6th SEM B.Sc. Botany

CALICUT UNIVERSITY

BIOTECHNOLOGY, MOLECULAR BIOLOGY AND BIOINFORMATICS

2019 ADMISSION

Prepared by

Ms. Visruthi Vijayan

Dr. Lins Simon

Syllabus

BOT6B11T: BIOTECHNOLOGY, MOLECULAR BIOLOGY AND BIOINFORMATICS

Lectures Hours per week: 8, Credit – 3, Internal: 20,

External: 80, Examination 3 Hours

Objectives:

- Analyze the role of biotechnology in daily life.
- Understand the basic aspects of bioinformatics.
- Explain the concepts in molecular biology.

BIOTECHNOLOGY

Module-1

1. The concept of biotechnology, landmarks in the history of biotechnology.
2. Recombinant DNA Technology: Gene cloning strategies – recombinant DNA construction – cloning vectors – plasmids pBR322, bacteriophage based vectors, Ti plasmids. Restriction endonucleases and ligases transformation and selection of transformants – using antibiotic resistances markers, southern blotting; PCR.
3. Different methods of gene transfer – chemically stimulated DNA uptake by protoplast, electroporation, microinjection, biolistics. Agrobacterium mediated gene transfer gene library, gene banks.

Module -II

2. 1. Application of Biotechnology in : a. Medicine - Production of human insulin, human growth hormone and b. Forensics - DNA finger printing. c. Agriculture - Genetically modified crops – Bt crops, Golden rice, d. Flavr Savr Tomato, Virus, herbicide resistant crops, Edible vaccines. e.Environment- Bioremediation-use of genetically engineered bacteria-super bug. f. Industry- Horticulture and Floriculture Industry, production of vitamins, amino acids and alcohol.

MOLECULAR BIOLOGY [1 ¼ hrs per week]

Module – I.

1. Nucleic acids - DNA – the genetic material; the discovery of DNA as the genetic material; bacterial transformation (Griffith's & Avery's experiments); Hershey and Chase experiment; Structure of DNA, Watson & Crick's Model, Types of DNA-(A,B,Z); Replication –semi conservative replication – Meselson and Stahl's experiment; Molecular mechanism of Replication. RNA- structure, types and properties.
2. Gene action - One gene - one enzyme hypothesis, one cistron one polypeptide hypothesis; concept of colinearity; modern concept of gene-cistorns, recons and mutons
3. Genetic code - Characters of genetic code
4. Central dogma protein synthesis; Transcription, post-transcriptional modification of RNA, translation; Teminism.
5. Gene regulation in prokaryotes - operon concept, (Lac operon, trp. operon)
6. Gene regulation in eukaryotes (brief account)
7. Mutation-spontaneous and induced; causes and consequences. Types of mutagens and their effects. Point mutations- molecular mechanism of mutationTransition, Transversion and substitution

BIOINFORMATICS Theory: 13 hrs. [3/4 hr. per week]

Module-I

1. IT in teaching, learning and research: Web page designing and web hosting. Academic web sites, e-journals, Open access initiatives and open access publishing, education software, academic services - INFLIBNET, NICNET, BRNET.
2. e-wastes and green computing
3. Futuristic IT - Artificial intelligence, virtual reality, bio-computing

Module – II

1. Introduction to Bioinformatics, scope and relevance. Brief History
2. Basics of Genomics, Proteomics and Comparative Genomics
3. Biological data bases and data bases, Nucleotide sequencedatabase – EMBL, Gen Bank, DDBJ; Protein sequence database – Swiss Prot, PDB, Uri Prot, PIR Organismal database – Human genome database Biodiversity database – Species 2000
4. Information retrieval from Biological database, sequence alignment types and tools: pair wise sequence alignment multiple sequence alignment, clustal W, BLAST

Module- III

1. Genomics: DNA sequencing Sangers procedure-automation of DNA sequencing, genome sequence assembly.
2. Genome projects – Major findings and relevance of the following genome projects – Human, Arabidopsis thaliana, Rice, Haemophilus influenza.
3. Proteomics : Protein sequencing- automation of sequencing, protein structure prediction and modelling (Brief account only)

Module- IV

A brief account on

1. Molecular phylogeny and phylogenetic trees.
2. Molecular visualization – use of Rasmol.
3. Molecular docking and computer aided drug design.

BIOTECHNOLOGY, MOLECULAR BIOLOGY, BIOINFORMATICS

Module I -Bioinformatics

IT in teaching, learning and research Access to learning the material

- Continuous learning
- Sharing of knowledge
- Using audio and visual material as learning aids
- Distance learning
- Proper record keeping
- Improves engagement
- Improves knowledge retention
- Encourages collaboration
- Students can learn useful life skills
- Virtual class room
- Educational software : software developed for the primary purpose of teaching is called educational software
  - Children's learning : software program developed for children is based on computer gaming concepts
  - Course ware : educational material kit for students
  - Classroom aid : software is projected onto a large whiteboard at the front of the class and it works simultaneously on a network of desktop computers in a classroom
  - Edutainment : in this computer games and educational software is merged together as a single product
  - Reference software: dictionaries and encyclopaedia
  - Software for specific purpose : includes spelling tutor , medical and health care
- Serve a pedagogical purpose to help an educator in disbursing information to students.
- Generally tends to have fairly conservative layouts
- Some are a simple one-page affair with contact information and a list of prominent publications
- Internet address ends with .ac
- In all foreign universities ,academic websites are granted to any professor who wishes to use the web as a tool for presenting his or her class information
- It can be used to enhance course presentation by offering online course syllabi , copies of presentation, links , reading assignments , feedback forums and class schedules

Web designing

- **Web pages** : files which can be displayed using a web browser application like internet explorer
- Web page created is accessed from other computer in internet through web server which holds the web
- pages
- **Web server** : another computer running web server software
- **Web site** : web pages possibly accessed through links connected to the home page
- **Web address** : place where the site files are saved which is used in the address field of the browser to access the site
- Web pages are of 2 types ,

 Static web pages : don't change the content or layout with every request to the web server. They change only when a web author manually updates them with a text editor or web editing tool

 Dynamic web pages/ interactive webpages : can adapt their content or appearance depending on the user's interactions , changes in data supplied by an application , or an evolution over time.

Designing of web pages

- One has to consider all the types of media elements that is being used to create web pages
- Media element neede : text , audio , graphics , animations , and video clips
- Common template is designed first
- Web designing software example : linux Ubuntu
- Design should match with the purpose of creating the web page , the field , target audience and the customer's taste
- Types of web sites,
  - Commercial web sites: end with .com
  - Educational websites: end with .edu
  - Organisational website: .org
  - News website: .com, .news
  - Personal website: .org , .com, .me
  - Blog

Web authoring

- Creating web pages with web authoring language
- Example : HTML – HyperText Markup Language
- HTML is a tahg language using markupsto describe the elements on the web page

Web hosting

- Service that makes the storage of web sites or homepages on the internet so that others can see it and interact with it
- Hosting companies operate with many web servers which store and serve websites to the world
- Web server is a special, highly reliable computer configured to run web serving applications
- Each server can be configured to host many different websites or to be totally available only for a particular customer
- Example : net4india.com
- Types ,
  - Shared hosting : depending on the web host , a physical web server can host a few hundred to even thousands of different websites at one time
  - Dedicated hosting : a specific web server is used by only one customer
  - Co-location hosting : customer owns the web server hardware and only housed their web server within the web hosting provider's secure data center.customer has full control over their web server
  - Reseller hosting: a web hosting provider offers web server storage to third part (reseller) at a discount price , who then resell the web server storage to their customers

e-journal

- Academic journals or intellectual magazines that can be accessed via electronic transmission
- Usually published in web

- Some are online-only journals , online versions of printed journals and some consist of the online modified version of a printed journal with videos and other interactive material
- Most commercial journals are subscription-based or allow pay-per-view access onlymost are published in HTML and or PDF format

Open access initiatives and open access publishing

- Open access : term used to describe published academic papers , books , reports and other periodicals. Those are electronically available to readers without financial or technological barriers
- Adavantages :
  - This will help scientists to develop collaborations with unknown collegues ,
  - Increases citation and usage of the published research
  - Work gets widely disseminated
  - Books , articles and art forms will enjoy world wide patronage
- India has done well in Open Access area by making 81 scientific journals accessible under OA
- Example : Digital Library of India : http://dli.iiit.ac.in
- Main aspects ,
  - OA self archiving / OA initiative : Also known as green road / route to OA. Here, authors publish in a subscription journal , but in addition make their articles freely accessible online , usually by depositing them in either an institutional repository such as PubMed Central
  - OA publishing :also known as gold road / route to OA), here, authors publish in open access journals that make their articles freely accessible online immediately upon publication

ACADEMIC SERVICES

INFLIBNET

- Information and Library Network (INFLIBNET) Centre is an Autonomous Inter-University Centre (IUC) of University Grants Commission (UGC)
- Involved in creating infrastructure for sharing of library and information resources and services among Academic and Research Institutions.
- Works collaboratively with Indian University Libraies.
- Involved in modernizing university libraries in India.
- Aid connecting them as well as information centres through a nationwide high speed data network using state-of-art technologies.
- Services includes financial support to universities , internet connectivity programme , training etc

INFLIBNET provides,

- Document delivery : providing interlibrary loans and document delivery services
- Bibliographics
  - Union Databases
  - Bibliographic Databases
  - Non-bibliographic Databases

Library

- SOUL software : Software for University Libraries having maximum number of installations in the libraries of academic and research institutions across the country
- Print archival: centre maintains a separate archival library consisting of print journals

NICNET

- Designed and implemented by NIC (National Informatics Centre) using state-of-the-art satellite based computer communication technology
- It is a premiere Science and Technology institution of the Government of India
- One of the largest VSAT(Very Small Aperture Terminal) networks
- Established in 1976 for providing e-Government/e-Governance solutions
- Example : AGMARK NET : Agriculture; application for driving license

BRNet

- Bio-Resource Network (BRNet) an initiative of Japan Science and Technology is a prototype portal site for biological information and Bio-resource information
- Possible to search bio-resource and its related information and classify , identify the Bio-Resource to construct the databases for Bio-Resource

Module II

BIOINFORMATICS

- Bioinformatics is the application of Information technology to store, organize and analyze the vast amount of biological data.
- The stored data is available in the form of sequences and structures of proteins and nucleic acids (the information carrier).
- The biological information of nucleic acids is available as sequences while the data of proteins is available as sequences and structures.
- Sequences are represented in single dimension where as the structure contains the three dimensional data of sequences.
- Bioinformatics is a field in which biology, mathematics, statistics, CS and IT are merged into a single discipline to process biological data.
- Complex machines are used to read in biological data at a much faster rate than before.
- The term "Bioinformatics" was invented by Paulien Hogeweg and Ben Hesper in 1970

Goals of Bioinformatics

- To uncover the wealth of Biological information hidden in the mass of sequence, structure, literature and biological data.
- It is being used now and in the foreseeable future in the areas of molecular medicine.
- It has environmental benefits in identifying waste and clean up bacteria.
- In agriculture, it can be used to produce high yield, low maintenance crops.

Fields of Bioinformatics

- Molecular Medicine
- Gene Therapy
- Drug Development
- Microbial genome applications
- Crop Improvement
- Forensic Analysis of Microbes
- Biotechnology
- Evolutionary Studies
- Bio-Weapon Creation

Module III

GENOMICS and PROTEOMICS

Genomics is the study of an organism's genome and the use of the genes.

It deals with the systematic use of genome information, associated with other data, to provide answers in biology, medicine, and industry.

Proteomics is the large-scale study of proteins, particularly their structures and functions.
Proteomics is much more complicated than genomics. Most importantly, while the genome is a rather constant entity, the proteome differs from cell to cell and is constantly changing through its biochemical interactions with the genome and the environment. One organism will have radically different protein expression in different parts of its body, in different stages of its life cycle and in different environmental conditions.

Module VI

Software and Tools of Bioinformatics

Primary sequence databases

- EMBL (European Molecular Biology Laboratory nucleotide sequence database at EBI, Hinxton, UK)
- GenBank (at National Center for Biotechnology information, NCBI, Bethesda, MD, USA)
- DDBJ (DNA Data Bank Japan at CIB , Mishima, Japan)

NCBI: National Center for Biotechnology information

Established in 1988 as a national resource for molecular biology information, NCBI creates public databases, conducts research in computational biology, develops software tools for analyzing genome data, and disseminates biomedical information - all for the better understanding of molecular processes affecting human health and disease.

EMBL Nucleotide Sequence Database (also known as EMBL-Bank) constitutes Europe's primary nucleotide sequence resource. Main sources for DNA and RNA sequences are direct submissions from individual researchers, genome sequencing projects and patent applications.

DDBJ (DNA Data Bank of Japan)

DNA data bank activities in earnest in 1986 at the National Institute of Genetics (NIG). DDBJ has been functioning as the international nucleotide sequence database in collaboration with EBI/EMBL and NCBI/GenBank.

DNA sequence records the organismic evolution more directly than other biological materials and ,thus, is invaluable not only for research in life sciences, but also human welfare in general. The databases are, so to speak, a common treasure of human beings. With this in mind, we make the databases online accessible to anyone in the world

Structural Bioinformatics

- Prediction of structure from sequence
- Secondary structure
- Homology modelling, threading
- Analysis of 3D structure
- Structure comparison/ alignment
- Prediction of function from structure - molecular mechanics/ molecular dynamics - prediction of molecular interactions, docking
- Structure databases (RCSB

Structural Bioinformatics

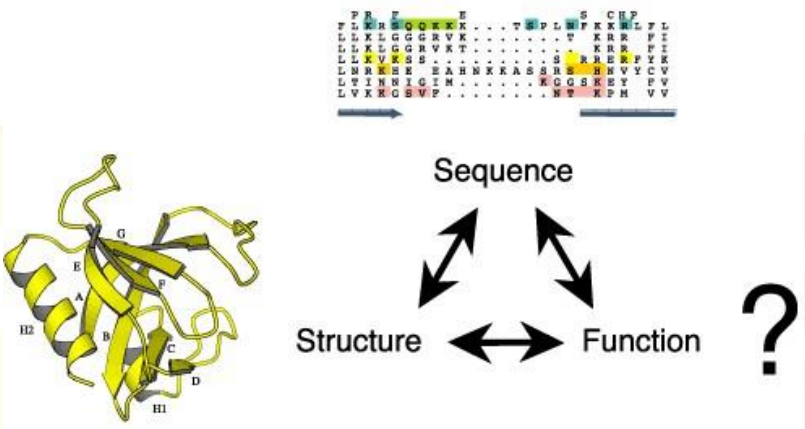
Prediction of structure from sequence

- Secondary structure
- Homology modelling, threading
- 3D prediction

Analysis of 3D structure

- Structure comparison/ alignment
- Prediction of function from structure

- Molecular mechanics/ molecular dynamics
- Prediction of molecular interactions, docking
- Structure databases (RCSB)



Sequence Alignment
Homology: Two (or more) sequences have a common ancestor.
Similarity: Two (or more) sequences are similar by some criterion, and it does not refer to any historical process.
◦FASTA

◦BLAST

Blast (Basic Local Alignment Search Tool)
 It is an algorithm for comparing biological sequences information, such as amino acid sequence of different proteins or the nucleotides of DNA sequences.
 BLAST is used to identify library sequences that resembles the query sequences
Software and Tools of Bioinformatics
Blast (Basic Local Alignment Search Tool)
 It is an algorithm for comparing biological sequences information, such as amino acid sequence of different proteins or the nucleotides of DNA sequences.
 BLAST is used to identify library sequences that resemble the query sequences.
Various versions:
Blastn: nucleotide sequences
Blastp: protein sequences
tBlastn: protein query - translated database
Blastx: nucleotide query - protein database
tBlastx: nucleotide query - translated database

Optimal Alignment: The alignment that is the best, given a defined set of rules and parameter values for comparing different alignments.
Global Alignment: An alignment that assumes that the two proteins are basically similar over the entire length of one another. The alignment attempts to match them to each other from end to end.
Local Alignment: An alignment that searches for segments of the two sequences that match well. There is no attempt to force entire sequences into an alignment, just those parts that appear to have good similarity.
Gaps & Insertions: In an alignment, one may achieve much better correspondence between two sequences if one allows a gap to be introduced in one sequence. Equivalently, one could allow an insertion in the other sequence. Biologically this corresponds to an mutation event.

Substitution matrix: A Substitution matrix describes the two residue types would mutate to each other in evolutionary time. This is used to estimate how well two residues of given types would match if they were aligned in a sequence alignment.
Gap Penalty: The gap penalty is used to help decide whether or not to accept a gap or insertion in an alignment when it is possible to achieve a good alignment residue to residue at some other neighboring point in the sequence.

Module I-Molecular Biology

Gene

Gene: A gene is a basic unit of heredity in a living organism. It is a locatable region of genomic sequence, corresponding to a unit of inheritance, which is associated with regulatory regions, transcribed regions, and or other functional sequence regions. Genes hold the information to build and maintain an organism'scells and pass genetic traits to offspring.

- Allele: Each gene can have different alleles.
- An allele (from the Greek allelos, meaning each other) is *one of two or more forms of the DNA sequence of a particular gene.* E.g. Diploid, Triploid etc.
- The vast majority of living organisms encode their genes in long strands of DNA.
- The most common form of DNA in a cell is in a double helix structure.
- RNA is common as genetic storage material in viruses, in mammals in particular RNA inheritance has been observed very rarely.

Central Dogma of Molecular Biology:

- The flow of genetic information in the cell starts at DNA, which replicates to form more DNA.
- Information is then 'transcribed" into RNA, and then it is "translated" into protein.
- The proteins do most of the work in the cell. Once information gets into protein, it can't flow back to nucleic acid.
- Johan Friedrich Miescher Swiss Biologist Isolated nuclei of white blood cells in 1869 and called it "Nuclein". The major component of "nuclein" is DNA. Protein is the other major component of nuclein.
- In 1889 Richard Altmann discovered that nuclein has acidic properties, and it became called nucleic acid.
- In 1938 Astbury and Bell published the first X-ray diffraction pattern of DNA.
- In 1953 Watson and Crick determined the structure of DNA.
- Wilhelm Roux: in 1883 speculated that chromosomes are the carriers of inheritance.
- Walter Sutton: Determined in 1903 that chromosomes carried units of heredity identified by Mendel.
- William Bateson in 1905 coined the term genetics.
- Wilhelm Johannsen Danish botanist coined the word "gene" ("gen" in Danish and German) in 1909 to describe these fundamental physical and functional units of heredity.

Chromosome Theory of Inheritance

- Genes are arranged in linear fashion on chromosome.
- The reason that certain traits tend to be inheritated together is that the genes governing these traits are on the same chromosome.
- Every gene has its place (locus).
- Diploid organism (human) normally have two copies of all chromosomes (except sex chromosomes)
- DNA recombination occurs in nature

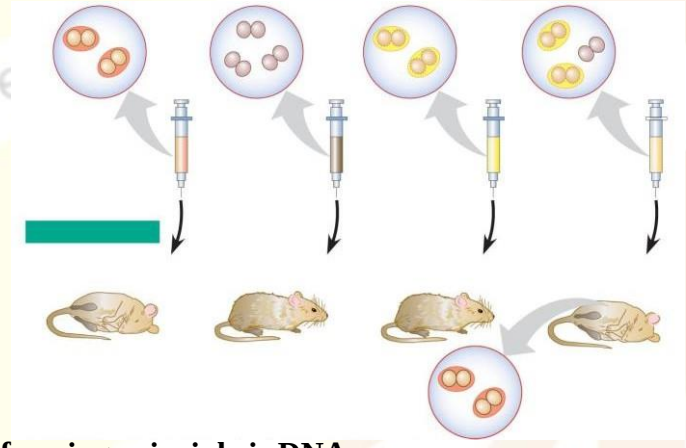
Griffith's Transformation Experiment

• The discovery of the genetic role of DNA in 1928

• Two strains of a bacterium, a pathogenic "S" and a harmless "R"

• Mixed heat-killed remains of the pathogenic strain with living cells of the harmless strain, some living cells became pathogenic

• He called this phenomenon transformation, now defined as a change in genotype and phenotype due to assimilation of foreign DNA

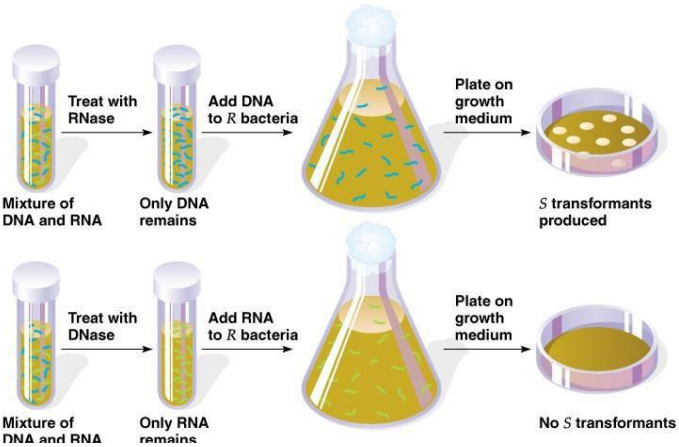


Evidence that Transforming principle is DNA

- In 1944, Oswald Avery, Maclyn McCarty, and Colin MacLeod announced that the transforming substance was DNA.
- Their conclusion was based on experimental evidence that only DNA worked in transforming harmless bacteria into pathogenic bacteria
- Many biologists remained skeptical, mainly because little was known about DNA
- Led by the earlier experiment of transfer genetic trait from one strain of bacteria to another

Evidence that Transforming principle is DNA

- In 1944, Oswald Avery, Maclyn McCarty, and Colin MacLeod announced that the transforming substance was DNA.
- Their conclusion was based on experimental evidence that only DNA worked in transforming harmless bacteria into pathogenic bacteria
- Many biologists remained skeptical, mainly because little was known about DNA Led by the earlier experiment of transfer genetic trait from one strain of bacteria to another.



Hershey-Chase Bacteriophage Experiment

• In 1952, Alfred Hershey and Martha Chase performed experiments showing that DNA is the genetic material of a phage known as T2

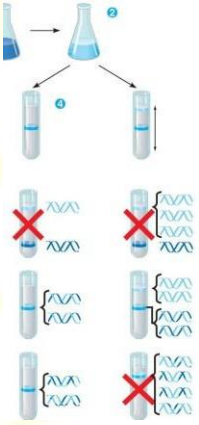
• To determine the source of genetic material in the phage, they designed an experiment showing that only one of the two components of T2 (DNA or protein) enters an *E. coli* cell during infection

• 32P is discovered within the bacteria and progeny phages, whereas 35S is not found within the bacteria but released with phage ghosts.

• They concluded that the injected DNA of the phage provides the genetic information

Meselson-Stahl Experiments

- Labeled the nucleotides of old strands with a heavy isotope of nitrogen (15N), new nucleotides were indicated by a lighter isotope (14N).
- The first replication in the 14N medium produced a band of hybrid (15N-14N) DNA, eliminating the conservative model.
- A second replication produced light and hybrid DNA, eliminating the dispersive model and supporting the semiconservative model.

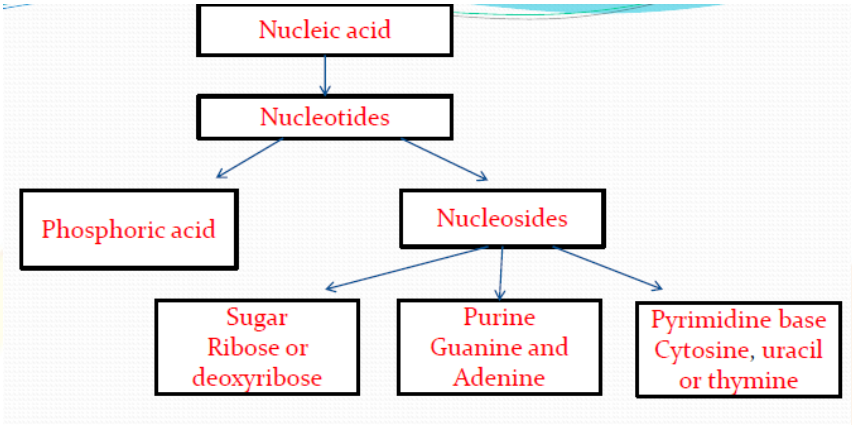


Additional Evidence That DNA Is the Genetic Material

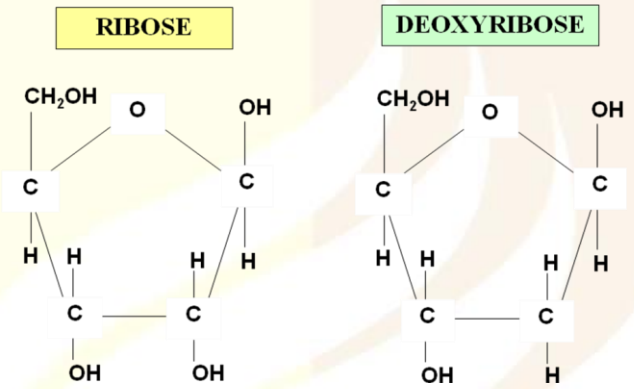
- 1947: Erwin Chargaff- DNA composition varies from one species to the next.
- By 1950s: DNA is a polymer of nucleotides, G=C, A=T
- Franklin's X-ray crystallographic images of DNA enabled Watson and Crick to deduce that DNA was helical
- The X-ray images also enabled Watson and Crick to deduce the width of the helix and the spacing of the nitrogenous bases
- The width suggested that the DNA molecule was made up of two strands, forming a double helix
- This evidence of diversity made DNA a more credible candidate for the genetic material

Nucleic acids

- Two types of nucleic acid are found
- Deoxyribonucleic acid (DNA)
- Ribonucleic acid (RNA)



Spot the difference



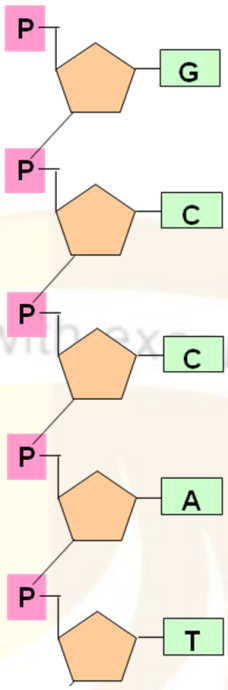
DNA structure

Sugar-phosphate backbone

- The nucleotides are all orientated in the same direction
- The phosphate group joins the 3rd Carbon of one sugar to the 5th Carbon of the next in line

ADDING IN THE BASES

- The bases are attached to the 1st Carbon
- Their order is important
- It determines the genetic information of the molecule



Double Helix Model of DNA Structure-James Watson and Francis Crick

- The "double helix" structure was first discovered by James D. Watson and Francis Crick in 1953.
- Watson, Crick & Wilkins won Nobel prize in 1962.
- Discovery of DNA Double Helix - Published in Nature April, 1953
- Deoxyribonucleic acid (DNA) is a nucleic acid that contains the genetic instructions used in the development and functioning of all known living organisms.
- The main role of DNA molecules is the long-term storage of information.
- DNA is often compared to a set of blueprints, since it contains the instructions needed to construct other components of cells, such as proteins and RNA molecules.
- The DNA segments that carry this genetic information are called genes, but other DNA sequences have structural purposes, or are involved in regulating the use of this genetic information.
- Chemically, DNA is a long polymer of simple units called nucleotides, with a backbone made of sugars and phosphate groups joined by ester bonds.
- Attached to each sugar is one of four types of molecules called bases.
- It is the sequence of these four bases along the backbone that encodes information.
- Two polynucleotide molecules (strands) diameter 2nm run in opposite directions (5' >3' and 3' > 5')
- Sugar phosphate backbone outside helix & bases inside helix; strands bonded by weak H_2 bonds
- Base pairing complementary; order of bases unique.
- Nucleotide base pairs are stacked at 3.4Å spacing
- Hydrogen bonding between purine and pyrimidine across the two strands
- Each base pair is rotated by 36^o with respect to the adjacent pair; 10 base pairs per helical turn;diameter of the double helix is 20Å.
- 2 internal helical grooves - Major and Minor Groove both are large enough to allow protein molecules to come into contact with the base pairs.

- In paired nucleotides the sugar component of one strand will face up while other will face down
- DNA strands are complementary & run anti-parallel with the 3'-OH terminus on one strand adjacent to the 5'-P terminus on the complementary strand.
- DNA & histone form nucleoprotein complex, which makes up 60 - 90% of the bulk of a chromosome.
- It represents the fundamental molecular basis of chromosome structure.
- Chemical analysis of isolated nucleoprotein DNA : histone 1 : 1

CHEMICAL STRUCTURE OF DNA

- The DNA chain is 22 to 26 A wide (2.2 to 2.6 nm).
- One nucleotide unit is 3.3 Ångstroms (0.33 nanometres) long.
- The largest human chromosome, chromosome number 1, is 220 million base pairs long.
- DNA does not usually exist as a single molecule, but instead as a tightly-associated pair of molecules.
- These two long strands entwine like vines, in the shape of a double helix.
- In general, a base linked to a sugar is called a nucleoside and a base linked to a sugar and one or more phosphate groups is called a nucleotide.
- The sugar in DNA is 2-deoxyribose, which is a pentose (five-carbon) sugar.
- The sugars are joined together by phosphate groups that form phosphodiester bonds between the third and fifth carbon atoms of adjacent sugar rings.
- DNA is stabilized by hydrogen bonds between the bases attached to the two strands.
- The four bases found in DNA are adenine (abbreviated A), cytosine (C), guanine (G) and thymine (T).
- Bases are classified into two; adenine and guanine are fused five- & six-membered heterocyclic compounds called purines
- Cytosine and thymine are six-membered rings called pyrimidines
- A fifth pyrimidine base, uracil U), usually takes the place of thymine in RNA & differs from it by lacking a methyl group on its ring.
- Uracil is not usually found in DNA and occurrs only as a breakdown product of cytosine.

MAJOR AND MINOR GROOVES

- The double helix is a right-handed spiral. As the DNA strands wind around each other, they leave gaps between each set of phosphate backbones, revealing the sides of the bases inside.
- There are two of these grooves twisting around the surface of the double helix: one groove, the major groove, is 22 Å wide and the other, the minor groove, is 12 Å wide.
- The narrowness of the minor groove means that the edges of the bases are more accessible in the major groove.
- As a result, proteins like transcription factors that can bind to specific sequences in double-stranded DNA usually make contacts to the sides of the bases exposed in the major groove.

BASE PAIRING

- Purines form hydrogen bonds to pyrimidines, with A bonding only to T, and C bonding only to G.
- Arrangement of 2 nucleotides binding together across the double helix is called a base pair.
- The double helix is also stabilized via forces generated by the hydrophobic effect and Pi stacking, which are not influenced by the sequence of the DNA.
- As hydrogen bonds are not covalent, they can be broken and rejoined relatively easily.
- The two strands of DNA in a double helix can therefore be pulled apart like a zipper, either by a mechanical force or high temperature.
- As a result of this complementarity, all the information in the double-stranded sequence of a DNA helix is duplicated on each strand, which is vital in DNA replication.
- Indeed, this reversible and specific interaction between complementary base pairs is critical for all the functions of DNA in living organisms.
- Two types of base pairs form different numbers of hydrogen bonds, AT forming two hydrogen bonds, and GC forming three hydrogen bonds.
- The GC base pair is therefore stronger than the AT base pair.
- As a result, it is both the percentage of GC base pairs and the overall length of a DNA double helix that determine the strength of the association between the two strands of DNA.
- Long DNA helices with a high GC content have stronger-interacting strands, while short helices with high AT content have weaker-interacting strands.
- When all the base pairs in a DNA double helix melt, the strands separate and exist in solution as two entirely independent molecules.

FORMS OF DNA

- DNA exists in several possible conformations.
- The conformations so far identified are: A-DNA, B-DNA, C-DNA, D-DNA, E-DNA,H-DNA,L-DNA,P-DNA, and Z-DNA.
- However, only A-DNA, B-DNA, and Z-DNA occur in biological systems.
- Which conformation DNA adopts depends on the sequence of the DNA, the amount and direction of supercoiling, chemical modifications of the bases and also solution conditions, such as the concentration of metal ions and polyamines.
- Of these three conformations, the "B" form described above is most common under the conditions found in cells & is the original DNA.
- In B form, the helix makes a turn every 3.4 nm, and the distance between two neighboring base pairs is 0.34 nm. Hence, there are about 10 base pairs per turn.
- The strands make major groove and minor groove, which may facilitate binding with specific proteins
- In a solution with higher salt concentrations or with alcohol added, the DNA structure may change to an A form, which is still right-handed, but every 2.3 nm makes a turn and there are 11 base pairs per turn.
- Another DNA structure is called the Z form, because its bases seem to zigzag.
- Z DNA is left-handed. One turn spans 4.6 nm, comprising 12 base pairs. The DNA molecule with alternating G-C sequences in alcohol or high salt solution tends to have such structure.
- The two alternative double-helical forms of DNA differ in their geometry and dimensions.

- A form is a wider right-handed spiral, with a shallow, wide minor groove and a narrower, deeper major groove.
- A form occurs under non-physiological conditions in dehydrated samples of DNA, while in the cell it may be produced in hybrid pairings of DNA and RNA strands, as well as in enzyme-DNA complexes.
- Segments of DNA where the bases have been chemically-modified by methylation may undergo a larger change in conformation and adopt the Z form.
- These unusual structures can be recognized by specific Z-DNA binding proteins and may be involved in the regulation of transcription.

B DNA

- This is the most common DNA conformation found *in vivo* . It is relatively narrow and more elongate than A or Z forms. Phosphodiester backbone runs in a smooth curve. The base pairs are almost perpendicular to the axis of the double helix. The wide major groove easily accessible to various polypeptides that affect level of transcription. Small kinks along the length of a B-DNA strand represent areas where a transcription factor is bound. B-DNA occur in actively transcribed genes

A DNA

- RNA-DNA hybrids and double-stranded RNA are most often found in this form. The helix is shorter and of greater diameter than in B-DNA. phosphodiester backbone runs in a smooth curve. Base pair H-bonds are slanted relative to the axis of the double helix. Major groove has become deep and narrow, making it less accessible to proteins that might affect transcription. Genes are not actively transcribed.

Z DNA

- Helix spirals to the left, instead of the right, as it does in B and A DNA. This isomer can form *in vivo*, but function is not fully understood. Helix is more narrow and elongate than either A-DNA or B-DNA. Major "groove" is so shallow as to no longer form a groove. Minor groove very narrow. Base pairs nearly perpendicular to axis of double helix. Phosphodiester backbone forms a zig-zagg pattern because of unusual sugar configuration. Genes in this configuration are believed to be transcriptionally inactive. {Z-DNA is stable *in vitro* when its cytosines are methylated (the 5-carbon H is replaced by CH3)}. Methylation protects DNA from the action of endonucleases.

SUPER HELICAL DNA/DNA SUPERCOIL

- DNA can be twisted like a rope in a process called DNA supercoiling
- With DNA in its "relaxed" state, a strand usually circles the axis of the double helix once every 10.4 base pairs, but if the DNA is twisted the strands become more tightly or more loosely wound.
- If the DNA is twisted in the direction of the helix, this is positive supercoiling, and the bases are held more tightly together.
- If they are twisted in the opposite direction, this is negative supercoiling, and the bases come apart more easily.
- In nature, most DNA has slight negative supercoiling that is introduced by enzymes called topoisomerases
- These enzymes are also needed to relieve the twisting stresses introduced into DNA strands during processes such as transcription and DNA replication.

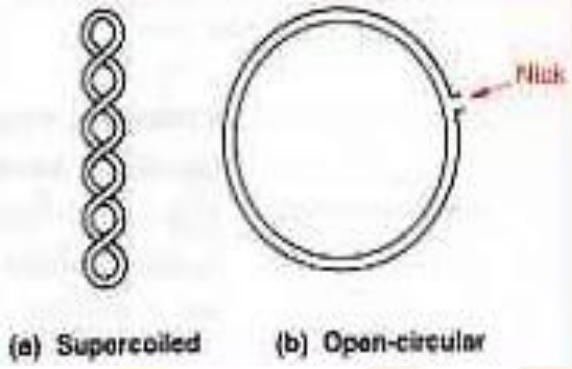
CIRCULAR DNA

- Circular DNA is a form of DNA that is found in bacteria and archea as well as in eukaryotic cells in the form of mitochondrial DNA.

- While the individual strands of a linear double helix represent two distinct and separable molecules, this need not be true for circular DNA.
- If the strands twist a number of times around one another in completing the DNA loop, then they are covalently joined into a single molecule.
- Majority of plasmids, prokaryotic and organellar (mitochondrial & plastid) DNAs, form a covalently closed structure, which is circular in shape and the genetic map derived from it will also be a circular genetic map.

NICKED DNA

- A nick is a point in a double stranded DNA molecule where there is no phosphodiester bond between adjacent nucleotides of one strand typically through damage or enzyme action. This can be caused by disruption of the phosphodiester bond in one of the chains of a double stranded nucleic acid by an endonuclease enzyme.



TRIPLE-HELIX DNA

- DNA triple helices form in a sequence-specific manner on polypurine : polypyrimidine tracts, which are widespread in mammalian genomes.
- The third strand lies in the major groove of an intact duplex and is stabilized by two Hoogsteen hydrogen bonds between third strand bases and the purines in the duplex.
- The third strand may consist of pyrimidines, or purines, depending on the nature of the target sequences.
- In the pyrimidine motif, a homopyrimidine oligonucleotide binds in a direction parallel to the purine strand in the duplex, with base triplets of T.A:T and C.G:C.
- In the alternate purine motif, a homopurine strand binds antiparallel to the purine strand, with base triplets of A.A:T and G.G:C.

SATELLITE DNA

- It consists of highly repetitive DNA, and is so called because repetitions of a short DNA sequence tend to produce a different frequency of the nucleotides adenine, cytosine, guanine and thymine, and thus have a different density from bulk DNA - such that they form a second or 'satellite' band when genomic DNA is separated on a density gradient.
- Types of satellite DNA: Satellite DNA, together with Minisatellite and Microsatellite DNA constitute the Tandem repeats.
- A repeated pattern can be between 1 base pair long (a mononucleotide repeat) to several thousand base pairs long, and the total size of a satellite DNA block can be several megabases without interruption.
- Most satellite DNA is localized to the telomeric or the centromeric region of the chromosome.
- The nucleotide sequence of the repeats is fairly well conserved across a species. However, variation in the length of the repeat is common in all organisms.

- There is remarkable variability in genome size among eukaryotes that has little correlation with organismal complexity, ploidy or no. of coding genes.
- Much of this variation is due to non-coding, tandemly repeated DNA.
- Indeed, a substantial fraction of the genomes of many eukaryotes is composed of repetitive DNA in which short sequences are tandemly repeated in small to huge arrays.
- Tandemly repetitive sequences, commonly known as "satellite DNAs" are classified into 3 major groups
- 1 – Satellites: They are very highly repetitive with repeat lengths of one to several thousand base pairs. These sequences typically are organized as large (up to 100 million bp) clusters in the heterochromatic regions of chromosomes, near centromeres and telomeres; these are also found abundantly on the Y chromsome.
- 2 – Minisatellites are moderately repetitive (20-50 times), tandemly repeated arrays of moderately-sized (9 to 100 bp, but usually about 15 bp) repeats, generally involving mean array lengths of 0.5 to 30 kb. They are found in euchromatic regions of the genome of vertebrates, fungi and plants and are highly variable in array size.
- 3 – Microsatellites are moderately repetitive, and composed of arrays of short (2-6 bp) repeats found in vertebrate, insect and plant genomes. The human genome contains at least 30,000 microsatellite loci located in euchromatin. Copy numbers are characteristically variable within a population, typically with mean array sizes on the order of 10 to 100.
- In general, satellite DNAs show exceptional variability among individuals, particularly with regard to the number of repeats at a given locus.
- Minisatellite loci are the most highly polymorphic sequence elements yet discovered in the human genome, and delineating the repeat lengths of these loci is the basis of most DNA typing systems used in forensic medicine

SELFISH DNA

- Selfish DNA are those sequences of DNA which, have two distinct properties: (1) the DNA sequence spreads by forming additional copies of itself within the genome; and (2) it makes no specific contribution to the reproductive success of its host organism.
- Richard Dawkins (1976) described it in the book *The Selfish Gene*.
- Selfish DNA can be considered an efficient replicator that follows another way of increasing in number.
- Selfish DNA is non-transcribed, non-coding, and contributes nothing to the well-being of the organism.
- In most cases selfish DNA is selectively neutral.
- Once it arises, selfish DNA is passively replicated and passed on from parent to offspring.
- Changes in its frequency in the population are due to genetic drift.
- Selfish DNA accumulates in parts of the genome where is does not interfere with genetic regulation and transcription.
- There are two types of selfish DNA:

  • Passive DNA. The sequence itself might not influence the genome, that it spreads among the genetic DNA and is retained; it is a passive kind of selfish DNA. It could accumulate as 'junk DNA' in the genome.

  • Parasitic DNA. A particular sequence might have a better than average chance of spreading through the genetic DNA; these sequences would be a more active, parasitic kind of selfish DNA and would proliferate and affect the genome, until checked by natural selection.

Ribonucleic acid (RNA)

- Ribonucleic acid, or RNA is one of the major biological macromolecules that are essential for all known forms of life (along with DNA and proteins)
- Each nucleotide in RNA contains a ribose sugar, with carbons numbered 1‘through 5'.
- A base is attached to the 1‘ position, in general, adenine (A), cytosine (C), guanine(G), and uracil (U).
- Back bone is sugar and phosphate group
- Nitrogenous bases linked to sugar moiety project from the backbone
- Nitrogenous bases are linked to pentose sugar through N-glycosidic linkage to form a nucleoside
- Phosphate group is linked with 3’OH of nucleoside through phosphoester linkage
- 2 nucleotides are linked through 3’-5’-phosphodiester linkage to form a dinucleotide
- More and more such groups will be linked to form a polynucleotide chain
- Such a polymer has a free phosphate moiety at 5’ end of ribose sugar and it is called as 5’-end of polynucleotide chain
- At other end, ribose has free 3’-OH group which is called as the 3’-end of polynucleotide chain
- In RNA, every nucleotide has an additional-OH present at 2’-position of ribose

RNA synthesis

- Synthesis of RNA is usually catalyzed by an enzyme—RNA polymerase
- By using DNA as a template
- The process is known as transcription
- There are also a number of RNA-dependent RNA polymerases that use RNA as their template for synthesis of a new strand of RNA
- A number of RNA viruses (such as poliovirus) use this type of enzyme to replicate their genetic material

Types of RNA

Messenger RNA (mRNA)

- Messenger RNA (mRNA) carries information about a protein sequence to the ribosomes, the protein synthesis factories in the cell
- It is coded so that every three nucleotides (a codon) correspond to one amino acid
- In eukaryotic cells, once precursor mRNA (premRNA) has been transcribed from DNA, it is processed to mature mRNA
- This removes its introns—non-coding sections of the pre-mRNA
- The mRNA is then exported from the nucleus to the cytoplasm, where it is bound to ribosomes and translated into its corresponding protein form with the help of tRNA
- mRNA – It carries genetic formation of DNA (Gene) for protein synthesis from nucleus to ribosome in the form of genetic code

Ribosomal RNA (rRNA)

- Ribosomal RNA (rRNA) is the catalytic component of the ribosomes.
- Eukaryotic ribosomes contain four different rRNA molecules: 18S, 5.8S, 28S and 5S rRNA.
- Three of the rRNA molecules are synthesized in the nucleolus, and one is synthesized elsewhere.
- In the cytoplasm, ribosomal RNA and protein combine to form a nucleoprotein called a ribosome.
- The ribosome binds mRNA and carries out.protein synthesis
- Several ribosomes may be attached to a single mRNA at any time.

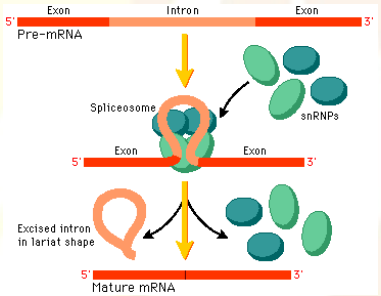
- Nearly all the RNA found in a typical eukaryotic cell is rRNA.
- It makes complex with proteins and form ribosomal subunits which provide space for protein synthesis ,single ribosomal RNA of smaller subunit helps correct orientation of mRNA during attachment with respect to P and A sites

Transfer RNA (tRNA)

- Transfer RNA (tRNA) is a small RNA chain of about 80 nucleotides
- It transfers a specific amino acid to a growing polypeptide chain at the ribosomal site of protein synthesis during translation
- It has sites for amino acid attachment and an anticodon region for codon recognition that binds to a specific sequence on the messenger RNA chain through hydrogen bonding
- Acts as adapter molecule, carries Amino Acid and drops it to particular location by recognising codon on mRNA by virtue of having anticodon

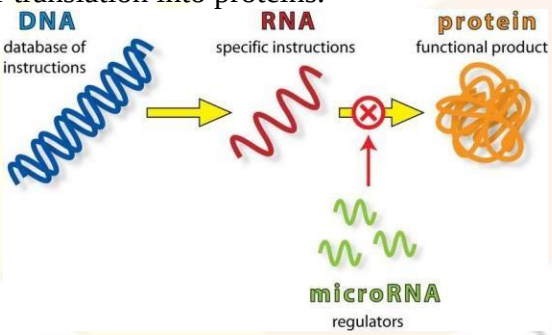
Small Nuclear RNAs (snRNAs)

- Sn RNA s are involved in the process of splicing (intron removal) of primary transcript to form mature m RNA. The Sn RNA s form complexes with proteins to form Ribonucleoprotein particles called snRNPs.



Micro RNAs (miRNAs)

- MicroRNAs, short non-coding RNAs present in all living organisms, have been shown to regulate the expression of at least half of all human genes.
- These single-stranded RNAs exert their regulatory action by binding messenger RNAs and preventing their translation into proteins.



Small Interfering RNAs (siRNAs)

- Small interfering RNA (siRNA) are 20-25 nucleotide-long double-stranded RNA molecules that have a variety of roles in the cell.
- They are involved in the RNA interference (RNAi) pathway, where it interferes with the expression of a specific gene by hybridizing to its corresponding RNA sequence in the target mRNA. This then activates the degrading mRNA.
- Once the target mRNA is degraded, the mRNA cannot be translated into protein.

Guide RNA (gRNA)

- It is RNA genes that function in RNA editing, found in mitochondria by inserting or deleting stretches of uridylates (Us).

- The gRNA forms part of editosome and contain sequences to hybridize to matching sequences in the mRNA to guide the mRNA modifications.

Complementary RNA (cRNA)

- Viral RNA that is transcribed from negative sense RNA and serves as a template for protein synthesis.

Negative sense RNA

- Viral RNA with a base sequence complementary to that of mRNA during replication it serves as a template to the transcription of viral complementary RNA

snoRNA – Plays role in gene silencing

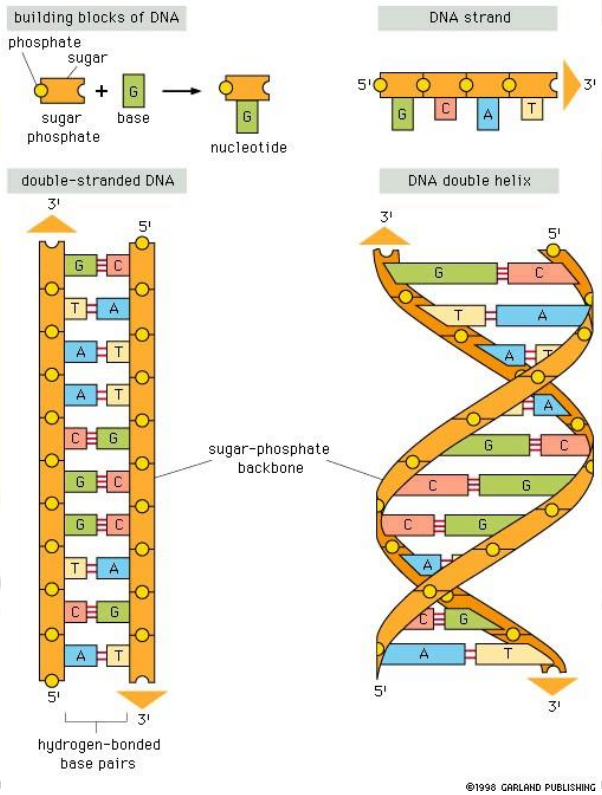
hnRNA - heterogeneous nuclear precursors & intermediates of mature mRNAs & other RNAs

Catalytic RNA –Ribozymes act as protein enzymes in catalyzing removal of intron, peptide bond formation etc

DNA Replication

-DNA is a Double Helix

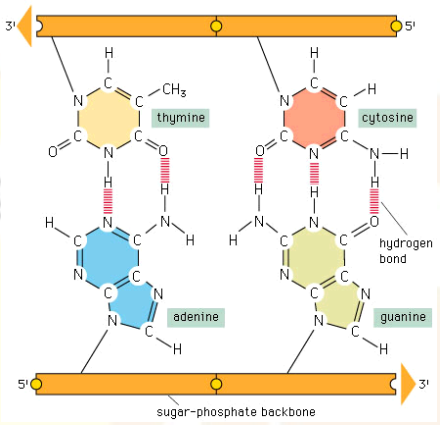
- Nucleotides
  - A, G, T, C
- Sugar and phosphate form the backbone
- Bases lie between the backbone
- Held together by
  H-bonds between the bases
  - A-T – 2 H bonds
  - G-C – 3 H bonds



H – Bonds

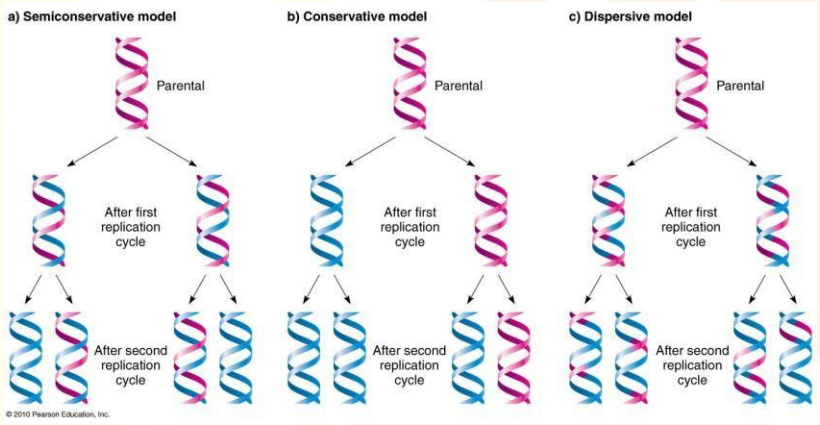
- **Base-pairing rules**
  - **A→T only (A→U if DNA-RNA hybrid)**
  - **G→C only**
- **DNA strand has directionality – one end is different from the other end**

- **2 strands are anti-parallel, run in opposite directions**
  - **Complementarity results**
  - **Important to replication**

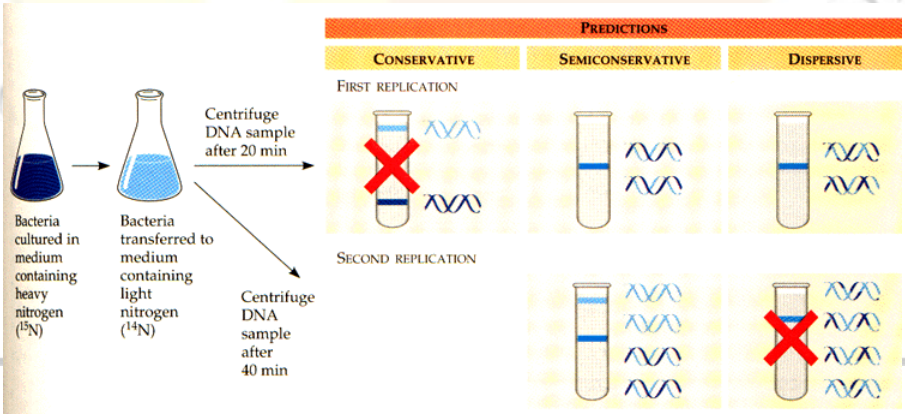


DNA Replication

- Purpose: cells need to make a copy of DNA before dividing so each daughter cell has a complete copy of genetic information
- 3 proposed Models of Replication



Meselson and Stahl Experiment



Semiconservative

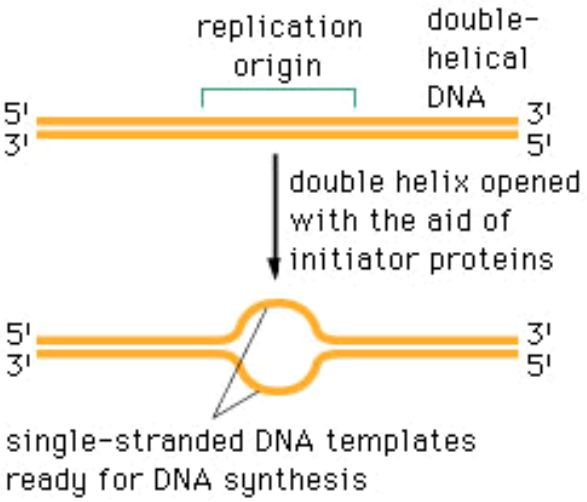
- Daughter DNA is a double helix with 1 parent strand and 1 new strand
- 1 strand serves as the template for new strand

DNA Template

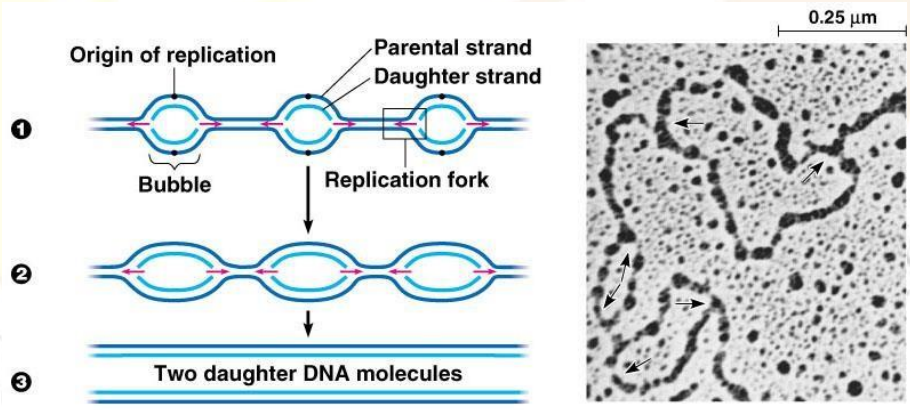
- Each strand of the parent DNA is used as a template to make the new daughter strand
- DNA replication makes 2 new complete double helices each with 1 old and 1 new strand

Replication Origin

- Site where replication begins
  - 1 in E. coli
  - 1,000s in human
- Strands are separated to allow replication machinery contact with the DNA
  - Many A-T base pairs because easier to break 2 H-bonds that 3 H-bonds



- Large team of enzymes coordinates replication

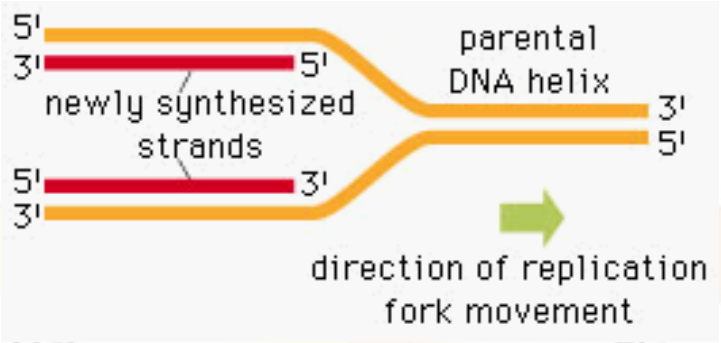


(a) In eukaryotes, DNA replication begins at many sites along the giant DNA molecule of each chromosome.

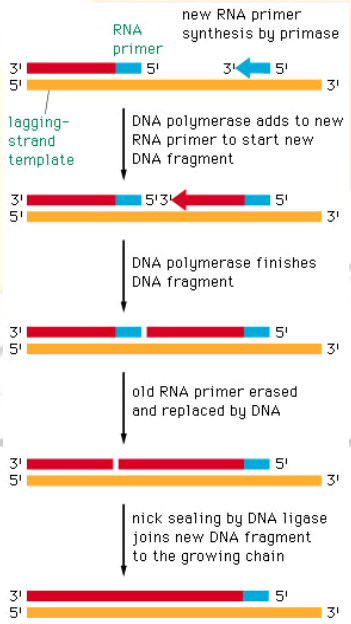
(b) In this micrograph, three replication bubbles are visible along the DNA of cultured Chinese hamster cells. The arrows indicate the direction of DNA replication at the two ends of each bubble (TEM).

- Bidirectional movement of the DNA replication machinery

How is DNA Synthesized?



- Replication fork
- Leading & Lagging strands
  - Simple addition of nucleotides along one strand, as expected
  - Called the leading strand
  - DNA polymerase reads 3' → 5' along the leading strand from the RNA primer
  - Synthesis proceeds 5' → 3' with respect to the new daughter strand
  - Other daughter strand is also synthesized 5'→3' because that is only way that DNA can be assembled
  - However the template is also being read 5'→3'
  - Compensate for this by feeding the DNA strand through the polymerase, and primers and make many short segments that are later joined (ligated) together
  - Called the lagging strand
- **Primers**
  - Simple addition of primer along leading strand
  - RNA primer synthesized 5' → 3', then polymerization with DNA
  - Many primers are needed along the lagging strand
  - 1 primer per small fragment of new DNA made along the lagging strand Called Okazaki fragments \



Replication: 1st step

- Unwind DNA
    - Helicase enzyme
        - unwinds part of DNA helix
        - stabilized by single-stranded binding proteins
            - PREVENTS DNA MOLECULE FROM CLOSING!
        - DNA gyrase
            - Enzyme that prevents tangling upstream from the replication fork

Replication: 2nd step

- RNA Primase
    - Adds small section of RNA (RNA primer) to the 3' end of template DNA
        - DNA polymerase 3 (enzyme that builds new DNA strand) can only add nucleotides to existing strands of DNA

Replication: 3rd step

- Build daughter DNA strand
    - add new complementary bases
    - With the help of the enzyme DNA polymerase III
    - DNA Polymerase:- An enzyme that catalyzes the addition of a nucleotide to the growing DNA chain
    - Nucleotide enters as a nucleotide tri-PO_4
    - 3'–OH of sugar attacks first phosphate of tri-PO_4 bond on the 5' C of the new nucleotide

Replication: 4th step

- Replacement of RNA primer by DNA
- Done by DNA polymerase I
- Other enzymes needed to excise (remove) the primers
- Nuclease – removes the RNA primer nucleotide by nucleotide
- Repair polymerase – replaces RNA with DNA
- DNA ligase – seals the sugar-phosphate backbone by creating phosphodiester bond
- Requires Mg^{2+} and ATP

Genomic library

DNA LIBRARY

- The term "library" can refer to a population of organisms, each of which carries a DNA molecule inserted into a cloning vector, or alternatively to the collection of all of the cloned vector molecules.
- Collection of DNA fragments that have been cloned into vectors so that researchers can identify and isolate the DNA fragments that interest them for further study.
- A gene library is a collection of different DNA sequences from an organism, which has been also called genomic libraries or gene banks.
- Cloned into a vector for ease of purification, storage and analysis.

Depending upon the source of the DNA used

- Genomic Library (Made from genomic DNA)

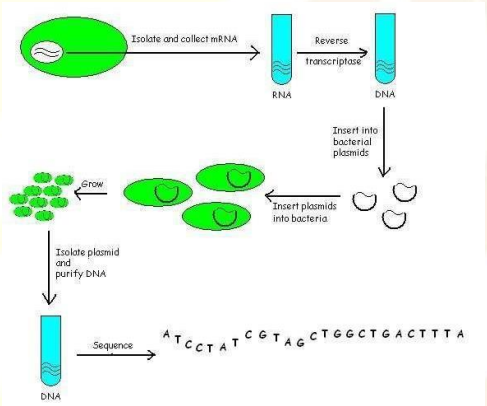
- cDNA library (Made from cDNA- copy of mRNA)

Construction of DNA library

- Size of the gene
- Capacity of the vector
- Molecular tools
- Vectors
- Genomic DNA library: Contains the whole genome of an organism
- Genome size is expressed in terms of number of base pairs.

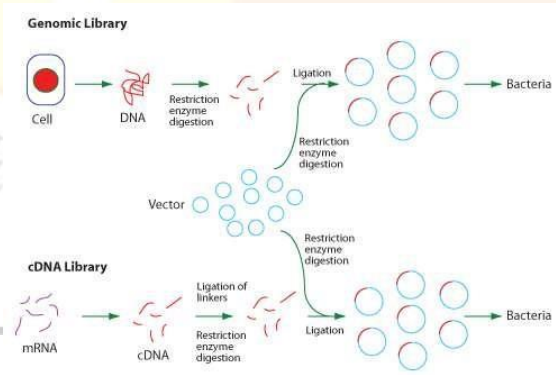
cDNA library

- cDNA library contains only complementary DNA molecules synthesized from mRNA molecules in a cell.
- cDNA is produced from fully transcribed mRNA found in the nucleus
- Therefore contains only the expressed genes of an organism.
- cDNA is created from a mature mRNA from a eukaryotic cell with the use of an enzyme known as reverse transcriptase.



Uses of cDNA library

- Used when reproducing eukaryotic genomes as the amount of information is reduced to remove the large number of non-coding regions from the library.
- To express eukaryotic genes in prokaryotes.
- Useful for subsequently isolating the gene that codes for that mRNA.



GENE EXPRESSION & REGULATION

Organisms adapt to environmental changes by altering gene expression. The regulation of gene expression of genes is necessary for the: - growth - development - differentiation & - very existence of the organism.

Types of Gene Expression: Mainly 2 types of gene expression & regulation:-

a. Positive regulation

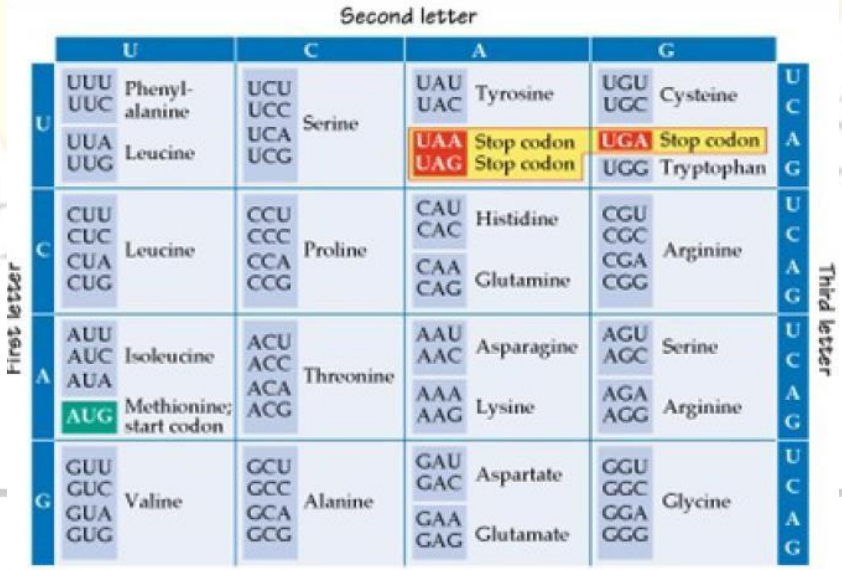
b. Negative regulation.

One gene-one enzyme Hypothesis

Earlier hypothesis proposed that one gene produces one enzyme or protein and "one gene-one enzyme" concept was introduced. It is now know that some enzymes and protein molecules are composed of two or more non-identical subunits, which cannot be explained by "one gene-one enzyme" theory & so it is not valid. The "cistron" is now considered as the genetic unit coding for the structure of the subunit of an enzyme or protein molecule, acting as the smallest unit of genetic expression. Hence, the "one gene-one enzyme" idea might be more accurately regarded as "one cistron-one subunit concept".

What is Genetic code???

- Genetic code is a dictionary that corresponds with sequence of nucleotides and sequence of amino acids.
- Genetic code is a set of rules by which information encoded in genetic material(DNA or RNA sequences) is translated into proteins by living cells.
- Term given By ″ Goerge Gamow ″
- Genetic code is a Dictionary consists of "Genetic words" called CODONS.
- Each codon consists of three bases (triplet)
- There are 64 codons.
- 61 codons code for 20 amino acids found in protein.
- 3 codons do not code for any amino acid.



Type of codon

- Sense Codons : The codon that code for amino acid are called sense codon.
  Signal Codons: Those codons that code for¬ signal during protein synthesis are called signal codons. For Example:- AUG, UAA, UAG & UGA . There are two types of signal codons.
  - Start codons- Initiating Codon.
  - Stop codons - Terminating Codon

Characteristic of the genetic code

- Triplet code
- Comma less
- Nonoverlapping code
- Degenerate code
- Universality of code
- Non ambiguous

Mechanism

Gene Expression in Prokaryotes

Operon: The concept of operon was introduced by Jacob and Monod in 1961. Operon is defined as a segment of a DNA strand⎫ consisting of: Structure genes: A cluster of several structural genes, which carries the codons which can be translated into proteins.

Operator genes: One operator gene which has an overall control over the process of translation.

Regulator gene: A third gene called regulator gene is located sometimes at a distance from the operator gene on the same DNA strand. Regulator gene transcribe m-RNA which⎫ synthesizes "repressor protein" molecules which regulate the transcription. P site (promoter site): is situated between operator gene & regulator gene.

Lactose (Lac) operon

The "lac operon" is an inducible catabolic operon of E.coli. It consists of:

1. Structural genes: It carries three structural genes 'Z', 'Y' & 'A'. Code respectively for "β-galactosidase", "galactoside permease" & "thiogalactoside transacetylase".
2. Functions:
   β-galactosidase: hydrolyzes lactose (β-galactoside) to galactose and glucose.
   Permease: responsible for the transport of lactose into the cell.
   Acetylase: coded by 'A' gene is not known properly.
   The structural genes Z, Y & A transcribe to form a single large m-RNA with three independent translation units for te synthesis of the three distinct enzymes. Such a m-RNA coding for more than one protein is¬ called "polycistronic m-RNA" which is characteristics in prokaryotes. "Lac-operon" Structure

Catabolic repression and lac operon

• Catabolite gene activator protein is a dimer and acts as a positive regulator of many catabolic operons like the "lac operon" in E.coli.

• Attachment of "RNA polymerase" to the promoter site requires the presence of CAP bound to cAMP.

• Absence of glucose in the cell activates "adenylate cyclase" which catalyses the synthesis of cAMP from ATP. The later binds to CAP to form a "CAP Camp complex".
• Unlike free CAP, this complex binds to promoter site immediately.

Mutations

Silent Mutations: Single nucleotide change-A to G, same amino acid is incorporated. Mutation goes unnoticed.

Missense mutations : Single nucleotide change A to C- different amino acid incorporated. Loss of functional capacity of protein.

Non sense Mutation: Single nucleotide change from C to T, stop codon is generated (In m RNA represented by UAG), premature termination of chain, may be incompatible with life.

Frame shift Mutation: Insertion or removal of a base can alter the reading frame with the resultant incorporation of different amino acids.

MODULE I- BIOTECHNOLOGY

- Biotechnology = bios (life) + logos (study of or essence) Literally 'the study of tools from living things'
- The word "biotechnology" was first used in 1917 by Karl Ereky –Hungarian agricultural engineer

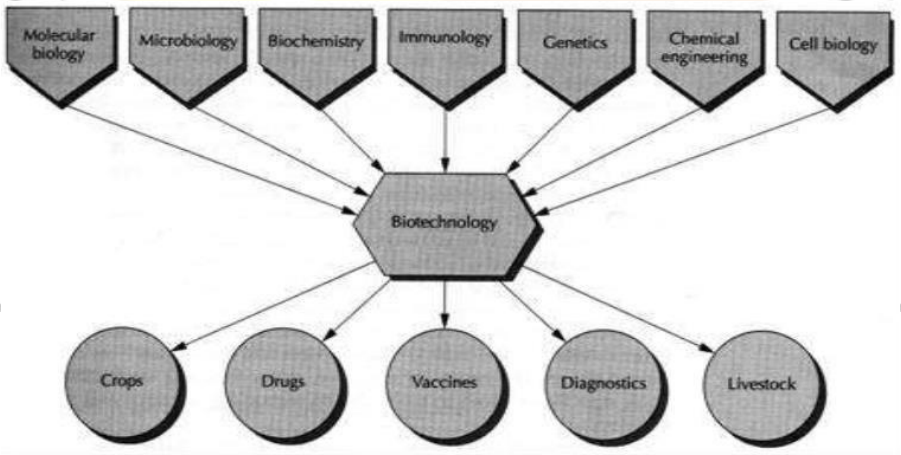
Definitions

- It is the use of biological process, organisms or systems to manufacture products intended to improve the quality of human life.
- **Biotechnology is the use of living systems and organisms to** develop or make useful products, or "any technological application that uses biological systems, living organisms or derivatives thereof, to make or modify products or processes for specific use.

Biotechnology is a multidisciplinarian in nature

• Cell and Molecular Biology
• Microbiology
• Genetics
• Physiology
• Biochemistry
• Immunology
• Virology

Biotechnology led to production of many products and provides many services for human welfare



Stages of biotechnology

- Ancient Biotechnology
  - Early history as related to food and shelter, including domestication
- Classical Biotechnology
  - Built on ancient biotechnology
  - Fermentation promoted food production
  - Medicine
- Modern Biotechnology
  - Manipulates genetic information in organism - genetic engineering

Principles of biotechnology

Two important techniques which enable development of modern biotechnology:

1. Alteration of chemistry of DNA & RNA to introduce into host organism to change phenotype of host- **Genetic engineering**

Recombinant DNA Technology: Genetic manipulation of bacteria, viruses, fungi, plants and animals, often forthe development of specific products

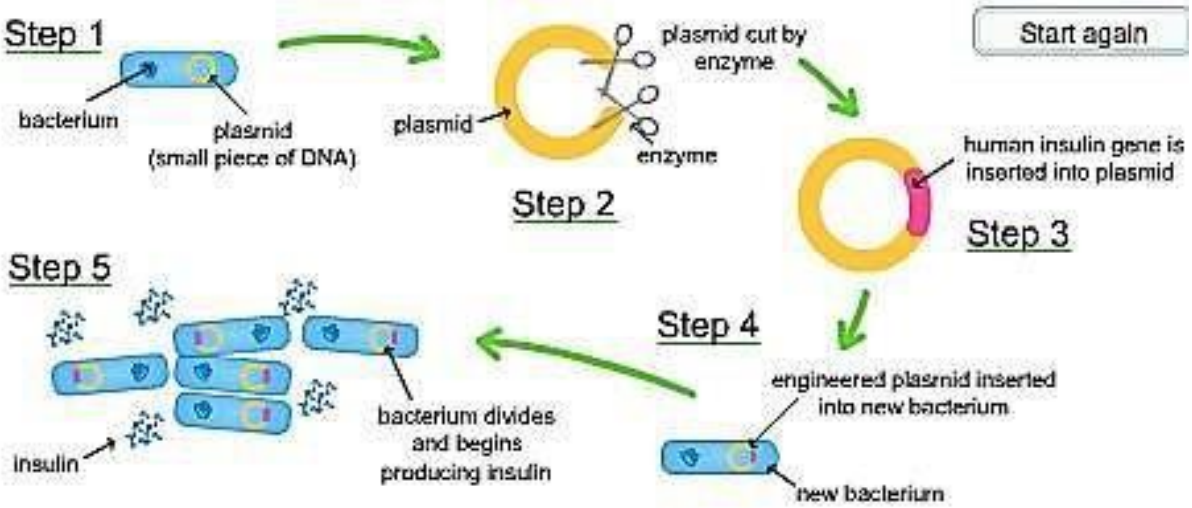
2. Maintenance of sterile ambience to enable growth of desired microbe/ eukaryotic cell in large quantities for manufacture of biotechnological products like *vaccine, enzymes, beverages, drugs etc.*- ***Chemical engineering***

IMPORTANT CONCEPTS

- **GENETIC ENGINEERING** : The techniques to alter the chemistry of genetic material (DNA and RNA) , to introduce these into host organisms and thus change the phenotype of the host organism
- **RESTRICTION ENZYMES**: Enzymes that are used to cut DNA segment at a specific site are called restriction enzymes. Eg.- EcoRI- isolated from Escherichia coli RY 13; R- name of strain, 'I' – order in which the enzyme were isolated from that strain of bacteria.
- **EXONUCLEASES:** Enzymes that remove nucleotides from the ends of the DNA molecule
- **ENDONUCLEASES:** Enzymes that make cuts at specific positions within the DNA molecule
- **PLASMID:** Autonomously replicating circular extra-chromosomal DNA of any bacteria
- **VECTORS:** May be plasmid DNA or viruses that can act as vehicle to transfer the piece of DNA attached to it.
- **PHAGE:** Virus that infects bacteria
- **Origin of Replication:** A specific DNA sequence which is responsible for initiating replication is called origin of replication.

Basic steps in GE

1. Identification of DNA with desirable gene
2. Introduction of the identified DNA into host
3. Maintenance of introduced DNA in the host and transfer of the DNA into its progeny



Biotechnology Timeline

- 1750 BC :The Sumerians brew beer
- 500 BC :Chinese use moldy soybean curds as an antibiotic to treat boils
- 100 AD: First insecticide-Powderd chrysanthemums(China)
- 1590: Janssen invents the microscope
- 1675 :Leeuwenhoek discovers cells (bacteria, red blood cells)
- 1797: First vaccination-Edward Jenner takes pus from a cowpox lesion, inserts it into an incision on a boy's arm
- 1830 :Proteins are discovered
- 1833 :The first enzymes are isolated
- 1855: The *Eschirium coli bacterium* is discovered
- 1859: Charles Darwin publishes *On the Origin of Species*
- 1864 :Louis Pasteur shows all living things are produced by other living things
- 1865 :The age of genetics begins
- 1902 Walter Sutton coins the term ‘gene’ - proposed that chromosomes carry genes
- 1910: Chromosomal theory of inheritance proved
- 1915:Phages were discovered: virus that infects bacteria
- 1928: Fleming discovers antibiotic properties of certain molds
- 1941:George Beadle and Edward Tatum propose that one gene makes one protein
- 1949: Sickle cell anaemia demonstrated to be molecular disease
- 1953 : James Watson and Francis Crick describe the double helical structure of DNA
- 1967 :The genetic code is cracked
- 1973 :Recombinant DNA technology begins
- 1975: First international conference on recombinant DNA technology
- 1975 :DNA sequencing discovered
- 1975: Monoclonal antibody technology introduced
- 1978:Genetic engineering to produce human insulin in *E.coli*
- 1978: Kary Mullis discovers PCR
- 1989 :The Human Genome Project begins
- 1990 :First use of gene therapy
- 1993 :FDA announces that transgenic food is safe
- 1994 The FLAVRSAVR tomato - first genetically engineered whole food
- 1996: First mammal cloned from adult cells- Ian Wilmut-Dolly

- 1997: First artificial chromosome
- 1999 :Completion of Drosophilia genome sequence
- 2001 Human genome project complete

Applications of biotechnology

- Production of new and improved crops/foods, industrial chemicals, pharmaceuticals and livestock
- Diagnostics for detecting genetic diseases
- Gene therapy
- Vaccine development (recombinant vaccines)
- Environmental restoration
- Protection of endangered species
- Conservation biology
- Bioremediation
- Forensic applications

Module I-What is DNA?

- DNA= Deoxyribo-Nucelic Acid
- DNA is a very large molecule, made up of smaller units called nucleotides.
- Each nucleotide has three parts: a sugar (ribose), a phosphate molecule, and a nitrogenous base.
- The nitrogenous base is the part of the nucleotide that carries genetic information.
- The bases found in DNA are four: adenine, cytosine, guanine, and thymine (ATP, CTP, GTP, and TTP)

Recombinant DNA Technology

Recombinant DNA technology procedures by which DNA from different species can be isolated cut and spliced together -- new "recombinant" molecules are then multiplied in quantity in populations of rapidly dividing cells (e.g. bacteria, yeast).

- DNA molecules that are extracted from different sources and chemically joined together.
- Production of a unique DNA molecule by joining together two or more DNA fragments not normally associated with each other
- DNA fragments are usually derived from different biological sources

Basic principles of rDNA technology:

- Generation of DNA fragments & selection of the desired piece of DNA.
- Insertion of the selected DNA into a cloning vector to create a rDNA or chimeric DNA.
- Introduction of the recombinant vectors into host cells.
- Multiplication & selection of clones containing the recombinant molecules.
- Expression of the gene to produce the desired product.

Recombinant DNA technology is based on a number of important things:

- A series of procedures used to recombine DNA segments.
- Under certain conditions, a recombinant DNA molecule can enter a cell and replicate.
- Recombinant DNA technology is one of the recent advances in biotechnology, which was developed by two scientists named Boyer and Cohen in 1973.
- The DNA is inserted into another DNA molecule called 'vector'.
- The recombinant vector is then introduced into a host cell, where it replicates itself, the gene is then produced.

Milestones in the development of Recombinant DNA Technology

- 1869: Miescher first isolate DNA from white blood cells harvested from pus-soaked bandages obtained from a nearby hospital.
- 1944: Avery provides evidence that DNA, rather than protein, carries the genetic information during bacterial transformation.
- 1953: Watson and Crick propose the double – helix for DNA structure based on X-ray results of Franklin and Wilkins
- 1955: Kornberg discovers DNA polymerase, the enzyme now used to produce labelled DNA probes
- 1961: Marmur and Doty discover DNA renatuuration, establishing the specificity and feasibility of nucleic acid hybridization reaction.
- 1962: Arber provides the first evidence for the existence of DNA restriction nucleases, leading to their purification and use in DNA sequence characterization by Nathans and H. Smith
- 1966:Nirenberg, Ochoa and Khorana elucidate the Genetic code
- 1967:Gellert discovers DNA ligase, the enzyme used to join DNA fragments together
- 1972-73: DNA cloning techniques are developed by the laboratories of Boyer, Cohen, Berg and their colleagues at Stanford University and the University of California at San Francisco.
- 1975: Southern develops gel-transfer hybridization for the detection of specific DNA sequences.
- 1975-1977 Sanger and Barrell and Maxam and Gilbert devlop rapid DNa sequencing methos.
- 1981-1982: Palmiter and Brinster produce transgenic mice; Spradling and Rubin produce transgenic fruit flies.
- 1985: Kary Mullis and Co-workers invent the polymerase chain reaction (PCR)
- 1990: Lipman and colleagues release BLAST, an algoritham used to search for homology between DNA and protein sequences.
- 1990: Simon and colleagues study how to efficiently how to efficiently use bacterial artificial chromosomes, BACs to carry large pieces of cloned human DNA for sequencing.

- 1991: Hood and Hunkapillar introduce new automated DNA sequence technology.
- 1995: Venter and colleagues sequence the first complete genome, that of the bacterium *Haemophilus influenzae*.

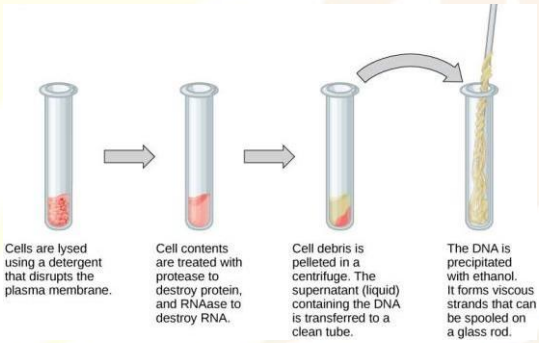
Goals of recombinant DNA technology

• To isolate and characterize a gene
• To make desired alterations in one or more isolated genes
• To return altered genes to living cells
• Artificially synthesize new gene
• Alternating the genome of an organism
• Understanding the hereditary diseases and their cure
• Improving human genome

Procedure of making rDNA

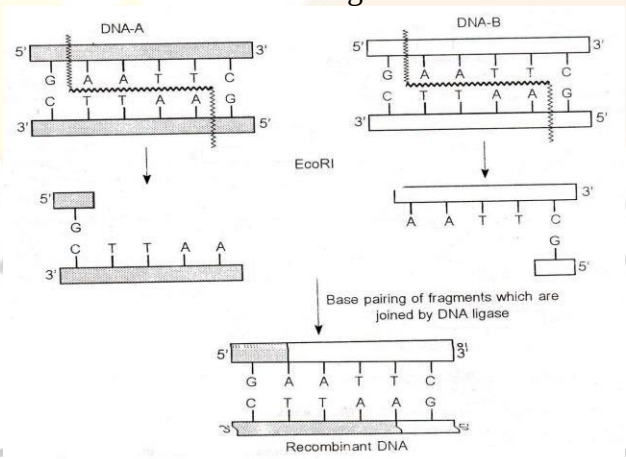
1. Identification of desired Gene/DNA
2. Isolating of DNA
3. Cutting of DNA
4. Joining of DNA
5. Amplifying of DNA

Isolating of DNA



Cutting of DNA

• DNA can be cut into large fragments by mechanical shearing.
• Restriction enzymes are the scissors of molecular genetics.



Restriction Endonucleases

- Enzymes for the manipulation of DNA.
- Are bacterial enzymes that can cut/split DNA at specific sites

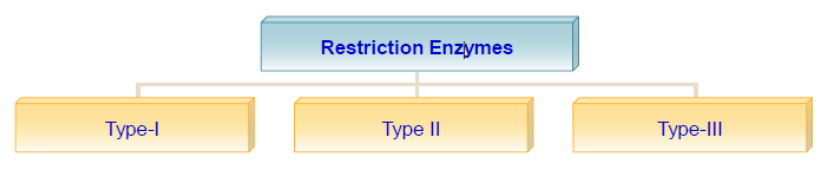
- These were first discovered in E.coli restricting the replication of bacteriophages,by cutting the viral DNA(The host E.coli DNA is protected from cleavage by addition of methyl groups).Thus, the enzymes that restrict the viral replication are known as restriction enzymes or restriction endonucleases.

Recognition sequences:

- Recognition sequence is the site where the DNA is cut by a restriction endonuclease.
- Restriction endonucleases can specifically recognise DNA with a particular sequence of 4-8 nucleotides & cleave.

Cleavage patterns:

- The cut DNA fragments by restriction endonucleases may have mostly sticky ends or blunt ends.
- DNA fragments with sticky ends are particularly useful for rDNA experiments,since single stranded sticky DNA ends can easily pair with any other DNA fragment having complementary



Type I

Don't Generate specific fragments

Require the presence of Mg+2,ATP and S-adenoyl methionane;the latter activates the enzyme.

Tracks along the DNA for a variable distance before breaking.

Type II

Cut within or immediately adjacent to target sequence

Generates specific fragments

Mg2+ essential

No ATP, No adenosyl methionine

Type III

Requires ATP and S-adenosyl methionine for Cleavage.

Make breaks in DNA 25 bp away from recognition site.

Examples of Restriction Enzymes

Microorganism	Restriction Enzyme Name	Restriction Site
Bacillus amyloliquefaciens H	*Bam*HI	G|G A T C C C C T A G|G
Brevibacterium albidum	*Bal*I	T G G|C C A A C C|G G T
Escherichia coli RY13	*Eco*RI	G|A A T T C C T T A A|G
Haemophilus aegyptius	*Hae*II	Pu G C G C|Py Py|C G G C Pu
Haemophilus aegyptius	*Hae*III	G G|C C C C|G G
Haemophilus influenzae R_d	*Hind*II	G T Py|Pu A C C A Pu|Py T G
Haemophilus influenzae R_d	*Hind*III	A|A G C T T T T C G A|A
Haemophilus parainfluenzae	*Hpa*I	G T T|A A C G A A|T T G
Haemophilus parinfluenzae	*Hpa*II	C|C G G G G C|C
Providencia stuartii 164	*Pst*I	C T G C A|G G|A C G T C
Streptomyces albus G	*Sal*I	G|T C G A C C A G C T|G

Joining DNA- DNA ligase

- These were originally isolated from viruses; also occur in E.coli & eukaryotic cells.
- The cut DNA fragments are covalently joined together by DNA ligases.
- DNA ligase joins the DNA fragments by forming a phosphodiester bond between the phosphate group of 5'-carbon of one deoxyribose with the hydroxyl group of 3'-carbon of another deoxyribose.

Enzymes used in recombinant DNA technology

Enzyme	Action
DNA ligase	Bind to DNA molecules
Type II restriction endonuclease	Cleaves DNA at specific sites
Reverse transcriptase Reverse transcriptase	Make a DNA copy of RNA molecule
DNA polymerase I	Fill single stranded gapes of DNA duplex
Polynycleotide Kinase	Adds a phosephate to the 5'-OH end of a polynucleotide
Terminal transferase	Adds homopolymer tails to the 3'-OH ends
Exonuclease III	Removes nucleotide residues from the 3' ends
Bacteriophage {lamda} exonuclease	removes nucleotides from the 5' ends
Alkaline phosphatase	Removes terminal phosphates

Vectors used in rDNA technology

A vector is an area of DNA that can join another DNA part without losing the limit for self replication

• Should be capable of replicating in host cell

• Should have convenient RE sites for inserting DNA of interest

• Should have a selectable marker to indicate which host cells received recombinant DNA molecule

• Should be small and easy to isolate

Plasmid vector

Plasmids are small, circular DNA molecules that are separate from the rest of the chromosome.

• They replicate independently of the bacterial chromosome.

• Useful for cloning DNA inserts less that 20 kb (kilobase pairs).

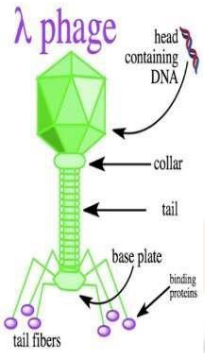
• Inserts larger than 20 kb are lost easily in the bacterial cell.

- Are extrachromosomal, double stranded, circular, self-replicating DNA molecules.
- Usually plasmids contribute to about 0.5%- 5.0% of the total DNA of bacteria.
- A few bacteria contain linear plasmids E.g., streptomyces sp, Borelia burgdorferi. E.g., pBR322,pUC
- The plasmids carries genes resistance for ampicillin & tetracycline that serve as markers for the identification of clones carrying plasmids.

Lamda phage vector

• Lamda phage vectors are recombinant infections, containing the phage chromosome in addition to embedded "outside" DNA.

• All in all, phage vectors can convey bigger DNA groupings than plasmid vectors.



Cosmid vector

• Cosmids are hybrids of phages and plasmids that can carry DNA fragments up to 45 kb.

• They can replicate like plasmids but can be packaged like phage lambda

Expression vectors

• Expression vectors are vectors that carry host signals that facilitate the transcription and translation of an inserted gene.

• They are very useful for expressing eukaryotic genes in bacteria.

Yeast artificial chromosomes (YAC)

• Yeast artificial chromosomes (YACS) are yeast vectors that have been engineered to contain a centromere, telomere, origin of replication, and a selectable marker.

• They can carry up to 1,000 kb of DNA.

• They are useful for cloning eukaryotic genes that contain introns.

Bacterial artificial chromosomes (BAC)

Bacterial artificial chromosomes (BACS) are bacterial plasmids derived from the F plasmid. They are capable of carrying up to 300 kb of DNA

Applications of rDNA technology

• Agriculture: growing crops of your choice (GM food), pesticide resistant crops, fruits with attractive colors, allbeing grown in artificial conditions

• Pharmacology: artificial insulin production, drug delivery to target sites

• Medicine: gene therapy, antiviral therapy, vaccination, synthesizing clotting factors

• Other uses: fluorescent fishes, glowing plants etc

Methods of Gene Transfer

Gene transfer is to transfer a gene from one DNA molecule to another DNA molecule. The directed desirable gene transfer from one organism to another and the subsequent stable integration & expression of foreign gene into the genome is referred as genetic transformation. Transient transformation occurs when DNA is not integrated into host genome. The transferred gene is known as transgene and the organism that develops after a successful gene transfer is known as transgenic.

Methods of Transformation

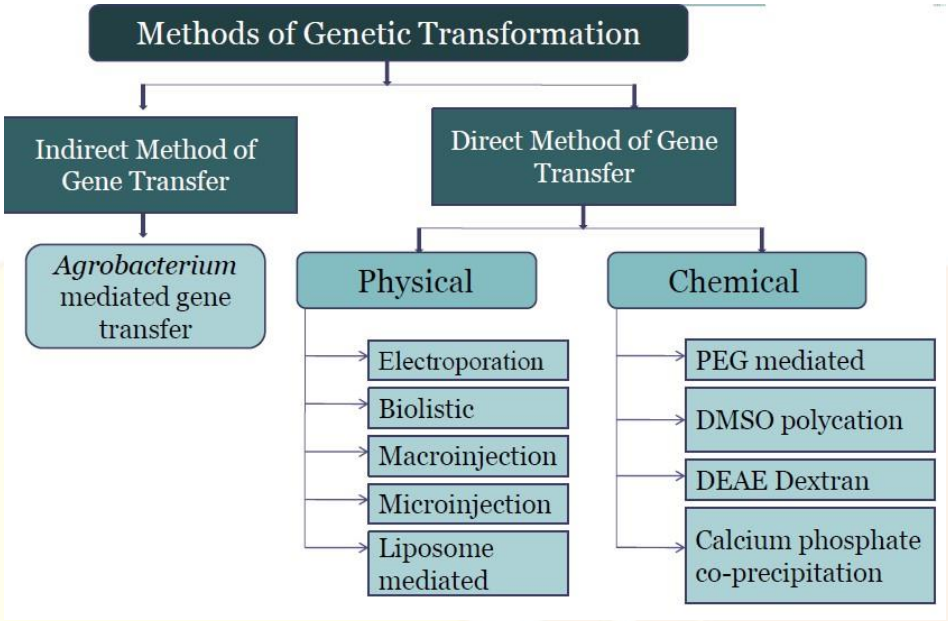
- There are mainly 2 methods of gene transfer:

Indirect (*Agrobacterium-mediated) gene transfer*

- Gene transfer is done by using the bacteria *Agrobacterium tumificiens.*

Direct gene transfer

- The gene is directly transferred into the host by using various techniques.



TECHNIQUES OF GENE ANALYSIS

Polymerase chain reaction

• PCR is a copying machine for DNA molecules
• Invented by Kary Mullis and his colleagues in the 1983
• Nobel prize 1993

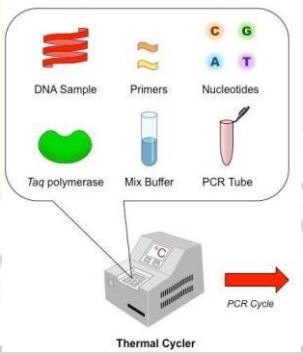
Purpose of PCR

• Polymerase chain reaction (PCR) is a technique used in molecular biology to amplify a single copy or a few copies of a segment of DNA across several orders of magnitude, generating thousands to millions of copies of a particular DNA sequence.
• Purpose of PCR Technique in vitro (test tube) is amplification of specific DNA sequences.
PCR allows a "target" DNA sequence to be selectively amplified, several million-fold in just a few hours.

- Exponential Amplification: The DNA sequence between primers **doubles after each cycle**

Requirements for PCR

- DNA Template
- Primers
- Nucleotide (dNTPs)
- Taq polymerase
- Mix Buffer
- PCR tube



DNA Template

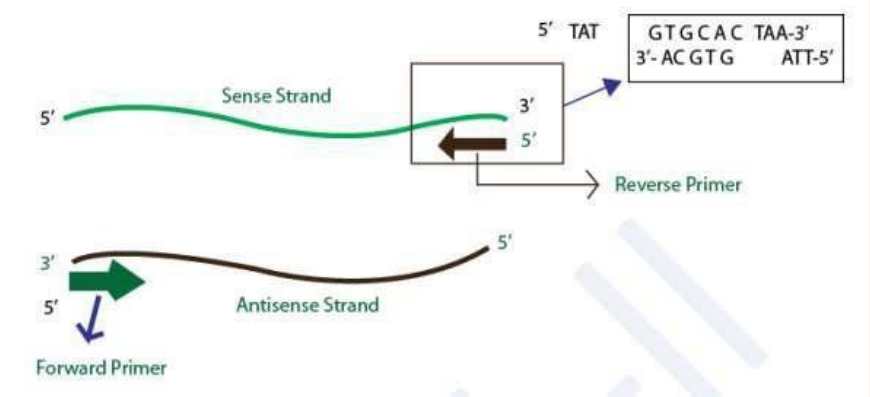
- DNA template is DNA target sequence.
- DNA template is the DNA molecule that contains the DNA region (segment) to be amplified, the segment we are concerned which is the target sequence.

Amount

- Human genomic DNA should be up to 500ng
- Bacterial DNA1-10ng
- Plasmid DNA0.1-1ng

Primers

- Typical primers are 18-28 bases in length
- Having 50-60% GC composition
- Have a balanced distribution of G/C and A/T rich domains
- Are not complementary to each other at the 3' ends to avoid primer- dimer forming artifacts.
- Not self complementary "Hairpin" formation
- Two primers must be designed for PCR; the forward primer and the reverse primer. The forward primer is complimentary to the 3' end of antisense strand (3'-5') and the reverse primer is complimentary to the 3' end of sense strand (5'-3'). If we consider the sense strand (5'-3') of a gene, for designing primers, then forward primer is the beginning of the gene and the reverse primer is the reverse-compliment of the 3' end of the gene.



Primer Design

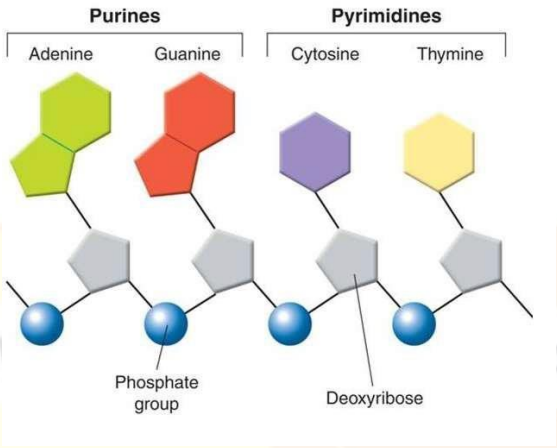
Primers should bind to template with good specificity and strength. If primers do not bind to correct template, wrong sequence will get amplified. Optimal primer sequences and appropriate primer concentrations are essential for maximal specificity and efficiency in PCR.

DNA polymerase

- DNA polymerase sequentially adds nucleotides complimentary to template strand at 3'-OH of the bound primers and synthesizes new strands of DNA complementary to the target sequence.
- The most commonly used DNA polymerase is Taq DNA polymerase (from Thermus aquaticus, a thermophillic bacterium) because of high temperature stability.

Nucleotides (dNTPs or deoxynucleotide triphosphates)

- Always use balanced solution of all four dNTPs to minimize polymerase error rate.
- All types of nucleotides are "building blocks" for new DNA strands and essential for reaction. It includes Adenine(A), Guanine(G), Cytosine(C), Thymine(T) or Uracil (U).



Buffer

The standard PCR buffer contains:

• Mg2+ ions in the buffer act as co-factor for DNA polymerase enzyme and hence are required for the reaction. Magnesium affects primer annealing and template denaturation, as well as enzyme activity. An excess of magnesium gives non-specific amplification products, while low magnesium yields lesser amount of desired product.

• Tris-HCl

• KCl

• Gelatin or Bovine Serum Albumin

PCR Master Mix

PCR Master Mix is a **premixed, ready-to-use** solution containing:

• Taq DNA Polymerase

• dNTPs

• MgCl2

• Reaction buffers

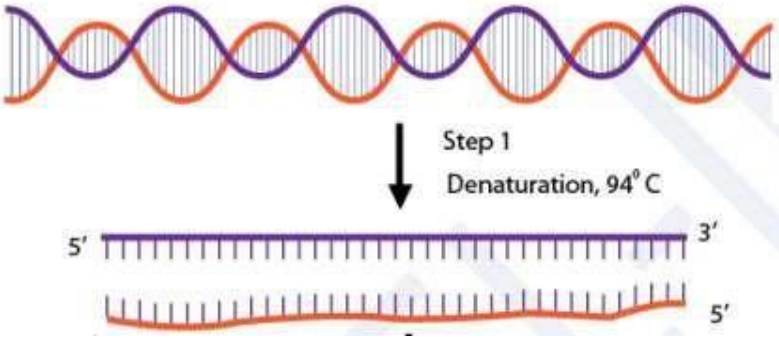
At **optimal concentrations for efficient amplification** of DNA templates by PCR.

PCR steps

- Each cycle includes three successive steps:
- There are three major steps in a PCR, which are repeated for 30 or 40 cycles.
- This is done on an automated cycler, which can heat and cool the tubes with the reaction mixture in a very short time.

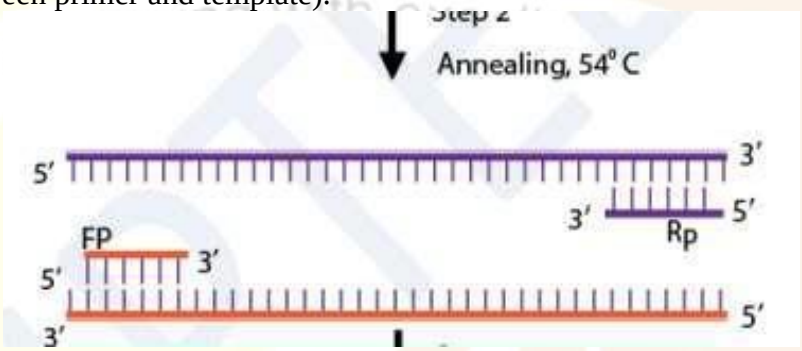
Denaturation at 94°C

- During the heating step (denaturation), the reaction mixture is heated to 94°C for 1 min, which causes separation of DNA double stranded. Now, each strand acts as template for synthesis of complimentary strand.



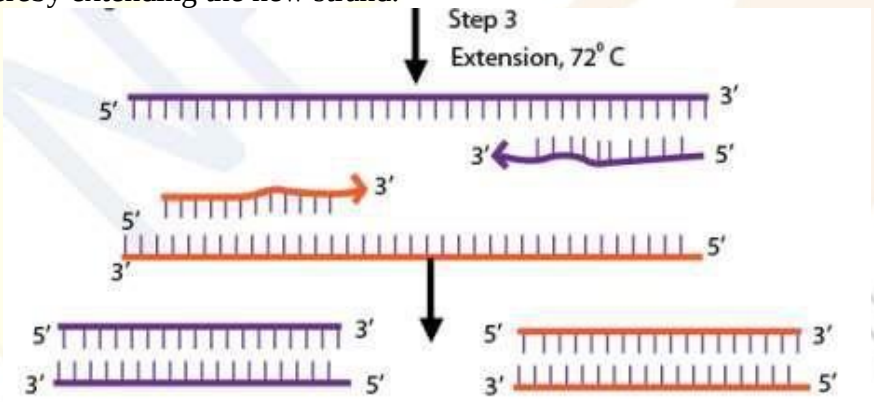
Annealing at 54°C

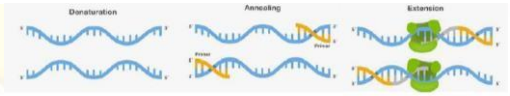
- This step consist of cooling of reaction mixture after denaturation step to 54°C, which causes hybridization (annealing) of primers to separated strand of DNA (template).
- The length and GC-content (guanine-cytosine content) of the primer should be sufficient for stable binding with template.
- Guanine pairs with cytosine with three hydrogen bonding adenine binds with thymine with two hydrogen bonds. Thus, higher GC content results in stronger binding. In case GC content is less, length may be increased to have stronger binding (more number of H bonding between primer and template).

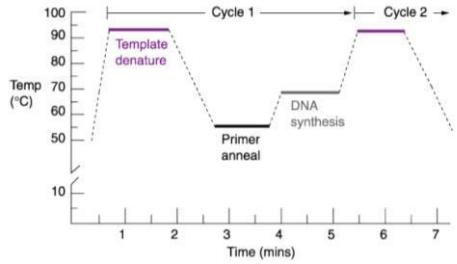


Extension at 72°C

- The reaction mixture is heated to 72°C which is the ideal working temperature for the Taq polymerase.
- The polymerase adds nucleotide (dNTP's) complimentary to template on 3' –OH of primers thereby extending the new strand.







Final hold

- First three steps are repeated 35-40 times to produce millions of exact copies of the target DNA. Once several cycles are completed, during the hold step, 4–15 °C temperature is maintained for short-term storage of the amplified DNA sample.

Number of Cycles

The number of cycles required for optimum amplification varies depending on the amount of the starting material.

• Most PCR should, therefore, include only 25 – 35 cycles. As cycle increases, nonspecific products can accumulate.

• After 20- 40 cycles of heating and cooling build up over a million copies of original DNA molecules.

Reaction Condition & Experimental Protocol

- Denaturing conditions are best at 94-95°C for 30-60 seconds. Lower temperatures may result in incomplete denaturation of target template and PCR products. Higher temperatures and a longer amount of time can lead to enzyme activity loss.

PCR machine

Gel Electrophoresis

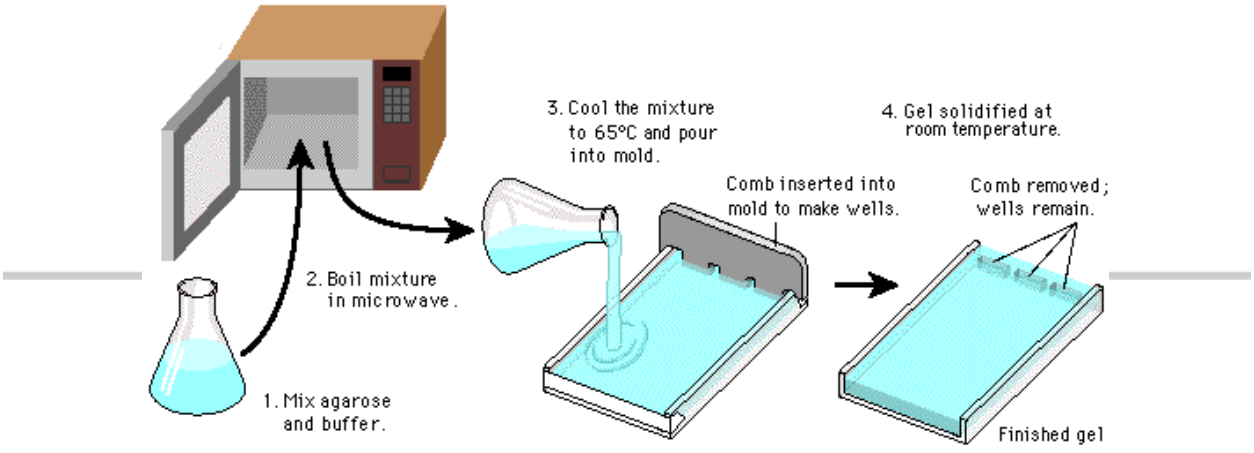
Agarose Gel Electrophoresis

• It is a method used in biochemistry and molecular biology to separate DNA, or RNA molecules based upon charge, size and shape.

• To determine the presence or amount of DNA

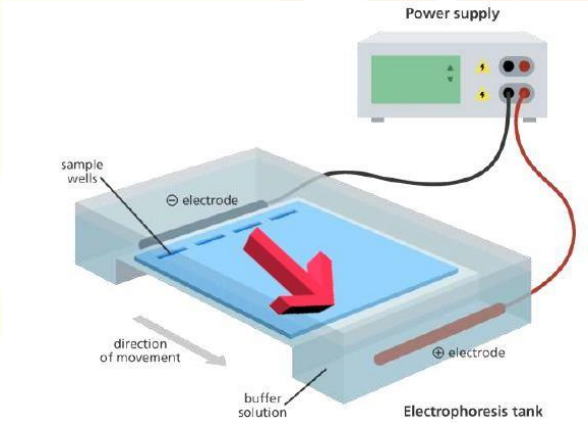
• To determine the sizes of DNA fragments

Preparation of Gel



Loading sample

Running the Electrophoresis



DNA has an overall negative charge. Because of Phosphate group.

Factors that affect the mobility of molecules in the gel

• Charge
• Shape
• Buffer conditions
• Gel concentration and
• Voltage
• Size

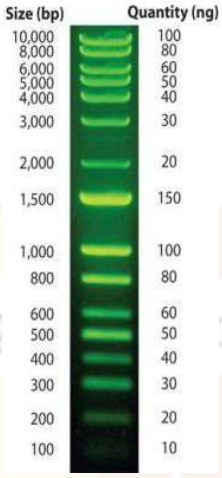
Reading results

Under the UV light/dyes are used:

- Ethidium Bromide Dye
- Green stain Dye
- Red Safe Dye

DNA ladder

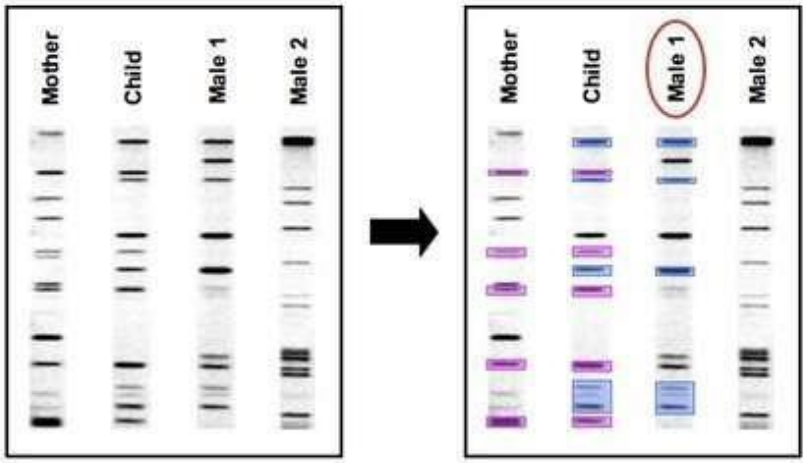
- DNA Ladder consists of known DNA sizes used to determine the size of an unknown DNA sample.



Module V-PCR Application

Paternity Testing

- Genetic material is inherited from both parents, half from mother and half from father.
- Matching genetic fingerprinting of suspected parents with that of children.
- DNA sample from buccal saliva or blood is collected and extracted from the alleged child. Then the extracted DNA is subjected to PCR, thousands of copies of amplified DNA is obtained.

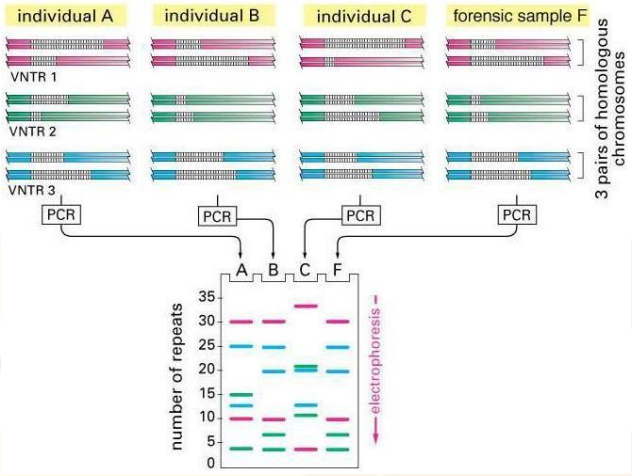


Diagnosis of Infectious Disease

- DNA from the infected parts of a person may be subjected to PCR with primer specific gene of the pathogen and diagnosis can be done on amplification of DNA.
- These PCR-based tests have advantage over conventional antibody-based diagnostic methods that determine the body's immune response to a pathogen.
- As patients can take weeks to develop antibodies against an contagious agent but this is fast and efficient.

DNA Fingerprinting

- PCR's main advantage in forensics is that forensic scientists can use it to amplify or make copies of regions of the genome that vary widely between different individuals, called VNTRs (variable number tandem repeats).
- By comparing the length of different VNTRs they can determine whether the sample may be a match with the suspect's DNA.



BLOTTING

- A blot, in molecular biology and genetics, is a method of transferring proteins, DNA or RNA, onto a carrier.
- The term "blotting" refers to the transfer of biological samples from a gel to a membrane and their subsequent detection on the surface of the membrane.
- Technique for transferring DNA ,RNA and Proteins onto a carrier so they can be separated, and often follows the use of a gel electrophoresis.

TYPES OF BLOTTING

Southern Blot: It is used to detect DNA
Northern Blot: It is used to detect RNA
Western Blot: It is used to detect protein

SOUTHERN BLOTTING

- A **Southern blot is a method used** in molecular biology for detection of a specific DNA sequence in DNA samples.
- Southern blotting combines transfer of electrophoresis -separated DNA fragments to a filter membrane and subsequent fragment detection by probe hybridization.
- The method is named after its inventor, the British biologist **Edwin Mellor Southern.**

PRINCIPLE

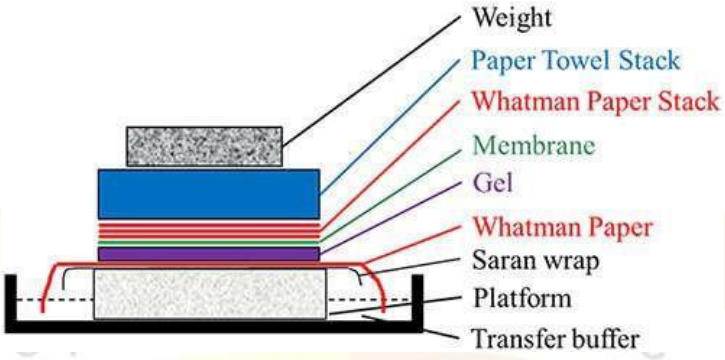
- **The key to this method is hybridization.**
- Hybridization: It is the process of forming a double stranded DNA molecule between a single-stranded DNA probe and a single-stranded target DNA.

There are 2 important features of hybridization:
• The reactions are specific-the probes will only bind to targets with a complementary sequence.
• The probe can find one molecule of target in a mixture of millions of related but non-complementary molecules.

STEPS INVOLVED IN SOUTHERN BLOTTING

1. Extract and purify DNA from cells.
2. DNA is restricted with enzymes.
3. Separated by electrophoresis.
4. Denature DNA.
5. Transfer to nitrocellulose paper.
6. Add labeled probe for hybridization to take place.
7. Wash off unbound probe.
8. Autoradiograph.

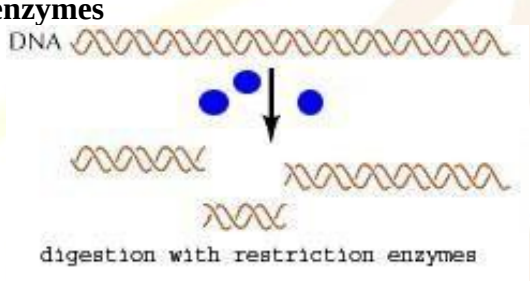
APPARATUS



1. Extract and purify DNA from cells

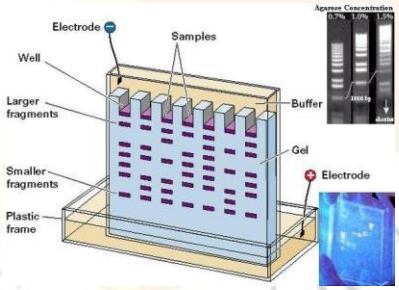
• Isolate the DNA in question from the rest of the cellular material in the nucleus.
• Incubate specimen with detergent to promote cell lysis.
• Lysis frees cellular proteins and DNA.
• Proteins are enzymatically degraded by incubation with proteinase.
• Organic or non-inorganic extraction removes proteins.
• DNA is purified from solution by alcohol precipitation.
• Visible DNA fibers are removed and suspended in buffer.

2. DNA is restricted with enzymes



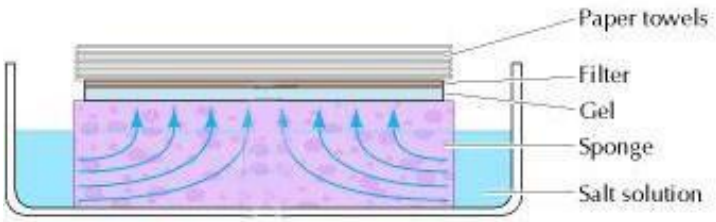
3. Separated by Gel electrophoresis

- The complex mixture of fragments is subjected to gel electrophoresis to separate the fragments according to size.



4. Denature DNA

- The restriction fragments present in the gel are denatured with alkali.
- This causes the double stranded to become single-stranded.
- DNA is then neutralized with NaCl to prevent re-hybridization before adding the probe.



5. Transfer to nitrocellulose paper

• Transfer the DNA from the gel to a solid support, ie, blotting.

• The blot is made permanent by:

– Drying at ~80°C

– Exposing to UV irradiation

6. Add labeled probe for hybridization

• The filter is incubated under hybridization conditions with a specific radio labeled DNA probe. The probe hybridizes to the complementary DNA restriction fragment.

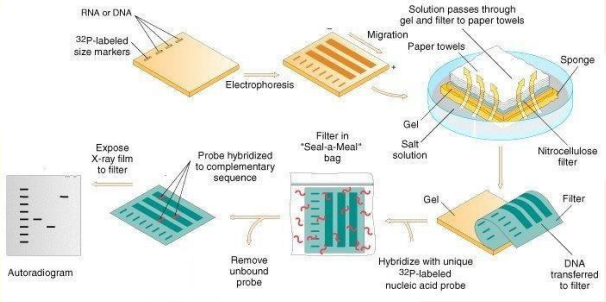
7. Wash off unbound probe

- Blot is incubated with wash buffers containing NaCl and detergent to wash away excess probe and reduce background.

8. Autoradiograph

• If the probe is radioactive, the particles emits when expose to X-ray film.

• There will be dark spots on the film wherever the probe bound.



APPLICATIONS

- To identify specific DNA in a DNA sample.
- To isolate desired DNA for construction of rDNA.
- Identify mutations, deletions, and gene rearrangements.
- Used in prognosis of cancer and in prenatal diagnosis of genetic diseases.
- In RFLP.
- Diagnosis of HIV-1 and infectious disease.
- In DNA fingerprinting.
- Paternity and Maternity Testing.
- Criminal Identification and Forensics.
- Personal Identification.

Ms. Visruthi Vijayan,
Dr. Lins Simon
Dept. of Botany

Principal
CPA College of Arts & Science