5 CHAPTER DNA and RNA

CLINICAL CASE

A 20-year-old male with the history of unilateral ptosis for the past 6 months and bilateral ptosis for the past 2 weeks, progressive exercise intolerance and bradycardia. Similar clinical history was found in his mother who is no more. The clinician called the genetic counsellor for help. What should be the response given by the genetic counsellor?

INTRODUCTION

Nucleotides

- DNA and RNA consist of nucleotides that form the building blocks.
- Nucleotide has three components:
  1. *Nitrogenous base:* May be purine (adenine and guanine) or pyrimidine (cytosine, thymine and uracil).
     Note: Pyrimidines are also found in thiamin.[NEXT]
  2. *Pentose sugar:* Ribose in RNA and deoxyribose in DNA.
  3. *Phosphate molecules:* May be one to three phosphate molecules.

Box 5.1: Facts about DNA.

- Xanthine and hypoxanthine are also purine bases but are not present in nucleic acids.
- Nucleoside = Nitrogen base + Pentose sugar.
- Nucleotide = Nitrogen base + Pentose sugar + Phosphate molecule.
- One kilobase (1 kb) length of nucleic acid consists of 1000 base pairs in double-stranded and 1,000 bases in single-stranded nucleic acid.[NEXT]
- New DNA synthesis occurs in interphase of cell division.[NEXT]

DEOXYRIBONUCLEIC ACID

- Deoxyribonucleic acid (DNA) is responsible for heredity and expression of characters.
- DNA is a polymer of deoxyribonucleotides.
- DNA consists of adenine, guanine, cytosine, thymine, deoxyribose sugar and phosphate molecules.

Watson and Crick Model

- DNA is a spiral twisting of two polyribonucleotide chains in the form of right-handed double helix structure **(Fig. 5.1)**.

- Both chains run in an antiparallel direction (one 3′ to 5′ and another 5′ to 3′). Base sequence in DNA molecule is always written from 5′ to 3′. NEXT
- Both the chains are held together by hydrogen bonds between nitrogen bases.
- Adenine (A) pairs with thymine (T) by two hydrogen bonds and cytosine (C) pairs with guanine (G) by three hydrogen bonds. Three hydrogen bonds between G and C provide *thermal stability* to DNA. NEXT

Box 5.2: Chargaff's rule.

- *Chargaff's rule:* For double-stranded DNA: A + G (purines) = C + T (pyrimidines). NEXT
- If the DNA contains A + G ≠ C + T, then the DNA is single-stranded. NEXT

Keys: A = adenine, T = thymine, C = cytosine, G = guanine

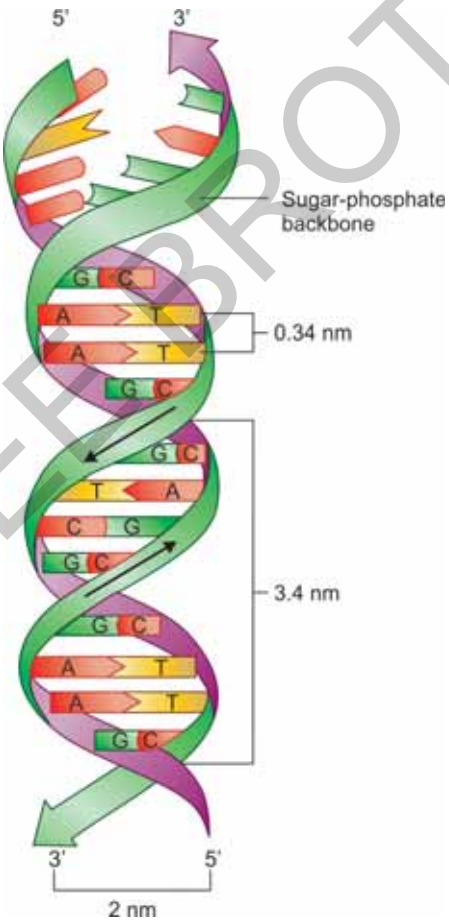


Fig. 5.1: Structure of DNA.

Keys: A = adenine, T = thymine, C = cytosine, G = guanine

Box 5.3: DNA specification.

- The length of each of helix: 3.4 nm.
- The width of helix: 2 nm.
- Each helix = 10 base pairs.
- The distance between adjacent base pairs: 3.4 nm.

RIBONUCLEIC ACID

- Ribonucleic acid (RNA) is a polymer of ribonucleotides that consists of the ribose sugar, adenine, guanine, cytosine, *uracil* and phosphate molecules.
- *There are three types of RNA:* Messenger RNA (mRNA), ribosomal RNA (rRNA) and transfer RNA (tRNA).

Transfer RNA (tRNA)

- Constitutes 20% of total RNA.
- Carrier of amino acids to ribosomes in protein synthesis.
- Each tRNA is specific for amino acid (some amino acids can be carried by more than one type of tRNA).
- An average number of nucleotides in tRNA varies between *74 to 95 nucleotides.* NEXT
- The tRNA contains unusual bases such as pseudouracil, thymidine. NEXT
- The tRNA has five arms as follows **(Fig. 5.2)**:
  1. *Acceptor (CCA) arm:* It has 3′ end with base sequence cytosine-cytosine-adenine. It binds with amino acids.
  2. *TΨC arm:* It binds tRNA to the ribosome.
  3. *Variable arm:* Useful for species' identification.
  4. *Anticodon arm:* It has *anticodons* and it binds with specific codon of mRNA during translation. NEXT
  5. *D arm:* It acts like recognition site for aminoacyl-tRNA synthetase enzyme that adds specific amino acids to tRNA.

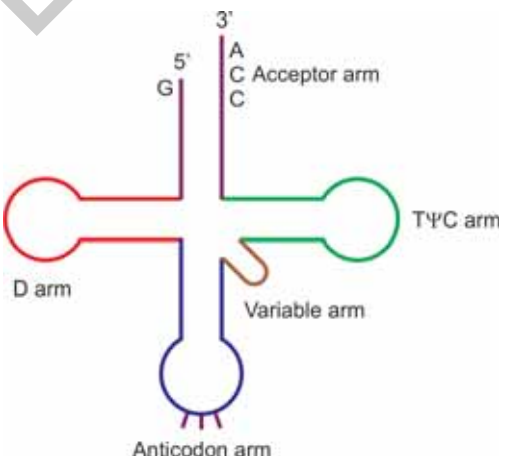


Fig. 5.2: Transfer RNA.

Keys: Ψ = pseudouracil, A = adenine, C = cytosine, G = guanine

Ribosomal RNA

- Constitutes 60–70% of total RNA.
- rRNA combines with proteins to form ribosomes.
- rRNA is produced in the *nucleolus*.NEXT
- Each ribosome has two subunits: Larger subunit (60S) and smaller subunit (40S).
- *Function:* It helps in protein synthesis by providing a site for interaction between mRNA and tRNA.
- The rRNA have peptidyl transferase activity.NEXT

Messenger RNA

- mRNA is synthesized in the nucleus from DNA by transcription. mRNA is a complimentary copy of the single-stranded DNA.NEXT
- Synthesized mRNA moves to the cytoplasm and at the ribosome, it gets translated to form protein.

Structure ***(Fig. 5.3)***

- mRNA has 7-methyl GTP cap at 5′ end (for protection from 5′ exonuclease).NEXT
- mRNA has poly-A tail of 22–250 AMP at 3′ end (for stability and protection from 3′ exonuclease action).NEXT
- mRNA has codons (sequence of *three bases*) for identification of specific anticodons of tRNA.NEXT

Some Interesting FactsNEXT

- A, U, G and C bases in mRNA form *64 triplets (codons)*.
- *61* triplets are codons as they recognize tRNA anticodon.
- Three triplets (*UAG, UGA, UAA*) are *nonsense codons* or chain *termination* codons as they cannot recognize any tRNA.
- The three stop codons have been given names: UAG is amber, UGA is opal (sometimes also called umber), and UAA is ochre.
- AUG codon is chain initiation codon.
- Methionine and tryptophan have only one codon for each of them.
- Reverse transcriptase converts single-stranded DNA into double-stranded RNA.
- Degeneracy of codon indicates more than one codon for single amino acid. Genetic code cannot be universal or overlapping.
- RNAs are less stable than DNA due to the presence of two hydroxyl groups. Hence, DNA is selected for genetic information.
- If there is an insertion in the region of the intron of the DNA, it will not be expressed and the resultant protein will be normal.
- DNA synthesis occurs only in S phase, whereas RNA and protein synthesis occurs in all phases of cell cycle.

Fig. 5.3: Messenger RNA.
(GTP, guanosine triphosphate; A, adenosine)

DIFFERENCES BETWEEN DNA AND RNA (TABLE 5.1)

Q. Write the differences between DNA and RNA.

Table 5.1: Differences between DNA and RNA.

DNA	RNA
• Double stranded	• Single stranded
• Present in nucleus or mitochondrion	• Present in the cytoplasm
• Sugar: Deoxyribose	• Sugar: Ribose
• Base pairs: A, T, G, C	• Base pairs: A, U, G, C
• Chargaff's rule follows (A = T, G = C)	• Does not follow Chargaff's rule
• Self-replicating	• Synthesized from DNA
• Alkali resistant	• Destroyed by alkali

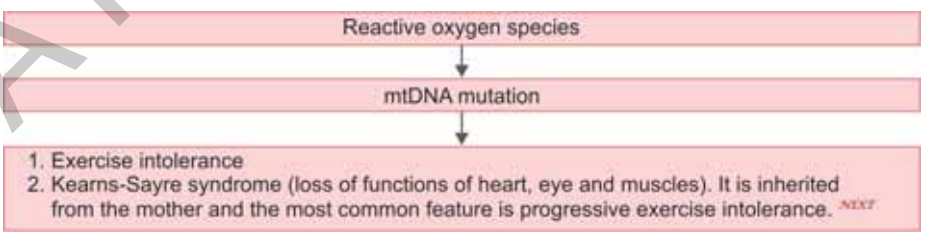
SATELLITE DNA

- It is noncoding repeat DNA sequence.
- In the human satellite, DNA is present in heterochromatin region of chromosome 1, 9, 16, long arm of chromosome Y, and satellite bodies of chromosome 13, 14, 15, 21 and 22.
- It is responsible for coding ribosomal and transfer RNAs.

MITOCHONDRIAL DNA

Q. Write a short note on mitochondrial DNA (mtDNA).

- mtDNA is located in the mitochondrion and it constitutes approximately 1% of total cellular DNA. NEXT
- All humans receive mtDNA from the mother as the only ovum provides cytoplasm and mitochondrion to the zygote (embryo). NEXT
- *Evolution:* mtDNA entered in eukaryotic cells through bacterial infection (endosymbiotic theory of evolution).
- mtDNA is circular double-stranded DNA. NEXT
- mtDNA codes for adenosine triphosphate synthase, cytochrome C oxidase, cytochrome B, NADH dehydrogenase.
- DNA polymerase γ is required for replication of mtDNA. NEXT
- Mutations:



REPLICATION OF DNA

- DNA replication is a process by that DNA copies itself to synthesize identical daughter molecules of DNA.

- Multiple origins of replication is the characteristic feature of eukaryotic cell DNA replication.
- Synthesis of DNA is a semiconservative method: Each of the strand from parental DNA acts as a template strand for DNA synthesis and each one of the daughter DNA molecule receive one of the template strand and another one is newly synthesized.
- Origin of replication: It is a particular sequence in a genome at which replication is initiated.
- Replication fork: It is a structure formed within the long helical DNA during DNA replication.
- ***Helicase*** breaks hydrogen bonds, and forms replication fork that has two DNA strands.
- DNA molecule has two strands:
  - *Leading strand*: It is strand or template of a DNA that is oriented in 3′ to 5′ direction towards the replication fork. It is replicates continuously by DNA polymerase.
  - *Lagging strand*: It runs in 5′ to 3′ direction towards the replication fork.
- *RNA primer* is a segment of RNA that initiates DNA synthesis and RNA primer is complimentary to DNA template strand.
- **DNA polymerases**
  These are enzymes involved in DNA synthesis. They are of 5 types:
  1. DNA polymerase α: Responsible for synthesis of RNA primer.
  2. DNA polymerase β: Responsible for DNA repair.
  3. DNA polymerase γ: Responsible for mitochondrial DNA replication.
  4. DNA polymerase δ: Responsible for replication on leading strand of DNA.
  5. DNA polymerase ε: Responsible for synthesis on lagging strand.
- DNA polymerase carries the individual nucleotide to the site of DNA replication. It builds the new DNA strand by matching complimentary nucleotide on the parent strand.

Stages of DNA Replication

- It occurs in three stages: Initiation, elongation, and termination **(Fig. 5.4)**.

Initiation

- Helicase enzymes separates parental strands of DNA.
- Replication protein A binds to exposed single-stranded template of DNA. These open strands form replication fork (Y-shaped structure).
- In human (eukaryotic) DNA, there are simultaneously hundreds to thousands of locations of replication forks that allow rapid DNA replication.

Elongation

- Primase enzyme forms a complex with DNA polymerase.
- This complex produces RNA primer.
- After formation of RNA by adding 20–30 nucleotides on formation of short stretch of DNA attached with RNA primer, the primer-enzyme complex dissociates.
- *Note:* DNA polymerase cannot initiate new strand synthesis.
- It can only add new nucleotides at 3′ end of existing DNA strand.
- Binding of replication factor C (RFC): The RFC binds the elongated primer and act as clamp loader. It also help in assembly of proliferating cell nuclear antigen molecules.
- As DNA polymerase can synthesize new strands only in 5′ to 3′ directions, two newly-synthesized strands grow in opposite directions. Leading strand is synthesized continuously. Lagging strand is synthesized in pieces called Okazaki fragments. Each Okazaki fragment has its own RNA primer.

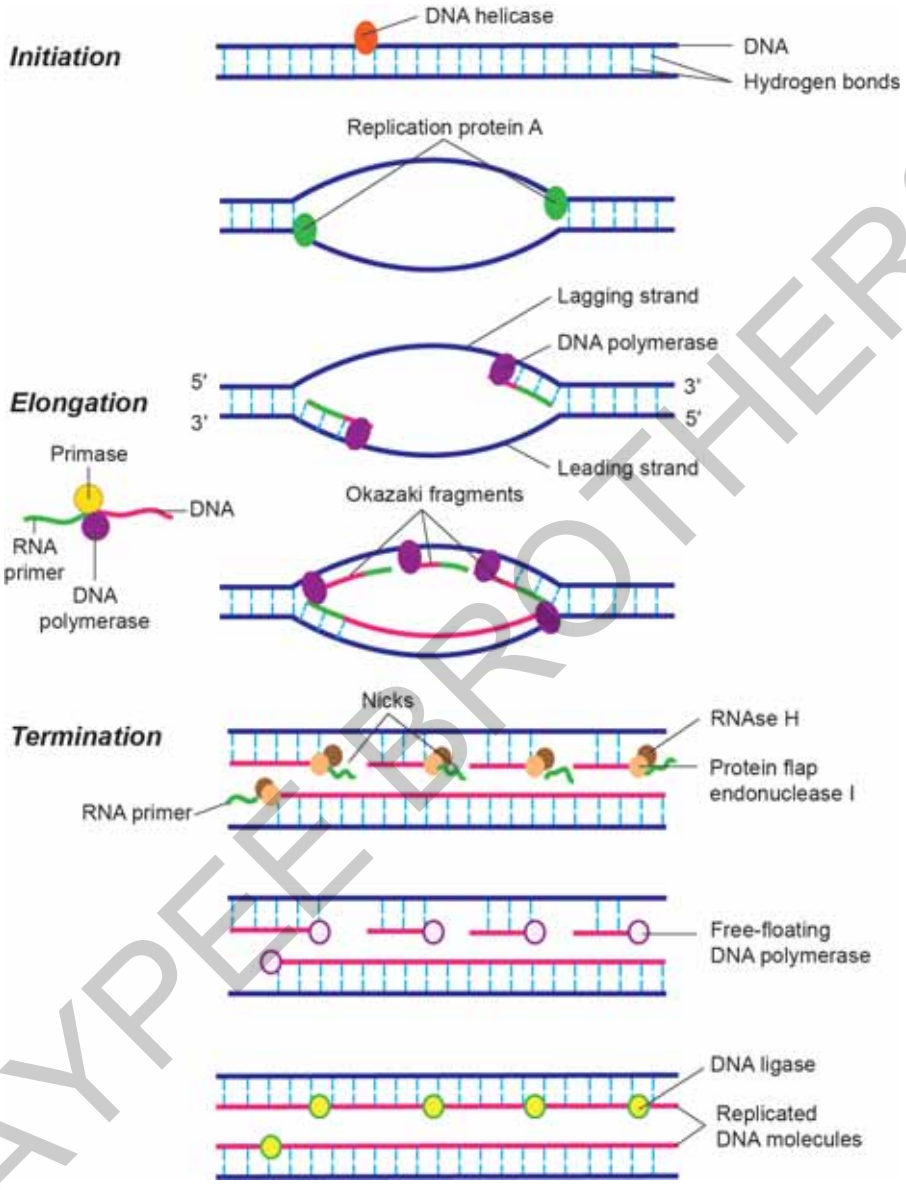


Fig. 5.4: DNA replication.

Termination

- *Replication bubbles*: Human (eukaryotic) DNA replication begins at multiple sites simultaneously. Each site produces bubble of duplicated DNA.
- After some time replication of adjacent bubbles reach to each other and DNA polymerase stops functioning leaving behind *nicks* between newly synthesized DNA fragments.
- Protein flap endonuclease 1 and RNAse H remove RNA primers leaving behind gaps of unreplicated template DNA.

- Free-floating DNA polymerase extends the DNA strands over these gaps.
- DNA ligase joins the nicks by forming sugar-phosphate back bone and thus, DNA replication completes.

Some Interesting Facts

- Anticancer drugs such as adriamycin, etoposide, doxorubicin inhibit human topoisomerase whereas, 6-mercaptopurin, 5-fluorouracil act as nucleotide analog and thus, inhibits DNA synthesis.
- DNA replication occurs in S-phase of cell cycle and requires 8–10 hours.
- Telomere prevent loss of DNA at the end of chromosome during DNA replication, specifically on removal of RNA primer. Thus, telomeres prevents fusion of ends of chromosomes.
- Telomeres contain repeated TTAGGG sequences.
- Telomeres are maintained by telomere terminal transferase (telomerase).

Box 5.4: DNA repair.

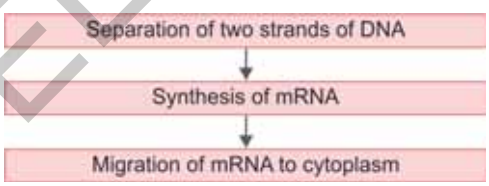
- Cell has the capacity to repair damaged DNA.
- DNA may get damaged by improper DNA replication, radiation, free radicals, and chemicals. The DNA damage is called mutation.
- DNA damage includes single base alteration, two base alteration, chain breaks, and cross-linkages.
- DNA repair is achieved by the following mechanisms:
  - Base excision-repair: It involves removal of base by glycosylase.
  - Nucleotide excision repair: It involves removal of DNA fragment and replacement.
  - Mismatch repair: It involves removal of strand with exonuclease and replacement.
  - Double strand break repair: It involves unwinding, alignment and ligation.
- DNA polymerase also perform proofreading function and may initiate DNA repair.[NEXT]

TRANSCRIPTION AND TRANSLATION

Transcription

Process of mRNA synthesis from DNA.

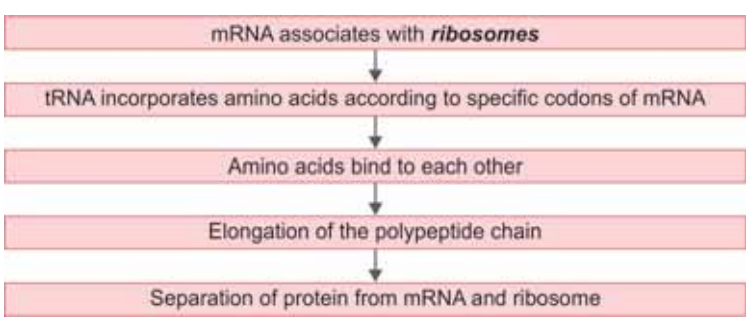
Process



Translation

The process of protein synthesis from mRNA.

Process



Box 5.5: Facts about translation.

- Ribozymes are RNA fragment with catalytic activity.NEXT
- Kozak sequence occurs in eukaryotic DNA and plays a role in translation.NEXT
- Aminoacyl-tRNA synthetase is fidelity enzyme in protein synthesis. It is required for the attachment of specific amino acids to the corresponding tRNA.NEXT

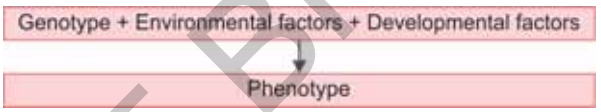
GENE

Q. Write a short note on the gene.

- Genes are defined as working subunits of DNA that are composed of a specific sequence of codons and may get expressed in the phenotypic form (protein).
- Gene is a region of DNA that encodes specific function.
- Gene = Exons + Introns **(Fig. 5.5)**
- *Exon:* Part of a gene that encodes into matured mRNA.
- *Intron:* Part of a gene that is removed by RNA splicing during mRNA maturation.
- Human chromosomes have 30,000 genes.
- Genes are either structural genes (produces proteins) or regulatory genes (promote or inhibit activity of other genes).

Genotype

Genes made up of different DNA sequences (total genes of a cell) are called genotype.



Flanking Region

- It is a region of DNA that contains promoter gene and a termination codon.
- It is essential for the beginning of mRNA synthesis, but it does not get transcribed into mRNA.
- It is also essential for the termination of protein synthesis.
- At 5′ end, flanking possess promoter zone with TATA box and CAT box.
- TATA box is called as Goldberg-Hogness box. It is the region that helps in transcription. It has 5′TATA.....TAA 3′ DNA sequence.
- CAT box is also called as CCAAT box. It also helps in transcription.
- At 3′ end, the flanking region possesses termination codon (TAA).

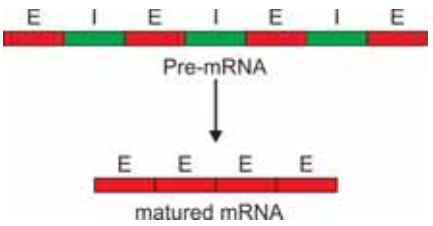


Fig. 5.5: Introns and exons.

Transcription

- It is the process of mRNA synthesis.
- Actinomycin D inhibits transcription. NEXT
- *Steps* ***(Fig. 5.6)****:*

Binding of RNA polymerase to promoter region

↓

Separation of two strands of DNA
(Coding strand of DNA runs in 3′–5′ direction, match with the RNA transcript that encodes proteins) NEXT

↓

Beginning of mRNA formation by transcription factor in 5′–3′ direction with the addition of proper nucleotides

↓

Release of RNA polymerase and immature mRNA

↓

Splicing (removal of introns) to form mature mRNA with the help of small nuclear RNA (snRNAs) NEXT

- *Poly-A tail:* It is present at 3′ end of mRNA degradation and prevents mRNA degradation.
- *Template strand:* It is a DNA strand (3′ to 5′) that forms mRNA.
- *Transcription factor:* Protein that attracts RNA polymerase to the promoter region.
- The intron is the segment of the gene that is not represented in the matured mRNA. NEXT
- Cistron is the smallest fundamental unit coding for the DNA synthesis. NEXT

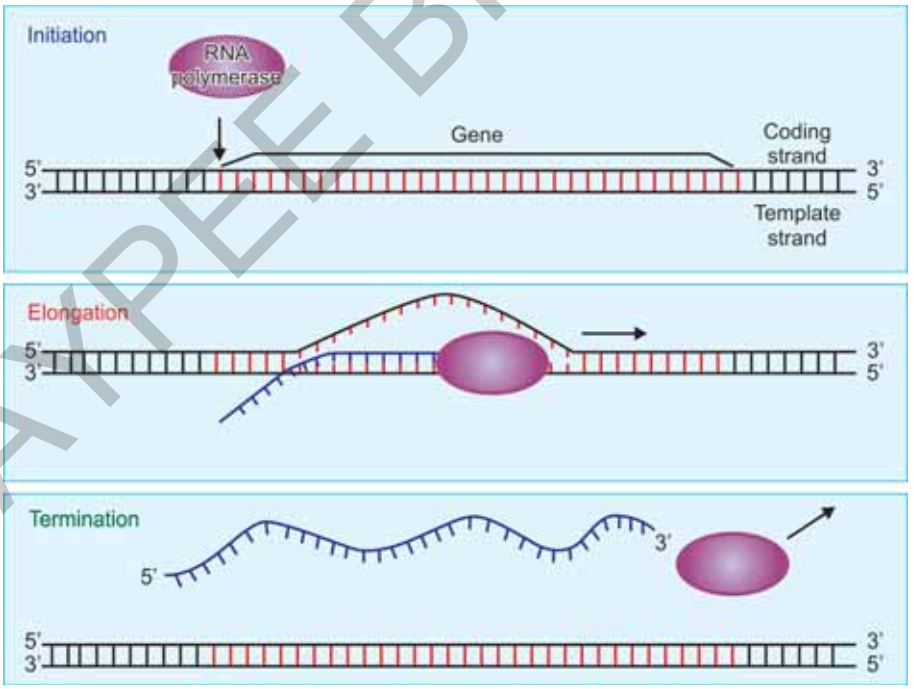


Fig. 5.6: Transcription.

- Gene = Intron + Exon; and Cistron = all exons of a single gene.
- Replication and transcription are the similar processes as these involve the formation of the phosphodiester bonds in elongation of the chain in 5′–3′ direction.[NEXT]

Processing of mRNA (Post-transcriptional Modifications)[NEXT]

- *mRNA capping:* Methylation of 5′ end (addition of methyl group)—act as an initiation site for translation. It helps in the attachment of mRNA to the ribosome.
- The addition of poly-A tail for stabilization of mRNA occurs in the cytoplasm.
- *Trimming:* Splicing out interns and leaving only exons.

Box 5.6: Initiation factors and alternate splicing.

Initiation factors
In eukaryotes, initiation factor is regulated by GTP to GDP transformation.[NEXT]

Alternate splicing
One gene can synthesise more than one protein due to alternate splicing **(Fig. 5.7)**.

RNA Polymerase[NEXT]

- It is DNA-dependent RNA polymerase.
- It produces primary transcript RNA from DNA.
- It has alpha (for initiation of transcription), beta (for initiation of elongation), beta 1 (for non-specific DNA binding) and sigma (for promoter binding) subunits.

Translation

- It is the process of protein synthesis from mRNA.
- *Process:*
  - *Initiation:*
    - The 5′ end of mRNA and initiator tRNA attaches to the small subunit of the ribosome.
    - The start codon is AUG (anticodon on tRNA is UAC and carry methionine).[NEXT]
    - Large ribosomal subunit attaches to small subunit.
  - *Elongation:*
    - Small ribosomal subunit further moves one codon on mRNA and new tRNA binds with mRNA.
    - Peptidyl transferase forms a peptide bond between new and previous amino acid and soon large subunit also moves towards small one.

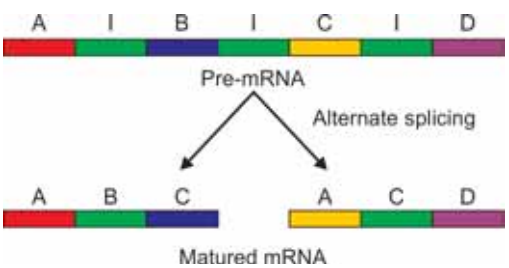


Fig. 5.7: Alternate splicing.

  - *Termination:*
    - The process of elongation continues till small unit reaches to the termination codon.
    - Polypeptide chain gets detached from ribosome and forms folding to develop tertiary protein structure.
- *Post-translational modifications includes*:NEXT
  - Trimming
  - Glycosylation
  - Covalent alteration
  - Phosphorylation
  - Hydroxylation.

Factors Regulate Gene Expression

- *Operator gene:* Controls the expression of structural genes that lies adjacent to operator genes on some chromosomes.
  Operon = Operator gene + Structural genes.
- *Regulator genes:* Controls operator genes through repressor genes.
- *Transcription factors:* Controls binding of RNA polymerase at promoter genes. Transcription factor may be enhancer (increases transcription) or silencer (decreases transcription).
- For lac operon model, lactose, allolactose, isopropyl thiogalactoside, low glucose and CAP-cAMP are positive regulators or inducers. *LacI* gene and high glucose concentration are negative regulators or repressors.NEXT

GENE MAPPING

Q. Write a short note on gene mapping.

Gene mapping is the creation of a genetic map, taking DNA fragments and assigning them to different chromosomes.

Definition

It is a method used to identify the locus of the gene and distances between genes.

Gene mapping is divided into two groups:

1. *Genetic mapping:* Uses the technique to construct the positions of genes, for example, pedigree charting.
2. *Physical mapping:* Uses the molecular biology techniques to examine DNA molecules directly.

For gene mapping, genetic markers are required.

Genetic Marker

- It is a gene or DNA sequence with a known location on a chromosome.
- It is useful to identify the presence of a gene and mutation in the gene to establish a relationship between the disease and genetic mutation.

Types of Genetic Markers

- Restriction fragment length polymorphism (RFLP)
- Amplified fragment length polymorphism (AFLP)
- Single nucleotide polymorphism (SNP) and so on.

Physical Mapping

It utilizes genetic markers for one of the following methods:

1. *Restriction mapping:* It locates the position of recognisable sequences for restriction endonucleases on a DNA molecule.
2. *Fluorescent in situ hybridization (FISH):* It locates marker on a chromosome by direct hybridization of the probe to DNA.
3. *Sequence-tagged site (STS) mapping:* It maps position of the short sequence by polymerase chain reaction (PCR).

Uses of Gene Mapping

- Identification of the gene responsible for heritable diseases or cancer.
- To establish a relationship between gene and the disease.
- To establish a relationship between the variation of the gene and the disease, drug resistance or phenotypic changes.

Box 5.7: Human Genome Project.

Q. Write a short note on Human Genome Project.

- It was an international scientific research project to locate the genes in the human genome and explore the details of genes.
- The human genome is a complete set of genetic information for humans.
- 1986: Human Genome Project started.
- 2003: Human Genome Project (sequencing) completed.

Advantages of Human Genome Project

- Ability to locate genes that are responsible for locating diseases.
- Can be used for gene therapy.

Outcomes

- Human genome contains 22 autosomes and X- and Y-chromosomes.
- 6 feet DNA
- 30,000 genes
- 3 billion nucleotide pairs
- Average gene contain 3,000 bases.
- Only 3% genome encodes for proteins and rest of it junks DNA.
- Chromosome 1 has highest genes (2,968) and chromosome Y has the lowest (231).
- Mostly DNA of the individual differs from each other by SNPs.
- The approximate size of the diploid human genome is 3×10^9 base pairs (3 billion bp). NEXT

GENE BANK

It is the collection of DNA molecules that possess complete genetic information of an organism.

Indian National Gene Bank

- National Bureau of Plant Genetic Resources (NBPGR) maintains Indian National Gene Bank.
- It preserves dehydrated plant seeds, tissue cultures, synthetic seeds, germplasms and so on.

Some Interesting Facts

- During DNA replication (copying), most DNA polymerases (mostly polymerase II) can check their work (proofreading). Polymerases remove wrong (incorrectly paired) nucleotide and replace the nucleotide right away before continuing with DNA synthesis.
- Ultraviolet radiation induces the formation of cyclobutane pyrimidine-pyrimidine dimers.
- Xeroderma pigmentosa is the example of *pyrimidine dimer* formation and *mismatch* repair or low activity of excision repair process.
- Leucine zipper motif is a mediator for binding of the regulatory protein to DNA.
- Jacob and Monod elucidated an operon model. An operon is a functioning unit of genomic DNA containing a cluster of genes under the control of a single promoter.
- Zinc finger is a nuclear receptor.

SOLUTION FOR CLINICAL CASE

Genetic counsellor should explain that this is a suspected case of Kearns-Sayre syndrome that shows mtDNA inheritance. As an individual receives mtDNA from the mother, this disease can be suspected in the mother also. Its diagnosis mostly depends on clinical analysis, changes on muscle biopsy (ragged-red fibers with the modified Gomori trichrome stain). Immunohistochemical staining shows increased succinate dehydrogenase and absence of cytochrome C oxidase. Deletion of mtDNA is present in 90% of the cases that can be detected using PCR and RFLP.