Biyani's Think Tank

Concept based notes

Molecular Biology and Biotechnology

(Botany-II)

[B.Sc. Part-III]

Neha Joshi
Lecturer
Deptt. of Science
Biyani Girls College, Jaipur

Published by :
Think Tanks
Biyani Group of Colleges

Sector-3, Vidhyadhar Nagar,
Jaipur-302 023 (Rajasthan)
Ph : 0141-2338371, 2338591-95 • Fax : 0141-2338007
E-mail : acad@biyanicolleges.org
Website :www.gurukpo.com; www.biyanicolleges.org

ISBN :

Edition : 2011
Price :

Leaser Type Setted by :
Biyani College Printing Department

Preface

I am glad to present this book, especially designed to serve the needs of the students. The book has been written keeping in mind the general weakness in understanding the fundamental concepts of the topics. The book is self-explanatory and adopts the "Teach Yourself" style. It is based on question-answer pattern. The language of book is quite easy and understandable based on scientific approach.

Any further improvement in the contents of the book by making corrections, omission and inclusion is keen to be achieved based on suggestions from the readers for which the author shall be obliged.

I acknowledge special thanks to Mr. Rajeev Biyani, *Chairman* & Dr. Sanjay Biyani, *Director* (*Acad.*) Biyani Group of Colleges, who are the backbones and main concept provider and also have been constant source of motivation throughout this Endeavour. They played an active role in coordinating the various stages of this Endeavour and spearheaded the publishing work.

I look forward to receiving valuable suggestions from professors of various educational institutions, other faculty members and students for improvement of the quality of the book. The reader may feel free to send in their comments and suggestions to the under mentioned address.

Author

Syllabus

Section A

History of molecular biology: work of Chargaff, Watson and crick model of DNA, Meselson and Stahl replication experiment ; Hershey and Chase experiment, Chromatin structure and gene expression, S. Benzer and gene concept. Kary Mullis and Polymerase chain reaction, Application of PCR technique, an overview of DNA fingerprinting and its use.

Section B

Central dogma , Reverse transcriptase and its application, Transcription in eukaryotes, RNA processing, capping, splicing and polyadenylation, Translation, initiation, elongation and termination. Jacob-Monod and Lac operon, Negative and positive control, attenuation and antitermination, structure of promoter.

Section C

Biotechnology : Functional definition. Basic aspects of Plant tissue culture, basal medium, media preparation and aseptic culture technique. Concept of cellular totipotency.Differentiation and morphogenesis. Micropropagation and synthetic seeds. Protoplast culture and somatic hybridization. Anther culture for androgenic haploid. Ovule and embryo culture and their application.

Section D

Recombinant DNA technology: techniques used in rDNA technology. Restriction enzymes. Vectors for gene transfer, Plasmids and cosmids, cDNA library, gene amplification ;Transgenic plants.

Contents

Section-A

History of Molecular Biology

Q1. What is the significance of chargaff's rule and what does it tell you about the double helix?

Ans. Erwin Chargaff proposed two main rules in his lifetime which were appropriately named Chargaff's rules. The first and best known achievement was to show that in natural DNA the number of guanine units equals the number of cytosine units and the number of adenine units equals the number of thymine units. In human DNA, for example, the four bases are present in these percentages: A=30.9% and T=29.4%; G=19.9% and C=19.8%. This strongly hinted towards the base pair makeup of the DNA. Most researchers had previously assumed that deviations from equimolar base ratios (G = A = C = T) were due to experimental error, but Chargaff documented that the variation was real, with [C + G] typically being slightly less abundant. He was able to do this with the newly developed paper chromatography and ultraviolet spectrophotometer. Chargaff's research would later help the Watson and Crick laboratory team to deduce the double helical structure of DNA.

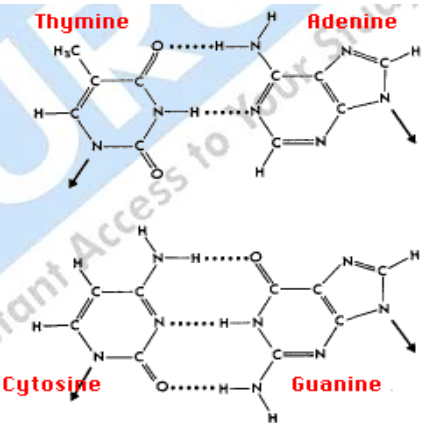


Figure: Chargaff's rule

The second of Chargaff's rules is that the composition of DNA varies from one species to another, in particular in the relative amounts of A, G, T, and C bases. Such evidence of molecular diversity, which had been presumed absent from DNA, made DNA a more credible candidate for the genetic material than protein.

Q2. How did watson and crick describe the structure of DNA?

Ans. The most widely accepted model for the structure of DNA molecule was proposed by Watson and Crick in 1953 (who won the Nobel Prize for Medicine in 1962). According to their model, the DNA has the following structural characteristics.

i. Molecule: The DNA molecule is a double helix. The molecule is formed by two *antiparallel* polynucleotide strands which are spirally coiled round each other in a right-handed helix. The two strands are held together by hydrogen bonds. The double stranded helical molecule has alternate major (or deep) and minor grooves.

ii. Structure of each strand: Each strand is a long polynucleotide of deoxyribonucleotides. The backbone of the strand is formed by alternately arranged deoxyribose sugar and phosphate molecules which are joined by thephosphodiester linkages. Each sugar in the strand has one base horizontally attached to it at carbon-1. It can be any one of the four: A, T, G or C. These four N-bases can occur in any possible sequence along the length of a strand.

iii. Complementary nature of the strands: The two strands are complementary to each other with regards to the arrangement of the bases in the two strands. For example, where adenine (a purine) occurs in one strand, thymine (a pyrimidine) is present in the corresponding position in the opposite strand and vice versa. Similarly, wherever guanine (a purine) is present in one strand, the other strand has cytosine (a pyrimidine) opposite to it and vice versa. Thus, in the double helix, purines and pyrimidines exist in base pairs, i.e., (A and T) and (G and C). As a result, if the base sequence of one strand of DNA is known, the base sequence of its complementary strand can be easily deduced.

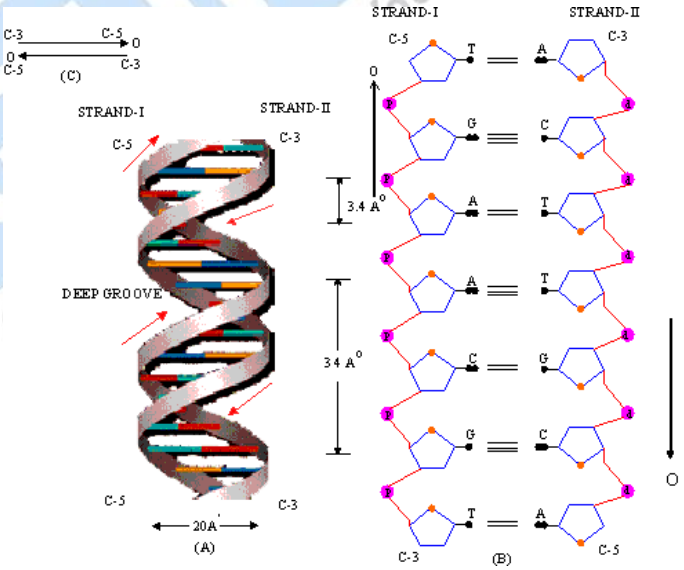


Figure: Structure of DNA (Watson and Crick model) (A) DNA double helix. (B) Detailed structure of the two strands. (C) C-5 and C-3 ends and antiparallel nature of strands.

vi. Complementary base pairing: In each pair, the two bases of the opposite strands are joined by hydrogen bonds. A and T are joined by two hydrogen bonds, while G and C are joined by three hydrogen bonds. This is called complementary base pairing. The two strands are thus held together all along their lengths by these hydrogen bonds.

v. Purine:Pyrimidine ratio: Because of the fixed or complementary base pairing in the DNA molecule, the total number of A is equal to the total number of T and the total number of G is equal to the total number of C. In other words, (A+G) = (T+C). Hence, purines: pyrimidines ratio is 1:1.

vi. C-3 and C-5 ends of the strand: In each strand one end of the strand has one free phosphate group on carbon-5 of the sugar molecule. This is the end of the strand is called C-5 (or 5') end. The other end of the strand has a free -OH on carbon-3 of the sugar molecule. This is called C-3 (or 3') end of the strand.

vii. Antiparallel nature of strands: The two strands are oppositely oriented and hence are called antiparallel. This means, the 3' end of one strand is adjacent to the 5' end of the other strand. This is because; the phosphate-sugar linkages run in opposite directions in the two strands.

viii. Dimensions: The diameter of the DNA double helix is 20 A^o. The length of each complete spiral (turn or pitch) of the molecule measures 34 A^o. 10 pairs ofnucleotides are present in each complete spiral. Therefore, each nucleotide in the strand occupies a distance of $3.4A^0$.

Q3. What was the Meselson-Stahl experiment and it's hypothesis?

Ans. The Meselson–Stahl experiment was an experiment by Matthew Meselson and Franklin Stahl in 1958 which supported the hypothesis that DNA replication was semiconservative. Semiconservative replication means that when the double stranded DNA helix was replicated, each of the two double stranded DNA helices consisted of one strand coming from the original helix and one newly synthesized. The semi-conservative nature of DNA replication was confirmed with the help of an experiment.

They marked the DNA in *Esherichia coli* with heavy isotope of nitrogen (15N) and then traced it for several following generations of the *E. coli* progeny in a medium with ^{15}N. When DNA is extracted from these cells and centrifuged on a salt density gradient, the DNA separates out at the point at which its density equals that of the salt solution. The DNA of the cells grown in ^{15}N medium had a higher density than cells grown in normal ^{14}N medium. After that, *E. coli* cells with only ^{15}N in their DNA were transferred to a ^{14}N medium and were allowed to divide; the progress of cell division was monitored by measuring the optical density of the cell suspension.

DNA was extracted periodically and was compared to pure ^{14}N DNA and ^{15}N DNA. After one replication, the DNA was found to have close to the intermediate density. Since conservative replication would result in equal amounts of DNA of the higher and lower densities (but no DNA of an intermediate density), conservative replication was excluded. However, this result was consistent with both semiconservative and dispersive replication. Semiconservative replication would result in double-stranded DNA with one strand of ^{15}N DNA, and one of ^{14}N DNA, while dispersive replication would result in double-stranded DNA with both strands having mixtures of ^{15}N and ^{14}N DNA, either of which would have appeared as DNA of an intermediate density.

The authors continued to sample cells as replication continued. DNA from cells after two replications had been completed was found to consist of equal amounts of DNA with two different densities, one corresponding to the intermediate density of DNA of cells grown for only one division in^{14}N medium, the other corresponding to DNA from cells grown exclusively in ^{14}N medium. This was inconsistent with dispersive replication, which would have resulted in a single density, lower than the intermediate density of the one-generation cells, but still higher than cells grown only in ^{14}N DNA medium, as the original ^{15}N DNA would have been split evenly among all DNA strands. The result was consistent with the semiconservative replication hypothesis

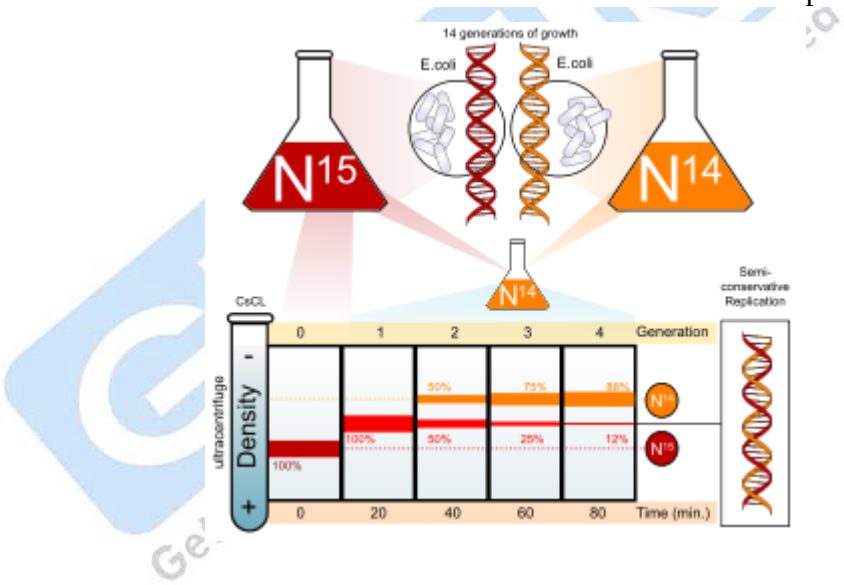


Figure: Meselson and Stahl Experiment

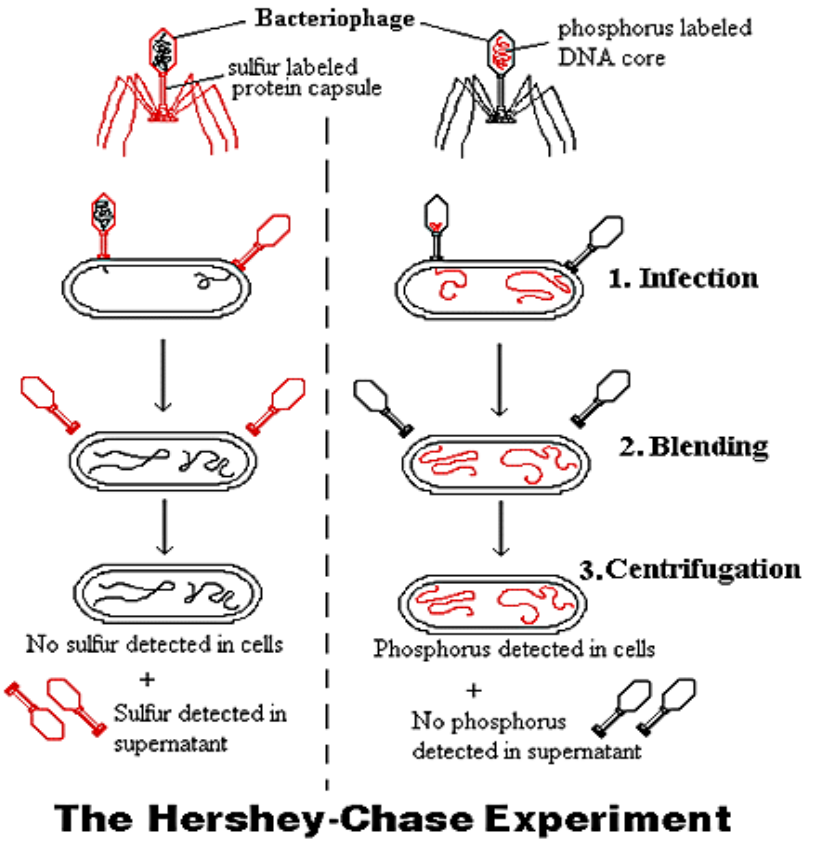
Q4. How did Hershey and Chase's research with bacteriophages help confirm DNA is the genetic material?

Ans. The Hershey–Chase experiment was conducted in 1952 by Alfred Hershey and Martha Chase, which helped to confirm that DNA was the genetic material. In their experiments, Hershey and Chase showed that when bacteriophages (composed of DNA and protein) infect bacteria, their DNA enters the host bacterial cell, but most of their protein does not.

Hershey and Chase were testing two competing hypotheses. The first was that DNA was the genetic material; the other was that protein was the genetic material. Since phosphorus is contained in DNA but not amino acids, radioactive phosphorus-32 was used to label the DNA contained in the T2 phage. Radioactive sulfur-35 was used to label the protein sections of the T2 phage, because sulfur is contained in amino acids but not DNA. They accomplished the incorporation of the radioactive elements into the bacteriophages by adding the isotopes into separate media and allowing bacteria to grow in these media for 4 hours before introducing the bacteriophages. When the bacteriophages infected the bacteria, the progeny contained the radioactive isotopes in their structures. This was done once for the sulfur labeled phages and once for phosphorus-labeled phages. Hershey and Chase were also able to prove that the DNA from the phage is inserted into the bacteria shortly after the virus attaches to its host. Using a high speed blender they were able to force the bacteriophages from the bacterial cells after adsorption. The lack of ^{32}P labeled DNA remaining in the solution after the bacteriophages had been allowed to adsorb to the bacteria showed that the phage DNA was transferred into the bacterial cell. The presence of almost all the radioactive ^{35}S in the solution showed that the protein coat that protects the DNA before adsorption stayed outside the cell.

However, once the DNA was inserted, it did not remain in the cytoplasm of the bacterial cell but was incorporated into the bacteria (later determined to be inserted into bacterial DNA). The infected bacterial cells were frozen to cause cell lysis and the release of most cellular material. Upon examination, almost the phage's entire labeled DNA was found to have been incorporated into the bacterial cells.

Hershey and Chase concluded that DNA, not protein, was the genetic material. They determined that a protective protein coat was formed around the bacteriophage, but that the internal DNA is what conferred its ability to produce progeny inside a bacteria. They determined this from the amount of radioactive material remaining outside of the cell. Only 20% of the ^{32}P remained outside the cell, demonstrating that it was incorporated with DNA in the cell's genetic material. All of the ^{35}S in the protein coats remained outside the cell, showing it was not incorporated into the cell, and that protein is not the genetic material. Hershey and Chase concluded that protein was not the hereditary genetic material.



The Hershey-Chase Experiment

Q5. Describe in detail semi conservative mode of DNA replication?

Ans. This is the most common method of DNA replication. It takes place in the nucleus where the DNA is present in the chromosomes. Replication takes place in the S-phase (synthesis phase) of the interphase nucleus. The deoxyribose nucleotides needed for the formation of the new DNA strands are present in the nucleoplasm.

At the time of replication, the two strands of DNA first separate. Each strand then acts as a template for the formation of a new strand. A new strand is constructed on each old strand, and two exactly identical double stranded DNA molecules are formed. In each new DNA molecule, one strand is old (original) while the other is newly formed. Hence, Watson and Crick described this method as semi-conservative replication.

The various steps involved in this process are summarized as follows:

i. The mechanism of replication starts at a specific point of the DNA molecule. This is called the origin.
ii. At the origin, the DNA strand breaks because of an incision (nick). This is made by an enzyme called incision enzyme (endonuclease).
iii. The hydrogen bonds joining the two strands are broken by the enzyme.

iv. The two strands start unwinding. This takes place with the help of a DNA unwinding protein. The two polynucleotide strands are thus separated.

v. The point where the two strands separate appears like a fork or a y-shape. This is described as a replicating fork.

vi. A new strand is constructed on each old strand. This takes place with the help of a small RNA primer molecule which is complimentary to the DNA at that point. Each old DNA strand acts as a template (site) for the construction of new strand. The RNA primer attaches itself to the old strand and attracts the enzymes which add new nucleotides. The deoxyribose nucleotides are present in the surrounding nucleoplasm. Appropriate nucleotides are selected from the nucleoplasm, and are attached by H-bonds to their respective complementary bases on the old strand. A new DNA strand is thus constructed opposite to each old strand.

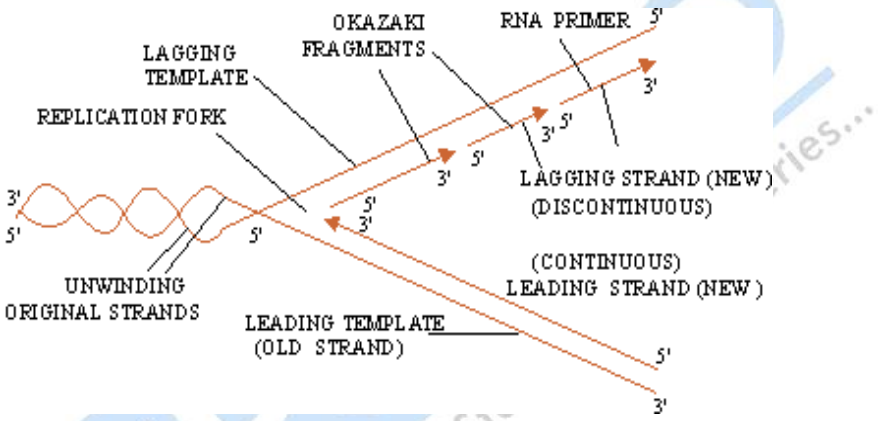


Figure: Replication Fork

vii. The formation of new complementary strand always begins at the 3' end of the template strand (original strand) and progresses towards the 5' end (i.e in 3'→5' direction). Since the new strand is antiparallel to the template strand, it is obvious that the new strand itself is always developed in the, 5'→3' direction. For this reason when the two original strands separate (then with respect to the origin of separation), one acts as 3'→5' template while the other acts as 3'→5' template. Of the two, the replication of 3'→5' template begins first. Hence the new strand formed on it is called the leading strand. The other template (5'→3') must begin replication at the fork and progress back toward the previously transcribed fragment. The new strand formed on it is called the lagging strand.

viii. Replication of the lagging strand takes place in small fragments called Okazaki fragments. These are then connected together by the enzyme polynucleotide ligase.

ix. Replication may take place in only one direction on the DNA helix (unidirectional) or in two directions (bidirectional).

x. At the end of the process, two double stranded DNA molecules are formed from the original DNA molecule.

xi. In each newly formed DNA, one strand is old while the other is new. Hence, it is described as semi-conservative replication.

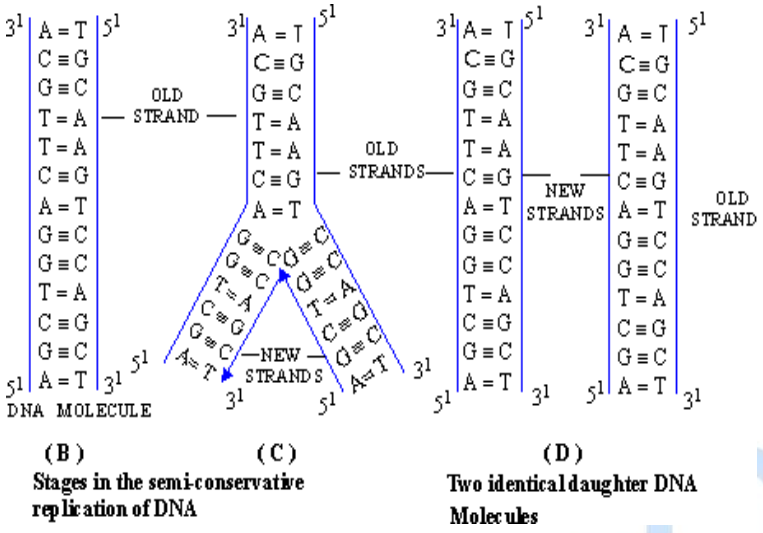


Figure: Stages in the semi-conservative method of DNA replication in eukaryotes

Q6. How DNA is organized into Chromatin in eukaryotes?

Ans. Chromatin is the combination of DNA and proteins that make up the contents of the nucleus of a cell. The primary functions of chromatin are; to package DNA into a smaller volume to fit in the cell, to strengthen the DNA to allow mitosis and meiosis and prevent DNA damage, and to control gene expression and DNA replication. The primary protein components of chromatin are histones that compact the DNA. Chromatin is only found in eukaryotic cells: prokaryotic cells have a very different organization of their DNA which is referred to as genophore (not chromatin).

The structure of chromatin depends on several factors. The overall structure depends on the stage of the cell cycle: during interphase the chromatin is structurally loose to allow access to RNA and DNA polymerases that transcribe and replicate the DNA. The local structure of chromatin during interphase depends on the genes present on the DNA: DNA coding genes that are actively transcribed ("turned on") are more loosely packaged and are found associated with RNA polymerases (referred to as euchromatin) while DNA coding inactive genes ("turned off") are found associated with structural proteins and are more tightly packaged (heterochromatin). Epigenetic chemical modification of the structural proteins in chromatin also alter the local chromatin structure, in particular chemical modifications of histone proteins by methylation and acetylation. As the cell prepares to divide, i.e. enters mitosis or meiosis, the chromatin packages more tightly to facilitate segregation of the chromosomes during anaphase. During this stage of the cell cycle this makes the individual chromosomes in many cells visible by optical microscope.

In general terms, there are three levels of chromatin organization:

1. DNA wraps around histone proteins forming nucleosomes; the "beads on a string" structure (euchromatin).

2. Multiple histones wrap into a 30 nm fibre consisting of nucleosome arrays in their most compact form (heterochromatin).
3. Higher-level DNA packaging of the 30 nm fibre into the metaphase chromosome (during mitosis and meiosis).

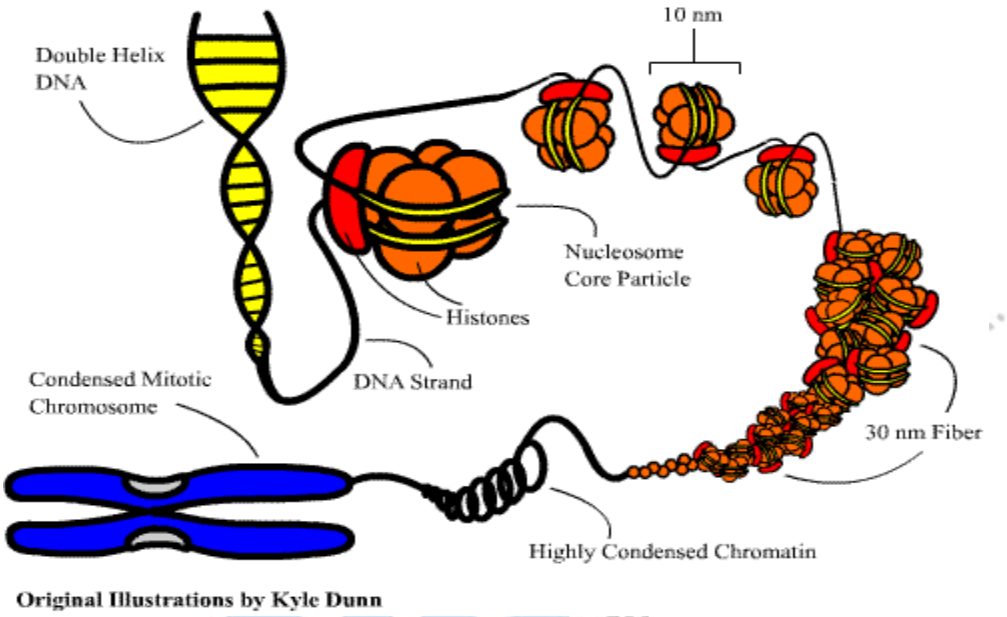


Figure: Chromatin structure

Q7. What is gene expression? Describe in detail various steps of gene expression.

Ans. Gene expression is the process by which information from a gene is used in the synthesis of a functional gene product. These products are often proteins, but in non-protein coding genes such as ribosomal RNA (rRNA), transfer RNA (tRNA) or small nuclear RNA (snRNA) genes, the product is a functional RNA. The process of gene expression is used by all known life - eukaryotes (including multicellular organisms), prokaryotes (bacteria and archae), possibly induced by viruses to generate the macromolecular machinery for life. Several steps in the gene expression process may be modulated, including the transcription, RNA splicing, translation, and post-translational modification of a protein. Gene regulation gives the cell control over structure and function, and is the basis for cellular differentiation, morphogenesis and the versatility and adaptability of any organism. Gene regulation may also serve as a substrate for evolutionary change, since control of the timing, location, and amount of gene expression can have a profound effect on the functions (actions) of the gene in a cell or in a multicellular organism.

The gene expression includes the following major steps:

a) Transcription

The gene itself is typically a long stretch of DNA which carries genetic information encoded by genetic code. Every molecule of DNA consists of two strands, each of them having 5' and 3' ends oriented in anti-parallel direction. The coding strand contains the genetic information while template strand (non-coding strand) serves as a blueprint for the production of RNA. The production of RNA copies of the DNA is called transcription, and is performed by RNA polymerase, which adds one RNA nucleotide at a time to a growing RNA strand. This RNA is complementary to the template 3' → 5' DNA strand, which is itself complementary to the coding 5' → 3' DNA strand. Therefore, the resulting 5' → 3' RNA strand is identical to the coding DNA strand with the exception that thymines (T) are replaced with uracils (U) in the RNA. A coding DNA strand reading "ATG" is transcribed as "AUG" in RNA.

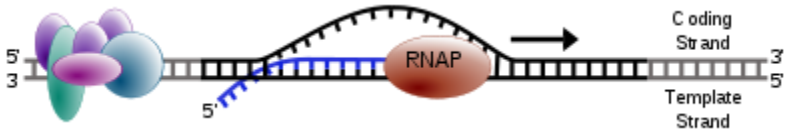


Figure: The process of transcription is carried out by RNA polymerase (RNAP), which uses DNA as a template and produces RNA

b) RNA processing

While transcription of prokaryotic protein-coding genes creates messenger RNA (mRNA) which is ready for translation, transcription of eukaryotic genes leaves a primary transcript of RNA (pre-mRNA), which first has to undergo series of modification to become a mature mRNA. These include 5' capping, 3' cleavage, polyadenylation and RNA splicing which aids in mRNA export to cytoplasm and translation re-initiation.

Extensive RNA processing may be an evolutionary advantage made possible by the nucleus of eukaryotes. In prokaryotes transcription and translation happen together whilst in eukaryotes the nuclear membrane separates the two processes giving time for RNA processing to occur.

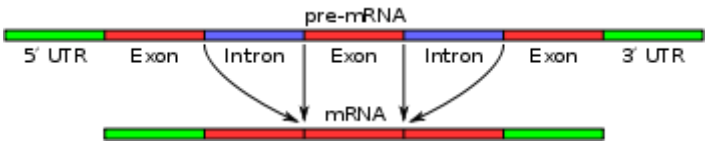


Figure: Simple illustration of exons and introns in pre-mRNA and the formation of mature mRNA by splicing. The UTRs are non-coding parts of exons at the ends of the mRNA.

c) non-coding RNA maturation

In most organisms non-coding genes (ncRNA) are transcribed as precursors which undergo further processing. In the case of ribosomal RNAs (rRNA), they are often transcribed as a pre-rRNA which contains one or more rRNAs, the pre-rRNA is cleaved and modified (2′-O-methylation and pseudouridine formation) at a specific sites by approximately 150 different small nucleolus-restricted RNA species, called snoRNAs. SnoRNAs associate with proteins, forming snoRNPs. In the case of transfer RNA (tRNA), for example, the 5' sequence is removed by RNase P, whereas the 3' end is removed by the tRNase Z enzyme and the non-templated 3' CCA tail is added by a nucleotidyl transferase. Even snRNAs and snoRNAs themselves undergo series of modification before they become part of functional RNP complex. This is done either in the nucleoplasm or in the specialized compartments called Cajal bodies.

d) RNA export

In eukaryotes most mature RNA must be exported to the cytoplasm from the nucleus. While some RNAs function in the nucleus, many RNAs are transported through the nuclear pores and into the cytosol. Notably this includes all RNA types involved in protein synthesis. In some cases RNAs are additionally transported to a specific part of the cytoplasm, such as a synapse; they are then towed by motor proteins that bind through linker proteins to specific sequences (called "zipcodes") on the RNA.

e) Translation

For some RNA (non-coding RNA) the mature RNA is the final gene product. In the case of messenger RNA (mRNA) the RNA is an information carrier coding for the synthesis of one or more proteins. mRNA carrying a single protein sequence (common in eukaryotes) is monocistronic whilst mRNA carrying multiple protein sequences (common in prokaryotes) is known as polycistronic.

Every mRNA consists of three parts - 5' untranslated region (5'UTR), protein-coding region or open reading frame (ORF) and 3' untranslated region (3'UTR). Coding region carries information for protein synthesis encoded by genetic code into form of triplets. Each triplet of nucleotides of the coding region is called codon and corresponds to a binding site complementary to an anticodon triplet in transfer RNA. Transfer RNAs with the same anticodon sequence always carry identical type of amino acid. Amino acids are then chained together by the ribosome according to order of triplets in the coding region. The ribosome helps transfer RNA to bind to messenger RNA and takes the amino acid from each transfer RNA and makes a structure-less protein out of it. Each mRNA molecule is translated into many protein molecules, on average ~900 in mammals.

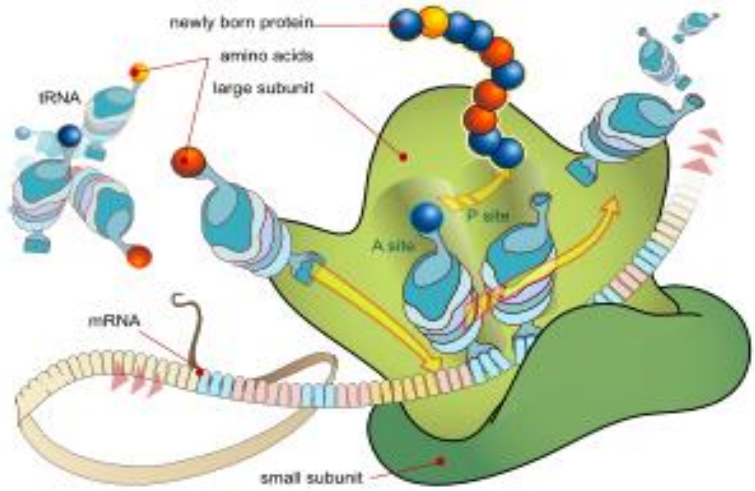


Figure: During the translation, tRNA charged with amino acid enters the ribosomes and aligns with the correct mRNA triplet. Ribosomes then adds amino acid to growing protein chain.

f) Folding

The polypeptide folds into its characteristic and functional three-dimensional structure from random coil. Each protein exists as an unfolded polypeptide or random coil when translated from a sequence of mRNA to a linear chain of amino acids. This polypeptide lacks any developed three-dimensional structure. Amino acids interact with each other to produce a well-defined three-dimensional structure, the folded protein (the right hand side of the figure), known as the native state. The correct three-dimensional structure is essential to function, although some parts of functional proteins may remain unfolded.

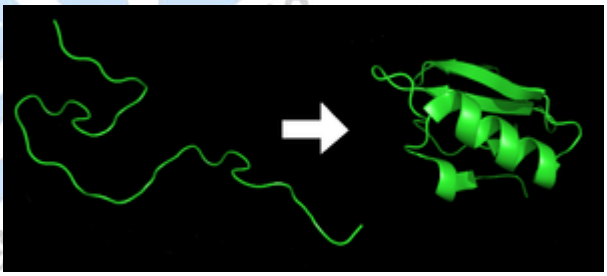


Figure: Protein before (left) and after (right) folding.

g) Protein transport

Many proteins are destined for other parts of the cell than the cytosol and a wide range of signalling sequences are used to direct proteins to where they are supposed to be. In prokaryotes this is normally a simple process due to limited compartmentalisation of the cell. However in eukaryotes there is a great variety of different targeting processes to ensure the protein arrives at the correct organelle.

Q8. Explain classical and modern concept of gene.

Ans. Classical concept of gene:

Classical concept of gene was introduced by sutton (1902) and was elaborated by Morgan (1913). Bidge (1923), Muller (1927) and others which outlined as follows.

i) Genes are discrete particles inherited in Mendelian fashion that occupies a definite locus in the chromosome and responsible for expression of specific phenotypic character.

ii) Number of genes in each organism is more than the number of chromosomes; hence several genes are located on each chromosome.

iii) The genes are arranged in a single linear order like beads on a string.

iv) Each gene occupies specific position called locus.

v) If the position of gene changes, character changes.

vi) Genes can be transmitted from parent to off springs.

vii) Genes may exist in several alternate formed called alleles.

viii) Genes are capable of combined together or can be replicated once during a cell division.

ix) Genes may under for sudden changes in position and composition called mutation.

x) Genes are capable of self duplication producing their own exact copies.

Modern concept of gene:

S. Benzer (1957) coined different terms for different nature of gene and genetic material in relation to the chromosome on the basis of genetic phenomena to which they involve.

i) Genes as unit of transmission or cistron:

The part of DNA specifying a single polypeptide chain is termed as cistron. A cistron can have 100 nucleotide pairs in length to 30,000 nucleotide pairs. It transmits characters from one generation to other as unit of transmission.

ii) Genes as unit of recombination or recon:

The smallest segment of DNA capable of being separated and exchange with other chromosome is called recon. A recon consists of not more than two pairs of uncleotides.

iii) Gene as unit of mutation or muton:

Muton is the smallest unit of genetic material which when changed or mutated produce a phenotypic trait. Thus muton is delimited to a single nucleotide.

Q9. Describe in detail the Polymerase chain reaction. Also explain the applications of PCR.

Ans. The polymerase chain reaction (PCR) developed by Kary Mullis developed in 1983. It is a scientific technique in molecular biology to amplify a single or a few copies of a piece of DNA across several orders of magnitude, generating thousands to millions of copies of a particular DNA sequence. PCR is now a common and often indispensable technique used in medical and biological research labs for a variety of applications. These include DNA cloning for sequencing, DNA-based phylogeny, or functional analysis of genes; the diagnosis of hereditary diseases; the identification of genetic fingerprints (used in forensic sciences and paternity testing); and the detection and diagnosis of infectious diseases. In 1993, Mullis was awarded the Nobel Prize in Chemistry along with Michael Smith for his work on PCR.

The method relies on thermal cycling, consisting of cycles of repeated heating and cooling of the reaction for DNA melting and enzymatic replication of the DNA. Primers (short DNA fragments) containing sequences complementary to the target region along with a DNA polymerase are key components to enable selective and repeated amplification. As PCR progresses, the DNA generated is itself used as a template for replication, setting in motion a chain reaction in which the DNA template is exponentially amplified. PCR can be extensively modified to perform a wide array of genetic manipulations.

PCR is used to amplify a specific region of a DNA strand (the DNA target). Most PCR methods typically amplify DNA fragments of up to ~10 kilo base pairs (kb), although some techniques allow for amplification of fragments up to 40 kb in size.

A basic PCR set up requires several components and reagents. These components include:

- *DNA template* that contains the DNA region (target) to be amplified.
- Two *primers* that are complementary to the 3' (three prime) ends of each of the sense and anti-sense strand of the DNA target.
- *Taq polymerase* or another DNA polymerase with a temperature optimum at around 70 °C. This enzyme is originally isolated from the bacterium *Thermus aquaticus*. This DNA polymerase enzymatically assembles a new DNA strand from DNA building-blocks, the nucleotides, by using single-stranded DNA as a template and DNA oligonucleotides (also called DNA primers), which are required for initiation of DNA synthesis.
- *Deoxynucleoside triphosphates* (dNTPs; nucleotides containing triphosphate groups), the building-blocks from which the DNA polymerase synthesizes a new DNA strand.

- *Buffer solution*, providing a suitable chemical environment for optimum activity and stability of the DNA polymerase.
- *Divalent cations*, magnesium or manganese ions; generally Mg^{2+} is used, but Mn^{2+} can be utilized for PCR-mediated DNA mutagenesis, as higher Mn^{2+}concentration increases the error rate during DNA synthesis
- *Monovalent cation* potassium ions.

Procedure

Typically, PCR consists of a series of 20-40 repeated temperature changes, called cycles, with each cycle commonly consisting of 2-3 discrete temperature steps, usually three. The cycling is often preceded by a single temperature step (called *hold*) at a high temperature (>90°C), and followed by one hold at the end for final product extension or brief storage. The temperatures used and the length of time they are applied in each cycle depend on a variety of parameters. These include the enzyme used for DNA synthesis, the concentration of divalent ions and dNTPs in the reaction, and the melting temperature (Tm) of the primers.

- *Initialization step*: This step consists of heating the reaction to a temperature of 94–96 °C (or 98 °C if extremely thermostable polymerases are used), which is held for 1–9 minutes. It is only required for DNA polymerases that require heat activation by hot-start PCR.

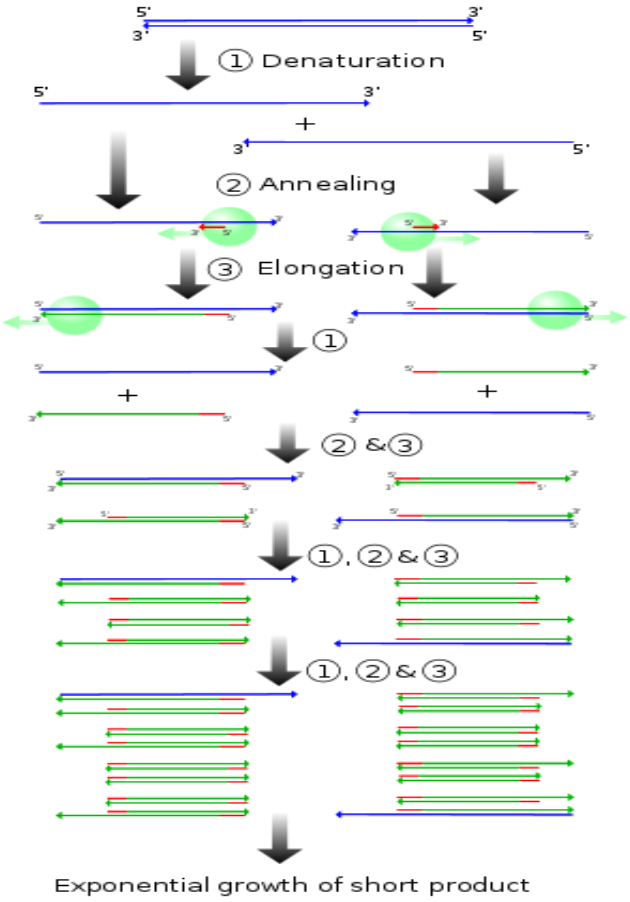


Figure: Schematic drawing of PCR cycle. (1) Denaturing at 94–96 °C. (2) Annealing (3) Elongation at 72 °C. Four at ~65 °C cycles are shown here.

- *Denaturation step*: This step is the first regular cycling event and consists of heating the reaction to 94–98 °C for 20–30 seconds. It causes DNA melting of the DNA template by disrupting the hydrogen bonds between complementary bases, yielding single-stranded DNA molecules.

- *Annealing step*: The reaction temperature is lowered to 50–65 °C for 20–40 seconds allowing annealing of the primers to the single-stranded DNA template. Typically the annealing temperature is about 3-5 degrees Celsius below the Tm of the primers used. Stable DNA-DNA hydrogen bonds are only formed when the primer sequence very closely matches the template sequence. The polymerase binds to the primer-template hybrid and begins DNA synthesis.

- *Extension/elongation step*: The temperature at this step depends on the DNA polymerase used; Taq polymerase has its optimum activity temperature at 75–80 °C, and commonly a temperature of 72 °C is used with this enzyme. At this step the DNA polymerase synthesizes a new DNA strand complementary to the DNA template strand

by adding dNTPs that are complementary to the template in 5' to 3' direction, condensing the 5'-phosphate group of the dNTPs with the 3'-hydroxyl group at the end of the nascent (extending) DNA strand. The extension time depends both on the DNA polymerase used and on the length of the DNA fragment to be amplified. As a rule-of-thumb, at its optimum temperature, the DNA polymerase will polymerize a thousand bases per minute. Under optimum conditions, i.e., if there are no limitations due to limiting substrates or reagents, at each extension step, the amount of DNA target is doubled, leading to exponential (geometric) amplification of the specific DNA fragment.

- *Final elongation*: This single step is occasionally performed at a temperature of 70–74 °C for 5–15 minutes after the last PCR cycle to ensure that any remaining single-stranded DNA is fully extended.

- *Final hold*: This step at 4–15 °C for an indefinite time may be employed for short-term storage of the reaction.

Applications of PCR

a) Selective DNA isolation

PCR allows isolation of DNA fragments from genomic DNA by selective amplification of a specific region of DNA. This use of PCR augments many methods, such as generating hybridization probes for Southern or northern hybridization and DNA cloning, which require larger amounts of DNA, representing a specific DNA region. PCR supplies these techniques with high amounts of pure DNA, enabling analysis of DNA samples even from very small amounts of starting material.

Other applications of PCR include DNA sequencing to determine unknown PCR-amplified sequences in which one of the amplification primers may be used in Sanger sequencing, isolation of a DNA sequence to expedite recombinant DNA technologies involving the insertion of a DNA sequence into a plasmid or the genetic material of another organism. Bacterial colonies (*E. coli*) can be rapidly screened by PCR for correct DNA vector constructs. PCR may also be used for genetic fingerprinting; a forensic technique used to identify a person or organism by comparing experimental DNAs through different PCR-based methods.

Some PCR 'fingerprints' methods have high discriminative power and can be used to identify genetic relationships between individuals, such as parent-child or between siblings, and are used in paternity testing. This technique may also be used to determine evolutionary relationships among organisms.

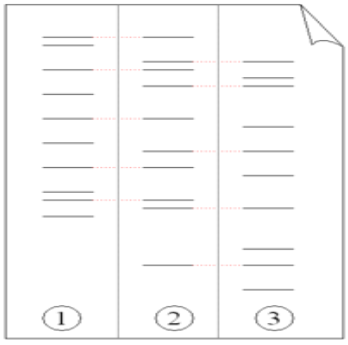


Figure: Electrophoresis of PCR-amplified DNA fragments. (1) Father. (2) Child. (3) Mother. The child has inherited some, but not the entire fingerprint of each of its parents, giving it a new, unique fingerprint.

b) Amplification and quantification of DNA

Because PCR amplifies the regions of DNA that it targets, PCR can be used to analyze extremely small amounts of sample. This is often critical for forensic analysis, when only a trace amount of DNA is available as evidence. PCR may also be used in the analysis of ancient DNA that is tens of thousands of years old. These PCR-based techniques have been successfully used on animals, such as a forty-thousand-year-old mammoth, and also on human DNA, in applications ranging from the analysis of Egyptian mummies to the identification of a Russian tsar.

Quantitative PCR methods allow the estimation of the amount of a given sequence present in a sample—a technique often applied to quantitatively determine levels of gene expression. Real-time PCR is an established tool for DNA quantification that measures the accumulation of DNA product after each round of PCR amplification.

c) PCR in diagnosis of diseases

PCR permits early diagnosis of malignant diseases such as leukemia and lymphomas, which is currently the highest-developed in cancer research and is already being used routinely. PCR assays can be performed directly on genomic DNA samples to detect translocation-specific malignant cells at a sensitivity that is at least 10,000-fold higher than that of other methods. PCR also permits identification of non-cultivatable or slow-growing microorganisms such as mycobacteria, anaerobic bacteria, or viruses from tissue culture assays and animal models. The basis for PCR diagnostic applications in microbiology is the detection of infectious agents and the discrimination of non-pathogenic from pathogenic strains by virtue of specific genes.

Viral DNA can likewise be detected by PCR. The primers used need to be specific to the targeted sequences in the DNA of a virus, and the PCR can be used for diagnostic analyses or DNA sequencing of the viral genome. The high sensitivity of PCR permits virus detection soon after infection and even before the onset of disease. Such early detection may give physicians a significant lead in treatment. The amount of virus ("viral load") in a patient can also be quantified by PCR-based DNA quantitation techniques.

Q10. Define DNA fingerprinting along with its applications.

Ans. The DNA profiling technique was first reported in 1984 by Sir Alec Jeffreys at the University of Leicester in England. DNA fingerprinting is a technique employed by forensic scientists to assist in the identification of individuals by their respective DNA profiles. DNA profiles are encrypted sets of numbers that reflect a person's DNA makeup, which can also be used as the person's identifier. They are able to determine whether two DNA samples are from the same person, related people, or non-related people. Scientists use a small number of sequences of DNA that are known to vary among individuals a great deal, and analyze those to get a certain probability of a match. It is used in, for example, parental testing and criminal investigation.

Although 99.9% of human DNA sequences are the same in every person, enough of the DNA is different to distinguish one individual from another, unless they are monozygotic twins. DNA profiling uses repetitive ("repeat") sequences that are highly variable, called variable number tandem repeats (VNTRs), particularly short tandem repeats (STRs). VNTR loci are very similar between closely related humans, but so variable that unrelated individuals are extremely unlikely to have the same VNTRs.

DNA fingerprinting Process:

The process begins with a sample of an individual's DNA (typically called a "reference sample"). The most desirable method of collecting a reference sample is the use of a buccal swab, as this reduces the possibility of contamination. Other methods may use a sample of blood, saliva, semen, or other appropriate fluid or tissue from personal items (e.g. toothbrush, razor, etc.) or from stored samples (e.g. banked sperm or biopsy tissue).

Samples obtained from blood relatives (biological relative) can provide an indication of an individual's profile, as could human remains which had been previously profiled. A reference sample is then analyzed to create the individual's DNA profile using one of a number of techniques, discussed below. The DNA profile is then compared against another sample to determine whether there is a genetic match.

a) Restriction fragment length polymorphism (RFLP) analysis

The first methods for finding out genetics used for DNA profiling involved restriction enzyme digestion, followed by Southern blot analysis. Although polymorphisms can exist in the restriction enzyme cleavage sites, more commonly the enzymes and DNA probes were used to analyze VNTR loci. However, the Southern blot technique is laborious, and requires large amounts of undegraded sample DNA. Also, Karl Brown's original technique looked at many minisatellite loci at the same time, increasing the observed variability, but making it hard to discern individual alleles (and thereby precluding parental testing). These early techniques have been supplanted by PCR-based assays.

b) Polymerase chain reaction (PCR) analysis

With the invention of the polymerase chain reaction (PCR) technique, DNA profiling took huge strides forward in both discriminating power and the ability to recover information from very small (or degraded) starting samples. PCR greatly amplifies the amounts of a specific region of DNA, using oligonucleotide primers and a thermostable DNA polymerase. Early assays such as the HLA-DQ alpha reverse dot blot strips grew to be very popular due to their ease of use, and the speed with which a result could be obtained. However they were not as discriminating as RFLP. It was also difficult to determine a DNA profile for mixed samples, such as a vaginal swab from a sexual assault victim. Fortunately, the PCR method was readily adaptable for analyzing VNTR, particularly STR loci.

c) Short tandem repeats (STR) analysis

The method of DNA profiling used today is based on PCR and uses short tandem repeats (STR) a type of VNTR. This method uses highly polymorphic regions that have short repeated sequences of DNA (the most common is 4 bases repeated, but there are other lengths in use, including 3 and 5 bases). Because unrelated people almost certainly have different numbers of repeat units, STRs can be used to discriminate between unrelated individuals. These STR loci (locations on a chromosome) are targeted with sequence-specific primers and amplified using PCR. The DNA fragments that result are then separated and detected using electrophoresis. There are two common methods of separation and detection, capillary electrophoresis (CE) and gel electrophoresis.

Each STR is polymorphic, but the number of alleles is very small. Typically each STR allele will be shared by around 5 - 20% of individuals. The power of STR analysis comes from looking at multiple STR loci simultaneously. The pattern of alleles can identify an individual quite accurately. Thus STR analysis provides an excellent identification tool. The more STR regions that are tested in an individual the more discriminating the test becomes.

d) Amplified fragment length polymorphism (AmpFLP)

Another technique, AmpFLP, or amplified fragment length polymorphism was also put into practice during the early 1990s. This technique was also faster than RFLP analysis and used PCR to amplify DNA samples. It relied on variable number tandem repeat (VNTR) polymorphisms to distinguish various alleles, which were separated on a polyacrylamide gel using an allelic ladder (as opposed to a molecular weight ladder). Bands could be visualized by silver staining the gel. One popular locus for fingerprinting was the D1S80 locus. As with all PCR based methods, highly degraded DNA or very small amounts of DNA may cause allelic dropout (causing a mistake in thinking a heterozygote is a homozygote) or other stochastic effects. In addition, because the analysis is done on a gel, very high number repeats may bunch together at the top of the gel, making it difficult to resolve. AmpFLP analysis can be highly automated, and allows for easy creation of phylogenetic trees based on comparing individual samples of DNA. Due to its relatively low cost and ease of set-up and operation, AmpFLP remains popular in lower income countries.

e) Y-chromosome analysis

Recent innovations have included the creation of primers targeting polymorphic regions on the Y-chromosome (Y-STR), which allows resolution of a mixed DNA sample from a male and female and/or cases in which a differential extraction is not possible. Y-chromosomes are paternally inherited, so Y-STR analysis can help in the identification of paternally related males.

d) Mitochondrial analysis

For highly degraded samples, it is sometimes impossible to get a complete profile of the STRs. In these situations, mitochondrial DNA (mtDNA) is sometimes typed due to there being many copies of mtDNA in a cell, while there may only be 1-2 copies of the nuclear DNA. Forensic scientists amplify the HV1 and HV2 regions of the mtDNA, then sequence each region and compare single-nucleotide differences to a reference. Because mtDNA is maternally inherited, directly linked maternal relatives can be used as match references, such as one's maternal grandmother's daughter's son. mtDNA can be obtained from such material as hair shafts and old bones/teeth.

Application of DNA fingerprinting:

1. Paternity and Maternity:

Because a person inherits his or her VNTRs from his or her parents, VNTR patterns can be used to establish paternity and maternity. The patterns are so specific that a parental VNTR pattern can be reconstructed even if only the children's VNTR patterns are known (the more children produced, the more reliable the reconstruction). Parent-child VNTR pattern analysis has been used to solve standard father-identification cases as well as more complicated cases of confirming legal nationality and, in instances of adoption, biological parenthood.

2. Criminal Identification and Forensics:

DNA isolated from blood, hair, skin cells, or other genetic evidence left at the scene of a crime can be compared, through VNTR patterns, with the DNA of a criminal suspect to determine guilt or innocence. VNTR patterns are also useful in establishing the identity of a homicide victim, either from DNA found as evidence or from the body itself.

3. Personal Identification:

The notion of us ing DNA fingerprints as a sort of genetic bar code to identify individuals has been discussed, but this is not likely to happen anytime in the foreseeable future. The technology required to isolate, keep on file, and then analyze millions of very specified VNTR patterns is both expensive and impractical. Social security numbers, picture ID, and other more mundane methods are much more likely to remain the prevalent ways to establish personal identification.

Section-B

Q1. Write short notes on:

a) **Central Dogma**
b) **Reverse transcriptase**
c) **Lac operon**
d) **Negative and Positive control**

Ans. **a) Central Dogma.** The dogma is a framework for understanding the transfer of sequence information between sequential information-carrying biopolymers, in the most common or general case, in living organisms. There are 3 major classes of such biopolymers: DNA and RNA (both nucleic acids), and protein. The general transfers describe the normal flow of biological information: DNA can be copied to DNA (DNA replication), DNA information can be copied into mRNA(transcription), and proteins can be synthesized using the information in mRNA as a template (translation).

DNA replication

As the final step in the Central Dogma, to transmit the genetic information between parents and progeny, the DNA must be replicated faithfully. Replication is carried out by a complex group of proteins called helicase that unwind the superhelix, unwind the double-stranded DNA helix, and, using DNA polymerase and its associated proteins, copy or replicate the master template itself so the cycle can repeat DNA → RNA → protein in a new generation of cells or organisms.

Table of the 3 classes of information transfer suggested by the dogma

General	Special	Unknown
DNA → DNA	RNA → DNA	Protein → DNA
DNA → RNA	RNA → RNA	Protein → RNA
RNA → protein	DNA → protein	Protein → protein

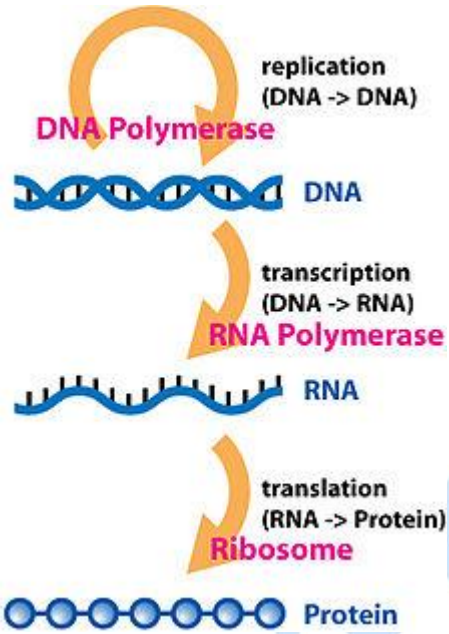
General transfers of biological sequential information

Transcription

Transcription is the process by which the information contained in a section of DNA is transferred to a newly assembled piece of messenger RNA (mRNA). It is facilitated by RNA polymerase and transcription factors. In eukaryote cells the primary transcript (pre-mRNA) is often processed further via alternative splicing. In this process, blocks of mRNA are cut out and rearranged, to produce different arrangements of the original sequence.

Translation

Eventually, this mature mRNA finds its way to a ribosome, where it is translated. In prokaryotic cells, which have no nuclear compartment, the process of transcription and translation may be linked together. In eukaryotic cells, the site of transcription (the cell nucleus) is usually separated from the site of translation (the cytoplasm), so the mRNA must be transported out of the nucleus into the cytoplasm, where it can be bound by ribosomes. Complexes of initiation factors and elongation factors bring aminoacylated transfer RNAs (tRNAs) into the ribosome-mRNA complex, matching the codon in the mRNA to the anti-codon in the tRNA, thereby adding the correct amino acid in the sequence encoding the gene. As the amino acids are linked into the growing peptide chain, they begin folding into the correct conformation. Translation ends with a UAA, UGA, or UAG stop codon. The nascent polypeptide chain is then released from the ribosome as a mature protein.



b) Reverse transcriptase. Reverse transcriptase, also known as RNA-dependent DNA polymerase, is a DNA polymerase enzyme that transcribes single-stranded RNA into single-stranded DNA. It also helps in the formation of a double helix DNA once the RNA has been reverse transcribed into a single strand cDNA. Normal transcription involves the synthesis of RNA from DNA; hence, reverse transcription is the *reverse* of this. The major reverse transcriptases include:

- HIV-1 reverse transcriptase from human immunodeficiency virus type 1 (PDB 1HMV)
- M-MLV reverse transcriptase from the Moloney murine leukemia virus
- AMV reverse transcriptase from the avian myeloblastosis virus
- Telomerase reverse transcriptase that maintains the telomeres of eukaryotic chromosomes

Reverse transcriptases have two activities:

- **DNA polymerase activity:** In the retroviral life cycle, reverse transcriptase copies only RNA, but, as used in the laboratory, it will transcribe both single-stranded RNA and single-stranded DNA templates with essentially equivalent efficiency. In both cases, an RNA or DNA primer is required to initiate synthesis.

- **RNase H activity:** RNase H is a ribonuclease that degrades the RNA from RNA-DNA hybrids, such as are formed during reverse transcription of an RNA template.

This enzyme functions as both an endonuclease and exonuclease in hydrolyzing its target.

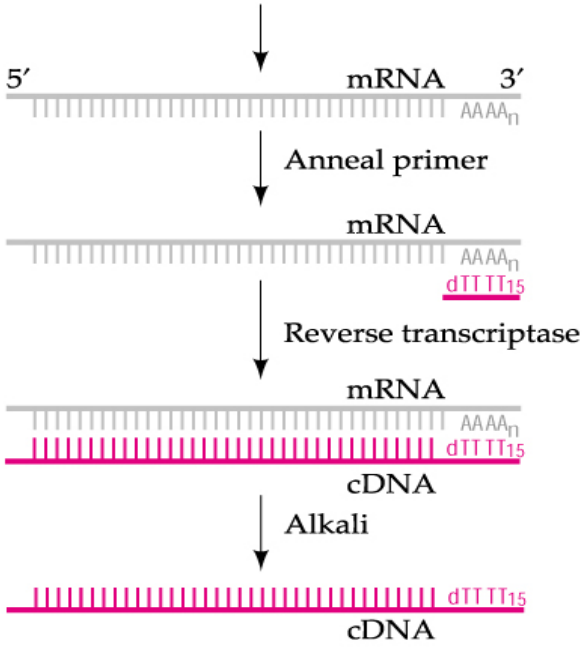


Figure: Process of reverse transcription

Applications

i) Antiviral drugs. As HIV uses reverse transcriptase to copy its genetic material and generate new viruses (part of a retrovirus proliferation circle), specific drugs have been designed to disrupt the process and thereby suppress its growth. Collectively, these drugs are known as reverse transcriptase inhibitors and include the nucleoside and nucleotide analogues zidovudine (trade name Retrovir), lamivudine (Epivir) and tenofovir (Viread), as well as non-nucleoside inhibitors, such as nevirapine (Viramune).

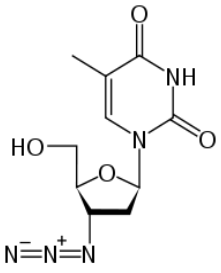


Figure: The molecular structure of zidovudine (AZT), a drug used to inhibit HIV reverse transcriptase

ii) Molecular biology. Reverse transcriptase is commonly used in research to apply the polymerase chain reaction technique to RNA in a technique called reverse

transcription polymerase chain reaction (RT-PCR). The classical PCR technique can be applied only to DNA strands, but, with the help of reverse transcriptase, RNA can be transcribed into DNA, thus making PCR analysis of RNA molecules possible. Reverse transcriptase is used also to create cDNA libraries from mRNA. The commercial availability of reverse transcriptase greatly improved knowledge in the area of molecular biology, as, along with other enzymes, it allowed scientists to clone, sequence, and characterize DNA.

Reverse transcriptase has also been employed in insulin production. By inserting eukaryotic mRNA for insulin production along with reverse transcriptase into bacteria, the mRNA can insert itself into the prokaryote's genome, and large amounts of insulin can be created, sidestepping the need to harvest pig pancreas and other such traditional sources. Inserting eukaryotic DNA (instead of mRNA) into bacteria would not work because it is fragmented, with introns, and would not transcribe successfully using the bacteria's ribosomes.

c) Lac Operon. The *lac* operon is an operon required for the transport and metabolism of lactose in *Escherichia coli* and some other enteric bacteria. The *lac* operon is regulated by several factors including the availability of glucose and of lactose.

The *lac* operon allows for the effective digestion of lactose. The cell can use lactose as an energy source by producing the enzyme β-galactosidase to digest that lactose into glucose and galactose. However, it would be inefficient to produce enzymes when there is no lactose available, or if there is a more readily-available energy source available such as glucose. The *lac* operon uses a two-part control mechanism with the *lac* repressor, which halts the production in the absence of lactose and the Catabolite activator protein (CAP), which assists in production in the absence of glucose.

Structure of the lac operon

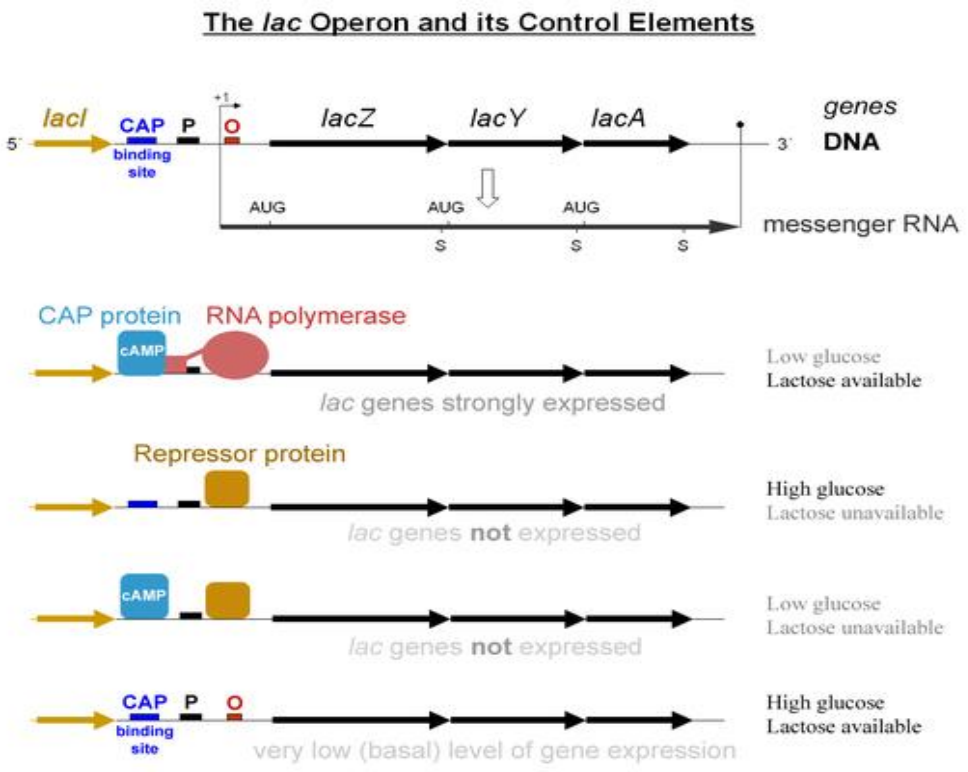
The *lac* operon consists of three structural genes, and a promoter, a terminator, regulator, and an operator. The three structural genes are: *lacZ*, *lacY*, and *lacA*.

- *lacZ* encodes β-galactosidase (LacZ), an intracellular enzyme that cleaves the disaccharide lactose into glucose and galactose.
- *lacY* encodes β-galactoside permease (LacY), a membrane-bound transport protein that pumps lactose into the cell.

- *lacA* encodes β-galactoside transacetylase (LacA), an enzyme that transfers an acetyl group from acetyl-CoA to β-galactosides.

Specific control of the *lac* genes depends on the availability of the substrate lactose to the bacterium. The proteins are not produced by the bacterium when lactose is unavailable as a carbon source. The *lac* genes are organized into an operon; that is, they are oriented in the same direction immediately adjacent on the chromosome and are co-transcribed into a single polycistronic mRNA molecule. Transcription of all genes starts with the binding of the enzyme RNA polymerase (RNAP), a DNA-binding protein, which binds to a specific DNA binding site, the *promoter*, immediately upstream of the genes. From this position RNAP proceeds to transcribe all three genes (*lacZYA*) into mRNA. The first control mechanism is the regulatory response to lactose, which uses an intracellular *regulatory protein* called the *lactose repressor* to hinder production of β-galactosidase in the absence of lactose. The *lacI* gene coding for the repressor lies nearby the *lac* operon and is always expressed (*constitutive*). If lactose is missing from the growth medium, the repressor binds very tightly to a short DNA sequence just downstream of the promoter near the beginning of *lacZ* called the *lac operator*. The repressor binding to the operator interferes with binding of RNAP to the promoter, and therefore mRNA encoding LacZ and LacY is only made at very low levels. When cells are grown in the presence of lactose, however, a lactose metabolite called allolactose, which is a combination of glucose and galactose, binds to the repressor, causing a change in its shape. Thus altered, the repressor is unable to bind to the operator, allowing RNAP to transcribe the *lac* genes and thereby leading to high levels of the encoded proteins.

The second control mechanism is a response to glucose, which uses the Catabolite activator protein (CAP) to greatly increase production of β-galactosidase in the absence of glucose. Cyclic adenosine monophosphate (cAMP) is a signal molecule whose prevalence is inversely proportional to that of glucose. It binds to the CAP, which in turn allows the CAP to bind to the CAP binding site (a 16 bp DNA sequence upstream of the promoter on the left in the diagram below), which assists the RNAP in binding to the DNA. In the absence of glucose, the cAMP concentration is high and binding of CAP-cAMP to the DNA significantly increases the production of β-galactosidase, enabling the cell to hydrolyze (digest) lactose and release galactose and glucose.



e) Negative and Positive control. Gene regulatory proteins are proteins that recognize and bind to specific short stretches of double-helical DNA and thereby determine which of the thousands of genes in a cell will be transcribed. The mode of control may be negative or positive.

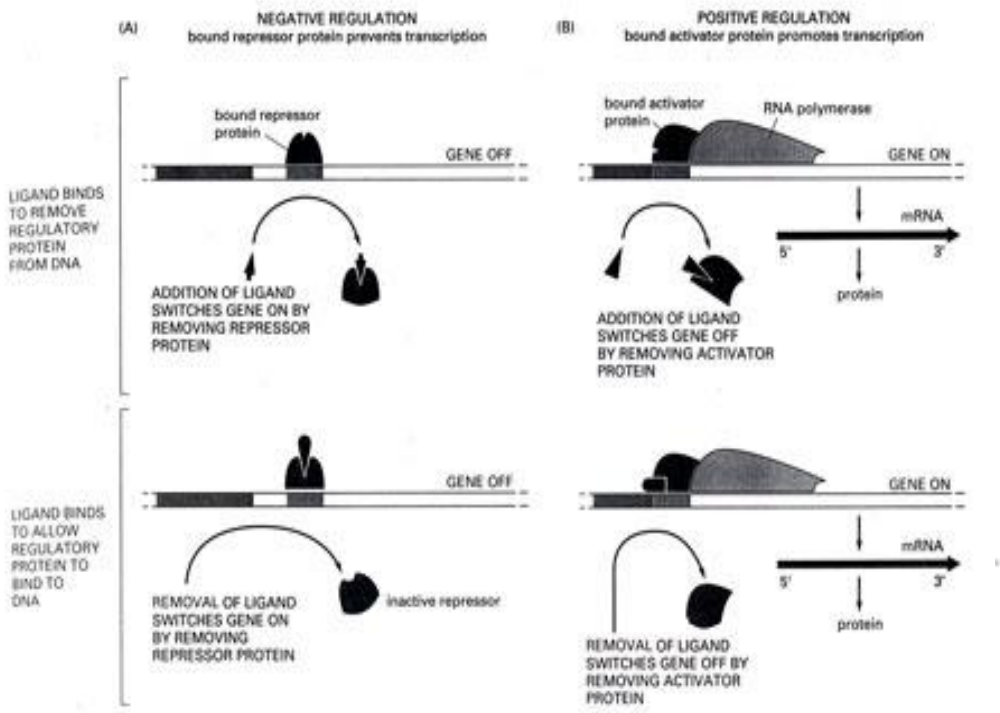
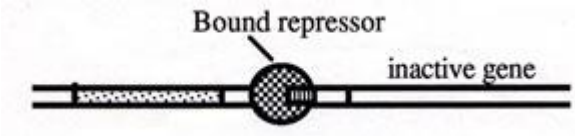


Figure: A ligand is defined as any molecule that binds to a specific site on a protein molecule. The binding of a specific ligand to an allosteric protein causes the protein to reversibly change its shape, from an active to an inactive form, or vice-versa. In the case of the lac operon (upper left), the ligand is the inducer (e.g. allolactose); in a repressible operon (bottom left), the ligand is a corepressor.

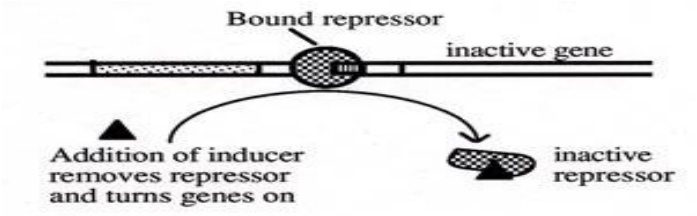
By definition, negative control occurs when a repressor protein is involved, that is when a regulator protein is active, DNA-binding repressor prevents RNA polymerase from binding and turning the genes off. Positive control is the exact opposite of negative control: instead of interfering with the initiation of transcription, it enhances transcription. Positive control occurs when an active, DNA-binding regulatory protein binds to DNA and assists the binding of RNA polymerase and therefore facilitates transcription. Such regulator proteins are called transcriptional activators.

Negative control: Bound repressor protein prevents transcription.

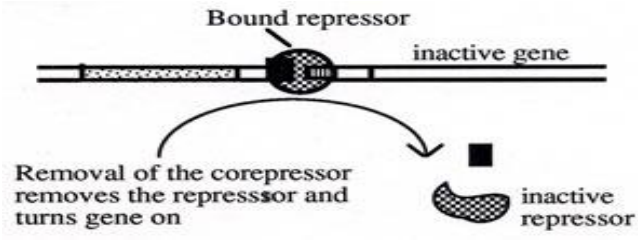


In negative control, the bacterial gene repressor protein binds to the operator near the promoter and thereby inhibits transcription of specific genes.

Examples: An inducible operon:



A repressible operon:

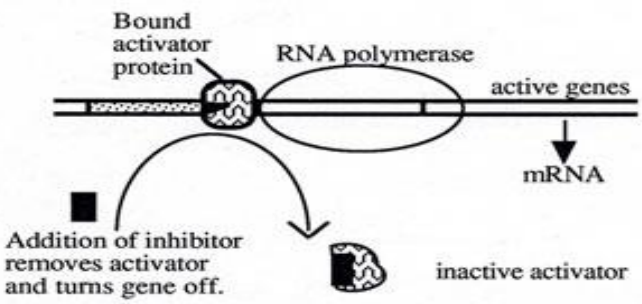


Because in both cases the binding of the regulatory protein suppresses transcription, this type of gene control is called *negative control*.

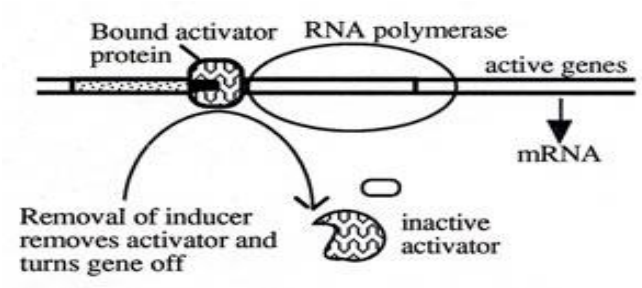
Positive gene control: In positive gen control, a gene activator protein facilitates the action of RNA polymerase.

Examples:

Activator protein facilitates transcription unless removed by inhibitor:



Activator protein facilitates transcription unless inducer is removed and activator can no longer bind:



Because in both cases more transcription occurs in the presence of the activator protein than in its absence, this type of regulation is called *positive control*.

Q2. Describe in detail the process of transcription.

Ans. **Transcription** is the process of creating a complementary RNA copy of a sequence of DNA. Both RNA and DNA are nucleic acids, which use base pairs of nucleotides as a complementary language that can be converted back and forth from DNA to RNA by the action of the correct enzymes. During transcription, a DNA sequence is read by RNA polymerase, which produces a complementary, antiparallel RNA strand. As opposed to DNA replication, transcription results in an RNA complement that includes uracil (U) in all instances where thymine (T) would have occurred in a DNA complement. Transcription can be explained easily in 4 or 5 steps, each moving like a wave along the DNA.

1. RNA polymerase unwinds/"unzips" the DNA by breaking the hydrogen bonds between complementary nucleotides.
2. RNA nucleotides are paired with complementary DNA bases.
3. RNA sugar-phosphate backbone forms with assistance from RNA polymerase.
4. Hydrogen bonds of the untwisted RNA+DNA helix break, freeing the newly synthesized RNA strand.
5. If the cell has a nucleus, the RNA is further processed (addition of a 3' poly-A tail and a 5' cap) and exits through to the cytoplasm through the nuclear pore complex.

Eukaryotic transcription is more complex than prokaryotic transcription. For instance, in eukaryotes the genetic material (DNA), and therefore transcription, is primarily localized to the nucleus, where it is separated from the cytoplasm (in which translation occurs) by the nuclear membrane. Eukaryotes have three nuclear RNA polymerases, each with distinct roles and properties:

Name	Transcribed	Type
RNA Polymerase I (Pol I, Pol A)	nucleolus	Larger ribosomal RNA (rRNA) (28S, 18S, 5.8S)
RNA Polymerase II	nucleus	messenger RNA (mRNA) and most small

(Pol II, Pol B)		nuclear RNAs (snRNAs)
RNA Polymerase III (Pol III, Pol C)	nucleus (and possibly the nucleolus-nucleoplasm interface)	transfer RNA (tRNA) and other small RNAs (including the small 5S rRNA)

Major steps involved in transcription

Pre-initiation. RNA polymerase and the initiation of transcription, requires the presence of a core promoter sequence in the DNA. Promoters are regions of DNA that promote transcription and, in eukaryotes, are found at -30, -75, and -90 base pairs upstream from the transcription start site. Core promoters are sequences within the promoter that are essential for transcription initiation. RNA polymerase is able to bind to core promoters in the presence of various specific transcription factors.

The most characterized type of core promoter in eukaryotes is a short DNA sequence known as a TATA box, found 25-30 base pairs upstream from the TSS. The TATA box, as a core promoter, is the binding site for a transcription factor known as TATA-binding protein (TBP), which is itself a subunit of another transcription factor, called Transcription Factor II D (TFIID). After TFIID binds to the TATA box via the TBP, five more transcription factors and RNA polymerase combine around the TATA box in a series of stages to form a preinitiation complex. One transcription factor, DNA helicase, has helicase activity and so is involved in the separating of opposing strands of double-stranded DNA to provide access to a single-stranded DNA template. However, only a low, or basal, rate of transcription is driven by the preinitiation complex alone. Other proteins known as activators and repressors, along with any associated coactivators or corepressors, are responsible for modulating transcription rate.

Thus, preinitiation complex contains: 1. Core Promoter Sequence 2. Transcription Factors 3. DNA Helicase 4. RNA Polymerase 5. Activators and Repressors The transcription preinitiation in archaea is, in essence, homologous to that of eukaryotes, but is much less complex. The archaeal preinitiation complex assembles at a TATA-box binding site; however, in archaea, this complex is composed of only RNA polymerase II, TBP, and TFB (the archaeal homologue of eukaryotic transcription factor II B (TFIIB)).

Initiation. In bacteria, transcription begins with the binding of RNA polymerase to the promoter in DNA. RNA polymerase is a core enzyme consisting of five subunits: 2 α subunits, 1 β subunit, 1 β' subunit, and 1 ω subunit. At the start of initiation, the core enzyme is associated with a sigma factor that aids in finding the appropriate -35 and -10 base pairs downstream of promoter sequences.

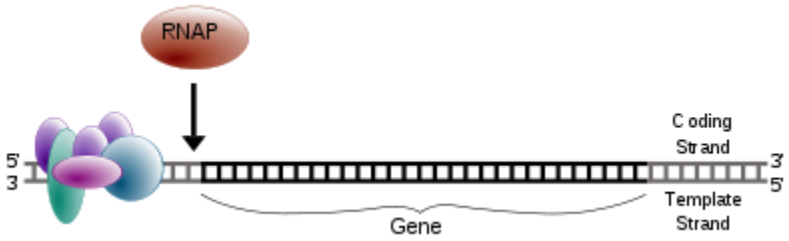


Figure: Simple diagram of transcription initiation. RNAP = RNA polymerase

Transcription initiation is more complex in eukaryotes. Eukaryotic RNA polymerase does not directly recognize the core promoter sequences. Instead, a collection of proteins called transcription factors mediate the binding of RNA polymerase and the initiation of transcription. Only after certain transcription factors are attached to the promoter does the RNA polymerase bind to it. The completed assembly of transcription factors and RNA polymerase bind to the promoter, forming a transcription initiation complex. Transcription in the archaea domain is similar to transcription in eukaryotes.

Promoter clearance. After the first bond is synthesized, the RNA polymerase must clear the promoter. During this time there is a tendency to release the RNA transcript and produce truncated transcripts. This is called *abortive initiation* and is common for both eukaryotes and prokaryotes. Abortive initiation continues to occur until the σ factor rearranges, resulting in the transcription elongation complex (which gives a 35 bp moving footprint). The σ factor is released before 80 nucleotides of mRNA are synthesized. Once the transcript reaches approximately 23 nucleotides, it no longer slips and elongation can occur. This, like most of the remainder of transcription, is an energy-dependent process, consuming adenosine triphosphate (ATP).

Promoter clearance coincides with phosphorylation of serine 5 on the carboxy terminal domain of RNA Pol in eukaryotes, which is phosphorylated by TFIIH.

Elongation. One strand of the DNA, the *template strand* (or noncoding strand), is used as a template for RNA synthesis. As transcription proceeds, RNA polymerase traverses the template strand and uses base pairing complementarity with the DNA template to create an RNA copy. Although RNA polymerase traverses the template strand from 3' → 5', the coding (non-template) strand and newly-formed RNA can also be used as reference points, so transcription can be described as occurring 5' → 3'. This produces an RNA molecule from 5' → 3', an exact copy of the coding strand (except that thymines are replaced with uracils, and the nucleotides are composed of a ribose (5-carbon) sugar where DNA has deoxyribose (one less oxygen atom) in its sugar-phosphate backbone). Unlike DNA replication, mRNA transcription can involve multiple RNA polymerases on a single DNA template and multiple rounds of

transcription (amplification of particular mRNA), so many mRNA molecules can be rapidly produced from a single copy of a gene.

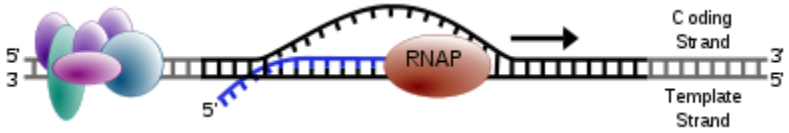


Figure: Simple diagram of transcription elongation

Elongation also involves a proofreading mechanism that can replace incorrectly incorporated bases. In eukaryotes, this may correspond with short pauses during transcription that allow appropriate RNA editing factors to bind. These pauses may be intrinsic to the RNA polymerase or due to chromatin structure.

Termination. Bacteria use two different strategies for transcription termination. In Rho-independent transcription termination, RNA transcription stops when the newly synthesized RNA molecule forms a G-C-rich hairpin loop followed by a run of Us. When the hairpin forms, the mechanical stress breaks the weak rU-dA bonds, now filling the DNA-RNA hybrid. This pulls the poly-U transcript out of the active site of the RNA polymerase, in effect, terminating transcription. In the "Rho-dependent" type of termination, a protein factor called "Rho" destabilizes the interaction between the template and the mRNA, thus releasing the newly synthesized mRNA from the elongation complex. Transcription termination in eukaryotes is less understood but involves cleavage of the new transcript followed by template-independent addition of *A*s at its new 3' end, in a process called polyadenylation.

Figure: Simple diagram of transcription termination

Q3 Explain post-transcriptional modification in RNA.

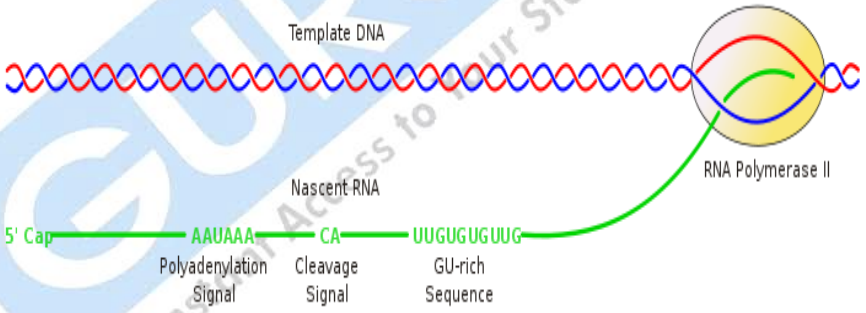
Ans. **Post-transcriptional modification** is a process in cell biology by which, in eukaryotic cells, primary transcript RNA is converted into mature RNA. A notable example is the conversion of precursor messenger RNA into mature messenger RNA (mRNA), which includes splicing and occurs prior to protein synthesis. This process is vital for the correct translation of the genomes of eukaryotes as the human primary RNA transcript that is produced as a result of transcription contains both exons, which are coding sections of the primary RNA transcript and introns, which are the non coding sections of the primary RNA transcript.

mRNA processing

The pre-mRNA molecule undergoes three main modifications. These modifications are 5' capping, 3' polyadenylation, and RNA splicing, which occur in the cell nucleus before the RNA is translated.

a) 5' Processing

Capping. Capping of the pre-mRNA involves the addition of 7-methylguanosine (m^7G) to the 5' end. To achieve this, the terminal 5' phosphate requires removal, which is done with the aid of a phosphatase enzyme. The enzyme guanosyl transferase then catalyses the reaction, which produces the diphosphate 5' end. The diphosphate 5' prime end then attacks the α phosphorus atom of a GTP molecule in order to add the guanine residue in a 5'5' triphosphate link. The enzyme (guanine-N^7-)-methyltransferase ("cap MTase") transfers a methyl group from S-adenosyl methionine to the guanine ring. This type of cap, with just the (m^7G) in position is called a cap 0 structure. The ribose of the adjacent nucleotide may also be methylated to give a cap 1. Methylation of nucleotides downstream of the RNA molecule produce cap 2, cap 3 structures and so on. In these cases the methyl groups are added to the 2' OH groups of the ribose sugar. The cap protects the 5' end of the primary RNA transcript from attack by ribonucleases that have specificity to the 3'5' phosphodiester bonds.



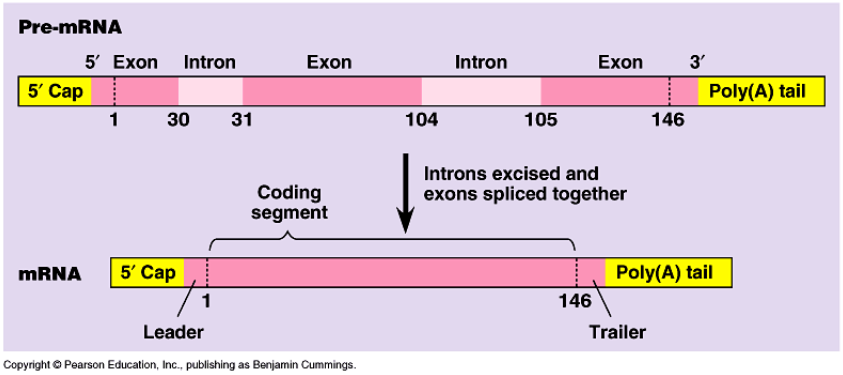
b) 3' Processing

Cleavage and Polyadenylation. The pre-mRNA processing at the 3' end of the RNA molecule involves cleavage of its 3' end and then the addition of about 200 adenine residues to form a poly (A) tail. The cleavage and adenylation reactions occur if a polyadenylation signal sequence (5'- AAUAAA-3') is located near the 3' end of the pre-mRNA molecule, which is followed by another sequence, which is usually (5'-CA-3'). The second signal is the site of cleavage. A GU-rich sequence is also usually present further downstream on the pre-mRNA molecule. After the synthesis of the sequence elements, two multisubunit proteins called cleavage and polyadenylation

specificity factor (CPSF) and cleavage stimulation factor (CStF) are transferred from RNA Polymerase II to the RNA molecule. The two factors bind to the sequence elements. A protein complex forms that contains additional cleavage factors and the enzyme Polyadenylate Polymerase (PAP). This complex cleaves the RNA between the polyadenylation sequence and the GU-rich sequence at the cleavage site marked by the (5'-CA-3') sequences. Poly (A) polymerase then adds about 200 adenine units to the new 3' end of the RNA molecule using ATP as a precursor. As the poly (A) tails is synthesised, it binds multiple copies of poly(A) binding protein, which protects the 3'end from ribonuclease digestion.

c) Splicing

RNA splicing is the process by which introns, regions of RNA that do not code for protein, are removed from the pre-mRNA and the remaining exons connected to re-form a single continuous molecule. Although most RNA splicing occurs after the complete synthesis and end-capping of the pre-mRNA, transcripts with many exons can be spliced co-transcriptionally. The splicing reaction is catalyzed by a large protein complex called the spliceosome assembled from proteins and small nuclear RNA molecules that recognize splice sites in the pre-mRNA sequence. Many pre-mRNAs, including those encoding antibodies, can be spliced in multiple ways to produce different mature mRNAs that encode different protein sequences. This process is known as alternative splicing, and allows production of a large variety of proteins from a limited amount of DNA.



Q4 What is translation? Describe in detail the major steps of translation.

Ans. Translation is the third stage of protein biosynthesis. In translation, messenger RNA (mRNA) produced by transcription is decoded by the ribosome to produce a specific amino acid chain, or polypeptide, that will later fold into an active protein. In Bacteria, translation occurs in the cell's cytoplasm, where the large and small subunits of the ribosome are located, and bind to the mRNA. In Eukaryotes, translation occurs across the membrane of the endoplasmic reticulum in a process called vectorial synthesis. The ribosome facilitates decoding by inducing the binding of tRNAs with complementary anticodon sequences to that of the mRNA. The tRNAs carry specific

amino acids that are chained together into a polypeptide as the mRNA passes through and is "read" by the ribosome in a fashion reminiscent to that of a stock ticker and ticker tape.

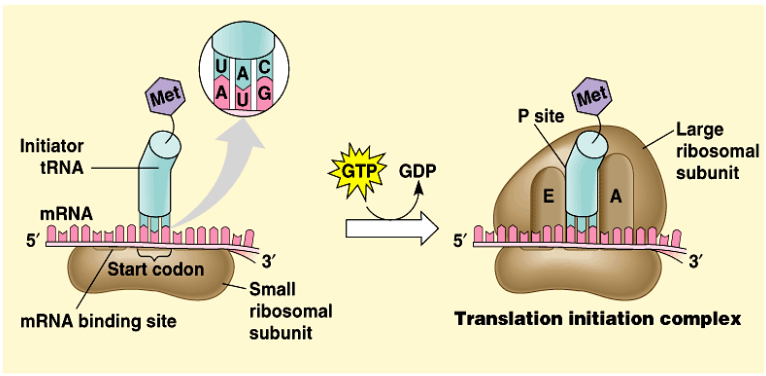
Translation proceeds in four phases: activation, initiation, elongation and termination (all describing the growth of the amino acid chain, or polypeptide that is the product of translation). Amino acids are brought to ribosomes and assembled into proteins.

In activation, the correct amino acid is covalently bonded to the correct transfer RNA (tRNA). The amino acid is joined by its carboxyl group to the 3' OH of the tRNA by an ester bond. When the tRNA has an amino acid linked to it, it is termed "charged". Initiation involves the small subunit of the ribosome binding to the 5' end of mRNA with the help of initiation factors (IF). Termination of the polypeptide happens when the A site of the ribosome faces a stop codon (UAA, UAG, or UGA). No tRNA can recognize or bind to this codon. Instead, the stop codon induces the binding of a release factor protein that prompts the disassembly of the entire ribosome/mRNA complex.

The Steps of Translation

1. Initiation

- The small subunit of the ribosome binds to a site "upstream" (on the 5' side) of the start of the message.
- It proceeds downstream (5' -> 3') until it encounters the start codon AUG. (The region between the mRNA cap and the AUG is known as the 5'-untranslated region [5'-UTR].)
- Here it is joined by the large subunit and a special initiator tRNA.
- The initiator tRNA binds to the P site (shown in pink) on the ribosome.
- In eukaryotes, initiator tRNA carries methionine (Met). (Bacteria use a modified methionine designated fMet.)

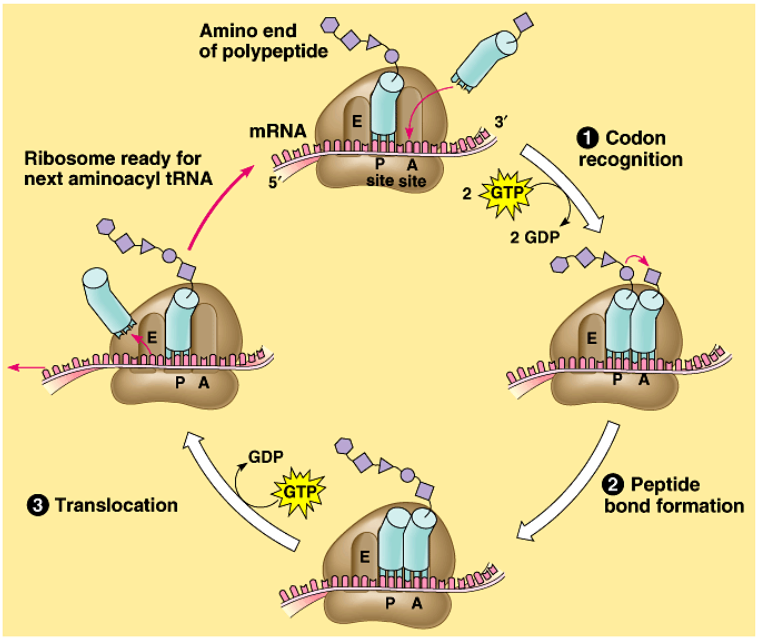


2. Elongation

- An aminoacyl-tRNA (a tRNA covalently bound to its amino acid) able to base pair with the next codon on the mRNA arrives at the A site (green) associated with:
    - an elongation factor (called EF-Tu in bacteria; EF-1 in eukaryotes)
    - GTP (the source of the needed energy)
- The preceding amino acid (Met at the start of translation) is covalently linked to the incoming amino acid with a peptide bond (shown in red).
- The initiator tRNA is released from the P site.
- The ribosome moves one codon downstream.
- This shifts the more recently-arrived tRNA, with its attached peptide, to the P site and opens the A site for the arrival of a new aminoacyl-tRNA.
- This last step is promoted by another protein elongation factor (called EF-G in bacteria; EF-2 in eukaryotes) and the energy of another molecule of GTP.

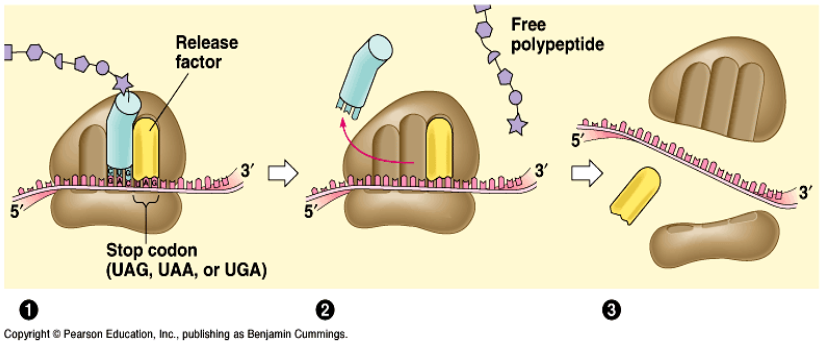
The initiator tRNA is the only member of the tRNA family that can bind directly to the P site. The P site is so-named because, with the exception of initiator tRNA, it binds only to a peptidyl-tRNA molecule; that is, a tRNA with the growing peptide attached.

The A site is so-named because it binds only to the incoming aminoacyl-tRNA; that is the tRNA bringing the next amino acid. So, for example, the tRNA that brings Met into the interior of the polypeptide can bind only to the A site.



3. Termination

- The end of translation occurs when the ribosome reaches one or more STOP codons (UAA, UAG, UGA). (The nucleotides from this point to the poly(A) tail make up the 3'-untranslated region [3'-UTR] of the mRNA.)
- There are no tRNA molecules with anticodons for STOP codons.
- However, protein release factors recognize these codons when they arrive at the A site.
- Binding of these proteins —along with a molecule of GTP — releases the polypeptide from the ribosome.
- The ribosome splits into its subunits, which can later be reassembled for another round of protein synthesis.



Section-C

Biotechnology

Q1. Write short notes on:

a) **Biotechnology**

b) **Plant tissue culture and its application**

c) **Tissue Culture Laboratory**

d) **Nutrient Media for plant tissue culture**

e) **Totipotency**

f) **Differentiation**

g) **Morphogenesis**

h) **Protoplast Culture**

i) **Hybrid and Cybrid**

j) **Somatic embryogenesis**

k) **Anther Culture**

l) **Embryo culture**

Ans. **a) Biotechnology** (sometimes shortened to "**biotech**") is a field of applied biology that involves the use of living organisms and bioprocesses in engineering, technology, medicine and other fields requiring bioproducts. Biotechnology also utilizes these products for manufacturing purpose. Modern use of similar terms includes genetic engineering as well as cell- and tissue culture technologies. The concept encompasses a wide range of procedures (and history) for modifying living organisms according to human purposes — going back to domestication of animals, cultivation of plants, and "improvements" to these through breeding programs that employ artificial selection and hybridization. By comparison to biotechnology, bioengineering is generally thought of as a related field with its emphasis more on higher systems approaches (not necessarily altering or using biological materials *directly*) for interfacing with and utilizing living things. The United Nations Convention on Biological Diversity defines biotechnology as:

"Any technological application that uses biological systems, living organisms, or derivatives thereof, to make or modify products or processes for specific use."

In other terms: "Application of scientific and technical advances in life science to develop commercial products" is biotechnology. Biotechnology draws on the pure biological sciences (genetics, microbiology, animal cell culture, molecular biology, biochemistry, embryology, cell biology) and in many instances is also dependent on knowledge and methods from outside the sphere of biology (chemical engineering, bioprocess engineering, information technology, biorobotics). Conversely, modern biological sciences (including even concepts such as molecular ecology) are intimately entwined and dependent on the methods developed through biotechnology and what is commonly thought of as the life sciences industry.

b)Plant tissue culture is a collection of techniques used to maintain or grow plant cells, tissues or organs under sterile conditions on a nutrient culture medium of known composition. Plant tissue culture is widely used to produce clones of a plant and the method is known as micropropagation. Different techniques in plant tissue culture may offer certain advantages over traditional methods of propagation, including:

- The production of exact copies of plants that produce particularly good flowers, fruits, or have other desirable traits.
- To quickly produce mature plants.
- The production of multiples of plants in the absence of seeds or necessary pollinators to produce seeds.
- The regeneration of whole plants from plant cells that have been genetically modified.
- The production of plants in sterile containers that allows them to be moved with greatly reduced chances of transmitting diseases, pests, and pathogens.
- The production of plants from seeds that otherwise have very low chances of germinating and growing, i.e.: orchids and nepenthes.
- To clean particular plant of viral and other infections and to quickly multiply these plants as 'cleaned stock' for horticulture and agriculture.

Plant tissue culture relies on the fact that many plant cells have the ability to regenerate a whole plant (totipotency). Single cells, plant cells without cell walls (protoplasts), pieces of leaves, or (less commonly) roots can often be used to generate a new plant on culture media given the required nutrients and plant hormones.

Applications

Plant tissue culture is used widely in plant science; it also has a number of commercial applications. Applications include:

- Micropropagation is widely used in forestry and in floriculture. Micropropagation can also be used to conserve rare or endangered plant species.
- A plant breeder may use tissue culture to screen cells rather than plants for advantageous characters, e.g. herbicide resistance/tolerance.

- Large-scale growth of plant cells in liquid culture in bioreactors for production of valuable compounds, like plant-derived secondary metabolites and recombinant proteins used as biopharmaceuticals.
- To cross distantly related species by protoplast fusion and regeneration of the novel hybrid.
- To cross-pollinate distantly related species and then tissue culture the resulting embryo which would otherwise normally die (Embryo Rescue).
- For production of doubled monoploid (dihaploid) plants from haploid cultures to achieve homozygous lines more rapidly in breeding programmes, usually by treatment with colchicine which causes doubling of the chromosome number.
- As a tissue for transformation, followed by either short-term testing of genetic constructs or regeneration of transgenic plants.
- Certain techniques such as meristem tip culture can be used to produce clean plant material from virused stock, such as potatoes and many species of soft fruit.
- micropropagation using meristem and shoot culture to produce large numbers of identical individuals.
- production of identical sterile hybrid species can be obtained

c) Tissue Culture Laboratory. For the tissue culture a good laboratory, similar to a microbiological laboratory, is required which must have facilities for: *(i)* nutrient medium preparation, sterilization, cleaning and storage of supplies, *(ii)* aseptic condition for working the living materials, *(iii)* a controlled environmental conditions for growth and development of cultures, *(iv)* observation and evaluation of culture as hoped, and *(v)* recording the observations made during the experiment.

Nutrient media could be prepared in a separate room where sufficient space is available for keeping and weighing the chemicals, and putting glassware. Vitamins and growth hormones are carefully weighed. Stock solutions of chemicals are kept in refrigerator to avoid contamination. For sterilization of glassware, working tables, nutrient media and plant materials different techniques are applied. Most common methods of sterilization are:

(i) *Use of Chemicals :* For example, chromic acid, mercuric chloride (0.1%), sodium hypochlorite or calcium hypochlorite (each 0.5%), alcohol (70%), etc. are used for sterilization of glassware, working table and living materials.

(ii) *Use of Dry Heat:* Glassware and metal instruments are kept in oven. Sterilization is done by keeping high temperature between 200-300°C for 1h.

(iii) *Use of Wet heat or Steam:* Distilled wafer, nutrient media, etc.are sterilized in an autoclave at 121°C for 30 min.

(iv) *Ultrafiltration :* Only those chemicals *e.g.* vitamins, growth hormones, etc., which are unstable at high temperature are sterilized by Seitz filter, milipore membrane filter, etc.

(v) *By Ultraviolet Light:* It is used in the inoculation chamber or laminar flow (Fig. 8.1).

(vi) *Using Antibiotics:* These are used only in case of animal cell culture, not for plant cells. Inoculation (transfer of sterilized biological material on sterile nutrient medium in an aseptic condition) of plant or animal material is done in an

inoculation chamber/laminar flow where ultraviolet lamp, 70% ethyl alcohol, spirit lamp, inoculation needle, glassware, etc. are kept in advance.

The glassware (culture tubes or Erlenmeyer conical flasks) containing sterilized nutrient medium after inoculation with plant material, are plugged with non-absorbent sterile cotton, and finally kept in growth chamber. Environmental conditions such as temperature, moisture and light are controlled as needed. For cell suspension culture, orbital shaker or aeration instrument is placed in growth chamber. Generally temperature is set between 25--27°C. Environmental conditions may vary according to plant species and nature of experiment. Light intensity, quality and photoperiod (light and dark cycles) are also regulated. Other laboratory accessories are dissection microscope, compound microscope, centrifuge, first aid kit, etc. which are required during *in vitro* culture of plant material.

e) **Nutrient Media. Composition and Preparation** Vital activity of a cell is the absorption of nutrients through cell membrance and rapid proliferation into innumerable cells. White (1934) observed the unlimited growth of isolated root tissues when provided with nutrient medium containing inorganic salts, sucrose, vitamins, growth hormone and a few amino acids. Composition of different nutrient media is given in Table 8.1.

Inorganic Minerals

Inorganic nutrients include macronutrients *(e.g.* nitrogen, phosphorus, potassium, calcium, magnesium and sulphur) in the form of salts in large amount and microelements *(e.g.* boron, molybdenum, copper, zinc, manganese, iron and chloride). A concentrated stock solution is prepared in advance and finally added to medium as required. To overcome the problem of solubility, the stock solution of iron is prepared in a chelated form as the sodium salt of ferric ethylenediamme tetra acetate (Fe-EDTA).

Growth Hormones

Several growth hormones are known which stimulate the biological activity in cultured materials. Cytokinins promote cell division and regulate growth and development similar to kinetin (6-furfuryl aminopurine). Auxin resembles indole acetic acid (IAA) and stimulates shoot elongation. Gibberellins are of less importance, however, GA_3 is used in apical meristem (Morel and Muller, 1964). The most widely used cytokinins are adenin, kinetin, zeatin, benzyladenin; and auxins are IAA, NAA (a -Naphthalene acetic acid) 2,4-D. For the induction of callus the amount of kinetin should be 0.1 mg/liter.

Organic Constituents

The organic compounds serve a source of carbon and energy. They are used in high concentration, *e.g.* 20-30g/liter. Sucrose and D-glucose (carbohydrates) are commonly used; but glycerol and myoinositol are also the principal source of carbon.

Other complex organic compounds are peptone, yeast extract, malt extract, coconut water, tomato juice, etc.

Vitamins

Vitamins are required in trace amount as they catalyze the enzyme system of the cells. Vitamin B_1 (thiamine) is the most commonly used vitamin for all plant tissue cultures. Other groups of vitamins which stimulate growth are niacin (nicotinic acid) vitamin B_2 (riboflavin) vitamin B_6 (pyrodoxin) (Gamborg *et al.*, 1976), vitamin C (ascorbic acid), vitamin H (biotin) and vitamin B_{12} (cyanocobalamin) (Huang and Murashige, 1977).

Amino acids

Although nitrogen sources are present in the inorganic salts, yet various amino acids and amides are used in plant tissue culture media as given by Huang and Murashige (1977). The most widely used amino acids are L-aspartic acid, L-asparagin, L-glutamic acid, L-glutamine and L-arginin.

e) Totipotency is the ability of a single cell to divide and produce all the differentiated cells in an organism, including extraembryonic tissues. Totipotent cells include spores and zygotes. In some organisms, cells can dedifferentiate and regain totipotency. For example, a plant cutting or callus can be used to grow an entire plant.

Human development begins when a sperm fertilizes an egg and creates a single totipotent cell (zygote). In the first hours after fertilization, this cell divides into identical totipotent cells, which can later develop into any of the three germ layers of a human (endoderm, mesoderm, or ectoderm) and into cells of the cytotrophoblast layer or syncytiotrophoblast layer of the placenta. After reaching the 16-cell stage, the totipotent cells of the morula differentiate into cells that will eventually become either the Blastocyst's Inner cell mass or outer trophoblasts. Approximately four days after fertilization and after several cycles of cell division, these totipotent cells begin to specialize. The inner cell mass, the source of embryonic stem cells, is pluripotent, not totipotent.

f)Differentiation. Cellular differentiation is the process by which a less specialized cell becomes a more specialized cell type. Differentiation occurs numerous times during the development of a multicellular organism as the organism changes from a simple zygote to a complex system of tissues and cell types. Differentiation is a common process in adults as well: adult stem cells divide and create fully differentiated daughter cells during tissue repair and during normal cell turnover. Differentiation dramatically changes a cell's size, shape, membrane potential, metabolic activity, and responsiveness to signals. These changes are largely due to highly controlled modifications in gene expression. With a few exceptions, cellular differentiation almost never involves a change in the DNA sequence itself. Thus,

different cells can have very different physical characteristics despite having the same genome.

A cell that is able to differentiate into all cell types of the adult organism is known as *pluripotent*. Such cells are called embryonic stem cells in animals and meristematic cells in higher plants. A cell that is able to differentiate into all cell types, including the placental tissue, is known as *totipotent*. In mammals, only the zygote and subsequent blastomeres are totipotent, while in plants many differentiated cells can become totipotent with simple laboratory techniques.

g) Morphogenesis (from the Greek *morphê* shape and *genesis* creation, literally, "beginning of the shape"), is the biological process that causes an organism to develop its shape. It is one of three fundamental aspects of developmental biology along with the control of cell growth and cellular differentiation.

The process controls the organized spatial distribution of cells during the embryonic development of an organism. Morphogenesis can take place also in a mature organism, in cell culture or inside tumor cell masses. Morphogenesis also describes the development of unicellular life forms that do not have an embryonic stage in their life cycle, or describes the evolution of a body structure within a taxonomic group. Morphogenetic responses may be induced in organisms by hormones, by environmental chemicals ranging from substances produced by other organisms to toxic chemicals or radionuclides released as pollutants, and other plants, or by mechanical stresses induced by spatial patterning of the cells.

h) Protoplast culture.

Isolation of protoplasts

Protoplasts (cell minus cell wall) is the biologically active and most significant material of cells. When cell wall is mechanically or enzymatically removed the isolated protoplast is known as "naked plant cell" on which most of recent researches are based.

Plant cell wall acts as physical barrier and protects cytoplasm from microbial invasion and environmental stress. It consists of a complex mixture of cellulose, hemicellulose, pectin, lignin, lipids, protein, etc. For dissolution of different components of the cell wall it is essential to have the respective enzymes.

Until suitable methods were developed, protoplasts were isolated by cutting the plasmolysed plant tissues and releasing protoplast through deplasmolysis of cells. Cooking (1960) for the first time isolated the protoplasts of plant tissues by using cell wall degrading enzymes viz., cellulase, hemicellulase, pectinase, and protease extracted from a saprophytic fungus *Trichoderma viride.* Later on protoplasts were cultured *in vitro.*

Microorganisms are well equipped with a system to produce substrate specific extracellular enzymes, the extent of which depends on the genetic variability of the

specific species and strains. However, the basic techniques of isolation and culture of protoplast are given:

(i) *Surface Sterilization of Leaf Samples*: Mature leaves are collected from healthy plants which are washed in tap water to remove adhering soil particles and sterilized with sodium hypochlorite solution.

(ii) *Rinsing in Suitable Osmoticum*: After 10 min, sample is properly washed with sterile distilled water or MS medium adjusted to a suitable pH and buffer to maintain osmotic pressure. Washing should be done for about 6 times to remove the traces of sodium hypochlorite.

(iii) *Plasmolysis of Cells*: The lower epidermis covered by thin wax cuticle is removed with a forcep. Stripping should be done from midrib to margin of lamina. The stripped surface of leaf is kept in mannitol solution (13% W/V) for 3 hours to allow plasmolysis of cells.

(iv) *Peeling of Lower Epidermis:* Thereafter, about 1 gm leaves are peeled off and transferred into enzyme mixture already sterilized through a Seitz filter (0.45 mm). This facilitates the penetration of enzyme into tissue within 12-18 hours at 25°C.

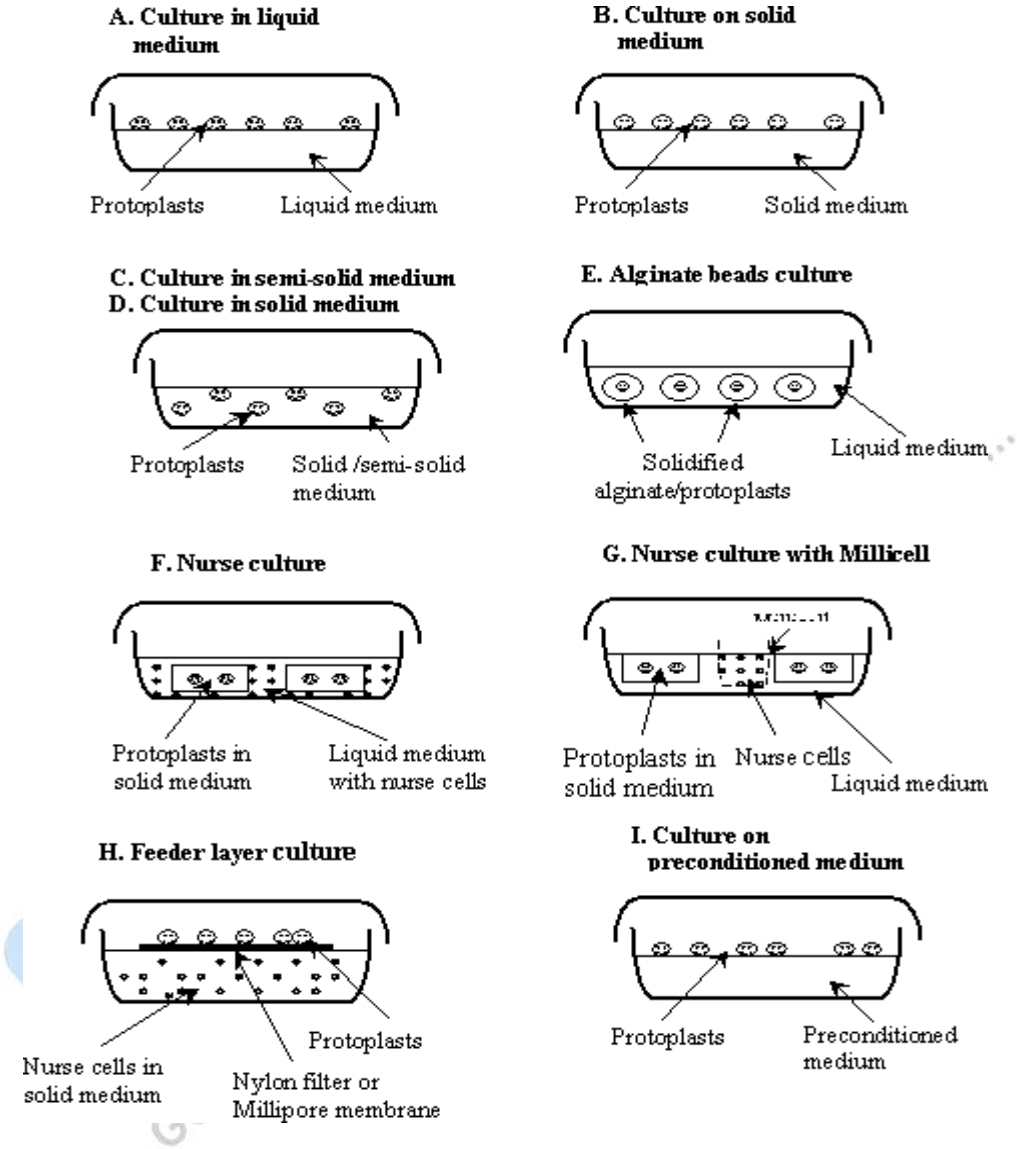
(v) *Isolation and Purification of Portoplasts*: Leaf debris are removed with forcep, and enzyme solution containing protoplasts are filtered with a nylon mesh (45mm). Filtrate is centrifuged at 75 X g for 5 min and supernatant is decanted. Again a fresh MS medium plus 13% mannitol is added to centrifuge. Repeated washing with nutrient medium, centrifugation and decantation are done for about three time. Finally specific concentration of protoplast suspension is prepared.

Protoplast culture and regeneration

From the protoplast solution of known density (about 10^5 protoplast/ml) about 1 ml suspension is poured on sterile and cooled down nutrient medium in Petri dishes. The plates are incubated at 25°C in a dim white light.

Culture in liquid media: For culturing protoplasts in liquid media, protoplasts with a density of 10^5-10^6 per millilitre are suspended in liquid MS media, with or without shaking.

Figure: The different protoplast culture techniques



Culture on solid medium: For culture on solid medium protoplasts are directly plated at high density on a gelatinous medium or on a membrane which covers the gelatinous medium. Such media are composed of:

Culture in semi-solid medium: For culture on semi-solid medium protoplasts are cultivated on medium containing 0.6% agarose with other required nutrients.

Culture embedded in solid medium: For embedment in solid medium (Figure 2D) protoplasts at twice the desired concentration (2×10^6) are placed in a twofold concentrated culture medium: MS minor nutrients (× 2), MS major nutrients, MS

vitamins (× 2), 20 mg/l ascorbic acid, 40 g/l sucrose, 10 µM 2,4-D, but with 0.55 M D-mannitol. This is followed by mixing with an equal volume of solution containing a gelling gum at 0.6%, 0.55 M D-mannitol, previously melted and maintained at 55°C [45].

Culture in alginate or solid medium floating in liquid medium (*bead type culture*): Protoplasts are cultured in a solid medium as in section 3.6.4, or in calcium alginate beads. For this purpose, the alginate solution (2.8% alginic acid, 0.4 M mannitol) is first autoclaved. Protoplasts are added to the alginate solution to obtain a concentration of 1.0×10^6 protoplasts per millilitre. Later, drops of the protoplast-alginate mixture are polymerised in 0.4 M mannitol/50 mM $CaCl_2$ solution, and transferred into protoplast liquid culture medium (see below) in small Petri dishes. Pieces of solid medium containing protoplasts are transferred into liquid MS medium.

Nurse cultures: Protoplasts can be maintained in MS media containing living cells as nurse in liquid medium.

Feeder-layer culture of living cells: Feeder cultures are prepared one day before protoplast isolation by mixing 3 ml of feeder cell suspension with 100 ml of a medium containing essential nutrients. Feeder cells are then separated from overlaying protoplasts by a 10 mm nylon filter or 0.22 mm Millipore membrane.

Culture on preconditioned medium: This cultivation technique resembles the feeder-layer technique, except that the preconditioned medium is the cell culture filtered through a Whatman filter, followed by mixing the filter-sterilized medium with the solidifying medium

The protoplasts regenerate a cell wall, undergo cell division and form callus. The callus can also be subcultured. Embryogenesis begins from callus when it is placed on nutrient medium lacking mannitol and auxin. The embryo develops into seedlings and finally mature plants.

i) Hybrid and Cybrid.

Fusion of cytoplasm of two protoplasts results in coalescence of cytoplasms. The nuclei of two protoplasts may or may not fuse together even after fusion of cytoplasms. The binucleate cells are known as heterokaryon or heterocyte. When nuclei are fused the cells are known as hybrid or synkaryocyte and when only cytoplasms fuse and genetic information from one of the two nuclei is lost is known as cybrid *i.e.* cytoplasmic hybrid or heteroplast. There are some genetic factors which are carried in cytoplasmic inheritance, instead of nuclear genes, for example, male sterility in some plants. Susceptibility and resistance to some of the pathotoxins and drug are controlled by cytoplasmic genes. Therefore, production of cybrids which contain the mixture of cytoplasms but only one nuclear genome can help in transfer of cyloplasmic genetic information from one plant to another. Thus, informations of cybrid can be applicable in plant breeding experiments.

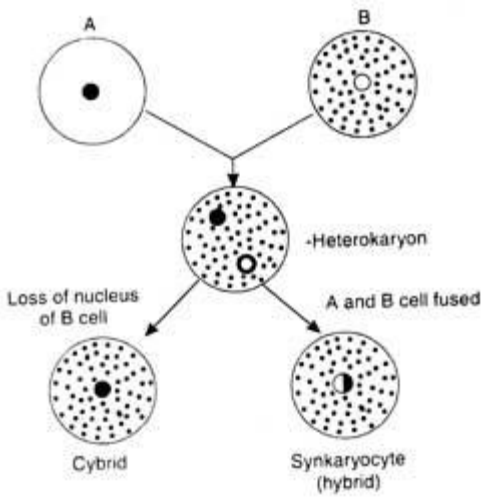


Figure. Hybrid/cybrid production through protoplast fusion.

j) Somatic embryogenesis. Embryo production is a characteristic feature of the flowering plants. However, such structures (embryoids) have also been artificially induced in cultured plant tissues, besides zygote. Somatic embryogenesis was first induced in suspension culture (Stewart *et al.* 1958) and callus culture (Reinert, 1959) of carrot. More than 30 plant families are known so far where somatic embryoids have been induced (Raghavan, 1976; Ammirato, 1983).

Somatic embryogenesis can be initiated in two ways : *(i)* by inducing embryogenic cells within the preformed callus, and *(ii)* directly from preembryonic determined cell, (without callus) which are ready to differentiate into embryoids (Sharp *et al* 1980). In the first case, ambryoids are initiated in callus from superficial cell aggregates where cells contain a large vacuole, dense cytoplasm, large starch granules and nucleus (Me Willian *et al.*, 1974).

Two nutritional media of different composition are required to obtain embryoids. First medium contains auxin to initiate embryogenic cells. Second medium lacks auxin or reduced level of auxin is needed for subsequent development of the embryonic cells into embryoids and plantlets. In both the cases reduced amount of nitrogen is required (Ammirato, 1983). The embryogenic cells pass through 3 different stages *e.g.* globular, heart shaped and torpedo shaped, to form embryoids.

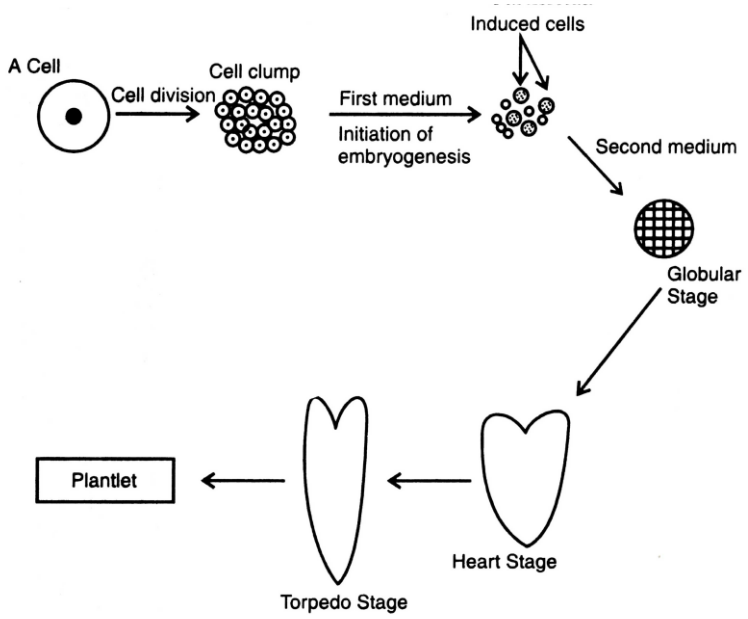


Figure. Events of somatic embryogenesis

These embryoids can be separated and isolated mechanically by using glassbeads. When embryoids reach torpedo stage they are transferred to filter paper bridge (a sterile and pluged culture tube containing about 10 ml MS liquid medium supplemented with Kinetin (0.2 mg/lit) and sucrose (2% W/V) on which Whatman No.1. Filter paper is placed to make a bridge (Dodds and Roberts, 1985). Some plants in which somatic embryogenesis has been induced *in vitro* are *Atropa belladona, Brassica oleracea, Carica papaya, Coffea arabica, Citrus cinensis, Daucus carrota, Nicotiana tabacum, Pinus ponderosa and Saccharuni officinarum.* Applications of this process include: clonal propagation of genetically uniform plant material; elimination of viruses; provision of source tissue for genetic transformation; generation of whole plants from single cells called protoplasts; development of synthetic seed technology.

k) Anther Culture. Anther, a male reproductive organ, is diploid in chromosome numbers. As a result of microsporogenesis, tetrads of microspores are formed from a single spore mother cell. Anter culture is the process of using anthers to culture haploid plantlets. The aim of anther and pollen culture is to get haploid plants by induction of embryogenesis. Haploid plants have single complete set of chromosomes that in turn may be useful for the improvement of many crop plants (Sunderland, 1979).

The technique was discovered in 1964 by Guha and Maheshwari. This technique can be used in over 200 species, including tomato, rice, tobacco, barley, and geranium. Some of the advantages which make this a valuable method for obtaining haploid plants are:

- the technique is fairly simple
- it is easy to induce cell division in the immature pollen cells in some species
- a large proportion of the anthers used in culture respond (induction frequency is high)
- haploids can be produced in large numbers very quickly.

Sunderland (1979) has described that the anthers to be cultured should be one of the three categories *i.e.* premitotic, mitotic and postmitotic. In premitotic anthers, where the microspores have completed meiosis but not started first pollen division, the best response is achieved *e.g. Hordeum vulgare.* Haploid plants are very useful in *(i)* direct screening of recessive mutation because in diploid or polyploid screening of recessive mutation is not possible, and *(ii)* development of homozygous diploid plants following chromosome doubling of haploid plant cells.

Culturing techniques Anthers are superficially sterilized and washed with double distilled sterile water. They are excised from the flower buds and their proper developmental stages are determined under microscope. On confirmation of a stage, *(a)* the anthers are directly transferred on nutrient agar or liquid medium where induction of embryogenesis occurs, or *(b)* the pollen grains are aseptically removed from the anthers and cultured on liquid medium

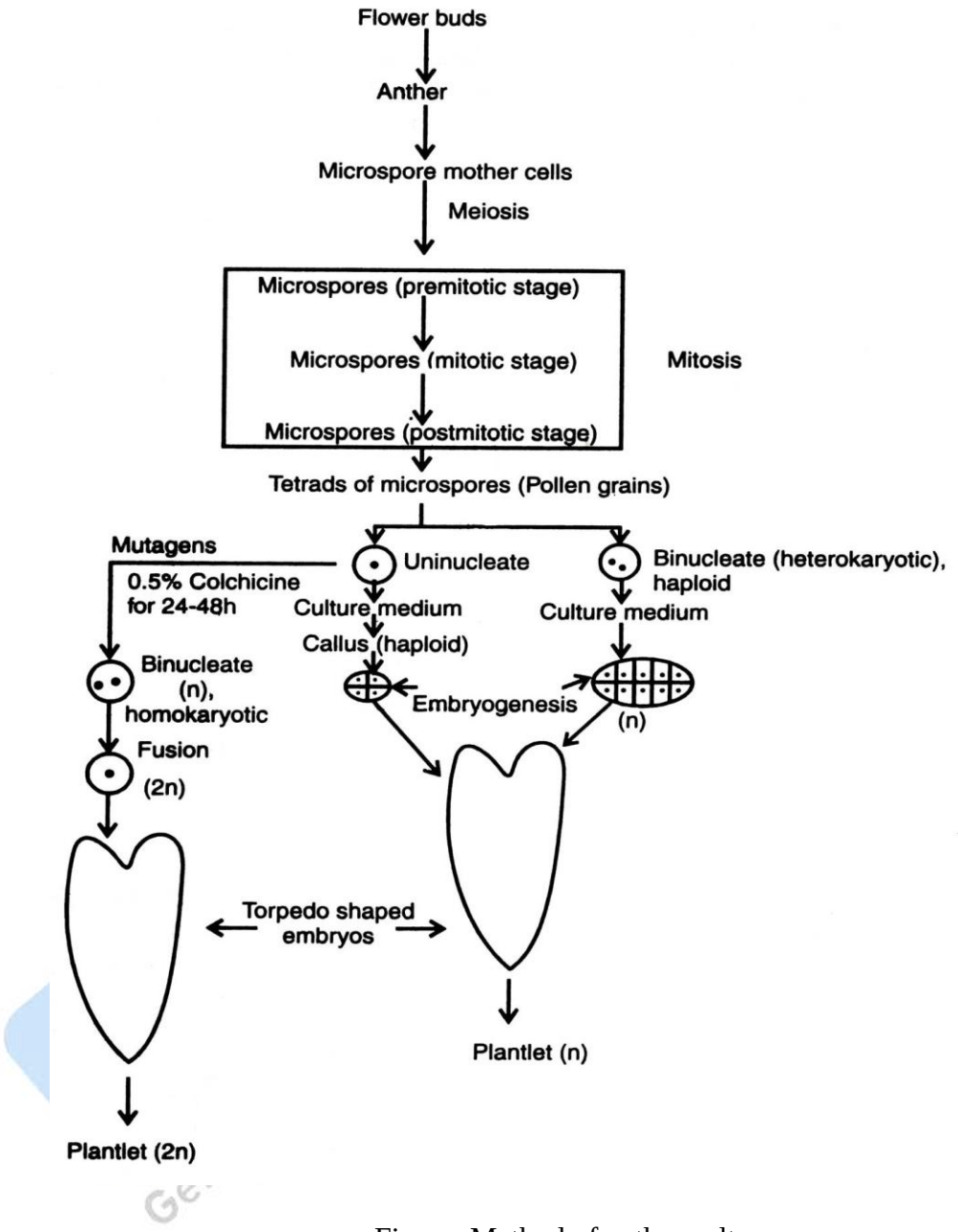


Figure: Method of anther culture

m) **Embryo culture.** In addition to root, shoot, and pollen culture, embryo culture has also been done for the production of haploid plants. Embryo culture is used for the recovery of plants from distinct crosses. Embryo culture is useful where embryo fails to develop due to degeneration of embryonic tissues. It is being used extensively in the extraction of haploid barley *(Hordeum vulgare)* from the crosses *H. vulgare* x *H. bulbosum.* Embryo culture is also a routine technique employed in orchid propagation and in breeding of those species that show dormancy. Das and Barman (1992) developed the method of regeneration of tea shoots from embryo callus. The embryo callus produced somatic embryoids

within 8 weeks of culture in the second medium which differentiated into buds after 2 weeks. Several shoots with 4-6 leaves developed after 16 weeks of culture.

Culturing method

The general method of embryo culture follows the following steps.

(i) Pluck healthy and mature fruits from the field and wash thoroughly in running water for about an hour.

(ii) Surface sterilize with 0.01% Tween-20 for 15 min, rinse seeds several times with distilled water and finally treat with 0.01% $HgCl_2$ solution for 10-15 min. Finally rinse it for six times with sterile distilled water.

(iii) Break seeds aseptically and isolate the embryo.

(iv) Culture embryo on callus proliferation medium. Supplement the basal medium of Murashige and Skoog (1962) with different combinations and concentrations of sugar, vitamins, hormones and other growth adjuvants for callus proliferation and shoot regeneration.

(v) Incubate the cultures at 22-25°C under a 16 h photoperiod of 2000 lux luminous intensity.

(vi) After two weeks of inoculation the embryo begins to swell on callus proliferation medium. Distinct callus growth is observed after 4 weeks.

(vii) After 8 weeks of inoculation transfer the callus on shoot regeneration medium. Within 4 weeks of transfer into second medium the callus turns green and produces soft spongy tissue. Some of these tissues are differentiated into embryoids.

(viii) The embryoids produce cluster of budlets when subcultured onto shoot regeneration medium. The budlets grow into shoots and produce 2-3 leaf appendages within 12 weeks. Thereafter, they are separated into individual shoots and then subcultured into a fresh medium of the same composition until shoots develop.

Q2. Explain the process of Micropropagation along with its advantages and disadvantages.

Ans. Micropropagation is the practice of rapidly multiplying stock plant material to produce a large number of progeny plants, using modern plant tissue culture methods. Micropropagation is used to multiply novel plants, such as those that have been genetically modified or bred through conventional plant breeding methods. It is also used to provide a sufficient number of plantlets for planting from a stock plant which does not produce seeds, or does not respond well to vegetative reproduction.

Methods

Establishment. Micropropagation begins with the selection of plant material to be propagated. Clean stock materials that are free of viruses and fungi are important in

the production of the healthiest plants. Once the plant material is chosen for culture, the collection of explant(s) begins and is dependent on the type of tissue to be used; including stem tips, anthers, petals, pollen and others plant tissues. The explant material is then surface sterilized, usually in multiple courses of bleach and alcohol washes and finally rinsed in sterilized water.

This small portion of plant tissue, sometimes only a single cell, is placed on a growth medium, typically containing sucrose as an energy source and one or more plant growth regulators (plant hormones). Usually the medium is thickened with agar to create a gel which supports the explant during growth. Some plants are easily grown on simple media but others require more complicated media for successful growth; The plant tissue grows and differentiates into new tissues depending on the medium. For example, media containing cytokinins are used to create branched shoots from plant buds. and it happens in a vegetative form

Multiplication. Multiplication is the taking of tissue samples produced during the first stage and increasing their number. Following the successful introduction and growth of plant tissue, the establishment stage is followed by multiplication. Through repeated cycles of this process, a single explant sample may be increased from one to hundreds or thousands of plants. Depending on the type of tissue grown, multiplication can involve different methods and media. If the plant material grown is callus tissue, it can be placed in a blender and cut into smaller pieces and recultured on the same type of culture medium to grow more callus tissue. If the tissue is grown as small plants called plantlets, hormones are often added that cause the plantlets to produce many small offshoots that can be removed and recultured.

Pretransplant. Plantlets transferred to soil (with vermicompost) from plant media. This process is done for acclimatization of plantlets to the soil as they were previously grown in plant media. After growing for some days the plantlets are transferred to the field. This stage involves treating the plantlets/shoots produced to encourage root growth and "hardening." It is performed *in vitro*, or in a sterile "test tube" environment. "Hardening" refers to the preparation of the plants for a natural growth environment. Until this stage, the plantlets have been grown in "ideal" conditions, designed to encourage rapid growth. Hardening typically involves slowly weaning the plantlets from a high-humidity, low light, warm environment to what would be considered a normal growth environment for the species in question. This is done by moving the plants to a location high in humidity.

Transfer from culture. In the final stage of plant micropropagation, the plantlets are removed from the plant media and transferred to soil or (more commonly) potting compost for continued growth by conventional methods.

Advantages

Micropropagation has a number of advantages over traditional plant propagation techniques:

- The main advantage of micropropagation is the production of many plants that are clones of each other.
- Micropropagation can be used to produce disease-free plants.
- Micropropagation produces rooted plantlets ready for growth, saving time for the grower when seeds or cuttings are slow to establish or grow.
- It can have an extraordinarily high fecundity rate, producing thousands of propagules while conventional techniques might only produce a fraction of this a number.
- It is the only viable method of regenerating genetically modified cells or cells after protoplast fusion.
- It is useful in multiplying plants which produce seeds in uneconomical amounts, or when plants are sterile and do not produce viable seeds or when seed can't be stored (vgr. recalcitrant seeds).

Disadvantages

Micropropagation is not always the perfect means of multiplying plants. Conditions that limits its use include:

- It is *very* expensive, and can have a labor cost of more than 70%
- A monoculture is produced after micropropagation, leading to a lack of overall disease resilience, as all progeny plants may be vulnerable to the same infections.
- An infected plant sample can produce infected progeny. This is uncommon if the stock plants are carefully screened and vetted to prevent culturing plants infected with virus or fungus.
- Not all plants can be successfully tissue cultured, often because the proper medium for growth is not known or the plants produce secondary metabolic chemicals that stunt or kill the explant.
- Sometimes plants or cultivars do not come true to type after being tissue cultured. This is often dependent on the type of explant material utilized during the initiation phase or the result of the age of the cell or propagule line.
- Some plants are very difficult to disinfest of fungal organisms.

Section-D

Q1. Write short notes on:

a) Restriction enzymes
b) Plasmids
c) Cosmids
d) pBR322
e) cDNA library

Ans. a) Restriction enzymes.

A **Restriction Enzyme** (or **restriction endonuclease**) is an enzyme that cuts double-stranded DNA at specific recognition nucleotide sequences known as restriction sites. Such enzymes, found in bacteria and archaea, are thought to have evolved to provide a defense mechanism against invading viruses. Inside a bacterial host, the restriction enzymes selectively cut up *foreign* DNA in a process called *restriction*; host DNA is methylated by a modification enzyme (a methylase) to protect it from the restriction enzyme's activity. Collectively, these two processes form the restriction modification system. To cut the DNA, a restriction enzyme makes two incisions, once through each sugar-phosphate backbone (i.e. each strand) of the DNA double helix.

Types

Restriction endonucleases are categorized into four general groups (Types I, II, III and IV) based on their composition and enzyme cofactor requirements, the nature of their target sequence, and the position of their DNA cleavage site relative to the target sequence. All types of enzymes recognise specific short DNA sequences and carry out the endonucleolytic cleavage of DNA to give specific double-stranded fragments with terminal 5'-phosphates. They are summarised as below:

- Type I enzymes (EC 3.1.21.3) cleave at sites remote from recognition site; require both ATP and S-adenosyl-L-methionine to function; multifunctional protein with both restriction and methylase (EC 2.1.1.72) activities.
- Type II enzymes (EC 3.1.21.4) cleave within or at short specific distances from recognition site; most require magnesium; single function (restriction) enzymes independent of methylase.
- Type III enzymes (EC 3.1.21.5) cleave at sites a short distance from recognition site; require ATP (but doesn't hydrolyse it); S-adenosyl-L-methionine stimulates reaction but is not required; exist as part of a complex with a modification methylase (EC 2.1.1.72).
- Type IV enzymes target methylated DNA.

Type I

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

Type II

Typical type II restriction enzymes differ from type I restriction enzymes in several ways. They are a dimer of only one type of subunit; their recognition sites are usually undivided and palindromic and 4–8 nucleotides in length, they recognize and cleave DNA at the same site, and they do not use ATP or AdoMet for their activity—they usually require only Mg^{2+} as a cofactor. These are the most commonly available and used restriction enzymes. Type II restriction enzymes (e.g. BcgI and BplI) are multimers, containing more than one subunit. They cleave DNA on both sides of their recognition to cut out the recognition site.

Type III

Type III restriction enzymes (e.g. EcoP15) recognize two separate non-palindromic sequences that are inversely oriented. They cut DNA about 20-30 base pairs after the recognition site. These enzymes contain more than one subunit and require AdoMet and ATP cofactors for their roles in DNA methylation and restriction, respectively. They are components of prokaryotic DNA restriction-modification mechanisms that protect the organism against invading foreign DNA.

Examples of restriction enzymes include:

Enzyme	Source	Recognition Sequence	Cut
EcoRI	*Escherichia coli*	5'GAATTC 3'CTTAAG	5'---G AATTC---3' 3'---CTTAA G---5'
EcoRII	*Escherichia coli*	5'CCWGG 3'GGWCC	5'--- CCWGG---3' 3'---GGWCC ---5'
BamHI	*Bacillus amyloliquefaciens*	5'GGATCC 3'CCTAGG	5'---G GATCC---3' 3'---CCTAG G---5'
HindIII	*Haemophilus influenzae*	5'AAGCTT 3'TTCGAA	5'---A AGCTT---3' 3'---TTCGA A---5'

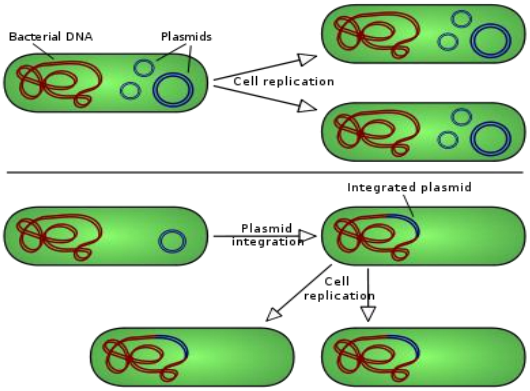
b) Plamid.

Plasmid is a DNA molecule that is separate from, and can replicate independently of, the chromosomal DNA. They are double-stranded and, in many cases, circular. Plasmids usually occur naturally in bacteria, but are sometimes found in eukaryotic organisms (e.g., the *2-micrometre ring* in *Saccharomyces cerevisiae*). Plasmid sizes vary from 1 to over 1,000 kbp. The number of identical plasmids in a single cell can range anywhere from one to even thousands under some circumstances. The term *plasmid* was first introduced by the American molecular biologist Joshua Lederberg in 1952. Plasmids are considered "replicons", capable of autonomous replication within a suitable host. Plasmid host-to-host transfer requires direct, mechanical transfer by conjugation or changes in host gene expression allowing the intentional uptake of the genetic element by transformation. Therefore, plasmids provide a mechanism for horizontal gene transfer within a population of microbes and typically provide a selective advantage under a given environmental state. Plasmids may carry genes that provide resistance to naturally occurring antibiotics in a competitive environmental niche, or the proteins produced may act as toxins under similar circumstances.

Vectors

Plasmids used in genetic engineering are called vectors. Plasmids serve as important tools in genetics and biotechnology labs, where they are commonly used to multiply (make many copies of) or *express* particular genes. Many plasmids are commercially available for such uses. The gene to be replicated is inserted into copies of a plasmid containing genes that make cells resistant to particular antibiotics and a multiple cloning site (MCS, or polylinker), which is a short region containing several commonly used restriction sites allowing the easy insertion of DNA fragments at this location. Next, the plasmids are inserted into bacteria by a process called *transformation*. Then, the bacteria are exposed to the particular antibiotics. Only bacteria that take up copies of the plasmid survive, since the plasmid makes them resistant. In particular, the protecting genes are expressed (used to make a protein) and the expressed protein breaks down the antibiotics. In this way, the antibiotics act as a filter to select only the modified bacteria. Now these bacteria can be grown in large amounts, harvested, and lysed (often using the alkaline lysis method) to isolate the plasmid of interest.

Another major use of plasmids is to make large amounts of proteins. In this case, researchers grow bacteria containing a plasmid harboring the gene of interest. Just as the bacterium produces proteins to confer its antibiotic resistance, it can also be induced to produce large amounts of proteins from the inserted gene. This is a cheap and easy way of mass-producing a gene or the protein it then codes for, for example, insulin or even antibiotics. However, a plasmid can contain inserts of only about 1–10 kbp. To clone longer lengths of DNA, lambda phage with lysogeny genes deleted, cosmids, bacterial artificial chromosomes, or yeast artificial chromosomes are used.



There are two types of plasmid integration into a host bacteria: Non-integrating plasmids replicate as with the top instance, whereas episomes, the lower example, integrate into the host chromosome.

Applications

i) Disease models

Plasmids were historically used to genetically engineer the embryonic system cells of rats in order to create rat genetic disease models. The limited efficiency of plasmid-based techniques precluded their use in the creation of more accurate human cell models. However, developments in Adeno-associated virus recombination techniques, and Zinc finger nucleases, have enabled the creation of a new generation of isogenic human disease models.

ii) Gene therapy

The success of some strategies of gene therapy depends on the efficient insertion of therapeutic genes at the appropriate chromosomal target sites within the human genome, without causing cell injury, oncogenic mutations (cancer) or an immune response. Plasmid vectors are one of many approaches that could be used for this purpose. Zinc finger nucleases (ZFNs) offer a way to cause a site-specific double-strand break to the DNA genome and cause homologous recombination. This makes targeted gene correction a possibility in human cells. Plasmids encoding ZFN could be used to deliver a therapeutic gene to a pre-selected chromosomal site with a frequency higher than that of random integration.

c) Cosmids.

A **cosmid**, first described by Collins and Hohn in 1978, is a type of hybrid plasmid (often used as a cloning vector) that contains *cos sequences*, DNA sequences originally from the Lambda phage. Cosmids can be used to build genomic libraries.

Cosmids are able to contain 37 to 52 kb of DNA, while normal plasmids are able to carry only 1–20 kb. They can replicate as plasmids if they have a suitable origin of replication: for example SV40 ori in mammalian cells,

ColE1 ori for double-stranded DNA replication or f1 ori for single-stranded DNA replication in prokaryotes. They frequently also contain a gene for selection such as antibiotic resistance, so that the transfected cells can be identified by plating on a medium containing the antibiotic. Those cells which did not take up the cosmid would be unable to grow.

Unlike plasmids, they can also be packaged in phage capsids, which allows the foreign genes to be transferred into or between cells by transduction. Plasmids become unstable after a certain amount of DNA has been inserted into them, because their increased size is more conducive to recombination. To circumvent this, phage transduction is used instead. This is made possible by the *cohesive ends*, also known as *cos* sites. In this way, they are similar to using the lambda phage as a vector, but only that *all* the lambda genes have been deleted with the exception of the *cos* sequence.

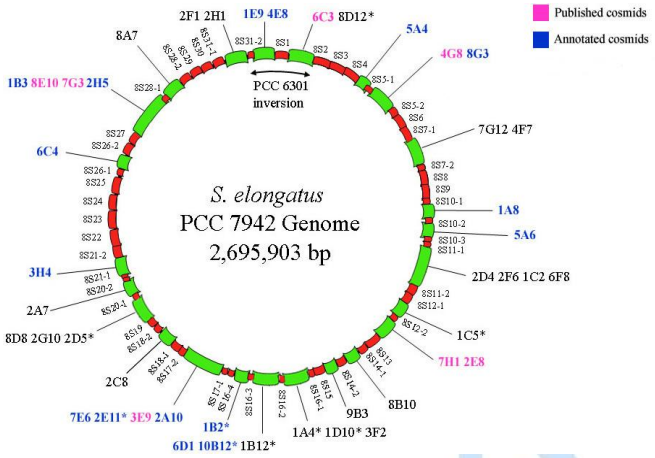
Cos sequences are ~200 base pairs long and essential for packaging. They contain a *cosN* site where DNA is nicked at each strand, 12bp apart, by terminase. This causes linearization of the circular cosmid with two "cohesive" or "sticky ends" of 12bp. (The DNA must be linear to fit into a phage head.) The *cosB* site holds the terminase while it is nicking and separating the strands. The *cosQ* site of next cosmid (as rolling circle replication often results in linear concatemers) is held by the terminase after the previous cosmid has been packaged, to prevent degradation by cellular DNases.

Cosmid features and uses

Cosmids are predominantly plasmids with a bacterial oriV, an antibiotic selection marker and a cloning site, but they carry one, or more recently two cos sites derived from bacteriophage lambda. Depending on the particular aim of the experiment broad host range cosmids, shuttle cosmids or 'mammalian' cosmids (linked to SV40 oriV and mammalian selection markers) are available. The loading capacity of cosmids varies depending on the size of the vector itself but usually lies around 40–45 kb. The cloning procedure involves the generation of two vector arms which are then joined to the foreign DNA. Selection against wildtype cosmid DNA is simply done via size exclusion. Cosmids, however, always form colonies and not plaques. Also the clone density is much lower with around 105 - 106 CFU per µg of ligated DNA.

After the construction of recombinant lambda or cosmid libraries the total DNA is transferred into an appropriate E.coli host via a technique called in vitro packaging. The necessary packaging extracts are derived from E.coli cI857 lysogens (red- gam- Sam and Dam (head assembly) and Eam (tail assembly) respectively). These extracts will recognize and package the recombinant molecules in vitro, generating either mature phage particles (lambda-based vectors) or recombinant plasmids contained in phage shells

(cosmids). These differences are reflected in the different infection frequencies seen in favour of lambda-replacement vectors. This compensates for their slightly lower loading capacity. Phage library are also stored and screened easier than cosmid (colonies) libraries.

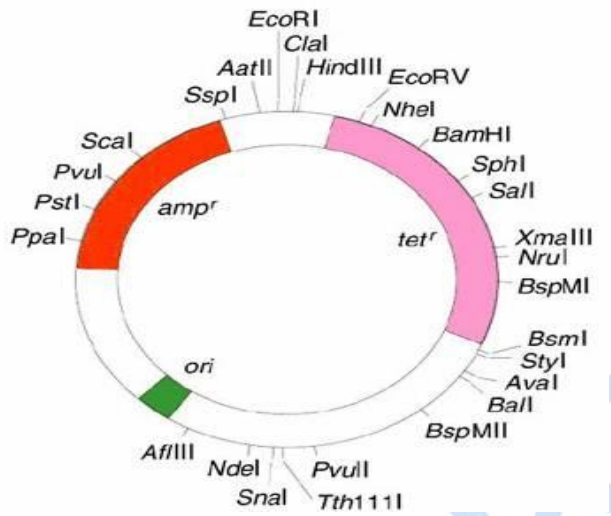


d) pBR322.

pBR322 is a plasmid and was the first widely-used *E. coli* cloning vectors. Created in 1977, it was named after its Mexican creators, p standing for plasmid, and BR for Bolivar and Rodriguez. pBR322 is 4361 base pairs in length and contains a replicon region (source plasmid pMB1), the amp^R gene, encoding the ampicillin resistance protein (souce plasmid RSF2124) and the tet^R gene, encoding the tetracycline resistance protein (source plasmid pSC101). The plasmid has unique restriction sites for more than forty restriction enzymes. 11 of these 40 sites lie within the tet^R gene. There are 2 sites for restriction enzymes HindIII and ClaI within the promoter of the tet^R gene. There are 6 key restriction sites inside the amp^R gene. The origin of replication or *ori* site in this plasmid is pMB1 (a close relative of ColE1). The ori encodes two RNAs (RNAI and RNAII) and one protein (called Rom or Rop).

The circular sequence is numbered such that 0 is the middle of the unique EcoRI site and the count increases through the tet genes. The ampicillin resistance gene is a penicillin beta-lactamase. Promoters P1 and P3 are for the beta-lactamase gene. P3 is the natural promoter, and P1 is artificially created by the ligation of two different DNA fragments to create pBR322. P2 is in the same region as P1, but it is on the opposite strand and initiates transcription in the direction of the tetracycline resistance gene.

The vector pBR322 with exemplified restriction sites.



e) cDNA library.

A **cDNA library** is a combination of cloned cDNA (complementary DNA) fragments inserted into a collection of host cells, which together constitute some portion of the transcriptome of the organism. cDNA is produced from fully transcribed mRNA found in the nucleus and therefore contains only the expressed genes of an organism. Similarly, tissue specific cDNA libraries can be produced. In eukaryotic cells the mature mRNA is already spliced, hence the cDNA produced lacks introns and can be readily expressed in a bacterial cell. While information in cDNA libraries is a powerful and useful tool since gene products are easily identified, the libraries lack information about enhancers, introns, and other regulatory elements found in a genomic DNA library.

cDNA Library Construction

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

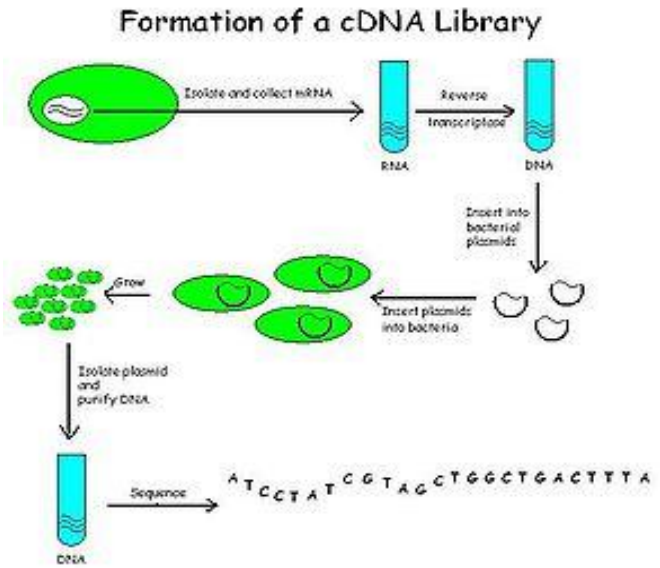


Figure: Formation of a cDNA library.

mRNA extraction

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

cDNA construction

Once mRNA is purified, *oligo-dT* (a short sequence of deoxy-thymine nucleotides) is tagged as a complementary primer which binds to the poly-A tail providing a free 3'-OH end that can be extended by reverse transcriptase to create the complementary DNA strand. Now, the mRNA is removed by using a RNAse enzyme leaving a single stranded cDNA (sscDNA). This sscDNA is converted into a double stranded DNA with the help of DNA polymerase. However, for DNA polymerase to synthesize a complementary strand a free 3'-OH end is needed. This is provided by the sscDNA itself by generating a *hair pin loop* at the 3' end by coiling on itself. The polymerase extends the 3'-OH end and later the loop at 3' end is opened by the scissoring action of S_1 *nuclease*. Restriction endonucleases and DNA ligase are then used to clone the sequences into bacterial plasmids. The cloned bacteria are then selected, commonly through the use of antibiotic selection. Once selected, stocks of the bacteria are created which can later be grown and sequenced to compile the cDNA library.

cDNA Library uses

cDNA libraries are commonly used when reproducing eukaryotic genomes, as the amount of information is reduced to remove the large numbers of non-coding regions from the library. cDNA libraries are used to express eukaryotic genes in prokaryotes.

Prokaryotes do not have introns in their DNA and therefore do not possess any enzymes that can cut it out during transcription process. cDNA do not have introns and therefore can be expressed in prokaryotic cells. cDNA libraries are most useful in reverse genetics where the additional genomic information is of less use. Also, it is useful for subsequently isolating the gene that codes for that mRNA.

Q2. What is recombinant DNA technology? Explain it with its applications.

Ans. **Recombinant DNA** (rDNA) molecules are DNA sequences that result from the use of laboratory methods (molecular cloning) to bring together genetic material from multiple sources, creating sequences that would not otherwise be found in biological organisms. Recombinant DNA is possible because DNA molecules from all organisms share the same chemical structure; they differ only in the sequence of nucleotides within that identical overall structure. Consequently, when DNA from a foreign source is linked to host sequences that can drive DNA replication and then introduced into a host organism, the foreign DNA is replicated along with the host DNA. Recombinant DNA molecules are sometimes called chimeric DNA, because they are usually made of material from two different species, like the mythological chimera.

The DNA sequences used in the construction of recombinant DNA molecules can originate from any species. For example, plant DNA may be joined to bacterial DNA, or human DNA may be joined with fungal DNA. In addition, DNA sequences that do not occur anywhere in nature may be created by the chemical synthesis of DNA, and incorporated into recombinant molecules. Using recombinant DNA technology and synthetic DNA, literally any DNA sequence may be created and introduced into any of a very wide range of living organisms.

Proteins that result from the expression of recombinant DNA within living cells are termed **recombinant proteins**. When recombinant DNA encoding a protein is introduced into a host organism, the recombinant protein will not necessarily be produced. Expression of foreign proteins requires the use of specialized expression vectors and often necessitates significant restructuring of the foreign coding sequence

Creating recombinant DNA

Construction of recombinant DNA, in which a foreign DNA fragment is inserted into a plasmid vector. In this example, the gene indicated by the white color is inactivated upon insertion of the foreign DNA fragment.

Molecular cloning is the laboratory process used to create recombinant DNA. It is one of two widely-used methods (along with polymerase chain reaction, abbr. PCR used to direct the replication of any specific DNA sequence chosen by the experimentalist. The fundamental difference between the two methods is that molecular cloning involves replication of the DNA within a living cell, while PCR replicates DNA in the test tube, free of living cells.

Formation of recombinant DNA requires a cloning vector, a DNA molecule that will replicate within a living cell. Vectors are generally derived from plasmids or viruses, and represent relatively small segments of DNA that contain necessary genetic signals for replication, as well as additional elements for convenience in inserting foreign DNA, identifying cells that contain recombinant DNA, and, where appropriate, expressing the foreign DNA. The choice of vector for molecular cloning depends on the choice of host organism, the size of the DNA to be cloned, and whether and how the foreign DNA is to be expressed. The DNA segments can be combined by using a variety of methods, such as restriction enzyme/ligase cloning or Gibson assembly.

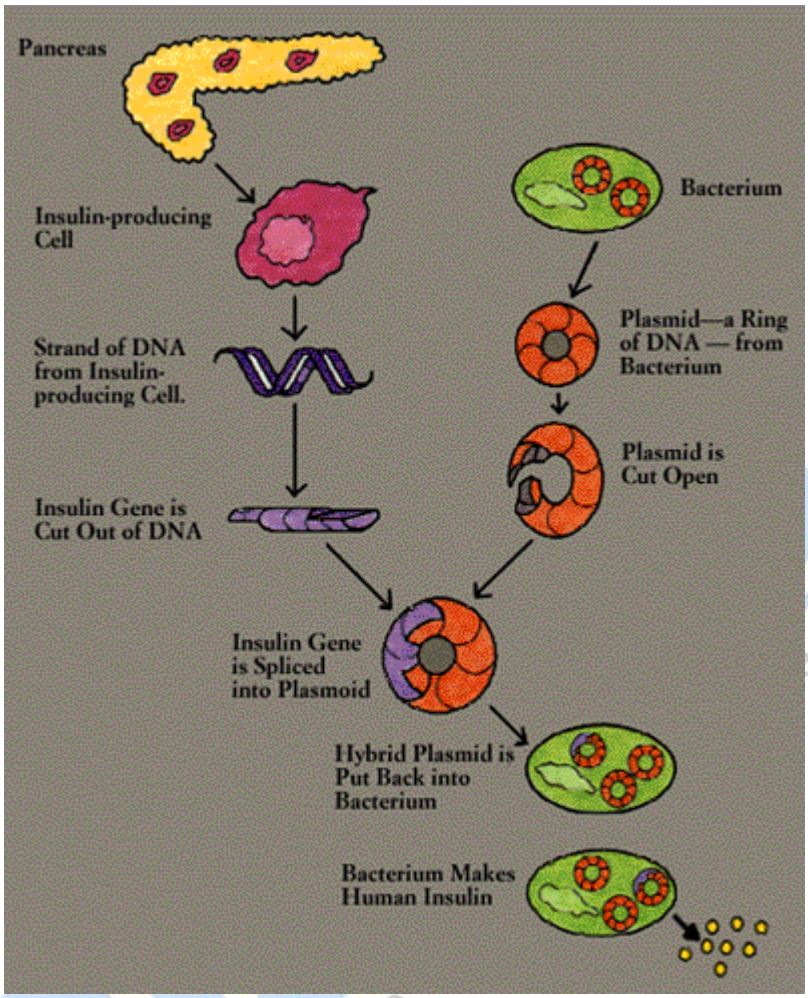
In standard cloning protocols, the cloning of any DNA fragment essentially involves seven steps: (1) Choice of host organism and cloning vector, (2) Preparation of vector DNA, (3) Preparation of DNA to be cloned, (4) Creation of recombinant DNA, (5) Introduction of recombinant DNA into the host organism, (6) Selection of organisms containing recombinant DNA, (7) Screening for clones with desired DNA inserts and biological properties.

Applications of recombinant DNA technology

Recombinant DNA is widely used in biotechnology, medicine and research. Recombinant DNA is used to identify, map and sequence genes, and to determine their function. rDNA probes are employed in analyzing gene expression within individual cells, and throughout the tissues of whole organisms. Recombinant proteins are widely used as reagents in laboratory experiments and to generate antibody probes for examining protein synthesis within cells and organisms.

Many additional practical applications of recombinant DNA are found in human and veterinary medicine, in agriculture, and in bioengineering. Some specific examples are identified below.

Recombinant human insulin. Recombinant insulin has almost completely replaced insulin obtained from animal sources (e.g. pigs and cattle) for the treatment of insulin-dependent diabetes. A variety of different recombinant insulin preparations are in widespread use.



- **Recombinant human growth hormone** (HGH, somatotropin). Growth hormone is administered to patients whose pituitary glands generate insufficient quantities to support normal growth and development. Before recombinant HGH became available, HGH for therapeutic use was obtained from pituitary glands of cadavers.
- **Recombinant blood clotting factor VIII**. Recombinant factor VIII is a blood-clotting protein that is administered to patients with forms of the bleeding disorder hemophilia, who are unable to produce factor VIII in quantities sufficient to support normal blood coagulation.
- **Recombinant hepatitis B vaccine**. Prevention of hepatitis B infection is controlled through the use of a recombinant hepatitis B vaccine, which contains a form of the hepatitis B virus surface antigen that is produced in yeast cells. The development of the recombinant subunit vaccine was an important and necessary development because hepatitis B virus, unlike other common viruses such as polio virus, cannot be grown in vitro. Vaccine information from Hepatitis B Foundation
- **Diagnosis of infection with HIV**. Each of the three widely-used methods for diagnosing HIV infection has been developed using recombinant DNA. The

antibody test (ELISA or western blot) uses a recombinant HIV protein to test for the presence of antibodies that the body has produced in response to an HIV infection. The DNA test looks for the presence of HIV genetic material using reverse transcriptase polymerase chain reaction (RT-PCR).

- **Golden rice** is a recombinant variety of rice that has been engineered to express the enzymes responsible for β-carotene biosynthesis. This variety of rice holds substantial promise for reducing the incidence of vitamin A deficiency in the world's population.
- **Herbicide-resistant crops** Commercial varieties of important agricultural crops (including soy, maize/corn, sorghum, canola, alfalfa and cotton) have been developed which incorporate a recombinant gene that results in resistance to the herbicide glyphosate (trade name *Roundup*), and simplifies weed control by glyphosate application.
- **Insect-resistant crops**. Bacillus thuringeiensis is a bacterium that naturally produces a protein (Bt toxin) with insecticidal properties. The bacterium has been applied to crops as an insect-control strategy for many years, and this practice has been widely adopted in agriculture and gardening.

Q3. What is transgenic plant? How they can be developed?

Ans. A transgenic crop is a genetically modified organism (GMO). Transgenic indicates that a transfer of genes has occurred using recombinant DNA technology. Generally a transgenic crop contains one or more genes that have been inserted artificially either from an unrelated plant or from different species altogether.

Importance of transgenic plants

Plants have been modified for many reasons over the years but the most common purpose is to produce the best products possible. Before recombinant DNA technology, it was possible to produce some improved products, although it was very difficult to do and many things could not be improved or changed. An improved product may involve changing the color or the size of a plant, making it more appealing to the buyer's eye. Or it may be more practical, allowing increased tolerance to cold, frost, or drought; all making a crop easier to grow in a constantly changing environment. Goals like these are unattainable with traditional selection programs and thus transgenics is now preferred when trying to come up with new products to improve sales. One of the hottest areas right now is the modification of crops to increase the resistance of plants to insects and diseases they may carry. Improving resistance to these diseases and insects also reduces the need for herbicides and pesticides. This makes the plants safer for the consumer and allows the farmer to save money on chemicals. In conclusion, transgenic crops are an economically safer method of producing crop products, which makes them appealing and potentially profitable.

Production of transgenic plants

For many years plant breeding entailed the selection of the finest plants to get the best crops. In those days, variation occurred through induced mutation or hybridization

where two or more plants were crossed. Selection occurred through nature, using a "selection of the fittest" concept, where only the seeds best adapted to that environment succeeded. For example, farmers selected only the biggest seeds with non-shattering seed heads, assuming these to be the best. Today, scientists can not only select, but also create crops by inserting genes to make a seeds bare any trait desired. In order to make a transgenic crop, there are five main steps: extracting DNA, cloning a gene of interest, designing the gene for plant infiltration, transformation, and finally plant breeding.

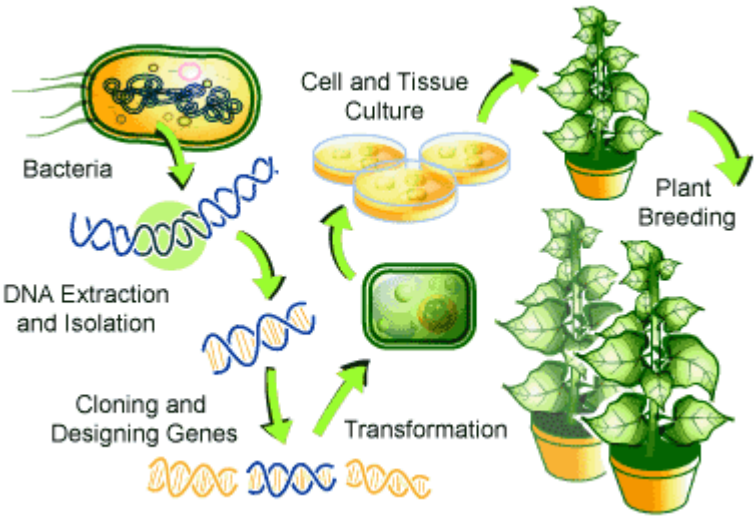


Figure. Overview of how transgenic crops are created.

To understand this process, one must first known a bit about DNA (deoxyribonucleic acids). DNA is the universal programming language of all cells and stores their genetic information. It contains thousands of genes, which are discrete segments of DNA that encode the information necessary to produce and assemble specific proteins. All genes require specific regions in order to be utilized (or expressed) by a cell. These regions include:

1. A promoter region, which signals where a gene begins and it used to express the gene;

2. A termination sequence, which signals the end of a gene;

3. And the coding region, which contains the actual gene to be expressed.

All these regions together allow a gene to create a protein. Once a gene is transcribed into a protein, it can then function as an enzyme to catalyze biochemical reactions or as a structural unit of a cell, both of which will contribute to the appearance of a particular trait in that organism.

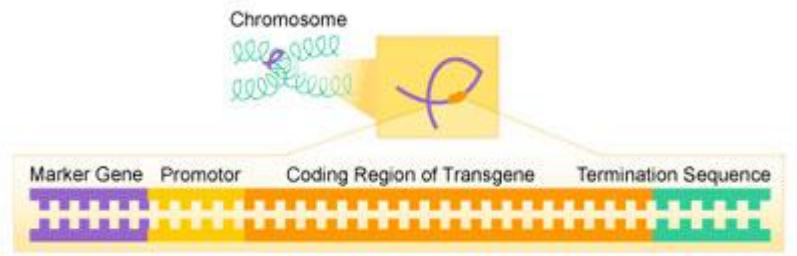


Figure. Gene Regions.

All species are capable of turning DNA into protein through a process known as translation. This capability makes it possible to artificially put genes from one organism into another-a process generally termed transgenics. But just isolating random DNA and inserting it into another organism is not practical. We must first know what particular segments of DNA, and in particular what genes, to insert. Unfortunately, with reference to producing new crops, not much is known about which genes are responsible for increased plant yield, tolerance to different stresses and insects, color, or various other plant characteristics. Much of the research in transgenics is now focused on how to identify and sequence genes contributing to these characteristics.

Genes that are determined to contribute to certain traits then need to be obtained in a significant amount before they can be inserted into another organism. In order to obtain the DNA comprising a gene, DNA is first extracted from cells and put into a bacterial plasmid. A plasmid is a molecular biological tool that allows any segment of DNA in be put into a carrier cell (usually a bacterial cell) and replicated to produce more of it. A bacterial cell (i.e. *E. coli*) that contains a plasmid can put aside and used over and over again to produce copies of the gene the researcher is interested in, a process that is generally referred to as "cloning" the gene. The word "cloning" referring to how many identical copies of the original gene can now be produced at will. Plasmids containing this gene can be used to modify the gene in any way the researcher sees fit, allowing novel effects on the gene trait to be produced.

Once the gene of interest has been amplified, it is time to introduce it into the plant species we are interested in. The nucleus of the plant cell is the target for the new transgenic DNA. There are many methods of doing this but the two most common methods include the "Gene Gun" and *Agrobacterium* method.

The "Gene Gun" method, also known as the micro-projectile bombardment method, is most commonly used in species such as corn and rice. As its name implies, this procedure involves high velocity micro-projectiles to deliver DNA into living cells using a gun. It involves sticking DNA to small micro-projectiles and then firing these into a cell. This technique is clean and safe. It enables scientists to transform organized tissue of plant species and has a universal delivery system common to many tissue types from many different species1. It can give rise to un-wanted side

effects, such as the gene of interest being rearranged upon entry or the target cell sustaining damage upon bombardment. Nevertheless, it has been quite useful for getting transgenes into organisms when no other options are available.

The *Agrobacterium* method involves the use of a soil-dwelling bacteria known as *Agrobacterium tumefaciens*, which has the ability to infect plant cells with a piece of its DNA. The piece of DNA that infects a plant is integrated into a plants chromosome through a tumor-inducing plasmid (Ti plasmid), which can take control of the plant's cellular machinery and use it to make many copies of its own bacterial DNA. The Ti plasmid is a large circular DNA particle that replicates independently of the bacterial chromosome.

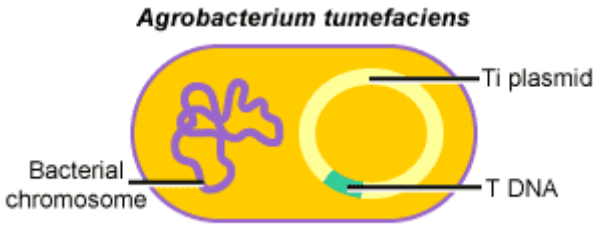


Figure. Transfer DNA on a plasmid in *Agrobacterium*

The importance of this plasmid is that it contains regions of transfer DNA (tDNA), where a researcher can insert a gene, which can be transferred to a plant cell through a process known as a floral dip. A floral dip involves dipping flowering plants into a solution of *Agrobacterium* carrying the gene of interest, followed by the transgenic seeds being collected directly from the plant. This process is useful in that it is a natural method of transfer and therefore thought of as a more acceptable technique. In addition, *Agrobacterium* is capable of transferring large fragments of DNA very efficiently without substantial rearrangements, followed by maintaining high stability of the gene that was transferred. One of the biggest limitations of *Agrobacterium* is that not all important food crops can be infected by this bacterium.

Advantages and Disadvantages of Transgenic Crops

The use of transgenic crops has been an issue for many years. Many concerns have been raised and these generally fall into two categories:

1. A concern about what affect genetically modified material could have on human health. For example, transgenic crops have been suggested to cause allergies in some people, although it is uncertain whether transgenic crops are the source of this reaction. Furthermore the antibiotic resistance genes placed in these crops has been suggested to cause resistance to antibiotics leading to super bugs that cannot be killed with antibiotic treatments. The idea of a population being uncomfortable with ingesting DNA that originated from

another source, such as a virus or bacteria, must also be considered when thinking about producing transgenic crops. However, to date, there is no evidence of the DNA from transgenic crops being any different from the DNA ingested from conventional crops.

2. A concern about whether transgenic crops cause damage to the natural environment. One example includes pollen from transgenic corn, which has been suggested to kill the Monarch butterfly larvae. It has been shown that hybrid corn expresses a bacterial toxin in its pollen, which is then dispersed over 60 meters by wind. In this range, the corn pollen is deposited on other plants near cornfields where it can be ingested by non-target organisms including the monarch butterfly. These butterflies have been found to eat less, have a slower growth rate and higher death rate. A second example is the hybridization of crops with nearby weeds. This could cause these weeds to attain resistance to herbicides or other things that we have been trying to avoid for many years. Genes that provide resistance to viral disease or other traits allowing them to survive in their environment could end up benefiting weed populations around a crop field. This trait could make that population more difficult to control. To date, there has been little evidence to support this theory.

On other side of the coin are the notions that support the use of transgenic crops. The potential benefits of which are quite obvious, including such things as increased yields (to feed a growing population), decreasing the use of pesticides (to save the environment and the cost of pesticides), and the production of novel crops (such as providing crops with increased nutritional value). Being able to retrofit any crop to our desires is a powerful concept, especially with the changing climates of today.

Key Terms

vacuole -- Membrane-bound fluid-filled space within a cell. In most plant cells, there is a single large vacuole filling most of the cell's volume. Some bacterial cells contain gas vacuoles.

Plant breeding

Allele -One of a pair or series of forms of a gene that are alternative in inheritance. Differing alleles can reasult in differing traits.

amino Acids-The basic building block of proteins derived from peptides and polypeptides. Amino acids form the shape and structure of cells. They play a large role both in protein development and metabolism.

Complete Dominance-A type of dominance where the dominant allele completely masks the effect of the recessive allele**.**

Dominant-An allele that determines phenotype even when heterozygous. Also the trait controlled by that allele.

Enzymes-A globular protein with an active site bound to substrate molecules which help to catalyze a reaction by holding molecules in the correct spatial conformation for the reaction to take place.

Gamete- A cell that fuses with another cell during fertilization in organisms that reproduce sexually**.**

Genome-The entirety of an organisms hereditary information encoded in DNA and RNA.

Heredity-Something that is passed on through generations genetically.

Heterozygous-A gene pair having different alleles in the two chromosome sets of the diploid individual; for example, Aa or, A1A2

Homozygous-A diploid gene pair having identical alleles in both copies; for

example, AA or, aa.

Incomplete Dominance-a type of inheritance in which one allele is not completely dominant over the other allele

Phenotype-Any observable characteristic or trait of an organism

sex-linked Inheritance-The inheritance pattern of a phenotype corresponding to the sex chromosomes (usually the X chromosome in XY species).

Variation-The occurrence of differences among individuals as a result of differences in their genetic composition or the environment in which they were raised. Genetic diversity within a species.

II. Nucleic Acids

DNA (deoxyribonucleic acid)

The polymer constructed of successive nucleotides linked by phosphodiester bonds. Some 3 x 10^9 nucleotides are contained in the human haploid genome. During interphase, DNA exists in a nucleoprotein complex containing roughly equal amounts of histones and DNA, which interacts with nuclear matrix proteins. This complex is folded into a basic structure termed a nucleosome containing approximately 150 base pairs. From this highly ordered structure, DNA replication requires a complex process of nicking, unfolding, replication, and splicing. In contrast, gene transcription requires nucleosomal re-organization such that sites critical for the binding of transcriptional machinery reside at internucleosomal junctions.

Branched chain DNA (b-DNA)

A method that exploits the formation of branched DNA to provide a sensitive and specific assay for viral RNA or DNA. The assay is performed in a microtiter format, in which partially homologous oligodeoxynucleotides bind to target to create a branched DNA. Enzyme-labeled probes are then bound to the branched DNA, and light output from a chemiluminescence substrate is directly proportional to the amount of starting target RNA. Standards provide quantitation. The assay displays a 4 log dynamic range of detection, with greater sensitivity to changes in viral load than RT-PCR-based assays. It has been employed to quantitate levels of HIV, HCV,

and HBV.

RNA (ribonucleic acid)

Three varieties of RNA are easily identified in the mammalian cell. Most abundant is ribosomal RNA (rRNA), which occurs in two sizes, 28S (approximately 4600 nucleotides) and 18S (approximately 1800 nucleotides); together they form the basic core of the eukaryotic ribosome. Messenger RNA (mRNA) is the term used to describe the mature form of the primary RNA transcript of the individual gene once it has been processed to eliminate introns and to contain a polyadenylated tail. mRNA links the coding sequence present in the gene to the ribosome, where it is translated into a polypeptide sequence. Transfer RNA (tRNA) is the form of RNA used to shuttle successive amino acids to the growing polypeptide chain. A tRNA molecule contains an anti-codon, a three-nucleotide sequence by which the tRNA molecule recognizes the codon contained in the mRNA template, and an adapter onto which the amino acid is attached.

Codon

Three successive nucleotides on an mRNA that encode a specific amino acid in the polypeptide. Sixty-one codons encode the 20 amino acids, leading to codon redundancy, and three codons signal termination of polypeptide synthesis.

ORF (open reading frame)

The term given to any stretch of a chromosome that could encode a polypeptide sequence, i.e., the region between a methionine codon (ATG) that could serve to initiate protein translation, and the inframe stop codon downstream of it. Several features of the ORF can be used to judge whether it actually encodes an expressed protein, including its length, the presence of a "Kozak" sequence upstream of the ATG (implying a ribosome might actually bind there and initiate protein translation), whether the ORF exists within the coding region of another gene, the presence of exon/intron boundary sequences and their splicing signals, and the presence of upstream sequences that could regulate expression of the putative gene.

Plasmids

Autonomously replicating circular DNA that are passed epigenetically between bacteria or yeast. In order to propagate, plasmids must contain an origin of replication. Naturally occurring plasmids transfer genetic information between

hosts; of these, the genes encoding resistance to a number of antibiotics are the most important clinically. The essential components of plasmids are used by investigators to introduce genes into bacteria and yeast and to generate large amounts of DNA for manipulation.

Phage

A virus of bacteria, phage such as lambda have been used to introduce foreign DNA into bacteria. Because of its infectious nature, the transfection (introduction) efficiency into the bacterial host is usually two orders of magnitude greater for phage over that of plasmids.

Cosmid

By combining the elements of phage and plasmids, vectors can be constructed that carry up to 45 kb of foreign DNA.

cDNA

A complementary copy of a stretch of DNA produced by recombinant DNA technology. Usually, cDNA represents the mRNA of a given gene of interest.

Telomere

A repeating structure found at the end of chromosomes, serving to prevent recombination with free-ended DNA. Telomeres of sufficient length are required to maintain genetic integrity, and they are maintained by telomerase.

CpG

This under-represented (i.e. < 1/16 frequency) dinucleotide pair is a "hotspot" for point mutation. CpG dinucleotides are often methylated on cytosine. Should Me-C undergo spontaneous deamination, uracil arises, which is then repaired by cellular surveillance mechanisms and altered to thymidine. The net result is a C to T mutation.

III. Enzymes of Recombinant DNA Technology

A.Nucleases

A number of common tools of recombinant DNA technology have been developed from the study of the basic enzymology of bacteria and bacteriophage. For example, most unicellular organisms have defense systems to protect against the invasion of foreign DNA. Usually, they specifically methylate their own DNA and then express restriction endonucleases to degrade any DNA not appropriately modified. From such systems come very useful tools. Today, most restriction endonucleases (and most other enzymes of commercial use) are highly purified from either natural or recombinant sources and are highly reliable. Using these tools, the manipulation of DNA and RNA has become routine practice in multiple disciplines of science.

Exonuclease

An enzyme that digests nucleic acids starting from the 5' or 3' terminus and extending inward.

Endonuclease

An enzyme that digests nucleic acids from within the sequence. Usually, specific sequences are recognized at the site where digestion begins.

Isoschizomer

Restriction endonucleases that contain an identical recognition site but are derived from different species of bacteria (and hence have different names).

Restriction endonuclease

These enzymes are among the most useful in recombinant DNA technology, capable of introducing a single cleavage site into a nucleic acid. The site of cleavage is dependent on sequence; recognition sites contain from 4 to 10 specific nucleotides. The resultant digested ends of the nucleic acid chain may either be blunt or contain a 5' or 3' overhang ranging from 1 to 8 nucleotides.

Ribonuclease

These enzymes degrade RNA and exist as either exonucleases or endonucleases. The three most commonly used ribonucleases are termed RNase A, RNase T1, and RNAse H (which degrades duplex RNA or the RNA portion of DNA•RNA hybrids).

Ribozymes

are based on a catalytic RNA characterized by a hammerhead-like secondary structure, and by introducing specific sequences into its RNA recognition domain, destruction of specific mRNA species can be accomplished. Ribozymes thus represent a tool to eliminate expression of specific genes, and are being tested in several hematological disease states, including neoplasia. A highly specific RNA sequence can generate secondary structure by virtue of intrachain base pairing. "Hairpin loops" and "hammer head" structures serve as examples of such phenomena. When the proper secondary structure forms, such RNA molecules can bind a second RNA molecule (e.g. an mRNA) at a specific location (dependent on an approximately 20-nucleotide recognition sequence) and cleave at a specific GUX triplet (where X = C, A, or U). These molecules will likely find widespread use as tools for specific gene regulation or as antiviral agents but are evolutionarily related to RNA splicing, which in its simplest form is autocatalytic.

B.Polymerases

DNA polymerase

The enzyme that synthesizes DNA from a DNA template. The intact enzyme purified from bacteria (termed the holoenzyme) has both synthetic and editing functions. The editing function results from nuclease activity.

Klenow fragment

A modified version of bacterial DNA polymerase that has been modified so that only the polymerase function remains; the 5'→3' exonuclease activity has been eliminated.

Thermostabile polymerases

The prototype polymerase, Taq, and newer versions such as Vent and Tth polymerase are derived from microorganisms that normally reside at high temperature. Consequently, their DNA polymerase enzymes are quite stable to heat denaturation, making them ideal enzymes for use in the polymerase chain reaction.

RNA polymerase II This enzyme is used by mammalian cells to transcribe structural genes that result in mRNA. The enzyme interacts with a number of other proteins to correctly initiate transcription, including a number of general factors,

and tissue-specific and induction-specific enhancing proteins.

RNA polymerase III This enzyme is used by the cell to transcribe ribosomal RNA genes.

Kinases These enzymes transfer the γ-phosphate group from ATP to the 5' hydroxyl group of a nucleic acid chain.

Viral-derived kinases These enzymes are utilized in recombinant DNA technology to transfer phosphate groups (either unlabeled or ^{32}P-labeled) to oligonucleotides or DNA fragments. The most commonly used kinase is T4 polynucleotide kinase.

Mammalian protein kinases

These enzymes transfer phosphate groups from ATP to either tyrosine, threonine, or serine residues of proteins. These enzymes are among the most important signaling molecules present in mammalian cell biology.

Farnesyl protein transferase (FTPase)

FTPase adds 15 carbon farnesyl groups to CAAX motifs, such as one present in ras, allowing their insertion into cellular membranes.

Terminal deoxynucleotidyl

This lymphocyte-specific enzyme normally transfers available (random) nucleotides to the 3' end of a growing nucleic acid chain. In recombinant DNA technology, these enzymes can be used to add a homogeneous tail to a piece of DNA, thereby allowing its specific recognition in PCR reactions or in cloning efforts.

Ligases

These enzymes utilize the γ-phosphate group of ATP for energy to form a phosphodiester linkage between two pieces of DNA. The nucleotide contributing the 5' hydroxyl group to the linkage must contain a phosphate, which is then linked to the 3' hydroxyl group of the growing chain.

DNA methylases

These enzymes are normally part of a bacterial host defense against invasion by foreign DNA. The enzyme normally methylates endogenous (host) DNA and

thereby renders it resistant to a series of endogenous restriction endonucleases. In recombinant DNA work, methylation finds use in cDNA cloning to prevent subsequent digestion by the analogous restriction endonuclease.

Reverse transcriptase

This enzyme, first purified from retrovirus-infected cells, produces a cDNA copy from an mRNA molecule if first provided with an antisense primer (oligo dT or a random primer). This enzyme is critical for converting mRNA into cDNA for purposes of cloning, PCR amplification, or the production of specific probes.

Topoisomerase

A homodimeric chromosomal unwinding enzyme that introduces a double-stranded nick in DNA, which allows the unwinding necessary to permit DNA replication, followed by religation. Inhibition of topoisomerases leads to blockade of cell division, the target of several chemotherapeutic agents (e.g., etoposide).

Telomerase

A specialized DNA polymerase that protects the length of the terminal segment of a chromosome. Should the telomere become sufficiently shortened (by repeated rounds of cell division), the cell undergoes apoptosis. The holoenzyme contains both a polymerase and an RNA template; only the latter has been characterized, although the gene for the enzymatic activity has recently been cloned.

IV. Molecular Methods

A number of molecular techniques have found widespread application in the biomedical sciences. This section of the glossary provides general concepts and is not intended to convey adequate details. The interested reader is referred to the excellent handbook of J. Sambrook and coworkers (Molecular Cloning, A Laboratory Manual, 2nd Ed., CSH Laboratory Press, 1989).

Maxam-Gilbert sequencing

A method to determine the sequence of a stretch of DNA based on its differential cleavage pattern in the presence of different chemical exposures. A nucleic acid chain can be cleaved following G, A, C, or C and T by exposure of ^{32}P-labeled DNA

to neutral dimethylsulfate, dimethylsulfate-acid, hydrazine-NaCl-piperidine or hydrazine-piperidine alone, respectively.

Dideoxynucleotide (ddN) chain termination sequencing

Also termed "Sanger sequencing," this method relies on the random incorporation of dideoxynucleotides into a growing enzyme-catalyzed DNA chain. As no 3' hydroxyl group is present on the ddN, chain synthesis halts following its incorporation into the chain. If ^{32}P or ^{35}S nucleotides are also incorporated into the reaction, a family of DNA fragments will be generated that can be visualized on a polyacrylamide gel. This method is presently the most commonly used chemistry to determine the sequence of DNA.

DNAse footprinting

This technique depends on the ability of protein specifically bound to DNA to block the activity of the endonuclease DNAse I. ^{32}P-labeled DNA is mixed with nuclear proteins, which potentially contain specific DNA-binding proteins, and the reaction is then subjected to limited DNAse digestion. If a given site of DNA is free of protein, it will be cleaved by the DNAse. In contrast, regions of DNAse specifically bound by proteins (transcription factors or enhancers) will be protected from digestion. The resultant mixture of DNA fragments from control and protein-containing reactions are then separated on a polyacrylamide gel. As the site of ^{32}P labeling of the original DNA fragment is known, sites that were protected from DNAse digestion will be represented on the gel as a region devoid of that length fragment. Therefore, in comparison to naked DNA, regions that bind specific proteins will be represented as a "footprint."

DNAse hypersensitivity site mapping

This technique is designed to uncover regions of DNA that are in an "active" transcriptional state. It depends on the hypersensitivity of such sites (because of the lack of the highly compact nucleosome structure) to limited digestion with DNAse. Intact nuclei are subjected to limited DNAse digestion. The resultant large DNA fragments are then extracted, electrophoretically separated, and hybridized with a ^{32}P-labeled probe from a known site within the gene of interest. If, for example, the probe were located at the site of transcription initiation, and should DNA fragments of 2 kb and 5 kb be detected with this probe, hypersensitive sites would thereby be

mapped to 2 kb and 5 kb upstream of the start of transcription initiation. By extrapolation, these sites would then be assumed important in the transcriptional regulation of the gene of interest, especially if such a footprint were only detected using cells that express that gene.

Mobility shift (or band shift) assays

Like DNAse footprinting, this technique is also utilized to determine whether a fragment of DNA binds specific proteins. ^{32}P-labeled DNA (either duplex oligonucleotides or small restriction fragments) are incubated with nuclear protein extracts and subjected to native acrylamide gel electrophoresis. Should specific DNA-binding proteins that recognize the oligonucleotide or restriction fragment probe be present in the nuclear extracts, a DNA-protein complex will be formed and its migration through the native gel will be retarded compared to the unbound DNA. Hence, the labeled band will be shifted to a more slowly migrating position. The specificity of their reaction can be demonstrated by also incubating, in separate reactions, competitor DNA that contains the presumed binding site or irrelevant DNA sequence.

S_1 nuclease analysis

This technique is used to identify the start of RNA transcription. The DNAse enzyme S_1 cleaves only at sites of single-stranded DNA. Therefore, if ^{32}P-labeled DNA is hybridized with mRNA, the resulting heteroduplex can be digested with S_1, and the resulting DNA fragment will be of length equivalent to the site at which the piece of DNA begins through the mature 5' end of the RNA.

RNAse protection assay

This assay is in many ways similar to the S_1 nuclease analysis. In this case, a ^{35}S- or ^{32}P-labeled antisense RNA probe is synthesized and hybridized with mRNA of interest. The duplex RNA is then subjected to digestion with RNAse A and T_1, both of which will cleave only single-stranded RNA. Following digestion, the remaining labeled RNA is size-fractionated, and the size of the protected RNA probe then gives an indication of the size of the mRNA present in the original sample. This assay can also be used to quantitate the amount of specific RNA in the original sample.

PCR (polymerase chain reaction)

This technique finds use in several arenas of recombinant DNA technology. It is based on the ability of sense and antisense DNA primers to hybridize to a cDNA of interest. Following extension from the primers on the cDNA template by DNA polymerase, the reaction is heat-denatured and allowed to anneal with the primers once again. Another round of extension leads to a multiplicative increase in DNA products. Therefore, a minute amount of cDNA can be efficiently amplified in an exponential fashion to result in easily manipulable amounts of cDNA. By including critical controls, the technique can be made quantitative. Important clinical examples of the use of PCR or reverse transcription PCR (see below) include (1) detection of diagnostic chromosomal rearrangements [e.g., bcr/abl in CML, t(15;17) in AML-M3, t(8;21) in AML-M2, or bcl-2 in follicular small cleaved cell lymphoma], or (2) detection of minimal residual disease following treatment. The level of sensitivity is one in 10^4 to 10^5 cells.

RT-PCR (reverse transcription PCR)

This technique allows the rapid amplification of cDNA starting with RNA. The first step of the reaction is to reverse-transcribe the RNA into a first strand cDNA copy using the enzyme reverse transcriptase. The primer for the reverse transcription can either be oligo dT, to hybridize to the polyadenylation tail, or the antisense primer that will be used in the subsequent PCR reaction. Following this first step, standard PCR is then performed to rapidly amplify large amounts of cDNA from the reverse transcribed RNA.

Nested PCR

By using an independent set of PCR primers located within the sequence amplified by the primary set, the specificity of a PCR reaction can be greatly enhanced. should the first PCR reaction yield a product of 600 nucleotides, a second PCR reaction using the first product as template and a different set of primers will produce a smaller, "nested" PCR product, the presence of which acts to confirm the identity of the primary product.

Real-time automated PCR

During PCR, a fluorogenic probe, consisting of an oligodeoxynucleotide with both reporter and quencher dyes attached, anneals between the two standard PCR primers. When the probe is cleaved during the next PCR cycle, the reporter is separated from the quencher so that the fluorescence at the end of PCR is a direct measure of the amplicons generated throughout the reaction. Such a system is

amenable to automation and gives precise quantitative information.

Allele-specific PCR

By using generic PCR primers flanking the immunoglobulin or T cell receptor genes, the precise rearranged gene characteristic of a B or T cell neoplasm can be amplified and sequenced. Once so obtained, new PCR primers can then be designed that are unique to the patient's tumor. Such allele-specific PCR can then be used to detect blood cell contamination by tumor and to detect minimal residual disease following therapy.

Southern blotting

This technique is used to detect specific sequences within mixtures of DNA. DNA is size-fractionated by gel electrophoresis and then transferred by capillary action to nitrocellulose or another suitable synthetic membrane. Following blocking of nonspecific binding sites, the nitrocellulose replica of the original gel electrophoresis experiment is then allowed to hybridize with a cDNA or oligonucleotide probe representing the specific DNA sequence of interest. Should specific DNA be present on the blot, it will combine with the labeled probe and be detectable by autoradiography. By co-electrophoresing DNA fragments of known molecular weight, the size(s) of the hybridizing band(s) can then be determined. For gene rearrangement studies, Southern blotting is capable of detecting clonal populations that represent approximately 1% of the total cellular sample.

Northern blotting

This modification of a Southern blot is used to detect specific RNA. The sample to be size-fractionated in this case is RNA and, with the exception of denaturation conditions (alkali treatment of the Southern blot versus formamide/formaldehyde treatment of the RNA sample for Northern blot), the techniques are essentially identical. The probe for Northern blotting must be antisense.

Western blotting

This technique is designed to detect specific protein present in a heterogenous sample. Proteins are denatured and size-fractionated by polyacrylamide gel electrophoresis, transferred to nitrocellulose or other synthetic membranes, and then probed with an antibody to the protein of interest. The immune complexes present on the blot are then detected using a labeled second antibody (for example,

a ^{125}I-labeled or biotinylated goat anti-rabbit IgG). As the original gel electrophoresis was done under denaturing and reducing conditions, the precise size of the target protein can be determined.

Southwestern blotting

This technique is designed to detect specific DNA-binding proteins. Like the Western blot, proteins are size-fractionated and transferred to nitrocellulose. The probe in this case, however, is a double-stranded labeled DNA that contains a putative protein-binding site. Should the DNA probe hybridize to a specific protein on the blot, that protein can be subsequently identified by autoradiography. This technique often suffers from nonspecificity, so that a number of critical controls must be included in the experiment for the results to be considered rigorous.

In situ hybridization

This technique is designed to detect specific RNA present in histological samples. Tissue is prepared with particular care not to degrade RNA. The cells are fixed on a microscope slide, allowed to hybridize to probe, and then washed and overlaid with photographic emulsion. Following exposure for one to four weeks, the emulsion is developed and silver grains overlying cells that contain specific RNA are detected. The most useful probes for this purpose are metabolically ^{35}S-labeled riboprobes generated by in vitro transcription of a cDNA using viral RNA polymerase. These probes give the lowest background and are preferable to using terminal deoxynucleotidyl transferase or alternative methods using ^{32}P as an isotope.

FISH (fluorescence in situ hybridization)

A general method to assign chromosomal location, gene copy number (both increased and decreased), or chromosomal rearrangements. Biotin-containing nucleotides are incorporated into specific cDNA probes by nick-translation. Alternatively, digoxigenin or fluorescent dyes can be incorporated by enzymatic or chemical methods. The probes are then hybridized with solubilized, fixed metaphase cells, and the copy number of specific chromosomes or genes are determined by counter-staining with fluorescein isothiocyanate (FITC)-labeled avidin or other detector reagents. The number and location of detected fluorescent spots correlates with gene copy number and chromosomal location. The method

also allows chromosomal analysis in interphase cells, allowing extension to conditions of low cell proliferation.

CGH (comparative genome hybridization)

In CGH, DNA is extracted from tumor and from normal tissues and differentially labeled with fluorescent dyes. Once the DNA samples are mixed and hybridized to normal metaphase chromosome spreads, chromosomal regions that are under-represented or over-represented in the tumor sample can be identified. This method can be applied to extremely small tumor samples (by using PCR methods) of formalin-fixed or frozen tissue. It has been applied to detect loss of chromosome 18q or 17p in colon cancer and is likely to be applied to hematologic malignancies. The sensitivity of the technique approaches 1 cell in 100.

Nick-translation

This technique is used to label cDNA to high specific activity for the purpose of probing Southern and Northern blots and screening cDNA libraries. The cDNA fragment is first nicked with a limiting concentration of DNAse, then DNA polymerase is used to both digest and fill in the resulting gaps with labeled nucleotides.

Random priming

This technique is also used to produce labeled cDNA probes and is dependent on using random 6- to 10-base oligonucleotides to sit down on a single-stranded cDNA and then using DNA polymerase to synthesize the complementary strand using labeled nucleotides. This technique usually produces more favorable results than nick-translation.

Riboprobes

These labeled RNA molecules are produced by first cloning the cDNA of interest into a plasmid vector that contains promoters for viral RNA polymerases. Following cloning, the viral RNA polymerase is added, and labeled nucleotides are incorporated into the resulting RNA transcript. This molecule is then purified and

used in probing reactions. Many such cloning vectors (for example, pGEM) have different RNA polymerase promoters on either side of the cloning site, allowing the generation of both sense and antisense probes from the same construct.

Mutagenesis, site-specific

Several methods are now available to intentionally introduce specific mutations into a cDNA sequence of interest. Most are based on designing an oligonucleotide that contains the desired mutation in the context of normal sequence. This oligonucleotide is then incorporated into the cDNA using DNA polymerase, either using a single-stranded DNA template (phage M13) or in a PCR format to produce a heteroduplex DNA containing both wild type and mutant sequences. Using M13, recombinant phage are then produced and mutant cDNA are screened for on the basis of the difference in wild type and mutant sequences; using the PCR format, the exponential amplification of the mutant sequence results in its overwhelming numerical advantage over wild type sequence, resulting in nearly all clones containing mutant sequence. Both of these methods require that the entire cDNA insert synthesized in vitro be sequenced in its entirety to guarantee the fidelity of mutagenesis and synthesis of the remaining wild type sequences.

Chromatography, gel filtration

This technique is designed to separate proteins based on their molecular weight. It is dependent on the exclusion of proteins from a matrix of specific size. Proteins that are too large to fit into the matrix of the gel bed run to the bottom of the column more quickly than smaller proteins, which are included in the volume of the matrix. Therefore, using appropriate size markers, the approximate molecular weight of a given protein can be determined and it can be separated from proteins of dissimilar size. Typical separation media for gel filtration chromatography include Sephadex and Ultragel.

Chromatography, ion exchange

This separation methodology depends on the preferential binding of positively charged proteins to a matrix containing negatively charged groups or a negatively charged protein binding to a matrix containing positively charged groups. Increases in the buffer concentration of sodium chloride are then used to break the ionic interaction between protein and matrix and elute off-bound proteins. Examples of

such separation media include DEAE and CM cellulose.

Chromatography, hydrophobic

This methodology separates proteins based on their hydrophobicity. Proteins preferentially bind to the matrix based on the strength of this interaction; proteins are then eluted off using solvents of increasing hydrophobicity. Separation media include phenyl-sepharose and octyl-sepharose.

Chromatography, affinity

This separation method depends on using any molecule that can preferentially bind to a protein of interest. Typical methodologies include using lectins (such as wheat germ or concanavalin A) to bind glycoproteins or using covalently coupled monoclonal antibodies to bind specific protein ligands.

Chromatography, high performance liquid (HPLC).

A general methodology to improve the separation of complex protein mixtures. The types of HPLC columns available are the same as for conventional chromatography, such as those based on size exclusion, hydrophobicity, and ionic interaction, but the improved flow rates resulting from the high pressure system provide enhanced separation capacity and improved speed.

Proteomics.

The general term used in the study of the display of all proteins present in cells under defined conditions. By deciphering which proteins are differentially displayed in tumor cells compared to their normal counterparts, or in cells stimulated to grow, vs. their quiescent state, one can determine the proteins that are responsible for the cellular phenotype. In essence, proteomics is to proteins what genomics is to genes.

DNA microarrays (gene expression arrays or gene chips)

Multiple (presently up to tens of thousands) gene fragments or oligonucleotides representing distinct genes spotted onto a solid support. Theoretically, microarrays could be used to determine the totality of the genome expressed in a given cell under specific growth conditions, if the entire genome were present on the microarray. At present, gene chips are available that represent about 1/3 of the

human genome. The microarray is hybridized with a labeled probe (either radioactive or fluoresceinated) representing all the mRNA species in a given cell grown under a certain condition. By comparing the hybridization patterns produced by probes produced from cells under two different growth conditions, one can determine which genes are increased and which are decreased in response to the growth stimulus. In a similar way, comparison of the expression profiles of a malignant cell type and its normal counterpart, potentially allows one to determine the genes responsible for transformation.

Yeast 2-hybrid screens

A strategy designed to determine the binding partners for a protein of interest. The gene (or a fragment of the gene) representing a protein of interest (the "bait") is fused in frame to DNA binding domain (DBD) of yeast transcription factor and then introduced into a yeast strain. A cDNA library is then constructed from the cells in which the bait is normally expressed, and fused in frame to the activation domain (AD) of the same yeast transcription factor. When the library is introduced into the yeast expressing the bait/DBD fusion, any yeast cell expressing a cDNA encoding a binding partner of the bait protein will have that cDNA/AD fusion protein bind to the bait/DBD fusion, bringing the AD and DBD together, thereby creating a fully functional transcription factor that now drives a reporter gene, allowing the yeast carrying such interacting proteins to be identified and the cDNA recovered.

V. Physiologic Gene Regulation

The regulation of gene expression is central to physiology. Complex organisms have evolved multiple mechanisms to accomplish this task. The first step in protein expression is the transcription of a specified gene. The rate of initiation and elongation of this process is the most commonly used mechanism for regulating gene expression. Once formed, the primary transcript must be spliced, polyadenylated, and transported to the cytoplasm. These mechanisms are also possible points of regulation. In the cytoplasm, mRNA can be rapidly degraded or retained, another potential site of control. Protein translation next occurs on the ribosome, which can be free or membrane-associated. Secreted proteins take the latter course, and the trafficking of the protein through these membranes and ultimately to storage or release makes up another important point of potential regulation. Individual gene expression is often controlled at multiple levels, making investigation and intervention a complex task.

Transcription

Transcription is the act of generating a primary RNA molecule from the double-stranded DNA gene. Regulation of gene expression is predominantly at the level of regulating the initiation and elongation of transcription. The enzyme RNA polymerase is the key feature of the system, which acts to generate the RNA copy of the gene in combination with a number of important proteins. There is usually a fixed start to transcription and a fixed ending.

TATA

Many genes have a sequence that includes this tetranucleotide close to the beginning of gene transcription. RNA polymerase binds to the sequence and begins transcription at the cap site, usually located approximately 30 nucleotides downstream.

Enhancer

An enhancer is a segment of DNA that lies either upstream, within, or downstream of a structural gene that serves to increase transcription initiation from that gene. A classical enhancer element can operate in either orientation and can operate up to 50 kb or more from the gene of interest. Enhancers are *cis*-acting in that they must lie on the same chromatin strand as the structural gene undergoing transcription. These *cis*-acting sequences function by binding specific proteins, which then interact with the RNA polymerase complex.

Silencer

These elements are very similar to enhancers except that they have the function of binding proteins and inhibiting transcription.

Initiation complex

This multi-protein complex forms at the site of transcription initiation and is composed of RNA polymerase, a series of ubiquitous transcription factors (TF II family), and specific enhancers and/or silencers. The proteins are brought together by the looping of DNA strands so that protein binding sites, which may range up to tens of kb apart, can be brought into close juxtaposition. Specific protein•protein interactions then allow assembly of the complex.

Polyadenylation

Following transcription of a gene, a specific signal near the 3' end of the primary transcript (AATAAA) signals that a polyadenine tail be added to the newly formed transcript. The tail may be up to several hundred nucleotides long. The precise function of the poly A tail is uncertain but it seems to play a role in stability of the mRNA and perhaps in its metabolism through the nuclear membrane to the ribosome.

Splicing

The primary RNA transcript contains a number of sequences that are not part of the mature mRNA. These regions are called introns and are removed from the primary RNA transcript by a process termed splicing. A complex tertiary structure termed a lariat is formed and the intron sequence is eliminated bringing the coding sequences (exons) together. Specific sequences within the primary transcript dictate the sites of intron removal.

Exons

These are the regions of the primary RNA transcript that, following splicing, form the mature mRNA species, which encodes polypeptide sequence.

Introns

These are the regions of the primary RNA transcript that are eliminated during splicing. Their precise function is uncertain. However, several transcriptional regulatory regions have been mapped to introns, and they are postulated to play an important role in the generation of genetic diversity (exon shuffling mechanism).

Nucleosomes

When linear, the length of a specific chromosome is many orders of magnitude greater than the diameter of the nucleus. Therefore, a mechanism must exist for folding DNA into a compact form in the interphase nucleus. Nucleosomes are complex DNA protein polymers in which the protein acts as a scaffold around which DNA is folded. The mature chromosomal structure then appears as beads on a string; within each bead (nucleosome) are folded DNA and protein. Nucleosome structure is quite fluid, and internucleosomal stretches of DNA are thought to be sites that are important for active gene transcription.

***Trans*-acting factors**

Proteins that are involved in the transcriptional regulation of a gene of interest.

***Cis*-acting factors**

These are regions at a gene either upstream, within, or downstream of the coding sequence that contains sites to which transcriptionally important proteins may bind. Sequences that contain 5 to 25 nucleotides are present in a typical *cis*-acting element.

Transcription factors

Specific proteins that bind to control elements of genes. Several families of transcription factors have been identified and include helix-loop-helix proteins, helix-turn-helix proteins, and leucine zipper proteins. Each protein includes several distinct domains such as activation and DNA-binding regions.

LCR (locus control region)

Cis-acting sites are occasionally organized into a region removed from the structural gene(s) they control. Such locus control regions (LCRs) are best described for the β globin and α globin loci. First recognized by virtue of clustering of multiple DNAse hypersensitive sites, the β globin LCR is required for high level expression from all of the genes and appears to be critical for their stage-specific developmental pattern of expression.

Protein translation

This term is applied to the assembly of a polypeptide sequence from mRNA.

KOZAK sequence

This five-nucleotide sequence resides just prior to the initiation codon and is thought to represent a ribosomal-binding site. The most consistent position is located three nucleotides upstream from the initiation ATG and is almost always an adenine nucleotide. When multiple potential initiation codons are present in an open reading frame, the ATG codon, which contains a strong consensus KOZAK sequence, is likely the true initiation codon.

Initiation codon

The ATG triplet is used to begin polypeptide synthesis. This is usually the first ATG codon, located approximately 30 nucleotides downstream of the site of transcription initiation (cap site). However, the context in which the ATG resides is also important (see KOZAK sequence).

Missense mutation

Mutation of the mRNA sequence to generate an altered codon, which results in an amino acid change, is termed a missense mutation.

Nonsense mutation

This type of mutation results in the generation of a premature termination codon and hence creates a truncated polypeptide.

Transcriptional regulation

Gene regulation is determined by the rate of transcriptional initiation. This usually results from alteration in the level of activity of trans-acting proteins, which, in turn, are regulated either by the amount of the transcriptionally active protein or by their state of activation.

Leucine zipper proteins

A family of DNA-binding proteins that require a dimeric state for activity and that dimerize by virtue of an alpha helical region that contains leucine at every seventh position. Because 3.4 amino acids reside in each turn of an alpha helix, the occurrence of leucine at every seventh position results in a strip of highly hydrophobic residues on one surface of the alpha helix. Such a domain on one polypeptide can intercollate with a similar domain on a second polypeptide, resulting in the formation of a stable homodimer or heterodimer. Examples of the leucine zipper family include the proto-oncogenes *c-jun* and *c-fos*.

Basic helix-loop-helix proteins

These transcriptional proteins are characterized by two alpha helical regions separated by a loop structure; this domain is involved in protein dimerization. Examples of this family of transcription factors include E12/E47 of the immunoglobulin promoter or Myo D of muscle cell regulation.

Helix-turn-helix

This family of transcriptionally active proteins depends on the helix-turn-helix motif for dimerization. Examples include the homeodomain genes such as the Hox family.

Master switch genes

These polypeptide products are thought to regulate a whole family of genes and result in a cell undergoing a new program of differentiation. An example of such a system is Myo D, in which activation is thought to lead to differentiation along the muscle cell lineage.

Zinc finger domain proteins

The presence of conserved histidine and cysteine residues allows chelation of a zinc atom and results in the formation of a loop structure called the zinc finger domain. This feature is present in a large family of transcriptionally active proteins such as the steroid hormone receptors.

Post-transcriptional regulation

Mechanisms of gene regulation that do not involve transcriptional enhancement or silencing and include altering the rate of mRNA degradation, the efficiency of translation or post-translational modification, or transportation of the polypeptide out of the cell.

Actinomycin D pulse experiments

The application of actinomycin D to actively metabolizing cells results in the cessation of new RNA transcription. Consequently, serial determinations of specific RNA levels will allow one to calculate the mRNA half life. Should this vary between control and stimulated conditions, evidence is garnered that a gene of interest is regulated at the level of mRNA stability.

Reporter genes

In order to determine how a gene promoter or enhancer works in vitro, that genetic

element is often linked to a gene for which a simple assay is readily available and whose regulation is not affected by post-transcriptional processes. Such reporter genes include chloramphenicol acetyl transferase, β galactosidase, and firefly luciferase. The first is the most commonly used reporter; however, more recent studies have emphasized the use of the latter two reporters, as these are more sensitive to minimal changes in promoter or enhancer activity.

CAT (chloramphenicol acetyl transferase)

The bacterial gene for chloramphenicol resistance, chloramphenicol acetyl transferase (CAT) is commonly used as a reporter gene for investigating physiologic gene regulation. The assay depends on the ability of transfected cellular cytoplasm to convert ^{14}C chloramphenicol to its acetylated form in the presence of acetyl CoA. The acetylated forms are separated from the ^{14}C substrate using thin layer chromatography.

β galactosidase

The presence of β galactosidase activity in the cytoplasm of transfected cells can be readily detected by its ability to convert a colorless substrate to a blue-colored product. This is usually assayed using a fluorimeter.

Luciferase

This gene, which is the most recent reporter gene to be used, has gained increasing acceptance because of its ease of assay and extreme sensitivity. The assay is based on the ability of the protein to undergo chemiluminescence and transmit light, detected with a luminometer.

Cyclins

A group of proteins that vary in expression throughout the cell cycle. Once a threshold level is attained, interaction with specific cellular kinases results in phosphorylation of critical components of the mitotic machinery. Several classes of cyclins (A through E) exist that regulate different aspects of the cell cycle (G_0, G_1, S, G_2, M). Altered expression of some cyclins is associated with hematologic malignancy, e.g., t(11;14) in mantle cell lymphoma leads to over-expression of cyclin

D_1, a G_1 phase cyclin.

Cdk (cyclin-dependent kinase)

A related group of cellular kinases, present in virtually all cells, that are regulated both positively and negatively by specific phosphorylation events and negatively by association with other proteins, and are dependent on cyclins, present only during certain phases of the cell cycle (cdk1-activated during G_2/M phase, cdk2-G_1/S phases, cdk4-G_1/S phases, cdk6-G_1 phase, cdk7-throughout the cell cycle).

CdkI (cdk inhibitors)

Proteins that inhibit the cdks by stoicheometric combination, arresting cells in G_1 phase, and include p27, p21 and the p16 Ink 4A family of proteins. The latter are implicated as tumor suppressor genes, as their deficiency in mice leads to rapid cellular proliferation and a high rate of spontaneous tumor development. Moreover, deficiency of p16 family members has been associated with numerous types of human tumors, including a fraction of cases of B cell ALL and T cell leukemia.

Proteosome

A large multiprotein complex designed to digest proteins that have been targeted for destruction, usually based on the presence of multiple sites of ubiquination. The proteosome is critical for many cellular processes, including cell growth, where it eliminates a series of brakes on cell cycle progression, or eliminates growth factor receptors following their internalization following ligand binding.

Ubiquitin

A small molecular weight protein that can be coupled to lysine residues of proteins targeted for destruction. There are ubiquitin ligases and deubiquinases, which add or remove ubiquitin from proteins, usually in a fairly specific manner. Once polyubiquinated, proteins are subject to destruction by the proteosome.

Immunoglobulin somatic hypermutation

Immunoglobulin variable region gene sequences are further diversified in mature B cells during clonal expansion that occurs following antigen stimulation. Mutations clustered within V regions typically involve nucleotide substitution, and less frequently small deletions or insertions. This event usually follows immunoglobulin

class switching, which by itself does little to alter Ig specificity; only the effector functions of the molecule are altered by the change from IgM to IgG, etc.

VI. Expression of Recombinant Proteins

In order to exploit the techniques of recombinant DNA research, one must possess a system to manufacture the protein of interest. After identifying the gene encoding the protein and obtaining a cDNA representation of it ("cloning"), the cDNA must be placed in a vector capable of driving high levels of RNA transcription in a host system capable of translating and appropriately modifying the polypeptide to produce fully functional protein. And just like obtaining a protein of interest from natural sources, one must purify protein from the final expression system. Because of the nature of the highly engineered systems and high levels of expression, this latter task is usually considerably easier using recombinant methods than from natural sources. The methods used to generate expression vectors are described in Section IV, but the methods to purify proteins are discussed in only a rudimentary way and are beyond the scope of this glossary.

Expression vector

A plasmid that contains all of the elements necessary to express an inserted cDNA in the host of interest. For a mammalian cell host, such a vector typically contains a powerful promoter coupled to an enhancer, a cloning site, and a polyadenylation signal. In addition, several expression vectors also contain a selectable marker gene such as DHFR or NeoR, which aids in the generation of stable cell lines. The plasmid also requires a bacterial origin of replication and an antibiotic resistance gene (AmpR) to allow propagation and expansion in a bacterial host.

Transfection

Once the expression vector has been assembled, it must be inserted into the host of interest. Several methods are available for such transfections and include calcium/phosphate/DNA complexes, DEAE Dextran, electroporation, liposome, and retrovirus-mediated gene transfer.

Calcium phosphate

This method relies on the production of a calcium/phosphate/DNA microprecipitate, which is then taken up by cells by pinocytosis. The method is very effective for a number of commonly used mammalian cell expression systems

including COS, BHK, 293, and CHO cells.

DEAE dextran

This method depends on the formation of a complex between the insoluble positively charged dextran and the DNA to be transfected. Like calcium phosphate, this method is highly successful with many cell types.

Electroporation

When cells are suspended in buffer between two electrodes, discharge of an electrical impulse momentarily creates pores in the cell membrane. During this time, DNA in solution is free to diffuse into the cells. This method is highly successful in transfecting a large number of cell types, including cells previously thought to be difficult to transfect with other methods, such as endothelial cells and fibroblasts.

Liposomes

By encapsulating the DNA to be transfected in an artificial lipid carrier, foreign DNA can be introduced into the cell. This method, like electroporation, has been successful in transfecting cells previously thought difficult to manipulate. Its only drawback is its expense.

Transduction

The act of transferring a foreign gene into a host genome.

VII. Experimental Gene Manipulation

Antisense oligonucleotides

By introducing short single-stranded deoxyribonucleic acids (ODN) into a cell, specific gene expression can be interrupted. Several mechanisms have been postulated to account for these results including interruption of ribosome binding to mRNA, enhanced degradation of mRNA mediated by the double-strand specific RNAseH, DNA triplex formation, and impairment of translation efficiency. Most successful attempts using antisense ODN have targeted sequences surrounding and

including the initiation codon. To reduce nuclease attack, the antisense ODN are often synthesized using an altered chemistry involving thiol rather than phosphodiester linkages.

Transgenic animals

By introducing an intact or manipulated gene into the germline of mice, the effects of promoter expression in specific cell lineages can be investigated. In contrast to highly artificial in vitro studies using reporter gene analysis, such transgenic animals provide an important in vivo model of gene function. The methods for production of transgenic mice have been extensively reviewed and are based on the microinjection of linear DNA into the pronucleus of a fertilized egg. Several types of experiments can be performed. First, the effect of aberrant expression of a gene can be investigated, as was recently performed by expressing GM-CSF in a wide variety of tissues. Second, the necessary elements for tissue- and developmental level-specific expression of a gene can be studied, as has been performed for the β-globin locus. Third, the tissue distribution of a specific gene can be determined by engineering a marker gene adjacent to a specific promoter. A specific example of this strategy employs a "suicide gene," the herpes virus thymidine kinase (TK). When animals carrying such genes are exposed to gancyclovir, cells expressing the promoter of interest will express TK, be killed, and be readily detected.

Gene knock-out experiments

Specific genes in the mammalian genome can now be targeted for interruption or correction based on the technique of homologous recombination. By generating DNA constructs that contain an interrupted gene of interest, or a corrected gene, in the setting of adequate flanking sequences to allow for targeting to the genetic locus of interest, the endogenous gene can be replaced or corrected. The methods involve introduction of the gene into an embryonic stem (ES) cell line, selection for subclones of cells that have had successful homologous recombination events, and then introduction of the ES subclone into the blastocyst of a developing embryo. A chimeric animal results, and should the newly introduced gene become part of the germline, it can be bred to the homozygous state. Using these techniques, investigators can now determine whether a single genetic locus is responsible for a given disease, determine the significance of specific cytokines or growth factors, and generate model systems useful investigation of human disease.

Gene knock-in experiments

A similar technology to knock-out strategy, but rather than simply obliterating function of the targeted gene, the knock-in is designed to replace the locus with a specific mutation of interest.

Homologous recombination

When a manipulated gene is introduced into a cell, it can be incorporated into the genome either randomly or at a specific locus. By incorporating sequences that normally flank the desired locus, a manipulated gene can be specifically (albeit rarely) introduced into the genome. Selection for this unlikely event can be enhanced by introduction of the herpes thymidine kinase (TK) gene into the original targeting construct. Should the construct be randomly incorporated into the genome, the TK gene will also be introduced, rendering the cell sensitive to gancyclovir. If homologous recombination occurs, the TK gene will be eliminated, as there are no homologous sequences at the specific genetic locus of interest and the resultant cell will be resistant to the antibiotic.

YAC (yeast artificial chromosome)

A yeast artificial chromosome (YAC) utilizes centromeric and telomeric elements from yeast chromosomes to construct genetic elements that can be propagated in yeast and transferred into mammalian cells. Such vehicles allow the introduction of up to 200 kb or more of genetic material into the host cells. YACs are now being used to study the physiologic regulation of large genetic loci such as the β-globin region of chromosome 11.

Contig

The jargon term used to describe the assembly of clones necessary to include all of the DNA in a specific stretch of chromosome. Such maps are usually assembled from overlapping YAC (yeast artificial chromosome) or BAC (bacterial artificial chromosome) clones. Once the "genome project" is complete, it will consist of 24 (very large) contigs (22 autosomal, an X and a Y).

Transposon

Naturally occurring genetic elements that are naturally easily removed and inserted into the genome, allowing for the recombination of genetic segments, giving rise to genetic diversity. These same elements can be utilized for gene therapy.

VIII. Gene Therapy

Gene therapy takes many forms. To treat malignancy, it may involve the insertion of an adjuvant substance (such as GM-CSF) into tumor cells to generate a tumor vaccine, transfer of a gene that renders tumor cells susceptible to eradication with an antitumor agent (e.g., herpes thymidine kinase), or insertion of a gene that makes bystander cells resistant to the effects of chemotherapy (e.g., MDR). For gene deficiencies, insertion of the wild type allele is the therapeutic goal. Obtaining cDNA for desired genes has become common. Insertion of the gene into target cells and high (adequate) level expression is more problematic. Several types of transfer vehicles have found use, including viral vectors and chemical agents.

Viral transduction vectors

Retroviral vectors are based on murine retroviruses. They can carry 6 to 7 kb of foreign DNA (promoter + cDNA) but suffer from the drawbacks of requiring the development of high titer packaging lines, requiring that target cells be dividing, and are subject to host cell down-modulation. Adenoviral vectors can be produced at high levels and do not require a dividing target cell, but they do not normally integrate, resulting in only transient expression. Adeno-associated viral vectors are defective parvoviruses that integrate into a non-dividing host cell at a specific location (19q). Disadvantages are genetic instability, small range of insert size (2–4.5 kb), and thus far, only transient expression.

Ecotropic vectors

Many retroviruses are host cell specific, i.e. they will only infect a specific species of cells. An example is the widely used Maloney virus, and its basis lies in the species-specific expression of the viral cell surface receptor.

Ecotropic viruses

Murine retroviruses that contain coat proteins that can only bind to murine cellular receptors.

Amphotropic viruses

Retroviruses whose coat proteins bind to a receptor found throughout multiple species, usually including man, making these vectors suitable for human use. Problems related to the level of receptor expression on cells of hematologic interest

(e.g. stem cells) remain for amphotropic viruses.

Pseudotyped virus

These take advantage of the powerful expression levels obtainable by murine retroviral backbones, yet are packaged in an envelope that allows docking and uptake by human target cells. An example is the popular MFG vector that utilizes a murine leukemia retroviral backbone and an amphotropic packaging cell line to produce infectious particles.

Episomal

Episomal refers to gene therapy vectors that remain free in the target cell without being taken into the host genome.

Positional variegation

Refers to the observation that the site of vector integration into the genome often results in variable levels of gene expression.

Chimeraplasty

A technique of gene therapy dependent on construction of a DNA:RNA oligonucleotide hybrid that once introduced into a cell relies upon DNA repair mechanisms to introduce a (corrective) change in the targeted gene.

Chitosan-DNA

A chemical means of packaging foreign DNA to allow introduction into cells; the complexes exist as nanospheres and have been tested in factor IX deficiency in animals.

Long terminal repeat (LTR)

This segment of a retroviral genome carries the genetic information for both transcription of downstream viral structural genes and the mechanisms of viral replication. It is often used in retroviral applications to drive the exogenous therapeutic gene as it carries a powerful (but non-tissue specific) promoter.

Interference

The mechanisms by which infection of a cell by one virus excludes infection by others. Interference is often due to the cellular production of coat proteins, which bind to and block the cells' remaining viral receptors.

Nonviral transduction methods

Nonviral methods include polylysine-ligand DNA complexes, where the ligand (e.g., transferrin) allows access to the cell through normal receptor-mediated uptake, and phospholipid vesicles. Both methods suffer from not providing a mechanism for genomic integration, precluding long-term expression.

IX. Cloning and Library Screening

Obtaining cDNA representing a protein of interest is usually the first step in the process of applying the techniques of recombinant DNA research to an important physiologic question. A suitable cDNA library must first be constructed starting with RNA abundant (or as abundant as possible) in the transcripts for the gene of interest. Following library construction a probe must be developed that can specifically recognize the gene or cDNA of interest, or the expressed protein product of the specific cDNA.

First strand synthesis

The retroviral enzyme reverse transcriptase is used along with an antisense primer to produce a complementary DNA strand of mRNA extracted from a cellular source known to express the gene of interest. Two types of primers are used, either oligo dT, in which the poly A tail begins the cDNA synthesis, or random primers, in which a whole range of start sites will be used.

Second strand synthesis

The enzyme DNA polymerase is used to generate the sense strand of cDNA. Priming of the second strand can occur spontaneously, as the antisense first cDNA strand can form a hairpin loop at its 3' end bending back to prime second strand synthesis. Alternatively, a polynucleotide tail can be added to the first strand synthesis using terminal deoxynucleotide transferase, then second strand priming can occur using a synthetic oligonucleotide complementary to the TdT tail. Should the former technique be used, an extra step to nick the hairpin loop using the enzyme S1 nuclease would be required prior to inserting the cDNA into its vector.

cDNA blunting

First and second strand synthesis usually results in nonflush ends. To prepare the cDNA for insertion into a cloning vector, the ends must be made flush with one another. Such blunting reactions can be conducted with a DNA polymerase, such as the Klenow fragment of DNA polymerase I or T4 DNA polymerase.

Linkering

To efficiently insert the cDNA library into a cloning vector, synthetic duplex oligonucleotides that contain a restriction endonuclease site are attached to the blunted ends of the cDNA. A restriction endonuclease is chosen that rarely cuts DNA (such as the 8 bp recognition sequence for Not I, or if a more common restriction site is used such as Eco RI, the cDNA should first be methylated in order to prevent subsequent cDNA digestion with the enzyme) and is used to generate "sticky ends" on the cDNA.

cDNA library preparation

Once the cDNA has been prepared and sticky ends generated, the library is inserted into a convenient cloning vector. Because of high cloning efficiency, most cDNA libraries are constructed in a λ phage vector. Typically, if screening is to be performed using a monoclonal antibody, λ gt 11 is used. If screening is to be performed using oligonucleotide probes, λ gt 10 can be used. If larger DNA fragments are to be prepared, such as from genomic fragments of DNA, λ vectors that can accommodate up to 20 kb are available (e.g., λ Charon 4A).

Subtractive library

The purpose of generating a subtractive library is to enrich for cDNA that are expressed under one condition but are not expressed under a second condition. This facilitates screening for the cDNA of interest in that the complexity of the library is much reduced, requiring one to screen far fewer clones. At its extreme, investigators have used subtractive libraries to generate a very highly select group of clones (in the range of 100) and then have sequenced all of the resulting cDNA. The principle behind a subtractive library is the elimination of cDNA common to induced and control conditions. By eliminating such clones, only cDNA that are present under the induced conditions will remain in the library. Those techniques

depend on the differential elimination of duplex mRNA/cDNA or cDNA/cDNA hybrids, which form between genes expressed under both conditions, leaving the single-stranded mRNA or cDNA of interest.

RDA (representational difference analysis)

A molecular method to amplify genes that are expressed in an RNA sample of interest, that are not present, or present at very reduced levels, in a comparison RNA sample (e.g. cytokine induced and control cells). The method relies on RT-PCR amplification of the RNA that does not contain the gene(s) of interest to produce a "driver" cDNA, and RT-PCR to produce "tester" cDNA from the RNA population in which you hope to find new genes. After ligation of different oligonucleotides to the ends of each population, both are denatured and an excess of the driver is hybridized to the tester and PCR performed with primers that will amplify only sequences present in the tester that are not in the driver, thereby "removing" cDNA common to both populations. The resultant cDNA are enriched in uniquely expressed genes.

Directional cloning

To improve efficiency when screening functional expression libraries, many investigators construct cDNA libraries in which the proper coding orientation of the cDNA is maintained in the library. In conventional library preparation, the 5' and 3' ends of the DNA are identical; thus, cDNA can be inserted into the cloning vector in either orientation. If screening is dependent on the production of a functional protein, one-half of the library will be useless, as those cDNA inserted in an inverse orientation will not produce functional protein. Directional cloning is dependent on producing sticky ends that differ on the 5' and 3' termini. The cloning vector has the appropriate pair of complementary cloning sites.

Library screening

Three major methods are available to obtain cDNA of interest. The classic technique utilizes DNA probes (such as oligonucleotides or intact cDNA from a homologous gene) to screen cDNA libraries. An oligonucleotide probe is usually derived from a reverse translation of known protein sequence. By expressing cDNA as a fusion protein with β galactosidase, various antisera can be used to screen for fusion proteins encoded by the cDNA of interest. Finally, cDNA libraries may be constructed in cloning vectors that allow for expression of the cDNA insert in *E. coli* or a mammalian cell host. If a highly sensitive assay for the desired protein's

function can be developed, pools of cDNA clones can be expressed and then assayed together; a positive assay from a pool would allow one to subdivide into smaller pools and eventually at clonal density.

Reverse genetics

Often, large families of homologous proteins exist and multiple previously unknown members of the family can be obtained by screening cDNA libraries under low stringency using cDNA or oligonucleotide probes from regions highly conserved amongst members of the family. In this case, genes are identified before their function is known, a situation referred to as reverse genetics. Examples in hematology include identifying members of the tyrosine kinase family of receptor proteins using a probe derived from the conserved kinase domain of the cytoplasmic region of src or other tyrosine kinase proto-oncogenes, or the identification of transcription factors important in hematopoiesis using conserved motifs present in zinc finger or homeodomain proteins.

X. Oncogenesis and Anti-Oncogenes

Oncogenes have usually been identified in the context of a tumor-inducing virus. Such viral oncogenes (*v-onc*) are thought to be derived from host cells, but have been altered such that abnormal regulation of production or function has ensued during the transfer process. Subsequent reintroduction of the altered gene into a host cell leads to transformation. Proto-oncogenes, the normal cellular counterpart of viral oncogenes, can contribute to cellular transformation by mechanisms that disturb normal gene function. Such mechanisms include mutation (resulting in abnormal function), amplification (resulting in abnormal levels of expression), rearrangement (resulting in a new function), or promoter mutation (again resulting in abnormal levels of expression). Most or all proto-oncogenes are involved in normal cellular processes such as growth factor signal transduction, mitogenic signaling, or regulation of DNA transcription or cellular proliferation. The nomenclature convention is to indicate the cellular version of the proto-oncogene as "*c-onc*" and the viral version, which is transforming, as "*v-onc*." Most altered proto-oncogenes act in a dominant genetic fashion. Anti-oncogenes, or tumor suppressor genes, usually act in a recessive genetic fashion and function to slow processes involved in cellular proliferation. Most of the identified anti-oncogenes have been involved in gene transcription, presumably acting to enhanced differentiation programs over those of proliferation.

c-abl

This gene, present on human chromosome 9, encodes a tyrosine kinase whose role in normal hematopoiesis is unclear; however, its fusion to the BCR gene on human chromosome 22, the functional counterpart of the Ph1 chromosome strongly associated with the disease chronic myelogenous leukemia, eliminates the first two or three exons of *c-abl* and results in unregulated tyrosine kinase activity. The resultant fusion protein is either 210 kDa or 195 kDa. The latter version is more acutely transforming in experimental settings; it is also associated with acute lymphoblastic leukemia and with a worse prognosis in both disease settings. One of the ways in which the unregulated kinase activity may be manifest is through phosphorylation of SHC and/or GRB-2, adopter proteins necessary for coupling growth factor signals to ras.

c-jun

This proto-oncogene encodes a ~45 kDa transcription factor that is a member of the AP1 family of transcriptional proteins. *c-jun* must form dimers to function and does so through the leucine zipper motif. Although *c-jun-c-jun* homodimers do form, they do so with low affinity and are not thought to be critical in gene transcription. Rather, a second partner, usually *c-fos*, generates the transcriptionally active heterodimer.

c-fos

This ~62 kDa leucine zipper protein cannot homodimerize but rather functions in heterodimeric complex with *c-jun* and other members of the AP1 family of transcription factors.

c-myc

This proto-oncogene plays a critical role in hematopoietic cell proliferation. Like the leucine zipper protein, it too functions as a heterodimer. One of its partners is termed Max. The *myc*-related protein, Mad, also dimerizes with Max; the *myc*/Max complex stimulates proliferation, the Mad/Max complex inhibits myc-function. The importance of dysregulated myc function can be seen in Burkitt lymphoma in which a t(8;14) brings *myc*, on chromosome 8, into juxtaposition with the immunoglobulin enhancer on chromosome 14. Such upregulation of *myc* in a B lymphocyte setting results in a proliferative advantage and represents one important step in the genesis of this lymphoma. *Myc* has both leucine zipper and helix-loop-helix domains.

c-myb

This gene encodes a transcription factor not belonging to any other class previously described and is expressed primarily in immature hematopoietic cells and declines as cells differentiate. Forced expression of *c-myb* tends to block hematopoietic differentiation. Clinically, high levels of *myb* are noted in acute leukemia, and such patients are less likely to enter remission or tend to have a short remission duration.

c-rel

This gene belongs to the NF-κB family of transcription factors and can act to enhance or repress transcription from selected genes. This family of proteins includes p50 and its precursor p105, p65, p49 and its precursor p100, and Bcl-3, one of the IκB family.

IRF-1 (interferon regulatory factor-1)

IRF-1 is a transcription factor that activates the expression of IFN α and β and maps to chromosome 5q31.1. As it is thought to act as a tumor suppressor gene, its role in the pathologic consequences of the 5q- syndrome is under active investigation.

IRF-2 (interferon regulatory factor-2)

Interferon regulatory factor-2 is a gene which binds to a promoter element shared by IFN α and β and many IFN-inducible genes; unlike IRF-1, which stimulates such genes, IRF-2 represses transcription at the site. It is felt that the ratio of IRF-1 to IRF-2 might be a critical event in the regulation of cellular proliferation.

Rb

The prototypical tumor suppressor gene *Rb* behaves in a genetically recessive fashion. Elimination or inactivation of both *Rb* gene copies is required for manifestation of the tumorgenic phenotype, first recognized in children with retinoblastoma. Such children inherit only a single functional copy; subsequent mutagenic inactivation of the remaining allele results in tumor susceptibility. *Rb* acts to sequester a group of transcription factors, termed E2F, which regulate genes critical for DNA synthesis. Alterations of *Rb* alleles are found in approximately 30% of human acute leukemias.

SCL

This proto-oncogene, first identified in a stem cell leukemia at the site of t(1;14), is a member of the helix-loop-helix group of transcriptionally active proteins. The gene, also termed *Tal 1*, is expressed in erythroid and mast cell lineages but not in T cells. The association of t(1;14) with up to 25% of T cell ALL suggests that its ectopic expression is associated with transformation.

Bcl-1

This gene, located on chromosome 11 q13, was first identified at the site of translocation p(11;14)(q13;q32), has a strong association with central acinar/mantle cell lymphoma and functions in normal cells as the G_1 cyclin termed CCND1 or cyclin D_1. Normally, lymphocytes lack cyclin D_1 expression; its aberrant expression resulting from chromosomal translocation of the Bcl-1 locus to an immunoglobulin locus is thought to be associated with aberrant proliferation.

Bcl-2

This gene product normally functions to suppress programmed cell death (apoptosis). Its overexpression is associated with the most common molecular abnormality in non-Hodgkin's lymphoma, t(14;18)(q32;q21), present in 80% of follicular small cleaved cell lymphoma. Presumably, suppression of apoptosis leads to extended cell survival, a characteristic of low-grade lymphomas.

Bcl-3

This gene is a member of the IκB family. Presently, it is unclear how this protein acts in tumorigenesis, but it is likely that its involvement in transcriptional processes is critical.

Bcl-6

A zinc finger transcription factor, expression of which is altered in approximately one-third of diffuse B large cell lymphomas as a consequence of 3q27 translocations. Its target genes are unknown.

RAR (retinoic acid receptor)

The retinoic acid receptor is a member of the steroid hormone group of transcriptionally active proteins and contains a steroid hormone-binding domain, a zinc finger DNA-binding domain, and a transcriptional activation domain. RAR is

located at the t(15;17) present in the majority of cases of acute promyelocytic leukemia. Its fusion partner in the translocation is termed pml. Normally, RAR forms heterodimers with members of the RXR family of transcription factors.

p53

Wild-type p53 is a sequence-specific DNA-binding nuclear protein that acts to induce gene expression. Overall, the program of p53-activated genes is associated with suppression of cell growth, consistent with our understanding of the mechanisms of anti-oncogenes. Mutations of p53 may not only inactivate its growth-suppression function, but can actually generate a genetically dominant, functional oncogene. Human tumors associated with p53 mutations include those of hematopoietic tissues (e.g., 20% of myelomas), bladder, liver, brain, breast, lung, and colon. It is likely the most frequently mutated gene in human cancer.

ras

This gene encodes a critical signalling intermediate involved in the response to multiple growth factors. There are several related proteins (Ha-ras, Ki-ras, N-ras). N-ras and K-ras are mutated in many cancers, including 45% of myelomas and > 50% of CMML cases. Constitutive activation of ras can mimic chronic stimulation by the corresponding lineage-specific growth factor.

Hox 11

A homeobox containing transcription factor disrupted by translocation to the T cell receptor locus [t(10;14)] in 10% of cases of T cell ALL/lymphoblastic lymphoma. The Hox 11 gene is critical to the development of the spleen but its role in hematopoiesis is unclear.

Rhomb 2

Like Hox 11, Rhomb 2 is translocated in T cell ALL/lymphoma associated with t(11;14). Rhomb 1 may play a similar role in additional cases of T cell ALL. The Rhomb gene products are members of a family of transcription factors, but as Rhomb 2 and Rhomb 1 do not contain DNA-binding domains, they are thought to be involved in protein-protein interactions. Neither Rhomb 2 or 1 are normally expressed in T cells; transformation involving these genes, like SCL or Hox 11 is thought to be due to ectopic expression of the protein in T cells.

ALK (anaplastic lymphoma kinase)

A large proportion of Ki-1 positive lymphomas are characterized by a t(2;5). The breakpoint involves nucleophosmin, a ubiquitously expressed gene, and ALK. The chimeric mRNA and protein are thought to be responsible for transformation. ALK is a member of the insulin receptor family of transmembrane receptor kinases, which is not normally expressed in hematopoietic tissues; the fusion gene is no longer membrane bound, which may underlie its pathogenesis.

Evi-1

A transcription factor whose rearrangement in t(3;21) is implicated as contributing to MDS. Overexpression of evi-1 blocks differentiation in response to hematopoietic growth factors.

ETO

Located on chromosome 8, ETO is involved in t(8:21) of AML type M2. Based on the presence of two zinc-finger motifs. ETO possibly encodes a transcription factor, but its role in the pathogenesis of AML is unknown.

AML-1

Located on chromosome 21, AML-1 is the fusion partner of ETO in t(8:21). The gene is homologous to the runt gene of *Drosophila* and encodes a transcription factor. Normal hemopoietic targets include the CD13, GM-CSF, MPO, IL-3, and the T cell antigen receptor promoters. AML-1 binds as a heterodimer, partnered with CBFβ. It is unclear if its mechanism of action is to enhance aberrant transcription or to blunt transcription by acting in a dominant negative fashion.

CBFβ

Located on chromosome 16, CBFβ is one of the fusion partners in the inv(16) associated with AML type M4Eo. As with AML-1, it is unclear whether the altered transcription factor enhances or blocks transcription.

MLL

Located on chromosome 11, MLL (mixed lineage leukemia) is frequently altered in ALL, 1°AML, and especially in AML secondary to the use of topoisomerase II

inhibitors. MLL is homologous to the trithorax gene of *Drosophila* and displays many features of a transcription factor and of a DNA methyl transferase.

Tel

A helix-loop-helix transcription factor fused to the PDGFβ-receptor in CMML t(5:12) and to other genes in AML or MDS. Like most other translocation oncogenes, the mechanism of leukemogenesis is unknown. More recently, a Tel/AML1 fusion gene representing a t(12;21) has been found in a large number of cases of childhood ALL. As the translocation is not detected by routine cytogenetics, molecular analysis (FISH, etc.) is required to identify this favorable chromosomal rearrangement.

DEK

Located on chromosome 6, DEK is involved in t(6:9) of AML. This translocation is usually seen in young patients and carries a poor prognosis. Its normal function is unknown, but DEK localizes to the nucleus.

CAN

Located on chromosome 9, CAN is part of t(6:9). CAN forms part of the nuclear pore. As it has two different fusion partners but a consistent phenotype, CAN is likely the critical component of t(6:9).

Fas (CD95 or Apo-1)

A transmembrane glycoprotein expressed on a wide variety of primitive and mature hematopoietic cells, which, upon binding to its natural ligand triggers programmed cell death.

NF-1

The gene responsible for neurofibromatosis. The normal protein functions to negatively regulate ras proteins, key intermediates in cytokine-induced cellular proliferation.

XI. Genetic Screening

X-linked methylation patterns

Several loci present on the X chromosome become highly methylated when inactive but remain unmethylated on the active X chromosome (Lyon hypothesis). Should a polymorphic site for a methylation-sensitive restriction endonuclease exist at such an X-linked locus, one can distinguish between the active and inactive X chromosome by the pattern of restriction endonuclease digestion of that gene. However, in order to be widely useful for determining clonality of hematopoietic cells, the allelic frequency must be close to equality. Several X-linked genes meet these criteria and include phosphoglycerate kinase (PGK), hypoxanthine phosphoribosyltransferase (HPRT), the human androgen receptor gene (HUMARA), and the hyper-variable DXS255 locus. Both Southern blotting and PCR methods can be applied to this type of analysis.

RFLP (restriction fragment length polymorphism)

If a mutation of one allele of a genetic locus either generates or destroys a restriction endonuclease site, the heterogeneity present within or very close to a gene of interest can be used to track which allele an individual has inherited from each parent. When genomic DNA is digested with a restriction enzyme that recognizes a polymorphic site and then hybridized with a probe specific for the gene of interest, the allelic pattern can be compared to that of a similar assessment of both parents. The presence of multiple family members allows a complete genetic pedigree to be constructed. For example, globin gene mutations such as sickle hemoglobin can be analyzed. The β6 mutation in hemoglobin, which results in Hgb S, destroys an Mst II site. Therefore, a larger than normal DNA fragment is generated by digestion of genomic DNA with Mst II, which can be easily detected by Southern blot hybridization. In this specific case, the Mst II polymorphism is absolutely specific for the mutant gene and family studies are not necessary. If the RFLP had not been specific for the mutation, but only existed close to the specific disease-producing mutation, then family studies would have been required to determine which pattern (presence or absence of restriction site) tracks with the mutant (disease) allele.

Allele-specific hybridization

If the nucleotide basis for a specific genetic abnormality is known, oligonucleotides specific for wild type and for mutant sequence can be designed and used to probe Southern blots of an individual's genomic DNA. The pattern of hybridization thus gives specific information regarding which alleles are present. In a polymorphic disease such as β thalassemia (in which multiple mutations can give rise to the same disease phenotype), multiple probes might be required to detect all possible causes.

In addition, new mutations causing the same disease would be missed. However, should a specific probe prove useful for one population group or be positive in one family member, that probe becomes very useful for the individual under study.

Reverse allele-specific hybridization

This automated variant of allele-specific hybridization couples unlabeled synthetic oligonucleotides specific for a wild type or mutant sequence to a solid support that is then allowed to bind genomic sequences of the locus of interest, which have been amplified by PCR. The use of highly stringent conditions of hybridization allows differential binding of the amplified DNA to the wild type or mutant specific oligonucleotide and thereby allows genotypic determination of the individual.

Competitive oligonucleotide hybridization

Mutant or wild type-specific oligonucleotide primers are used in a PCR reaction with genomic DNA. The primers and stringency of PCR are chosen so that single-based mismatches between genomic DNA and PCR primer fail to yield an amplified product. Thus, the PCR detection of a locus-specific product allows the genotyping of the individual.

Color complementation assay

This method is an advancement over competitive oligonucleotide hybridization in that the wild type and mutation specific PCR primers are labeled with different color fluorescent tags, and both are used in a PCR reaction with genomic DNA. When highly stringent conditions are met, the fluorescent colors of the resultant PCR product indicate whether wild type, mutant, or both specific alleles were present in the original DNA sample.

Multiple Choice Questions

Q.1- A critical enzyme used directly in the synthesis of dTMP (thymidine) is

a) Carbamoyl phosphate

b) Aspartate Transcarbamoylase

c) Dihydrooratase

d) Thymidylate synthase

Q.2- In which phase of the cell cycle does DNA replication occur?

a) G_0

b) G_1

c) S

d) G_2

Q.3-A 10 month old baby boy presents with steatorrhea, recurrent pulmonary infections, GI upset and foul-smelling stool. Which of the following tests is undertaken to confirm your diagnosis?

a) Sweat test

b) Blood glucose

c) CBC

d) RFLP analysis

Q.4-The enzyme responsible for initiating DNA replication in prokaryotes is:

a) DNA polymerase I

b) DNA polymerase III

c) Polymerase beta

d) Primase

Q.5- The enzyme responsible for continuing DNA replication in prokaryotes, once it is initiated is:

a) DNA polymerase I

b) DNA polymerase III

c) polymerase beta

d) polymerase delta

Q.6- Antibiotics such as Ciprofloxacin and Flouroquinolines work by inhibiting a specific enzyme. This enzyme is normally necessary to relieve torsional strain that is caused by the unwinding of the helix. What is the name of this enzyme?

a) DNA ligase

b) Topoisomerase (DNA Gyrase)

c) single-stranded binding protein

d) primase

Q.7- Which of the following techniques is primarily undertaken to amplify DNA?

a) PCR

b) Microarrays

c) Northern Blotting

d) Southern Blotting

Q.8- In Pyrimidine Synthesis, Eukaryotes can use uracil to feedback inhibit which of the following enzymes?

a) Carbamoyl phosphate synthase

b) Aspartate Transcarbamoylase

c) Dihydrooratase

d) Thymidylate synthase

Q.9- All of the following are required for PURINE DE NOVO synthesis EXCEPT:

a) CO2

b) Glycine

c) Glutamine

d) Isoleucine

Q.10- Which of the following, in high concentrations, can overcome PURINE DE NOVO inhibition?

a) Ribose-5-phosphate

b) PRPP

c) 5-phosphoribosyl 1-amine

d) GDP

Q.11- A 40-year-old alcoholic comes in with severe pain in his big toe. You decide to administer Allopurinol to inhibit which of the following enzymes:

a) Xanthine Oxidase

b) Thymidine kinase

c) Adenosine Deaminase

) Adenine phosphoribosyl transferase

Q.12- All of the following are used in PCR EXCEPT:

a) Taq polymerase

b) Restriction enzymes

c) Oligonucleotide primers

d) Deoxynucleoside triphosphates

Q.13- The following are features of DNA replication EXCEPT:

a) Semi-conservative

b) Semi-discontinuous

c) unidirectional

d) chain growth in the 5′ –> 3′ direction

Q.14- The oncogene Ras binds:

a) ATP

b) GTP

c) Glucose

d) Hemoglobin

Q.15- Which out of the following mechanisms is involved in the production of variety of immunoglobulins each specific for a specific antigen?

a) Class switching

b) Gene amplification

c) Gene rearrangement

d) RNA editing

Q.16-Which out of the followings is a common enzyme for de novo as well as salvage pathway of purine biosynthesis?

a) Amidotransferase

b) PRPP synthetase

c) HGPRTase

d) Adenylosuccinate synthetase

Q.17-Which of the following does not have introns?

a) DNA

b) Non-processed pseudo genes

c) Processed m RNA

d) Primary RNA transcript

Q.18- Which of the following is NOT true of RNA synthesis?

a) The key enzyme is RNA polymerase

b) The energy is supplied by ring cleavage

c) The RNA sequence is complementary to the template strand of DNA

d) The RNA sequence is of opposite polarity to the template strand of DNA

Q.19- If the molar amount of G in a DNA sample is 20%, what is the molar amount of T in the sample?

a) 20%

b) 30%

c) 40%

d) 60%

Q.20- With respect to the LAC operon, if both glucose and lactose are present and glucose is low, which of the following is NOT true?

a) High CAP

b) increased uptake of lactose

c) low cAMP

d) increased transcription of the lac operon

Q.21- Which of the following subunits of the bacterial RNA polymerase is responsible for promoter recognition?

a) alpha

b) B

c) B′

d) sigma

Q.22-Which is true of the melting temperature of G-C pairs compared to A-T pairs in DNA?

a) The Tm are equal

b) Tm of G-C is less than the Tm of A-T

c) Tm of G-C is greater than the Tm of A-T

d) None of the above

Q.23- A 10-year-old boy comes to the ER after eating wild mushrooms. The poison associated with these mushrooms most likely DIRECTLY inhibits the synthesis of the following:

a) hn RNA

b) t RNA

c) DNA

d) m RNA

Q.24- DNA methylation is associated with:

a) CpG islands

b) CAT box

c) TATA box

d) increasing gene transcription

Q.25-Alternative splicing…

a) Creates protein from multiple segments of DNA on different chromosomes

b) Is the reason why the human genome is much more complex than other species

c) Creates different proteins from a single gene

d) is not tissue specific

Q.26-All of the following are involved in translating information into proteins EXCEPT:

a) rRNA

b) siRNA

c) tRNA

d) snRNA

Q.27-Which histone is NOT part of the nucleosome?

a) H1

b) H2A

c) H2B

d) H3

Q.28-Which out of the followings is an inhibitor of prokaryotic transcription?

a) Ciprofloxacin

b) Etoposide

c) Erythromycin

d) Rifampicin

Q.29- Choose the nucleoside analogue used as an anticancer drug out of the followings

a) Methotrexate

b) 6- Mercaptopurine

c) Vinblastin

d) Cytosine Arabinoside

Q.30- Which amino acid residue is in abundance in histones?

a) Arginine

b) Aspartic acid

c) Tryptophan

d) Phenyl alanine

Q.31– Which out of the following techniques is used for the detection of gene of interest –

a) Southern Blotting

b) Polymerase chain reaction

c) Northern Blotting

d) DNA Foot printing

Q.32–Which out of the followings is an example of post translational modification?

a) Splicing

b) Class switching

c) Subunit aggregation

d) Base modification

Q.33- Which of the following is a required substrate for purine biosynthesis?

a) 5- methyl thymidine

b) Ribose phosphate

c) Ara C

d) PRPP (5- phosphoribosylpyrophosphate)

Q.34-Triple repeat sequence disease occurs in:

a)Alzheimer's disease

b) Cystic fibrosis

c) Ataxia telangectasia

d) Huntington's chorea

Q.35-Northern blotting is used for separation of:

a) DNA

b) mRNA

c) Protein

d) Protein DNA interactions

Q.36.The term cistorn, muton and recon were introduced by
(A) Watson and Crick
(B) S. Benzer
(C) Meselson
(D) Morgan

Q36.Extranuclear genetic material is found in
(A) Plastid and nucleus
(B) Mitochondria and plastids
(C) Nucleus and cytoplasm
(D) Mitochondria and nucleus

Q37. The molecular formulae of deoxyribose sugar and ribose sugar respectively are
(A) $C_5 H_{10} O_4$ and $C_5H_{10}O_6$
(B) $C_5 H_{10} O_4$ and $C_5H_{10}O_5$
(C) $C_5 H_{10} O_5$ and $C_5H_{10}O_4$
(D) $C_5 H_{10} O_5$ and $C_6H_{10}O_4$

Q38. The nitrogen bases which pair with two hydrogen bonds are
(A) Adenine and thymine
(B) Adenine and Cytosine
(C) Cytosine and guanine
(D) Cytosine and adenine

Q39. DNA differs from RNA in
(A) Presence of deoxyribose sugar
(B) Presence of thymine base
(C) Property of replication
(D) All the above

Q40. DNA molecules makes a complete turn after every
(A) 20 Å
(B) 34 Å
(B) 3.4 Å
(D) 10 base pairs

Q41.The distance between two successive nitrogenous base pairs is
(A) 34 Å
(B) 36 Å
(C) 20 Å
(D) 3.4 Å

Q42.In nucleoside, nitrogen base is attached to pentose sugar at
(A) Carbon – 1 of pentose sugar
(B) Carbon – 2 of pentose sugar
(C) Carbon – 4 of pentose sugar
(D) Carbon – 5 of pentose sugar

Q43. If the strand of DNA has 35 nucleotide how many phosphodiester bonds would exist
(A) 34
(B) 35

(C) 24
(D) 70

Q44.. In eukaryotic DNA replication, lagging strand is formed by
(A) RNA fragments
(B) Okazaki fragments
(C) DNA fragments
(D) Nucleotide fragments

Q45.. The enzyme DNA polymerase can work only in
(A) 3¢ ® 5¢ direction
(B) 5¢ ® 3¢ direction
(C) Both the direction
(D) 5¢ ® 5¢ direction

Q46. Enzyme required for removing RNA primer during DNA replication is
(A) DNA primase
(B) DNA ligase
(C) DNA polymerase I
(D) DNA polymerase III

Q47. During DNA replication, the reunion or recoiling of separated DNA strand is prevented by
(A) Helix destabilizing protein
(B) Single strnad binding protein
(C) Rep protein
(D) Both (A) and (B)

Q48. The enzyme that cuts the bonds of DNA molecule at the origin of replication is
(A) Endonuclease
(B) DNA polymerase
(C) DNA gyrase
(D) DNA ligase

Q49. Which of the following enzyme is required to release the tension imposed by uncoiling of strands?
(A) Endonuclease
(B) DNA ligase
(C) DNA gyrase

(D) DNA helicase

Q50. The cellular composition of m-RNA is
(A) 5-10%
(B) 3-5%
(C) 10-20 %
(D) 70-80%

Q51. Formation of mRNA from DNA is called
(A) Transformation
(B) Transduction
(C) Traslation
(D) Transcription

Q52. The ratio of purines and pyrimidines in mRNA is not 1:1 because the nitrogenous bases are
(A) Unpaired
(B) Paired
(C) Paired only in loops
(D) Paired in stems

Q53. The codons which may present at 3¢ end of mRNA
(A) UAA
(B) UAG
(C) UGA
(D) Any one of these

Q54. Which of the following is not tool of genetic engineering?
(A) Vectors
(B) Enzymes
(C) Foreign DNA
(D) GMO

Q55. In recombinant DNA technology a plasmid vector is cleaved by
(A) Modified DNA ligase
(B) A heated alkaline solution
(C) The same enzyme that cleave the donor DNA
(D) The different enzyme other than that cleave the donor DNA

Q56. The most common plasmid vector used in genetic engineering is
(A) PBR 328
(B) PBR 322

(C) PBR 325
(D) PBR 330

Q57. 'Nif gene' for nitrogen fixation is cereal crops like wheat, jowar etc. is introduced by cloning
(A) Rhizobium meliloti
(B) Bacillus thuringiensis
(C) Rhizopus
(D) Rhizophora

Q58. Eco RI is an
(A) Ligase
(B) Polymerase
(C) Restriction enzyme
(D) Gyrase

Q59. The transgenic plant flavr savr tomato carries an artificial gene for
(A) Delay ripening process
(B) Longer shell life
(C) Added flavours
(D) All of these

Q60. Hirudin is obtained from the transgenic plant
(A) Brassica napus
(B) Hibiscus rosasinesis
(C) Raphanus sativus
(D) Vinca rosea

Q61. Bt Cotton is
(A) Cloned plant
(B) Transgenic plant
(C) Hybrid plant
(D) Mutated plant

Q62. Dolly sheep was genetically similar to
(A) The mother from which nucleated fertilized egg was taken
(B) The mother from which nuclear DNA of udder cell was taken
(C) The surrogate mother
(D) Both surrogate mother and nuclear donor mother

Q63. Genome is
(A) Genes on nuclear DNA
(B) Nuclear DNA + mitochondrial DNA
(C) Nuclear DNA + chloroplast DNA
(D) Nuclear DNA + Mitochondrial DNA + Chloroplast DNA

Q64. A technique of using very small metal particles coated with desired gene in the gene transfer is called
(A) Electroporation
(B) Microinjection
(C) Liposome
(D) Biolistics

Q65. The complete set of chromosomal and extrachromosomal genes of an organisms is called
(A) Genome
(B) Gene pool
(C) Gene bank
(D) Gene library

Q66. The study of all the proteins coded by the genome is called
(A) Proteome
(B) Proteomics
(C) Genome
(D) Protein formation

Q67. Sequencing of genomic DNA is included under
(A) Structural genomics
(B) Functional genomics
(C) Proteomics
(D) Transgenesis

Q68. Gene expression, regulation and phenotype production are studied in second phase of genome analysis called
(A) Structural genomics
(B) Functional genomics
(C) Proteomics
(D) Transmeiosis

Q69. A flowering plant lily have more DNA than

humans
(A) 10 times
(B) 15 times
(C) 18 times
(D) 13 times

Q70. In forensic science which of the following is used?
(A) Bacterial cloning
(B) DNA foot printing
(C) DNA fingerprinting
(D) DNA cloning

Q71. DNA fingerprinting is based on
(A) Occurance of VNTR's
(B) Knowledge of human karyotype
(C) Cloned DNA
(D) Recombinant DNA

Q72. VNTRs represnets-
(A) New terminal regions in DNA
(B) Functional genes in the DNA
(C) Split genes in the sample DNA
(D) Specific non-coding sequences with unique tandem repeats

Q73. Which ones produce androgenic haploids in anther cultures?
(A) Anther wall
(B) Tapetal layer of anther wall
(C) Connective tissue
(D) Young pollengrains

Q74. Variations observed during tissue culture of some plants are known as
(A) Clonal variations
(B) Somatic variations
(C) Somaclonal variations
(D) Tissue culture variations

Q75. Virus free plants can be obtained through

(A) Anitibiotic treatment
(B) Bordeaux micture
(C) Root tip culture
(D) Shoot tip culture

Q76. To raising of plants from a small tissue in culture is known as
(A) Macroproduction
(B) Micropropagation
(C) Tissue culture
(D) Mass production

Q77. Callus is
(A) Tissue that forms embryo
(B) an insoluble carbohydrate
(C) Unorganised actively dividing mass of cells maintained in culture
(D) Tissue that growth to form embryoid

Q78. Biopatents are _____.
(A) Right to use invention
(B) Right to use biological entities
(C) Right to use products
(D) Right to use process

Q79. African plant Pentadiplandra is used as_____.
(A) Low calories sweetner
(B) 2000 times sweeter agent
(C) Sweetner for diabetic patients
(D) All of these

Q80. Which organism was used as bioweapon derived from_____.
(A) Clostridium
(B) Yerstsinia pestis
(C) Fusarium species
(D) Green algae

Q81. A set standards used to regulate own or community activity in relation to biological world is
(A) Bionotency

(B) Biopiracy
(C) Biowar
(D) Bioethics

Q82. Biopiracy means
(A) Use of biopatents
(B) Thefts of plants and animals
(C) Stealing of bioresources
(D) Exploitation of bioresources without authentic permission

Q83. Bioethcs is related to
(A) Preventing biopiracy
(B) Regulation of unethical activities likegene cloning in animals
(C) Preventing theft of living materials
(D) Moral guidance to the problems in biology

Q84. Three dimensional shape of tRNA is
(A) L-shaped
(B) Clover leaf-like
(C) X-shaped
(D) Y-shaped

B.Sc. (Part II) EXAMINATION, 2011

Faculty of Science

[Also common with subsidiary Paper of B.Sc. (Hons.)Part-II]

(Three Year Scheme of 10+2+3 Pattern)

BOTANY

Second Paper – (Molecular Biology and Biotechnology)

Time: 3 hours **MM: 33**

1. No supplementary answer-book will be given to any candidate. Hence the candidates should write the answer precisely in the Main answer-book only.
2. All the parts of one question should be answered at one place in the answer-books. Once complete question should not be answered at different places in the answer-book.

All questions are compulsory.

(i) The term triplet code for genetic code was speculated by :
- (a) Waston and Crick
- (b) Nirenberg
- (c) Gamow
- (d) Conrat

(ii) Nuclesome is the unit of:
- (a) Protein
- (b) DNA
- (c) RNA
- (d) Nucleic acid

(iii) Protein helping in the opening of DNA double helix in front of replication fork is:
- (a) DNA gyrase
- (b) DNA polymerase
- (c) DNA ligase
- (d) DNA topisomerase.

(iv) Which resistant genes are found in P^{BR322}vector?
- (a) Only ampicillin resistant gene
- (b) Only tetracycline resistant gene
- (c) Both the above

(d) None of the above.

(v) The 'Central Dogma' concept was proposed by:
(a) Stahl
(b) Crick
(c) Watson
(d) Astbury.

(vi) The part of Ti plasmid transferred to host genome is:
(a) r-RNA
(b) C-DNA
(c) m-RNA
(d) T-DNA.

(vii) C-DNA synthesis takes place on:
(a) t-RNA
(b) r-RNA
(c) m-RNA
(d) DNA

(viii) Bt toxin contains:
(a) Enzyme
(b) Alkaloid
(c) Lipid
(d) Dry protein

(ix) Chemical synthesized by T-DNA is ;
(a) Lipid
(b) Hormone
(c) Vitamin
(d) Opines.

(x) Which substance secreted by plant activates VIR genes of Agroba terium ?
(a) Auxin
(b) Acetosyringone
(c) Cytokinin
(d) Tryptophan.

(xi) When was NBTB established ?
(a) 1982
(b) 1986
(c) 1988
(d) 1955

Fill in the blanks.

(xii) …… sequence is known as TATA Box.
(xiii) …….is called Amber Cordon.
(xiv) Generally …….. culture is used for production of bioactive molecules.
(xv) Liposomes artificial ……vesicles.
(xvi) ……..restriction enzymes are used in recombinant DNA technology.

Write short answers of the following.

(xvii) What is the facility of pollen culture?
(xviii) Where is 'Lal Bahadur Shastri Centre for Advanced Research in Biotechnology' located and when did it start functioning?
(xix) What are transposons ?
(xx) What is ICGEB ? Where is it located and when did it start functioning?

2. Write notes on the following :

a. Meselson and Stahl replication experiment
b. Hershey and Chase experiment
c. Chargaff rule.

Or

Write notes on following :

a. PCR technique and its applications
b. DNA fingerprinting and its uses.

3. Describe Central Dogma concept in detail.

Or

For gene regulation in prokaryotes describe lac operon in E-coli bacterium.

4. Write notes on the following :

a. The concept of Cellular Totipotency
b. Media preparation and aseptic culture technique.

Or

Describe somatic hybridization technique. Give a brief account of cybrids.

5. Define recombinant DNA technology and its various steps in detail.

Or

What are transgenic plants? Explain their merits and demerits with suitable examples.

B.Sc. (Part II) EXAMINATION, 2010

Faculty of Science

[Also common with subsidiary Paper of B.Sc. (Hons.)Part-II]

(Three Year Scheme of 10+2+3 Pattern)

BOTANY

Second Paper – (Molecular Biology and Biotechnology)

Time: 3 hours **MM: 33**

1. No supplementary answer-book will be given to any candidate. Hence the candidates should write the answer precisely in the Main answer-book only.
2. All the parts of one question should be answered at one place in the answer-books. Once complete question should not be answered at different places in the answer-book.

All questions are compulsory.

1.

(i) The term 'Molecular Biology' was first used by the scientist:
- (a) W. Weaver
- (b) J.D. Watson
- (c) F.H. C. Crick
- (d) G. Beadle

(ii) The scientist who proposed 'Central Dogma' concept is:
- (a) Astbury
- (b) Watson
- (c) Crick
- (d) Stahl

(iii) Among the following which basic amino acids are found in chromatin?
- (a) Arginine and Lysine
- (b) Methionine and Serine
- (c) Phenyl alanine and Proline
- (d) Tryptophan and Tyrosine

(iv) The scientist who first of all developed DNA fingerprinting technique is:
- (a) Alec Jeffreys
- (b) Eriscon
- (c) Garrod
- (d) Griffith

(v) Example of thermal stable DNA polymerase is:
(a) Taq
(b) P Flu
(c) Vent
(d) All the above

(vi) The first amino acid which is used in polypeptide chain initiation is:
(a) Glutamine
(b) Lysine
(c) Proline
(d) Methionine

(vii) The enzyme which is used for the hydrolysis of Lactose sugar is:
(a) Nuclease
(b) Protease
(c) Amylase
(d) β Galactosidase

(viii) The 'National Centre for Plant Biotechnology Research' is situated in:
(a) Karnal
(b) Bareilly
(c) Lucknow
(d) Delhi

(ix) Which plant cells lack tot potency?
(a) Parenchymatous
(b) Collenchymatous
(c) Mature Xylem vessels
(d) All of the above

(x) In the field of tissue culture technique the concept of synthetic seeds was given by the scientist:
(a) T. Murashige
(b) Skoog
(c) Styeward
(d) Morrel

(xi) BT gene containing cotton is known as:
(a) Killer cotton
(b) Egyptian cotton
(c) Hairy cotton
(d) Indian cotton

Fill in the blanks.

(xii) 'Folded Fibre Model for ultra structure of chromosome was propounded by ….
(xiii) The enzyme reverse transcriptase was isolated by……. and ….. scientists.
(xiv) Coding and Non-coding sequences respectively are called as …… and ……
(xv) The extra chromosomal DNA in bacterial cell is called as …….
(xvi) …….. is the most appropriate get which is used for development of encapsulated seeds.

Write short answers of the following.

(xvii) Write full forms of the following:
  (a) HGP
  (b) RAPD
(xviii) Define operon.
(xix) What are DNA probes
(xx) What do you mean by Chargraffs rule?

2. With suitable diagrams, describe double helical model for molecular structure of DNA, proposed by Watson and Crick.

Or

Write notes on any one [(a) or (b)], of the following:

(a) (i) Hershey and Chase experiment
  (ii) PCR technique and its applications.

Or

(b) (i) Gene concept
  (ii) DNA fingerprinting and its uses.

3. For gene regulation in prokaryotes, describe lac operon in E. coli bacterium.

Or

Write notes on:

(a) RNA polymerase (in eukaryotes)
(b) RNA splicing

4. Write an essay on plant tissue culture technique

Or

Write note on any one of the following:

(a) Protoplast culture and somatic hybridization
(b) Micropropagation and its advantages
(c) Another culture for androgenic haploid and their use in crop improvement.

5. Write notes on any one, [(a) or (b)] of the following:

(a) (i) Restriction enzyme and their role in recombinant DNA technology
(ii) cDNA library

Or

(b) (i) Plasmids and Cosmids as vectors for gene transfer
(ii) Achievements of recombinant DNA technology

Or

What are transgenic plants? With suitable examples describe herbicide resistant and insect resistant transgenic plants.

Bibliography

1. Molecular Cloning:a Laboratory Manual,J.Sambrook,E.F.Fritsch and I.Maniatis,Cold Spring harbor Laboratory Press,New York,2000.
2. Introduction to Practical Molecular Biology,P.D.Dabre,John Wiley & sojns Ltd.,Yourk,1988.
3. Molecular Biology LabFax. T.A. Brown (Ed.), bios Scientific Publishers Ltds., Oxford,1991.
4. Molecular biology of the Gene (4th Edition),J.D. Watson,N.H. Hopkins,J.W. Roberts,J.A. Steitz and A.M.
5. Molecular Cell biology (2nd Edition)J.Darnell,H.Lodish and D.Baltimore,Scientific American Books,USA,1994.
6. Molecular Biology of the Cell (2nd Edition)B.Alberts,D.Bray,J.Lewis,M.Raff,K.Roberts,and J.D. Watson, Garland publishing. Inc., New York,1994.
7. Gene VI(6th Edition)Benjamin Lewin.Oxford University Press.U.K.,1998.
8. Molecular Biology and biotechnology.A comprehensive desk reference.R.A. Meyers(Ed.) VCH Publishers,Inc.,New York,1995.
9. Genomes,T.S.Brown
10. Molecular Closing:aLaboratory Manual,J.Sambrook.E.F.Fritsch and T.Maniatis,Cold Spring Harbor Laboratory Press,New York,2000.
11. DNA Cloning:a practical Approach,D.M.Glover and B.D.Hames,IRL Press Oxford,1955.
12. Molecular and cellular Methods in Biology and Medicine,P.B. Kaufman,W.Wu.,D.Kim and L.J:Cseke,CRC Press.Florida,1995.
13. Methods in Enzymology Vol.152,Guide to Molecular Cloning Techniques,S.L.Berger and A.R.Kimmel,Academic press.Inc.San Dlogo,1998.
14. Methods in Enzymology Vol.185,GeneExpression Technology,D.V.Gooddol,Academic Press,Inc.,San Diego,1990.
15. DNA Science.A First Course in Recombinant Technology,D.A.Mickloss and G.A.Greyer,Cold Spring Harbor Laboratory Press,New York,1990
16. Molecular Biotechnology (2nd Edn.),S.B.Primorso,Blackwoll Scientific Publishers,Oxford,1994.
17. Milestones in Biotechnology.Classic papers on genetic Engineering,J.A.Davies and W.S.Roznikolf,Butterworth-Helnemann,Boston,1992.
18. Route Maps in Gene Technology,M.R.Walker and R.Repley,Blackwell Science Ltd.Oxford,1997.

19. Genetic Engineering.An Introduction to gene analysis and exploitation in eukaryotes.S.M.Kingsman and A.J.Kingsman,Blackwell Scientific Publications.Oxford,1998.
20. Molecular Biotechnology-Glick.