC's with their inserted DNA, on then taken up by yeast cells. As the yeast cells grow and divide, they amplify the YAC DAM, which can be isolated and used for DNA mapping and sequencing.

Large DNA segments, usually between 200 and 2,000 kilobases (kb), are cloned using yeast artificial chromosomes (YACs). Large DNA inserts can be maintained steadily because they are made to function and replicate inside yeast cells, imitating the chromosomes' natural behavior. YACs are very helpful in the investigation of big genes, complicated genomes, and genomic areas with intricate structural details. A genomic library can include a significantly smaller number of clones. The elements of YAC vectors are as follows:

1. coli origin of replication
2. Yeast origin of replication
3. Elements of eukaryotic yeast chromosome (centromere and telomere region)
4. Selection markers for both the hosts (Bacterial as well as Yeas

DNA pieces up to 3,000 kb in size and greater than 100 kb can be cloned using the YAC vector. Large gene cloning and the physical mapping of complicated genomes are two applications for YACs. YACs have the capacity to hold incredibly big DNA pieces, up to 2,000 kb in size. They are appropriate for cloning big genes or complete genomic areas due to their huge insert capacity. YACs have a replication origin that enables them to multiply on their own within yeast cells. Selectable yeast markers, namely auxotrophic markers (like URA3, LEU2, and HIS3), are commonly found in YACs and enable the selection of yeast cells that have successfully undergone transformation. These markers make up for shortcomings in the used yeast strain. Many genome mapping initiatives, such as the Human Genome Project, have made substantial use of YACs. They make it possible to clone sizable genomic sections, which can then be thoroughly examined and sequenced to produce comprehensive maps of whole genomes.

4.12. Mammalian artificial chromosome (MAC)

Mammalian Artificial Chromosomes (MACs) are very large genomic constructions prepared is yeasts or bacteria that can incorporate form 100kb to 10 Mb, including all the regulatory elements for a gene to be expressed and the centromeres for chromosomal segregation during cell duplication. Its

provides a means to introduce large payloads of genetics information into the cell is an autonomously replicating non-integrating format.

Engineered constructs called Mammalian Artificial Chromosomes (MACs) are intended to replicate and perform like real chromosomes in mammalian cells. Large DNA pieces, typically spanning 1 to 5 megabases (Mb), are cloned and manipulated with the aim of investigating gene function, expression, and genetic regulation within the setting of mammals. Large genetic elements can be inserted into mammalian cells by MACs while preserving chromosomal structure and function. The MAC have ability to replicate autonomously and segregate in mammalian cells. These MAC can be adjusted for investigations on the expression of big genes and their coding sections as well as DNA regulatory elements.

There are two fundamental methods for getting MACs ready. The first technique uses natural chromosomal fragmentation directed by the telomere. For instance, chromosome 21 has been used to create a human artificial chromosome (HAC). De novo assembly of cloned replication origins, telomeres, and centromeric regions occurs in vitro in another technique.Gene clusters or big genes are cloned using MACs so that their function and regulation can be studied in a mammalian environment. This helps to explain intricate genetic relationships and the expression of genes in mammalian cells.

4.13. Summary

In molecular biology, cloning vectors are essential instruments for cloning, modifying, and analyzing DNA fragments. By serving as carriers, these vectors enable researchers to introduce and spread foreign DNA inside of host cells. The phagemid vectors bridge the gap between sequencing vectors and high capacity cloning vectors. While the advanced functions of bacterial artificial chromosomes, binary bacterial artificial chromosomes, P1 derived artificial chromosomes, and yeast artificial chromosomes include expression investigations, plant transformations, and the creation of genomic libraries, they differ. Small, circular DNA molecules called plasmids are autonomously replicating in bacteria. Because of their effectiveness and simplicity, they are frequently employed for cloning small to medium-sized DNA fragments. Vectors such as pBR322 and pUC are examples.

Large DNA fragments up to 300 kb can be cloned using bacterial artificial chromosomes, or BACs. BACs contain components from the F plasmid, including centromeres, telomeres, and a low-copy number origin of replication. For initiatives involving genome mapping and sequencing, they are

essential. Yeast Artificial Chromosomes (YACs) may clone DNA segments ranging in size from 200 kb to 2 Mb and function as chromosomes within yeast cells. YACs, which resemble natural yeast chromosomes and are utilized in extensive gene and genomic research, have telomeres, centromeres, and a replication origin.

Mammalian Artificial Chromosomes (MACs) are chromosomes that are engineered to function in mammalian cells. They possess telomeres, centromeres, and replication origins, which enable them to mimic the properties of natural chromosomes. They are employed in functional genomics, gene therapy, and gene cloning. Cloning vectors are essential for genetic engineering, functional genomics, gene cloning, and gene expression research. They make it possible for scientists to investigate and introduce particular genes or genomic regions, which advances the fields of biotechnology, medicine, and genetics.

4.14. Terminal questions

Q.1. What are Plasmids? Discuss the types and structure of plasmids.

Answer: _____

Q.2. Discuss about pBR322 and pUC8

Answer: _____

Q.3. What are the bacteriophase? Write about Lambda (λ)-phage vector.

Answer: _____

Q.4. Write about Cosmids and Phasmids.

Answer: _____

_____

_____

Q.5. Discuss about bacterial artificial chromosome (BAC).

Answer: _____

_____

_____

Q.6. What is Yeast artificial chromosome (YAC)? Discuss it structure and applications.

Answer: _____

_____

_____

Q.7. Write short notes on

a) Mammalian artificial chromosome (MAC

b) M13 phage

Answer: _____

_____

_____

1.5. Further suggested readings

1. Robert Schlei, Genetics and Molecular Biology,2nd Edition
2. McGraw-Hill, Cell and Molecular Biology, Human Genetics: Concepts and Application, 9th Edition.
3. An Introduction to Genetic Engineering, Desmond S. T. Nicholl , Third Edition, University of the West of Scotland, Paisley, UK.
4. Source: Fundaments of molecular biology, J.K.pal and S.G. Ghaskadbi, oxford university press
5. Genetic engineering, Smita Rastogi and Neelam Pathak, oxford university press

Unit-5: cDNA and its applications

5.1. Introduction

In molecular biology, cDNA is synthesized from an RNA template using the enzyme reverse transcriptase (RTase). Many molecular biology operations, including gene expression investigations utilizing real-time PCR, begin with cDNA synthesis. cDNA is also used to clone genes and build cDNA libraries. cDNA library are crucial tools in molecular biology research. They provide a stable, manipulated copy of the m RNA content of cells, allowing researcher to investigate gene expression, splice variants, proteins interactions and more. A cDNA library is a collection of cDNA fragments, each of which is a cloned copy of messenger RNA (mRNA). The cDNA library is a potent and versatile tool with numerous applications in biotechnology. Unlike genomic DNA libraries, which include all of the DNA in an organisms genome (including non-coding sequences that are transcribed into mRNA and potentially translated into proteins. this makes temperature useful for studding proteins coding genes. It allows researchers to study expression of mRNA. The RNA molecules with in a cell can serve as templates for the creation of DNA molecules, these DNA molecules can then be combined in a single lactation to form a cDNA library.

1. It enhances the expression of eukaryotic genes in prokaryotes, facilitating transcription.
2. It isolates DNA sequences that code for a certain mRNA.

3. It is beneficial in reverse genetics when extra genomic information is not needed.

cDNA is formed from completely transcribed mRNA located in the nucleus and hence contains only an organism's expressed genes. cDNA helps to coordinate the investigation and containment of communicable disease outbreaks across multiple jurisdictions. Engage and collaborate with a wide spectrum of national and international partners to prevent and control communicable illnesses. cDNA is used to clone eukaryotic genes into prokaryotic cells. It is used to make proteins in a cell that would not normally develop on its own. It also aids in understanding gene activity in organisms' metabolic processes. There are some of the key principles behind DNA libraries.

1) **Represent gene expression:** a cDNA library represents the genes that are being actively expresses in a given cell at a particular point in time. Therefore it is valuable tool for determinations which genes are active within a particular cell.
2) **Reverse transcription:** the process of creating cDNA involves a reaction know as reverse transcription. In this reaction, the enzyme reverse transcriptase uses mRNA as a template to synthesize a complementary strand of DNA. The cDNA can then be amplified and cloned into a suitable vector, creation a library of cDNA clones.
3) **Clone selection:** by selecting and sequencing specific clones form the cDNA library, researchers can study the sequence and function for individual genes.
4) **Non- loading regions:** one advantage of cDNA library over genomic library is that cDNA library does not contain non coding regions. This is become mRNA, the template for creating cDNA does not include these non coding regions. This makes cDNA libraries a more effective tool for gene expression studies.
5) **Tissue specific:** cDNA libraries can be made form a specific tissues or from a particular development stage using such as library, one confined out which gene can expressed in that tissue or stage of development.
6) **Expression profiling:** they can also be used for comparative gene expression studies. By comparing cDNA libraries prepared determine changes in gene expression patterns.

The step-wise process of gene cloning

The cDNA contains only coding sequences that can be directly expressed in eukaryotic systems. Plasmid, cosmid, lambda phage, BAC (Bacterial artificial chromosomes), and YAC (Yeast artificial chromosome) are used to accommodate huge DNA segments. Before amplification, cDNA can be

safely stored overnight at 4°C or -20°C. To keep the sample for a longer period, add 0.5 µl of 10X NEBNext Cell Lysis Buffer to the cDNA after RT, rather than during the PCR stage. After cDNA amplification and SPRI cleanup, the cDNA fragments for Single Cell 3' and 5' libraries will range in size from 400 to 9000 bases (see figure below). cDNA synthesis is the process of synthesizing complementary DNA (cDNA) using reverse transcription. The technique uses mRNA or miRNA as a template, a reverse transcriptase enzyme, and a thermostable primer that complements the RNA template's 3' end. cDNA clones are created by synthesizing complementary DNA from mRNA and putting a duplex copy of it into a cloning vector, which is subsequently transformed into bacteria. This simple and effective cDNA cloning approach drastically decreases the amount of RNA and labor necessary to create large directionally cloned libraries. Directional libraries that maintain many of the beneficial characteristics of traditional bacterial expression vectors. To get the DNA fragment into a bacterial cell in a format that can be duplicated or expressed, it is first introduced into a cloning vector. A cloning vector is a fragment of DNA that can be permanently maintained in an organism and into which a foreign DNA fragment can be inserted for cloning. A cDNA library is a collection of cloned DNA sequences that are complementary to mRNA taken from an organism or tissue ('c' in cDNA stands for 'complementary'). Genomic DNA libraries include huge DNA fragments. CDNA libraries, on the other hand, are made up of reverse-transcribed mRNA that has been cloned. As a result, they lack DNA sequences in genomic locations where no expression occurs, such as introns.

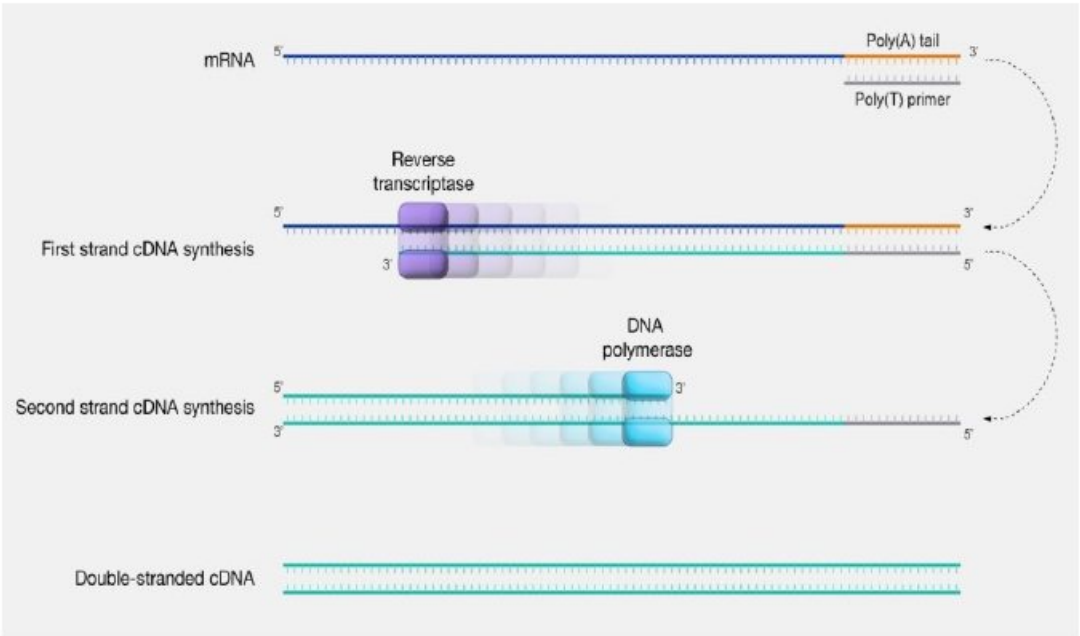
Objectives:

After reading this unit the learner will able to know

- The cDNA and its Synthesis from mRNA,
- The Cloning cDNA in plasmid vectors
- The Cloning cDNAin bacteriophage vectors,
- The Expression of cloned cDNA molecules,

5.2. cDNA

cDNA is an abbreviation for complementary DNA, or simply cDNA. cDNA is synthetic DNA that has been transcribed from a specific mRNA by the enzyme reverse transcriptase. DNA consists of both coding and non-coding sequences, whereas cDNA only contains coding sequences. Scientists routinely create and use cDNA for gene cloning and other scientific purposes.

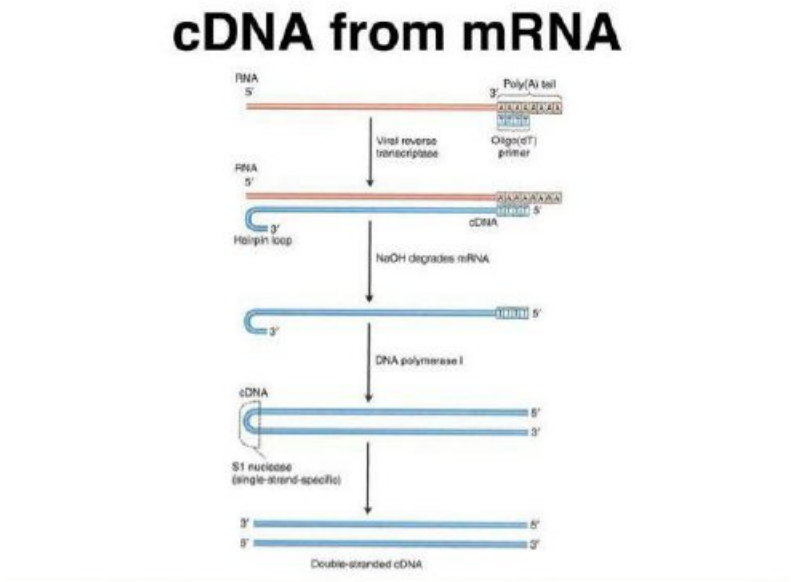


Each cell has the same set of genes, but various cell types activate different sets of genes while turning off others. The fact that only a few genes are expressed in any particular type of cell can be leveraged to create a library if the cloned material is mRNA rather than DNA. Only expressed genes are transcribed into mRNA, hence if mRNA is used as the starting material, the resulting clones will only include a subset of the cell's total number of genes. A mRNA-based cloning method is especially useful when the desired gene is overexpressed in a single cell type. For example, the gliadin gene is overexpressed in the cells of growing wheat seeds. More than 30% of the total mRNA in these cells codes for gliad. However, mRNAs are extremely lengthy molecules. These cannot be put into a cloning vector and are difficult to amplify in their natural form. As a result, the RNA-encoded data must be converted into a stable DNA duplex known as complementary DNA (cDNA) or copy DNA, which can then be inserted into a self-replicating vector. The resulting cDNA represents the expressed genes and information encoded in the mRNA of a particular tissue or organism, as well as a certain developmental stage. It takes into account mRNA levels and the diversity of splice isoforms in specific organs.

5.3. Synthesis of cDNA from mRNA

cDNA synthesis refers to the process of producing complementary DNA (cDNA) via reverse transcription. The process uses reverse transcriptase, a thermostable primer complementary to the 3' end of the RNA template, and mRNA or miRNA as the template. When an RNA template is reverse

transcribed, complementary DNA (cDNA) is produced. Reverse transcriptases (RTs) synthesize first-strand cDNA from an RNA template and a short primer complementary to the RNA's 3' end.



This cDNA can be used directly as a template to create the second DNA strands using PCR (Polymerase chain reaction) experiments. Reverse transcription plus PCR (RT-PCR) allows you to identify low abundance RNAs in a sample and generate the corresponding cDNA, making low copy gene cloning easier. Alternatively, DNA Polymerase I and DNA Ligase can be employed to double-strand the first-strand cDNA. It is feasible to clone directly from these reaction products without amplification. In this case, exogenous RNase H activity or activity from the RT is required. The conversion of mRNA sequences to ds cDNA offers several advantages, which include:

1. mRNA is intrinsically stable, but cDNA allows for long-term storage of these sequences.;
2. cDNA represents the cell's mRNA population (expressed genes). Thus, CDNA synthesis creates a permanent biochemical record of the cell during lysis.
3. The process of generating and cloning cDNA from a single source creates a technique for propagation. This strategy is substantially aided by the wide range of vectors compatible with hosts.
4. Additionally, cDNA molecules, both long and short, can be employed to screen members of complicated genomic DNA libraries.

This approach makes it easier to separate exon and intron sequences from genes' structural regions, as well as the flanking 5' and 3' sequences. This comparison also helps with the identification of 5' and 3'

splice sites, as well as information about alternatively spliced genes. The cDNA library also contains information on genes that are controlled throughout development, as well as genes that are particular to cells or tissues.

i. Prepare sample RNA for cDNA synthesis

The RNA serves as a template for the production of cDNA. Certain forms of RNA, such as messenger RNA (mRNA) and small RNAs like miRNA, can be increased when building cDNA libraries and miRNA profiles, but whole RNA is commonly used in cDNA synthesis for downstream applications such as RT-(q)PCR. When extracting, processing, storing, and using RNA in research, extreme caution must be exercised to ensure its integrity. The best approaches to prevent RNA degradation are to wear gloves, use pipettes with aerosol-barrier tips, use nuclease-free reagents and labware, and keep your workstation clean. There are numerous approaches for isolating and purifying RNA, depending on the nature of the source materials (blood, tissues, cells, plants, etc.) and the goals of the experiments.

ii. Remove genomic DNA preparing sample for cDNA synthesis

Trace amounts of genomic DNA (gDNA) can be isolated alongside RNA. Contaminating gDNA can interfere with reverse transcription, resulting in false positives, increased background, or decreased detection in sensitive applications like RT-qPCR. The standard method for removing gDNA is to add DNase I to isolated RNA samples. However, DNase I must be eliminated before cDNA synthesis since any remaining enzyme will destroy single-stranded DNA, affecting RT-PCR results. Unfortunately, DNase I inactivation approaches can typically (This is recombinant double stranded specific DNas for removal of containing genomic DNA form RNA preparations) cause RNA loss or damage. Double-strand-specific DNases, such as Invitrogen ezDNase Enzyme, can be used instead of DNase I to remove contaminated gDNA without damaging RNA or single-stranded DNAs. Their thermolabile feature enables straightforward inactivation at relatively low temperatures (e.g., 55°C) with no detrimental consequences. These double-strand-specific, thermolabile DNases provide quicker protocols, requiring only 2 minutes at 37°C before reverse transcription processes to complete gDNA digestion

iii. Select reverse transcriptase for improved cDNA yield

The majority of reverse transcriptase's utilized in molecular biology are generated from (the polygene encodes several proteins needed at lower levels for replication of virus, including the reverse

transcription and integrate enzymes. The polORF is not expressed as sepate proteins in most retrovirus, but rather is expressed as a part of larger Gag-Pro-Pol fusion proteins. the pol gene of the avian myeloblastosis virus (AMV) or the Moloney murine leukemia virus. The AMV reverse transcriptase was one of the first enzymes identified in the laboratory for cDNA synthesis. The enzyme has significant RNase H activity and degrades RNA in RNA: cDNA hybrids, producing shorter cDNA fragments (<5 kb). The MMLV reverse transcriptase became popular due to its monomeric form, which allowed for easier cloning and customization of the recombinant enzyme. Although MMLV (Moloney murine leukemia virus)is less thermos table than AMV reverse transcriptase, MMLV reverse transcriptase's decreased RNase H activity allows for more efficient synthesis of longer cDNA (<7 kb). To improve cDNA synthesis even more, MMLV reverse transcriptase has been designed with a mutant RNase H domain (RNaseH-), improved thermo stability (up to 55°C), and increased processivity. These characteristics can lead to enhanced cDNA length and yield, greater sensitivity, improved inhibitor resistance, and faster reaction times (Table 5.1).

Table 5.1: Common reverse transcriptases and their attributes

	AMV reverse transcriptase	MMLV reverse transcriptase	Engineered MMLV reverse transcriptase (e.g., Invitrogen SuperScript IV Reverse Transcriptase)
RNase H activity	High	Medium	Low
Reaction temperature (highest recommended)	42°C	37°C	55°C
Reaction time	60 min	60 min	10 min
Target length	≤5 kb	≤7 kb	≤14 kb
Relative yield (with challenging or suboptimal RNA)	Medium	Low	High

Source: https://www.thermofisher.com/in/en/home/life-science/pcr/reverse-transcription/5steps-cDNA.html

iv. Prepare reaction mix for cDNA synthesis

In addition to the enzyme, the primary reaction components for cDNA synthesis are buffer, dNTPs, DTT, RNase inhibitor, nuclease-free water, and primers (Figure 2).

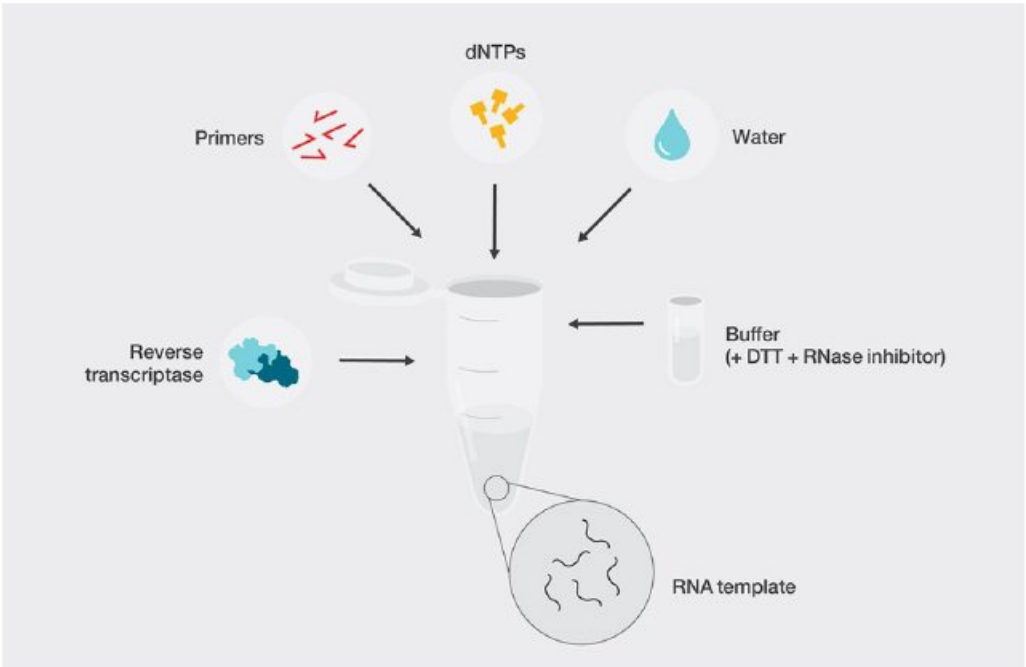


Table 5. 2: Reaction components for cDNA synthesis

Source: https://www.thermofisher.com/in/en/home/life-science/pcr/reverse-transcription/5steps-cDNA.html

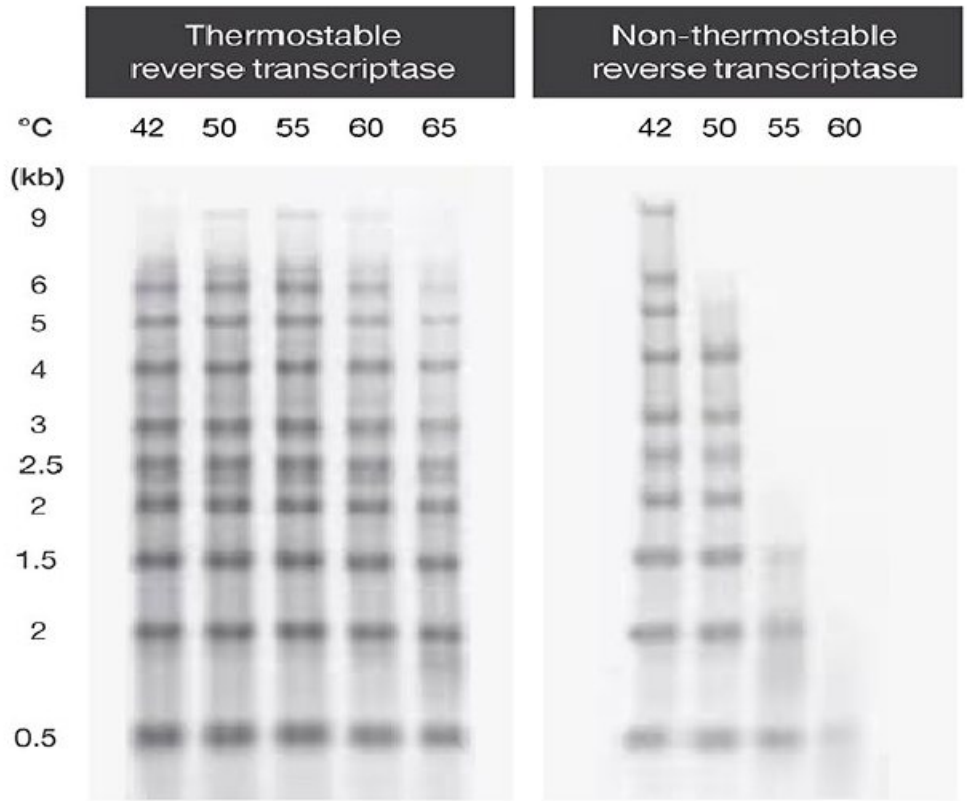
Component	Key features
RNA template	Maintaining RNA integrity is critical and requires special precautions during extraction, processing, storage, and experimental use. Total RNA is routinely used in cDNA synthesis for downstream applications such as RT-(q)PCR Messenger RNA (mRNA) and small RNAs such as miRNA) may be enriched for certain applications like cDNA library construction and miRNA profiling
Reaction buffer	Maintains a favorable pH and ionic strength for the reaction The supplied buffer may also contain additives to enhance the efficiency of reverse transcription
dNTPs	Generally, should be at 0.5–1 mM each, preferably at equimolar concentrations High-quality dNTPs, freshly diluted, are recommended to help ensure proficient reverse transcription
DTT	Reducing agent, often included for optimal enzyme activity Reaction efficiencies may be compromised if DTT or other additives precipitate; hence, reaction components should be dissolved and well mixed
RNase inhibitor	Often included in the reaction buffer or added to the reverse transcription reaction to prevent RNA degradation by RNases. RNase inhibitors may be: Co-purified during isolation

	Introduced during reaction setup A number of known RNases exist, and appropriate RNase inhibitors should be chosen based on their mode of actions and reaction requirements.
Water	Use DEPC-treated or nuclease-free water from a commercial source to minimize the risk of contaminating RNases. Nuclease-free water DEPC (diethylpyrocarbonate)-treated water Contaminating RNases cannot be removed by simple filtration, and autoclaved water is not adequate because RNases are heat stable.

Source: https://www.thermofisher.com/in/en/home/life-science/pcr/reverse-transcription/5steps-cDNA.html

v. Perform cDNA synthesis

Reverse transcription reactions have three major steps: primer annealing, DNA polymerization, and enzyme deactivation. The temperature and duration of these steps vary depending on the primer, target RNA, and reverse transcriptase employed. If the RNA is GC-rich or known to include secondary structures, an optional denaturation phase can be undertaken by heating the RNA-primer mix to 65°C for 5 minutes before chilling on ice for 1 minute. If utilizing random hexamers, we recommend incubating the reverse transcription reaction at room temperature (~25 °C) for 10 minutes to anneal and lengthen the primers. DNA polymerization is a vital stage in this process; the reaction temperature and time may differ depending on the reverse transcriptase employed. Reverse transcriptases differ in their thermostability, which impacts the ideal polymerization temperature. Using a thermostable reverse transcriptase allows for a higher reaction temperature (e.g., 50°C) to help denature RNA with high GC content or secondary structures while preserving enzyme activity (Figure 3). Thus, higher-temperature processes can result in increased cDNA yield, length, and representation.



Polymerization time is determined by a reverse transcriptase's processivity, which is the number of nucleotides integrated in a single binding event. For example, wild-type MMLV reverse transcriptase with low processivity frequently takes more than 60 minutes to generate cDNA. In comparison, a designed reverse transcriptase with high processivity might finish cDNA synthesis in as little as ten minutes.

5.4. Cloning cDNA in plasmid vectors

Contraction of cDNA involves the synthesis of complimentary DNA form mRNA and them inserting a duplex copy of that into a cloning vector, followed by transformation of bacteria. First strand synthesis, first one anneals an aligo primer onto the 3 poly A tail of a population of mRNA. Plasmids are still often employed in cDNA cloning, even though many professionals prefer to utilize a bateriophage vector system. The principle behind these techniques of cDNA cloning is that an mRNA population is isolated form a specific developmental stage should contain mRNA specific for any protein expressed during that stage. Thus if the mRNA can be isolated, thus gene can be studied. This is especially true in situations when screening a few numbers of clones is necessary to isolate the appropriate cDNA sequence. The three cDNA cloning techniques shown in Fig. 5.1 are typically used to join the cDNA fragments to the vector: blunt-end ligation, linker molecule usage, and homopolymer

tailing. While cDNA cloning is the preferred use for these techniques, genomic DNA can also be employed (see Section 5.1). A brief description of each of the three approaches will be given.

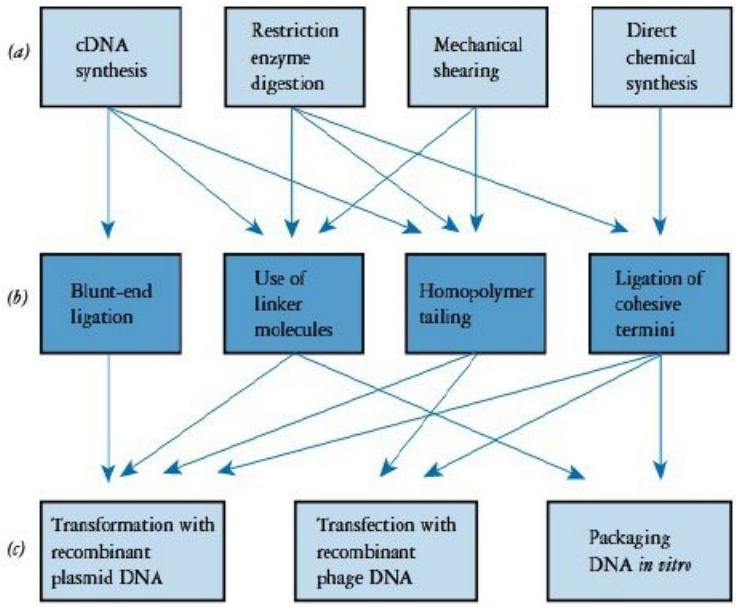


Fig 5.1: Routes available for cloning by what might now be considered 'traditional' methods. Possible methods available for three key stages of a cloning procedure are shown as follows: (a) the generation of DNA fragments, (b) joining to a vector, and (c) introducing the recombinant DNA into a host cell.

Source: An Introduction to Genetic Engineering, Desmond S. T. Nicholl , Third Edition, University of the West of Scotland, Paisley, UK

Blunt-end ligation (usually, a straight cut creates blunt ends or non over hanging ends. these ends can be joined using a DNA Ligases enzyme. the joining of two blunt ends is called blunt end ligation. this does not need maturing complimentary ends for ligation) is the process of connecting DNA molecules with blunt ends using ligase. In cDNA cloning, blunt ends can be created by using S1 nuclease or by filling in protruding ends using DNA polymerase. Blunt-end ligation is inefficient because there is no specific intermolecular interaction to hold the DNA strands together. In contrast, DNA ligase generates the necessary phosphodiester bonds to make recombinant DNA. To maximize the likelihood of two ends joining, utilize high amounts of participating DNAs.

The criteria for end ligation must be carefully considered. When vector DNA and cDNA are combined, many results may occur. The goal is for one cDNA molecule to connect with one vector molecule, creating a recombinant with a single insert. Insufficient concentrations can cause insert or vector DNAs to self-ligate, resulting in circular molecules or concatemers instead of bimolecular recombinants. To inhibit self-ligation, the vector is typically treated with a phosphatase, such as bacterial alkaline phosphatase or calf intestinal alkaline phosphatase. The vector and insert DNA concentrations are set to promote the generation of recombinants. One downside of blunt-end ligation is the lack of restriction enzyme recognition sequences at the cloning site, which can make it difficult to remove the insert from the recombinant. Linkers are self-complementary oligomers with recognition sequences for certain restriction enzymes. For example, 5′-CCGAATTCGG-3′ has the EcoRI recognition sequence (GAATTC) in its ds form. Fig. 5.2 shows an alternative method of cloning that involves adding sequences to the ends of DNA molecules using adaptors. These ss non-complementary oligomers can be employed with linkers. Annealing together creates a linker/adaptor with one blunt end and one sticky end. This can be added to cDNA for sticky-end cloning without the need for linker digestion. Homopolymer tailing is a popular and effective method for cloning cDNA. This approach involves (a homopolymer is a polymer made from many copies of a single repeating unit, for example, a number of glycolic acid molecules can be conjugated to creat the homopolymer) adding homopolymers of dA, dT, dG, or dC to a DNA molecule using the enzyme ter-minal transferase.

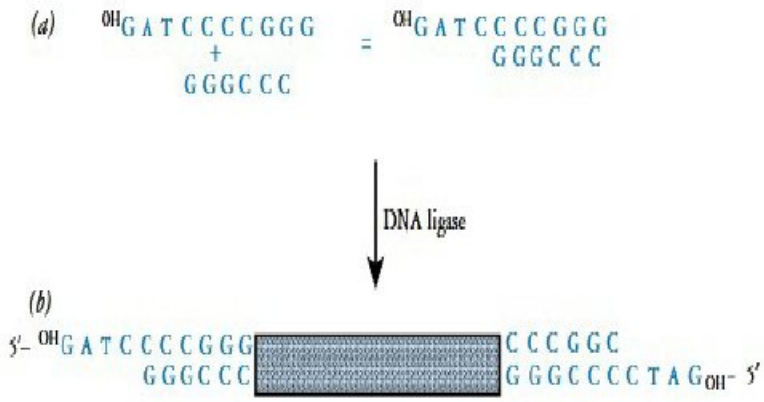


Fig.5.2: Use of adaptors. In this example a BamHI adaptor (5′-GATCCCCGGG-3′) is annealed with a single-stranded HpaII linker (3′-GGGCCC-5′) to generate a double-stranded sticky-ended molecule, as shown in (a). This is added to blunt-ended DNA using DNA ligase. The DNA

therefore gains protruding 5′ termini without the need for digestion with a restriction enzyme, as shown in (b). The 5' terminus of the adaptor can be dephosphorylated to prevent self-ligation.

Source: An Introduction to Genetic Engineering, Desmond S. T. Nicholl , Third Edition, University of the West of Scotland, Paisley, UK

Homopolymers offer two key benefits over other approaches for connecting DNA from diverse sources. They offer longer areas for annealing DNAs compared to cohesive termini formed by restriction enzyme digestion. The cDNA-vector hybrid is stable enough to be introduced into the host cell and ligated in vivo, eliminating the necessity for in vitro ligation. A second advantage is specificity. Because the vector and insert cDNAs have distinct but complementary 'tails', self-annealing is unlikely. This allows for the creation of bimolecular recombinants over a larger range of effective concentrations than other annealing/ligation procedures. An example of the use of homopolymer tailing is shown in Fig. 6.6.

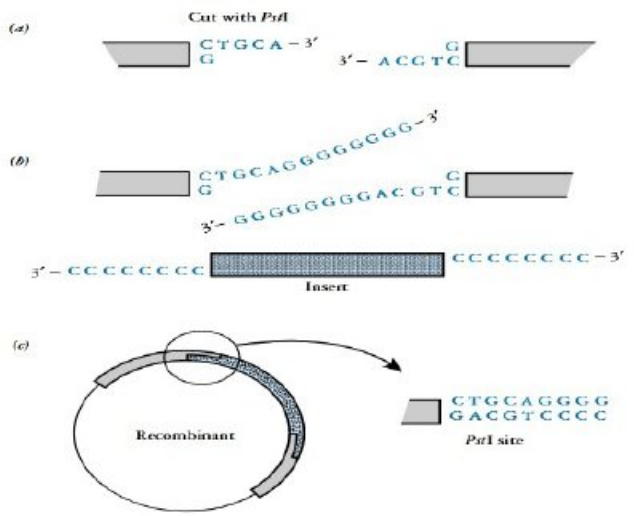


Fig.5.3: Homopolymer tailing. (a) The vector is cut with PstI, (The Pstl restriction/ modification (R/M) system has two components: a restriction enzyme that cleave foreign DNA, and a methyltransferase which protect native DNA strands by methylation of the adenine base inside the recognition sequence. the combination of both provide is a defense mechanism agisnt invading virus) which generates protruding 3′-OH termini. (b) The vector is then tailed with dG residues using terminal transferase. The insert DNA is tailed with dC residues in a similar way. (c) The dC and dG tails are complementary and the insert can therefore be annealed with the vector to generate a recombinant. The PstI sites are regenerated at the ends of the insert DNA, as shown.

Source: An Introduction to Genetic Engineering, Desmond S. T. Nicholl , Third Edition, University of the West of Scotland, Paisley, UK

5.5. Cloning cDNAin bacteriophage vectors

bateriophage λ cloning g vector λ ORF8 can be used for contraction of cDNA, libraries this methods utilized the synthesis of double stranded cDNA including printing of first strand, synthesis by aligo (dT). after complication of second strand synthesis a olegodeoxynucleotide linker is legated to the cDNA fragment, this linker which contains a Bam HI restriction site, will creat a Hind III restriction site when ligated to the 3 end of cDNA fragments. Subsequent treatment of methylated cDNA with Hind III and Bam HI endonucleases allows these fragments to be cloned directionally. Cloning cDNA into bacteriophage vectors is a molecular biology technique used to generate cDNA libraries, which are collections of cDNA clones that represent the mRNA population of a specific tissue or cell type. Bateriophage vectors, such as λ phage, offer bigger cloning capacity than plasmids and are particularly effective for building libraries with a broad range of genes. Here's a step-by-step overview of the procedure. The simple and efficient cDNA cloning system significantly reduces the amount of RNA and effort required for the preparation of large directionally cloned libraries. Directional libraries retain many of the useful features of existing bacterial expression vectors, relative to vector-encoded expression of signal.

i. Isolation of mRNA:

It begins with isolating mRNA from the tissue or cell type of interest. To extract total RNA, use phenol-chloroform extraction or a commercial RNA isolation kit, then purify it with oligo(dT) beads or columns.

ii. Synthesis of cDNA:

- ***First-strand synthesis:*** Reverse transcriptase uses an oligo(dT) primer or random primers to manufacture the first strand of cDNA from the mRNA template.
- ***Second-strand synthesis***: The second strand is synthesized using methods like the Gubler-Hoffman method, where RNase H partially degrades the RNA in the RNA-DNA hybrid, and DNA polymerase I synthesizes the second strand of DNA.
- ***cDNA end repair (if necessary):*** If the cDNA ends are not blunt, they can be treated with enzymes like T4 DNA polymerase or Klenow fragment (it is a large protein fragment produced when DNA polymerase I from E. coli is enzymatically cleaved the protease subtilisin. it is the large fragment

that contains 5' to 3' polymerase and 3' 5' exonuclease activity domains of the DNA polymerases I) to create blunt ends.

iii. Preparation of Bateriophage Vector:

Choose a bateriophage vector, commonly λ phage, that can accommodate large inserts (up to 15-20 kb). The bateriophage vector is generated by digesting it with the required restriction enzymes to produce ligation-compatible ends. Vectors, such as λgt10 or λZAP, can take DNA inserts with blunt or cohesive ends. The vector arms can be dephosphorylated to prevent self-ligation.

iv. Ligation of cDNA into the Bacteriophage Vector:

If the cDNA does not have compatible ends, synthetic oligonucleotide adaptors can be ligated to provide sticky ends that match the bacteriophage vector's digested ends.T4 DNA ligase is used to incorporate the cDNA into the bacteriophage vector, resulting in recombinant DNA molecules. The ligated DNA is subsequently packaged into phage particles using in vitro packaging extracts, which combine the recombinant DNA into functional phage particles capable of infecting E. coli.

v. Transformation of Host Cells:

Recombinant bacteriophages are employed to infect an E. coli host, typically a phage-tolerant strain such as E. coli C600 or LE392. Infected E. coli cells are placed on an agar plate. Each recombinant phage will generate a plaque, which is a clear spot on the bacterial lawn where the phage has lysed the cells.

vi. Screening of cDNA Library:

Nitrocellulose or nylon membranes are put on agar plates to collect phage particles from plaques.The membrane is then hybridized with a tagged probe (such as radioactive or fluorescently labeled DNA/RNA) that is complementary to the target cDNA. Positive plaques with hybridization indicate the presence of the cDNA of interest.Alternatively, PCR can be used to directly amplify cDNA inserts from phage plaques, allowing for the quick identification of clones that contain the desired gene.

vii. Subcloning (optional):

The cDNA inserts from positive phage clones can be removed and subcloned into plasmid vectors for further investigation or protein production. In systems like λZAP, the cDNA may be easily

extracted in vivo as a plasmid (phagemid) that can replicate in E. coli, simplifying the shift from phage-based to plasmid-based systems.

5.6. Expression of cloned cDNA molecules

Cloned cDNA molecules are expressed in a host organism by creating the protein represented by the cDNA, often via a plasmid or other vector system that drives gene expression. This is an important stage in understanding gene function, creating recombinant proteins for research or therapeutic purposes, and investigating gene control. Here's a thorough description of the procedure:

i. Preparation of the Expression Vector:

Choose an expression vector that is compatible with your host organism (such as E. coli, yeast, or mammalian cells). The vector should include: A powerful promoter for cDNA transcription. A ribosome binding site (RBS) that facilitates translation initiation (in prokaryotes). A selectable marker (such as an antibiotic resistance gene) for identifying cells that have taken up the vector. A host-appropriate replication origin. Any required regulatory elements, such as enhancers, transcription terminators, or signal sequences for protein secretion. The cDNA is placed into the vector in the proper direction, downstream of the promoter. This could include introducing restriction sites or employing recombination-based cloning procedures. If protein purification or visualization is needed, the insert should be in-frame with any necessary tags (such as His-tag or GFP).

ii. Transformation or Transfection of Host Cells:

Chemical transformation or electroporation can be used to introduce the recombinant expression vector into competent bacterial cells (for example, E. coli). The vector is transfected into mammalian or other eukaryotic cells using methods such as lipid-mediated transfection, electroporation, or viral transduction. Grow the transformed/transfected cells on selective media to determine which ones have successfully taken up the vector.

iii. Induction of Protein Expression:

In these systems, the promoter is always active, resulting in constant production of the cDNA. This is frequent in vectors with constitutive promoters, such as CMV (cytomegalovirus) in mammalian cells and T7 in E. coli. In some systems, expression is regulated by an inducible promoter, which is triggered when an inducer is added to the culture.

For example:

- ***IPTG induction:*** Used in E. coli with the lac operon or T7 promoter system.
- ***Tetracycline or Doxycycline:*** Used in mammalian cells with Tet-On/Tet-Off systems.
- ***Galactose induction:***Used in yeast with the GAL1 promoter.

Conditions such as temperature, inducer concentration, and timing are optimized to maximize expression and stability of the recombinant protein.

iv. Detection and Analysis of Expressed Protein:

Proteins produced from cDNA can be identified by SDS-PAGE (It is a continuous Electrophoretic system developed by Ulrich K. Laeminli which is commonly used a method ot proteins based their molecular weight) followed by staining (e.g., Coomassie blue) or Western blotting with antibodies specific to the protein or epitome tag. If the expressed protein is an enzyme, its activity can be determined using the proper substrates and assays. Fluorescence microscopy can be used to monitor protein expression that has been labeled with GFP or another fluorescent tag. Used to verify the identification and post-translational changes of the expressed protein.

v. Purification of the Recombinant Protein (optional):

If the protein was expressed with an affinity tag (such as His-tag or GST), affinity chromatography can be used to purify it. To get very pure protein, additional purification processes such as ion exchange or size-exclusion chromatography may be necessary.

5.7. cDNA library

A cDNA library is a collection of closed DNA sequences that are complimentary to the mRNA that was extracted from an organism or tissues. cDNA libraries have been broadly used to determine the expressed proteins of protein coding genes in eukaryotes. The contraction of cDNA library involved the extractions purification of mRNA. These mRNA are used as template for the synthesis of sDNA by the process of revere transcription in the presence of olego dT primer. The oligo dT primer binds with the poly-A tail of mRNA followed by synthesis of the fire strand of cDNA by using reverse transcriptase. A cell can synthesize several thousand proteins at once, all of which are translated from their respective mRNA molecules. CDNA libraries contain mRNA from a specific cell or tissue. Because mRNA is too unstable to use directly in cloning, cDNA molecules are generated for all

mRNAs and ligated to the vector before being cloned. The reverse transcriptase enzyme is used to generate cDNA. Reverse transcriptase is an RNA-dependent DNA polymerase that uses a mix of dNTPs to create a first strand DNA complementary to the mRNA template.

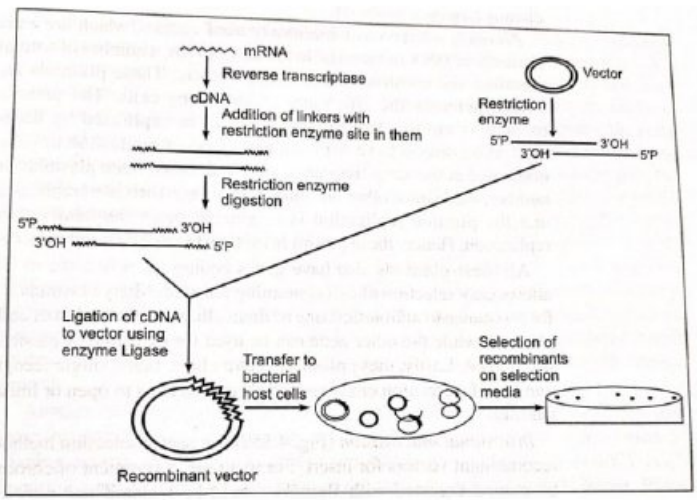


Fig. 5.4: cDNA cloning

Source: Fundaments of molecular biology, J.K.pal and S.G. Ghaskadbi, oxford university press

Because eukaryotic mRNAs have a poly A tail at the 3' end, a complementary oligo dT primer can be utilized. Such primers include a free 3'OH end, which reverse transcriptase uses to generate cDNA. Following the synthesis of the first strand, a second strand is added to produce double-stranded DNA. Because ligation of blunt-ended DNA fragments is inefficient, CDNA molecules' termini must be polished. One method is to connect linkers with internal restriction sites to cDNA and then construct cohesive termini by digesting it with the associated restriction enzymes before ligation (Fig.5.4). It is also feasible to create the DNA fragments that will be cloned using PCR as explained earlier. cDNA library have several disadvantages over genomic DNA libraries such as containing only expressed genes, lacking introns and non-coding regions, and allowing for the comparison of gene expression pattern across different tissue and developmental stage. A disadvantage of cDNA libraries is that it contains only sequences that are present in mature mRNA. Another disadvantage is that introns and other sequences that are altered during transcription are absent in a cDNA library.

A cDNA library is a collection of complementary DNA (cDNA) molecules that represent the mRNA population of a single tissue, cell type, or organism at a given time. This library captures a snapshot of the genes that were actively transcribed at the time of RNA isolation, making it an

invaluable resource for investigating gene expression, discovering new genes, and cloning specific genes of interest. Steps for Creating a cDNA Library:

i. **Isolation of mRNA:**

It begins with obtaining total RNA from the tissue or cells of interest. In eukaryotes, mRNA has a poly-A tail, thus it can be separated from total RNA using oligo(dT) beads or columns that bind to the poly-A tail.

ii. **Synthesis of cDNA:**

Reverse transcriptase, an enzyme that synthesizes DNA from an RNA template, is used to create the first strand of cDNA from pure mRNA. This synthesis can be initiated with an oligo(dT) primer that binds to the poly-A tail of mRNA or random primers. Following the synthesis of the first strand, RNase H normally degrades the RNA strand. DNA polymerase I is then utilized to create the second strand of cDNA, utilizing the first strand as a template. The outcome is a double-stranded cDNA molecule that corresponds to the original mRNA.

iii. **Modification of cDNA Ends (if necessary):**

If the cDNA ends are blunt, adaptors or linkers can be added to aid in cloning into a vector. Alternatively, restriction enzymes can be used to breakdown the cDNA, resulting in cohesive ends compatible with the cloning vector.

iv. **Cloning into a Vector:**

The cDNA is inserted into a linearized cloning vector, such as a plasmid or bacteriophage (e.g., λ phage). The vector used is determined on the application and host organism. Recombinant vectors containing cDNA inserts are transformed into a host organism (e.g., E. coli) or packaged in vitro into bacteriophages. Once inside the host cells, each cell (or phage) possesses a unique cDNA insert that represents the mRNA population of the donor material.

v. **Amplification and Storage of the Library:**

The altered cells or phages are cultivated to increase the library's size, allowing each clone to be stored and studied. The cDNA library is archived as a population of bacterial colonies, phage plaques, or extracted DNA for future use.

Applications of cDNA Libraries:

- **Gene Identification and Cloning:** cDNA libraries are used to discover and clone genes of interest. Researchers can use particular probes or antibodies to screen the library for clones that contain the gene of interest.
- **Expression Analysis:** By creating cDNA libraries from several tissues or developmental stages, researchers can compare gene expression patterns and uncover differentially expressed genes.
- **Functional Genomics:** cDNA libraries are a valuable resource for functional genomics research, allowing researchers to investigate the function of diverse genes by expressing the cDNAs in different systems.
- **Discovery of Novel Genes**: Screening a cDNA library can result in the discovery of previously unknown genes, especially if the source tissue or organism has not been thoroughly explored.
- **Protein Expression**: cDNA libraries can be used to create recombinant proteins in a variety of environments, allowing researchers to investigate protein function, interaction, and structure.
- **Evolutionary Studies**: cDNA libraries from different species or evolutionary stages can be compared to study the conservation and divergence of genes.
- Complementary DNA (cDNA) is a type of DNA generated from a messenger RNA (mRNA) template using the enzyme reverse transcriptase. Unlike genomic DNA, cDNA only represents the expressed genes within a cell because it is derived from mature mRNA that has been spliced to eliminate introns. This makes cDNA an extremely useful tool for analyzing gene expression, copying genes, and creating recombinant proteins.

1.8. Summary

The initial step in generating cDNA is to isolate mRNA from the desired tissue or cell type. This mRNA is reverse transcribed into single-stranded cDNA, which is subsequently converted to double-stranded cDNA. This double-stranded cDNA can be cloned into vectors and transmitted inside host species such as bacteria and yeast. A cDNA library is the result of a collection of clones, each containing a unique cDNA. This library contains the complete set of genes expressed in the original tissue or cell type, providing a snapshot of gene activity. cDNA has several applications in molecular biology. One notable application is the creation of cDNA libraries, which help in gene discovery and cloning. Researchers can scan these libraries for cDNAs that encode proteins of interest. This is especially significant for identifying and researching genes whose expression varies between different conditions, such as diseased and healthy tissues. Gene expression studies are another key application

for cDNA. Because cDNA corresponds to mRNA, it represents the genes that are now being transcribed. Scientists can learn about gene function and control by comparing cDNA libraries from different tissues, developmental stages, and situations. This is especially significant in fields like cancer, where knowing changes in gene expression is critical to determining disease progression. Another prominent application for cDNA is the production of recombinant proteins.

Once a cDNA encoding a protein of interest has been extracted, it can be inserted into an expression vector and introduced into a host organism such as bacteria, yeast, or mammal cells. The host cells subsequently generate and purify the protein, which is ultimately used in a range of applications including drug development, medicinal treatments, and industrial activities. For example, cDNA cloning and expression are used to create insulin and other therapeutic proteins. However, cDNA is a very valuable tool in modern molecular biology since it connects active genes throughout cells. Its applications in gene discovery, expression analysis, recombinant protein synthesis, gene therapy, and diagnostics make it essential for expanding our understanding of biology and developing breakthrough medical and biotechnological solutions.

1.9.Terminal questions

Q.1. What is the cDNA? Discuss application of cDNA.

Answer: ----------

Q.2. Discuss about Synthesis of cDNA from mRNA.

Answer: ----------

Q.3. Discuss about Cloning cDNA in plasmid vectors.

Answer: ----------

Q.4. Write about Cloning cDNAin bacteriophage vectors.

Answer: ----------

Q.5. Explain Expression of cloned cDNA molecules.

Answer: ----------

Q.6. Writ the Application of cDNA library.

Answer: ----------

1.10. Further suggested readings

1. Robert Schlei, Genetics and Molecular Biology,2nd Edition
2. McGraw-Hill, Cell and Molecular Biology, Human Genetics: Concepts and Application, 9th Edition.
3. An Introduction to Genetic Engineering, Desmond S. T. Nicholl , Third Edition, University of the West of Scotland, Paisley, UK.
4. Source: Fundaments of molecular biology, J.K.pal and S.G. Ghaskadbi, oxford university press
5. Genetic engineering, Smita Rastogi and Neelam Pathak, oxford university press

Unit 6: Cloning from Genomic DNA

Structure

6.1 Introduction

Genetic engineering, also called genetic modification, is the direct manipulation of an organism's genome using biotechnology. It is a set of technologies used to change the genetic makeup of cells, including the transfer of genes within and across species boundaries to produce improved or novel organisms. Genetic engineering has been applied in numerous fields including research, medicine, industrial biotechnology and agriculture. In research, GMOs are used to study gene function and expression through loss of function, gain of function, tracking and expression experiments. By knocking out genes responsible for certain conditions it is possible to create animal model

organisms of human diseases. As well as producing hormones, vaccines and other drugs, genetic engineering has the potential to cure genetic diseases through gene therapy. Chinese hamster ovary (CHO) cells are used in industrial genetic engineering. Additionally mRNA vaccines are made through genetic engineering to treat viruses such as COVID-19. The same techniques that are used to produce drugs can also have industrial applications such as producing enzymes for laundry detergent, cheeses and other products. Cloning is the process of producing individual organisms with identical genomes, either by natural or artificial means. In nature, some organisms produce clones through asexual reproduction; this reproduction of an organism by itself without a mate is known as parthenogenesis. In the field of biotechnology, cloning is the process of creating cloned organisms of cells and of DNA fragments. The artificial cloning of organisms, sometimes known as reproductive cloning, is often accomplished via somatic-cell nuclear transfer (SCNT), a cloning method in which a viable embryo is created from a somatic cell and an egg cell. In 1996, Dolly the sheep achieved notoriety for being the first mammal cloned from a somatic cell. Another example of artificial cloning is molecular cloning, a technique in molecular biology in which a single living cell is used to clone a large population of cells that contain identical DNA molecules. In bioethics, there are a variety of ethical positions regarding the practice and possibilities of cloning. The use of embryonic stem cells, which can be produced through SCNT, in some stem cell research has attracted controversy. Cloning has been proposed as a means of reviving extinct species. In popular culture, the concept of cloning-particularly human cloning-is often depicted in science fiction; depictions commonly involve themes related to identity, the recreation of historical figures or extinct species, or cloning for exploitation (e.g. cloning soldiers for warfare).

Objectives

After reading this unit, learner will able

- To know the concept of genetic engineering
- To know about genome, genomic DNA and genomic DNA library
- To know about construction of a genomic DNA Library
- To discuss the principle of DNA cloning and its components
- To discuss the DNA cloning methods and its applications

What is Genetic Engineering

- In simple words, genetic engineering can be described as the manual addition of a new DNA into an organism.
- It aids the addition of such traits that are not originally found in the organisms.

- ❖ Recombinant DNA is required to create Genetically Modified Organisms (GMO.)
- ❖ An area of chromosome (gene) is spliced.
- ❖ Genetic disorders in humans can be corrected using genetic engineering.
- ❖ Selective breeding has been in the world since ancient times.
- ❖ Jack Williamson used the word 'Genetic Engineering' in his science fiction novel Dragon's Island which was published in 1951.
- ❖ First recombinant DNA molecules were created by an American Biochemist, Paul Berg.

New DNA may be inserted in the host genome by first isolating and copying the genetic material of interest using molecular cloning methods to generate a DNA sequence, or by synthesizing the DNA and then inserting this construct into the host organism. Genes may be removed, or "knocked out", using a nuclease. Gene targeting is a different technique that uses homologous recombination to change an endogenous gene and can be used to delete a gene, remove exons, add a gene, or introduce point mutations.

Applications of Genetic Engineering

Medicine, research, industry and agriculture are a few sectors where genetic engineering applies. It can be used on various plants, animals and microorganisms. The first microorganism to be genetically modified is bacteria.

1. **In Medicine:** Genetic engineering can be applied to:

- Manufacturing of drugs
- Creation of model animals that mimic human conditions and,
- Gene therapy
- Human growth hormones
- Follicle-stimulating hormones
- Human albumin
- Monoclonal antibodies
- Antihemophilic factors
- Vaccines

In Research:

Genes and other genetic information from a wide range of organisms can be inserted into bacteria for storage and modification, creating genetically modified bacteria in the process.

In Industry:

- Transformation of cells in organisms with a gene coding to get a useful protein.
- Medicines like insulin, human growth hormone, and vaccines, supplements such as tryptophan, aid in the production of food (chymosin in cheese making) and fuels are produced using such techniques.

In Agriculture:

- Genetically modified crops are produced using genetic engineering in agriculture.
- Such crops are produced that provide protection from insect pests.
- It is used or can be used in the creation of fungal and virus-resistant crops.

❖ **Genetic engineering can be applied to other areas:**

- Conservation
- Natural area management
- Microbial art

Benefits of Genetic Engineering

- The production of genetically modified crops is a boon to agriculture.
- The crops that are drought-resistant, disease-resistant can be grown with it.
- As described earlier, genetic disorders can be treated.
- The diseases such as malaria, dengue can be eliminated by sterilising the mosquitoes using genetic engineering.
- Therapeutic cloning

Challenges of Genetic Engineering

- The production of genetically-engineered entities may result in an adverse manner and produce undesired results which are unforeseen.
- With the introduction of a genetically-engineered entity into one ecosystem for a desirable result, may lead to distortion of the existing biodiversity.
- Genetically-engineered crops can also produce adverse health effects.

- The concept of genetic-engineering is debated for its bioethics where community against it argues over the right of distorting or moulding the nature as per our needs.

Regulations in India

Genetic Engineering Appraisal Committee (GEAC) is the biotech regulator in India. It is created under the Ministry of Environment and Forests. Read more about GEAC in the linked article. There are five bodies that are authorized to handle rules noted under Environment Protection Act 1986 "Rules for Manufacture, Use, Import, Export and Storage of Hazardous Microorganisms/Genetically Engineered Organisms or Cells 1989". These are:

- Institutional Biosafety Committees (IBSC)
- Review Committee of Genetic Manipulation (RCGM)
- Genetic Engineering Approval Committee (GEAC)
- State Biotechnology Coordination Committee (SBCC) and
- District Level Committee (DLC)

Which are the genetically modified crops in India?

- Bt Cotton is the genetically modified crop that is under cultivation in India.
- Bt Brinjal was initially approved but later was blocked from production.
- GM Mustard is yet to be allowed for cultivated. It will be the first genetically modified food crop in the country.

6.2 What is a Genome?

A genome is an organism's complete set of DNA, including all of its genes. Each genome contains all of the information needed to build and maintain that organism. In humans, a copy of the entire genome is more than 3 billion DNA base pairs which are present in all cell nuclei. The nuclear genome contains protein-coding and non-coding genes, as well as other junk DNA and functional regions of the genome.Most eukaryotes are diploid, which means that each chromosome has two copies in the nucleus, but the term genome refers to just one copy of each chromosome.

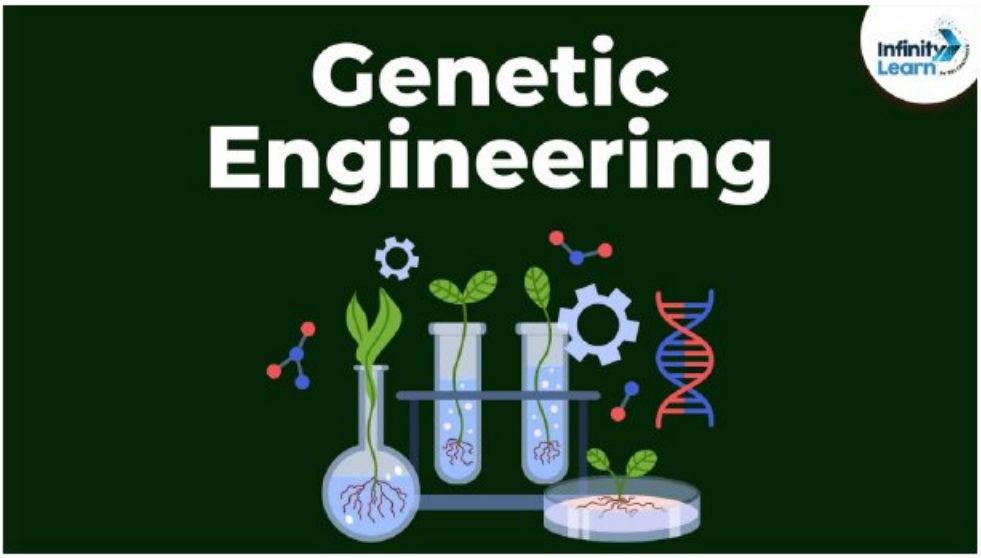


Fig.6.1: Genetic engineering

6.3 Genomic DNA

Genomic DNA constitutes the total genetic information of an organism. The genomes of almost all organisms are DNA, the only exceptions being some viruses that have RNA genomes. Genomic DNA molecules are generally large, and in most organisms are organized into DNA–protein complexes called chromosomes. The size, number of chromosomes, and nature of genomic DNA varies between different organisms. Viral DNA genomes are relatively small and can be single- or double-stranded, linear, or circular. All other organisms have double-stranded DNA genomes. Bacteria have a single, circular chromosome. In eukaryotes, most genomic DNA is located within the nucleus (nuclear DNA) as multiple linear chromosomes of different sizes. Eukaryotic cells additionally contain genomic DNA in the mitochondria and, in plants and lower eukaryotes, the chloroplasts. This DNA is usually a circular molecule and is present as multiple copies within these organelles.

Genomic deoxyribonucleic acid (abbreviated as gDNA) is chromosomal DNA, in contrast to extra-chromosomal DNAs like plasmids. Most organisms have the same genomic DNA in every cell; however, only certain genes are active in each cell to allow for cell function and differentiation within the body. gDNA predominantly resides in the cell nucleus packed into dense chromosome structures. Chromatin refers to the combination of DNA and proteins that make up chromosomes. When a cell is not dividing, chromosomes exist as loosely packed chromatin mesh.

The genome of an organism (encoded by the genomic DNA) is the (biological) information of heredity which is passed from one generation of organism to the next. That genome is transcribed to produce various RNAs, which are necessary for the function of the organism. Precursor mRNA (pre-mRNA) is transcribed by RNA polymerase II in the nucleus. Pre-mRNA is then processed by splicing to remove introns, leaving the exons in the mature messenger RNA (mRNA). Additional processing includes the addition of a 5' cap and a poly(A) tail to the pre-mRNA. The mature mRNA may then be transported to the cytosol and translated by the ribosome into a protein. Other types of RNA include ribosomal RNA (rRNA) and transfer RNA (tRNA). These types are transcribed by RNA polymerase I and RNA polymerase III, respectively, and are essential for protein synthesis. However 5s rRNA is the only rRNA which is transcribed by RNA Polymerase III.

6.4 Genomic DNA Library

A genomic library or gene bank is a complete collection of cloned DNA fragments that constitutes the entire genome of an organism. It represents all the genes – expressed, non-expressed, intron, exons, etc. Genomic libraries can be kept for many years and the copies can be used for research purposes.

6.4.1 Advantages of a Genomic Library

- Genomic libraries derived from eukaryotic organisms are critical for studying the genome sequence of a particular gene of interest.
- It is useful for prokaryotes with small genomes to identify a clone encoding a specific gene of interest.
- It helps researchers to explore more about an organism's genomic structure and function. It is also used to study genetic mutations.
- Pharmaceutically important genes can also be identified by this method.

6.4.2 Disadvantages of a Genomic Library

- It requires sophisticated software and a vast amount of computing power. Also, the process is prone to errors.
- Eukaryotic genome libraries with very large genomes contain many DNA that do not code for proteins, as well as non-coding DNA like repetitive DNA and regulatory regions, making them less than ideal.

6.4.3. Construction of a Genomic DNA Library

The Genomic Library or gene bank is constructed by a shotgun experiment where the entire genome of the cell is cloned in the form of random and unidentifiable clones. It uses the chain termination method (Sanger's sequencing) to sequence the DNA molecules. The DNA clones are produced by the following steps:

- Isolating the DNA fragments that are to be cloned and joining them with suitable vectors like the lambda phage.
- Now, it is introduced into the host cell at high efficiency to get a large number of independent clones.
- Finally, the desired clones are selected and used for the construction of the genomic library.

6.4.4. Steps to Construct a Genomic DNA Library

1. First, the purification of the desired eukaryotic cell nuclei is done and this is accomplished through protease digestion and organic extraction.
2. The derived genomic DNA is too large to be incorporated into a vector and must be fragmented into desired sizes. Both physical and enzymatic methods can be used to fragment DNA.
3. There are several vectors available for cloning large DNA fragments. Phage, bacterial artificial chromosome, yeast artificial chromosome, and other such vectors are suitable for larger DNA. The λ replacement vectors are the most preferred ones for the construction of a genomic DNA library.
4. Usually, the T4 DNA ligase enzyme is used to ligate the chosen DNA sequence into the vector.
5. The recombinant vectors and the insert combinations are grown in a bacterial host cell (E. coli). They replicate their genome along with the vector genome contained within them.
6. The collection of clones that contain all the sequences from the original source (including the sequence of interest) forms the gene bank or genomic library.

6.5. Screening of Clone

Each transformed bacterial host cell in a library will have only one vector with one DNA insert. The entire library can be plated over media on a filter. The filter, as well as colonies, are then hybridised, labelled with a probe and identified using detection methods such as autoradiography. Other techniques like PCR and the blue-white selection method can also be used to screen the clone. **DNA cloning is a method used to produce multiple identical copies of a DNA fragment within a cell. DNA cloning is also known as gene cloning or molecular cloning. All three terms are used**

interchangeably to describe the single technique. DNA cloning has opened up new possibilities in areas like genetic engineering and biomedical research. With DNA cloning, researchers can study gene function and explore different biological processes.

6.6. Principle of DNA Cloning

The principle of DNA cloning involves the production of multiple copies of a particular DNA fragment of interest. It involves inserting the desired DNA fragment into a cloning vector, usually a plasmid, to create a recombinant DNA molecule which is then introduced into host cells through transformation. The transformed cells are selected and cultured on selective media, allowing for the replication of the inserted DNA fragment. This results in the production of multiple copies of the desired DNA, which can be isolated for further analysis.

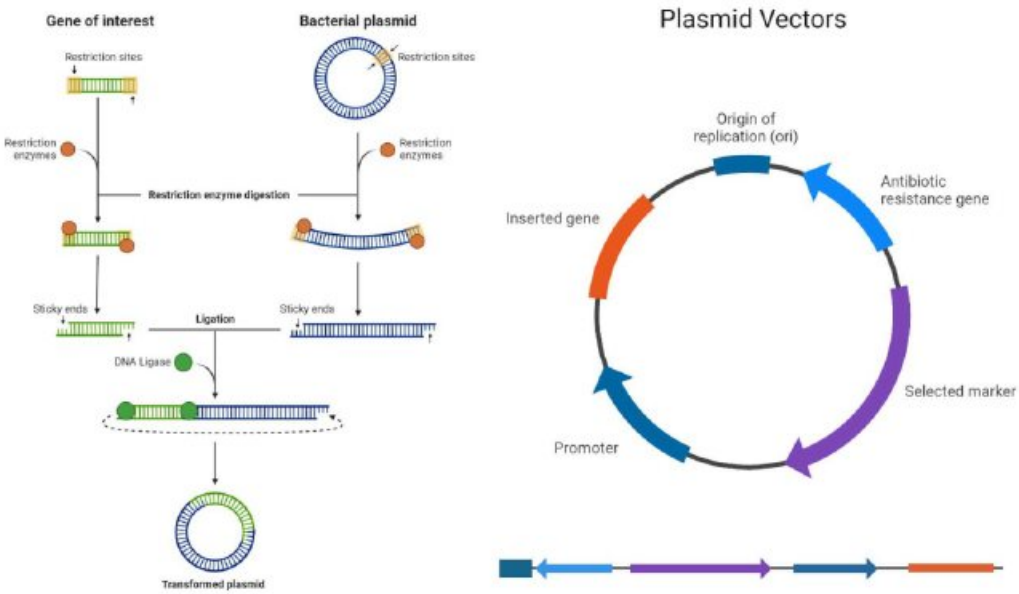


Fig.6.2: Restriction Enzymes Cloning Steps

6.6.1 Steps in DNA Cloning

The process of DNA cloning can be divided into the following five steps:

1. Preparation of gene of interest and vector

- The first step in DNA cloning is to obtain the gene of interest, which contains the desired DNA sequence to be cloned.

- For simple organisms, such as bacteria, the DNA fragment can be obtained by digesting the genomic DNA using restriction enzymes. More complex organisms, such as mammals, use alternative methods such as reverse transcription of mRNA or PCR amplification. In some cases, physical methods like sonication or shearing can be used to fragment DNA.
- The gene of interest is inserted into a vector and the most commonly used vector is the plasmid, a circular DNA molecule commonly found in prokaryotes.
- Both the vector and the gene of interest are cut using the same or compatible restriction enzymes. Restriction enzymes identify specific DNA sequences and cut the DNA at those sites.

2. Ligation of the gene of interest and vector

- After digestion with restriction enzymes, the vector and the gene of interest can be joined together to form recombinant DNA (rDNA) using the enzyme DNA ligase.
- DNA ligase works by recognizing and binding to the ends of the DNA fragments that are cut by the restriction enzymes. It then catalyzes the formation of new phosphodiester bonds, joining the DNA fragments.

3. Transformation

- The next step is transformation where the rDNA is introduced into a host cell that takes up and expresses the inserted gene of interest.
- Before the transformation, the host cells must be made competent, which means they are made capable of taking up DNA through their membrane.
- Different methods are used to make host cells competent. One common approach involves using cold calcium chloride, followed by a brief heat shock.
- Another method for achieving competency is through electroporation. In electroporation, the host cells are exposed to an electric field, which creates temporary pores in the cell membrane, making it more permeable to DNA molecules.
- Once the host cells have been made competent, the recombinant plasmid is mixed with the competent host cells. The transformed host cells then take up the recombinant plasmid and incorporate it into their own genetic material.

4. Selection/screening and culturing of transformed cells

- After the transformation of host cells with the recombinant plasmid, the next step is the selection or screening of the transformed cells.

- The transformed host cells are plated onto a nutrient agar medium that contains a specific antibiotic. The antibiotic is chosen based on the antibiotic resistance gene present in the recombinant plasmid.
- Successfully transformed cells will contain genes with antibiotic resistance, allowing them to grow and form colonies on the selective media.

5. Isolation of recombinant DNA

- Once colonies of transformed cells have formed on the agar plate, the rDNA from the culture can be isolated.
- A single colony of the transformed cells is selected from the agar plate and cultured in a liquid nutrient medium.
- During this process, the host cells will multiply, and the recombinant plasmid, along with its inserted gene of interest, will also be replicated, producing multiple copies of the rDNA.
- The isolated rDNA can be further analysed and used for various applications, such as protein expression or genetic engineering experiments.

6.7. Components of DNA Cloning

1. Cloning vector

- A cloning vector is a DNA molecule used as a carrier to insert a specific segment of foreign DNA into a host cell for cloning.
- In order to be used as a cloning vector, a DNA molecule must have specific characteristics. One important feature is the ability to replicate within the host cell. It should also be small, usually less than 10 kilobases (kb), for easy handling and stability.
- A cloning vector also requires a suitable cloning site and a selectable marker that is recognized by specific restriction enzymes to insert DNA fragments into the vector.

There are different types of cloning vectors. Some commonly used cloning vectors are:

1. **Plasmids** are circular, autonomously replicating DNA molecules widely used as cloning vectors. They can hold DNA inserts of sizes up to around 15 kb.
2. **Bacteriophages**, also known as phages, are viruses that infect bacteria. Bacteriophages like lambda (λ) and M13 are often used as cloning vectors.
3. **Cosmids** are hybrid vectors that combine the characteristics of both plasmids and bacteriophages. They are more stable than regular plasmids.
4. **Bacterial artificial chromosomes (BACs)** are large cloning vectors used for cloning DNA sequences in bacterial cells and they can hold DNA segments up to 350 kb.

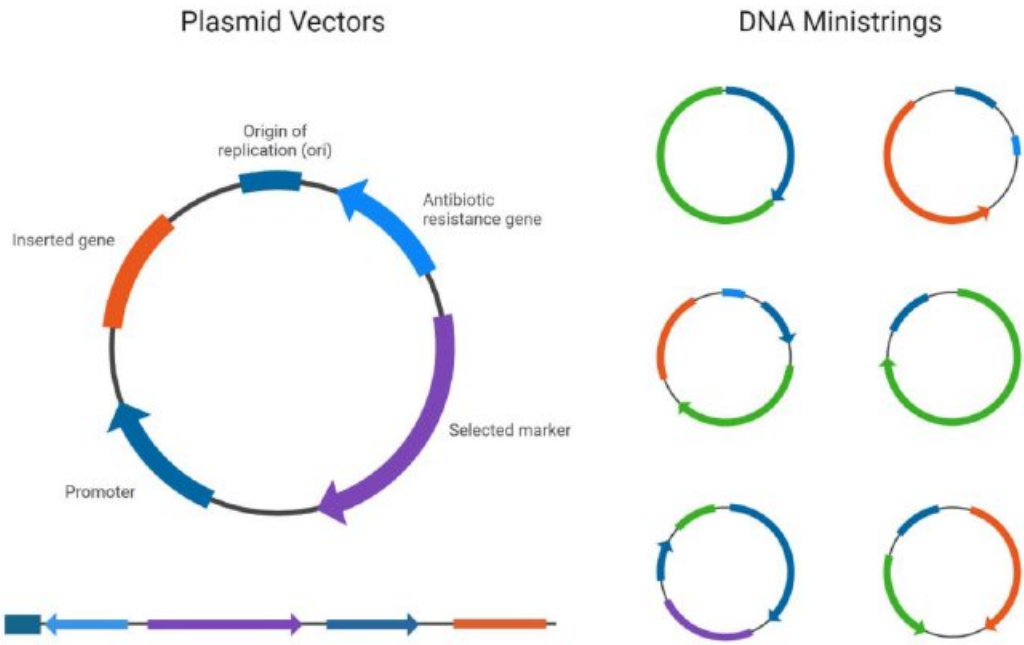


Fig.6.3: DNA Cloning Vectors

Yeast artificial chromosomes (YACs) are vectors used for cloning DNA fragments larger than 1 megabase (1 Mb) in size. They are commonly used in genome mapping and sequencing projects.

2. Restriction enzymes

- Restriction enzymes, also called restriction endonucleases, are enzymes produced by bacteria that recognize and cut DNA sequences at unique sites called recognition sites.
- Different restriction enzymes have different cutting patterns, which result in either sticky ends or blunt ends.
- Sticky ends result in overhanging single-stranded DNA sequences that can easily be joined with other DNA fragments that are cut by the same enzyme. On the other hand, blunt ends have no overhangs, requiring additional techniques or enzymes for successful ligation.

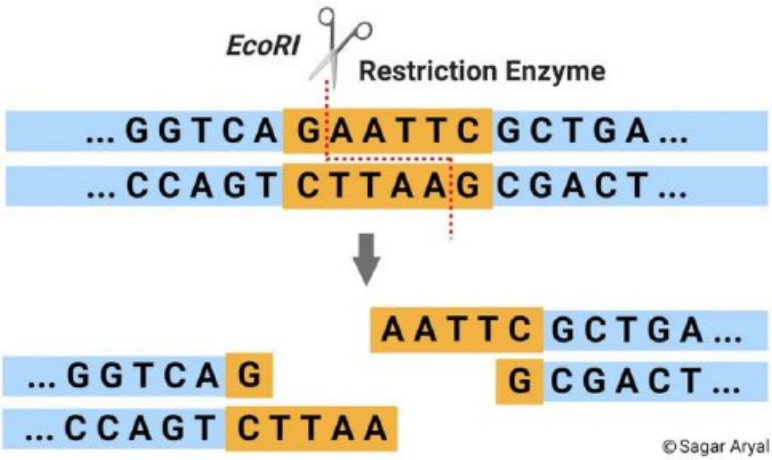


Fig.6.4: Restriction enzymes

Questions for Practice

Q. 1 What is a cDNA library?

Q. 2 How is the cDNA library constructed?

Q. 3 What is a cloning vector?

Q. 4 What is a Human Genome Project (HGP)?

6.8. DNA cloning methods

There are several methods of DNA cloning. Some of the popular cloning methods are:

1. Traditional Cloning

- Traditional cloning, also called restriction enzyme-based cloning, uses restriction enzymes to cut the DNA insert and vector at specific restriction sites.
- The DNA insert should not contain any internal restriction sites that are similar to the ones present on the plasmid as it could result in the production of unwanted smaller DNA fragments.
- After the DNA fragments have been cut by the restriction enzymes, DNA ligase is used to join the DNA insert with the vector.

2. PCR Cloning

- PCR cloning is a type of cloning that involves the direct ligation of DNA fragments, obtained through PCR amplification, into a vector without the need for cutting the insert using restriction enzymes.

- There are several types of PCR cloning methods. One popular method of PCR cloning is TA cloning.
- In TA cloning, Taq polymerase adds an adenine (A) residue to the 3′ ends of PCR products, creating "A-tailed" DNA fragments. These fragments are directly ligated with "T-tailed" vectors having thymidine (T) residues at their ends using DNA ligase.

3. Ligation-Independent Cloning (LIC)

- Ligation-independent cloning (LIC) is a method where specific short sequences are added to the ends of a DNA insert to match the sequences on a vector.
- The 3′ ends of the DNA fragment are trimmed using the enzymes with 3′ to 5′ exonuclease activity which creates cohesive ends between the DNA insert and the vector.
- The vector and insert molecules are combined. The resulting plasmid contains four single-stranded DNA nicks which are repaired by the host during transformation.
- An important advantage of LIC is that it maintains the original sequence integrity without introducing any additional elements.

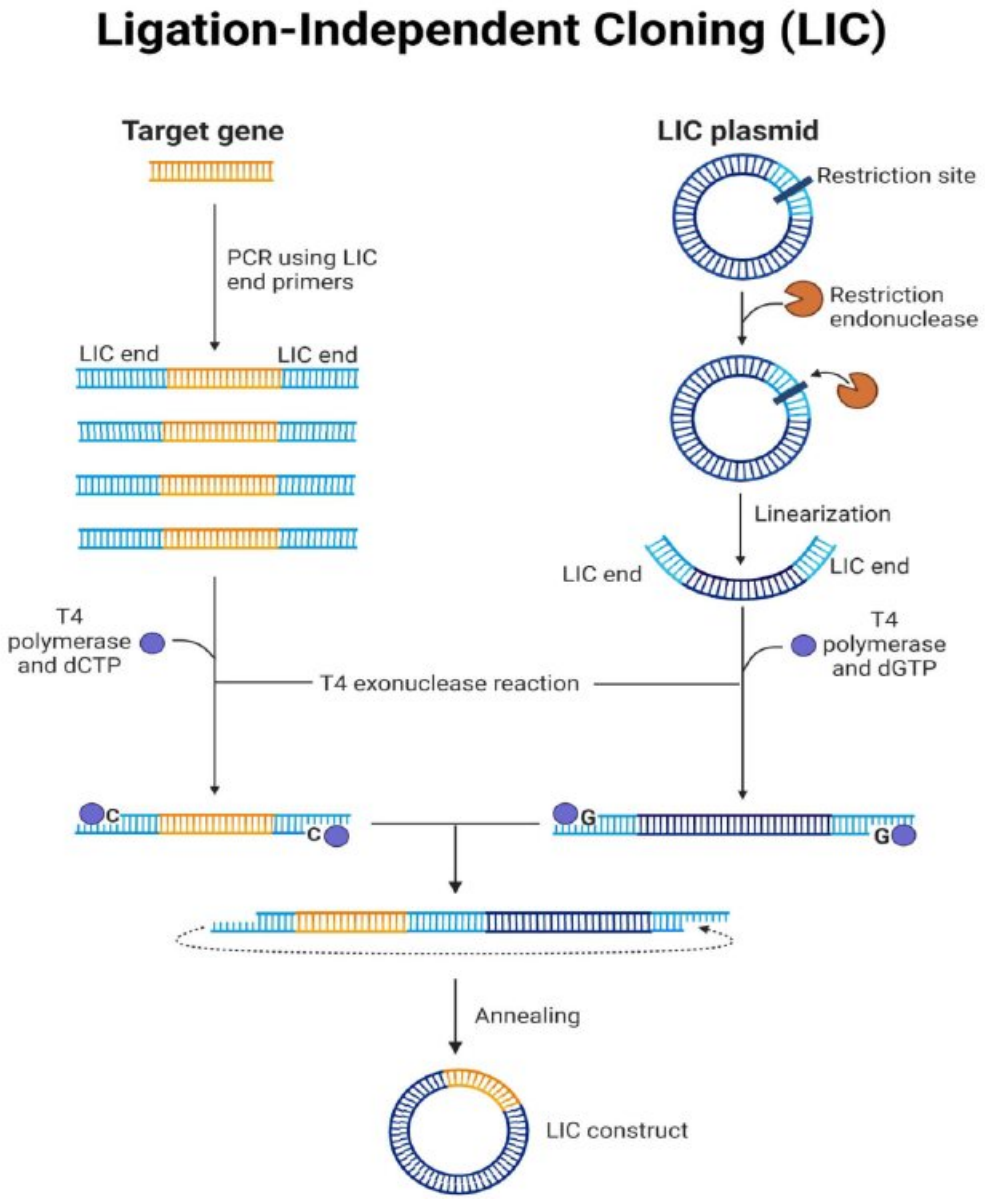


Fig. 6.5: Ligation Independent Cloning (LIC)

4. Seamless Cloning (SC)

- Seamless cloning (SC) is a method that relies on matching short sequences at the ends of a DNA fragment with corresponding short sequences on a vector. It is similar to the LIC method.
- In SC, an enzyme with 5′ to 3′ exonuclease activity is used to create 3′ overhangs on the DNA fragment.
- The advantage of SC over traditional cloning is that they enable the insertion of multiple DNA fragments into a vector.

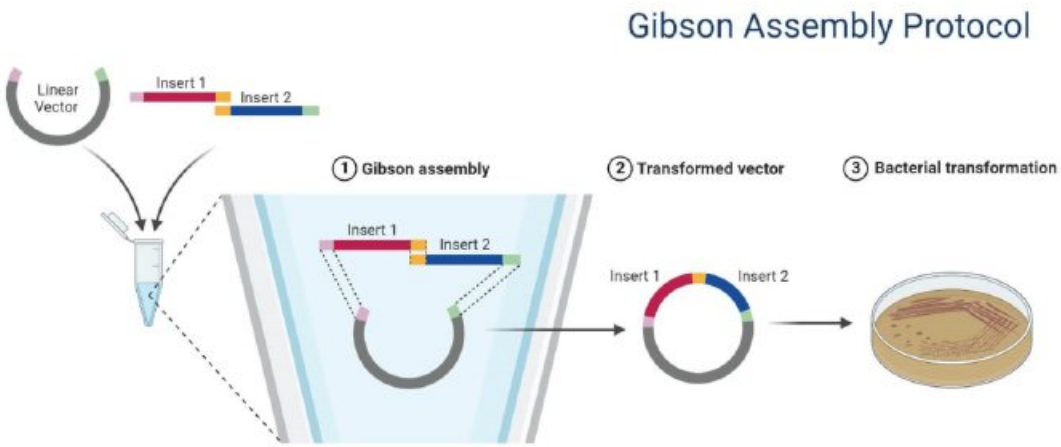


Fig.6.6: Gibson Assembly Cloning

5. Recombinational Cloning

- Recombinatorial cloning involves the use of site-specific DNA recombinases that facilitate the exchange and recombination of DNA fragments at specific recombination sites.
- The process begins by inserting a DNA fragment into an entry vector, creating an entry clone. Once the entry clone is obtained, it is recombined with a destination clone.
- Recombinatorial cloning provides an efficient way to create complex DNA constructs by allowing easy transfer of DNA fragments between different vectors through site-specific recombination.

6.9. Applications of DNA Cloning

DNA cloning has many applications in various fields of research. Some major applications along with examples of DNA cloning are:

- DNA cloning is useful for studying the gene functions of specific genes in different organisms. For example, the cloning of the green fluorescent protein (GFP) gene from jellyfish has allowed the visualization of protein expression in living cells.
- DNA cloning has been used for producing recombinant proteins in large quantities. For example, the cloning of the human insulin gene led to the large-scale production of insulin for the treatment of diabetes, reducing dependence on animal-derived insulin.
- DNA cloning plays an important role in genetic engineering to create genetically modified organisms (GMOs), which introduces desired genes into organisms to modify their traits. For

example, cloning genes to create genetically modified crops with improved traits such as pest resistance and higher yield.

- DNA cloning is also useful in gene therapy, where therapeutic genes are cloned and used to treat genetic diseases.
- DNA cloning techniques are also used in forensic analysis. Cloning specific DNA regions can be used in the amplification and analysis of genetic markers to determine an individual's identity in forensic investigations.

Challenges and Limitations of DNA Cloning

DNA cloning has brought significant advancements in various fields, but it also has its limitations that need to be considered. Some of the challenges and limitations are:

- Traditional DNA cloning can be time-consuming, especially when working with large DNA fragments. It may take several days to complete steps such as culturing and restriction digestion.
- Another limitation is the potential for contamination during the cloning process.
- DNA cloning can be costly and labour-intensive due to the reagents, enzymes, and equipment required.
- In order to ensure successful cloning, the compatibility between the insert and vector needs to be considered.

Ethical Considerations in DNA Cloning

DNA cloning raises several ethical concerns. To ensure ethical practices in DNA cloning, it is important to address and take into account these ethical concerns and considerations.

- One of the concerns is genetic modification, which raises questions about the potential consequences for organisms and ecosystems.
- Introducing cloned or genetically modified organisms (GMOs) into the environment can have unintended environmental impacts that need to be carefully assessed.
- Another ethical issue is patenting and commercialization of genetic resources, which may negatively impact scientific research and access to genetic information.
- Privacy of genetic information is also an important consideration, with concerns about confidentiality and potential misuse of individuals' genetic data. Informed consent is also crucial when human subjects are involved in cloning research.

6.10. Library amplification

Depending on the need for amplification, DNA library preparation methods can be categorized as PCR-free or PCR-based. In either method, care should be taken to follow protocols that yield highly diverse and representative libraries of input samples from different amounts to help generate high-quality data.

(i) PCR-free libraries

Since PCR amplification can contribute to GC bias, PCR-free library preparation is usually the preferred method to create libraries covering high-GC or high-AT sequences, to help ensure library diversity. Note that even with PCR-free library preparation methods, bias can be introduced during cluster generation and from the chemistry of the sequencing step itself. Compared to PCR-based methods, PCR-free libraries require higher input amounts of starting material (although improvements have been made in lowering the input requirements). This can be challenging in scenarios such as using limited or precious samples and highly degraded nucleic acids. With PCR-free libraries, accurate assessment of library quality and quantity may be difficult, compared to PCR-amplified libraries. Nevertheless, better representation and balanced coverage offered by PCR-free libraries make them attractive for the following applications:

- Studies of population-scale genomics and molecular basis of a disease
- Investigation of promoters and regulatory regions in the genome, which often are high in GC or AT content
- Whole-genome sequencing analysis and variant calling for single-nucleotide polymorphisms (SNPs) and small insertions or deletions (indels)

(ii) PCR-based libraries

The PCR-based method is a popular strategy for constructing NGS libraries, since it allows lower sample input and selective amplification of inserts with adapters at both ends. However, PCR can introduce GC bias, leading to challenges in data analysis. For example, GC bias may hinder *de novo* genome assembly and single-nucleotide polymorphism (SNP) discovery. A number of factors can impact GC bias, and the following factors should be considered to achieve balanced library coverage:

- PCR enzyme and master mix used
- Number of PCR cycles run, and cycling conditions
- PCR additives or enhancers in the reaction

With a given PCR enzyme or master mix, an increase in the number of PCR cycles usually increases GC bias. Therefore, a general recommendation is to run the minimum number of cycles that generates sufficient library yields for sequencing. Decreasing the number of PCR cycles also reduces PCR duplicates and improves library complexity. PCR duplicates are defined as sequencing reads resulting from two or more PCR amplicons of the same DNA molecule. Although bioinformatic tools are available to identify and remove PCR duplicates during data analysis, minimizing PCR duplicates is important for efficient use of the flow cell in sequencing. Other PCR artifacts can also result in reduced library quality and complexity. These artifacts include amplification bias (due to PCR stochasticity), nucleotide errors (from enzyme fidelity), and PCR chimeras (due to enzyme's template switching)

Library quantification approaches

Before NGS libraries are loaded onto the sequencer, they should be quantified and normalized so that each library is sequenced to the desired depth with the required number of reads. Concentrations of prepared NGS libraries can vary widely because of differences in the amount and quality of nucleic acid input, as well as the target enrichment method that may be used. While underclustering due to overestimated library concentrations can result in diminished data output, overclustering can result in low quality scores and problematic downstream analysis.

(i) Microfluidics-based quantitation

Microfluidic electrophoresis separates fragments in NGS libraries based on size and can estimate the quantity of different size ranges using a reference standard. More commonly, however, the results of fragment analysis obtained by this method are used in conjunction with the two other methods listed below for more accurate quantitation of NGS libraries.

(ii) Fluorometry-based quantitation

The fluorometric assay uses fluorescent dyes that bind specifically to double-stranded DNA (dsDNA) to determine library concentration. After a short incubation of samples with a dye, the samples are read in a fluorometer, and library concentrations are calculated by (built-in) analysis software. Although the workflow is simple and takes only a few minutes per sample, this method may not scale well above 20–30 samples because samples are often read one at a time. Nevertheless, flexible input volumes and short incubation times allow for quick and easy testing of prepared libraries

for concentrations. Since the measured concentration is for total dsDNA, the average size distribution of the libraries should be taken into account for accurate quantitation.

(iii) qPCR-based quantitation

The qPCR-based assay quantifies NGS libraries by amplifying DNA fragments with the P5 and P7 adapters. A qPCR standard curve is used to determine a broad range of library concentrations, even as low as femtomolar. Since the PCR primers are designed specifically to bind to the adapter sequences, the qPCR assays detect only properly adapted, amplifiable libraries that can form clusters during sequencing. Note, though, that qPCR can also amplify adapter dimers; therefore, melting curve analysis and/or fragment size analysis should be performed to assess specificity and accuracy of quantitation by qPCR. The final library concentration is calculated based on the following formula.

6.11. Summary

Under this unit we have discussed genomic DNA library with different steps, principle of DNA cloning with different components, applications of DNA cloning and library amplification etc. A genome is an organism's complete set of DNA, including all of its genes. Each genome contains all of the information needed to build and maintain that organism. A genomic library or gene bank is a complete collection of cloned DNA fragments that constitutes the entire genome of an organism. It represents all the genes – expressed, non-expressed, intron, exons, etc. Genomic libraries can be kept for many years and the copies can be used for research purposes.

DNA cloning is a method used to produce multiple identical copies of a DNA fragment within a cell. DNA cloning is also known as gene cloning or molecular cloning. All three terms are used interchangeably to describe the single technique. DNA cloning has opened up new possibilities in areas like genetic engineering and biomedical research. With DNA cloning, researchers can study gene function and explore different biological processes.

DNA testing can also be used to identify pathogens, identify biological remains in archaeological digs, trace disease outbreaks, and study human migration patterns. In the medical field, DNA is used in diagnostics, new vaccine development, and cancer therapy. DNA cloning can be used to make proteins such as insulin with biomedical techniques. It is used to develop recombinant versions of the non-functional gene to understand the functioning of the normal gene. This is applied in gene therapies also. It helps to analyse the effect of mutation on a particular gene.

6.12. Terminal Questions

Q. 1 What do you mean by genomic DNA? Explain it.

Answer:

Q. 2 Describe the advantages and disadvantages of a genomic library.

Answer:

Q. 3 Describe principle of DNA cloning with different steps involved in it.

Answer:

Q. 4 What are the different components of DNA Cloning?

Answer:

Q. 5 Explain applications of DNA Cloning.

Answer:

Q. 6 Write short note on the followings.

(i) Genomic DNA Library

(ii) Library amplification

Answer:

Q. 7 Write short note on the followings.

(i) Screening of clone

(ii) Genome

Answer:

6.13. Further Readings

1. Biochemistry- Lehninger A.L.
2. Biochemistry –J.H.Weil.
3. Biochemistry fourth edition-David Hames and Nigel Hooper.
4. Textbook of Biochemistry for Undergraduates - Rafi, M.D.
5. Biochemistry and molecular biology- Wilson Walker.

Uttar Pradesh Rajarshi Tandon
Open University

PGBCH - 118 N
Genetic Engineering

Block- 3

Genetic Engineering- II

Course Design Committee

Dr. (Prof.) Ashutosh Gupta, School of Science, UPRTOU, Prayagraj	Chairman
Prof. Prof. Umesh Nath Tripathi Department of chemistry Deen Dayal Upadhyaya Gorakhpur University, Gorakhpur	Member
Prof. S.I. Rizvi Department of Biochemistry University of Allahabad, Prayagraj	Member
Prof. Dinesh Yadav Department of Biotechnology Deen Dayal Upadhyaya Gorakhpur University, Gorakhpur	Member
Prof. Sharad Kumar Mishra Department of Biotechnology Deen Dayal Upadhyaya Gorakhpur University, Gorakhpur	Member
Dr. Ravindra Pratap Singh Assistant Professor (Biochemistry) School of Science, UPRTOU, Prayagraj	Member
Dr. Dharmveer Singh Assistant Professor (Biochemistry) School of Science, UPRTOU, Prayagraj	Course Coordinator

Course Preparation Committee

Dr. Gopal Dixit Assistant Professor Department of Botany, Upadhi Mahavidyalaya, Pilibhit, U.P.	**Author**	**Block-1-2**	Unit: 1-5
Dr. Arun Kumar Pandey Assistant Professor Department of Botany, PSMPG College, Maharajganj. U.P.	**Author**	**Block-2-3**	Unit: 6-9
Dr. Sadhana Singh Assistant Professor- Biochemistry School of Sciences, UPRTOU, Prayagraj	**Author**	**Block-4**	Unit: 10, 12
Dr. Anuradha Singh Assistant Professor- Botany School of Sciences, UPRTOU, Prayagraj	**Author**	**Block-4**	Unit: 11
Dr. Mohd. Khalid Masroor Retd. Associate Professor-Botany, University of Allahabad, U.P.	**Editor**	(Block- 01, 02, 03&04, Unit: 1, 2,3,4,5, 10, 11, &12)	
Dr. Rajiv Ranjan Associate Professor, MLKPG College, Balrampur, U.P.	**Editor**	(Block- 02& 03)	Unit: (6-9)
Dr. Dharmveer Singh Assistant Professor (Biochemistry) School of Sciences, UPRTOU, Prayagraj			**(SLM & Course Coordinator)**

PGBCH – 118, Genetic Engineering
©UPRTOU , 2024
ISBN :

Printed By: K.C.Printing & Allied Works, Panchwati, Mathura -281003.

Introduction

The following three units are included in the first block of genetic engineering are as:

Unit-7: This unit covers Polymerase chain reaction (PCR).The polymerase chain reaction enables investigators to obtain the large quantities of DNA that are required for various experiments and procedures in molecular biology, forensic analysis, evolutionary biology, and medical diagnostics. The basic PCR amplifies specific DNA sequences using cycles of denaturation, annealing, and extension. RT-PCR converts RNA to cDNA before amplification. PCR applications in genetic engineering include gene cloning, mutation detection, and expression analysis is discussed.

Unit-8: This unit covers the DNA Finger printing, DNA fingerprinting is a laboratory technique used to determine the probable identity of a person based on the nucleotide sequences of certain regions of human DNA that are unique to individuals. The History of finger printing, DNA markers, Minisatellites, Microsatellites, Methods and applications are discussed in this unit.

Unit-9: This unit covers the Screening and analysis of recombinants. Chromogenic substrates produce color changes in reactions to identify gene expression. Genetic selection methods isolate cells with desired traits. Nucleic acid probes detect specific sequences. Chromosome walking and jumping map genes by sequentially or bypassing regions. Screening cloned banks identifies clones of interest, while immunological screening detects expressed proteins. Blotting techniques (e.g., Southern, Northern, Western) analyze nucleic acids and proteins are discussed.

Unit- 7: Polymerase Chain Reaction (PCR)

Structure

7.1 Introduction

The development of the bioanalytical techniques brought a progressive discipline for which the future holds many exciting opportunities to further improvement. The main impact of bionalysis in the pharmaceutical industry is to obtain a quantitative measure of the drug and its metabolites. The purpose is to perform the pharmacokinetics, toxicokinetics, and bioequivalence and exposure response like pharmacokinetic/pharmacodynamic studies. Various bioanalytical techniques are performed in bioanalytical studies such as hyphenated techniques, chromatographic techniques, and ligand binding assays. This review extensively highlights the role of bioanalytical techniques and hyphenated

instruments in assessing the bioanalysis of the drugs. The model of a generalized instrumentation system necessitates only four parts: a measurand, a sensor, a signal processor, and an output display. More complicated instrumentation devices may also designate function for data storage and transmission, calibration, or control and feedback. However, at its core, an instrumentation systems convert's energy or information from a physical property not otherwise perceivable, into an output display that users can easily interpret.

The field of bioanalysis has matured significantly from early studies in drug metabolism using many simple and advanced techniques, and in today's Bioanalyst is well equipped to deal with the modern challenges. A bioanalytical method is a set of procedures involved in the collection, processing, storage, and analysis of a biological matrix for a chemical compound. Bioanalytical method validation (BMV) is the process used to establish that a quantitative analytical method is suitable for biochemical applications. Bioanlysis covers the quantitative measurement of Xenobiotics of drugs such as their metabolites, and biological molecules in unnatural locations or concentrations and Biotics like macromolecules, proteins, DNA, large molecule drugs, metabolites in biological systems. Bioanalysis is a progressive discipline for which the future holds many exciting opportunities to further improve sensitivity, specificity, accuracy, efficiency, assay throughput, data quality, data handling and processing, analysis cost and environmental impact.

Objectives

After reading this unit, the learner will able to understands:

- To know about Polymerase Chain Reaction (PCR) & its requirment
- To know the procedure of PCR with different steps involved in it.
- To discuss different types of PCR with its Applications
- To discuss RT-PCR, qPCR and hot-start PCR

7.2 Polymerase Chain Reaction (PCR)

7.2.1 What is PCR?

Is an ingenious technology in molecular biology to amplify a single or a few copies of a piece of DNA across several orders of magnitude, generating thousands to millions of copies of a particular DNA sequence in vitro. This PCR process, invented by Kary Mullis in 1984 in California. The purpose of a PCR is to make a huge number of copies of a gene. The invention of the PCR (Polymerase Chain Reaction) technique has resulted in a revolution for evolutionary biologists interested in genetic

questions. Suddenly there was a fast, robust and relatively inexpensive technique to get hold of genetic information from small samples of e.g. skin, blood or faeces. An advantage with PCR based molecular studies of DNA is that, once DNA is extracted and purified, the techniques are very similar regardless of the taxonomy of the study organisms.

The technique was made possible by the discovery of Taq polymerase, the DNA polymerase that is used by the bacterium (Thermus aquaticus) that was discovered from Yellow Stone National Park hot springs. This DNA polymerase is stable at the high temperatures need to perform the amplification, whereas other DNA polymerases become denatured. The method relies on thermal cycling instrument (thermal cycler), which provide program consisting of cycles of repeated heating and cooling of the reaction for DNA melting and enzymatic replication of the DNA

The polymerase chain reaction (PCR) is a method widely used to make millions to billions of copies of a specific DNA sample rapidly, allowing scientists to amplify a very small sample of DNA (or a part of it) sufficiently to enable detailed study. PCR was invented in 1983 by American biochemist Kary Mullis at Cetus Corporation. Mullis and biochemist Michael Smith, who had developed other essential ways of manipulating DNA, were jointly awarded the Nobel Prize in Chemistry in 1993.

Fig. 7.1 An older, three-temperature thermal cycler for PCR

PCR is fundamental to many of the procedures used in genetic testing and research, including analysis of ancient samples of DNA and identification of infectious agents. Using PCR, copies of very small amounts of DNA sequences are exponentially amplified in a series of cycles of temperature changes. PCR is now a common and often indispensable technique used in medical laboratory research for a broad variety of applications including biomedical research and forensic science.

The majority of PCR methods rely on thermal cycling. Thermal cycling exposes reagents to repeated cycles of heating and cooling to permit different temperature-dependent reactions specifically, DNA melting and enzyme-driven DNA replication. PCR employs two main reagents primers (which are short single strand DNA fragments known as oligonucleotides that are a complementary sequence to the target DNA region) and a thermostable DNA polymerase. In the first step of PCR, the two strands of the DNA double helix are physically separated at a high temperature in a process called nucleic acid denaturation. In the second step, the temperature is lowered and the primers bind to the complementary sequences of DNA. The two DNA strands then become templates for DNA polymerase to enzymatically assemble a new DNA strand from free nucleotides, the building blocks of DNA. As PCR progresses, the DNA generated is itself used as a template for replication, setting in motion a chain reaction in which the original DNA template is exponentially amplified.

Almost all PCR applications employ a heat-stable DNA polymerase, such as *Taq* polymerase, an enzyme originally isolated from the thermophilic bacterium *Thermus aquaticus*. If the polymerase used was heat-susceptible, it would denature under the high temperatures of the denaturation step. Before the use of *Taq* polymerase, DNA polymerase had to be manually added every cycle, which was a tedious and costly process.

Applications of the technique include DNA cloning for sequencing, gene cloning and manipulation, gene mutagenesis; construction of DNA-based phylogenies, or functional analysis of genes; diagnosis and monitoring of genetic disorders; amplification of ancient DNA; analysis of genetic fingerprints for DNA profiling (for example, in forensic science and parentage testing); and detection of pathogens in nucleic acid tests for the diagnosis of infectious diseases.

7.2.2 Requirements of PCR Technology

- The thermal cycler heats and cools the reaction tubes to achieve the temperatures required at each step of the reaction.

- DNA template that contains the DNA region (target) to be amplified, the source of DNA for the PCR amplification. We use a standard concentration at 25ng/μl. DNA purity is very important, because template contaminants (i.e. excesses of phenolic compounds, EDTA) may leads to PCR inhibition and give false-negative results.
- Pair of primers: short artificial DNA fragments containing sequences complementary to the target region, that are complementary to the 3' (three prime) ends of each of the sense(forward) and the 5' end of the anti-sense (reveres) strand of the DNA target, usually the length of 18-30 nt.

7.2.3. Master Mix Which Contains

- Taq DNA polymerase the enzyme that puts the free nucleotides together. It starts at the 3'end of the primer, and uses the complementary DNA strain as a template.
- Deoxynucleoside triphosphates (dNTPs, sometimes called "deoxynucleotide triphosphates"; nucleotides containing triphosphate groups), Free nucleotides (G, A, T,C) of which the artificial DNA copies are made, the building-blocks from which the DNA polymerase synthesizes a new DNA strand.
- Buffer solution, maintains pH and ionic strength of the reaction solution suitable for the activity of the enzyme.
- Mg^{++} ions - cofactor of the enzyme.
- Free nuclease water: must be present for the reaction to work.

7.2.4 Procedure:

Typically, PCR consists of a series of 25-40 repeated temperature changes, called cycles, each cycle of PCR includes steps for template denaturation, primer annealing and primer extension:

Initialization step:

This step consists of heating the reaction to a temperature of 94–96 °C (or 98 °C if extremely thermostable polymerases are used), which is held for 1–9 minutes. It is only required for DNA polymerases that require heat activation by hot-start PCR.

Denaturation step:

This step is the first regular cycling event and consists of heating the reaction to 94–98 °C for 30sec.-1min. It causes DNA melting of the DNA template by disrupting the hydrogen bonds between complementary bases, yielding single-stranded DNA molecules.

Annealing step:

The reaction temperature is lowered to 50–65 °C for 30 sec-1min allowing annealing of the primers to the single-stranded DNA template. Typically, the annealing temperature is about 3-5 degrees Celsius below the Tm of the primers used. Stable DNA-DNA hydrogen bonds are only formed when the primer sequence very closely matches the template sequence. The polymerase binds to the primer-template hybrid and begins DNA formation.

Extension/elongation step:

The temperature at this step depends on the DNA polymerase used; Taq polymerase has its optimum activity temperature at 75–80 °C, and commonly a temperature of 72 °C is used with this enzyme. At this step the DNA polymerase synthesizes a new DNA strand complementary to the DNA template strand by adding dNTPs that are complementary to the template in 5' to 3' direction, condensing the 5'-phosphate group of the dNTPs with the 3'-hydroxyl group at the end of the nascent (extending) DNA strand. The extension time depends both on the DNA polymerase used and on the length of the DNA fragment to be amplified. As a rule-of-thumb, at its optimum temperature, the DNA polymerase will polymerize a thousand bases per minute. Under optimum conditions, i.e., if there are no limitations due to limiting substrates or reagents, at each extension step, the amount of DNA target is doubled, leading to exponential (geometric) amplification of the specific DNA fragment.

Final elongation:

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

Final hold:

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

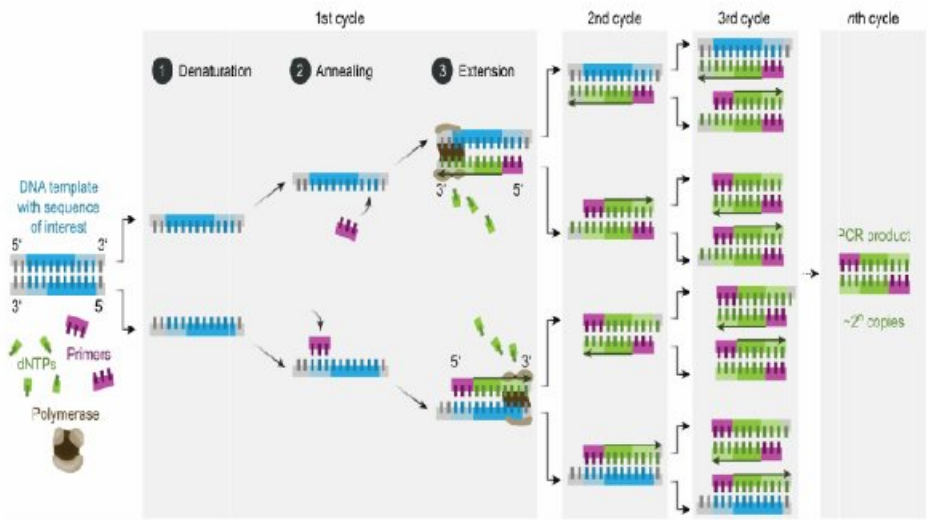
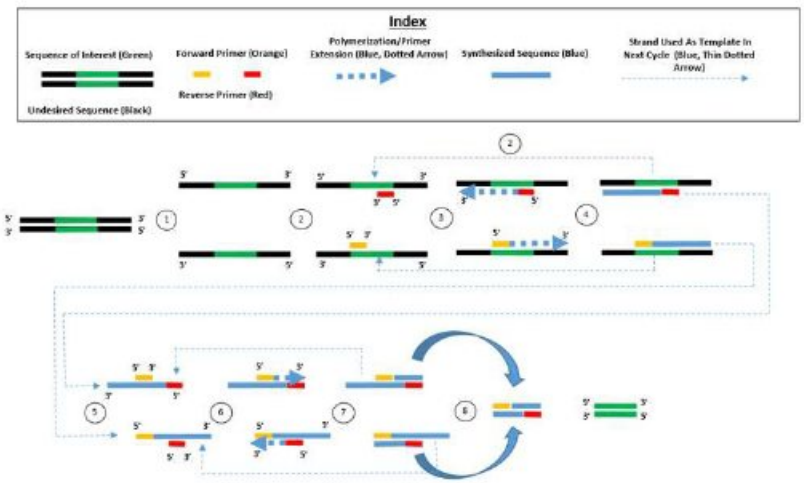


Fig. 7.2 Schematic drawing of a complete PCR cycle

To check whether the PCR successfully generated the anticipated DNA target region (also sometimes referred to as the amplimer or amplicon), agarose gel electrophoresis may be employed for size separation of the PCR products. The size of the PCR products is determined by comparison with a DNA ladder, a molecular weight marker which contains DNA fragments of known sizes, which runs on the gel alongside the PCR products.



1. The DNA double helix is melted apart at T> 90°C and its strands separate.
2. The temperature is decreased to slightly below the T_m of both the primers being used. Both primers bind to the available strands. These primers are supplied in excess to insure that the strands do not only come back and reanneal to one another.
3. Polymerization (extension) occurs via DNA Polymerase in the 5' to 3' direction on each strand.
4. Incorporated additional nucleotides give rise to new strands that extend past the sequence of interest.
5. The previously polymerized strands act as template for the other primer (if forward primer bound first, reverse primer now binds and vice versa).
6. Polymerization occurs via DNA Polymerase in the 5' to 3' direction on each strand, this time ending at the end of the sequence of interest.
7. Incorporated additional nucleotides give rise to new strands that only encode the sequence of interest.
8. The synthesized strands encoding the sequence of interest anneal to one another to form the end product.

Fig.7.3

7.2.4. PCR Steps :

- ✓ Initialization step: 94°C 10 min
- ✓ 35 Cycles:
- ✓ Denaturation step 94°C 30s
- ✓ Annealing step 55 °C 30s 1 cycle
- ✓ Extension step 72 °C 30s
- ✓ Final elongation : 72 °C 10 min

IMPORTANT REMINDERS:

- ✓ Always keep PCR reagents (master mix, primers, DNA template on ICE.
- ✓ Remember the centrifuge tubes are sterile so as soon as you are ready to start making the cell suspensions get your tubes and close them so that nothing contaminates your centrifuge tubes.
- ✓ Gloves are good to use to keep your tubes from getting contaminated.
- ✓ While making the cell suspension you should NOT flame the centrifuge tubes they might melt.
- ✓ Keep your tubes closed and make sure you do not cross contaminate with the tips of the micropipettes.
- ✓ Micropipettes come in different sizes and so do the pipette tips make sure you are using the correct pipette and tip for the amount you need to pick up.

Very Important

- ➢ When using the micropipettes be careful and avoid contaminating the micropipettes by slowly releasing the plunger.
- ➢ Dirty pipette tips should all be autoclaved so they are to be discarded in the tin cans with red bags.
- ➢ Please balance the micro centrifuge machine by placing another tube across from your tube both tubes should have approximately the same amount of liquid, remember you are balancing.

How do we see our results after running a PCR?

The most common way of seeing the results of a PCR is by running a gel electrophoresis. The PCR product is pipette into a special agar that will separate the DNA fragments according to their weight by using electricity.

Why "Polymerase"?

It is called "polymerase" because the only enzyme used in this reaction is DNA polymerase.

Calculating annealing temperature (Tm) of a primer:

$$Tm = 4\,(G + C) + 2\,(A + T)$$

Where G, C, A, T are the number of respective nucleotides. This formula can be used for primers that are no longer than 25 nt long. Nowadays, some specialized computer programs can do all calculations and determinations for proposed primers, e.g. Gene runner.

Questions for Practices

Q.1. What is PCR?

Q.2. What is the importance of PCR?

Q.3. what are the advantages of PCR?

7.3. PCR: Types and its Applications

PCR stands as a cornerstone in molecular biology, enabling scientists to magnify and analyze fragments of DNA with incredible precision. In this blog, we'll embark on understanding of PCR from its fundamental principles to the diverse techniques that have revolutionized research, diagnostics, and beyond.

Key Takeaways

1. PCR is a pivotal technique for DNA amplification, crucial in research, diagnostics, and forensics.
2. Variants like Hot Start PCR, RT-PCR, and qPCR enhance specificity, analyze RNA, and quantify nucleic acids.
3. Advances like High Fidelity Polymerase and Digital PCR offer precision in detecting mutations and low-abundance genes.

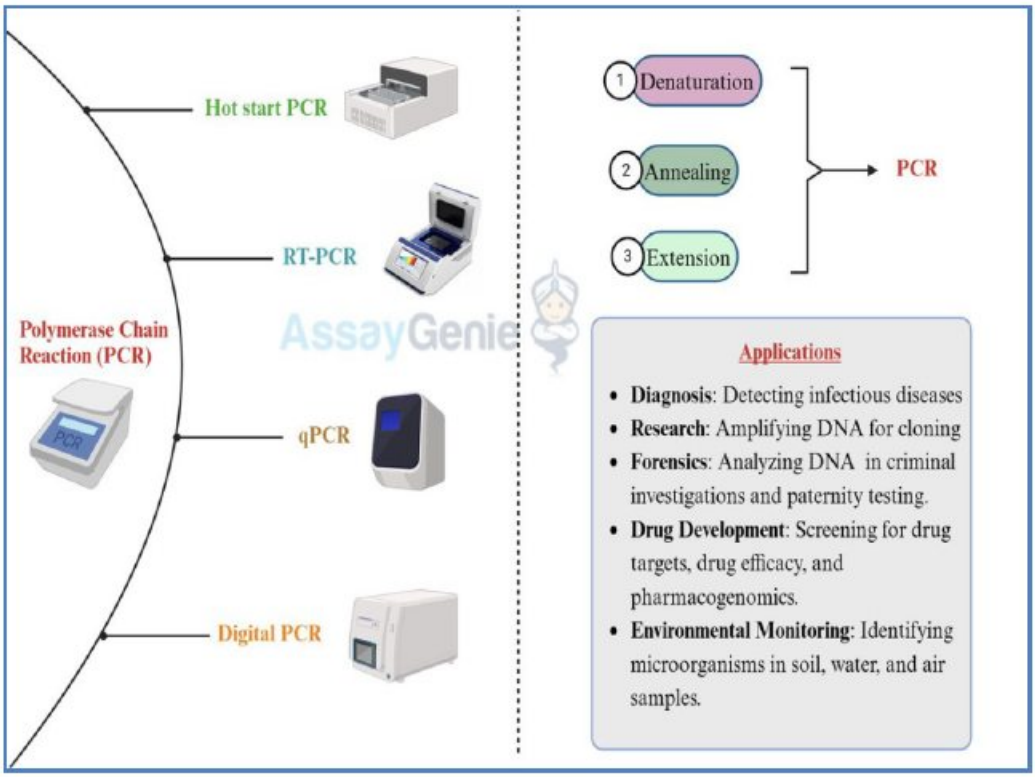


Fig.7.4: Types of PCR

7.3.1. Hot Start PCR

Hot-Start PCR is a modified version of the traditional polymerase chain reaction (PCR) technique that is designed to enhance the specificity and efficiency of DNA amplification. It addresses a common issue in standard PCR, where non-specific amplification can occur due to the premature activation of the DNA polymerase enzyme during the initial stages of the reaction setup. This can lead to the generation of undesired products and reduced overall specificity. In Hot-Start PCR, measures are taken to prevent the DNA polymerase from being active at low temperatures, such as during the initial setup and the initial heating steps of the PCR process. The goal is to prevent the amplification of non-specific products and enhance the amplification of the intended target sequence. This is particularly valuable when working with complex DNA templates or when dealing with samples that contain traces amounts of target DNA. The advantages of Hot-Start PCR include enhanced specificity, increased sensitivity and reduced optimization.

There are several methods used to implement the Hot-Start PCR technique:

Physical Separation:

One common approach involves physically separating the DNA polymerase from its DNA template and primers during the initial setup. This separation is achieved by using modified DNA polymerases or antibody-based techniques. For example, an antibody can be used to inhibit the

polymerase's activity until the reaction is heated, at which point the antibody denatures, allowing the polymerase to become active.

Chemical Modification:

Another method involves chemically modifying the DNA polymerase enzyme itself. Chemical modifications are added to the enzyme that renders it inactive at lower temperatures. As the reaction temperature increases during the initial heating steps, these modifications are removed, activating the enzyme for DNA amplification.

Hot-Start Taq Polymerase:

Some commercial DNA polymerases are designed to be inactive at room temperature or lower, but become active at elevated temperatures. These specialized enzymes, often referred to as "Hot-Start Taq Polymerases," possess antibody-mediated inhibition, reversible chemical modifications, or other mechanisms to prevent their activity until the reaction is properly heated.

High Fidelity Polymerase

Nucleotide matching errors can happen even though DNA polymerases amplify to the original template sequence quite precisely. In applications like cloning, mismatches can lead to shortened transcripts and proteins that have been incorrectly translated or are inactive. Polymerases with a "proofreading" activity have been found and incorporated into the workflow to prevent these mismatches. In *Pyrococcus furiosus*, the first proofreading polymerase, Pfu, was discovered in 1991. The 3' to 5' exonuclease activity of this Pfu enzyme. The exonuclease eliminates mismatched nucleotides at the 3' end of the strand as the DNA is amplified. The proper nucleotide is then substituted, and DNA synthesis keeps going.

The proper nucleoside triphosphate's binding affinity with the enzyme, where ineffective binding delays synthesis and permits the right replacement, is used to identify erroneous nucleotide sequences. Compared to Taq DNA polymerase, Pfu polymerase's proofreading activity leads to fewer errors in the final sequence. In order to further lower the mistake rate during DNA amplification, additional proofreading enzymes have been discovered recently, and the original Pfu enzyme has been modified.

7.3.2. RT-PCR

Reverse Transcription Polymerase Chain Reaction, commonly known as RT-PCR, is a powerful molecular biology technique that combines the principles of reverse transcription and polymerase chain reaction. This method is specifically designed to analyze and amplify RNA molecules, converting them into complementary DNA (cDNA) for further analysis.The process begins with the reverse transcription step, where an enzyme called reverse transcriptase synthesizes a complementary DNA strand (cDNA) using a single-stranded RNA molecule as a template. This step is

crucial because many biological processes involve RNA, such as gene expression and viral replication, and converting RNA to cDNA allows researchers to study these processes more easily.

Once the cDNA is generated, the polymerase chain reaction is employed to amplify specific target sequences. This involves repeated cycles of heating and cooling the reaction mixture. During the heating step (denaturation), the DNA strands are separated, creating single-stranded templates. In the cooling step (annealing), short DNA primers specifically designed to bind to the target cDNA sequences attach to the templates. Then, a heat-stable DNA polymerase enzyme extends the primers by adding complementary nucleotides, resulting in the synthesis of new DNA strands.

The end result of RT-PCR is an amplified amount of cDNA from the original RNA sample, allowing researchers to study gene expression levels, detect viral infections, analyze RNA sequences, and more. RT-PCR has proven invaluable in various fields, from medical diagnostics, where it's used to detect diseases like COVID-19, to molecular biology research, where it enables the exploration of gene function and regulation. Its ability to convert RNA into DNA and then amplify specific sequences has made RT-PCR an indispensable tool in modern molecular biology.

7.3.3. qPCR and RT-qPCR

For many applications, nucleic acids are detected, described, and quantified using quantitative PCR (qPCR). In RT-qPCR, RNA transcripts are frequently quantified by first reverse transcribed into cDNA, as previously mentioned, and then qPCR is performed. Denaturation, annealing, and elongation are three processes that are repeated to amplify DNA, much like in conventional PCR. But in qPCR, fluorescent labeling makes it possible to gather data as the PCR proceeds. Due to the variety of techniques and chemistries accessible, this technology has several advantages.

Fluorescent labeling, which uses a dsDNA binding dye, enables the measurement of the amplified DNA molecules in dye-based qPCR (usually green). The fluorescence is gauged over each cycle. The amount of DNA that has been duplicated causes the fluorescence signal to rise correspondingly. As a result, the DNA is measured "real-time". Only one target may be analyzed at a time with dye-based qPCR, and any ds-DNA found in the sample will cause the dye to bind. Many targets can be identified simultaneously in each sample using probe-based qPCR, but this needs the creation and development of a target-specific probe or probes, which are employed in addition to primers. There are many other sorts of probe designs, but the most popular kind is a hydrolysis probe that combines a fluorophore with a quencher. When the probe is still intact, fluorescence resonance

energy transfer (FRET) stops the fluorophore from emitting through the quencher. However, the probe is hydrolyzed during primer extension and amplification of the particular sequence it is linked to during the PCR reaction. An amplification-dependent rise in fluorescence is produced as a result of the probe's cleavage, which frees the fluorophore from the quencher.

As a result, the amount of the probe target sequence present in the sample directly correlates with the fluorescence signal from a probe-based qPCR process. In qPCR-based diagnostic tests, probe-based qPCR is frequently employed because it is more precise than dye-based qPCR.

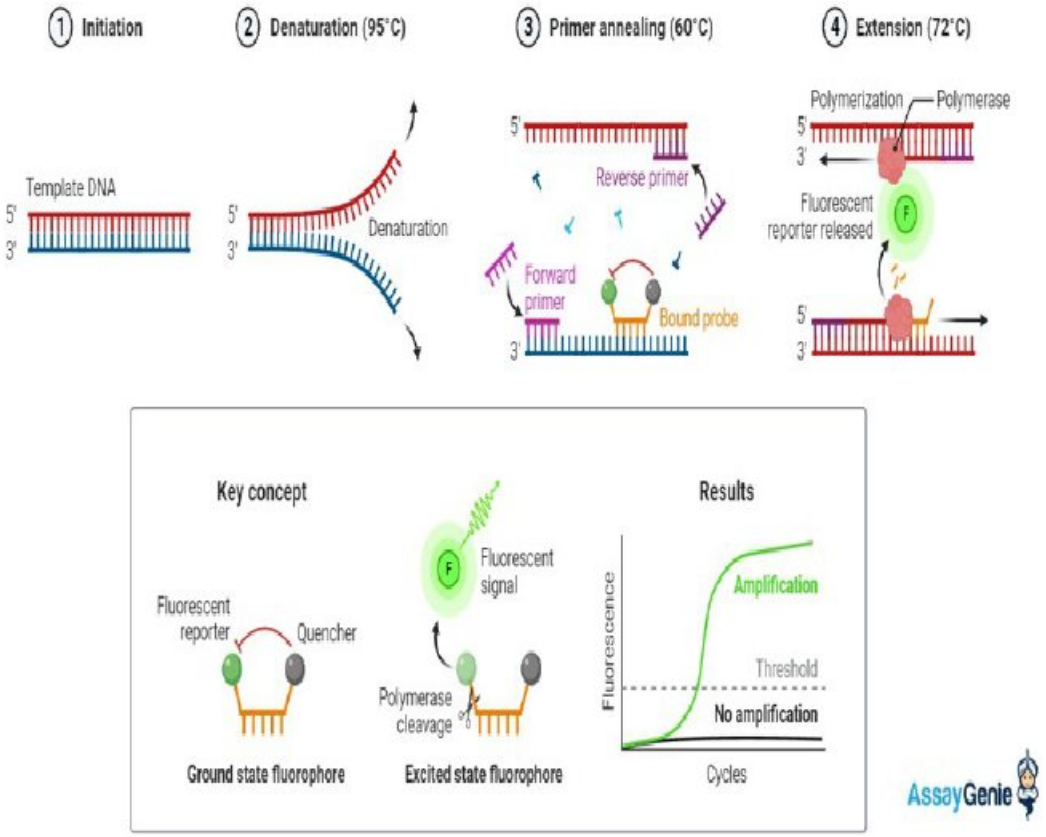


Fig7.5: Steps in qPCR

7.3.4 .Isothermal Amplification

For the PCR methods discussed above to precisely ramp up and down chamber temperatures for the denaturation, annealing, and extension phases, expensive thermo cycling equipment is needed.

Numerous methods have been devised that don't require such exact equipment and can be used inside the target cells or even in a straightforward water bath. These methods, which are referred to as isothermal amplification collectively, operate on the principles of exponential, linear, or cascade amplification.

Loop-mediated isothermal amplification, or LAMP, is the most well-known kind of isothermal amplification. LAMP amplifies template DNA or RNA using exponential amplification at 650C. Using DNA polymerase and four to six primers that are complementary to specific sections of the target DNA, LAMP creates new DNA. A "loop" structure might form in the freshly synthesized DNA as a result of two of these primers' complementary sequences recognizing and binding to sequences in the other primers. This structure then facilitates primer annealing in subsequent rounds of amplification. Numerous techniques, such as fluorescence, agarose gel electrophoresis, or colorimetry, can be used to see LAMP.

LAMP was a suitable alternative for SARS-CoV-2 testing in locations where clinical lab testing was not easily accessible, where sample storage and transport was not practical, or in labs that did not previously have PCR thermocycling equipment because it was simple to visualize and detect the presence or absence of product by colorimetric and did not require expensive equipment.

7.3.5 Digital PCR

Digital Polymerase Chain Reaction, or Digital PCR, is a cutting-edge molecular biology technique that takes the principles of traditional PCR to a new level of precision and sensitivity. Digital PCR is used to accurately quantify and analyze DNA or RNA molecules present in a sample, even when they are present in very low concentrations. In Digital PCR, the sample is partitioned into thousands of individual reactions, each containing a single molecule or a few molecules of the target DNA or RNA. This partitioning is achieved using microfluidic devices or specialized emulsion techniques. Each partitioned reaction acts as a miniature PCR reaction, with either the target DNA/RNA amplifying or remaining unamplified, depending on its presence or absence.

After amplification, the partitions are analyzed, and the number of positive partitions (those where amplification occurred) and negative partitions (those without amplification) is counted. This information is used to calculate the absolute quantity of the target molecules in the original sample. Digital PCR's ability to detect and quantify rare sequences with high precision makes it particularly

valuable for applications such as detecting genetic mutations, studying gene expression at low levels, and analyzing complex mixtures of DNA or RNA.

Compared to traditional quantitative PCR (qPCR), Digital PCR offers several advantages, including enhanced sensitivity, reduced susceptibility to PCR inhibitors, and improved accuracy for low-abundance targets. It's a powerful tool in various fields, including medical diagnostics, genetic testing, environmental monitoring, and more, where accurate quantification of nucleic acids is crucial. By providing a digital, absolute quantification of target molecules, Digital PCR contributes to more reliable and precise molecular analyses.

In conclusion, this exploration into Polymerase Chain Reaction (PCR) and its diverse types underscores the paramount significance of this revolutionary molecular biology technique. By adeptly amplifying and analyzing nucleic acids, PCR techniques enable us to unravel intricate genetic information and unravel previously inaccessible insights. Our investigation has delved into the intricacies of PCR variants, encompassing Reverse Transcription PCR (RT-PCR) and their multifaceted applications. The elucidation of these methodologies and their inherent capabilities serves to underscore the crucial role they play in modern molecular biology research and diagnostic

7.4 Applications of PCR:

Medicine: The PCR technique enables early diagnosis of malignant diseases.

- ✓ Classification of organisms
- ✓ Mutation detection
- ✓ Detection of pathogens
- ✓ Gene therapy
- ✓ Finger print
- ✓ Forensic science: PCR is very important for the identification of criminal
- ✓ Evolutionary studies: It plays an important role in phylogenetic analysis.

7.5. Summary

Under this unit we have discussed polymerase chain reaction (PCR), requirements of PCR technology, types and its applications etc. PCR ia a revolutionary method developed by Kary Mullis in the 1980s. PCR is based on using the ability of DNA polymerase to synthesize new strand of DNA complementary to the offered template strand. Because DNA polymerase can add a nucleotide only

onto a preexisting 3'-OH group, it needs a primer to which it can add the first nucleotide. This requirement makes it possible to delineate a specific region of template sequence that the researcher wants to amplify. At the end of the PCR reaction, the specific sequence will be accumulated in billions of copies (amplicons).

PCR amplification is an indispensable tool with various applications within medicine. Often, the process is used to test for the presence of specific alleles, such as prospective parents screening for genetic carriers. PCR amplification can also be used to diagnose the presence of disease directly and for mutations in the developing embryo. For example, the first time PCR was used in this way was for the diagnosis of sickle cell anemia through the detection of a single gene mutation.

Additionally, PCR has dramatically revolutionized the diagnostic potential for infectious diseases; it can rapidly determine the identity of microbes that were traditionally unable to be cultured or that required weeks for growth. Pathogens routinely detected using PCR include Mycobacterium tuberculosis, human immunodeficiency virus, herpes simplex virus, syphilis, etc. Moreover, qPCR is used to test the qualitative presence of microbes and quantify the bacterial, fungal, and viral loads.

7.6. Terminal Questions

Q. 1 What do you mean by PCR (Polymerase chain Reaction)? Explain it.

Answer:_____

_____

_____

Q. 2 Describe the different types of PCR (Polymerase Chain Reaction).

Answer:_____

_____

_____

Q. 3 Describe different applications of PCR.

Answer:_____

_____

_____

Q. 4 What are the qPCR and RT-qPCR? Explain it.

Answer:

Q. 5 Explain digital PCR with its applications.

Answer:

Q. 6 Write short note on the followings.

(i) RT-PCR (Real Time PCR)

(ii) Hot Start PCR

Answer:

Q. 7 Describe the requirements of PCR Technology.

Answer:

7.7. Further Readings

1. Biochemistry, Lehninger A.L.
2. Biochemistry, J.H.Weil.
3. Biochemistry fourth edition-David Hames and Nigel Hooper.
4. Textbook of Biochemistry for Undergraduates, Rafi, M.D.
5. Biochemistry and molecular biology- Wilson Walker.

Unit- 8: DNA Fingerprinting

Structure

8.1. Introduction

DNA fingerprinting, in genetics, method of isolating and identifying variable elements within the base-pair sequence of DNA (deoxyribonucleic acid). The technique was developed in 1984 by British geneticist Alec Jeffreys, after he noticed that certain sequences of highly variable DNA (known as minisatellites), which do not contribute to the functions of genes, are repeated within genes. Jeffreys recognized that each individual has a unique pattern of minisatellites (the only exceptions being multiple individuals from a single zygote, such as identical twins).

The procedure for creating a DNA fingerprint consists of first obtaining a sample of cells, such as skin, hair, or blood cells, which contain DNA. The DNA is extracted from the cells and purified. In Jeffreys's original approach, which was based on restriction fragment length polymorphism (RFLP) technology, the DNA was then cut at specific points along the strand with proteins known as restriction enzymes. The enzymes produced fragments of varying lengths that were sorted by placing them on a gel and then subjecting the gel to an electric current (electrophoresis): the shorter the fragment, the more quickly it moved toward the positive pole (anode). The sorted double-stranded DNA fragments were then subjected to a blotting technique in which they were split into single strands and transferred to a nylon sheet. The fragments underwent autoradiography in which they were exposed to DNA probes-pieces of synthetic DNA that were made radioactive and that bound to the minisatellites. A piece of X-ray film was then exposed to the fragments, and a dark mark was produced at any point where a radioactive probe had become attached. The resultant pattern of marks could then be analyzed.

Objectives

After reading this unit, the learner will able

- To know about characteristics of DNA, DNA markers& DNA fingerprinting
- To discuss principle and different methods in DNA Fingerprinting
- To discuss different steps involved in DNA Fingerprinting
- To know about different applications of DNA fingerprinting.

The assay developed by Jeffreys has been supplanted by approaches that are based on the use of the polymerase chain reaction (PCR) and so-called microsatellites (or short tandem repeats, STRs), which have shorter repeat units (typically 2 to 4 base pairs in length) than minisatellites (10 to more

than 100 base pairs in length). PCR amplifies the desired fragment of DNA (e.g., a specific STR) many times over, creating thousands of copies of the fragment. It is an automated procedure that requires only small amounts of DNA as starting material and works even with partially degraded DNA. Once an adequate amount of DNA has been produced with PCR, the exact sequence of nucleotide pairs in a segment of DNA can be determined by using one of several biomolecular sequencing methods. Automated equipment has greatly increased the speed of DNA sequencing and has made available many new practical applications, including pinpointing segments of genes that cause genetic diseases, mapping the human genome, engineering drought-resistant plants, and producing biological drugs from genetically altered bacteria.

An early use of DNA fingerprinting was in legal disputes, notably to help solve crimes and to determine paternity. Since its development, DNA fingerprinting has led to the conviction of numerous criminals and to the freeing from prison of many individuals who were wrongly convicted. However, making scientific identification coincide exactly with legal proof is often problematic. Even a single suggestion of the possibility of error is sometimes enough to persuade a jury not to convict a suspect. Sample contamination, faulty preparation procedures, and mistakes in interpretation of results are major sources of error. In addition, RFLP requires large amounts of high-quality DNA, which limits its application in forensics. Forensic DNA samples frequently are degraded or are collected postmortem, which means that they are lower-quality and subject to producing less-reliable results than samples that are obtained from a living individual. Some of the concerns with DNA fingerprinting, and specifically the use of RFLP, subsided with the development of PCR- and STR-based approaches.

8.2. DNA

DNA, organic chemical of complex molecular structure that is found in all prokaryotic and eukaryotic cells and in many viruses. DNA codes genetic information for the transmission of inherited traits. The chemical DNA was first discovered in 1869, but its role in genetic inheritance was not demonstrated until 1943. In 1953 James Watson and Francis Crick, aided by the work of biophysicists Rosalind Franklin and Maurice Wilkins, determined that the structure of DNA is a double-helix polymer, a spiral consisting of two DNA strands wound around each other. The breakthrough led to significant advances in scientists' understanding of DNA replication and hereditary control of cellular activities.

Each strand of a DNA molecule is composed of a long chain of monomer nucleotides. The nucleotides of DNA consist of a deoxyribose sugar molecule to which is attached a phosphate group

and one of four nitrogenous bases: two purines (adenine and guanine) and two pyrimidines (cytosine and thymine). The nucleotides are joined together by covalent bonds between the phosphate of one nucleotide and the sugar of the next, forming a phosphate-sugar backbone from which the nitrogenous bases protrude. One strand is held to another by hydrogen bonds between the bases; the sequencing of this bonding is specific—i.e., adenine bonds only with thymine, and cytosine only with guanine.The configuration of the DNA molecule is highly stable, allowing it to act as a template for the replication of new DNA molecules, as well as for the production (transcription) of the related RNA (ribonucleic acid) molecule. A segment of DNA that codes for the cell's synthesis of a specific protein is called a gene.

DNA replicates by separating into two single strands, each of which serves as a template for a new strand. The new strands are copied by the same principle of hydrogen-bond pairing between bases that exist in the double helix. Two new double-stranded molecules of DNA are produced, each containing one of the original strands and one new strand. This "semi conservative" replication is the key to the stable inheritance of genetic traits.

Within a cell, DNA is organized into dense protein-DNA complexes called chromosomes. In eukaryotes, the chromosomes are located in the nucleus, although DNA also is found in mitochondria and chloroplasts. In prokaryotes, which do not have a membrane-bound nucleus, the DNA is found as a single circular chromosome in the cytoplasm. Some prokaryotes, such as bacteria, and a few eukaryotes have extra chromosomal DNA known as plasmids, which are autonomous, self-replicating genetic material. Plasmids have been used extensively in recombinant DNA technology to study gene expression.

The genetic material of viruses may be single- or double-stranded DNA or RNA. Retroviruses carry their genetic material as single-stranded RNA and produce the enzyme reverse transcriptase, which can generate DNA from the RNA strand. Four-stranded DNA complexes known as G-quadruplexes have been observed in guanine-rich areas of the human genome.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."

8.3 Genetic marker

A genetic marker is a gene or DNA sequence with a known location on a chromosome that can be used to identify individuals or species. It can be described as a variation (which may arise due to mutation or alteration in the genomic loci) that can be observed. A genetic marker may be a short

DNA sequence, such as a sequence surrounding a single base-pair change (single nucleotide polymorphism, SNP), or a long one, like minisatellites.

For many years, gene mapping was limited to identifying organisms by traditional phenotypes markers. This included genes that encoded easily observable characteristics such as blood types or seed shapes. The insufficient number of these types of characteristics in several organisms limited the mapping efforts that could be done. This prompted the development of gene markers which could identify genetic characteristics that are not readily observable in organisms (such as protein variation).

8.3.1 Types

Some commonly used types of genetic markers are:

- RFLP (or Restriction fragment length polymorphism)
- SSLP (or Simple sequence length polymorphism)
- AFLP (or Amplified fragment length polymorphism)
- RAPD (or Random amplification of polymorphic DNA)
- VNTR (or Variable number tandem repeat)
- SSCP (or Single-strand conformation polymorphism)
- SSR Microsatellite polymorphism, (or Simple sequence repeat)
- SNP (or Single nucleotide polymorphism)
- STR (or Short tandem repeat)
- SFP (or Single feature polymorphism)
- DArT (or Diversity Arrays Technology)
- RAD markers (or Restriction site associated DNA markers)
- STS (using Sequence-tagged sites)

Molecular genetic markers can be divided into two classes: a) biochemical markers which detect variation at the gene product level such as changes in proteins and amino acids and b) molecular markers which detect variation at the DNA level such as nucleotide changes: deletion, duplication, inversion and/or insertion. Markers can exhibit two modes of inheritance, i.e. dominant/recessive or co-dominant. If the genetic pattern of homo-zygotes can be distinguished from that of hetero-zygotes, then a marker is said to be co-dominant. Generally co-dominant markers are more informative than the dominant markers.

8.3.2 Uses

Genetic markers can be used to study the relationship between an inherited disease and its genetic cause (for example, a particular mutation of a gene that results in a defective protein). It is known that pieces of DNA that lie near each other on a chromosome tend to be inherited together. This property enables the use of a marker, which can then be used to determine the precise inheritance pattern of the gene that has not yet been exactly localized.

Genetic markers are employed in genealogical DNA testing for genetic genealogy to determine genetic distance between individuals or populations. Uniparental markers (on mitochondrial or Y chromosomal DNA) are studied for assessing maternal or paternal lineages. Autosomal markers are used for all ancestry. Genetic markers have to be easily identifiable, associated with a specific locus, and highly polymorphic, because homozygotes do not provide any information. Detection of the marker can be direct by RNA sequencing, or indirect using allozymes.

Some of the methods used to study the genome or phylogenetics are RFLP, AFLP, RAPD, and SSR. They can be used to create genetic maps of whatever organism is being studied. There was a debate over what the transmissible agent of CTVT (canine transmissible venereal tumor) was. Many researchers hypothesized that virus like particles were responsible for transforming the cell, while others thought that the cell itself was able to infect other canines as an allograft. With the aid of genetic markers, researchers were able to provide conclusive evidence that the cancerous tumor cell evolved into a transmissible parasite. Furthermore, molecular genetic markers were used to resolve the issue of natural transmission, the breed of origin (phylogenetics), and the age of the canine tumor.

Genetic markers have also been used to measure the genomic response to selection in livestock. Natural and artificial selection leads to a change in the genetic makeup of the cell. The presence of different alleles due to a distorted segregation at the genetic markers is indicative of the difference between selected and non-selected livestock.

8.4 Minisatellites

In genetics, a minisatellite is a tract of repetitive DNA in which certain DNA motifs (ranging in length from 10–60 base pairs) are typically repeated two to several hundred times. Minisatellites occur at more than 1,000 locations in the human genome and they are notable for their high mutation rate and high diversity in the population. Minisatellites are prominent in the centromeres and telomeres of

chromosomes, the latter protecting the chromosomes from damage. The name "satellite" refers to the early observation that centrifugation of genomic DNA in a test tube separates a prominent layer of bulk DNA from accompanying "satellite" layers of repetitive DNA. Minisatellites are small sequences of DNA that do not encode proteins but appear throughout the genome hundreds of times, with many repeated copies lying next to each other. Minisatellites and their shorter cousins, the microsatellites, together are classified as VNTR (variable number of tandem repeats) DNA. Confusingly, minisatellites are often referred to as VNTRs, and microsatellites are often referred to as short tandem repeats (STRs) or simple sequence repeats (SSRs).

Structure

Minisatellites consist of repetitive, generally GC-rich, motifs that range in length from 10 to over 100 base pairs. These variant repeats are tenderly intermingled. Some minisatellites contain a central sequence (or "core unit") of nucleobases "GGGCAGGANG" (where N can be any base) or more generally consist of sequence motifs of purines (adenine (A) and guanine (G)) and pyrimidines (cytosine (C) and thymine (T)). Hypervariable minisatellites have core unit's 9–64 bp long and are found mainly at the centromeric regions. In humans, 90% of minisatellites are found at the sub-telomeric region of chromosomes. The human telomere sequence itself is a tandem repeat: TTAGGG TTAGGGTTAGGG.

Function

Minisatellites have been implicated [citation needed] as regulators of gene expression (e.g. at levels of transcription, alternative splicing, or imprint control). They are generally non-coding DNA but sometimes are part of possible genes. Minisatellites also constitute the chromosomal telomeres, which protect the ends of a chromosome from deterioration or from fusion with neighboring chromosomes.

8.5. Microsatellites

A microsatellite is a tract of repetitive DNA in which certain DNA motifs (ranging in length from one to six or more base pairs) are repeated, typically 5–50 times.Microsatellites occur at thousands of locations within an organism's genome. They have a higher mutation rate than other areas of DNA leading to high genetic diversity. Microsatellites are often referred to as short tandem repeats (STRs) by forensic geneticists and in genetic genealogy, or as simple sequence repeats (SSRs) by plant geneticists. Microsatellites and their longer cousins, the minisatellites, together are classified as VNTR

(variable number of tandem repeats) DNA. The name "satellite" DNA refers to the early observation that centrifugation of genomic DNA in a test tube separates a prominent layer of bulk DNA from accompanying "satellite" layers of repetitive DNA.

They are widely used for DNA profiling in cancer diagnosis, in kinship analysis (especially paternity testing) and in forensic identification. They are also used in genetic linkage analysis to locate a gene or a mutation responsible for a given trait or disease. Microsatellites are also used in population genetics to measure levels of relatedness between subspecies, groups and individuals.

History

Although the first microsatellite was characterized in 1984 at the University of Leicester by Weller, Jeffreys and colleagues as a polymorphic GGAT repeat in the human myoglobin gene, the term "microsatellite" was introduced later, in 1989, by Litt and Luty. The name "satellite" DNA refers to the early observation that centrifugation of genomic DNA in a test tube separates a prominent layer of bulk DNA from accompanying "satellite" layers of repetitive DNA. The increasing availability of DNA amplification by PCR at the beginning of the 1990s triggered a large number of studies using the amplification of microsatellites as genetic markers for forensic medicine, for paternity testing, and for positional cloning to find the gene underlying a trait or disease. Prominent early applications include the identifications by microsatellite genotyping of the eight-year-old skeletal remains of a British murder victim (Hagelberg et al. 1991), and of the Auschwitz concentration camp doctor Josef Mengele who escaped to South America following World War II (Jeffreys et al. 1992).

Structures, locations, and functions

A microsatellite is a tract of tenderly repeated (i.e. adjacent) DNA motifs that range in length from one to six or up to ten nucleotides (the exact definition and delineation to the longer minisatellites varies from author to author), and are typically repeated 5–50 times. For example, the sequence TATATATATA is a dinucleotide microsatellite, and GTCGTCGTCGTCGTC is a trinucleotide microsatellite (with A being Adenine, G Guanine, C Cytosine, and T Thymine). Repeat units of four and five nucleotides are referred to as tetra- and pentanucleotide motifs, respectively. Most eukaryotes have microsatellites, with the notable exception of some yeast species. Microsatellites are distributed throughout the genome. The human genome for example contains 50,000–100,000 dinucleotide microsatellites, and lesser numbers of tri-, tetra- and pentanucleotide microsatellites.Many are located in non-coding parts of the human genome and therefore do not produce proteins, but they can also be located in regulatory regions and coding regions.

Microsatellites in non-coding regions may not have any specific function, and therefore might not be selected against; this allows them to accumulate mutations unhindered over the generations and gives rise to variability that can be used for DNA fingerprinting and identification purposes. Other microsatellites are located in regulatory flanking or intronic regions of genes or directly in codons of genes – microsatellite mutations in such cases can lead to phenotypic changes and diseases, notably in triplet expansion diseases such as fragile X syndrome and Huntington's disease.

Telomeres are linear sequences of DNA that sit at the very ends of chromosomes and protect the integrity of genomic material (not unlike an aglet on the end of a shoelace) during successive rounds of cell division due to the "end replication problem". In white blood cells, the gradual shortening of telomeric DNA has been shown to inversely correlate with ageing in several sample types. Telomeres consist of repetitive DNA, with the hexanucleotide repeat motif TTAGGG in vertebrates. They are thus classified as minisatellites. Similarly, insects have shorter repeat motifs in their telomeres that could arguably be considered microsatellites.

Applications

Microsatellites are used for assessing chromosomal DNA deletions in cancer diagnosis. Microsatellites are widely used for DNA profiling, also known as "genetic fingerprinting", of crime stains (in forensics) and of tissues (in transplant patients). They are also widely used in kinship analysis (most commonly in paternity testing). Also, microsatellites are used for mapping locations within the genome, specifically in genetic linkage analysis to locate a gene or a mutation responsible for a given trait or disease. As a special case of mapping, they can be used for studies of gene duplication or deletion. Researchers use microsatellites in population genetics and in species conservation projects. Plant geneticists have proposed the use of microsatellites for marker assisted selection of desirable traits in plant breeding.

8.6. DNA Fingerprinting

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.

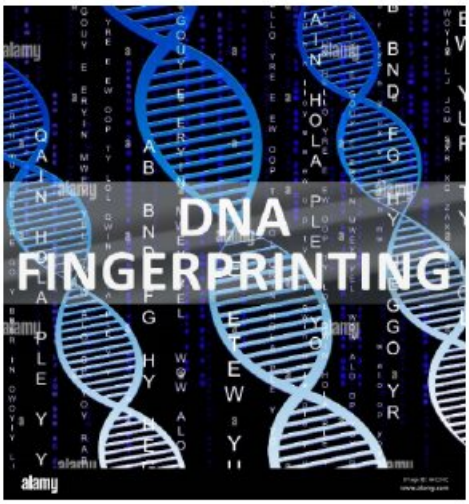


Fig.8.1 DNA fingerprinting

8.6.1 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.

8.6.2 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.

8.7. 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.

You can utilize DNA extraction strategies enrolled beneath,

1. Phenol-chloroform DNA extraction strategy
2. CTAB DNA extraction strategy
3. Proteinase K DNA extraction strategy

In any case, we emphatically prescribe utilizing a ready to go DNA extraction unit for DNA fingerprinting. The immaculateness and amount of DNA ought to be ~1.80 and 100ng, individually to play out the DNA test. Filter the DNA utilizing the DNA sanitization unit, if necessary. From that point onward, measure the DNA utilizing the UV-Visible spectrophotometer. Furthermore, perform one of the accompanying strategies recorded underneath.

Questions for Practice:

Q.1. What is the innovative advancement of DNA profiling and forensics?

Q.2. How is DNA fingerprinting done in criminal cases?

Q.3. Who discovered DNA fingerprinting technique?

Q.4. State two applications of DNA fingerprinting.

8.8. DNA Fingerprinting Strategies

Stage 3: Restriction Absorption, Enhancement or DNA Sequencing

Three regular strategies are utilized:

1. RFLP based STR investigation
2. PCR based investigation
3. Real-time PCR investigation

Stage 4: Analysis of Results

As we examined, 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.

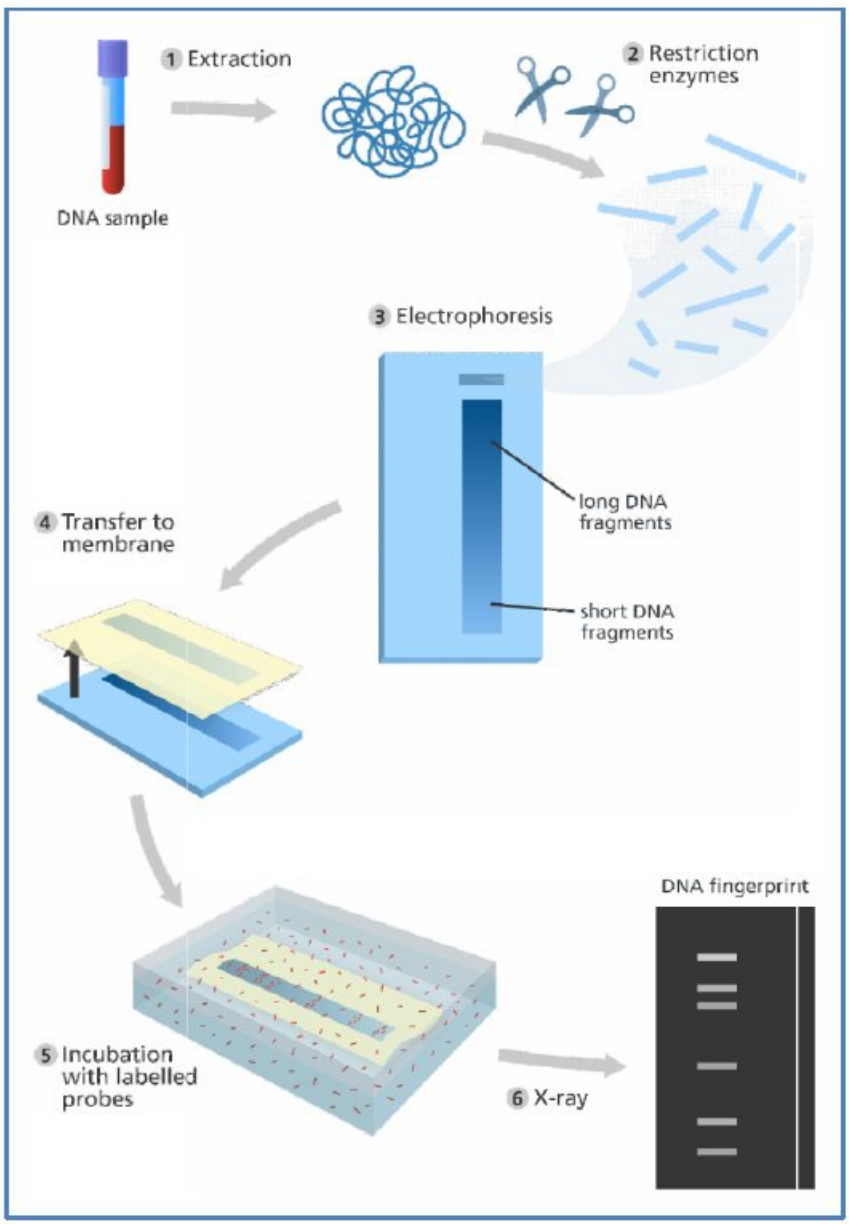


Fig. 8.2

- The DNA of every human being on the planet is 99.9% same. However, about 0.1% or 3 x 10^6base pairs (out of 3 x 10^9 bp) of DNA is unique in every individual.
- Human genome possesses numerous small non-coding but inheritable sequences of bases which are repeated many times. They do not code for proteins but make-up 95% of our genetic DNA and therefore called the, junk DNA.
- They can be separated as satellite from the bulk DNA during density gradient centrifugation and hence called satellite DNA.
- In satellite DNA, repetition of bases is in tandem. Depending upon length, base composition and numbers of tandemly repetitive units, satellite DNAs have subcategories like microsatellites and mini-satellites.
- Satellite DNAs show polymorphism. The term polymorphism is used when a variant at a locus is present with a frequency of more than 0.01 populations.
- Variations occur due to mutations. These mutations in the non-coding sequences have piled up with time and form the basis of DNA polymorphism (variation at genetic level arises due to mutations).
- The junk DNA regions are thus made-up of length polymorphisms, which show variations in the physical length of the DNA molecule.
- At specific loci on the chromosome the number of tandem repeats varies between individuals. There will be a certain number of repeats for any specific loci on the chromosome.
- Depending on the size of the repeat, the repeat regions are classified into two groups. Short tandem repeats (STRs) contain 2-5 base pair repeats and variable number of tandem repeats (VNTRs) have repeats of 9-80 base pairs.
- Since a child receive 50% of the DNA from its father and the other 50% from his mother, so the number VNTRs at a particular area of the DNA of the child will be different may be due to insertion, deletion or mutation in the base pairs.
- As a result, every individual has a distinct composition of VNTRs and this is the main principle of DNA fingerprinting.
- As single change in nucleotide may make a few more cleavage site of a given nucleotide or might abolish some existing cleavage site.
- Thus, if DNA of any individual is digested with a restriction enzyme, fragments pattern (sizes) will be produced and will be different in cleavage site position. This is the basics of DNA fingerprinting.

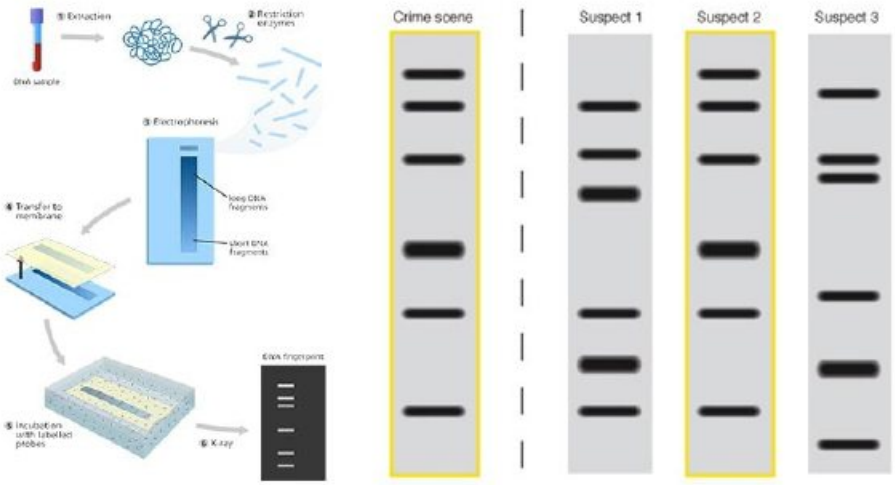


Fig. 8.3

8.9. Methods of DNA Fingerprinting

Restriction fragment length polymorphism (RFLP) and polymerase chain reaction (PCR) amplification of short tandem repeats (STRs) are two main DNA tests widely used for DNA fingerprinting.

A. Restriction fragment length polymorphism (RFLP)

1. The first step in this process is to isolate the DNA from the sample material to be tested. The sample size for RFLP test must be large enough to get the proper result.
2. Once the required size of the sample is available, the DNA is isolated from the sample and is subjected to restriction digestion using restriction enzymes.
3. The digested DNA sample is then separated by agarose gel electrophoresis, in which the DNA is separated based on the size.
4. The next step is transfer of separated DNA from gel slab onto the nitrocellulose membrane to hybridize with a labeled probe that is specific for one VNTR region (radio activity labeled complimentary sequence for VNTR region nucleotide sequence).
5. This technique of transferring and hybridizing DNA onto nitrocellulose membrane is known as southern blotting, a most widely used DNA detection technique by molecular biologists.
6. After the hybridization with the radioactive probes, the X- ray film is developed form the southern blotting and only the areas where the radioactive probe binds will show up on the film.

7. Now these bands when compared with the other known samples, will give the final result of the DNA fingerprinting.

Advantages

The RFLP is considered to be more accurate than the PCR, mainly because the size of the sample used more, use of a fresh DNA sample, and no amplification contamination.

Limitation

The RFLP, however, require longer time period to complete the analysis and is costly.

B. Polymerase Chain Reaction (PCR) amplification of short tandem repeats (STRs)

1. Thousands of copies of a particular variable region are amplified by **PCR** which forms the basis of this detection.
2. STR with a known repeat sequence is amplified and separated using gel-electrophoresis.
3. The distance migrated by the STR is examined.
4. For the amplification of STRs using PCR, a short synthetic DNA, called primers are specially designed to attach to a highly conserved common nonvariable region of DNA that flanks the variable region of the DNA.
5. By comparing the STR sequence size amplified by PCR with the other known samples, will give the final result of the DNA fingerprinting.

Advantages

- Small amount of specimen is sufficient for the test.
- Takes a shorter time to complete.
- Less costly.

Limitation

- Less accurate than RFLP.
- Possibility of amplification contamination.

8.10. 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.

- DNA Fingerprinting is used by scientists to distinguish between individuals of the same species using only samples of their DNA. It is a primary method for identifying an individual.

1. Forensic Science:

Biological materials used for DNA profiling are: Blood, Hair, Saliva, Semen, Body tissue cells etc. DNA isolated from the evidence sample can be compared through VNTR (Variable number of tandem repeats) prototype. It is useful in solving crimes like murder and rape.

2. Paternity and Maternity Determination:

A Person accedes to his or her VNTRs from his or her parents. Parent-child VNTR prototype analysis has been used to solve disputed cases. This information can also be used in inheritance cases, immigration cases.

3. Personal Identification:

It utilizes the concept of using DNA fingerprints as a sort of genetic bar code to pinpoint individuals.

4. Diagnosis of Inherited Disorders:

It is also useful in diagnosing inherited disorders in both prenatal and newborn babies. These disorders may include cystic fibrosis, haemophilia, Huntington's disease, familial Alzheimer's, sickle cell anemia, thalassemia, and many others.

5. Development of Cures for Inherited Disorders:

By studying the DNA fingerprints of relatives who have a history of some particular disorder, DNA prototypes associated with the disease can be ascertained.

6. Detection of AIDS:

By comparing the band of HIV "RNA" (converted to DNA using RTPCR) with the bands form by the man's blood, person suffering with AIDS can be identified.

7. Breeding Program:

Breeders conventionally use the phenotype to evaluate the genotype of a plant or an animal. As it is difficult to make out homozygous or heterozygous dominance from appearance, the DNA fingerprinting allows a fastidious and precise determination of genotype. It is basically useful in breeding race horses and hunting dogs.

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 bloods bunches with numerous alleles, and these alleles display a scope of predominance designs.

8.11. 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.

8.12. Summary

Under this unit we have summarized the concept of DNA, DNA fingerprinting, different steps involved in it, principle and methods with different applications etc.DNA fingerprinting is a laboratory technique used to determine the probable identity of a person based on the nucleotide sequences of certain regions of human DNA that are unique to individuals. DNA fingerprinting is used in a variety of situations, such as criminal investigations, other forensic purposes and paternity testing. In these

situations, one aims to "match" two DNA fingerprints with one another, such as a DNA sample from a known person and one from an unknown person.

DNA fingerprinting, also known as DNA profiling or genetic fingerprinting, is a sophisticated genetic identification technique that analyzes specific regions of an individual's DNA, known as short tandem repeats (STRs), to create a unique genetic profile. This profile can be used for various purposes, including molecular diagnostics, forensics investigations, paternity testing, and genealogical research, due to its ability to distinguish individuals based on their distinct DNA patterns." The precision and reliability of DNA fingerprinting have revolutionized genetics, molecular diagnostics and forensic science, providing an indispensable tool for identifying individuals and establishing relationships based on their genetic makeup.

8.13. Terminal questions

Q. 1 What do you mean by**unique features of DNA?** Explain it.

Answer:_____

_____

_____

Q. 2 Describe principle of DNA fingerprinting.

Answer:_____

_____

_____

Q. 3 Describe different methods of DNA fingerprinting.

Answer:_____

_____

_____

Q. 4 Write short notes on the following.

(a) DNA profiling

(b) DNA

Answer:_____

_____

_____

Q. 5 Write short notes on applications of DNA fingerprinting.

Answer:——

——

——————————————————————————————

8.14. Further readings

1. Biochemistry- Lehninger A.L.
2. Text book of Botany – Singh -Pande-Jain.
3. The elements of Botany- James Hewetson Wilson
4. Textbook of Biotechnology –H. K. Das
5. Biochemistry and molecular biology- Wilson Walker

Unit-9: Screening and Analysis of Recombinants

Structure

9.1. Introduction

Recombinant DNA technology describes a process of genetic engineering that uses enzymes and various laboratory techniques to isolate and manipulate genetic material. There are different ways to carry out manipulation in an organism's genome, all with the aim to improve certain characteristics, for instance to achieve greater stability of human proteins, but also to reduce toxicity on normal cells and the associated side effects. Recombinant monoclonal antibodies, for instance, are produced in vitro (respectively ex vivo), by cloning synthetic genes that have been recovered from monoclonal antibodies, generally from mammalian cells. They are then modified by means of genetic engineering to improve their specificity and reproducibility, as well as antigen binding properties.

Recombinant describes DNA, proteins, cells, or organisms that are created by combining genetic material from two different sources in order to alter their characterization. They are then tested

in clinical trials and used in different healthcare and biotechnology products, such as mrna vaccines. The synthetic genes used for recombinant antibodies are typically generated from a plasmid or from an integrated sequence in a stable cell line. Expression takes place in host cells, with the aim to alter the material and achieve translational modifications.

Recombination is used in the biotechnology sector to overcome weaknesses of polyclonal and traditional monoclonal antibodies in terms of quality and reproducibility. It serves several functions in organisms, including DNA repair in bacteria and eukaryotes, and ensures the correct alignment and segregation of chromosomes during meiosis in eukaryotes. Technologies to alter genetic material have been successfully applied to generate important proteins used in diagnostics, clinical trials and general health care.

Recombinant proteins are useful tools in understanding protein-protein interactions. They allow scientists to study protein interactions with other proteins or peptides, enzymes, small molecules, lipids, and nucleic acids. The choice of the host cell whose protein synthesis machinery will produce the protein will initiate the outline of the whole process and also impact the growth factor. Recombinant DNA molecules, generated for the first time half a century ago, are increasingly used in pharmaceutical and medical fields thanks to a number of benefits, such as their reliable expression and distribution as DNA sequences and plasmids.

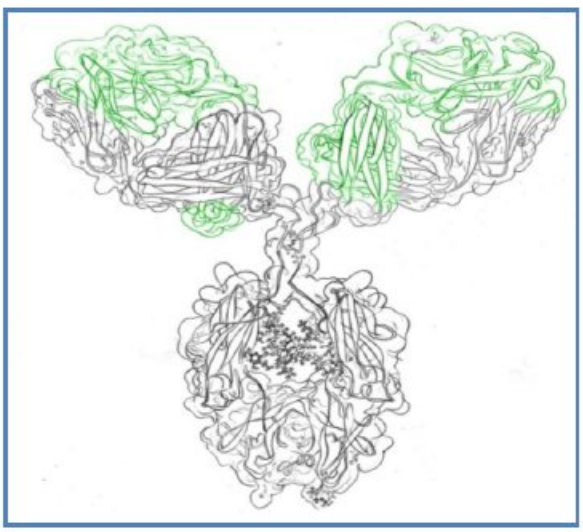


Fig.9. 1: Recombinant DNA

Objectives

After reading this unit, the learner will able

- To know about Chromogenic substrates & its uses

- To know the genetic selection, nucleic probe and hybridization probe
- To discuss chromosome jumping, chromosome walking and clone library screening
- To discuss genetic disorders and different blotting techniques

9.2. Chromogenic Substrates & its Uses

Chromogenic substrates are colorless soluble molecules consisting of chromospheres, a chemical group that, after enzymatic cleavage, releases colour and a specific enzymatic substrate. They are synthetically produced and are designed to possess selectivity similar to the natural substrate for the enzyme. These compounds are useful for enzymatic detection, as Chromogenic substrates specifically bind with the target enzyme. This reaction allows the enzyme to catalyze the separation of the chromospheres group, resulting in an insoluble product with a distinctive colour releasing the chromospheres that confirms the existence and activity of the enzyme under study. This colour change can be followed spectrophotometrically and is proportional to the proteolysis activity of the enzyme.

Chromogenic substrates facilitate the quantitative and qualitative identification of enzymes and proteins in laboratory experiments, thanks to their visible colour change. This chromatic transition, whose intensity can be quantified, allows the precise measurement of the enzymatic or protein target in Chromogenic assays. Its use is common in techniques such as Western blot, ELISA, immune histochemistry, enzymatic assays, and microbial detection in culture media. Chromogenic substrates aid in the qualitative and quantitative detection of enzymes or other proteins in biochemical experiments by producing a visible color (*chromo*) change. The basic mechanism of Chromogenic substrate-mediated detection is that the substrate produces a colored product that can be visually detected, when acted upon by a corresponding enzyme.

This enzyme may be the ultimate target of detection like in enzyme assays. Alternatively, the enzyme could be attached, or *conjugated*, to another molecule – mostly, an antibody – that detects and binds to the primary target protein in the sample. The color change produced by cleavage of a Chromogenic substrate can be visually deciphered through qualitative detection of the target enzyme or protein, the intensity of the color change can also be measured. This is how quantitative detection of the target protein or enzyme is done using a Chromogenic assay.

Uses

Common experimental techniques where Chromogenic substrates are used are Western blot, ELISA, immunohistochemistry (IHC), enzyme assays and microbial detection procedures. Chromogenic substrates offer a number of advantages over traditional media based on pH indicators.

A color change is seen in a colony after transportation and hydrolysis, and this reduces the potential number of false positives due to mutation by more than 10-fold, as fewer genes are involved in generating a signal.

9.3 Genetic selection

Genetic selection is the process by which certain traits become more prevalent in a species than other traits. It describes the forces that determine the traits seen in a population or species. Natural selection involves forces such as predation or sexual preference or even something seemingly as mundane as tail size. These traits are expressed through different versions of our genes or alleles. Artificial selection on the other hand, involves human interference in natural selection. Genetic testing increasingly informs decisions about whether to continue a pregnancy (prenatal genetic testing) or which IVF embryo to implant (pre-implantation genetic diagnosis). Genetic selection techniques allow parents at risk of transmitting serious genetic conditions to have unaffected and genetically related children. But what counts as serious is controversial, and our perceptions of which conditions should be prevented are socially constructed and change over time.

Genetic selection is a powerful strategy for generating pure or nearly pure populations of differentiated cell types. In order to develop this strategy, a gene that is expressed specifically in the desired cell type must be identified. This gene can be specific to a terminally differentiated cell type or to a lineage-restricted stem cell. The promoter of the specific gene is then joined to a gene that encodes resistance to an antibiotic drug such as neomycin (G 418), hygromycin, bleomycin, or zeocin. This construct is transfected into stem cells that are then caused to differentiate. If the differentiating cells are treated with the antibiotic drug, all cells that cannot express the fusion gene construct will die as in given figure. An example of genetic selection is the production of a highly pure (>99%) population of cardiomyocytes from differentiated mouse ES cells using the α-cardiac myosin heavy chain promoter joined to a cDNA coding for neomycin resistance.

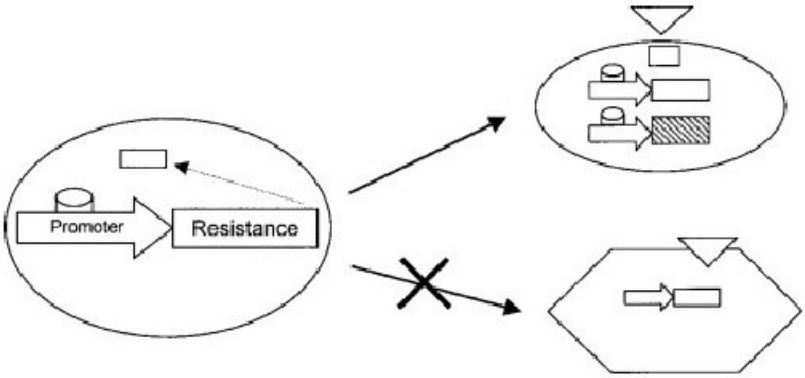


Fig.9.2: Genetic Selection (Genetic selection used to obtain a specific population of differentiated cells)

Disability rights advocates point out that when prenatal screening identifies a fetus with Down syndrome, this often results in terminating the pregnancy, even though people with Down syndrome often thrive and live happily. Driven by genetic testing companies, prenatal genetic screening has expanded rapidly to include a wide range of conditions. There are also new companies that rank IVF embryos according to their future risk of conditions like cancer, heart disease, or schizophrenia – a scientifically and ethically questionable practice.

Genetic screening is the process of using scientific techniques to test an embryo for the presence of a negative genetic code. For example, scientists can screen for genetic codes related to Tay-Sachs disease or other diseases. Genetic screening is also referred to as genetic selection, but there are two main types of genetic selection. The first is artificial selection, or genetic screening, discussed above. In this process, alleles are the most significant element of human genetics. Alleles are different versions of the same genes. Each individual has two alleles of each gene. Genetic screening can select which allele will be expressed. If both alleles for a gene are the same, then an individual is homozygous. If they are different, the person is heterozygous.

If a person is homozygous, then that allele will be the one expressed in the person's phenotype. In a heterozygous individual, the dominant allele of the gene will override the recessive allele. If two dominant alleles are present, then they are referred to as codominant alleles. The various genes in an individual's DNA are called their genotype. The physical manifestations of genes are referred to as a person's phenotype. Codominant alleles are both expressed in the phenotype. For example, people with both A and B blood alleles have an AB blood type.

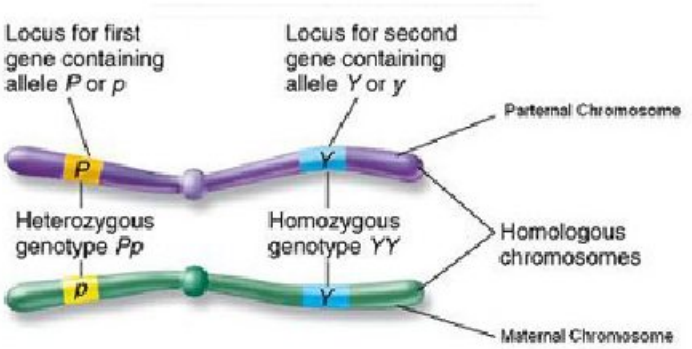


Fig.9.3: Genetic Selection Types, Impacts

9.4 Nucleic Acid Probe

Nucleic acid probe technology is increasingly being used in basic research in veterinary microbiology and in diagnosis of infectious diseases of veterinary importance. This review presents an overview of nucleic acid probe methodology and its applications in veterinary infectious diseases. The major applications of nucleic acid probes include detection of pathogens in clinical samples, especially those organisms which are fastidious and difficult to cultivate, differentiation of virulent from a virulent organisms and vaccine strains from wild type isolates, typing of microorganisms mapping genes, screening libraries of cloned DNA for specific genes, detection of latently infected or carrier animals, study of mechanisms of pathogenesis, epidemiological studies and food safety.

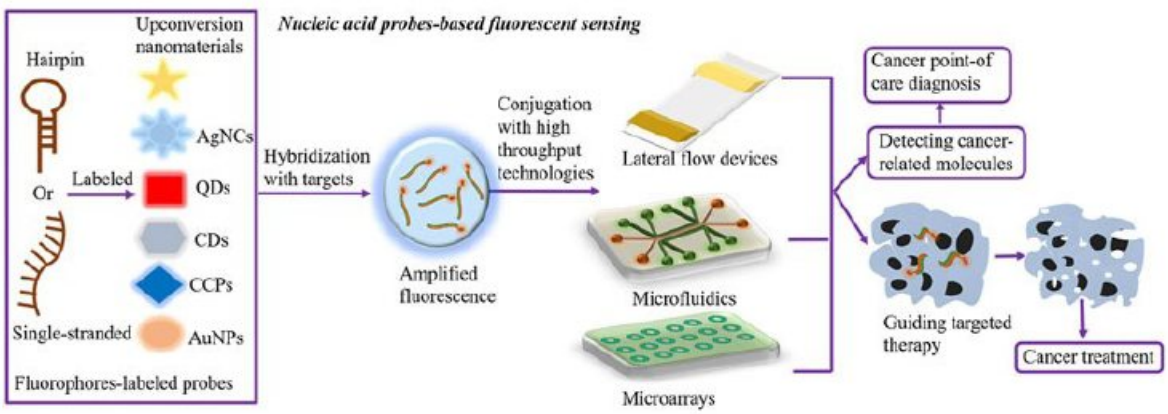


Fig.9.4: The Application of Nucleic Acid Probe–Based

A variety of nucleic acid probes are available commercially as kits for detection of microorganisms that infect humans. Some of these can also be used to detect the same pathogens in specimens from dogs or cats. Probes are available for use in the clinical microbiology laboratory to

identify organisms that have grown in culture (e.g., Gen-Probe Inc., San Diego), and these probes are widely used in human clinical microbiology laboratories.

Probe hybridization can be performed on a nitrocellulose membrane (solid-phase hybridization); on formalin-fixed, paraffin-embedded sections mounted on a microscope slide (in situ hybridization); or in solution (liquid-phase hybridization). For solid-phase hybridization, the probe is reacted with microorganism DNA that has been immobilized on the membrane. Unbound probe is washed away, and the bound probe is detected using fluorescence, chemiluminescence, radioactivity, or color development (in the same way that bound antibody or antigen is detected in an ELISA or immunofluorescent antibody assay). For in situ hybridization, formalin-fixed, paraffin-embedded specimens are sectioned and mounted on a special slide. The sections are deparaffinized, dried, and incubated with a solution that contains the probe, so both the presence and the location of the target pathogen within tissues can be identified. In situ hybridization assays that include a fluorescent-labeled probe are referred to as fluorescence in situ hybridization (FISH) assays.

Because liquid-phase hybridization occurs in solution, unbound probe cannot be washed away. To overcome this problem, a chemiluminescent acridinium ester label is attached to the probe. A subsequent chemical hydrolysis step selectively degrades only unbound probe. On addition of peroxides, the intact (hybridized) probe then emits light.

Although not yet widely used for veterinary applications, peptide nucleic acid (PNA) probes are now increasingly available to detect target DNA. PNA probes are uncharged peptides that mimic DNA and bind to complementary DNA sequences just as a nucleic acid probe would.12, 13 PNA probes lack the net negative charge of nucleic acid probes; therefore, the electrostatic repulsion that normally occurs when two negatively charged DNA strands hybridize does not occur. The result is a more stable and specific binding of the probe to its target, which in turn can be associated with increased assay sensitivity and specificity. Branched DNA assays and hybrid capture assays are highly sensitive hybridization methods that include steps to intensify the signal generated from probe hybridization.3,4 They are not yet widely used in veterinary medicine.

9.5 Chromosome Walking

Chromosome walking is a technique, which is widely used for the discovery for new genes and sequencing of chromosome. For this the chromosome is isolated using pulsed field gel electrophoresis and digested partially to result in overlapping fragments. The fragments are ligated in plasmid and transformed in cells. The transformed cells are allowed to form colonies. In chromosome walking a

colony is picked up and the insert is used as a probe for screening colonies having overlapping fragments. When the process is repeated multiple times, a complete sequence of chromosome can be obtained.

When a probe is used for the identification of a gene sequence in a genomic library, the probe may hybridize with a number of clones, each carrying a part of a large gene fragmented during preparation of genomic library. If we obtain partial digests (by digesting the DNA only partially) from the genome, different genomes (from large number of cells) may give fragments which have overlapping sequences, because sites cleaved in different genomes of the same organism, will differ being random. Since none of these fragments may have its entire sequence represented in the probe, overlapping sequences may be used to construct the original genomic sequence. Identification of fragments with an overlapping sequence may be a key to the reconstruction or characterization of large chromosome regions. This is achieved by the technique popularly called chromosome walking.

Chromosome walking is a method of positional cloning used to find, isolate, and clone a particular allele in a gene library. Chromosome Walking was developed by Welcome Bender, Pierre Spierer, and David S. Hogness in the early 1980's. There are nearly half a dozen positional cloning tests that are done prior to a chromosome walk. Each clone in the cosmic library has a DNA insert of 50 KB. The walking starts at the closest gene that has already been identified, known as a marker gene.

Once the markers on either side of an unmapped sequence are found, the chromosome walk can begin from one of the markers. Each successive gene in the sequence is tested repeatedly, known as overlap restrictions and mapped for their precise location in the sequence. Eventually, walking through the genes reaches the mutant gene in an unmapped sequence that binds to a fragment of a gene of that particular disease. The testing on each successive clone is complex, time-consuming, and varied by species. This series of overlapping clones could for example consist of Bacterial Artificial Chromosomes.

9.5.1 Hybridization Probe

A more straightforward approach thus is to use the insert DNA from the starting clone as a hybridization probe to screen all the other clones in the library. Positive hybridization signals that are given by clones, whose inserts overlap with the probe, are used as new probes to continue the walk. There are about 96 clones that a library consists of and each clone contains a different insert. A probe may have a genome wide repetition of sequences. This can be reduced by blocking the repeat sequence with pre-hybridization with unlabeled genomic DNA. But this isn't that affective solution especially in the case when high capacity vectors such as BACs or YACs are used in the walk. Therefore for

chromosome walks with human DNA which have a high rate of repetition, intact inserts are not used in general. Instead the probe is taken from the end of an insert which has a lesser chance of repetition. The walk can also be sped up by using the PCR instead of hybridization.

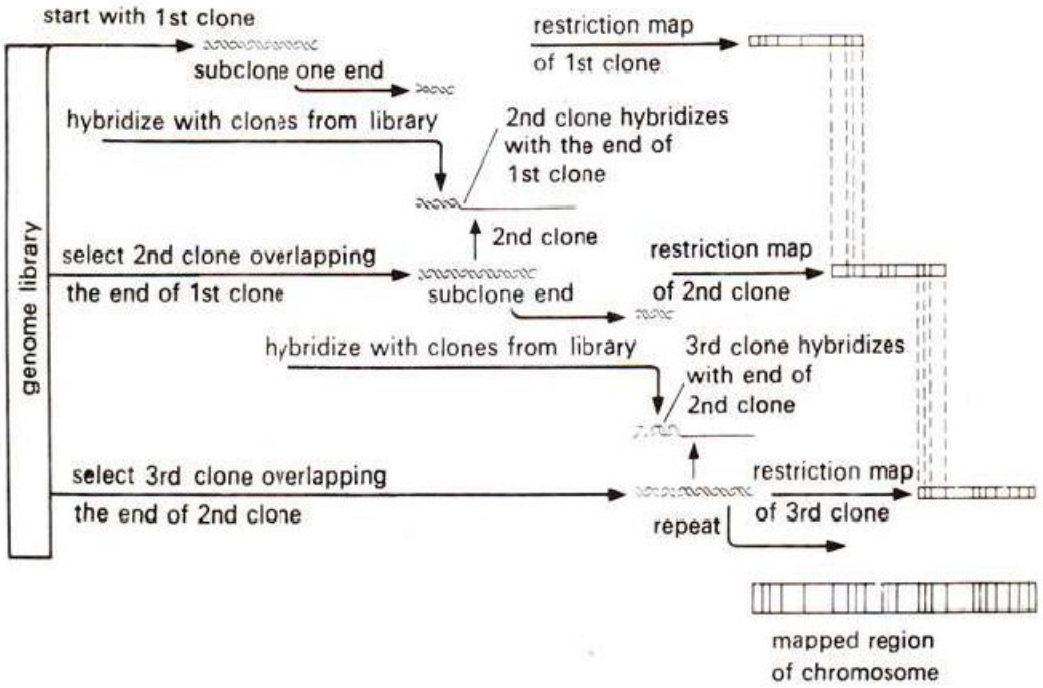


Fig.9.5:The technique of walking through successive hybridization between chromosomes overlapping genomic clones.

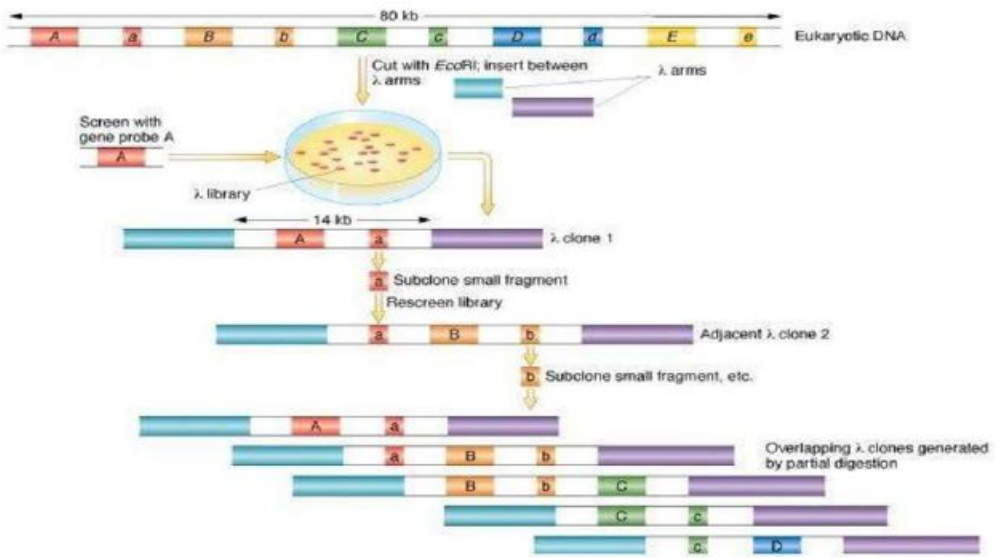


Fig.9.6: Cloning strategy of chromosome walking process

The technique of chromosome walking involves the following steps:

1. From the genomic library select a clone of interest (identified by a probe) and subclone a small fragment from one end of the clone (there is a technique available to subclone a fragment from the end);
2. The subcloned fragment of the selected clone may be hybridized with other clones in the library and a second clone hybridizing with the subclone of the first clone is identified due to presence of overlapping region;
3. The end of the second clone is then subcloned and used for hybridization with other clones to identify a third clone having overlapping region with the subcloned end of the second clone;
4. Third clone identified as above is also subcloned and hybridized with clones in the same manner and the procedure may be continued;
5. Restriction map of each selected clo.ne may be prepared and compared to know the regions of overlapping, so that identification of new overlapping restriction sites will amount to walking along the chromosome or along a long chromosome segment.
6. Regions of chromosome approaching 1000kb have been mapped following the above technique. Restriction maps of entire chromosomes can be prepared in this manner following the technique of chromosome walking.

Applications

- This technique can be used for the analysis of genetically transmitted diseases, to look for mutations.
- Chromosome Walking is used in the discovery of single-nucleotide polymorphism of different organisms.

Disadvantages

- There is a limitation to the speed of chromosome walking because of the small size of the fragments that are to be cloned.
- Another limitation is the difficulty of walking through the repeated sequence that are scattered through the gene.
- If the markers were too far away, it simply was not a viable option.
- Additionally, chromosome walking could easily be stopped by unclonable sections of DNA.

- A solution to this problem was achieved with the advent of chromosome jumping (Marx, 1989), which allows the skipping of unclonable sections of DNA.

9.6 Chromosome jumping

Chromosome jumping is a tool of molecular biology that is used in the physical mapping of genomes. It is related to several other tools used for the same purpose, including chromosome walking. Chromosome jumping is used to bypass regions difficult to clone, such as those containing repetitive DNA that cannot be easily mapped by chromosome walking, and is useful in moving along a chromosome rapidly in search of a particular gene. Unlike chromosome walking, chromosome jumping is able to start on one point of the chromosome in order to traverse potential distant point of the same chromosome without cloning the intervening sequences. The ends of a large DNA fragment is the target cloning section of the chromosome jumping while the middle section gets removed by sequences of chemical manipulations prior to the cloning step.

Process

Chromosome jumping enables two ends of a DNA sequence to be cloned without the middle section. Genomic DNA may be partially digested using restriction endonuclease and with the aid of DNA ligase, the fragments are circularized at low concentration. From a known sequence, a primer is designed to sequence across the circularized junction. This primer is used to jump 100 kb-300 kb intervals: a sequence 100 kb away would have come near the known sequence on circularization, it permits jumping and sequencing in an alternative manner. Thus, sequences not reachable by chromosome walking can be sequenced.Chromosome walking can also be used from the new jump position (in either direction) to look for gene-like sequences, or additional jumps can be used to progress further along the chromosome. Combining chromosome jumping to chromosome walking through the chromosome allows bypassing repetitive DNA for the search of the target gene.

Library

Chromosome jumping library is different from chromosome walking due to the manipulations executed before the cloning step. In order to construct the library of chromosome jumping, individual clones originate from random points in the genome (general jumping libraries first basic protocol) or from the termini of specific restriction fragments (specific jumping libraries alternate protocol) should be identified.

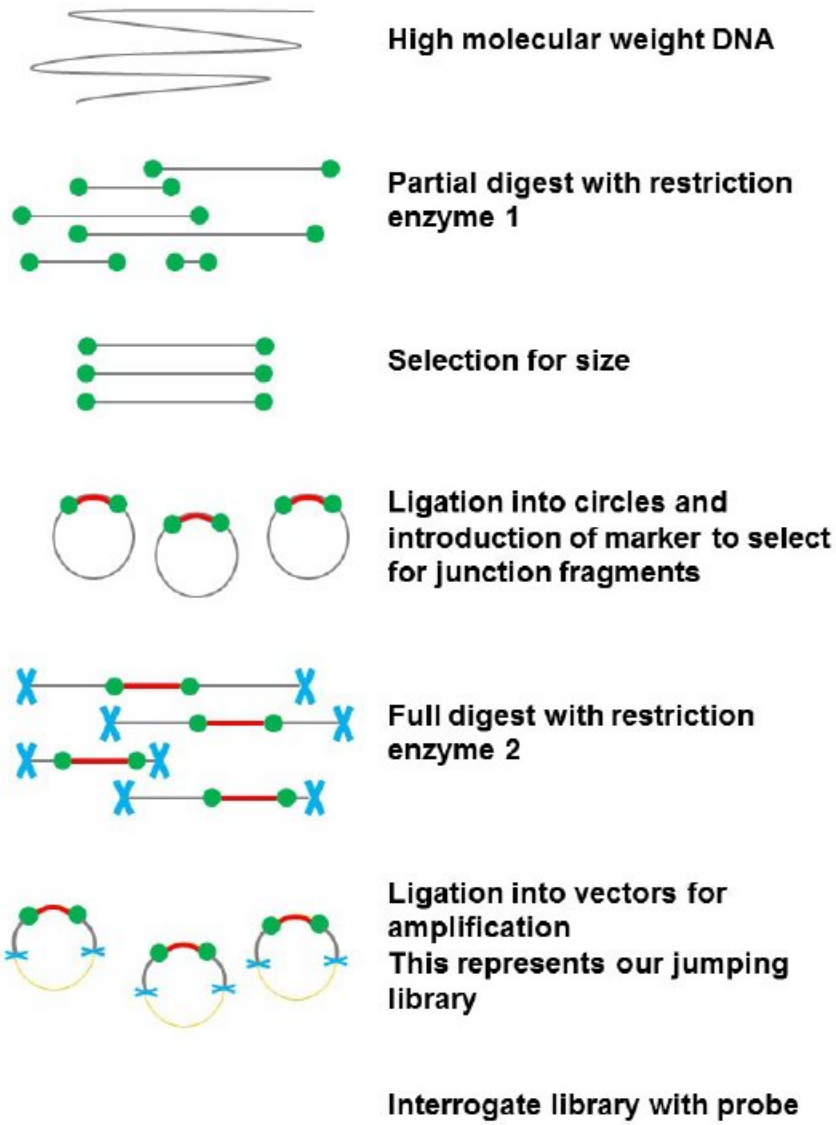


Fig.9.7: Method for creating a chromosome jumping library.

Clone Library Screening

The result of any cloning experiment that begins with total DNA from a specific source is a **library** of clones. Thus cloning mRNA through a cDNA intermediate generates a **cDNA library** that represents the mRNA from that specific developmental stage. In contrast, if you clone total human DNA for example into a lambda or cosmid vector, you obtain a **genomic library** that contains (with a probability of 95-99%) all the sequences from the human genome or DNA complement. Lambda or cosmid libraries are typically used for genomic libraries because you generally can clone an entire gene containing both the coding sequence and regulatory elements on a single clone. One of the key elements required to identify a gene during cloning is a **probe**. A probe is normally a cloned piece of

DNA that contains a portion of the sequence for which you are searching. You typically will make the probe radioactive and add it to a solution. Filters containing immobilized clones are then bathed in the solution. The principal behind this step is that the probe will bind to any clone containing sequences similar to those found on the probe. This binding step is called hybridization.

- **Probe** - a nucleic acid (usually DNA) that is complementary to a specific gene or nucleic acid sequence of interest; when a cDNA library is being screened an antibody can be used to identify the protein that is being expressed by the insert of the clone
- **Homologous Probe** - a probe that is exactly complementary to the nucleic acid sequence for which you are searching; ex., a human cDNA used to search a human genomic library
- **Heterologous Probe** - a probe that is similar to, but not exactly complementary to the nucleic acid sequence for which you are searching; ex., a mouse probe used to search a human genomic library

Screening a cDNA or Genomic Library

1. immobilize members of the library onto a nylon membrane and denature them so that they are single-stranded
2. prepare a radiolabelled probe and denature it to make it single-stranded
3. hybridize the probe to the library of clones
4. wash the excess probe and expose an X-ray film
5. isolate the positive clone and analyze

The hybridization step is performed at a **non-stringent** temperature that ensures the probe will bind to any clone containing a similar sequence. At the same time some non-specific hybridization will occur because some of the clones will contain limited, but not significant similarity to the probe. The washing step is performed at a **stringent** temperature that is high enough to wash the probe off all clones to which it has bound in a non-specific manner. But it is important that the temperature is not so high that it washes the probe off of clones that contain sequences that are similar or identical to the probe itself. Therefore, consideration about the source of the probe (homologous or heterologus) determine the temperature at which the washing step is performed.

NotI-digested DNA

One example to build a library is a classified as a rare-cutting restriction end nuclease such as NotI. In order to construct and characterize a library based from NotI-digested human DNA, random clones were analyzed by restriction mapping. Due to the wide distribution of fragment sizes made by

the complete digestion with NotI, the library was constructed into two fractions, low and high plasmid concentration. Clones that possessed unique end fragments were then analyzed by hybridization to Pulse Field Gradient (PFG) Southern blots. Examining the results gathered for single and double digests of human DNA with enzymes NotI, BssHII, and NruI, a restriction map with 850 kb was region containing the linking and jumping clones were created. Furthermore, NotI fragments of 250 and 350 kb jumps were evident in the two end clones derived corresponding to genetic distances of 0.25 and 0.35 cM.

9.7 Advantages and disadvantages

Allows more rapid movement through the genome compared to other techniques, such as chromosome walking. Able to travel across chromosomal regions containing unclonable sequences in bacterial hosts. Thirdly, this technique can be used to generate genomic markers with known chromosomal locations. Combination of jumping and linking jumping libraries to walking offers possibility of directional walking and might allow the analysis of longer regions in parallel mapping strategies. Reduces the complexity of libraries to be screened and constructed of mammalian genome. However, despite these advantages, chromosome jumping is still restricted by the capacity of the cloning vector which is the distance of the ends of the two fragments which can be approximately hundreds of kilobases. Additionally, because the jump does not clone the intervening DNA, chromosome walking would have to be done to identify all the genes present in the DNA. Regardless, it is still deemed to be beneficial due to the possibility to jump over hundred kilobases in comparison to chromosome walking.

9.8. Genetic disorders

Chromosome jumping libraries help address the complication of standard cloning techniques with large molecular distances. This process allowed the possibility to use the chromosome jumping library for other genetic disorders that requires 100 kilobases jumps. Particularly for genetic disorders such as cystic fibrosis, its gene is located in human chromosome 7, was able to utilize the chromosome jumping library to search for a jumping clone, met ontogeny. Identification of the cystic fibrosis was complicated due to it existing in eukaryotic genes that are composed with coding (exons) and non-coding (introns) segments, where introns are small in size making them difficult for detection. Another struggle in recognizing cystic fibrosis gene is because mammalian cells contains variety of repetitive DNA that can lead to incorrect cloning and blockage of DNA Replication and can cause instability. Both these complications, traditional cloning techniques are unable to process because large yield of

exons would have to be visible to produce a signal for the cystic fibrosis gene to be identified and DNA would have to be free of any repetitive elements.

9.8 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.

Blotting techniques are used to detect and analyze three types of biological macromolecules: DNA, RNA and proteins. Results of a blotting experiment tell you whether a macromolecule of a specific sequence is present in your sample or not. And, if it indeed is present, its abundance in the sample can also be quantified. Specifically, Southern, Northern and Western blotting is done to analyze DNA, RNA and protein samples respectively. 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, northern and 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 end nucleases are used to cut high-molecular-weight DNA strands into smaller

- ✓ Fragments, which are then electrophoreses 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 within 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 angel), 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 specific sequence whose presence in the target DNA is to be determined. The probe DNA is labeled so that it can be detected, usually by incorporating radioactivity or tagging themolecule 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.

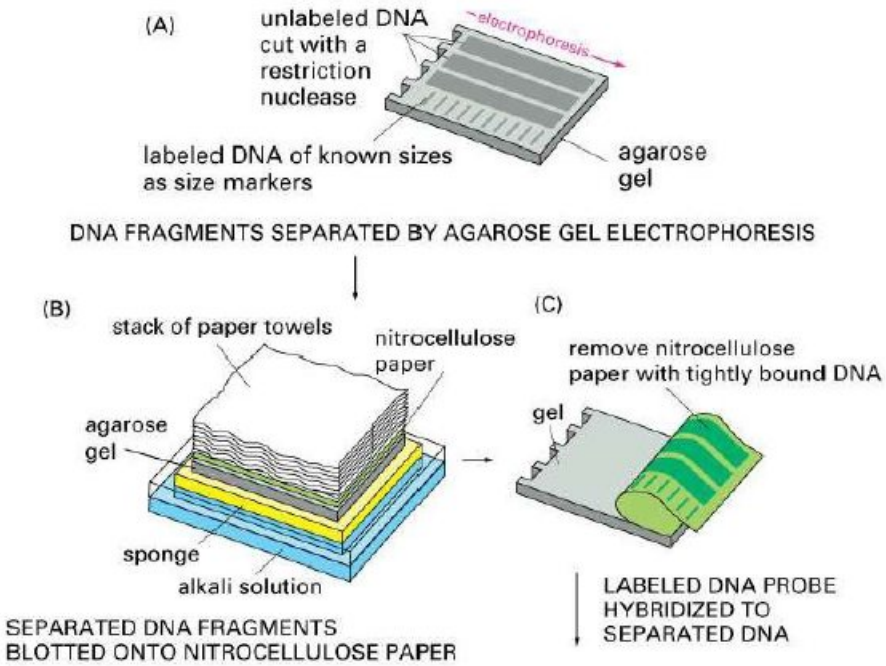


Fig.9.8: Different steps in blotting

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 beencut by the restriction enzyme will produce a single band on a Southern blot, whereas multiplebands 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 hybridizationconditions (ie, increasing the hybridization temperature or decreasing salt concentration) maybe used to increase specificity and decrease hybridization of the probe to sequences that areless 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 in1977 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 form amide 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.

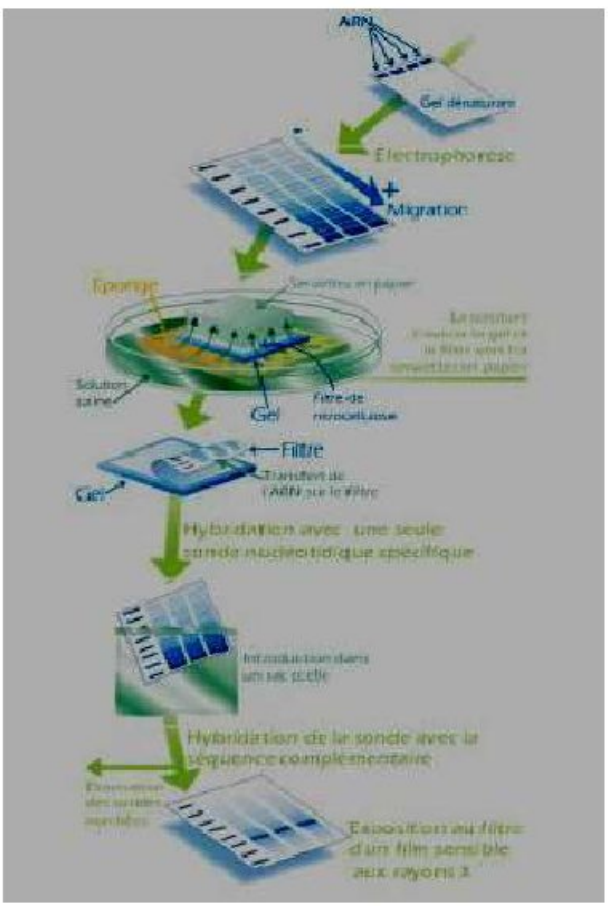


Fig.9.9: Different steps in blotting

Important 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 ontogenesis 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 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.

9.9. 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 (diethylpyro carbonate).

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 reproved for years after blotting.

Reverse northern blot

A variant of the procedure known as the reverse northern blot is occasionally used. In this procedure, the substrate nucleic acid (that is affixed to the membrane) is a collection of isolated DNA fragments, and the probe is RNA extracted from a tissue and radioactively labeled. These of DNA microarrays that have come into widespread use in the late 1990s and early 2000sis more akin to the reverse procedure, in that they involve the use of isolated DNA fragments affixed to a substrate, and hybridization with a probe made from cellular RNA. Thus the reverse procedure, though originally uncommon, enabled northern analysis to evolve into gene expression profiling, in which many (possibly all) of the genes in an organism may have their expression monitored.

Western blot

The western blot (alternatively, immunoblot) is used to detect specific proteins in a given sample of tissue homogenate or extract. The method originated from the laboratory of George Stark at Stanford. The name western blot was given to the technique by W. Neal Burnette.Steps in a western blot

Tissue preparation

Samples may be taken from whole tissue or from cell culture. In most cases, solid tissues are first broken down mechanically using a blender (for larger sample volumes), homogenizer (smaller volumes) or sonication. Assorted detergents, salts and buffers may be employed to encourage lysis of cells and to solubilize proteins. Protease and phosphates inhibitors are often added to prevent the digestion of the sample by its own enzymes. Tissue preparation is often done at cold temperatures to avoid protein denaturing. A combination of biochemical and mechanical techniques, including various types of filtration and centrifugation can be used to separate different cell compartments and organelles.

Gel electrophoresis

The proteins of the sample are separated using gel electrophoresis. Separation of proteins may be by isoelectric point (pI), molecular weight, electric charge or a combination of these factors. SDS-PAGE (SDS polyacrylamide gel electrophoresis) maintains polypeptides in a denatured state once they have been treated with strong reducing agents to remove secondary and tertiary structure and thus allows separation of proteins by their molecular weight. Sampled proteins become covered in the negatively charged SDS and move to the positively charged electrode through the acryl amide mesh of the gel. Smaller proteins migrate faster through this mesh and the proteins are thus separated according to size. The concentration of acryl amide determines the resolution of the gel - the greater the acryl

amide concentration the better the resolution of lower molecular weight proteins. The lower the acrylamide concentration the better the resolution of higher molecular weight proteins. Proteins travel only in one dimension along thegel for most blots.Samples is loaded into wells in the gel. One lane is usually reserved for a marker or ladder, commercially available mixture of proteins having defined molecular weights, typically stained so as to form visible, colored bands. When voltage is applied along the gel, proteins migrate into it at different speeds. These different rates of advancement separate into bands within each lane. It is also possible to use a two-dimensional (2-D) gel which spreads the proteins from single sample out in two dimensions. Proteins are separated according to is electric point (pHat which they have neutral net charge) in the first dimension, and according to their molecular weight in the second dimension.

Transfer

In order to make the proteins accessible to antibody detection, they are moved from within the gel onto a nitrocellulose or polyvinylidene difluoride (PVDF) membrane similar to Southern blot DNA transfer. Another method for transferring the proteins is called electro blotting and uses an electric current to pull proteins from the gel into the PVDF or nitrocellulose membrane. The proteins move from within the gel onto the membrane while maintaining the organization they had within the gel. As a result of this "blotting" process, the proteins are exposed on a thin surface layer for detection. Protein binding is based upon hydrophobic interactions, as well as charged interactions between the membrane and protein. Nitrocellulose membranes are cheaper than PVDF, but are far more fragile and do not stand up well to repeat probings.

The uniformity and overall effectiveness of transfer of protein from the gel to the membrane can be checked by staining the membrane with Coomassie or Ponceau S dyes. Ponceau S is the more common of the two, due to Ponceau S's higher sensitivity and its water solubility makes it easier to subsequently distain and probe the membrane.

Blocking

Since the membrane has been chosen for its ability to bind protein and both antibodies and the target are proteins, steps must be taken to prevent interactions between the membrane and the antibody used for detection of the target protein. Blocking of non-specific binding is achieved by placing the membrane in a dilute solution of protein - typically Bovine serum albumin (BSA) or on-fat dry milk (both are inexpensive), with a minute percentage of detergent such as Tween20. The protein in the dilute solution attaches to the membrane in all places where the target proteins have not attached. Thus, when the antibody is added, there is no room on the membrane for it to attach other than on the

binding sites of the specific target protein. This reduces "noise" in the final product of the Western blot, leading to clearer results, and eliminates false positives.

Detection

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

Two step

1. Primary antibody

Antibodies are generated when a host species or immune cell culture is exposed to the protein of interest. Normally, this is part of the immune response; whereas here they are harvested and used as sensitive and specific detection tools that bind the protein directly. After blocking, a dilute solution of primary antibody (generally between 0.5 and 5micrograms/ml) is incubated with the membrane under gentle agitation. Typically, the solution is composed of buffered saline solution with a small percentage of detergent, and sometimes with powdered milk or BSA. The antibody solution and the membrane can be sealed and incubated together for anywhere from 30 minutes to overnight. It can also be incubated at different temperatures, with warmer temperatures being associated with more binding, both specific (to the target protein, the signal) and non-specific (noise).

2. Secondary antibody

After rinsing the membrane to remove unbound primary antibody, the membrane is exposed to another antibody, directed at a species-specific portion of the primary antibody. This is known as a secondary antibody, and due to its targeting properties, tends to be referred to as "ant mouse," "anti-goat," etc. Antibodies come from animal sources (or animal sourced hybridomacultures); an anti-mouse secondary will bind to just about any mouse-sourced primary antibody. This allows some cost savings by allowing an entire lab to share a single source of mass-produced antibody, and provides far more consistent results. The secondary antibody is usually linked to biotin or to a reporter enzyme such as alkaline phosphates or horseradish peroxides. This means that several secondary antibodies will bind to one primary antibody and enhance the signal.

Most commonly, a horseradish peroxides-linked secondary is used in conjunction with a chemiluminescent agent, and the reaction product produces luminescence in proportion to the amount

of protein. A sensitive sheet of photographic film is placed against the membrane, and exposure to the light from the reaction creates an image of the antibodies bound to the blot. A cheaper but less sensitive approach utilizes a 4-chloronaphthol stain with 1% hydrogenperoxide; reaction of peroxide radicals with 4-chloronaphthol produces a dark brown stain thatcan be photographed without using specialized photographic film. As with the ELISPOT and ELISA procedures, the enzyme can be provided with a substrate molecule that will be converted by the enzyme to a colored reaction product that will be visible on the membrane. A third alternative is to use a radioactive label rather than an enzyme coupled to the secondary antibody, such as labeling an antibody-binding protein like Staphylococcus Protein A with a radioactive isotope of iodine. Since other methods are safer, quicker and cheaper this method is now rarely used.

One step

Historically, the probing process was performed in two steps because of the relative ease of producing primary and secondary antibodies in separate processes. This gives researchers and corporations huge advantages in terms of flexibility, and adds an amplification step to the detection process. Given the advent of high-throughput protein analysis and lower limits of detection, however, there has been interest in developing one-step probing systems that would allow the process to occur faster and with less consumables. This requires a probe antibody which both recognizes the protein of interest and contains a detectable label, probes which are often available for known protein tags. The primary probe is incubated with the membrane in amanner similar to that for the primary antibody in a two-step process, and then is ready for direct detection after a series of wash steps.

Analysis

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

Colorimetric detection

The colorimetric detection method depends on incubation of the western blot with a substrate that reacts with the reporter enzyme (such as peroxides) that is bound to the secondary antibody. This converts the soluble dye into an insoluble form of a different color that precipitates next to the enzyme

and thereby stains the membrane. Development of the blot is then stopped by washing away the soluble dye. Protein levels are evaluated through densitometry (how intense the stain is) or spectrophotometric.

Chemiluminescent detection

Chemiluminescent detection methods depend on incubation of the western blot with a substrate that will luminescent when exposed to the reporter on the secondary antibody. The light is then detected by photographic film, and more recently by CCD cameras which captures a digital image of the western blot. The image is analyzed by densitometry, which evaluates the relative amount of protein staining and quantifies the results in terms of optical density. Newer software allows further data analysis such as molecular weight analysis if appropriate standards aroused.

Radioactive detection

Radioactive labels do not require enzyme substrates, but rather allow the placement of medical-ray film directly against the western blot which develops as it is exposed to the label and creates dark regions which correspond to the protein bands of interest (see image to the right).The importance of radioactive detections methods is declining, because it is very expensive, health and safety risks are high and ECL provides a useful alternative.

Fluorescent detection

The fluorescently labeled probe is excited by light and the emission of the excitation is then detected by a photo sensor such as CCD camera equipped with appropriate emission filters which captures a digital image of the western blot and allows further data analysis such as molecular weight analysis and a quantitative western blot analysis. Fluorescence is considered to be among the most sensitive detection methods for blotting analysis.

Secondary probing

One major difference between nitrocellulose and PVDF membranes relates to the ability of each to support "stripping" antibodies off and reusing the membrane for subsequent antibody probes. While there are well-established protocols available for stripping nitrocellulose membranes, the sturdier PVDF allows for easier stripping, and for more reuse before background noise limits experiments. Another difference is that, unlike nitrocellulose, PVDF must be soaked in 95%ethanol, isopropanol or methanol before use. PVDF membranes also tend to be thicker and more resistant to damage during use.

2-D Gel Electrophoresis

2-dimensional SDS-PAGE uses the principles and techniques outlined above. 2-D SDS-PAGE, as the name suggests, involves the migration of polypeptides in 2 dimensions. For example, in the first dimension polypeptides are separated according to isoelectric point, while in the second dimension polypeptides are separated according to their molecular weight. The is electric point of a given protein is determined by the relative number of positively (e.g. lysine and argentine) and negatively (e.g. glutamate and aspartame) charged amino acids, with negatively charged amino acids contributing to a high is electric point and positively charged amino acids contributing to a low isoelectric point. Samples could also be separated first under non reducing conditions using SDS-PAGE and under reducing conditions in the second dimension, which breaks apart disulfide bonds that hold subunits together. SDS-PAGE might also be coupled with urea-PAGE for a 2-dimensional gel.

In principle, this method allows for the separation of all cellular proteins on a single large gel. Major advantage of this method is that it often distinguishes between different isoforms of particular protein - e.g. a protein that has been phosphorylated (by addition of a negatively charged group). Proteins that have been separated can be cut out of the gel and then analyzed mass spectrometry, which identifies the protein.

Eastern blotting

It is a technique to detect protein post translational modification and is an extension of the biochemical technique of western blotting. Proteins blotted from two dimensional SDS-PAGE gel on to a PVDF or nitrocellulose membrane are analyzed for post-translational protein modifications using probes specifically designed to detect lipids, carbohydrate, phosphomoieties or any other protein modification. The technique was developed to detect protein modifications in two species of Ehrlichia-E.muris and IOE. Cholera toxin B subunit (which detects lipids), Concanavalin A (which detects glucose moieties) and nitrophospho molybdate-methyl green (detects phosphoproteins) were used to detect protein modifications. The technique showed that the antigenic proteins of then on-virulent Immures are more post-translationally modified than the highly virulent IOE. The technique was conceptualized by S. Thomas while working on sandal spike phytoplasma and developed at the Dept. of Pathology, University of Texas Medical Branch, Galveston, Texas, while working on the intracellular bacteria, Ehrlichia.

9.10. Significance

Most of the proteins that are translated from mRNA undergo modifications before becoming functional in cells. The modifications collectively, are known as post-translational modifications

(PTMs). The nascent or folded proteins, which are stable under physiological conditions, are then subjected to a battery of specific enzyme-catalyzed modifications on the side chains orbackbones.Post-translational protein modifications includes: acetylation, acylation (myristoylation, palmitoylation), alkylation, arginylation, biotinylation, formylation, glutamylation, glycosylation, glycylation, hydroxylation, isoprenylation, lipoylation, methylation, nitroalkylation,phosphopantetheinylation, phosphorylation, prenylation, selenation, S-nitrosylation, sulfation,transglutamination and ubiquitination (sumoylation).

Post-translational modifications occurring at the N-terminus of the amino acid chain play animportant role in translocation across biological membranes. These include secretory proteins inprokaryotes and eukaryotes and also proteins that are intended to be incorporated in various cellular and organelle membranes such as lysosomes, chloroplast, mitochondria and plasmamembranes. Expression of post translated proteins is important in several diseases.

Applications of Blotting and Hybridization Techniques

- Southern blotting technique is widely used to find specific nucleic acid sequence present indifferent plant species.
- Northern blotting technique is widely used to find gene expression and regulation ofspecificgenes.
- By using blotting technique we can identify infectious agents present in the sample.
- We can identify inherited disease.
- It can be applied to mapping restriction sites in single copy gene.

Disadvantages of Blotting and Hybridization Techniques

- ✓ The process is a complex, cumbersome and time consuming one.
- ✓ It requires electrophoretic separation.
- ✓ Only one gene or RNA can be analyzed at a time.
- ✓ Gives information about presence of DNA, RNA or proteins but does not give information about regulation and gene interaction.

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 analyzing. None of the blot methods require electrophoresis prior to blotting and hybridization. Hybridization of cloned DNA without electrophoretic separation is called as dot blotting.

Plaque or Colony Blotting Techniques

This method was first developed by Grinstein's and Hogness (1975). This method is used to identify which colony of bacteria contains the DNA ofinterest among thousands. In this procedure, the bacterial colonies to be screened aretransferred onto nitrocellulose or nylon membrane by using replica plating. Due to the negative charge of the cell surface, some cells bind to the nitrocellulose membrane. Then the membrane is placed in a solution of 0.5 N NaOH to break the cell surface, convert dsDNA to ssDNA and to bind DNA to the membrane. Later, the membrane is transferred to absolution containing protease solution after neutralizing with neutralization solution.

The DNA is fixed tightly to membrane by either W cross linking or oven baking. This membrane is used for hybridization with a probe and analyzed by using autoradiography or biotin method for positive hybridization. A colony whose DNA print (as replica plating provides a replica print master plate colony on the membrane) gives a positive hybridization can be picked from the master plate. Plaque blotting is similar to colony blotting; the only difference is that instead of bacterial colony, a plaque is transferred onto the membrane. Benton and Davis developed this method in 1977.The greatest advantage of this method is that several identical DNA prints can be easily made from a single master plate containing bacteria/plaques which are to be made.

Dot Plot Assay Techniques

This method is widely used to hybridize DNA from a single cell type against a wide variety of probes, for example, for a viral infection which cannot be identified by normal conventional methods or if we want to know what all genes are expressed in a single cell type (e.g. brain cell). Cell type or cells that are to be screened are placed on the membrane as dot in the order of rows and columns. Then the cells are denatured by using enzymes or detergents (SDS) and DNA is fixed by using W - cross link or oven baking. This membrane is then used for hybridization by using probes (which are specific to a gene).

9.11. Summary

Under this unit we have summarized the concept of Chromogenic substrates, genetic selection, nucleic acid probes, chromosome walking & jumping and various types of blotting techniques with applications etc. Chromogenic substrates are colorless soluble molecules consisting of a chromophore, a chemical group that, after enzymatic cleavage, releases colour – and a specific enzymatic substrate. They are synthetically produced and are designed to possess a selectivity similar to the natural substrate for the enzyme. Genetic selection is the process by which certain traits become more prevalent in a species than other traits. It describes the forces that determine the traits seen in a

population or species. Natural selection involves forces such as predation or sexual preference or even something seemingly as mundane as tail size.

Blotting techniques involve in the separation (via electrophoresis) and transfer of DNA, RNA, or proteins onto a blotting membrane. The target DNA is then attached to a molecule in order to aid detection. Southern blotting is used to evaluate for specific DNA sequences and may be used in identification of genetic mutations and in forensics. Northern blotting focuses on RNA sequences and is helpful in assessing gene expression. Western blotting identifies proteins and antibodies and has applications in diagnosing infectious diseases, protein abnormalities (such as prion disease), and autoimmune conditions.

9.12. Terminal Questions

Q. 1 What do you mean by Chromogenic substrates? Explain it.

Answer:

Q. 2 Describe the different types of blotting techniques.

Answer:

Q. 3 Describe different applications of blotting and hybridization techniques.

Answer:

Q. 4 What are nucleic acid probe? Explain it.

Answer:

Q. 5 Write short note on the followings.

(i) Genetic selection

(ii) Hybridization Probe

Answer:

Q. 6 Write short note on the followings.

(i) Chromosome jumping

(ii) Chromosome Walking

Answer:

Q. 7 Write the short note on the clone library screening.

Answer:

9.13. Further Readings

1. Biochemistry- Lehninger A.L.
2. Biochemistry –J.H.Weil.
3. Biochemistry fourth edition-David Hames and Nigel Hooper.
4. Textbook of Biochemistry for Undergraduates - Rafi, M.D.
5. Biochemistry and molecular biology- Wilson Walker.

Uttar Pradesh Rajarshi Tandon
Open University

PGBCH - 118 N
Genetic Engineering

Block- 4

Genetic Engineering- III

Course Design Committee

Dr. (Prof.) Ashutosh Gupta, School of Science, UPRTOU, Prayagraj	Chairman
Prof. Prof. Umesh Nath Tripathi Department of chemistry Deen Dayal Upadhyaya Gorakhpur University, Gorakhpur	Member
Prof. S.I. Rizvi Department of Biochemistry University of Allahabad, Prayagraj	Member
Prof. Dinesh Yadav Department of Biotechnology Deen Dayal Upadhyaya Gorakhpur University, Gorakhpur	Member
Prof. Sharad Kumar Mishra Department of Biotechnology Deen Dayal Upadhyaya Gorakhpur University, Gorakhpur	Member
Dr. Ravindra Pratap Singh Assistant Professor (Biochemistry) School of Science, UPRTOU, Prayagraj	Member
Dr. Dharmveer Singh Assistant Professor (Biochemistry) School of Science, UPRTOU, Prayagraj	Course Coordinator

Course Preparation Committee

Dr. Gopal Dixit Assistant Professor Department of Botany, Upadhi Mahavidyalaya, Pilibhit, U.P.	**Author**	**Block-1-2**	Unit: 1-5
Dr. Arun Kumar Pandey Assistant Professor Department of Botany, PSMPG College, Maharajganj. U.P.	**Author**	**Block-2-3**	Unit: 6-9
Dr. Sadhana Singh Assistant Professor- Biochemistry School of Sciences, UPRTOU, Prayagraj	**Author**	**Block-4**	Unit: 10, 12
Dr. Anuradha Singh Assistant Professor- Botany School of Sciences, UPRTOU, Prayagraj	**Author**	**Block-4**	Unit: 11
Dr. Mohd. Khalid Masroor Retd. Associate Professor-Botany, University of Allahabad, U.P.	**Editor**	(Block- 01, 02, 03&04, Unit: 1, 2,3,4,5, 10, 11, &12)	
Dr. Rajiv Ranjan Associate Professor, MLKPG College, Balrampur, U.P.	**Editor**	(Block- 02& 03)	Unit: (6-9)
Dr. Dharmveer Singh Assistant Professor (Biochemistry) School of Sciences, UPRTOU, Prayagraj			**(SLM & Course Coordinator)**

PGBCH – 118, Genetic Engineering
©UPRTOU , 2024
ISBN :

Printed By: K.C.Printing & Allied Works, Panchwati, Mathura -281003.

Introduction

The following three units are included in the Fourth block of genetic engineering are as:

Unit-10: This unit covers Genetic engineering in action. Analysis of gene structure and function, Techniques- Gel retardation, DNA foot printing, DNA finger printing, Primer, extension, S1 mapping, and Dot-blot analysis are discussed in this unit. Gene structure and function analysis involves techniques like gel retardation to study protein-DNA interactions, DNA footprinting to map binding sites, DNA fingerprinting for genetic identification, primer extension to determine transcription start sites, S1 mapping for mapping nucleases-sensitive regions, and dot-blot analysis for detecting specific DNA or RNA.

Unit-11: This unit covers the Immuno-technology. The Hybridoma technology, Monoclonal antibodies and its production, Antibody engineering, Uses of monoclonal antibodies are discussed. Hybridoma technology fuses myeloma cells with antibody-producing B cells to create hybrid cells that produce monoclonal antibodies. Antibody engineering involves modifying these antibodies for improved specificity or functionality.

Unit-12: This unit covers the transgenic animals, transgenic animal's vs cloned animals, producing transgenic animals, and applications of transgenic animals are mentioned here. Transgenic animals have foreign genes inserted into their genomes, whereas cloned animals are genetically identical to the donor animal. Producing transgenic animals involves gene insertion into embryos, while cloning uses somatic cell nuclear transfer. Applications include disease models, drug production, and studying gene function and expression.

Unit 10- Genetic Engineering in Action

Content

10.1. Introduction

Understanding the complicate details of gene structure and function is fundamental to the fields of genetics and molecular biology. This understanding not only provides insights into the fundamental processes that govern cellular operations but also paves the way for advancements in medical research, biotechnology, and evolutionary studies.

Genetic engineering, the artificial manipulation, modification, and recombination of DNA or other nucleic acid molecules, in order to modify an organism or population of organisms. the term

genetic engineering is generally used to refer to the methods of recombinants DNA technology, which emerged from basic research in microbial genetics. The techniques employed in genetic engineering have led to the production of medically important products, including human insulin, human growth hormone, and hepatitis B Vaccine, as well as to the development of genetically modified organisms such as disease resistant plants.

The analysis of gene structure and function involves a variety of sophisticated techniques designed to unravel the complex interactions between DNA, RNA, and proteins. One of the essential techniques in this regard is the gel retardation assay, which helps in identifying DNA-protein interactions by measuring the mobility of DNA fragments through a gel matrix. Similarly, DNA foot printing provides a more refined method to pinpoint the exact locations where proteins bind to DNA. DNA fingerprinting, on the other hand, offers a powerful tool for identifying genetic variations among individuals, which is crucial for applications ranging from forensic investigations to genetic disease research.

Genetic engineering has been used successfully to develop novel genes of economic importance that can be used to improvethe genetics of crop plants. Genetic engineering is the targeted addition of a foreign gene or genes into the genomes of an organism. The genes may be isolated from one organism and transferred to another or may be genes of one species that are modified and reinserted into the same species. The new genes, commonly referred to as transgenes, one inserted into a plant by transformation. The inserted genes hold information that will give the organism a trait.

Additionally, techniques like primer design and primer extension are vital for amplifying and analyzing specific DNA sequences, thus facilitating detailed study of gene expression and mutation. S1 mapping and dot-blot analysis further extend our ability to explore gene structure by allowing the detection and quantification of specific nucleic acid sequences within complex mixtures. Crop genetic improvement (plant breeding) is an important tool but has limitations. Genetic engineering physically removes the DNA from one organism and transfers the gene(s) for one or few traits into another. A basic explanation of the five steps for genetically engineering acrop is provided. The five steps are:

1-Locating an organism with specific trait and extracting its DNA.

2-Cloning a gene that controls the trait.

3- Designing a gene to express in a specific manner.

4- Transformation, inserting the gene into the cells of a crop plant.

5- Cross the transgene intoan elite background.

Each of these methodologies contributes uniquely to our comprehensive understanding of gene functions and their regulation. By integrating these techniques, researchers can delve deeper into genetic mechanisms, paving the way for innovations in diagnostics, therapeutics, and fundamental biological research. In this unit learners will get to know into these techniques in detail, elucidating their principles, applications, and the critical insights they provide into the molecular basis of life.

Objectives

- To understand Gene Structure and Function
- To understand Gel Retardation Technique
- To get knowledge on DNA Foot Printing Technique
- To get knowledge on DNA Foot Printing Technique
- To understand S1 Mapping Technique and Dot-blot analysis

10.2. Analysis of Gene Structure and Function

Introduction to Gene Structure and Function

A gene is a basic unit of inheritance passed on from male and female parents to their children. Genes contain DNA, which is made up of sequences that determine the physical and biological traits of each person. Genes are building block of life. They contain information for making specific molecules and proteins that allow human cells to function and that control how the body grows and operates. They also lead to the expression of particular physical characteristics and traits like hair or eye color. Most genes code for specific proteins which have different functions throughout the body and allow humans to live, grow and reproduce.Genes are made up of sections of DNA. DNA is made up of chemical building blocks called nucleotide. A gene consists of four different nucleotide bases, which can be sequenced in different ways. Different sequences of bases determine different instructions, which account for various physical traits, like changes in genes can lead to incorrectly formed proteins that cannot function correctly. Thesesare called gene mutations and may lead to genetic disorders.

Genes are fundamental units of heredity that encode information for the synthesis of proteins and, in some cases, functional RNA molecules. The structure of a gene and the way it functions are intricately linked and crucial for understanding biological processes at a molecular level. Gene structures typically include various components such as exons, introns, promoters, enhancers, and regulatory sequences. The function of a gene involves transcription, RNA processing, and translation into proteins, which then participate in cellular processes.

10.2.1. Gene Structure

Components of a Gene

1. **Exons and Introns:** Exons are the coding regions of a gene that are expressed and translated into proteins, while introns are non-coding regions interspersed between exons (**Figure 10.1a and b**). During the initial transcription phase, both exons and introns are copied into a precursor mRNA (pre-mRNA). Introns are later removed during RNA splicing, leaving only the exons in the mature mRNA.

Characteristics of Exons:

Exons are coding sections of an RNA transcript, or the DNA encoding it, that are translated into protein.

- ✓ **Coding Sequences:** Exons are the portions of a gene that are retained in the final messenger RNA (mRNA) transcript after splicing. They contain the coding sequences that are translated into proteins.
- ✓ **Presence in Mature mRNA:** Exons are spliced together from the precursor mRNA (pre-mRNA) to form the mature mRNA that is eventually translated into protein.
- ✓ **Variability:** The number and length of exons can vary significantly between different genes and between different organisms. Some genes may have only one exon, while others may have multiple exons separated by introns.

Functions of of Exons:

- ✓ **Protein Coding:** The primary function of exons is to provide the coding sequences that dictate the amino acid sequence of the protein. Each exon usually corresponds to a specific domain or functional part of the protein.

✓ **Alternative Splicing:** Exons play a key role in alternative splicing, a process that allows a single gene to produce multiple protein isoforms. By including or excluding different exons, cells can generate proteins with varying functions and properties.

Characteristics of Introns:

An intron is any nucleotide sequence within a gene that is not expressed or operative in the final RNA product. The word intron is derived from the term intragenic region i.e. aregion inside a gene. The term intron refers to both the DNA sequence within a gene and the corresponding RNA sequence in RNA transcript.

✓ **Non-Coding Sequences:** Introns are non-coding regions interspersed between exons within a gene. They are transcribed into pre-mRNA but are removed during RNA splicing.

✓ **Variable Length:** Introns can vary greatly in length. Some introns are relatively short, while others can be very long.

✓ **Presence in Pre-mRNA:** Introns are initially present in the pre-mRNA transcript but are not included in the mature mRNA that is translated into protein.

Functions of Introns:

✓ **Regulation of Gene Expression:** Introns can influence gene expression through several mechanisms. For example, they may contain regulatory elements such as enhancers or silencers that affect the transcription of neighboring exons.

✓ **Alternative Splicing:** Introns are essential for alternative splicing. They provide the necessary sequences that are spliced out in various combinations to generate different mRNA isoforms from the same gene.

✓ **Facilitation of RNA Processing:** Some introns contain sequences that are involved in the regulation of RNA splicing and other post-transcriptional modifications. For example, some introns contain branch points that are crucial for the splicing reaction.

✓ **Evolutionary Significance:** Introns may contribute to genomic evolution by providing sites for recombination and genetic variability. Their presence can influence the evolution of gene functions and the development of new proteins.

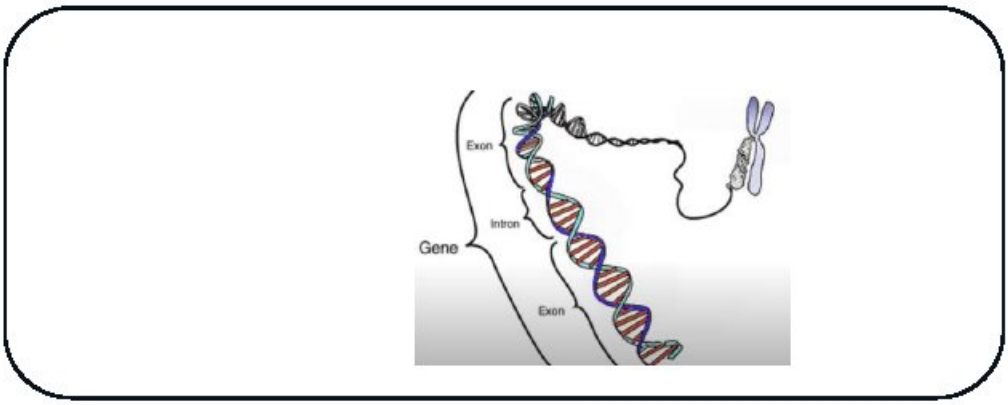


Figure 10.1a Structure of gene https://www.youtube.com/watch?v=_asGjfCTLNE

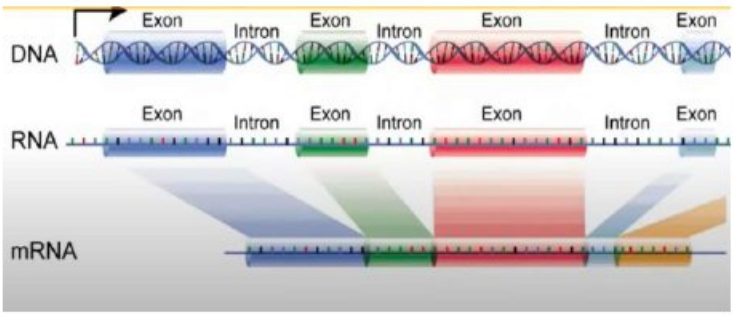


Figure 10.1b Exons Vs Introns https://www.youtube.com/watch?v=_asGjfCTLNE

Promoters: The promoter is a regulatory sequence located upstream of the coding region. It serves as the binding site for RNA polymerase and various transcription factors, initiating the transcription process (**figure 10.2a, b**). The strength and specificity of a promoter influence the level of gene expression. The function of promoters in gene expression can be summarized as follows:

- **Initiation of Transcription:** Promoters serve as the binding sites for RNA polymerase and essential transcription factors. They mark the starting point for transcription, enabling RNA polymerase to synthesize mRNA from the DNA template. Without a functioning promoter, transcription cannot commence.
- **Regulation of Gene Expression:** Promoters regulate the rate of gene transcription. They can be classified as strong or weak based on their ability to attract RNA polymerase and transcription factors. Strong promoters facilitate high levels of transcription, while weak promoters result in lower levels of gene expression.
- **Tissue-Specific and Developmental Regulation:** Some promoters are active only in specific tissues or developmental stages. This tissue-specific regulation ensures that genes are expressed only where and when they are needed, contributing to proper development and function of different cell types.

➢ **Response to Environmental and Cellular Signals:** Promoters can respond to various internal and external signals, such as stress or nutrient availability. This allows cells to adapt their gene expression profiles in response to changing conditions, ensuring that appropriate genes are activated or repressed as needed.

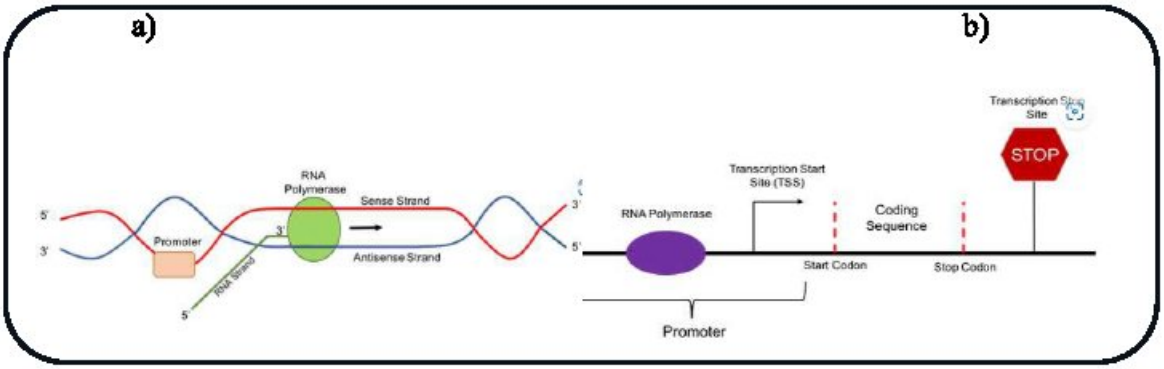


Figure.10.2 a. Structure of Promoter; **Figure.10.2b.** Prompter regions (https://www.addgene.org/mol-bio-reference/promoters/)

2. **Enhancers and Silencers:** Enhancers are distal regulatory elements that increase the likelihood of transcription by interacting with the promoter region, often through DNA looping mechanisms (**figure 10.3**). Conversely, silencers are sequences that repress gene transcription.

Enhancers are regulatory DNA sequences that play a crucial role in the control of gene expression by increasing the transcription of associated genes. Unlike promoters, which are located upstream of the transcription start site, enhancers can be situated far away from the gene they regulate, often thousands of base pairs upstream or downstream, or even within introns of other genes. They function by binding to specific transcription factors and forming loops that bring them into close proximity with the promoter region of the gene. This interaction enhances the ability of RNA polymerase to initiate transcription, thus boosting gene expression. Enhancers are highly versatile and can be active in a tissue-specific or developmental stage-specific manner, allowing for precise and regulated gene expression. Their activity is often influenced by various signaling pathways, making them essential for the dynamic regulation of genes in response to environmental and physiological changes.

Gene silencers are regulatory elements within DNA that play a crucial role in gene expression by repressing or silencing the transcription of specific genes. Unlike promoters, which initiate gene expression, silencers function as binding sites for transcriptional repressors

or other regulatory proteins that inhibit the transcription machinery's access to the gene. These elements can be located far from the genes they regulate, often interacting with promoters through DNA looping mechanisms to exert their effects. Gene silencers are essential for controlling gene activity in various biological processes, including development, differentiation, and response to environmental stimuli. They contribute to the precise regulation of gene expression necessary for maintaining cellular function and overall organismal health.

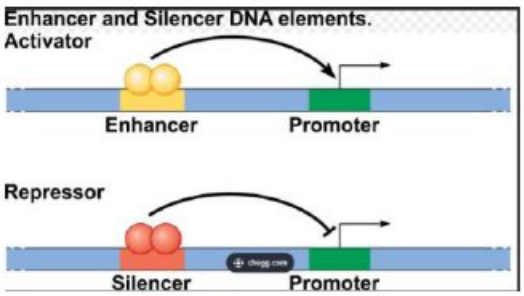


Figure:10.3: Enhancers andsilencers elements (https://www.chegg.com/homework-help/questions-and-answers/activator-enhancer-promoter-repressor-silencer-promoter-enhancer-silencer-dna-elements-reg-q60936769)

✓ **Regulatory Sequences:** A regulatory sequence is a segment of nuclei acid molecule which is capable of increasing or decreasing the expression of specific genes within an organism. Regulation of gene expression is an essential feature of all living organism and viruses. These include various binding sites for transcription factors and other proteins that regulate gene expression. They can be located within introns, exons, or in intergenic regions.Besides promoters, enhancers, and silencers, there are several other types of regulatory sequences that can influence gene expression:

✓ **Insulators**: These sequences function to block or insulate the effects of enhancers or silencers on adjacent genes. They can create boundaries that separate different regulatory domains, preventing unintended interactions between regulatory elements of neighboring genes.

✓ **Response Elements**: Found within promoters or enhancers, response elements are specific sequences that respond to various external signals, such as hormones or environmental factors. They allow genes to be regulated in response to changes in the cell's environment.

✓ **Splicing Regulatory Elements**: These include exonic splicing enhancers (ESEs) and exonic splicing silencers (ESSs) found within exons. They influence the splicing of pre-mRNA, affecting which exons are included in the final mRNA transcript and thereby influencing protein diversity.

✓ **Promoter-Proximal Elements**: These are regulatory sequences located close to the core promoter but outside of the canonical promoter region. They help regulate the efficiency of transcription initiation by interacting with transcription factors and other regulatory proteins.

✓ **Long Non-Coding RNA (lncRNA) Regulatory Elements**: These are sequences that produce long non-coding RNAs, which can interact with DNA, RNA, or proteins to regulate gene expression. lncRNAs can influence chromatin structure and gene expression in various ways.

✓ **Gene-Specific Enhancer/Silencer Modules**: These are specialized types of enhancers and silencers that act in a gene-specific manner. They can be found in introns, exons, or intergenic regions and are tailored to regulate specific genes or gene clusters.

✓ **CpG Islands**: These are regions rich in CpG dinucleotides (cytosine followed by guanine) often found near gene promoters. Methylation of CpG islands can lead to gene silencing and is an important mechanism of epigenetic regulation.

✓ **UTR Regulatory Sequences**: The 5' and 3' untranslated regions (UTRs) of mRNA contain regulatory sequences that control mRNA stability, localization, and translationefficiency. These include elements like the iron response element (I RE) and the AU-rich element (ARE).

Each of these regulatory sequences contributes to the precise control of gene expression, ensuring that genes are expressed at the correct levels and times.

11.2.2. Functional Domains

Genes can be further classified based on their functional domains. For example, transcription factors often have DNA-binding domains that specifically recognize and bind to particular DNA sequences, influencing the transcription of target genes.

1. **DNA-Binding Domains (DBDs)**: These domains allow transcription factors to bind specifically to DNA sequences, known as DNA motifs or recognition sites, in the promoter or enhancer regions of target genes.

Examples:

✓ **Zinc Finger Domains**: Contain cysteine and histidine residues that coordinate zinc ions. They bind to specific DNA sequences by inserting loops of the protein into the major groove of the DNA. Example: **Kruppel-like factors (KLFs).**

✓ **Helix-Turn-Helix (HTH) Domains**: Consist of two α-helices connected by a short turn. The second helix fits into the major groove of the DNA, allowing specific binding. Example: **Homeobox (HOX) proteins**.

✓ **Leucine Zipper Domains**: Composed of leucine residues at regular intervals, forming a coiled-coil structure that facilitates dimerization and DNA binding. Example: **c-Fos/c-Jun**.

✓ **Basic Helix-Loop-Helix (bHLH) Domains**: Include a basic region that interacts with DNA and a helix-loop-helix motif that facilitates dimerization. Example: **Myc family proteins**.

2. **Activation Domains**: These domains interact with other proteins or components of the transcription machinery to promote the initiation of transcription. They do not directly bind to DNA but rather help recruit or stabilize the transcriptional machinery.

Examples:

✓ **Acidic Activation Domains**: Rich in acidic residues (e.g., glutamic acid) that interact with components of the transcriptional machinery. Example: **Activator protein 1 (AP-1)**.

✓ **Glutamine-Rich Domains**: Rich in glutamine residues, these domains interact with other transcriptional co-activators. Example: **CREB-binding protein (CBP)**.

3. **Repression Domains**:These domains are responsible for recruiting co-repressors or interacting with histone deacetylases to inhibit transcription.

Examples: Repression Domains: Often contain motifs that bind to co-repressors or other proteins that modify chromatin structure to inhibit gene expression. Example: **N-CoR (Nuclear receptor co-repressor)**.

4. **Dimerization Domains**: Facilitate the formation of protein dimers, which is often necessary for the transcription factor to bind DNA and regulate target genes effectively.

Examples:

✓ **Leucine Zipper**: Facilitates dimerization of proteins, often forming functional transcription factor complexes. Example: **c-Fos/c-Jun**.

✓ **bHLH Dimerization**: Allows formation of heterodimers or homodimers that bind DNA. Example: **Myc-Max**.

5. **Nuclear Localization Signals (NLS)**: These sequences direct the transcription factor to the nucleus where it can interact with DNA and the transcriptional machinery. **Examples: SV40 Large T-antigen NLS.**

Role in Gene Regulation

Transcription factors with these domains regulate gene expression through several mechanisms:

- **Binding Specificity**: The DNA-binding domains ensure that transcription factors bind to specific sequences in the DNA, thereby targeting specific genes for regulation.
- **Transcriptional Activation or Repression**: Activation and repression domains determine whether a gene is turned on or off by interacting with the transcriptional machinery or other regulatory proteins.
- **Chromatin Remodeling**: By recruiting co-activators or co-repressors, transcription factors can influence chromatin structure, making DNA more or less accessible for transcription.
- **Complex Formation**: Many transcription factors function as part of larger protein complexes, and dimerization or multimerization domains facilitate these interactions.

In conclusion, the functional domains within transcription factors are essential for their role in regulating gene expression. Each domain contributes to a specific aspect of transcriptional control, from recognizing and binding to DNA to interacting with other proteins that influence the transcriptional process.

10.2.3. Gene Function

a. Transcription and RNA Processing

The process of gene expression begins with transcription, where a gene's DNA sequence is copied into a complementary RNA sequence by RNA polymerase (**figure 10.4 a, b**). This pre-mRNA undergoes several processing steps:

- **Capping:** A 7-methylguanylate (7mG) cap is added to the 5' end of the pre-mRNA, protecting it from degradation and assisting in ribosome binding for translation.

- **Splicing:** Introns are removed, and exons are joined together to form the mature mRNA. This splicing process can result in different isoforms of a protein through alternative splicing.

- **Polyadenylation:** A poly-A tail is added to the 3' end of the mRNA, enhancing stability and facilitating export from the nucleus to the cytoplasm.

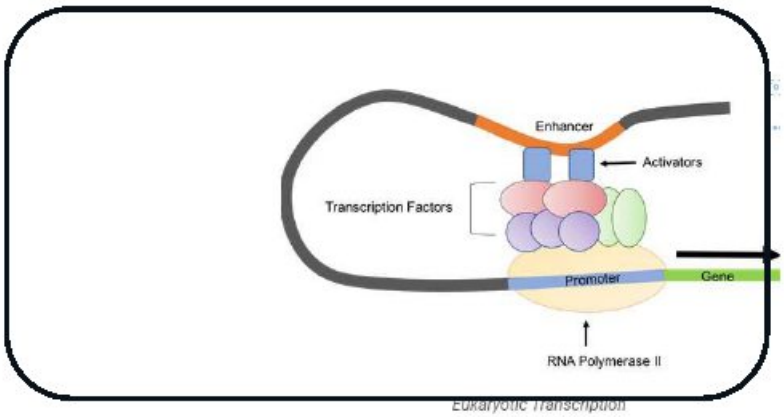


Figure 10.4a: Process of Transcription (https://www.addgene.org/mol-bio-reference/promoters/)

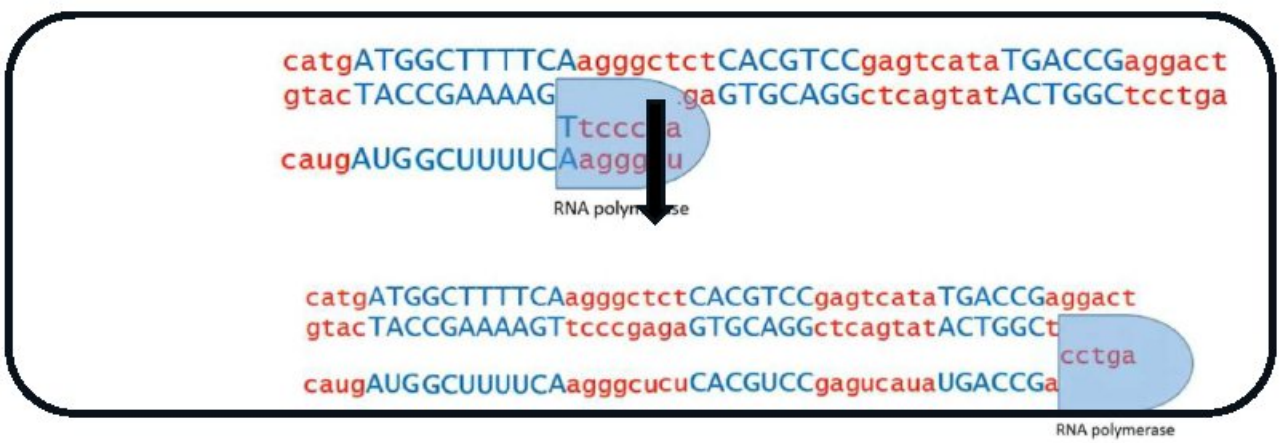


Figure10.4b. Process of Transcription (https://www.youtube.com/watch?v=_asGjfCTLNE)

b. Translation

In the cytoplasm, the mature mRNA is translated into a protein by ribosomes. Translation involves several steps:

- **Initiation:** The ribosome assembles at the start codon of the mRNA, and the first tRNA molecule binds to the start codon, bringing the first amino acid.

- **Elongation:** tRNA molecules sequentially add amino acids to the growing polypeptide chain according to the codon sequence on the mRNA.

- **Termination:** The ribosome reaches a stop codon, and the newly synthesized protein is released. The protein then undergoes folding and post-translational modifications to become functionally active.

SAQ 1(Self-assessment question)

Fill in the blanks-

1. …….. are the coding regions of a gene that are expressed and translated into proteins, while …… are non-coding regions interspersed between exons.

2. During RNA splicing, …….. are removed from the precursor mRNA (pre-mRNA), leaving only the …… in the mature mRNA.

3. The …….. is a regulatory sequence located upstream of the coding region that serves as the binding site for RNA polymerase and various ……..

4. …….. are distal regulatory elements that increase the likelihood of transcription by interacting with the …….. region through …….. mechanisms.

5. …….. are sequences that repress gene transcription by binding …….. or other regulatory proteins that inhibit the transcription machinery's access to the gene.

6. …….. are sequences that function to block or insulate the effects of …….. or …….. on adjacent genes, creating boundaries between different regulatory domains.

7. The …….. of mRNA contains regulatory sequences that control mRNA ……, localization, and ……..

8. …….. contain cysteine and histidine residues that coordinate zinc ions and bind to specific DNA sequences by inserting loops of the protein into the major groove of the DNA. An example is ……..

9. …….. are rich in acidic residues and interact with components of the transcriptional machinery to promote the …….. of transcription. An example is ……..

10. During ....., the ribosome assembles at the start codon of the mRNA, and the first ..... molecule binds to the start codon, bringing the first amino acid.

10.3 Gel Retardation Technique

The Gel Retardation Technique, also known as the Electrophoretic Mobility Shift Assay (EMSA), is a laboratory method used to study protein-DNA or protein-RNA interactions. A fragment of DNA with a protein bound to it migrates slower in a native polyacrylamide gel than the DNA fragment alone. It helps researchers identify and characterize the binding of proteins to specific nucleicacid sequences. Here's a basic overview of how it works:

Principle:

When a protein binds to a DNA or RNA fragment, the complex formed usually has a different electrophoretic mobility compared to the free nucleic acid. This is because the bound nucleic acid-protein complex is larger and has a different charge distribution compared to the free nucleic acid.

Procedure:

1. *Preparation of Samples*:

- Nucleic Acids: Typically, a labeled DNA or RNA fragment is used. The label can be a radioactive isotope or a non-radioactive tag (e.g., biotin or fluorescent dye).
- Proteins: The protein of interest is prepared, often including various concentrations or different proteins to study specific interactions.

2. Binding Reaction:

- Mix the labeled nucleic acid with the protein(s) in a reaction buffer that supports binding. This buffer may include salts, proteins, and other components that stabilize the interaction.

3. Gel Electrophoresis:

- Load the reaction mixture onto a gel (usually a polyacrylamide gel). The gel is typically run under non-denaturing conditions, meaning the proteins and nucleic acids are not denatured during the electrophoresis process.

4. Detection:

- After electrophoresis, the gel is treated to detect the labeled nucleic acids. If using radioactive labels, the gel can be exposed to a film. For non-radioactive labels, the gel might be exposed to a chemiluminescent or fluorescent substrate.

5. Analysis:

- The resulting bands are compared to identify the shift in mobility of the nucleic acid bands. Free nucleic acids will migrate differently compared to nucleic acids bound to proteins. The presence of a slower-migrating band indicates a protein-DNA or protein-RNA complex.

Applications:

- Identifying Protein Binding Specificity: Determine which proteins bind to specific DNA or RNA sequences.
- Characterizing Protein-DNA/RNA Interactions: Analyze the strength and nature of protein interactions with nucleic acids.
- Studying Transcription Factors: Investigate how transcription factors bind to promoter regions or other regulatory sequences.

The EMSA is a powerful technique for understanding molecular interactions in many biological processes, including gene regulation and cellular responses to environmental signals.

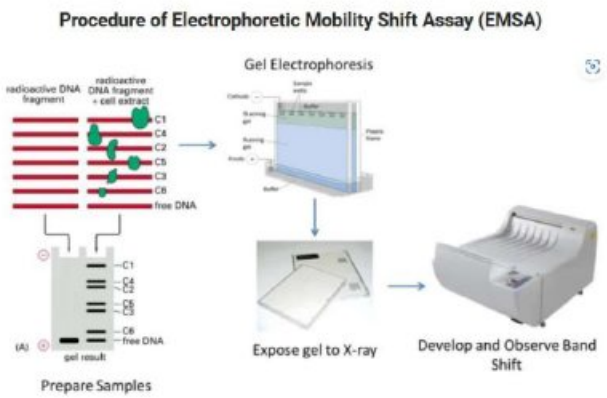


Figure 10.5: Gel Retardation Technique(https://thesciencenotes.com/electrophoretic-mobility-shift-assay-emsa-principle-procedure-application-troubleshooting-controls-limitations/)

SAQ2

1. The Gel Retardation Technique, also known as the ....., is used to study ..... or ..... interactions.
2. When a protein binds to a DNA or RNA fragment, the complex formed usually has a different ..... compared to the free nucleic acid because the bound nucleic acid-protein complex is ..... and has a different .....
3. In the ....., the labeled nucleic acid is mixed with the ..... in a reaction buffer that supports ....., which may include salts, proteins, and other stabilizing components.
4. During ....., the reaction mixture is loaded onto a gel and run under ....., meaning that the proteins and nucleic acids are not ..... during the process.
5. After electrophoresis, the gel is treated to detect the labeled nucleic acids. If using ....., the gel can be exposed to a ..... For non-radioactive labels, the gel might be exposed to a ..... or ..... substrate.

10.4 DNA Foot printing Technique

DNA foot printing is a technique used to identify the specific regions of DNA that are bound by proteins, such as transcription factors, regulatory proteins, or other DNA-binding molecules (**figure 10.6**). It helps to map out protein-DNA interactions by showing which parts of the DNA are protected from cleavage or degradation by the bound proteins. Here's an overview of how DNA footprinting works and its applications:

Principle:

The DNA footprinting technique is based on the principle that proteins binding to DNA will protect the DNA from being cut by specific nucleases or other chemical agents. By comparing the pattern of DNA cleavage with and without the bound protein, researchers can identify the regions of DNA that are protected and thus bound by the protein.

Procedure:

1. **Labeling DNA:**

- The DNA fragment of interest is labeled at one end, typically using radioactive or fluorescent markers.

2. **Binding Reaction:**

- The labeled DNA is incubated with the protein of interest. This allows the protein to bind to its specific DNA sequence.

3. **Cleavage or Digestion:**

- The DNA-protein complex is then treated with a nuclease (e.g., DNase I) or chemical reagents that cleave the DNA. If the protein is bound to the DNA, the regions of DNA bound by the protein will be protected from cleavage.

4. **Termination of Reaction:**

- After the cleavage or digestion, the reaction is stopped, and the DNA is purified to remove any remaining protein and other reaction components.

5. **Separation and Analysis:**

- The DNA is separated by gel electrophoresis to analyze the cleavage pattern. The resulting DNA fragments are visualized by autoradiography if radioactive labels were used or by other detection methods if non-radioactive labels were used.

6. **Interpretation:**

- The presence of protected regions indicates where the protein was bound. In the gel, these protected regions appear as "footprints" where the normal pattern of DNA cleavage is missing.

Types of DNA Footprinting:

1. **Chemical Footprinting:**

- Uses chemical reagents like hydroxylamine or dimethyl sulfate to modify DNA bases. The modification protects the DNA from further chemical cleavage, and the footprint is visualized by analyzing the pattern of cleavage.

2. **Enzymatic Footprinting:**

- Uses nucleases such as DNase I to cleave the DNA. The footprint is detected by the absence of cleavage in regions where the protein is bound.

Applications:

- **Mapping DNA-Binding Sites:** Identify and map the exact regions of DNA where proteins bind.
- **Characterizing Transcription Factors:** Study how transcription factors interact with promoter regions or other regulatory elements.
- **Understanding Gene Regulation:** Investigate the mechanisms by which proteins regulate gene expression by binding to specific DNA sequences.

DNA footprinting is a valuable technique for studying molecular interactions and understanding the regulatory mechanisms that control gene expression and cellular functions.

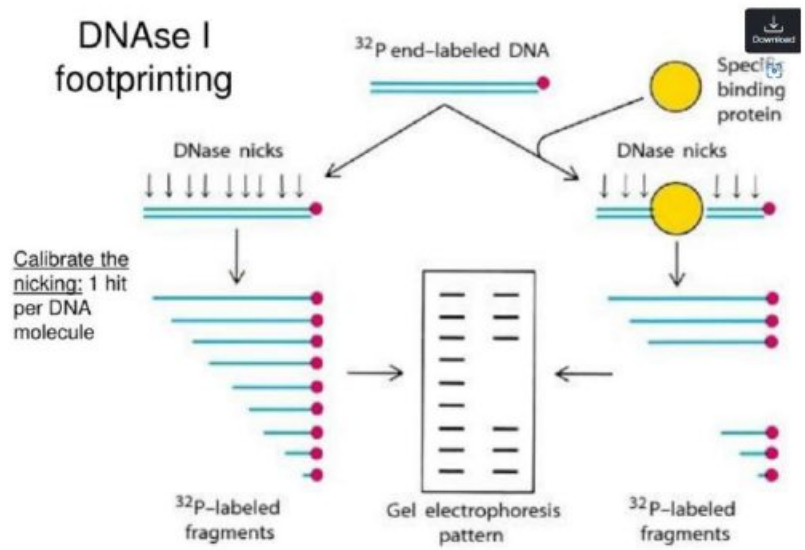


Fig.10.6: DNA Footprinting Technique(https://www.slideserve.com/odetta/enzymes-for-manipulating-dna)

SAQ 3

---

1. ..... is a technique used to identify the specific regions of DNA that are bound by **proteins**, such as transcription factors or regulatory proteins.

2. The principle of DNA footprinting is based on the idea that proteins binding to DNA will ..... the DNA from being cut by specific ..... or chemical agents, revealing which parts of the DNA are protected.

3. During the ..... or digestion step, the DNA-protein complex is treated with a nuclease (e.g., DNase I) or chemical reagents that ..... the DNA. Regions bound by the protein will be ..... from cleavage, showing up as "footprints" in the resulting pattern.

10.5. DNA Fingerprinting

DNA fingerprinting, also known as DNA profiling, is a technique used to identify individuals based on their unique DNA patterns (**figure 10.7**). It's widely used in forensic science, paternity testing, and genetic research. Here's an overview of the DNA fingerprinting process and its applications:

Principle:

DNA fingerprinting relies on the fact that each individual has a unique DNA sequence (except for identical twins). By analyzing specific regions of the DNA that vary between individuals, a unique DNA profile can be created for each person.

Key Steps in DNA Fingerprinting:

1. **Sample Collection:**

   - Collect DNA samples from biological sources such as blood, saliva, hair, or other tissues.

2. **DNA Extraction:**

   - Isolate DNA from the collected samples using chemical or mechanical methods. This step involves breaking open cells and separating DNA from proteins and other cellular components.

3. **DNA Quantification:**

   - Measure the amount of DNA extracted to ensure there is enough for the fingerprinting process.

4. **DNA Amplification (PCR):**

- Use Polymerase Chain Reaction (PCR) to amplify specific regions of the DNA. PCR targets regions known for variability between individuals, such as short tandem repeats (STRs) or variable number tandem repeats (VNTRs).

5. **DNA Fragment Separation:**

- Separate the amplified DNA fragments using gel electrophoresis or capillary electrophoresis. This process sorts DNA fragments based on their size.

6. **Visualization and Analysis:**

- Stain the DNA fragments with a dye or use a fluorescent marker. Visualize the separated fragments using techniques like autoradiography (for radioactive labels) or fluorescence detection.

7. **Pattern Comparison:**

- Compare the DNA profiles from the sample with known profiles or reference samples. In forensic cases, this involves matching the profile from a crime scene with potential suspects. In paternity tests, it involves comparing the profiles of the child, mother, and alleged father.

Applications:

1. **Forensic Science:**

- **Crime Scene Investigation:** Identify suspects or victims based on DNA left at a crime scene.

- **Missing Persons:** Help in identifying missing persons by comparing their DNA with available samples.

- **Paternity Testing:**

- **Establish Parentage:** Confirm biological relationships by comparing the DNA profiles of the child and the alleged parents.

2. **Genetic Research:**

- **Genetic Mapping:** Study genetic variation and associations with diseases or traits.

- **Population Genetics:**Analyze genetic diversity within and between populations.

3. **Identification of Human Remains:**

- **Disaster Victim Identification:** Identify victims of natural disasters, accidents, or other mass casualty events.

Types of DNA Markers:

- **Short Tandem Repeats (STRs):** Short, repetitive sequences of DNA. STR profiling is commonly used in forensic DNA fingerprinting.

- **Variable Number Tandem Repeats (VNTRs):** Longer repetitive sequences that vary in number between individuals.

Advantages:

- **High Specificity:** Each individual (except for identical twins) has a unique DNA profile.

- **Reliable:** Provides accurate results when protocols are followed correctly.

- **Versatile:** Can be used for a variety of applications, from forensic investigations to ancestry research.

Limitations:

- **Quality of Sample:** The quality and quantity of the DNA sample can affect the accuracy of the results. Degraded or contaminated samples may lead to inconclusive results.

- **Privacy Concerns:** DNA profiling raises ethical and privacy issues regarding the storage and use of genetic information.

DNA fingerprinting has revolutionized forensic science and many other fields by providing a powerful tool for individual identification and genetic analysis.

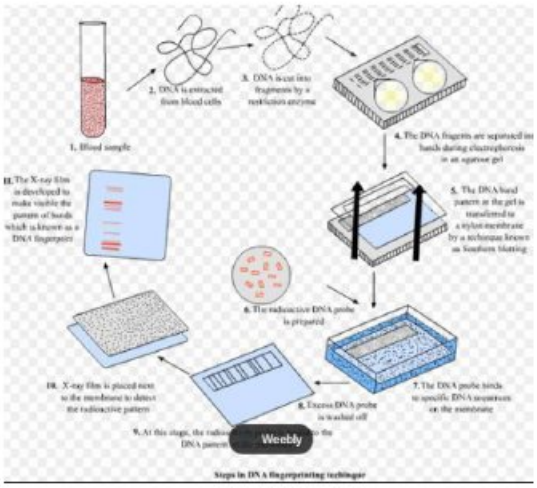


Fig.10.7: DNA fingerprinting Technique(https://allaboutdnafingerprinting.weebly.com/steps-of-dna-fingerprinting.html)

SAQ4

1. DNA fingerprinting, also known as ....., is a technique used to identify individuals based on their unique ....., which vary between individuals (except for identical twins).
2. In the DNA fingerprinting process, ... is used to ..... specific regions of the DNA, such as ..... or ....., which are known to vary between individuals.

10.6. Primer Design

Primers are short DNA sequences used to initiate DNA synthesis in PCR (Polymerase Chain Reaction). Proper primer design is crucial for amplifying specific gene regions, enabling detailed analysis of gene function and mutation.

Designing primers is a crucial step in many molecular biology techniques, including Polymerase Chain Reaction (PCR), qPCR, and sequencing. Primers are short, single-stranded DNA sequences that initiate the DNA synthesis process. Here's a guide to designing effective primers:

1. Primer Design Principles:

a. Length:

- **Optimal Length:** Typically, 18-25 nucleotides. Longer primers (25-30 nucleotides) can increase specificity but may also lead to secondary structures.

b. Melting Temperature (Tm):

- **Optimal Tm:** Aim for a Tm between 55-65°C. Primers with a Tm that is too high or too low can affect the efficiency of the reaction.
- **Tm Calculation:** Use the formula Tm=2(A+T)+4(G+C)T_m = 2(A + T) + 4(G + C)Tm=2(A+T)+4(G+C) for a rough estimate. Many online tools can calculate Tm more accurately, taking into account salt concentration and other factors.

c. GC Content:

- **Optimal Range:** 40-60% GC content. A higher GC content can increase the Tm and improve the stability of the primer-template binding.

d. Specificity:

- **Avoiding Homology:** Ensure primers do not have significant homology to each other (to avoid primer-dimer formation) or to other regions of the genome (to avoid non-specific amplification).
- **Avoid Repeats:** Avoid long stretches of a single nucleotide or repeated sequences, which can lead to secondary structures or non-specific binding.

e. Primer Pairing:

- **Complementarity:** Ensure that the forward and reverse primers are not complementary to each other, as this can lead to primer-dimer formation.

f. 3' End Stability:

- **Specific Binding:** Ensure the 3' end of the primer is stable and specific to the target sequence to promote accurate extension.

2. <u>Practical Considerations</u>:

a. Design Tools:

- **Online Tools:** Use software or online tools such as Primer3, NCBI Primer-BLAST, or commercial software like OligoCalc to assist in primer design and check for potential issues.

b. Amplicon Size:

- **Target Size:** For PCR, design primers to amplify a specific range of amplicon sizes, usually between 100-1000 base pairs. Ensure the size of the target amplicon is appropriate for your application.

c. Avoid Secondary Structures:

- **Hairpins and Dimers:** Check primers for potential secondary structures like hairpins or primer-dimers using tools like the Primer3 web interface or other in silico tools.

d. Include Controls:

- **Positive Controls:** Design primers to include known sequences as positive controls to validate the performance of your PCR or other assays.
- **Negative Controls:** Design primers for negative controls to ensure no non-specific amplification occurs.

3. <u>Steps for Designing Primers</u>:

a. Define the Target Sequence:

- Identify the region of DNA you want to amplify or sequence.

b. Use a Primer Design Tool:

- Input the target sequence into a primer design tool. Specify parameters such as primer length, Tm, and desired amplicon size.

c. Review Primer Suggestions:

- Evaluate the suggested primers for potential issues, such as self-complementarity or off-target binding.

d. Test and Optimize:

- Synthesize the primers and test them in PCR or other assays. Optimize conditions if needed, including adjusting the annealing temperature or buffer conditions.

e. Validate Results:

- Confirm that the primers work as expected by analyzing the PCR products or sequencing results.

Example Workflow:

1. **Target Sequence:** 5'-AGCTGAGCTGACT-3'
2. **Forward Primer Design:** Choose a sequence starting 18-25 bases from the 5' end.
3. **Reverse Primer Design:** Select a complementary sequence on the opposite strand, ensuring it is 18-25 bases long.
4. **Tm Calculation:** Check the melting temperature of each primer and ensure they are compatible.
5. **Test in PCR:** Use the designed primers in PCR, adjust conditions, and analyze the amplicon for specificity and yield.

Effective primer design is critical for the success of molecular biology experiments. By following these guidelines, you can enhance the accuracy, efficiency, and reliability of your assays.

10.7. Primer Extension (PCR)

Primer extension is a key step in Polymerase Chain Reaction (PCR) and related techniques, where DNA polymerase synthesizes new DNA strands by extending from primers. This process amplifies specific DNA sequences, enabling detailed analysis and study. Here's a breakdown of the primer extension process within PCR:

1. Overview of Primer Extension in PCR

Primer Extension: The step where DNA polymerase extends the primers to synthesize new DNA strands complementary to the target sequence.

2. PCR Process:

PCR involves a series of temperature cycles, each with specific steps:

1. Denaturation:
   - Temperature: Typically, 94-98°C.

- Purpose: Denature the double-stranded DNA into single strands by breaking the hydrogen bonds between base pairs.

2. Annealing:

- Temperature: Typically, 50-65°C, depending on the primers' Tm.
- Purpose: Allow primers to bind (anneal) to their complementary sequences on the single-stranded DNA. The temperature is lower than the denaturation temperature to enable primer binding.

3. Extension (Elongation):

- Temperature: Typically, 68-75°C, depending on the DNA polymerase used.
- Purpose: DNA polymerase extends the primers by adding nucleotides to the 3' end of the primer, synthesizing new DNA strands complementary to the template strand.

3. Detailed Steps for Primer Extension:

a. Primer Binding:

Binding: Primers are short sequences that bind specifically to the complementary sequences flanking the target DNA region. This binding occurs during the annealing phase.

b. DNA Polymerase Action:

- Extension: DNA polymerase extends the primer by adding nucleotides to the growing DNA strand. The enzyme synthesizes DNA in the 5' to 3' direction, using the original DNA strand as a template.
- Types of DNA Polymerases: Commonly used enzymes include Taq polymerase, which is thermostable and can withstand the high temperatures used during PCR.

c. Reaction Conditions:

- Buffer: The PCR reaction mix includes a buffer to maintain optimal pH and ionic conditions.
- dNTPs: Deoxynucleotide triphosphates (dATP, dCTP, dGTP, dTTP) are the building blocks for DNA synthesis.
- Magnesium Ions: Magnesium chloride (MgCl2) is often included to enhance the activity of DNA polymerase.

d. Amplification:

- Cycles: The extension step is repeated in each cycle of PCR, resulting in exponential amplification of the target DNA region.

4. Optimization Tips:

a. Primer Design:

- Specificity: Ensure primers are specific to the target sequence to minimize non-specific amplification.
- Melting Temperature (Tm): Primers should have similar Tm values for optimal annealing.

b. Enzyme Selection:

- Thermostability: Choose a DNA polymerase that can withstand the high temperatures used in PCR. Taq polymerase is commonly used due to its thermostability.

c. Reaction Conditions:

- Optimize Concentrations: Adjust concentrations of primers, DNA template, dNTPs, and MgCl2 to improve yield and specificity.

d. Cycle Conditions:

- Temperature and Time: Optimize annealing temperature and extension time based on the length of the amplicon and the DNA polymerase used.

6. Applications of Primer Extension in PCR**:**

a. Gene Amplification: Amplify specific genes or DNA regions for further analysis or cloning. b. **Genetic Analysis**: Study genetic variations, mutations, or gene expression. c. Diagnostic Testing: Detect specific DNA sequences associated with diseases or pathogens. d. Sequencing: Prepare DNA samples for sequencing by amplifying target regions. Primer extension is central to PCR and similar techniques, enabling the specific amplification of DNA sequences for a wide range of applications in research, diagnostics, and forensic science.

SAQ 5

1. In PCR, ... is the step where ..... synthesizes new DNA strands by adding nucleotides to the 3' end of the primers, extending them to create complementary DNA strands.

2. The ..... step in PCR involves heating the reaction to temperatures typically between ..... to separate the double-stranded DNA into single strands by breaking the hydrogen bonds between base pairs.

3. During the ..... phase of PCR, the temperature is typically set between ....., which allows the primers to bind specifically to their complementary sequences on the single-stranded DNA.

4. ..... is a commonly used DNA polymerase in PCR due to its ....., enabling it to withstand the high temperatures used during the PCR process.

10.8 S1 Mapping

S1 Mapping is a technique used to study RNA-DNA interactions and to determine the transcriptional start sites and boundaries of gene expression (**figure 10.8**). It is a lab technique used to the 5' end of an RNA transcript within a mixture by using the S1 nuclease. This method is used to quantify RNAs, to map the positions of introns and to identify the locations of 5' and 3' ends of mRNAs on cloned DNA templates. Here's an overview of the S1 Mapping process and its applications:

Principle

S1 Mapping is based on the principle that S1 nuclease can degrade single-stranded DNA or RNA, but not double-stranded DNA. When RNA hybridizes with its complementary DNA, the regions of RNA that are unprotected (not hybridized) are susceptible to degradation by S1 nuclease, whereas the protected regions remain intact.

Procedure

Preparation of RNA and DNA Templates:

- RNA: Obtain the RNA of interest, often extracted from cells or tissues.
- DNA Template: Use a DNA fragment that includes the target region where the RNA might bind. This DNA is usually labeled with a radioactive or fluorescent marker.

Hybridization:

- Mixing: Incubate the labeled DNA fragment with the RNA sample to allow the RNA and DNA to hybridize.
- Conditions: Ensure the reaction conditions (such as temperature and buffer) support stable RNA-DNA hybrid formation.

S1 Nuclease Treatment:

- Enzyme Addition: Add S1 nuclease to the reaction mixture. This enzyme specifically degrades single-stranded nucleic acids.
- Digestion: The enzyme will degrade the single-stranded, unprotected regions of the RNA or DNA, leaving behind the RNA-DNA hybrids intact.

Separation and Analysis:

- Stopping the Reaction: Terminate the nuclease reaction by adding a stop solution.
- Purification: Purify the remaining nucleic acids to remove any enzymes and contaminants.
- Separation: Separate the products using gel electrophoresis to distinguish between the degraded and intact nucleic acids.
- Detection: Use autoradiography, fluorescence, or other detection methods to visualize the results.

Interpretation:

Mapping: Analyze the gel to determine the positions where the RNA was protected by the DNA. This information helps identify the regions of RNA binding and can provide insights into the transcriptional start sites or boundaries of gene expression.

Applications

- Mapping Transcription Start Sites: Identify where transcription begins for specific genes, helping to map promoter regions and regulatory elements.
- Determining RNA-DNA Interactions:Understand how RNA binds to specific regions of DNA, which can be useful for studying regulatory mechanisms and gene expression.

- **Characterizing RNA Structure: Investigate the secondary structure of RNA molecules by analyzing protected regions and mapping interactions with DNA.**
- **Studying Mutations and Variants:Assess how mutations in DNA affect RNA binding and gene expression, which can be important for understanding genetic diseases.**

Advantages

Specificity: Allows precise mapping of RNA-DNA interactions.

Resolution: Provides detailed information on transcription start sites and RNA binding sites.

Versatility: Can be applied to various types of RNA and DNA sequences.

Limitations

Sensitivity: Requires high-quality RNA and DNA samples to obtain reliable results.

Complexity: The technique can be complex and requires careful optimization of conditions.

S1 Mapping is a valuable tool in molecular biology for understanding gene regulation, RNA-DNA interactions, and the structural features of RNA and DNA.

Fig. 10.8: S1 Mappingtechnique (https://www.nationaldiagnostics.com/2011/08/19/s1-mapping/)

SAQ 6

---

1. **Hybridization**: Incubate the labeled DNA fragment with the RNA sample to allow the RNA and DNA to ……. This ensures that RNA and DNA form complementary RNA-DNA hybrids.

2. **S1 Nuclease Treatment:** Add ..... to the reaction mixture. This enzyme specifically degrades single-stranded nucleic acids, such as the unprotected regions of RNA or DNA.

10.9. Dot-blot Analysis

Dot-blot analysis is a simple and efficient technique used to detect and quantify specific molecules, such as proteins, nucleic acids, or small molecules, on a solid support membrane. It is similar to Western blotting and Northern blotting but involves spotting samples directly onto the membrane without separating the components by gel electrophoresis (**figure 10.9**). Here's an overview of how dot-blot analysis works and its applications:

1. Principle:

In dot-blot analysis, samples containing the target molecules are spotted directly onto a solid support membrane (usually nitrocellulose or PVDF). The membrane is then probed with specific antibodies or nucleic acid probes that bind to the target molecules. The presence of the target is detected using various methods, such as chemiluminescence, fluorescence, or colorimetric detection.

2. Procedure:

a. Sample Preparation:

- **Preparation:** Prepare samples containing the target molecules. For proteins, this might involve cell lysates, purified proteins, or tissue extracts. For nucleic acids, it could be DNA or RNA samples.

b. Spotting:

- **Membrane:** Place the membrane (nitrocellulose or PVDF) onto a blotting apparatus or support.
- **Spotting Samples:** Using a pipette or spotting device, deposit small volumes of the sample onto the membrane in discrete spots. Allow the spots to dry.

c. Blocking:

- **Blocking Solution:** Incubate the membrane with a blocking solution (e.g., BSA, non-fat milk) to prevent non-specific binding of antibodies or probes to the membrane.

d. Probing:

- **Primary Antibody/Probe:** Incubate the membrane with a primary antibody (for proteins) or nucleic acid probe (for DNA/RNA) that specifically binds to the target molecule.
- **Washing:** Wash the membrane to remove unbound antibodies or probes.

e. Detection:

- **Secondary Antibody/Probe:** Incubate with a secondary antibody (if using primary antibodies) conjugated to a detectable marker (e.g., enzyme, fluorophore).
- **Detection:** Visualize the bound secondary antibody or probe using the appropriate detection method:
- **Colorimetric:** Uses substrates that produce a color change (e.g., BCIP/NBT for alkaline phosphatase).
- **Chemiluminescent:** Uses substrates that emit light upon reaction (e.g., ECL for horseradish peroxidase).
- **Fluorescent:** Uses fluorescent dyes or tags (e.g., fluorescence microscopy or imaging).

f. Analysis:

- **Quantification:** Compare the intensity of the spots to a standard or control to quantify the amount of target molecule.

3. Applications:

a. Protein Detection:

- **Quantification:** Measure the levels of specific proteins in various samples.
- **Validation:** Confirm the expression of recombinant proteins or the presence of proteins in different conditions.

b. Nucleic Acid Detection:

- **Gene Expression:** Detect specific DNA or RNA sequences.
- **Mutation Analysis:** Identify genetic variations or mutations.

c. Small Molecule Detection:

- **Screening:** Detect and quantify small molecules or metabolites in biological samples.

d. Cross-Contamination Checks:

- **Quality Control:** Verify the presence of specific proteins or nucleic acids in samples, ensuring no cross-contamination.

4. Advantages:

- **Simplicity:** Requires less equipment and preparation compared to gel-based methods.
- **Speed:** Faster than Western or Northern blotting because it omits electrophoresis.
- **Quantitative:** Allows for semi-quantitative analysis of target molecules.

5. Limitations:

- **Resolution:** Cannot resolve multiple target molecules in a complex mixture as effectively as gel-based methods.
- **Specificity:** Requires highly specific antibodies or probes to ensure accurate detection.

Dot-blot analysis is a versatile and straightforward technique for detecting and quantifying specific molecules in a wide range of applications, from basic research to clinical diagnostics.

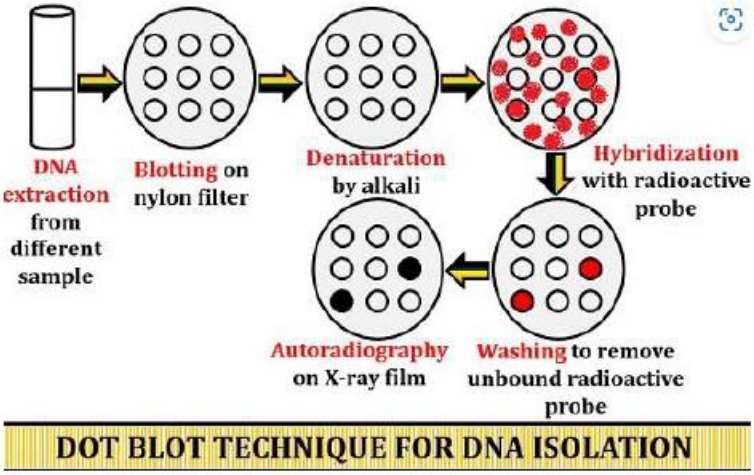


Figure 10.9. Dot Blot technique https://biologyreader.com/dot-blot-technique.html

SAQ 7

1. In dot-blot analysis, the target molecules are spotted directly onto a solid support membrane, which is usually made of _____ or _____.
2. The detection of the target molecules in dot-blot analysis can be achieved using methods such as _____, _____, or _____.
3. During the dot-blot procedure, the membrane is first incubated with a _____ solution to prevent non-specific binding of antibodies or probes.
4. In dot-blot analysis, after spotting the samples on the membrane, they are allowed to _____ before further processing.
5. For protein detection, dot-blot analysis involves probing with _____ that specifically binds to the target protein.
6. The secondary _____ or _____ used in dot-blot analysis is conjugated to a detectable marker to visualize the presence of the target molecule.
7. To quantify the amount of target molecule in dot-blot analysis, the intensity of the spots is compared to a _____ or _____.
8. One advantage of dot-blot analysis is its _____ compared to gel-based methods, making it faster and simpler.
9. A limitation of dot-blot analysis is its _____ in resolving multiple target molecules in a complex mixture compared to gel-based methods.
10. In the dot-blot procedure, the membrane is purified to remove any remaining _____ and contaminants after S1 nuclease treatment.

10.10 Summary

The analysis of gene structure and function involves a suite of techniques designed to unravel the complexities of genetic information. Understanding gene structure—comprising exons, introns, and regulatory elements—alongside the processes of transcription and translation, is crucial for advancing our knowledge of molecular biology and genetics. Techniques such as gel electrophoresis, DNA footprinting, and PCR play pivotal roles in exploring these aspects, leading to significant discoveries in genetics, medicine, and biotechnology. Each method offers unique insights into how genes are organized, expressed, and regulated, contributing to a comprehensive understanding of genetic function and its impact on organisms.

10.11. Suggested Readings

1. "Molecular Biology of the Gene" by James D. Watson, Tania A. Baker, Stephen P. Bell, Alexander Gann, Michael Levine, and Richard Losick
2. "Molecular Cell Biology" by Harvey Lodish, Arnold Berk, Chris A. Kaiser, and Monty Krieger
3. "Genetics: A Conceptual Approach" by Benjamin A. Pierce
4. "Principles of Gene Manipulation and Genomics" by S. B. Primrose and R. Twyman
5. "Molecular Cloning: A Laboratory Manual" by Michael R. Green and Joseph Sambrook

10.12. Terminal Questions

Q.1. Define the primary components of a gene and describe the function of each component in gene expression.

Answer: ———

Q.2. . Explain the role of promoters and enhancers in the regulation of gene transcription?

Answer: ———

Q.3. Discuss the significance of exons and introns in gene structure.

Answer: _____

_____

_____

Q.4. Describe the process of transcription initiation.

Answer: _____

_____

_____

Q.5. Compare and contrast gel electrophoresis and DNA footprinting techniques.

Answer: _____

_____

_____

Q.6. What is DNA fingerprinting?

Answer: _____

_____

_____

Q.7. What is S1 mapping, and how is it used to study gene expression?

Answer: _____

_____

_____

Answers to SAQs (Self-assessment Questions)

SAQ1

1. **Exons** are the coding regions of a gene that are expressed and translated into proteins, while **introns** are non-coding regions interspersed between exons.

2. During RNA splicing, **introns** are removed from the precursor mRNA (pre-mRNA), leaving only the **exons** in the mature mRNA.
3. The **promoter** is a regulatory sequence located upstream of the coding region that serves as the binding site for RNA polymerase and various **transcription factors**.
4. **Enhancers** are distal regulatory elements that increase the likelihood of transcription by interacting with the **promoter** region through **DNA looping** mechanisms.
5. **Silencers** are sequences that repress gene transcription by binding **transcriptional repressors**or other regulatory proteins that inhibit the transcription machinery's access to the gene.
6. **Insulators** are sequences that function to block or insulate the effects of **enhancers** or **silencers** on adjacent genes, creating boundaries between different regulatory domains.
7. The **5' UTR** of mRNA contains regulatory sequences that control mRNA **stability**, localization, and **translation efficiency**.
8. **Zinc Finger Domains** contain cysteine and histidine residues that coordinate zinc ions and bind to specific DNA sequences by inserting loops of the protein into the major groove of the DNA. An example is **Kruppel-like factors**.
9. **Acidic Activation Domains** are rich in acidic residues and interact with components of the transcriptional machinery to promote the **initiation** of transcription. An example is **Activator protein 1 (AP-1).**
10. During **translation**, the ribosome assembles at the start codon of the mRNA, and the first **tRNA** molecule binds to the start codon, bringing the first amino acid.

SAQ2

1. The Gel Retardation Technique, also known as **the Electrophoretic Mobility Shift Assay (EMSA)**, is used to study **protein-DNA** or **protein-RNA**interactions.
2. When a protein binds to a DNA or RNA fragment, the complex formed usually has a different **electrophoretic mobility** compared to the free nucleic acid because the bound nucleic acid-protein complex is **larger** and has a different **charge distribution**.
3. In the binding reaction, the labeled nucleic acid is mixed with the **protein**(s) in a reaction buffer that supports **binding**, which may include salts, proteins, and other stabilizing components.
4. During **gel electrophoresis**, the reaction mixture is loaded onto a gel and run under **non-denaturing conditions**, meaning that the proteins and nucleic acids are not **denatured** during the process.

5. After electrophoresis, the gel is treated to detect the labeled nucleic acids. If using **radioactive labels**, the gel can be exposed to a **film**. For non-radioactive labels, the gel might be exposed to a **chemiluminescent** or **fluorescent** substrate.

SAQ 3

1. **DNA footprinting** is a technique used to identify the specific regions of DNA that are bound by **proteins**, such as transcription factors or regulatory proteins.
2. The principle of DNA footprinting is based on the idea that proteins binding to DNA will **protect** the DNA from being cut by specific **nucleases** or chemical agents, revealing which parts of the DNA are protected.
3. During the **cleavage** or digestion step, the DNA-protein complex is treated with a nuclease (e.g., DNase I) or chemical reagents that **cleave** the DNA. Regions bound by the protein will be **protected** from cleavage, showing up as "footprints" in the resulting pattern.

SAQ 4

1. DNA fingerprinting, also known as **DNA profiling**, is a technique used to identify individuals based on their unique **DNA patterns**, which vary between individuals (except for identical twins).
2. In the DNA fingerprinting process, **Polymerase Chain Reaction (PCR)** is used to **amplify** specific regions of the DNA, such as **short tandem repeats (STRs)** or **variable number tandem repeats (VNTRs)**, which are known to vary between individuals.

SAQ 5

1. In PCR, **primer extension** is the step where **DNA polymerase** synthesizes new DNA strands by adding nucleotides to the 3' end of the primers, extending them to create complementary DNA strands.
2. The **Denaturation** step in PCR involves heating the reaction to temperatures typically between **94-98°C** to separate the double-stranded DNA into single strands by breaking the hydrogen bonds between base pairs.
3. During the **Annealing** phase of PCR, the temperature is typically set between **50-65°C**, which allows the primers to bind specifically to their complementary sequences on the single-stranded DNA.
4. **Taq polymerase** is a commonly used DNA polymerase in PCR due to its **thermostability**, enabling it to withstand the high temperatures used during the PCR process.

SAQ 6

1. **Hybridization**: Incubate the labeled DNA fragment with the RNA sample to allow the RNA and DNA to **hybridize**. This ensures that RNA and DNA form complementary RNA-DNA hybrids.
2. **S1 Nuclease Treatment**: Add **S1 nuclease** to the reaction mixture. This enzyme specifically degrades single-stranded nucleic acids, such as the unprotected regions of RNA or DNA.

SAQ 7

1. In dot-blot analysis, the target molecules are spotted directly onto a solid support membrane, which is usually made of **nitrocellulose** or **PVDF**.
2. The detection of the target molecules in dot-blot analysis can be achieved using methods such as **chemiluminescence**, **fluorescence**, or **colorimetric**.
3. During the dot-blot procedure, the membrane is first incubated with a **blocking** solution to prevent non-specific binding of antibodies or probes.
4. In dot-blot analysis, after spotting the samples on the membrane, they are allowed to **dry** before further processing.
5. For protein detection, dot-blot analysis involves probing with **antibodies** that specifically binds to the target protein.
6. The secondary **antibody** or **probe** used in dot-blot analysis is conjugated to a detectable marker to visualize the presence of the target molecule.
7. To quantify the amount of target molecule in dot-blot analysis, the intensity of the spots is compared to a **standard** or **control**.
8. One advantage of dot-blot analysis is its **simplicity** compared to gel-based methods, making it faster and simpler.
9. A limitation of dot-blot analysis is its **resolution** in resolving multiple target molecules in a complex mixture compared to gel-based methods.
10. In the dot-blot procedure, the membrane is purified to remove any remaining **enzymes** and contaminants after S1 nuclease treatment.

UNIT-11: IMMUNO TECHNOLOGY

Structure

11.1. Introduction

Immune technology is a technology based on applications of cells and molecules of the immune system. The exquisite specificity of antigens antibody interactions has led to the development of a variety of immunological techniques. These techniques have played a vital role in diagnosing disease, monitoring the level of the humoral immune response, and identifying molecules of biological or medical interest. Advances in immunodiagnostic technologies provide the basis for developing antigen detection plateform capable of meeting stringent requirements for sensitivity, specificity, assay speed, robustness, and simplicity. Antibody based microarray is setting a novel proteomics technology setting a new standard for molecular profiling of non fractionated complex proteomics. Improvements in the affinity, specificity and mass production of antibodies will dictate the success or failure of a given immunoassay technology.

Immuno-technology encompasses a wide range of techniques that harness the unique capabilities of antibodies across medicine, diagnostics, and research. Central to this field is hybridoma technology, pioneered by Köhler and Milstein in the 1970s. Immunotechnology is an important arm of biotechnology, constituting the indusrial scale application of immunological procedures in disease

diagnosis, to produce vaccines, for mass immunization to prevent prevalent diseases and producing immunological therapeutic agent to cure the afflieted.

This method involves fusing antibody-producing B cells with immortalized myeloma cells to generate hybridomas capable of producing monoclonal antibodies with exceptional specificity for a single antigen epitope. Monoclonal antibodies, produced through hybridoma technology, play crucial roles in therapeutic and diagnostic applications by precisely targeting antigens on cancer cells, pathogens, and other specific targets. Furthermore, antibody engineering techniques such as chimerization, humanization, affinity maturation, and the creation of antibody fragments like Fab and scFv have further refined antibody properties, improving their stability, specificity, and functionality. These advancements are pivotal in the development of personalized therapies, enhancing drug delivery systems, and advancing the accuracy of diagnostic tools. In essence, immuno-technology continues to drive significant progress in healthcare and scientific research by enabling precise targeting and tailored solutions for the treatment and diagnosis of various diseases.

Immunology is an experimental science which cuts across many discipline such as molecular biology, physical chemistry, biochemistry, microbiology, pharmacology, electronices, instrumentation and others. Today immunology is a very complex and sophisticated area of biology, which has become one of the most versatile research tools in biology and medicine as well as a powerful weapon in the armory of prevention and management of several viral and bacterial diseases.

Objectives

After study this topic you will be able to know

- To understand the hybrydoma technology
- To know about monoclonal antibody
- To understand the process of antibody engineering
- To know the uses of monoclonal antibody

11.2. Hybridoma technology

Hybridoma technology is a well established method for producing monoclonal antibodies (mAbs) that target specific antigens. This technique involves fusing short lived antibody-producing B cells with immortal myeloma cells, resulting in hybridoma cell lines that continuously produce large

quantities of a single type of mAb. These favored hybridoma lines can be cryopreserved for long-term mAb production, ensuring a steady supply.

This techeniquie was discovered in 1975 by Georges Kohler and Cesar Milstein and earned them the Nobel Prize in Physiology or Medicine in 1984. Hybridoma technology revolutionized biotechnology by enabling the creation of immortal hybrid cells capable of producing mAbs specific to desired antigens. By cloning individual hybrid cells, the first hybridoma cell lines were established, laying the foundation for diverse applications in research and medicine. Hybridomas are widely preferred for mAb production due to their ability to consistently generate highly pure, sensitive, and specific antibodies. This method supports various fields including toxicology, animal biotechnology, medicine, pharmacology, and molecular biology. Monoclonal antibodies derived from hybridoma technology are invaluable in diagnostic, imaging, and therapeutic applications, including the targeted detection of tumor antigens and specific therapeutic actions against cancer cells using radiolabeled mAbs.

Sino Biological utilizes hybridoma technology to offer a cost-effective service for generating mouse monoclonal antibodies, delivering purified antibodies within 60 days. Hybridoma technology enables the mass production of monoclonal antibodies, identical antibodies derived from a single clone of cells. The process begins with immunizing a mammal, typically a mouse, with a specific antigen to provoke an immune response. B cells, which produce antibodies targeting the antigen, are isolated from the spleen of the immunized animal.

These B cells are fused with immortal myeloma cancer cells, chosen for their ability to replicate indefinitely. This fusion creates hybrid cells called hybridomas, possessing the antibody-producing capability of B cells and the replicative longevity of myeloma cells. The fusion process typically employs techniques like electrofusion or chemical protocols using agents such as polyethylene glycol. After fusion, the hybridoma cells are cultured in a medium that selects for fused cells and inhibits the growth of unfused myeloma cells, ensuring only hybridomas survive. This medium often includes HAT (hypoxanthine-aminopterin-thymidine), which selectively kills unfused myeloma cells lacking the enzyme HGPRT.

The surviving hybridomas are then screened to identify those producing antibodies with the desired specificity. Screening methods include ELISA (enzyme-linked immunosorbent assay),

immunocytochemistry, western blotting, immunoprecipitation-mass spectrometry, and flow cytometry. This ensures selection of hybridomas that produce monoclonal antibodies specific to the antigen.

This technology offers numerous advantages, namely (1) Precise antigen targeting (2) A never ending supply of consistent antibodies (3) High sensitivity and specificity for use in biological assays (4) Elimination of the need for animal models (in vitro method) (5) Utilization in therapeutic and diagnostic treatments, vaccine creation, and chemothearapy.

Nevertheless, the technology also has few limitations. (1).Long production time (2).Resource intensive and expensive workflow (3).Nit suitable for generating short peptides and fragments antigens (4).Susceptibility to contamination and poor cell viability (5).Risk of virus contamination and disease transmission (6).Absence of stable myeloma cells for human antibody production.

Selected hybridoma cells are cloned to produce large quantities of identical daughter cells, each capable of secreting the specific monoclonal antibody. These cells are cultured in nutrient-rich media supplemented with factors like interleukin-6 to support growth and antibody production. The final step involves cloning desired hybridoma cells to obtain a stable cell population and growing the culture to collect large amounts of monoclonal antibodies. This can be achieved through one of the two methods. (1) In vitro growth of hybrid cells in tissue culture. (2) In vivo growth following inoculation of hybridoma cells into a mouse"s abdomen.

Ultimately, the monoclonal antibodies harvested from these hybridomas can be used for various applications in research, diagnostics, and therapy due to their uniformity and specificity. This technology, pioneered by César Milstein and Georges J. F. Köhler in 1975, revolutionized biomedical research and earned them the Nobel Prize in Physiology or Medicine in 1984.

Owing to their high specificity, the antibodies produced by hybrydoma technology have a wide range of diagnostic applications including the following-

1.Enzyme linked immunosorbent (ELISA), HIV antibodies, hepatitis B surface antigen, and pregnency harmone.

2. Immuno fluorescence assay (IFA); detecting autoimmune disorders, influenza virus, and chlamydia trachomatis.

3.Western blot; analyzing cancer biomarkers.

4. Flow cytometry; assessing immune cells in HIV, Leukemia, and Lymphoma.

5. Immunohistochemistry (IHC); analyzing cancer biomarkers.

6. Rapid antigen tests; detecting malaria. dengue. zika virus and COVID-19.

11.3. Monoclonal Antibodies

Antibodies naturally present in our blood helps in fighting infections. It is also known as immunoglobulin. Antibodies are Y shaped in structure.These are glycoproteins.These are generated by plasma cells. Monoclonal antibody (MAB) therapies, created in labs, replicate these natural antibodies. Each MAB therapy consists of numerous copies of a single type of antibody.

Structure of antibody

1. **Two heavy chains and two light chains**; these chains are linked by disulfide bonds.
2. **Variable Region**: The tips of the Y shape, known as the Fab (fragment antigen-binding) region, are specific to each antibody and can bind to a unique antigen.
3. **Constant Region**: The stem of the Y, or Fc (fragment crystallizable) region, interacts with other immune cells and systems, such as those that help destroy or clear the pathogen.

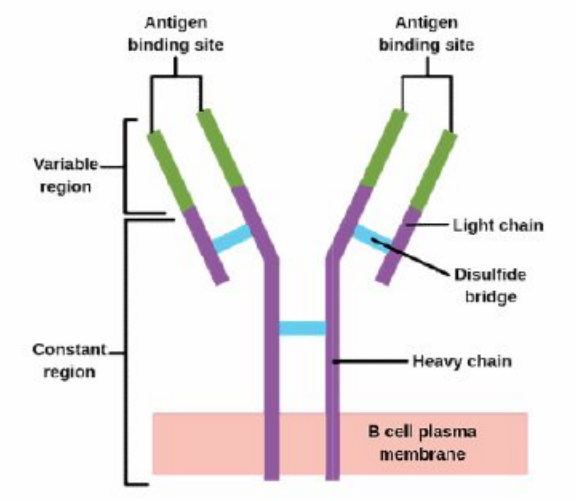


Fig.11.1; Antibody; Source; https://labster-image-manager.s3.amazonaws.com/v2/ABL/73fdc78e-cbf6-4021-bd7d-2c183935c306/ABL_

Types of Antibodies

1. **IgG**: The most common antibody in blood and extracellular fluid, crucial for long-term protection and can cross the placenta to protect the fetus.
2. **IgM**: The first antibody produced in response to an infection; it is primarily found in blood and lymphatic fluid.
3. **IgA**: Found mainly in mucosal areas such as the gut, respiratory tract, and urogenital tract, as well as in saliva, tears, and breast milk.
4. **IgE**: Involved in allergic reactions and defense against parasitic infections; binds to allergens and triggers histamine release.
5. **IgD**: Mostly found on the surface of immature B-lymphocytes and in the respiratory tract; its precise function is less well understood but is thought to play a role in initiating immune responses.

Function of antibody

1. **Binding to Antigens**: Antibodies have regions that specifically bind to antigens, marking them for destruction or neutralization.
2. **Neutralization**: By binding to pathogens, antibodies can prevent them from entering or damaging cells.
3. **Opsonization**: Antibodies can tag pathogens for destruction by phagocytes.
4. **Activation of Complement**: The antibody-antigen complex can activate the complement system, leading to the destruction of pathogens.

Antibodies are key players in both the innate and adaptive immune responses and are fundamental to the effectiveness of vaccines and immunotherapies.

Monoclonal antibodies (moAbs or mAbs) are synthetic proteins that copy the immune system's natural antibodies. Numerous MABs are used to treat cancer, each functioning uniquely, with some having multiple modes of action. These proteins are exactly designed to bind only to specific targets in the body, such as cancer cell antigens. Each cancer cells posess a specific antigen called

tumour marker, on their memberanes. There are various types of monoclonal antibodies, each tailored to bind to a distinct antigen. They are instrumental in diagnosing and treating a wide range of diseases, including specific cancers. Monoclonal antibodies (mAbs) are specific to a particular antigen, allowing them to target tumors without harming other cells. Monoclonal antibodies can function independently or serve as carriers for delivering drugs, toxins, or radioactive agents directly to cancerous cells.

Fig;11.2 ;Monoclonal antibody; Source; https://assets.technologynetworks.com/production/dynamic/images/content/376377/efficient-removal-of-aggregates-from-monoclonal-antibodies

Types of monoclonal antibody

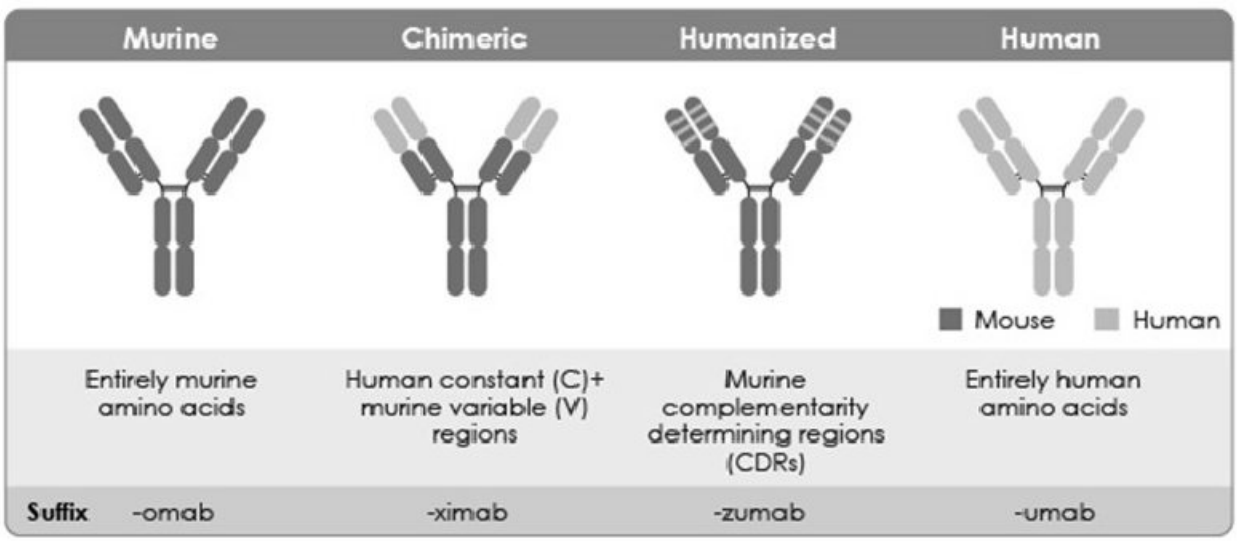
Naked monoclonal antibodies (mAbs) are antibodies that operate independently without any drug or radioactive material attached. They are extensively employed in cancer treatment, primarily binding to antigens present on cancer cells. However, they can also target antigens on non-cancerous cells or free-floating proteins. Naked mAbs achieve their therapeutic effects through diverse mechanisms.

Conjugated monoclonal antibodies (mAbs) are attached to chemotherapy drugs or radioactive particles to act as targeted vehicles for delivering these substances directly to cancer cells. They navigate through the bloodstream until they locate and bind to the specific target antigen. This

targeted approach ensures precise delivery of the toxic substance to the intended site. Conjugated mAbs are alternatively referred to as tagged, labeled, or loaded antibodies.

Monoclonal antibodies are classified into four types based on their composition:

1. **Murine antibodies**: Derived from mouse proteins, treatments using these antibodies are named with the suffix "-omab".
2. **Chimeric antibodies**: Composed of a combination of mouse and human proteins, treatments using these antibodies are named with the suffix "-ximab".
3. **Humanized antibodies**: Made from small portions of mouse proteins fused with human proteins, treatments using these antibodies are named with the suffix "-zumab".
4. **Human antibodies**: Comprising entirely of human proteins, treatments using these antibodies are named with the suffix "-umab".



Fig;11.3 Source; https://www.caltagmedsystems.co.uk/information/wp-content/uploads/image-62.png

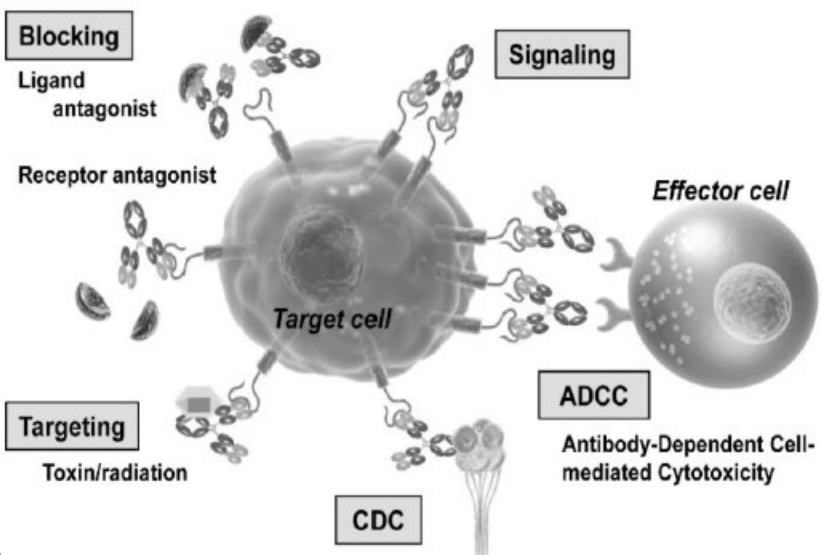
Work Mechanism of monoclonal antibody

Monoclonal antibody drugs operate through various mechanisms designed to target and treat diseases. Here's how they work:

Target Recognition: Monoclonal antibodies coat cancer cells, aiding immune system cells in identifying and attacking them effectively.

1. **Cell-Membrane Destruction:** Some antibodies trigger immune responses that dismantle the outer membrane of cancer cells.
2. **Inhibition of Growth Signals:** Certain antibodies block connections between cancer cells and growth-promoting proteins crucial for their survival and proliferation.
3. **Angiogenesis Inhibition:** By blocking interactions necessary for new blood vessel formation, these drugs starve tumors of their blood supply, inhiboit their growth.
4. **Immune Activation:** Monoclonal antibodies disrupt proteins that regulate immune cell activity, allowing immune responses to target cancer cells more effectively.
5. **Direct Cell Destruction:** Specific antibodies attach to cancer cells, initiating processes within the cell that lead to self-destruction.
6. **Radiation Delivery:** Engineered antibodies can carry radioactive particles directly to cancer cells, minimizing radiation exposure to healthy tissues.
7. **Chemotherapy Delivery:** Similarly, some antibodies transport chemotherapy drugs specifically to cancer cells, sparing healthy cells.
8. **Cellular Interactions:** Combining antibodies that bind to both cancer and immune cells can enhance immune system attacks against cancer cells.

So, monoclonal antibodies are versatile tools that leverage various mechanisms to target and treat diseases, offering precise and often targeted therapies while minimizing damage to healthy Tissues.



Fig;11.4 Work Mechanism of monoclonal antibody; Source;https://d3i71xaburhd42.cloudfront.net/21880a941de52e62f98cfd413e62df14c9824ed4/3-

Production of monoclonal antibody

The production of monoclonal antibodies (mAbs) involves a series of well-defined steps that ensure the generation of highly specific antibodies against a particular antigen. Here's a detailed overview of the process:

Immunization

↓

Isolation of Spleen Cells

↓

Fusion with Myeloma Cells

↓

Selection and Screening

↓

Single Cell Cloning

↓

Cell Culture and Expansion

↓

Harvesting

↓

Purification

↓

Characterization

↓

Formulation

↓

Packaging

↓

Distribution

Each step is determining in ensuring the production of high-quality monoclonal antibodies, used in a wide range of applications from research to therapeutic treatments.

1. Antigen Selection

- **Identification:** The antigen of interest is identified. This could be a protein, peptide, carbohydrate, or other molecule.

2. Immunization

- **Animal Selection:** Typically, mice are used due to their robust immune response.
- **Immunogen Preparation:** The antigen is purified and prepared for injection into the chosen animal.

3. Fusion of B Cells with Myeloma Cells

- **Isolation of B Cells:** Spleen cells containing B cells specific to the antigen are isolated from the immunized animal.
- **Fusion:** B cells are fused with myeloma cells (cancerous B cells that are immortal and can divide indefinitely) to create hybridoma cells.

4. Screening and Cloning

- **Screening:** Hybridoma cells are screened to identify clones that produce antibodies specific to the antigen.
- **Cloning:** Individual hybridoma cells are isolated and cultured to ensure monoclonality (production of antibodies by a single clone).

5. Expansion and Production

- **Culture:** Monoclonal hybridoma cells are cultured in bioreactors or flasks containing suitable growth medium.
- **Antibody Production:** Antibodies are secreted into the culture medium.

6. Antibody Harvesting and Purification

- **Harvesting:** The culture medium is collected to recover antibodies.
- **Purification:** Antibodies are purified using techniques such as chromatography to remove cell debris, proteins, and other contaminants.

7. Characterization

- **Quality Control:** Purified antibodies undergo diligent characterization to confirm specificity, affinity, and functionality.
- **Testing:** Various tests are conducted to ensure the antibodies meet desired quality standards.

8. Formulation and Packaging

- **Formulation:** Antibodies are formulated into a suitable buffer or solution for stability and functionality.
- **Packaging:** Packaged into vials, syringes, or other containers depending on their intended use (research, diagnostics, therapeutics).

9. Storage and Distribution

- **Storage:** Antibodies are stored under appropriate conditions (e.g., temperature-controlled environments) to maintain stability.
- **Distribution:** Distributed to end-users such as laboratories, clinics, or pharmaceutical companies.

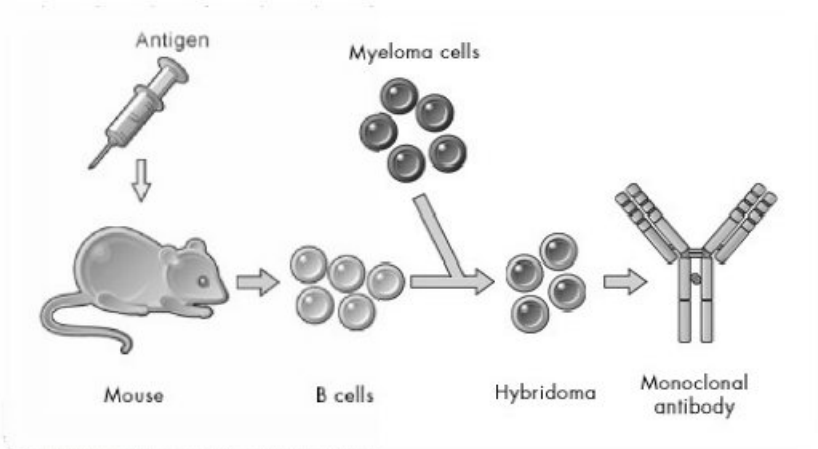


Fig 11.5 Production of monoclonal antiboidy; Source, https://abgenex.com/images/content/production_illust.jpg

Examples of monoclonal antibody

Certainly! Monoclonal antibodies (mAbs) are designed to target specific antigens with high precision, making them valuable tools in medicine, research, and diagnostics. Here are some examples of monoclonal antibodies along with their targets and applications:

Therapeutic Monoclonal Antibodies

1. **Trastuzumab (Herceptin):**
   - **Target**: HER2 (human epidermal growth factor receptor 2).
   - **Use**: Treatment of HER2-positive breast cancer.
2. **Rituximab (Rituxan):**
   - **Target**: CD20 antigen on B cells.
   - **Use**: Treatment of non-Hodgkin lymphoma, chronic lymphocytic leukemia, and certain autoimmune diseases.
3. **Adalimumab (Humira):**
   - **Target**: Tumor necrosis factor-alpha (TNF-alpha).
   - **Use**: Treatment of autoimmune diseases such as rheumatoid arthritis, psoriasis, and inflammatory bowel disease.
4. **Pembrolizumab (Keytruda):**
   - **Target**: PD-1 (programmed cell death protein 1).
   - **Use**: Immunotherapy for various cancers, including melanoma, non-small cell lung cancer, and head and neck squamous cell carcinoma.
5. **Bevacizumab (Avastin):**
   - **Target**: Vascular endothelial growth factor (VEGF).
   - **Use**: Treatment of various cancers, including colorectal cancer, lung cancer, and glioblastoma.

Diagnostic Monoclonal Antibodies

1. **Anti-Human Chorionic Gonadotropin (hCG) Antibodies:**

- **Use**: Detection of pregnancy in home pregnancy tests and clinical laboratories.

2. **Anti-CD3 Antibodies**:
   - **Use**: Diagnosis of T-cell lymphomas and leukemias.

Research and Laboratory Use

1. **Anti-FLAG Antibodies**:
   - **Use**: Detection and purification of proteins tagged with the FLAG epitope in biochemical and molecular biology research.
2. **Anti-HA Antibodies**:
   - **Use**: Detection and purification of proteins tagged with the hemagglutinin (HA) epitope.
3. **Anti-GFP Antibodies**:
   - **Use**: Detection and localization of proteins fused with green fluorescent protein (GFP) in cellular and molecular biology research.

These examples illustrate the diverse applications of monoclonal antibodies in targeting specific molecules for therapeutic, diagnostic, and experimental purposes. Each monoclonal antibody is carefully engineered to recognize a particular antigen, allowing for highly targeted treatments and precise diagnostic tests.

SAQ

1. Hybridoma technology has been successfully used in .....(synthesis of monoclonal antibody, synthesis of somatic cells).

2. Myeloma cells are type of.....cells.

3 .Antibodies are .....shaped.

4. Hybridoma cells are formed by fusion of .....and .....cells.

11.4. Antibody engineering

Antibody engineering consists of modifying monoclonal antibody (mAb) sequences and/ or structures to either enhance or dampen their functions. Monoclonal Abs have revolutionized the field of diagnosis and immunotherapy for the treatment of a variety of diseases, particularly in cancer therapy.

Antibodies are vital proteins in the immune system, crucial for detecting and neutralizing threats such as bacteria and viruses by recognizing specific protein patterns in the body. They play essential roles in clinical diagnosis and therapy. The FDA approved the first engineered antibody for marketing in 1997, employing phage display to select antigen-specific binders from transgenic animals like mice and rabbits.

Antibody engineering modifies monoclonal antibody (mAb) structures to enhance or diminish their functions, significantly advancing immunotherapy and diagnostics. These innovations are pivotal in preventing cancer and treating a range of viral and microbial diseases. Unlike polyclonal antibodies, which vary in concentration and see less clinical use, engineering efforts primarily target monoclonal antibodies due to their precise targeting capabilities in treatment.

A challenging issue remains the production of therapeutic in Abs and Ab derived drugs with the highest objectives response rate in patients and the lowest toxicity. Therefore Ab engineering is a major translational research topic which causes at producing highly specifie and effective mAbs, with optimal processing, stability and tolerance.

Antibody engineering represents an advanced discipline dedicated to integrating technology and techniques for combating diseases caused by microorganisms and viruses. It focuses on manipulating molecular chains, aiming to optimize the sequence of molecules for enhanced therapeutic outcomes. In recent years, antibody engineering has emerged as a leading method for selecting, designing, and producing molecules, pivotal not only in fundamental research but also in pharmaceutical applications. Its potential impact is particularly promising in the treatment of cancers and related diseases. Key discoveries in antibody research continue to drive advancements in engineering techniques, fostering the development of novel therapeutic strategies.

Techniques in Antibody Engineering

Antibody engineering includes the introduction of the antibody combining site (variable regions) into a host of a architectures including bi and multi specific formats that further impact the theraopeutic properties leading to further advantages and successes in patient treatment. Antibody engineering is a complex field that involves various techniques to modify and enhance antibodies for specific applications. Here are some key techniques used in antibody engineering:

1. Recombinant DNA Technology

- **Gene Cloning:** Involves cloning the DNA encoding the antibody into expression vectors to produce antibodies in various systems (e.g., bacteria, yeast, mammalian cells).
- **Chimeric Antibodies:** Combine variable regions from one species with constant regions from another species to create antibodies with desired properties (e.g., humanized chimeric antibodies).

2. Humanization

- **CDR Grafting:** Transfer the complementarity-determining regions (CDRs) of a non-human antibody to a human antibody framework to reduce immunogenicity while retaining antigen specificity.
- **Framework Optimization:** Modify the antibody framework regions to improve stability and reduce immunogenicity while maintaining binding affinity.

3. Phage Display

- **Library Construction:** Create large libraries of antibody fragments displayed on the surface of bacteriophages (viruses that infect bacteria).
- **Selection (Panning):** Expose the phage library to the target antigen to isolate and identify phages displaying antibodies with high affinity for the antigen.

4. Yeast Display

- **Antibody Display:** Display antibody fragments on the surface of yeast cells.
- **Screening and Selection:** Use yeast display to screen and select antibodies with desirable properties through iterative rounds of binding and selection.

5. Transgenic Animals

- **Production of Human Antibodies:** Use genetically modified animals (e.g., mice) that have human antibody genes to produce fully human antibodies.
- **Humanized Mice:** Use mice with a humanized immune system to generate antibodies with human-like properties.

6. Hybridoma Technology

- **Monoclonal Antibody Production:** Fuse spleen cells from immunized animals with myeloma cells to create hybridomas that produce specific monoclonal antibodies.
- **Screening:** Identify and isolate hybridomas that produce antibodies with the desired specificity and affinity.

7. Single-Cell Sequencing

- **Antibody Gene Sequencing:** Sequence the antibody genes from individual B cells to identify and produce antibodies with unique specificities and properties.
- **High-Throughput Screening:** Use sequencing data to rapidly produce and screen large numbers of antibodies.

8. CRISPR/Cas9 Gene Editing

The CRIPR/Cas9 adits genes by precisely cutting DNA and then harnessing natural DNA repair process to modify the gene in the desired manner. The system has two components , the Cas9 enzyme and a guide RNA. Rapidaly translating a revolutionary technology with the goals of developing transformative therapies. Cas9 is a crisper associated (Cas) endonuclease or enzyme that acts as molecular scissors to cut DNA at a location specified by a guide RNA.

- **Genome Editing:** Use CRISPR/Cas9 technology to edit antibody genes directly in cell lines or animals to improve antibody properties or create custom antibodies.
- **Gene Knockout/Knockin:** Create knockout or knockin models to study the effects of specific genes on antibody production and function.

9. Antibody Engineering Libraries

Immune antibody libraries are created starting with lymphocyte RNA from indivisuals previously exposed to the desired antigen. Such libraries are biased towards high affinity binders for the desired antigens source they contain heavy and light chains that have undergone in vivo affinity interaction.

- **Synthetic Libraries:** Create synthetic libraries of antibody variants to explore a wide range of antibody sequences and identify those with desired properties.
- **Diverse Libraries:** Use various methods to create diverse antibody libraries for broad screening and discovery.

11.5. Site-Directed Mutagenesis

It is an invaluable tool to modify genes and study the structural and functional properties of a protein, based on the structure, function, catalegtic mechanism, and catalytic residues of enzymes. Site directed mutagenesis includes single and combinational mutations. It is usually analyzed by bioinformatic methods. Single site directed mutagenesis and multiple mutations have been used to expected and simplify methods for mutagenesis. The properties of enzymes can be improved markedly by the combination of site directed mutagenesis with other methods.

- **Antibody Modification:** Introduce specific mutations into the antibody genes to enhance properties such as affinity, stability, or specificity.
- **Functional Analysis:** Analyze the impact of mutations on antibody function and characteristics.

11.6. Bispecific Antibodies

A bispecific monoclonal antibody (BsMAb, BsAb) is an artificial protein that can simultaneously bind to two different types of antigen or two different epitopes on the same antigen. Naturally occuring antibidies typically only target one antigen. BsAbs can be manufactured in several stuctural formats. BsAbs can be designed to recreat and activate immune cells, to interfere with receptor signaling and inactivate signaling legands, and to force association of protein complexes. BsAbs have been explored for cancer immunotherapy , drug delivery and alzheimers disease.

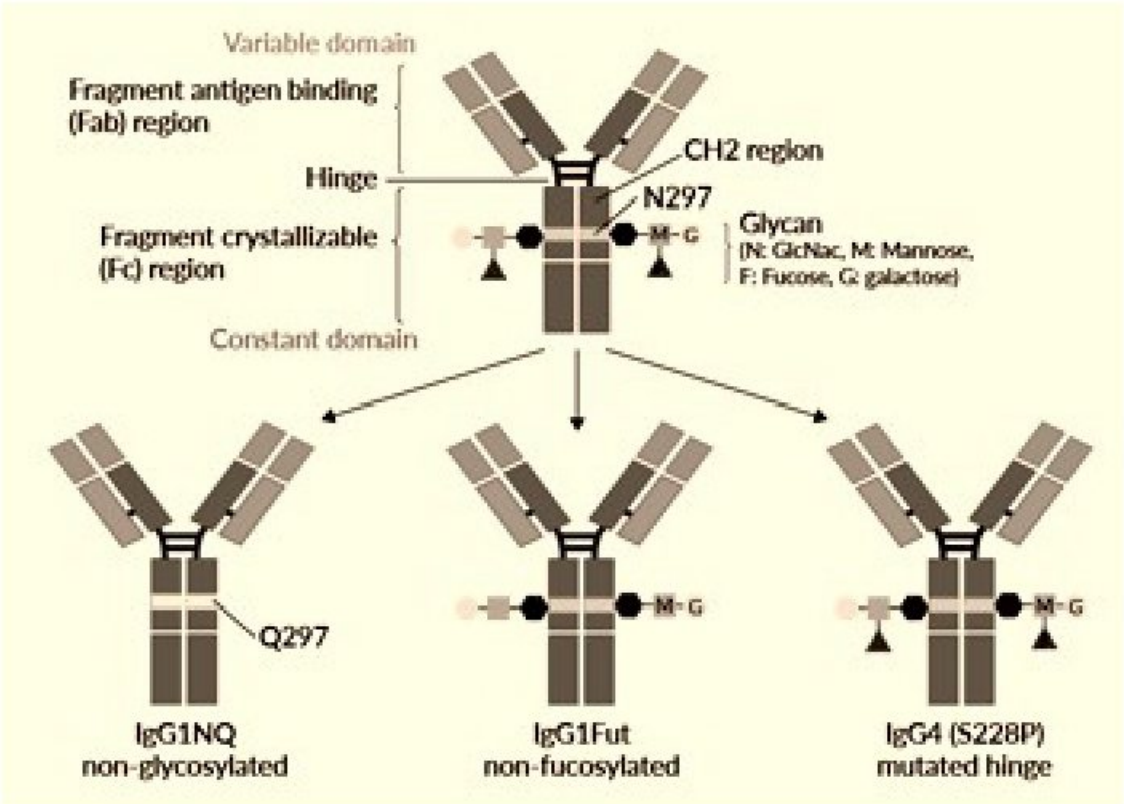
- **Dual Targeting:** Engineer antibodies to bind to two different antigens simultaneously, enhancing therapeutic efficacy or targeting multiple pathways.

- **Crosslinking:** Use different formats (e.g., single-chain variable fragments, diabodies) to achieve bispecificity.

11.7. Antibody-Drug Conjugates (ADCs)

In recent years, antibody drug conjugales (ADCs) may serve as a novel therapeuctic modality for many cancer patients. Indeed three key elements define an ADc, the antibody directed against a specific tumor antigen, the cytotoxin payload and the cleavable/uncleavable linker connecting the payload to the antibody specific target recognition and effective toxin makes ADCs veritable magic bullet. These magic bullets deliver treatment that is both highly targeted and highly potent seeking out and attacking cancerous cells throughout the body while to avoiding healthy cells reasonable selection of target antibody, linked payload and their rational combinations, as well as conjugation methods are crucial for an effective treatment.

- **Conjugation:** Attach cytotoxic drugs or other therapeutic agents to antibodies to deliver targeted therapy directly to cancer cells or other diseased cells.
- **Linker Chemistry:** Develop and optimize linkers that connect the drug to the antibody to ensure stability and effective delivery.



Fig;11.6 Antibody engineering

;Saurce;https://www.invivogen.com/sites/default/files/pictures/antibody-isotypes-strategies-small.png

Need of antibody engineering

Antibody engineering is crucial in medicine and biotechnology because it enhances several key characteristics of antibodies:

1. **Affinity and Specificity:** By improving the strength of binding to target antigens and ensuring specificity, engineered antibodies can precisely target disease markers or therapeutic targets.
2. **Stability:** Engineering antibodies enhances their ability to withstand degradation in various environments, ensuring they remain effective and functional over longer periods.
3. **Functionality:** Modification of effector functions such as antibody-dependent cellular cytotoxicity and complement activation can be customized to enhance therapeutic efficacy or diagnostic applications.

4. **Half-life:** Prolonging the circulation time of antibodies in the body improves their therapeutic potential by extending the duration of action and reducing the frequency of dosing.
5. **Production:** Optimizing production methods, particularly through recombinant DNA technology in cell lines, enables efficient and scalable production of engineered antibodies for widespread clinical and research use.

Applications

Engineered antibodies play a critical role across multiple domains, from therapeutics to diagnostics and research. In therapeutics, they are instrumental in targeting and treating diseases like cancer, autoimmune disorders, and infectious diseases. Engineered antibodies can precisely recognize disease markers or pathogens, modulate immune responses, and deliver therapeutic payloads directly to affected cells, enhancing treatment efficacy and reducing side effects. A biomarker may be used to see how well the body responds to a treatment for a disease condition.

In diagnostics, engineered antibodies are indispensable for detecting specific biomarkers or pathogens in various diagnostic tests. They serve as vital components in immunoassays, point-of-care tests, and molecular diagnostics, enabling accurate disease detection, monitoring disease progression, and guiding treatment decisions.

Furthermore, in research, engineered antibodies are essential tools for advancing biological and medical knowledge. They enable scientists to study and manipulate target molecules with high specificity, facilitating applications such as imaging cellular structures, identifying and purifying proteins, and elucidating complex biochemical pathways. Their versatility and ability to bind selectively to molecular targets make them invaluable in developing new therapeutic strategies and understanding disease mechanisms.

Overall, engineered antibodies represent a keystone of modern healthcare and biomedical research, driving innovation, improving diagnostic accuracy, and paving the way for more effective treatments and discoveries in medicine. Eg.Trastuzumab, known as Herceptin, is a monoclonal antibody employed in breast cancer therapy, specifically targeting the HER2/neu receptor. It plays a crucial role in inhibiting the overexpression of HER2/neu, a characteristic found in certain breast cancers. Rituximab targets the CD20 antigen present on B-cells, used predominantly in treating non-Hodgkin lymphoma and other B-cell malignancies by selectively eliminating these cancerous cells.

Adalimumab, an anti-TNF-alpha antibody, is pivotal in managing autoimmune diseases such as rheumatoid arthritis and Crohn's disease, reducing inflammation by binding to TNF-alpha and modulating immune responses effectively. These antibodies exemplify targeted therapies that have revolutionized treatment approaches, offering improved outcomes and quality of life for patients across various medical conditions.

SAQ

1. What is the primary purpose of antibody engineering?

A) To increase the natural diversity of antibodies

B) To create antibodies with enhanced specificity and affinity for target antigens

C) To produce antibodies in large quantities

D) To reduce the cost of antibody production

2. Which technology is commonly used to generate monoclonal antibodies?

A) Phage display

B) ELISA

C) PCR

D) Western blotting

3. In the context of antibody engineering, what does "humanization" refer to?

A) The process of modifying antibodies to have a human-like structure

B) Increasing the antigen-binding affinity of antibodies

C) Altering the antibody to recognize new antigens

D) Converting antibodies into small molecules

4. Which region of an antibody is primarily responsible for antigen binding?

A) Fc region

B) Constant region

C) Variable region

D) Hinge region

5. What is a common application of chimeric antibodies in clinical settings?

A) Diagnostic imaging

B) Gene therapy

C) Cancer treatment

D) Vaccine development

6. In the context of antibody engineering, what does "Fab" stand for?

A) Fragment antigen-binding

B) Fully active binding

C) Fragment antigen-binding

D) Functional antigen-binding

6. Which method is used to generate a diverse library of antibodies for screening?

A) Hybridoma technology

B) Phage display

C) Yeast display

D) All of the above

11.8. Uses of monoclonal antibody

Monoclonal antibodies (mAbs) are highly versatile tools known for their precise binding capabilities, finding application across various fields. Here are their primary uses:

Therapeutic Applications:

- **Cancer Treatment:** Targeting specific antigens on cancer cells to deliver drugs, toxins, or radioactive substances directly (e.g., Trastuzumab for HER2-positive breast cancer).
- **Autoimmune Diseases:** Modulating immune responses by targeting inflammatory molecules in conditions like rheumatoid arthritis and Crohn's disease (e.g., Infliximab for TNF-alpha).
- **Infectious Diseases:** Neutralizing or blocking pathogens such as viruses (e.g., COVID-19 treatments), bacteria, and fungi.
- **Eye Disorders:** Inhibiting abnormal blood vessel growth in conditions like age-related macular degeneration (e.g., Ranibizumab targeting VEGF).
- **Other Uses:** Treating asthma, preventing transplant rejection, and more.

Diagnostic Uses:

- **Medical Imaging:** Utilizing monoclonal antibodies labeled with radioactive tracers for precise imaging in procedures like PET scans (e.g., Prostate-specific membrane antigen for prostate cancer).
- **Laboratory Tests:** Crucial in diagnostics such as ELISA, immunohistochemistry, and flow cytometry to detect and quantify proteins or cells in patient samples.

Research Applications:

- **Basic Research:** Essential for studying protein functions and distributions within cells and tissues.
- **Drug Development:** Serving as models for developing new therapies and gaining insights into disease mechanisms.

Other Applications:

- **Toxin Delivery:** Delivering therapeutic payloads directly to target cells by linking monoclonal antibodies with toxins or drugs, minimizing collateral damage.
- **Vaccine Production:** Playing a critical role in the development and quality control of vaccines, particularly for viral vaccines.

In short, monoclonal antibodies play significant roles in clinical practice, diagnostics, and research, driving advancements in personalized medicine and disease management.

Side effects of monoclonal antibody

Monoclonal antibody drugs commonly cause the following side effects:

- **Allergic Reactions:** Such as hives or itching.
- **Flu-like Symptoms:** Including chills, fatigue, fever, and muscle aches.
- **Gastrointestinal Disturbances:** Such as nausea, vomiting, and diarrhea.
- **Skin Reactions:** Such as rashes.
- **Low Blood Pressure.**

Some other serious Side Effects:

Although rare, serious side effects of monoclonal antibody therapy may include:

- **Infusion Reactions:** Severe allergic-like reactions that can rarely be lethal. Patients may receive medications to prevent allergic reactions before starting treatment. Infusion reactions typically occur during or shortly after treatment, necessitating close monitoring by healthcare providers. Patients may need to remain at the treatment facility for observation for several hours.
- **Cardiovascular Issues:** Some monoclonal antibodies may increase the risk of high blood pressure, congestive heart failure, or heart attacks.
- **Pulmonary Complications:** Certain antibodies are associated with a higher risk of inflammatory lung disease.

- **Skin Complications:** Sores and rashes that can lead to serious infections. In severe cases, sores may develop on the mucosal lining of the cheeks and gums.
- **Bleeding:** Some monoclonal antibody drugs carry a risk of internal bleeding.

In summary, while monoclonal antibodies are effective treatments for various conditions, patients and healthcare providers should be vigilant about potential side effects, both common and serious. Early recognition and management of these side effects are pivotal for ensuring safe and effective treatment.

11.9. Summary

Immuno-technology continues to drive significant progress in healthcare and scientific research by enabling precise targeting of disease markers and facilitating the development of innovative therapeutic and diagnostic solutions. Its ongoing advancements promise to further enhance medical treatments and diagnostic capabilities, ultimately improving patient outcomes and advancing our understanding of diseases. Immuno-technology, a field leveraging the unique properties of antibodies, has seen significant advancements due to hybridoma technology, monoclonal antibodies, and antibody engineering.

Hybridoma Technology: Developed by Köhler and Milstein in the 1975s, this technique fuses antibody-producing B cells with immortal myeloma cells to create hybridomas that produce monoclonal antibodies (mAbs). These mAbs are highly specific to a single antigen epitope, allowing for precise targeting in therapeutic and diagnostic applications. Hybridoma technology has revolutionized biotechnology by providing a reliable method for large-scale production of mAbs, which are used in various fields including medicine, research, and diagnostics.

Monoclonal Antibodies: These are identical antibodies derived from a single clone of cells and are crucial in both research and therapy. Structurally, antibodies consist of two heavy chains and two light chains linked by disulfide bonds, with variable regions specific to antigens and constant regions interacting with other immune components. Types of monoclonal antibodies include naked, conjugated, and engineered variants, each tailored to specific therapeutic or diagnostic purposes. For instance, Trastuzumab targets HER2 in breast cancer, while Adalimumab is used for autoimmune diseases like rheumatoid arthritis.

Antibody Engineering: This field involves modifying antibody structures to enhance their therapeutic and diagnostic functions. Techniques include recombinant DNA technology, humanization, phage display, and CRISPR/Cas9 gene editing. These methods improve antibody affinity, stability, and functionality, leading to better-targeted therapies and diagnostic tools. Applications of engineered antibodies are extensive, ranging from cancer treatment and autoimmune disease management to diagnostic assays and research.

Uses of Monoclonal Antibodies: They are versatile in therapeutic applications (e.g., cancer treatment, autoimmune diseases, and infectious diseases), diagnostic uses (e.g., medical imaging and laboratory tests), and research (e.g., studying protein functions and drug development). Engineered antibodies, in particular, play a pivotal role in personalized medicine by targeting specific disease markers and improving treatment outcomes.

Side Effects: Common side effects of monoclonal antibody therapy include allergic reactions, flu-like symptoms, gastrointestinal disturbances, and skin reactions. Rare but serious side effects can involve infusion reactions, cardiovascular issues, pulmonary complications, and bleeding.

In conclusion, immuno-technology, through hybridoma technology, monoclonal antibodies, and antibody engineering, has significantly advanced the fields of medicine and research, providing precise tools for diagnosis, treatment, and therapeutic development.

11.10. Terminal Questions

Q.1 What do you understand by hybridoma techenology ?

Answe.....

.....

.....

Q.2 what are monoclonal antibody?

Answer

.....

.....

.....

Q.3 Explain the advantage and disadvantage of monoclonal antibody.

Answer.

.....

.....

.....

Q.4 What are myloma cells ?

Answer

.....

.....

.....

Q.5 Describe the production of monoclonal antibody ?

Answer

.....

.....

.....

Q.6 Why are myeloma cells used in hybridoma ?

Answer

.....

.....

.....

Q.7 What are monoclonal antibody ?

Answer

.....

.....

.....

Q.8 How monoclonal antibodies can be used to diagnose cancer ?

Answer.....

.....

.....

Q.9 What do you understand by antibody engineering ?

Answer

.....

.....

.....

Q.10 Write short notes on following

a. Antibody

b. Myeloma cells

c. Hybridoma techenology

d. Monoclonal antibody

e. Antibody engineering

11.7 Answes

SAQ

1. synthesis of monoclonal antibody

2.type of tumer cells

3. Y shaped

4. myeloma cells and B cells

5.B

6. A

7. A

8. C

9. C

10. C

11. D

11.11. Further suggested readings

1. Robert Schlei, Genetics and Molecular Biology,2nd Edition
2. McGraw-Hill, Cell and Molecular Biology, Human Genetics: Concepts and Application, 9th Edition.
3. Desmond S. T. Nicholl, An Introduction to Genetic Engineering Third Edition, University of the West of Scotland, Paisley, UK
4. Robert Schleif, Genetics and Molecular Biology, second edition, The Johns Hopkins University Baltimore, Maryland
5. NPTEL – Bio Technology – Genetic Engineering & Applications
6. T.A. Brown, Gene Cloning And Dna Analysis An Introduction, Sixth Edition, University Of Manchester Manchester.

Unit 12: Transgenic Animals

Contents

12.1 Introduction

Transgenic animals are GMOs whose genetic makeup has been altered through genetic engineering to express genes from different species. This process involves inserting foreign DNA into an animal's genome, leading to new traits or characteristics not present in the original species. A transgenic animal is one that carries a foreign gene that has been deliberately inserted into its genome.Transgenic animals provide a chance to produce animals that are a source of useful human therapeutic protein like growth hormone, insulin, human growth lactoferrin etc.transgenicanimals can be easily produced by transforming stem cells, growing in vitro, with the desired gene constructed by homologous recombination. Successfully transfected cells can be used in somatic cell nuclear transfer (SCNT) to produce transgenic animals or directly inject the gene constructed into an embryo using micromanipulation.Transgenic animals are a significant tool in biotechnology, with applications in research, medicine, agriculture, and industry. Creating transgenic animals starts with identifying and isolating a specific gene of interest. This gene is then inserted into the animal's genome using techniques like microinjection, retrovirus-mediated gene transfer, or embryonic stem cell-mediated gene transfer. Once the gene is integrated into the genome, it can be expressed and passed on to future generations. One major reason for developing transgenic animals is to advance scientific research.

These animals are valuable models for studying human diseases, enabling researchers to explore genetic disorders, develop new treatments, and test drug efficacy. For example, transgenic mice with human genes are extensively used in cancer research, providing insights into tumor growth and progression. In agriculture, transgenic animals are engineered to boost productivity and disease resistance. Examples include cows that produce nutritionally enhanced milk, pigs resistant to swine fever, and chickens less prone to avian influenza. Transgenic animals can nowadays easily be produced through CRISPR/Cas9 genome editing technique. These innovations benefit the agricultural industry and contribute to food security and animal welfare. Additionally, transgenic animals are crucial in biopharmaceutical production. They can be engineered to produce therapeutic proteins, antibodies, and other biologically active substances in their milk, blood, or eggs. This approach offers a cost-effective and scalable method for producing complex pharmaceuticals essential for treating various medical conditions. However, the creation and use of transgenic animals raise ethical, environmental, and safety concerns. Ethical discussions focus on animal welfare and potential unintended environmental impacts. Robust regulatory frameworks and rigorous testing are necessary to address these concerns and ensure the responsible development and application of transgenic technology. Transgenic animals are those in which an exogenous gene was artificially gene was artificially inserted and stably incorporated into the genome of every cell of the organisms and that can be transmitted to their descendants. The first transgenic mouse has the genome sequence of the Maloney Leukaemia virus inserted into its genome,

In brief, transgenic animals are at the forefront of genetic engineering and biotechnology, with vast potential to advance scientific knowledge, enhance agricultural practices, and transform medical treatments. As this field progresses, it is vital to balance innovation with ethical considerations and regulatory oversight.

Objectives

- To understand the definition and concept of transgenic animals
- To get an idea of creating transgenic animals
- To understand the idea of cloned animals
- To find out the differences between transgenic and cloned animals
- To explore the applications of transgenic animals

12.2 Transgenic Animals

Transgenic animals are organisms that have had foreign genes deliberately inserted into their genome. This technology is widely used in scientific research, agriculture, and medicine to study gene function, develop new treatments, and improve livestock traits. Transgenic animals represent a powerful tool in biotechnology with numerous applications across various fields. While they offer significant benefits, ethical and safety considerations must be carefully managed to ensure responsible use.

Definition

Transgenic animals are those that have been genetically modified to carry genes from other species. The foreign genes are inserted into their genomes using various biotechnological methods.

Methods of Creating Transgenic Animals

1. **Microinjection**: DNA is directly injected into the pronucleus of a fertilized egg.
2. **Retrovirus-Mediated Gene Transfer**: Retroviruses are used as vectors to transfer genetic material into the host genome.
3. **Embryonic Stem Cell-Mediated Gene Transfer**: Embryonic stem cells are modified and then introduced into early embryos.
4. **CRISPR-Cas9**: A modern and precise gene-editing technology that allows for specific modifications to the DNA sequence.

Applications

1. **Biomedical Research**:
   - **Disease Models**: Transgenic animals can be used to model human diseases, allowing researchers to study the disease process and test new treatments.
   - **Gene Function Studies**: By inserting or knocking out specific genes, scientists can learn about their roles and interactions.
2. **Agriculture**:
   - **Improved Traits**: Livestock can be genetically modified for better growth rates, disease resistance, or improved nutritional content.

- **Bioreactors**: A bioreactor is any manufactured device or system that supports a biologically active environment. A bioreactor is a vessel in which a chemical process is carried out which involves organisms or biochemically active substances derived from such organisms. Animals can produce pharmaceutical proteins in their milk, eggs, or blood.

3. **Environmental**:

- **Bioremediation**: Bioremediation broadly refers to any process wherein a biological system, living or dead, is employed for removing environmental pollutants from air, water, soil, gases, industrial affluents etc, in natural or artificial settings. Transgenic animals can be designed to help clean up environmental pollutants.
- **Conservation**: Genetic modification might help preserve endangered species by increasing genetic diversity or resistance to diseases.

Examples of Transgenic Animals

1. **Mice**: Widely used in research to study genes involved in cancer, diabetes, obesity, heart diseases, and neurological disorders.
2. **Rabbits**: Used for producing human proteins in milk for pharmaceutical purposes.
3. **Fish**: Modified for increased growth rates, such as the Aqua vantage salmon.
4. **Pigs**: Modified for improved organ transplantation compatibility with humans (xenotransplantation).

Ethical Considerations

The creation and use of transgenic animals raise several ethical issues:

1. **Animal Welfare**: Concerns about the welfare of animals that may suffer due to genetic modifications.
2. **Environmental Impact**: Potential risks of transgenic animals affecting wild populations if they escape.
3. **Human Health**: Ensuring that products derived from transgenic animals are safe for human consumption.

4. **Moral and Religious Beliefs**: Different cultures and religions have varying views on the manipulation of animal genomes.

SAQ1: Transgenic Animals

---

1. Transgenic animals are organisms that have had _____ genes deliberately inserted into their genome.
2. The foreign genes in transgenic animals can come from _____ species, and the process is known as _____ _____.
3. Methods of creating transgenic animals include microinjection, retrovirus-mediated gene transfer, embryonic stem cell-mediated gene transfer, and -.
4. Transgenic animals are used in biomedical research, agriculture, and _____.
5. An example of a transgenic animal used for producing human proteins in milk is the _____.
6. Ethical considerations for transgenic animals include concerns about animal welfare, environmental impact, human health, and _____ and _____ beliefs.

12.3 Cloned Animals

Cloned animals are organisms that are genetically identical copies of another organism. A genetic copy of an animal is produced by replacing the nucleus of an unfertilized ovum with the nucleus of a body (somatic) cell from the animal to form an embryo . They are produced through a process called cloning, which involves creating an exact genetic replica of a donor animal. This is achieved by transferring the DNA from a donor cell into an egg cell that has had its nucleus removed. The egg cell then develops into an embryo that is implanted into a surrogate mother and carried to term. The resulting animal is a clone, genetically identical to the donor animal from which the original DNA was taken.

History and Techniques

1. **Dolly the Sheep**: The first mammal to be cloned from an adult somatic cell using a technique called somatic cell nuclear transfer (SCNT). Born in 1996, Dolly marked a significant breakthrough in cloning technology.
2. **Somatic Cell Nuclear Transfer (SCNT)**: In this method, the nucleus of a somatic cell (a non-reproductive cell) is transferred to an egg cell whose nucleus has been

removed. The egg cell then develops into an embryo that is genetically identical to the donor of the somatic cell.

3. **Embryo Splitting**: This involves splitting a developing embryo at an early stage to create two or more embryos that are genetically identical. This technique is similar to the natural process that produces identical twins.

Applications

Agriculture: Cloning is used to produce animals with desirable traits, such as high milk production in cows or fast growth in livestock.

1. **Conservation**: Cloning can help preserve endangered species by creating genetic duplicates of individuals to increase population numbers.
2. **Medical Research**: Cloned animals are used to study diseases and potential treatments. For example, cloning can produce animals with specific genetic mutations to study genetic disorders.
3. **Pets**: There is a niche market for cloning pets, allowing owners to have a genetic copy of their beloved animals.

Ethical and Biological Considerations

1. **Ethics**: Cloning raises ethical concerns regarding animal welfare, the potential for reduced genetic diversity, and the implications of cloning humans.
2. **Health Issues**: Cloned animals often face health problems. Many clones do not survive gestation, and those that do often suffer from various abnormalities and shortened lifespans.
3. **Biodiversity**: Relying on cloning for agriculture can reduce genetic diversity within livestock populations, making them more vulnerable to diseases.

Examples of Cloned Animals

1. **Dolly the Sheep**: The first mammal cloned from an adult cell.Dolly the Sheep was the first mammal to be cloned from an adult somatic cell using the process of nuclear transfer. Here are some key points about Dolly:

✓ **Birth and Creation**: Dolly was born on July 5, 1996, at the Roslin Institute in Scotland. She was cloned by a team of scientists led by Ian Wilmut and Keith Campbell.

✓ **Process**: The process used to create Dolly is called somatic cell nuclear transfer (SCNT). In this process, the nucleus of an adult somatic cell (a mammary gland cell, in Dolly's case) was transferred into an enucleated egg cell (an egg cell that had its own nucleus removed). The reconstructed egg cell was then stimulated to divide and develop into an embryo, which was subsequently implanted into a surrogate mother sheep.

✓ **Significance**: Dolly's creation was a significant scientific breakthrough as it demonstrated that it was possible to reprogram a specialized adult cell to develop into an entirely new organism. This challenged previously held beliefs about cell differentiation and the irreversible nature of cell specialization.

✓ **Lifespan and Health**: Dolly lived for 6.5 years, which is relatively short for a sheep. She was euthanized on February 14, 2003, after developing a progressive lung disease and severe arthritis. There was speculation about whether cloning had caused these health issues, but no definitive conclusion was reached.

✓ **Legacy**: Dolly's creation opened up new avenues in biotechnology and medicine, including the potential for cloning in agriculture, the development of genetically modified organisms (GMOs), and advancements in regenerative medicine and therapeutic cloning.

2. **Cc (Carbon Copy) the Cat**: The first cloned pet cat.Cc (Carbon Copy or CopyCat) was the first cloned pet cat. Here are the key points about Cc:

✓ **Birth and Creation**: Cc was born on December 22, 2001, at Texas A&M University. She was cloned by a team of scientists led by Dr. Duane Kraemer and Dr. Mark Westhusin as part of the "Missyplicity Project," which aimed to clone a dog but later expanded to include other pets.

✓ **Process**: Similar to Dolly the Sheep, Cc was created using somatic cell nuclear transfer (SCNT). The nucleus of a cell from a donor cat named Rainbow was transferred into an enucleated egg cell. The reconstructed egg was then implanted into a surrogate mother cat.

✓ **Differences from the Donor**: Despite being a clone, Cc did not look exactly like her donor Rainbow. While Rainbow was a calico cat, Cc had a different coat pattern due to the way calico coloration is determined by random X-chromosome inactivation. This

highlighted that clones can have different phenotypic traits due to genetic and environmental factors.

✓ **Significance**: Cc's successful cloning demonstrated the feasibility of cloning pets, opening up possibilities for pet owners to clone their beloved animals. This also underscored the genetic complexities involved in cloning, particularly regarding traits influenced by multiple factors.

✓ **Lifespan and Health**: Cc lived a healthy life, surpassing 18 years and passing away in 2020. She gave birth to several litters of kittens, indicating that cloned animals can reproduce naturally and lead normal lives.

3. **Snuppy the Dog**: Snuppy was the first cloned dog, a significant milestone in the field of animal cloning. Here are the key points about Snuppy:

✓ **Birth and Creation**: Snuppy was born on April 24, 2005, at Seoul National University in South Korea. He was cloned by a team of scientists led by Woo Suk Hwang and Byeong Chun Lee.

✓ **Process**: Snuppy was created using somatic cell nuclear transfer (SCNT). The nucleus from a somatic cell of an adult Afghan Hound named Tai was transferred into an enucleated egg cell. The reconstructed egg was then implanted into a surrogate mother dog.

✓ **Significance**: Snuppy's creation marked the first successful cloning of a dog, a species that had proven particularly challenging to clone due to their reproductive biology. This achievement demonstrated the potential for cloning in canines, which had previously been considered very difficult.

✓ **Lifespan and Health**: Snuppy lived a relatively healthy life and died at the age of 10 on May 2015. He was noted to have been in good health for most of his life, suggesting that cloned dogs can lead normal lives similar to non-cloned dogs.

✓ **Research and Legacy**: Snuppy's cloning was a significant scientific breakthrough, paving the way for further research into canine genetics, breeding, and cloning. The techniques developed during Snuppy's cloning have been applied in various studies aimed at understanding genetic diseases and improving cloning efficiency.

4. **Cloned Cattle**: Used in agricultural practices to replicate animals with superior traits.Cloned cattle have been produced for various purposes, including agriculture, medical research, and the preservation of endangered species. Here are key points about cloned cattle:

- ✓ **First Cloned Cow**: The first cloned cow, named "Gene," was produced in 1997 by scientists at the Roslin Institute in Scotland, the same institute where Dolly the Sheep was cloned. This achievement demonstrated that cloning technology could be applied to larger mammals used in agriculture.
- ✓ **Cloning Process**: The process of cloning cattle, like other mammals, typically involves somatic cell nuclear transfer (SCNT). In SCNT, the nucleus of a somatic cell from the donor animal is transferred into an enucleated egg cell. The reconstructed egg cell is then stimulated to divide and develop into an embryo, which is implanted into a surrogate cow.

Future of Cloning

1. **Advancements**: Research continues to improve cloning efficiency and reduce associated health problems.
2. **CRISPR and Gene Editing**: Combining cloning with gene editing technologies like CRISPR may offer new possibilities for genetic research and therapy.
3. **Regulations**: The future of cloning will likely involve ongoing discussions and regulations to address ethical, environmental, and health-related concerns.

Cloning remains a controversial and rapidly developing field with the potential for significant impact on science, medicine, and agriculture.

SAQ2: Cloned Animals

---

1. Cloned animals are organisms that are genetically _____ copies of another organism.
2. Dolly the Sheep, born in 1996, was the first mammal cloned from an adult somatic cell using a technique called _____ _____ _____ _____.
3. Cloning can help preserve _____ species by creating genetic duplicates of individuals to increase population numbers.
4. Cloned animals often face health problems, including developmental abnormalities and _____ lifespans.

5. An example of a cloned pet cat is _____ (Carbon Copy) the Cat.

12.4 Transgenic Animals Vs Cloned Animals

S. No	Transgenic animals	Cloned animals
Definition	Transgenic animals are those that have had foreign genes deliberately inserted into their genome. These foreign genes can come from other species, and the process is known as genetic engineering(Figure 12.1).	Cloned animals are those that are genetically identical copies of another animal. This is achieved by copying the entire genome of an organism(Figure 12.1).
Genetic Modification	Transgenic animals have genes from other species inserted into their genome	Cloned animals are exact genetic copies of the donor animal
Purpose	Transgenic animals are mainly used for research, agricultural improvements, and pharmaceutical production	Cloned animals are used for agricultural, conservation, and medical research purposes
Creation Process	Transgenic animals are created by inserting foreign genes into their DNA	Cloned animals are created by copying the entire genome of another animal
Health Issues	Comparatively face less significant health problems than transgenic animals	Cloned animals often face more significant health problems than transgenic animals, including developmental abnormalities and shorter lifespans
Examples	GloFish: Fish that have been modified to express fluorescent proteins. EnviroPig: Pigs that produce less phosphorus in their manure. Knockout Mice: Mice with specific genes turned off for studying gene function.	Dolly the Sheep: The first mammal cloned from an adult somatic cell. Cc (Carbon Copy) the Cat: The first cloned pet cat. Snuppy the Dog: The first cloned dog

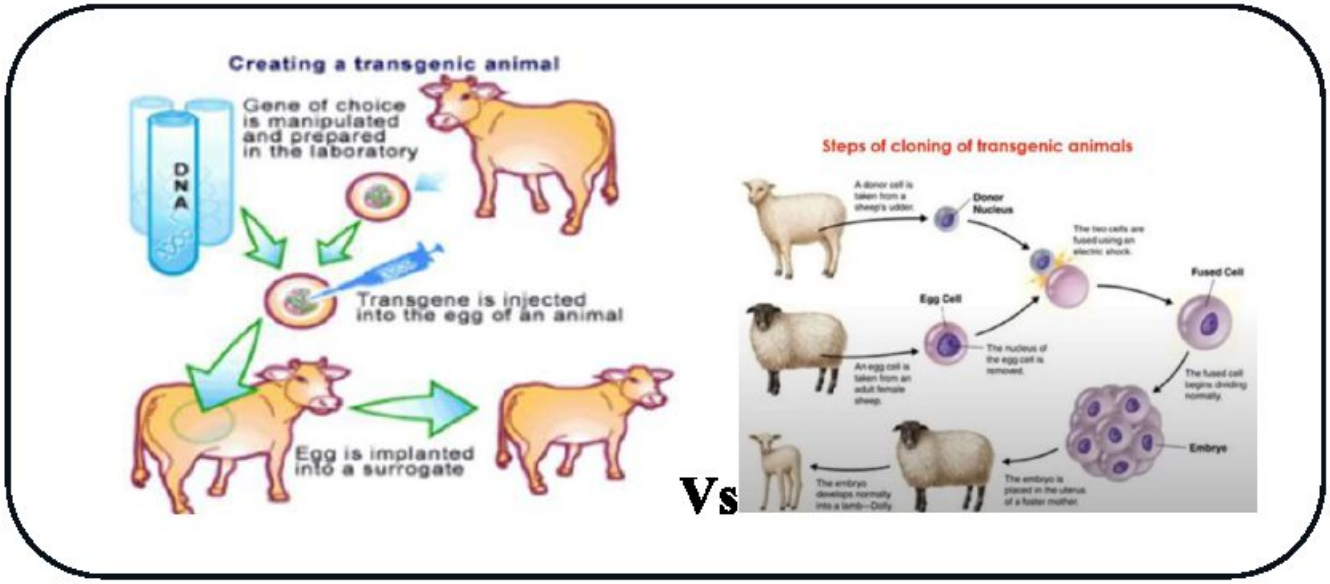


Figure 12.1 Pictorial representation of Transgenic (https://www.slideserve.com/jadyn/genetic-engineering) vs cloning of animals(https://www.youtube.com/watch?v=-KYq9sjCP6g)

SAQ3: Transgenic Animals Vs Cloned Animals

---

1. Transgenic animals are those that have had _____ genes deliberately inserted into their genome, whereas cloned animals are _____ identical copies of another animal.
2. Transgenic animals are mainly used for research, agricultural improvements, and _____ production.
3. Cloned animals often face more significant health problems than transgenic animals, including developmental _____ and shorter _____.

12.5 Producing Transgenic Animals

Producing transgenic animals involves sophisticated techniques (**Figure. 12.2a; 12.3a, b**) and careful consideration of ethical and regulatory standards, aiming to advance scientific research, improve agriculture, and develop new medical therapies.

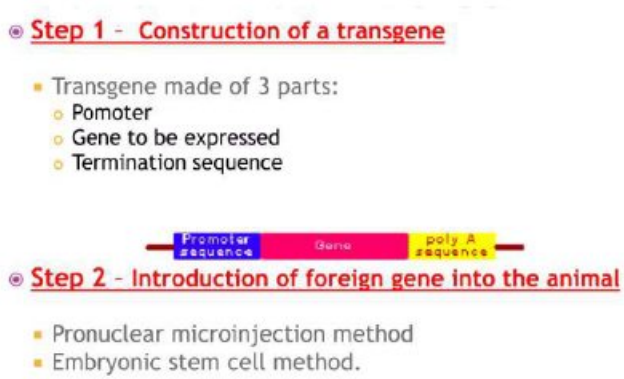


Figure. 12.2a- Methodology for producing transgenic animals (https://www.slideshare.net/slideshow/transgenic-animals-174586923/174586923#4)

a) Gene Identification and Isolation:

- **Target Gene Selection:** Identify the gene of interest that needs to be introduced into the animal's genome. This gene is usually associated with a specific trait, disease, or protein production.
- **Gene Isolation:** Isolate and clone the target gene using techniques such as polymerase chain reaction (PCR) or restriction enzyme digestion.

b) Vector Construction:

- **Vector Selection:** Choose an appropriate vector (such as a plasmid) to carry the target gene into the animal's cells. The vector often includes a promoter to ensure gene expression and a selectable marker for identifying successfully modified cells.
- **Gene Insertion:** Insert the isolated gene into the vector using molecular cloning techniques.

c) Gene Delivery Methods:

- **Microinjection:** Directly inject the vector containing the target gene into the pronucleus of a fertilized egg (**Figure. 12.2b**). This is a common method for creating transgenic mice.

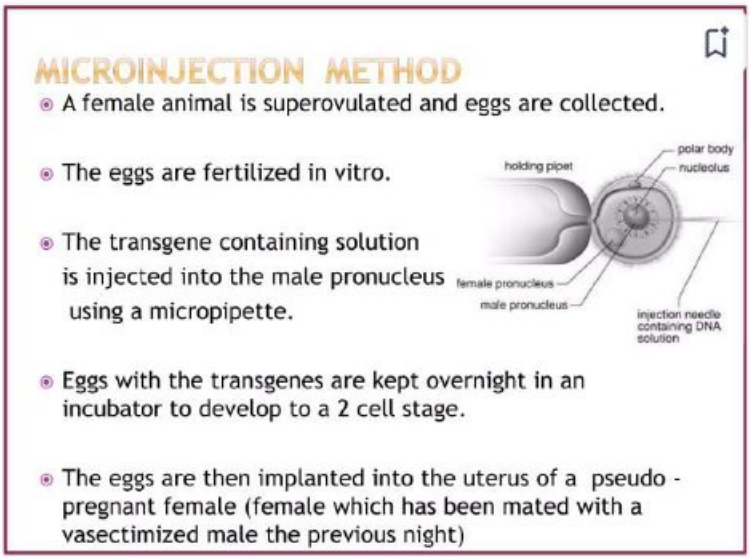


Figure. 12.2b-Microinjection for creating transgenic animal (https://www.slideshare.net/slideshow/transgenic-animals-174586923/174586923#5)

- **Viral Vectors**: Use modified viruses to deliver the target gene into animal cells. Retroviruses and lentiviruses are often used for their ability to integrate into the host genome.
- **Electroporation**: Apply an electrical field to cells to increase cell membrane permeability, allowing the vector to enter the cells. This method is often used with embryonic stem cells.
- **Gene Gun**: Shoot microscopic particles coated with the DNA into the target cells. This method is more commonly used in plant transgenics but can be applied to some animal cells.

d) Selection and Screening:

- **Selection**: Use selectable markers (such as antibiotic resistance) to identify and select cells that have successfully integrated the transgene.
- **Screening**: Verify the presence and correct integration of the transgene using techniques like PCR, Southern blotting, or fluorescence in situ hybridization (FISH).

✓ Transgenic progenies are screened by PCR to examine the site of the incorporation of the gene

✓ Some transgene may not be expressed if integrated into the transcriptionally inactive site.

e) Embryo Transfer:

- **Recipient Preparation:** Prepare a surrogate mother by synchronizing her estrous cycle with the donor animal.
- **Embryo Implantation:** Transfer the genetically modified embryos into the uterus of the surrogate mother for development.

f) Breeding and Establishing Transgenic Lines:

- **Breeding:** Breed the offspring to produce homozygous transgenic animals. This may require several generations of selective breeding.
- **Characterization:** Thoroughly characterize the transgenic animals to ensure stable integration and expression of the transgene. This involves genetic, phenotypic, and functional analysis.

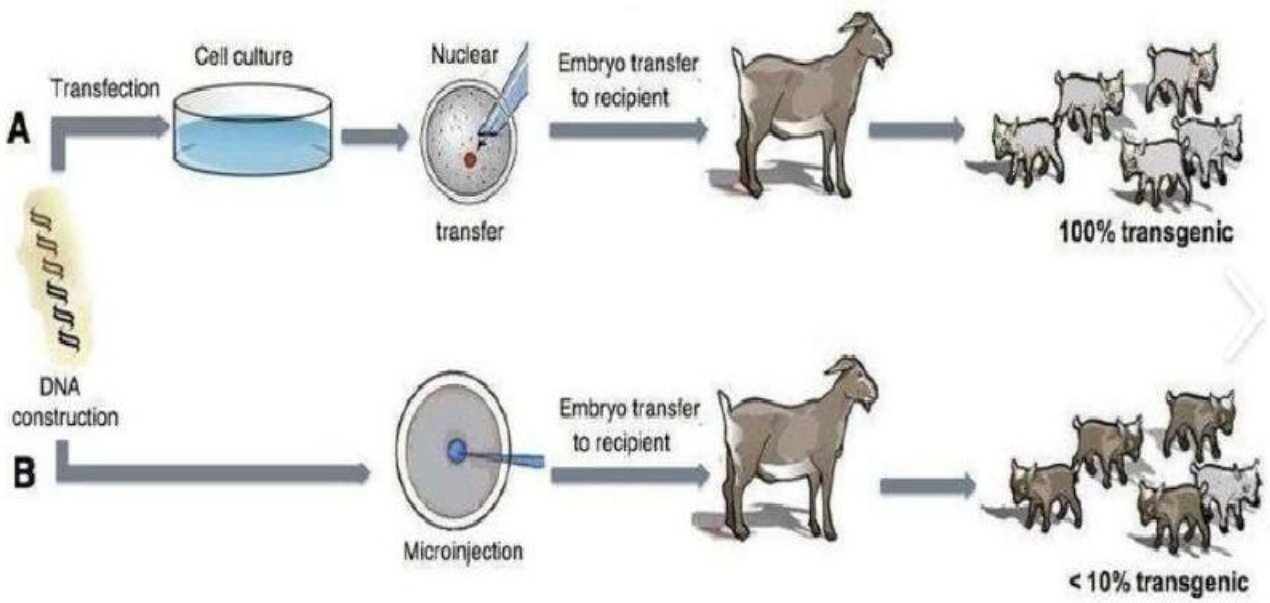


Figure 12.3aSteps involved in producing transgenic animals (goat)A) somatic cell nuclear transfer (SCNT) and pronuclear microinjection (B) (DOI: 10.1590/S1516-89132011000500010)

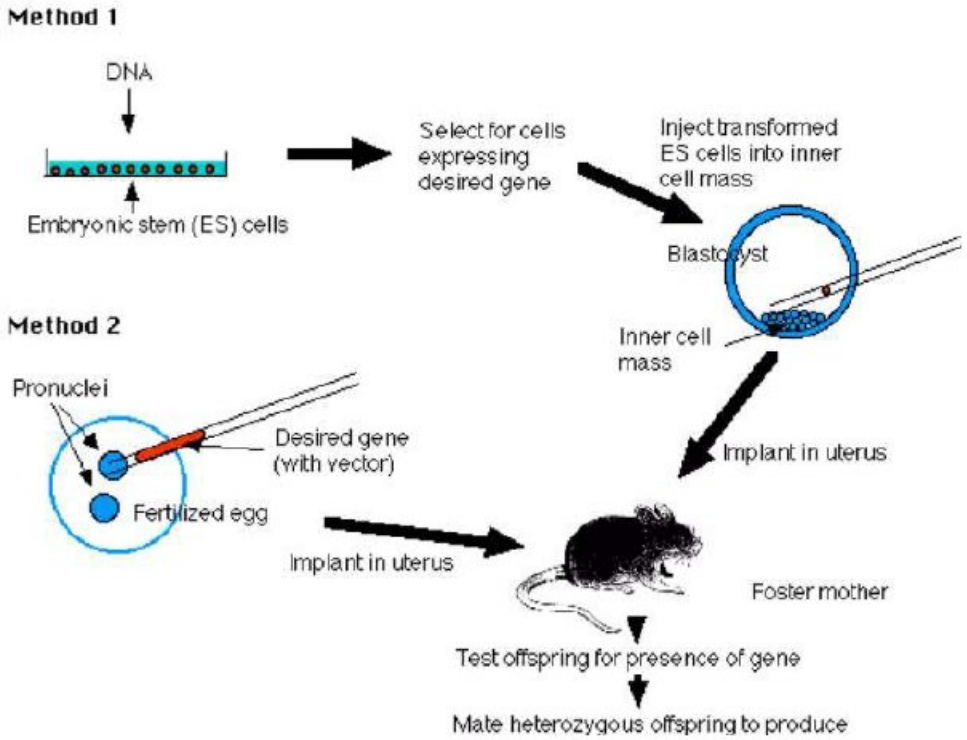


Figure 12.3bSteps involved in producing transgenic animals (mouse).https://www.slideshare.net/slideshow/transgenic-animals-174586923/174586923#9

SAQ4: Producing Transgenic Animals

1. Target gene selection involves identifying the gene of interest that needs to be introduced into the animal's _____.
2. The vector often includes a _____ to ensure gene expression and a _____ marker for identifying successfully modified cells.
3. Directly injecting the vector containing the target gene into the pronucleus of a fertilized egg is a common method for creating transgenic _____.
4. Screening involves verifying the presence and correct integration of the transgene using techniques like _____, Southern blotting, or fluorescence in situ hybridization (FISH).

12.6Making Transgenic Cow (Case study)

Scientists employ a variety of techniques to produce transgenic cows, including DNA cloning, restriction enzyme digests, ligation, polymerase chain reaction (PCR), transfection, nuclear transfer, and in vitro embryo production. At AgResearch, scientists have successfully created transgenic cows

that generate additional proteins in their milk through the following seven steps (**Figure 12.5**) (https://www.sciencelearn.org.nz/resources/857-techniques-used-to-make-transgenic-cows).

Step 1: Gene ConstructDesigning

The construct can be designed to be inserted randomly into the genome of the animal, which is called transgenesis by addition, or can be designed to be inserted into the genome at a specific targeted site into the correct position of a determined chromosome, which is called transgenesis by homologous recombination. The initial step involves designing a gene construct, which is a unit of DNA containing the following components (https://www.sciencelearn.org.nz/resources/857-techniques-used-to-make-transgenic-cows):

✓ *Antibiotic resistance gene*: This gene allows for the selection of cells that have successfully incorporated the gene construct. These genes are often located on plasmids or transposons and can be transferred from cell to cell by conjugation, transformation, or transduction. This gene exchange allows the resistanceto rapidly spread throughout a population of bacteria and among different species of bacteria.

✓ *Tissue-specific promoter sequence*: Tissue specific promoters are available as native or composite promoters, consists of a single fragment from the 5' region of a given gene. Each of them comprises a core promoter and its natural 5' UTR. In some areas the 5' UTR contains an intron. This sequence signals the start of gene expression specifically in the targeted tissue, such as mammary cells in lactating cows.

✓ *Desired gene*: This is the gene of interest, such as bovine casein or human myelin basic protein, which is intended to be expressed.

✓ *Stop sequence*: This sequence marks the end of the gene's information, ensuring proper termination of protein synthesis.

Step 2: Transgene Sourcing

Traditionally, the gene would have been extracted from the DNA of the source organism. Nowadays, if the desired gene sequence is known, it can be synthesized in a lab. Specialized companies can create custom genes within a couple of weeks.

Step 3. Gene Construct Making

The gene is typically provided in a vector, a small DNA molecule, often a plasmid, that can accommodate a foreign DNA segment(**Figure 12.4**). Vectors facilitate the transfer of the gene between labs, its storage, manipulation, or use in transforming bacteria to amplify the gene of interest. Vectors contain multiple restriction enzyme sites, also known as multiple cloning sites. These sites allow for the insertion and excision of the gene using restriction enzymes. Once the gene is excised from the vector, it is inserted into the multiple cloning site of the gene construct through a process called ligation.

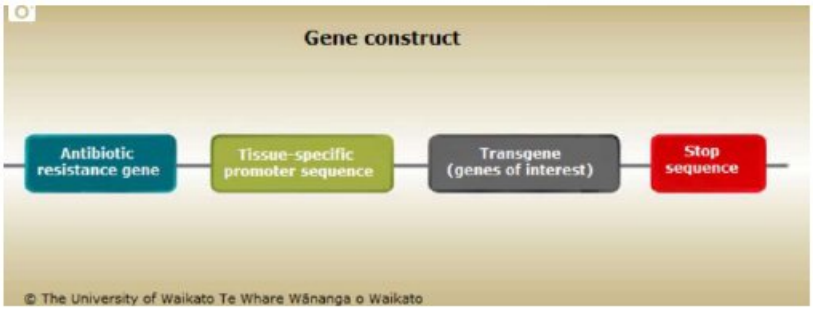


Figure 12.4 How to make gene construct: https://www.sciencelearn.org.nz/resources/857-techniques-used-to-make-transgenic-cows

Step 4: Transfecting Bovine Cells

The gene construct is introduced into the genome of a bovine (cow) cell through a process called transfection. During transfection, the cell membrane is temporarily permeabilized to allow the DNA to enter. This can be achieved by applying an electrical pulse (electroporation) or by using chemical agents. Once the gene construct is inside the cell, it can integrate into the cell's genome within the nucleus.

Step 5: Selecting for Transgene-Positive Cells

Post-transfection, an antibiotic is applied to identify and select the bovine cells that have successfully incorporated the gene construct. Transgenic bovine cells survive the antibiotic treatment due to the presence of the antibiotic resistance gene, while cells without the construct perish. Additionally, the presence of the transgene in the surviving cells is confirmed using polymerase chain reaction (PCR).

Step 6: Creating a Transgenic Embryo via Nuclear Transfer

A transgenic embryo is produced by transferring the nucleus from a single transgenic bovine cell into an enucleated egg cell, using a technique known as nuclear transfer. This process can lead to the development of a whole transgenic animal from the modified cell. The transgenic bovine cell is fused with a bovine oocyte from which the chromosomes have been removed (enucleated oocyte). An

electrical pulse is applied to facilitate the fusion of the two cells. After fusion, the transgenic cell's chromosomes are reprogrammed to direct the development of the oocyte into an embryo. Within 7 days, the transgenic embryo will develop to around 150 cells and can then be transferred into a recipient cow for further development to term.

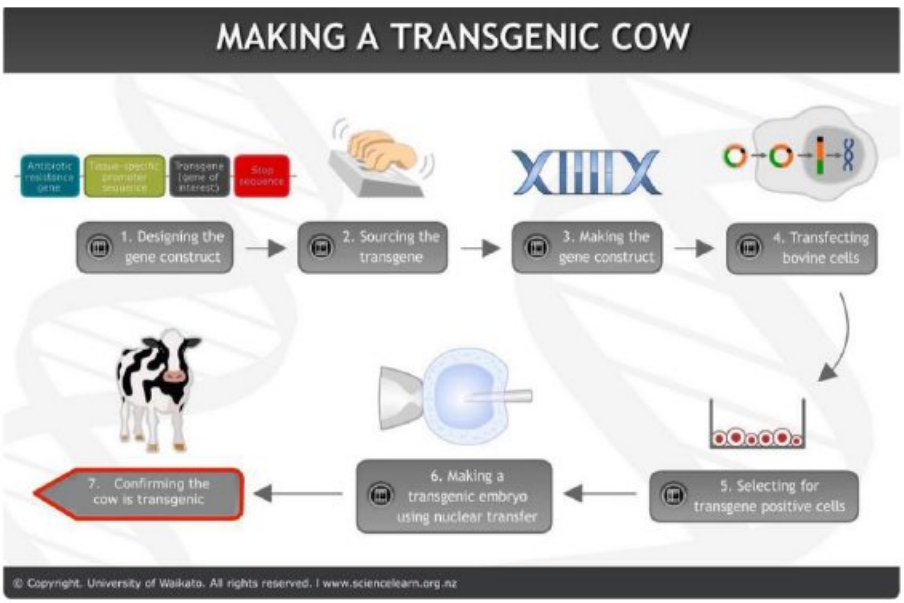


Figure 12.5 How to make transgenic cow, A case study: https://www.sciencelearn.org.nz/resources/857-techniques-used-to-make-transgenic-cows

Step 7: Confirming the Cow is Transgenic

If the embryo develops to full term, the cow will give birth to a calf after approximately 9 months. To confirm that the calf is transgenic, scientists use the following methods (**Figure 12.6**):

- **PCR**: To detect the presence or absence of the transgene.
- **Quantitative PCR (q-PCR)**: To quantify the number of transgene copies.
- **Fluorescent In Situ Hybridization (FISH)**: To visualize the location of the transgene on the chromosome and verify if it has integrated into multiple chromosomes.

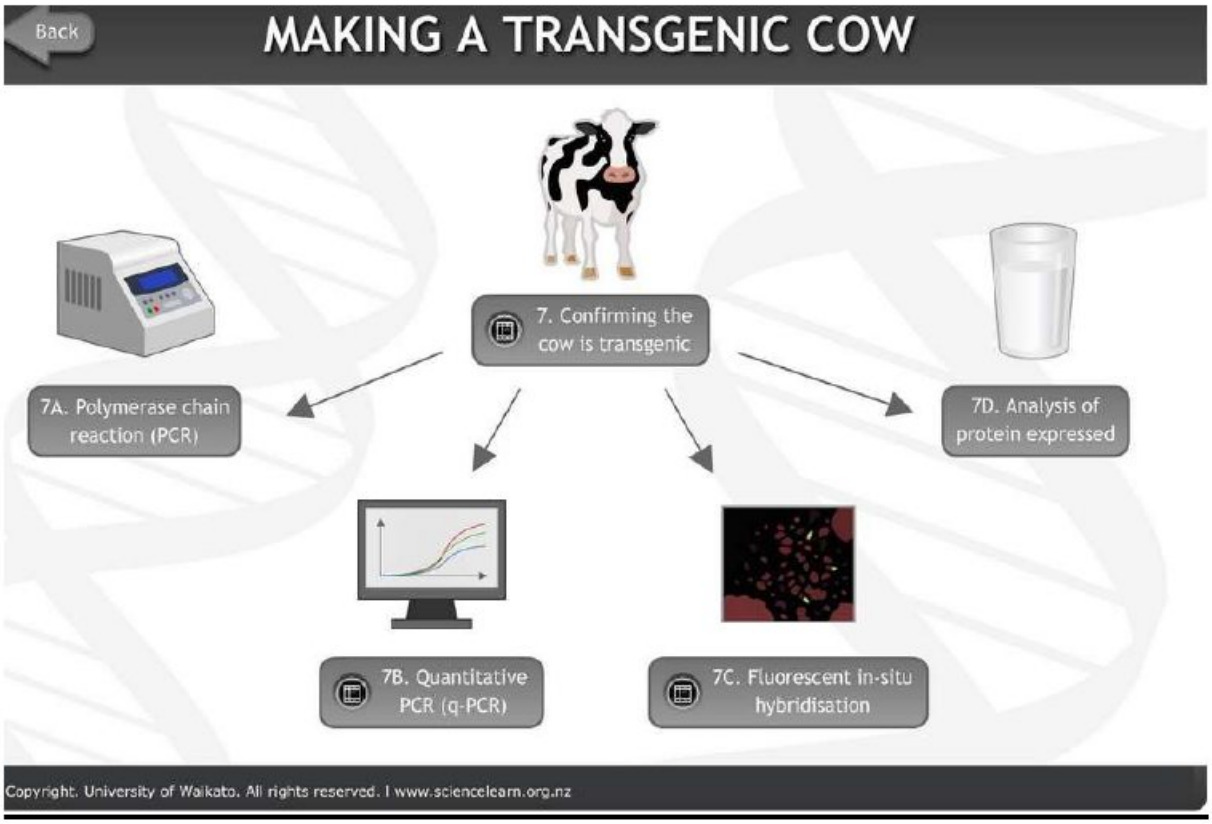


Figure 12.6 How to confirm a cow is transgenic?

https://www.sciencelearn.org.nz/resources/857-techniques-used-to-make-transgenic-cows

12.7 Applications of Transgenic Animals

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. By harnessing genetic engineering, transgenic animals provide valuable tools for advancing science, improving agriculture, producing medicines, and addressing environmental challenges. Below are few applications:

- **Disease Models**: Transgenic animals, such as mice, are engineered to carry genes associated with human diseases. These models aid researchers in studying disease mechanisms and developing treatments. For instance, transgenic mice with human Alzheimer's genes are used to understand disease progression and test new therapies.
- **Drug Development**: By expressing human proteins, transgenic animals enable more accurate testing of pharmaceuticals. This process helps in evaluating drug efficacy and safety before human clinical trials.
- **Gene Function Studies**: Scientists use transgenic animals to investigate the role of specific genes in development, physiology, and pathology. This research can reveal the functions of unknown genes and their contributions to health and disease.

a) Agricultural Improvements:

➢ **Enhanced Growth and Productivity**: Transgenic animals can be modified to grow faster, produce more milk, or yield better meat quality. For example, transgenic salmon have been developed to grow faster than their non-transgenic counterparts.

➢ **Disease Resistance**: Livestock can be engineered to resist common diseases, reducing the need for antibiotics and improving animal welfare. An example is transgenic pigs that are resistant to African swine fever.

➢ **Nutritional Quality**: Genetic modifications can enhance the nutritional content of animal products. Transgenic animals can produce milk with higher levels of beneficial fatty acids or eggs enriched with vitamins.

b) Biopharming: It is the production of pharmaceutical proteins in genetically engineeredplants.

➢ **Production of Therapeutic Proteins**: Transgenic animals can be used to produce human proteins, such as insulin, in their milk, blood, or eggs. This method, known as "biopharming," offers a cost-effective and scalable way to produce pharmaceuticals.

➢ **Clotting Factors**: Transgenic goats and cows have been engineered to produce clotting factors in their milk, which are essential for treating haemophilia patients.

➢ **Antibodies and Vaccines**: Transgenic chickens have been developed to lay eggs containing human antibodies or vaccines, streamlining the production process for these critical therapeutic agents.

c) Xenotransplantation: Xenotransplantation or heterologous transplant is the transformation of living cells, tissues, or organs from one species to another species. Such cells, tissues or organs are called xenografts or Xenotransplants.

➢ **Organ Donation**: Transgenic pigs are being developed with modified genes to make their organs more compatible with the human immune system. This research aims to address the shortage of human organs available for transplantation, potentially saving many lives.

d) Environmental Benefits:

➢ **Bioremediation**: Transgenic animals can be engineered to process and detoxify environmental pollutants. For example, transgenic fish have been created to detect and respond to water pollution, serving as living indicators of environmental health.

SAQ5: Applications of Transgenic Animals

---

1. Transgenic animals, such as mice, are engineered to carry genes associated with human _____.

2. By expressing human _____, transgenic animals enable more accurate testing of pharmaceuticals.

3. Transgenic animals can be modified to grow faster, produce more milk, or yield better meat quality. An example is transgenic _____ that have been developed to grow faster.

4. Transgenic animals can produce milk with higher levels of beneficial _____ acids or eggs enriched with _____.

5. Transgenic goats and cows have been engineered to produce clotting factors in their milk, which are essential for treating _____ patients.

6. Transgenic chickens have been developed to lay eggs containing human _____ or vaccines.

7. Transgenic pigs are being developed with modified genes to make their organs more compatible with the human _____ system.

8. Transgenic fish have been created to detect and respond to water _____, serving as living indicators of environmental health.

12.8 Summary

Transgenic and cloned animals represent pivotal advancements in biotechnology, each with distinct methodologies and applications. Transgenic animals are genetically engineered to carry genes from other species, allowing researchers to study gene functions, develop new treatments, and enhance agricultural traits. The creation of transgenic animals involves methods such as microinjection, retrovirus-mediated gene transfer, embryonic stem cell-mediated gene transfer, and CRISPR-Cas9 technology. These animals play crucial roles in biomedical research for disease modeling and gene function studies, agriculture for improved growth and disease resistance, and environmental applications like bioremediation. Examples include GloFish, EnviroPig, and knockout mice. Ethical considerations include animal welfare, environmental impact, human health, and moral beliefs. In contrast, cloned animals are exact genetic replicas of another animal, created through techniques like somatic cell nuclear transfer (SCNT) and embryo splitting, with Dolly the Sheep being the first mammal cloned from an adult somatic cell. Cloning is used in agriculture to replicate desirable traits, conservation to preserve endangered species, medical research to study genetic disorders, and pet cloning. However, cloned animals often face health issues such as developmental abnormalities and shorter lifespans. Ethical concerns encompass the welfare of cloned animals, reduced genetic diversity, and the implications of human cloning. Transgenic animals differ from cloned animals in that they carry foreign genes, while cloned animals are genetic duplicates of the donor. The production of transgenic animals involves gene identification and isolation, vector construction, gene delivery methods, selection and screening, embryo transfer, and breeding to establish transgenic lines. Applications include disease modeling, drug development, gene function studies, agricultural improvements, biopharming for therapeutic proteins, and environmental benefits like bioremediation. Both transgenic and cloned animals offer significant benefits in research, agriculture, and medicine, but their development and use must be carefully managed to address ethical, environmental, and health-related concerns.

12.9 Suggested Readings

- **Transgenic Animal Technology**: A Laboratory Handbook" by Carl A. Pinkert: This comprehensive handbook covers the methods and techniques used in the production of transgenic animals.
- **Principles of Cloning** by Jose Cibelli, Robert P. Lanza, Keith H.S. Campbell, and Michael D. West: This book provides an in-depth look at the science of cloning, which is closely related to the creation of transgenic animals.
- **Animal Biotechnology**: Science-Based Concerns by the National Research Council: This book discusses the scientific and ethical concerns associated with animal biotechnology, including the creation of transgenic animals.

12.10Terminal Questions

Q.1. What do you understand by transgenic animals?

Answer:----------

Q.2. Write few examples of transgenic animals.

Answer:----------

Q.3. How do you differentiate betweentransgenic animals and cloned animals?

Answer:----------

Q.4. What are the various methods of producing transgenic animal?

Answer:----------

Q.5. What are the different steps involved in producing transgenic animals?

Answer:----------

Answers

SAQ1: Transgenic Animals

1. Transgenic animals are organisms that have had **foreign** genes deliberately inserted into their genome.
2. The foreign genes in transgenic animals can come from **different** species, and the process is known as **genetic engineering**.
3. Methods of creating transgenic animals include microinjection, retrovirus-mediated gene transfer, embryonic stem cell-mediated gene transfer, and **CRISPR-Cas9**.
4. Transgenic animals are used in biomedical research, agriculture, **and pharmaceutical production.**
5. An example of a transgenic animal used for producing human proteins in milk is the **transgenic goat**.
6. Ethical considerations for transgenic animals include concerns about animal welfare, environmental impact, human health, and **cultural** and **religious** beliefs.

SAQ2: Cloned Animals

1. Cloned animals are organisms that are genetically **identical** copies of another organism.
2. Dolly the Sheep, born in 1996, was the first mammal cloned from an adult somatic cell using a technique called **somatic cell nuclear transfer**.
3. Cloning can help preserve **endangered** species by creating genetic duplicates of individuals to increase population numbers.
4. Cloned animals often face health problems, including developmental abnormalities and **shortened** lifespans.
5. An example of a cloned pet cat is **CC** (Carbon Copy) the Cat.

SAQ3:Transgenic Animals Vs Cloned Animals

1. Transgenic animals are those that have had **foreign** genes deliberately inserted into their genome, whereas cloned animals are **genetically** identical copies of another animal.
2. Transgenic animals are mainly used for research, agricultural improvements, and **pharmaceutical** production.
3. Cloned animals often face more significant health problems than transgenic animals, including developmental **abnormalities** and shorter **lifespans**.

SAQ4: Producing Transgenic Animals

1. Target gene selection involves identifying the gene of interest that needs to be introduced into the animal's **genome**.

2. The vector often includes a **promoter** to ensure gene expression and a **selectable** marker for identifying successfully modified cells.

3. Directly injecting the vector containing the target gene into the pronucleus of a fertilized egg is a common method for creating transgenic **animals**.

4. Screening involves verifying the presence and correct integration of the transgene using techniques like **PCR**, Southern blotting, or fluorescence in situ hybridization (FISH).

SAQ5: Applications of Transgenic Animals

1. Transgenic animals, such as mice, are engineered to carry genes associated with human **diseases**.

2. By expressing human **proteins**, transgenic animals enable more accurate testing of pharmaceuticals.

3. Transgenic animals can be modified to grow faster, produce more milk, or yield better meat quality. An example is transgenic **salmon** that have been developed to grow faster.

4. Transgenic animals can produce milk with higher levels of beneficial **fatty** acids or eggs enriched with **omega**-3.

5. Transgenic goats and cows have been engineered to produce clotting factors in their milk, which are essential for treating **hemophilia** patients.

6. Transgenic chickens have been developed to lay eggs containing human **proteins** or vaccines.

7. Transgenic pigs are being developed with modified genes to make their organs more compatible with the human **immune** system.

8. Transgenic fish have been created to detect and respond to water **pollution**, serving as living indicators of environmental health.