DAY THIRTY

Molecular Basis of Inheritance

Learning & Revision for the Day

- DNA as a Genetic Material
- Structure of DNA
- Packaging of DNA
- DNA Replication
- RNA
- Central Dogma
- Transcription
- Genetic Code
- One Gene-One Enzyme Hypothesis
- Translation
- Gene Expression
- Regulation of Gene Expression
- Human Genome Project
- DNA Fingerprinting

Genetic material is the substance which controls the inheritance of traits from one generation to next. As living matter is made up of biochemicals, therefore, genetic material should be biochemical in nature.

Following are the criteria that a molecule must fulfil to act as a genetic material.

- It should be able to produce its replica.
- It should be chemically and structurally stable.
- It should provide the scope for slow changes that are necessary for evolution.
- It should be able to express itself in the form of Mendelian characters.

DNA as a Genetic Material

The various indirect and direct evidences in support of DNA as a genetic material are as follows

(i) **Friedrich Miescher** in 1869, explained that nucleus possesses deoxyribose nucleic acid (DNA) that has capacity to replicate to produce similar copies of it. It carries genetic information from one generation to next.

(ii) **Fredcrick Griffith** (1928) conducted a series of experiments with *Streptococcus pneumoniae,* the bacterium causing pneumonia.

- He observed two strains of this bacterium, i.e. one forming smooth colonies with capsule (S-type) and the other forming rough colonies without capsule (R-type).
- He concluded that the R-strain bacteria had somehow been transformed by the heat-killed S-strain bacteria, which must be due to the transfer of the genetic material (transforming principle).

(iii) The unequivocal proof that DNA is the genetic material came from the experiments of **Alfred Hershey** and **Martha Chase** (1952). They worked with viruses that infect bacteria called **bacteriophages**.

They made two different preparations of the phage. In one, the DNA was made radioactive with ^{32}P and in the other, the protein coat was made radioactive with ^{35}S.

The phage containing radioactive DNA was radioactive that was found in the bacterial cells indicating that the DNA is the genetic material.

Structure of DNA

DNA acts as the genetic material in most of the organisms. It is a long polymer of deoxyribonucleotides. Each nucleotide has three components, i.e. a nitrogenous base, a pentose sugar (deoxyribose) and a phosphate group.

Double helix model of DNA was proposed by **Watson** and **Francis Crick** (1953) based on X-ray diffraction data, produced by **Maurice Wilkins** and **Rosalind Franklin.**

The important features of this model are given below

- It contains two polynucleotide chains. Its backbone constitutes sugar-phosphate and the bases project inside.
- There are two types of nitrogen bases, i.e. purines (adenine and guanine) and pyrimidines (cytosine, uracil and thymine).
- The bases in the two strands are held together by hydrogen bonds forming base pairs.
- Adenine pairs with thymine through two hydrogen bonds and guanine with cytosine through three hydrogen bonds.
- **Erwin Chargaff** proposed base equivalence rule or **Chargaff's rule** for a double-stranded DNA. It is not applicable to RNA.
- It states that, the proportion of adenine is always equal to that of thymine and the proportion of guanine is always equal to that of cytosine, i.e.

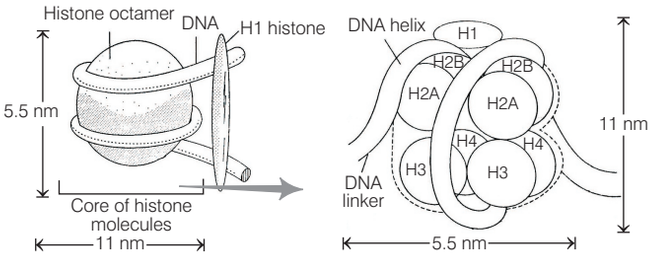
$$A + G = T + C \text{ or } \frac{A+G}{T+C} = 1$$

- The two chains have an anti-parallel polarity, i.e. one chain has a $5' \rightarrow 3'$ polarity, while the other has $3' \rightarrow 5'$ polarity.
- The two chains are coiled in a right-handed fashion and the pitch of the helix is 3.4 nm.
- There are about 10 base pairs in each turn with 0.34 nm distance between two base pairs.
- The plane of one base pair stacks over the other in the double helix.

NOTE
- Out of the pyrimidines, cytosine is common for both DNA and RNA, while thymine is present in DNA and uracil is present in RNA.
- The diameter of DNA is 2 nm.

Packaging of DNA

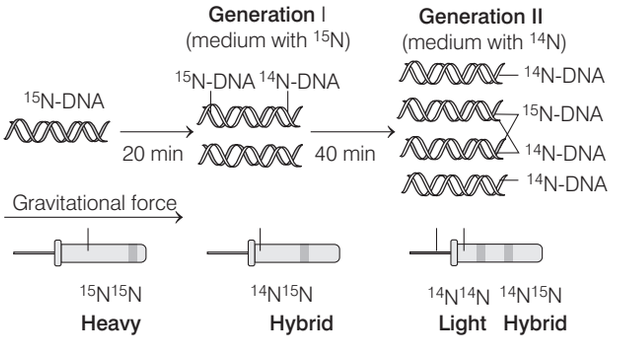
- **Prokaryotes**, such as *E. coli,* do not have a well-defined nucleus and the DNA is scattered throughout the cytoplasm.
- DNA (being negatively charged) is held with some basic proteins (that have positive charges) in a cytoplasmic region termed as nucleoid.
- In **eukaryotes,** there is a set of positively charged, basic proteins called **histones** found in nucleus.
- Histones are rich in the basic amino acid residues lysine and arginine. Both the amino acid residues carry positive charges in their side chains. Histones are organised to form a unit of eight molecules called as **histone octamer** (H2A, H2B, H3, H4).
- The negatively charged DNA is wrapped around the positively charged histone octamer to form a structure called **nucleosome** which contains 200 bp of DNA helix.
- The DNA connecting the two adjacent nucleosomes is called **linker DNA** and it bears **H1 histone** protein.
- Nucleosome constitutes the repeating unit of a structure in nucleus called **chromatin**.
- The nucleosomes in chromatin are seen as 'beads-on-string' when viewed under electron microscope.
- The chromatin fibres condense at metaphase stage of cell division to form chromosomes.
- The packaging of chromatin at higher level requires additional set of proteins called **Non-Histone Chromosomal** (NHC) proteins.
- In a nucleus, certain regions of the chromatin are loosely packed and they stain lighter than the other regions, these are called **euchromatin** (transcriptionally more active).
- The other regions are tightly packed and they stain darker and are called **heterochromatin** (transcriptionally less active).



Structure of nucleosome

DNA Replication

- In this process, the two strands of the parent DNA molecule separate and complementary strands are synthesised using old strands as template, thus each progeny strand has one old and one new strand.
- It is a semiconservative process, because out of two progeny strands, present in DNA helix, one is the old strand and other is the new one.
- **Meselson** and **Stahl** (1958) provided the evidence in support of semiconservative replication of DNA.
- They grew *E. coli* on ^{15}N (i.e. a heavy isotope of ^{14}N) and observed that 50% of entire DNA in generation-I possess ^{15}N while remaining 50% have only ^{14}N.
- This showed that one parent strand was conserved in daughter while the second strand was freshly prepared.

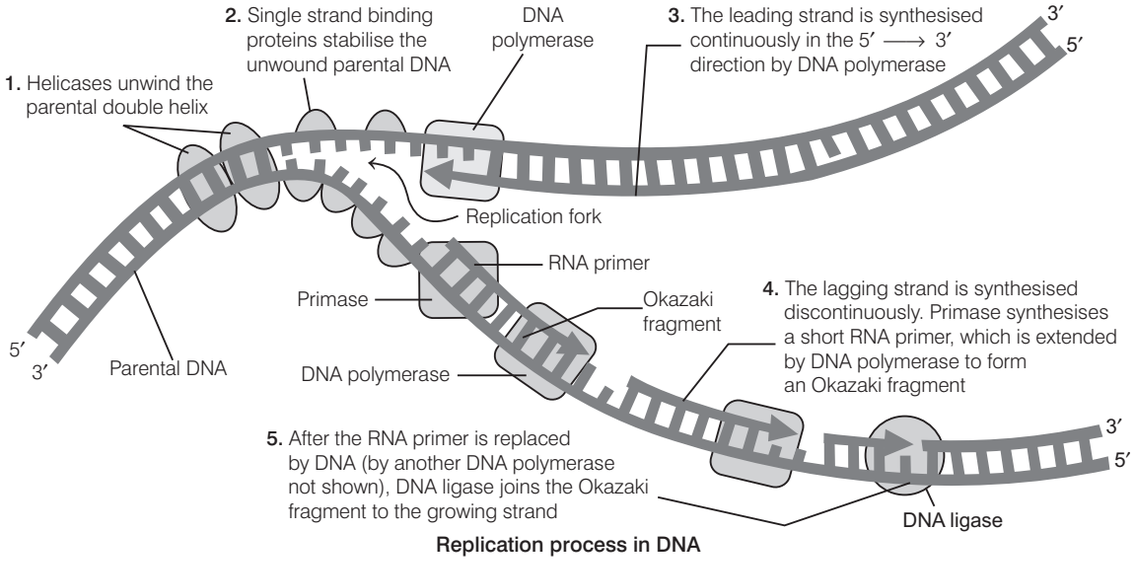


(Separation of DNA by centrifugation)

Meselson and Stahl's experiment

The process of DNA replication requires the following components

- DNA template, a primer (usually RNA), Deoxyribonucleotide triphosphates (i.e. *d* ATP, *d* GTP, *d* TTP, *d* CTP), Mg^{2+}, DNA unwinding proteins (helicase and gyrase), RNA polymerase to synthesise the RNA primer, polynucleotide ligase (a joining enzyme) and DNA polymerase (a DNA synthesising enzyme).
- DNA replication is bidirectional.
- A replicon is the segment of DNA, which is capable to undergo replication, independent of the other segments of DNA.
- Each replicon has an **origin of replication**, at which DNA replication starts and a terminus, at which replication stops.
- The two strands of DNA double helix unwind with the help of DNA unwinding proteins (e.g. DNA gyrase and DNA helicase).
- Unwinding produces two Y-shaped forks at origin, one fork is located at each end of the origin. When replication begins, these forks become **replication forks**.
- The unwinding of the strands is brought out by the action of a superhelix relaxing protein, whereas the, Single Strand Binding (SSB) proteins stabilise the unwound parental DNA.
- Initiation of DNA synthesis requires an RNA primer which is synthesised on DNA template by an enzyme primase.
- The free 3′—OH end of RNA primer RNA provides the initiation point for DNA polymerase, which in turn requires a free 3′—OH end of a pre-existing polynucleotide for the initiation of DNA replication.
- Replication of 3′ → 5′ strand of a DNA molecule proceeds continuously therefore, the 3′→5′ strand of DNA molecule is known as leading strand. The replication of 5′ → 3′ strand of the DNA molecule is discontinuous and it is known as lagging strand.
- Replication of lagging strand generates small nucleotide fragments in 5′ → 3′ direction called **Okazaki fragments**.



Replication process in DNA

- Different types of DNA polymerases in prokaryotes and their respective functions are tabulated below

Activities	DNA Polymerase-I	DNA Polymerase-II	DNA Polymerase-III
5′→ 3′ polymerase	Present	Present	Present
3′→ 5′ exonuclease	Present	Present	Present
5′→ 3′ exonuclease	Present	Absent	Absent
Functions	DNA repair, excision of RNA primers and fills gaps	DNA repair	DNA replication. It is the real replicase.

- The different types of DNA polymerases in eukaryotes and their functions are as follows
  (i) DNA polymerase- α replicates the leading strand, while DNA polymerase- δ synthesises the lagging strand.
  (ii) DNA polymerase- β and ε are nuclear DNA repair enzymes, whereas DNA polymerase-γ is found in **mitochondria** and **chloroplasts**.

RNA

- RNA is known to be genetic material in some viruses, e.g. retroviruses.
- The essential life processes such as metabolism, translation, splicing, etc., have evolved around RNA, even before DNA, RNA was evolved as a genetic material.
- RNA acts as a genetic material as well as a catalyst.
- The 2′ —OH group of ribonucleotides is a reactive group, that makes RNA to act as a catalyst.
- But, RNA being a catalyst is reactive and hence, unstable.
- Therefore, DNA has evolved from RNA with chemical modifications that made it more stable.
- In RNA, each molecule has three components as in DNA.
- The sugar is ribose, which has an additional OH group on the 2′-position.
- The nucleosides and nucleotides of RNA are called ribonucleosides and ribonucleotides, respectively.

Types of RNA

There are three types of RNA *viz*., *m*RNA, *r*RNA and *t*RNA.

(i) **Messenger RNA** (*m*RNA) or nuclear RNA is a polymer of ribonucleotide which is always single-stranded. It carries genetic information in cytoplasm for the synthesis of proteins and acts as template.

(ii) **Ribosomal RNA** (*r*RNA) occurs in ribosomes which are nucleoprotein molecules.

It plays structural and catalytic role during translation.

(iii) **Transfer RNA** (*t*RNA) or soluble RNA or adapter RNA are smallest molecules containing 75 to 80 nucleotides.

- Most accepted model for *t*RNA structure is clover leaf model which was given by **Robert Holley**.
- This structure contains five arms each with a loop and a stem.
  - **Acceptor stem** is without loop, contains 7 base pairs and 4 unpaired bases.
  - **D arm** contains 3-4 base pairs and 7-11 nitrogenous bases, which are unpaired.
  - **Anticodon arm** contains 5 base pairs and 7 unpaired nitrogenous bases.
  - **Variable arm** may or may not contain stem.
  - **T ψ C arm** contains 5 base pairs and 7 unpaired nitrogenous bases.

Other types of RNA are

(i) **Small nuclear RNA** (s*n*RNA) is small sized RNA found in the nucleus. They take part in splicing, *r*RNA processing and *m*RNA processing.

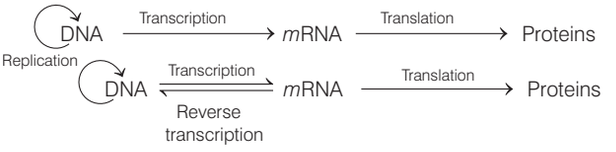
(ii) **Small cytoplasmic RNA** (s*c*RNA) is small sized RNA present freely in the cytoplasm. It helps in binding of ribosomes to ER for producing secretory proteins.

(iii) **Heterogenous RNA** (*hn*RNA) is present in nucleus. It has large molecules and acts as precursor of *m*RNA.

(iv) Ribozymes are RNA molecules having enzymatic or catalytic activity.

Central Dogma

- It is the flow of information from DNA to *m*RNA (transcription) and then decoding the information present in *m*RNA in the formation of polypeptide chain or protein (translation).
- The concept of central dogma was forwarded by **Crick** in 1958 that proposed unidirectional flow of information from DNA to RNA and then to protein.
- **Temin** (1970) and **Baltimore** (1970) reported that double-stranded RNA of Rous Sarcoma Virus (RSV) operates on a central dogma reverse (inverse flow of information).
- RNA of these viruses first synthesises DNA through reverse transcription or teminism.



Diagrammatic representation of the concepts about the flow of transcriptional information from DNA

Transcription

- The process of copying genetic information from one strand of the DNA into RNA is termed as transcription.
- The principle of complementarity governs the process of transcription as well, (except of adenine that pairs with uracil instead of thymine).
- However, in the process of replication the total DNA of an organism gets duplicated, while in transcription only a specific segment of one of the strands of DNA is copied into RNA.

A transcription unit is defined by three regions in DNA

(i) **Promoter** is the specific sequence of 20-200 bases, where the RNA polymerase binds to DNA.

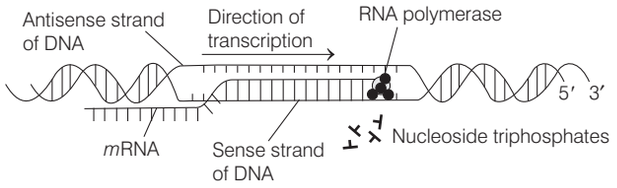
- In prokaryotes, this site contains **Pribnow box** which is a consensus sequence, TATAATG that orients RNA polymerase so that synthesis proceeds from left to right.
- In eukaryotes, the corresponding consensus sequence is TATA AAA or TATA box or **Hogness box**.

(ii) **Structural gene** is a component of DNA strand with 3′ → 5' polarity. This strand of DNA is called template strand or antisense strand. The other strand with polarity 5' → 3' is called sense or coding strand because genetic code present in this strand is similar to genetic code of *m*RNA.

(iii) **Terminator** is the region present downstream to structural gene at 3' end of coding strand. It defines the end of transcription process.

Steps of Transcription

- The segment of DNA involved in transcription is called **cistron.**
- **RNA polymerase** binds to a specific sequence of bases on the DNA of the gene to be expressed called promoter and initiates transcription by associating transiently with initiation factor (σ).
- In *E. coli*, the enzyme RNA polymerase consists of five different polypeptide chains α, α′ β, β′ and σ.
- The holoenzyme (with five units) has a molecular weight of 450000 Da.
- The enzyme without σ (sigma) subunit is called core enzyme.
- The DNA begins to unwind and the strands begin to separate.
- The RNA polymerase begins to move along one strand of the exposed DNA (the sense strand), thus linking ribonucleotides together in the order specified by the sequence of bases on DNA.
- Transcription continues until the RNA polymerase reaches a 'stop' message on the DNA (a terminator).
- Here, RNA polymerase associates transiently with the termination factor (ρ) to terminate the transcription.
- The unzipped DNA closes back, the enzymes drop off and the messenger RNA is released into the nucleus prior to the next stage in the production of a protein.



Schematic representation of sense and antisense strand during transcription

- In **eukaryotes**, there are two additional complexities

(i) There are at least three RNA polymerases in the nucleus (in addition to the RNA polymerase found in the organelles) which show clear cut division of labour. The RNA polymerase-I transcribes *r*RNAs (28S, 18S and 5.8S).

- The RNA polymerase-III is responsible for transcription of *t*RNA, 5*sr*RNA and *sn*RNAs (small nuclear RNAs).
- The RNA polymerase-II transcribes precursor of *m*RNA, the heterogeneous nuclear RNA (*hn*RNA).

(ii) The second complexity is that the primary transcripts contain both the exons and the introns.

Hence, it is subjected to a process called **splicing**, where the introns are removed and exons are joined in a defined order. *hn*RNA undergoes additional processing called as **capping** and **tailing**.

NOTE

- **Spliceosome** is large molecular complex of *sn*RNA and proteins in the nucleus of eukaryotic cells. It helps in the splicing of introns.
- **In capping**, an unusual nucleotide (methyl guanosine triphosphate) is added to the 5′-end of *hn*RNA.
- **In tailing**, adenylate residues (200-300) are added at 3′-end of *hn*RNA in an independent manner.
- It is the fully processed *hn*RNA, that is called *m*RNA, it is transported out of the nucleus for translation.

Differences between Prokaryotic and Eukaryotic Transcription

Prokaryotic Transcription	Eukaryotic Transcription
Occurs in the cytoplasm.	Occurs in the nucleus.
There is no definite phase for its occurrence.	Takes place in the G_1 and G_2 phases of cell cycle.
A single RNA polymerase synthesises all the three types of RNA, i.e. *m*RNA, *t*RNA, *r*RNA.	Three RNA polymerases I, II and III synthesise *r*RNA, *m*RNA and *t*RNA, respectively.
Coupled transcription/RNA molecules are released and processed in the cytoplasm.	Coupled transcription/RNA molecules are released and processed in the nucleus.
RNA polymerase is a complex of 5 polypeptides.	RNA polymerases are complexes of 10-15 polypeptides.
Transcriptional unit has one or more genes.	Transcriptional unit has only one gene.
The *m*RNA primary transcript has fewer surplus nucleotides.	The *m*RNA primary transcript has a large number of surplus nucleotides.

Genetic Code

- The triplet sequence of nucleotides on *m*RNA, which stores information for linking amino acid in a definite sequence during protein synthesis is called genetic code.
- The term 'genetic code' was given by **George Gamow**.
- The first clue to codon assignment was given by **MW Nirenberg** and **JH Mathaei**.
- Later on, **HG Khorana** deciphered genetic code by developing a new technique.
- **FHC Crick** (1965) proposed **Wobble hypothesis** according to which, the third base of a codon is not very important and the specificity of a codon is particularly determined by the first two bases.

- The third base pair (Wobble) is loosely paired with the corresponding base in anticodon of *t*RNA.
- There are 64 codons, out of which AUG (rarely GUG) is called initiation codon.
- UAA, UGA, UAG are called **termination** or **non-sense codons** because these codons do not code for any amino acids, so they stop the growth of polypeptide chain.
- The genetic code is triplet (AAG, GUU, etc), degenerate, non-overlapping, commaless, non-ambiguous and universal.

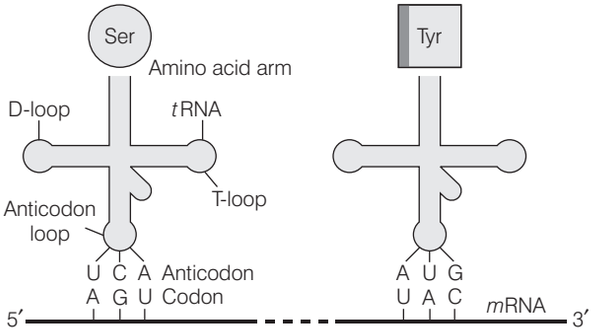
First Letter	Second Letter U	Second Letter C	Second Letter A	Second Letter G	Third Letter
U	UUU, UUC: Phe; UUA, UUG: Leu	UCU, UCC, UCA, UCG: Ser	UAU, UAC: Tyr; UAA: Stop; UAG: Stop	UGU, UGC: Cys; UGA: Stop Codon; UGG: Trp	U C A G
C	CUU, CUC, CUA, CUG: Lue	CCU, CCC, CCA, CCG: Pro	CAU, CAC: His; CAA, CAG: Gln	CGU, CGC, CGA, CGG: Arg	U C A G
A	AUU, AUC, AUA: Ile; AUG: Met	ACU, ACC, ACA, ACG: Thr	AAU, AAC: Asn; AAA, AAG: Lys	AGU, AGC: Ser; AGA, AGG: Arg	U C A G
G	GUU, GUC, GUA, GUG: Val	GCU, GCC, GCA, GCG: Ala	GAU, GAC: Asp; GAA, GAG: Glu	GGU, GGC, GGA, GGG: Gly	U C A G

One Gene-One Enzyme Hypothesis

- It was proposed by **Beadle** and **Tatum** (1948), according to which, a particular gene controls the synthesis of specific enzyme.
- Beadle and Tatum conducted experiment on pink bread mould (*Neurospora crassa*) and stated that each gene has the information to produce one enzyme. This hypothesis was discarded because some regions of DNA produce *t*RNA or *r*RNA rather than *m*RNA (which produces a protein).

Translation

- The decoding of the message of *m*RNA in the form of protein is purely unidirectional process called **translation.**
- It takes place in ribosomes.
- In this process, *t*RNA acts as adapter molecule because it has an anticodon loop that has bases complementary to the code and it also has an amino acid acceptor end to which it binds to amino acids.
- *t*RNAs are specific for each amino acid. For initiation, there is another specific *t*RNA that is referred to as initiator *t*RNA. There are no *t*RNAs for stop codons.



*t*RNA—The adapter molecule

The process of translation occurs in the following three steps

1. **During initiation,** *m*RNA binds to small unit of ribosome in the presence of initiation factor.
   - In prokaryotes, several ribosomes may attach to single *m*RNA to form a chain called **polyribosomes** or **polysomes** which help to produce number of copies of same polypeptide.
   - Protein synthesis mostly starts with AUG codon on *m*RNA, which codes for methionine (formlylated methionine in case of prokaryotes).
   - The *t*RNA first binds to P-site of ribosome and then the large subunit of ribosome gets attached to the small subunit forming a complex.
   - The next codon on the *m*RNA determines the next *t*RNA anticodon, which binds to it and to the next amino acid in the polypeptide.
2. **During elongation**, *t*RNA forms hydrogen bonds with the second codon on *m*RNA.
   - This *t*RNA comes on site 'A' of ribosome and peptide bond is formed between the first amino acid and second amino acid.
   - The enzyme peptidyl transferase catalyses the reaction.
   - The *t*RNA detaches from the ribosome and return to cytoplasm to pick up another amino acid.
   - Now the second amino acid shifts to P-site and the process of shifting of amino acid from A-site to P-site is called **translocation**.
3. **During termination**, codon UAA, UAG or UGA comes to A-site and hence, the polypeptide synthesis stops. The process of initiation, elongation and termination involve the mediation of several factors and GTP for energy. The ribosome releases the *m*RNA and the newly formed polypeptide (further processing of the polypeptide may occur to make it functional).

Gene Expression

- It is the process by which information contained in genes is decoded to produce other molecules that determine the phenotypic traits of organisms.
- The process is initiated, when the information contained in the base sequence of DNA is copied into a molecule of RNA.

- The processed RNA molecule is used to specify the order in which amino acids are joined together to form a polypeptide chain.
- **Johannsen** in 1909 introduced a term 'gene' as an elementary unit of inheritance. Gene can be defined as

Functional unit – Cistron – Structural unit of DNA which determines the synthesis of complete polypeptide. It is considered equivalent to structural gene.

Recon	–	Unit of recombination
Muton	–	Unit of mutation.

Regulation of Gene Expression

Gene regulation is the mechanism of switching off and switching on, the gene depending upon the requirement of cells and the state of the development.

Control of Gene Expression in Prokaryotes (Operons)

The hypothesis of this regulation was given by F Jacob and J Monod and is known as Operon model. The theory was given on the basis of study of *lac* (lactose) operon in *E. coli*. The operon consists of four main components. These are as follows:

(i) **Regulator** (i) **1200 bp** Regulator is responsible for the synthesis of protein called repressor. The active repressor is seen in inducible system, while inactive repressor is seen in repressible system.

(ii) **Promoter** (ρ) **30 bp** It is the segment at which RNA polymerase binds. It initiates the transcription of structural gene and control the rate of *m*RNA synthesis.

(iii) **Operator** (o) **35 bp** This segment of DNA impose control over the transcription. This region works like 'on' and 'off' switch for protein synthesis.

(iv) **Structural gene** (z, y and a) **3063 bp**, **800 bp, 800 bp** This region of the DNA codes for the synthesis of proteins. These determine the primary structure of polypeptide.

Out of these four components, the first three genes among the above produce three main compounds, i.e. repressor, inducer and corepressor.

Repressor has the capacity to bind on operator gene only after activation by corepressor. Another protein, inducer have the capacity to bind to operator as well as repressor.

Types of Operon

On the basis of their activity principles, the operons are of two types:

(a) **Inducible system** This operon is inactive in normal case and can be activated by inducer molecule, which is synthesised by regulator. It shows two possibilities.

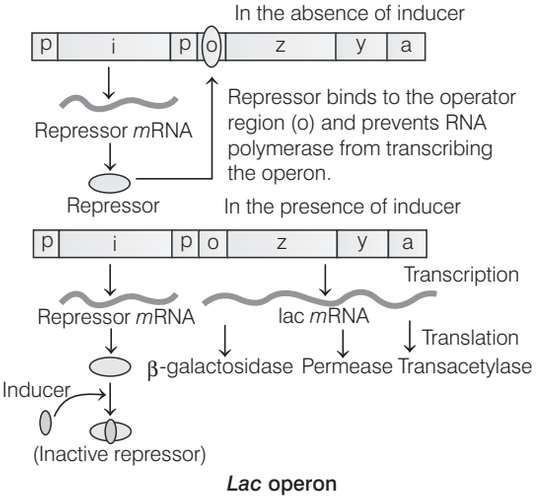
- **When inducer is absent** In this, regulator gene produce active repressor, which bind to operator gene and blocks transcription and protein synthesis.
- **When inducer is present** Regulator gene produce active repressor, which forms inducer-repressor complex. Thus, it does not bind to operator gene and transcription and translation goes on.

(b) **Repressible system** This operon is active in normal cases and could be inactivated by corepressor. It shows two types of possibilities.

- **When corepressor absent** In this, regulator gene produce aporepressor which does not have affinity for operator gene. So, it does not bind to operator to block the transcription and translation.
- **When corepressor present** The aporepressor produced is combined with corepressor to activate it. This active repressor then binds to operator gene and blocks both transcription and translation.

Lac Operon

- Lactose or *lac* operon of *E.coli* is an example of inducible operon. Lactose is known to be the inducer and the substrate for enzyme β-galactosidase.
- If lactose is provided as the carbon source in the growth medium, in the absence of preferred carbon source such as glucose, the lactose is transported to the cells by the action of enzyme permease and induces the operon.



***Lac* operon**

Inducing of Operon by Lactose

- The '*i*' gene synthesises the repressor of the operon which have tendency to repressor bind with operator region of the operon, thus prevent RNA polymerase from transcribing of operon. It can be inactivated by the interaction with the inducer (lactose) or allolactose.
- If the repressor is inactivated, RNA polymerase can have access to the promoter and transcription continues. Regulation of *lac* operon by repressor is referred to as negative regulation. *Lac* operon can work under the control of positive regulation also.

- **When lactose is absent** The '*i*' gene regulates and produces repressor *m*RNA in the absence of lactose. The repressor protein binds to the operator region of the operon and as a result it prevents RNA polymerase to bind with the operon, thus blocking transcription. The operon will get switched off in this situation.
- **When lactose is present** Lactose acts as an inducer here and binds to the repressor. Thus, it forms an inactive repressor. The repressor fails to bind to the operator region.
  Therefore the RNA polymerase binds to the operator and transcripts *lac m*RNA. In this situation, operon will be switched on.

*Lac m*RNA is known to be polycistronic, which produces all three enzymes, e.g. β-galactosidase, permease and transacetylase.

Control of Gene Expression in Eukaryotes

In eukaryotes, the most accepted theory is Operon-Operator Model of Britton-Davidson (1969). According to this model, the eukaryotic operon contains four basic type of genes

(i) **Sensor** These gene segments are sensitive to cellular environment.

(ii) **Interogator** These acts as carrier of signal from sensor to receptor.

(iii) **Receptor** The signal is received by these genes. These are associated with producer.

(iv) **Producer** These are output control centre.

The gene regulation occurs at various levels:

- At the level of transcription
- At the level of RNA processing and splicing
- At the level of translation.

Human Genome Project (HGP)

- HGP was 13 year project that was launched in the year 1990 and completed in 2003.
- The project was coordinated by US Department of Energy and the National Institute of Health.
- The goals of HGP are as follows
  - Identify all the genes in human DNA.
  - Determine the sequences of the 3 billion base pairs present in human DNA.
  - Store the information in databases.
  - Improve the tools for data analysis.
  - Transfer the technologies to other sectors.
  - Address the ethical, legal and social issues that may arise from this project.
- The salient features considered in human genome are listed below
  - The human genome contains 3164.7 million nucleotides (base pairs).
  - The size of genes varies. The longest gene of human body is duchenne muscular dystrophy on X- chromosome and the smallest gene is Testis Determining Factor (TDF) on Y-chromosome.
  - The total number of genes is estimated at 30000 and 99.9% of the nucleotides are the same in all humans.
  - Only less than 2% of the genome codes for proteins.
  - The part of DNA called satellite DNA contains repeated sequences. The moderately repeated sequences (150-300 bp) include transposons and Alu elements. Minisatellite sequences are 11-60 bp long hypervariable repeat sequences. They are VNTRs used in DNA fingerprinting.
  - Scientists have identified about 1.4 million locations, where DNA differs in single base in human beings. These are called Single Nucleotide Polymorphisms (SNPs).
- The various advantages of HGP are as follows
  - Knowledge of the effects of variations of DNA among individuals can provide the ways to diagnose.
  - Treat and even prevent a number of diseases/disorders that affect human beings.
  - It provides clues to the understanding of human biology.

DNA Fingerprinting

- It involves identification of differences in repetitive DNA and it was first discovered by Alec Jeffreys in 1985-86.
- These repetitive DNA are separated from the bulk genomic DNA during density gradient centrifugation.
- The bulk DNA forms major peak during centrifugation and the small peaks are formed by repetitive or satellite DNA.
- These sequences do not code for any proteins but they constitute a large portion of genome.
- The technique of DNA fingerprinting involves the following steps
  - **DNA isolation** involves the extraction of DNA from the cells in a high speed centrifuge.
  - **Amplification** in which many copies of the extracted DNA can be made by the use of polymerase chain reaction.
  - **Digestion** of DNA by restriction endonucleases.
  - **Separation** of DNA fragments by electrophoresis.
  - **Blotting** involves the transfer of separated DNA fragments to synthetic membranes (like nylon or nitrocellulose).
  - **Hybridisation** is done with the help of a radiolabelled VNTR probe (small segments of DNA which help to detect the presence of a gene in a long DNA sequence). These proteins target a specific nucleotide sequence that is complementary to them.
  - **Autoradiography** involves the detection of hybridised DNA fragments in the form of bands of different size. These bands give a characteristic pattern for individual DNA.

DAY PRACTICE SESSION 1

FOUNDATION QUESTIONS EXERCISE

1 First code of every *m*RNA is
(a) either AUG or GUG
(b) either AUG or UAG
(c) either GUG or UAA
(d) either UGA or UAA

2 The final proof for DNA as the genetic material came from the experiments of ➝ NEET 2017
(a) Griffith
(b) Hershey and Chase
(c) Avery, MacLeod and McCarty
(d) Har Gobind Khorana

3 Transformation was discovered by ➝ CBSE-AIPMT 2014
(a) Meselson and Stahl
(b) Hershey and Chase
(c) Griffith
(d) Watson and Crick

4 Antiparallel strands of a DNA molecule means that
(a) one strand turns anticlockwise
(b) the phosphate groups of two DNA strands at their ends, share the same position
(c) the phosphate groups at the start of two DNA strands are in opposite position (pole)
(d) one strand turns clockwise

5 In AGCT bases of DNA, hydrogen bonds and base pairing occur between
(a) A = T, G ≡ C (b) A ≡ T, G = C
(c) A = G, T ≡ C (d) A ≡ G, T = C

6 The following ratio is generally constant for a given species
(a) A + G / C + T (b) T + C / G + A
(c) G + C / A + T (d) A +G / T + C

7 What are the structures called that give an appearance as 'beads' on string in the chromosomes when viewed under electron microscope? ➝ CBSE-AIPMT 2011
(a) Genes (b) Nucleotides
(c) Nucleosomes (d) Base pairs

8 Nucleosome core is made of
(a) H1, H2A, H2B and H3
(b) H1, H2A, H2B and H4
(c) H1, H2A, H2B, H3 and H4
(d) H2A, H2B, H3 and H4

9 The association of histone H1 with a nucleosome indicates ➝ NEET 2017
(a) transcription is occurring
(b) DNA replication is occurring
(c) the DNA is condensed into chromatin fibre
(d) the DNA double helix is exposed

10 *Escherichia coli* fully labelled with N^{15} is allowed to grow in N^{14} medium. The two strands of DNA molecule of the first generation bacteria have
(a) different density and do not resemble parent DNA
(b) different density but resemble parent DNA
(c) same density and resemble parent DNA
(d) same density but do not resemble parent DNA

11 Which of the following is capable of self-replication?
(a) Enzymes (b) Amino acid
(c) DNA (d) Protein

12 DNA replication in bacteria occurs ➝ NEET 2017
(a) during S-phase (b) within nucleolus
(c) prior to fission (d) just before transcription

13 The experimental proof for semiconservative replication of DNA was first shown in a ➝ NEET 2018
(a) plant (b) bacterium (c) fungus (d) virus

14 Ligase is an enzyme responsible for
(a) renaturation of DNA (b) Proofreading
(c) joining of DNA bits (d) breaking of DNA

15 Semiconservative model of DNA replication was first demonstrated by
(a) Watson and Crick (b) Meselson and Stahl
(c) Herbert Taylor (d) Har Gobind Khorana

16 The enzyme which helps to cut one strand of DNA duplex to release tension of coiling of two strands is
(a) DNA ligase
(b) DNA polymerase-I
(c) topoisomerase
(d) swielases (helicase or unwindases)

17 The unwinding of DNA helix is carried out by the enzyme
(a) DNA ligase (b) DNA helicase
(c) DNA polymerase (d) Topoisomerase

18 The strand synthesised continuously and discontinuously are called
(a) leading strand and lagging strand, respectively
(b) lagging strand and leading strand, respectively
(c) template strand and leading strand, respectively
(d) non - template strand and lagging strand, respectively

19 Taylor conducted the experiments to prove semiconservative mode of chromosome replication on
(a) *Vinca rosea* ➝ NEET-II 2016
(b) *Vicia faba*
(c) *Drosophila melanogaster*
(d) *E. coli*

20 Which of the following RNAs should be most abundant in animals cell? ➜ NEET 2017

(a) *r*RNA (b) *t*RNA
(c) *m*RNA (d) *m*RNA

21 Which of the following *r*RNAs acts as structural RNA as well as ribozyme in bacteria? ➜ NEET-II 2016

(a) 5*sr*RNA (b) 18 *sr*RNA
(c) 23 *sr*RNA (d) 58 *sr*RNA

22 Which one of the following is not applicable to RNA ? ➜ CBSE-AIPMT 2015

(a) Complementary base pairing
(b) 5' phosphoryl and 3' hydroxyl ends
(c) Heterocyclic nitrogenous bases
(d) Chargaff's rule

23 Ribosomal RNA is actively synthesised in ➜ CBSE-AIPMT 2012

(a) lysosomes (b) nucleolus
(c) nucleoplasm (d) ribosomes

24 Removal of RNA polymerase - III from nucleoplasm will affect the synthesis of ➜ CBSE-AIPMT 2012

(a) *t* RNA (b) *hn*RNA (c) *m*RNA (d) *r*RNA

25 Select the correct match.

(a) TH Morgan – Transduction
(b) F_2 × Recessive parent – Dihybrid cross
(c) Ribozyme – Nucleic acid
(d) G Mendel – Transformation

26 The DNA in nucleolar organiser region codes for *r*RNA is called

(a) *m*DNA (b) *r*DNA
(c) *t*DNA (d) All of these

27 Central dogma of modern biology is

(a) Proteins → RNA → DNA
(b) RNA → DNA → Proteins
(c) DNA → RNA → Proteins
(d) RNA → Proteins → DNA

28 Which one of the following makes use of RNA as a template to synthesise DNA?

(a) Reverse transcriptase
(b) DNA dependant RNA polymerase
(c) DNA polymerase
(d) RNA polymerase

29 Reverse transcription is the process of formation of DNA from RNA. It was discovered by

(a) Watson and Crick
(b) Khorana
(c) Temin and Baltimore
(d) Beadle and Tatum

30 Reverse transcriptase is

(a) RNA dependent RNA polymerase
(b) DNA dependent RNA polymerase
(c) DNA dependent DNA polymerase
(d) RNA dependent DNA polymerase

31 The diagram shows an important concept in the genetic implication of DNA. Fill in the blanks *A* to *C*

DNA $\xrightarrow{A}$ *m* RNA $\xrightarrow{B}$ Protein $\xrightarrow[C]{\text{Proposed by}}$ ➜ NEET 2013

(a) *A* – Transcription , *B* – Replication, *C* – James Watson
(b) *A* – Translation, *B* – Transcription, *C* – Erwin Chargaff
(c) *A* – Transcription, *B* – Translation, *C* – Francis Crick
(d) *A* – Translation, *B* – Extension, *C* – Rosalind Franklin

32 DNA dependent RNA polymerase catalyses transcription on the strand of the DNA which is called the ➜ NEET-II 2016

(a) template strand (b) coding strand
(c) alpha strand (d) anti-strand

33 Which one of the following is incorrect matched?

(a) Transcription — Writing information from DNA to *t*RNA
(b) Translation — Using information in *m*RNA to make protein
(c) Repressor protein — Binds to operator to stop enzyme synthesis
(d) Operon — Structural genes, operator and promoter

34 AGGTATCGCAT is a sequence from the coding strand of a gene. What will be the corresponding sequence of the transcribed *m*RNA? ➜ NEET 2018

(a) ACCUAUGCGAU
(b) UGGTUTCGCAT
(c) AGGUAUCGCAU
(d) UCCAUAGCGUA

35 Which one of the following is not a part of a transcription unit in DNA ? ➜ CBSE-AIPMT 2012

(a) The inducer (b) A terminator
(c) A promoter (d) The structural gene

36 If one strand of DNA has the nitrogenous base sequence as ATCTG, what would be the complementary RNA strand sequence? ➜ CBSE-AIPMT 2012

(a) TTAGU (b) UAGAC
(c) AACTG (d) ATCGU

37 During transcription, the nucleotide sequence of the DNA strand that is being coded is ATACG, then the nucleotide sequence in the *m*RNA would be

(a) TATGC (b) TCTGG
(c) UAUGC (d) UATGG

38 The sequence of bases that orient RNA polymerase, so that synthesis proceeds in left to right direction is called

(a) Pribnow box (b) promoter box
(c) regulator box (d) operator box

39 Hogness box or TATA box has base sequences

(a) GATAGA (b) TATA AA
(c) GAGAGA (d) GATGAT

40 In which direction *m*RNA is synthesised on DNA template?

(a) 5′→ 3′ (b) 3′→ 5′
(c) Both (a) and (b) (d) All of these

41 During transcription, RNA polymerase holoenzyme binds to a gene promoter and assumes a saddle-like structure. What is it's DNA binding sequence?
(a) TTAA (b) AATTA
(c) CACC (d) TATA

42 Molecular basis of organ differentiation depends on the modulation in transcription by
(a) RNA polymerase (b) ribosome
(c) transcription factor (d) anticodon

43 If the base sequence on template strand is ACTGCTA then, the complementary strand will have
(a) ACTCACA (b) TGACGAT
(c) ATGACGTA (d) ACTGCTA

44 The *m*RNA sequences in eukaryotes, which do not code for anything are called
(a) introns (b) interferons
(c) exons (d) endons

45 Removal of introns and joining of exons in a defined order during transcription is called
(a) looping (b) inducing
(c) slicing (d) splicing

46 Genetic code is said to be degenerate because
(a) codons degenerate very quickly
(b) one amino acid is coded by more than one codon
(c) one codon codes for more than one amino acid
(d) None of the above

47 Non-sense codon takes part in
(a) terminating message of gene controlled protein synthesis
(b) formation of unspecified amino acids
(c) conversion of sense DNA into non-sense one
(d) releasing *t*RNA from polypeptide chain

48 Which one of the following is the starter codon?
➜ NEET-I 2016
(a) UGA (b) UAA (c) UAG (d) AUG

49 What is not true for genetic code? ➜ CBSE-AIPMT 2009
(a) A codon in *m*RNA is read in a non-contiguous fashion
(b) It is nearly universal
(c) It is degenerate
(d) It is unambiguous

50 Whose experiments cracked the DNA and discovered unequivocally that a genetic code is a triplet?
➜ CBSE-AIPMT 2009
(a) Nirenberg and Matthaei
(b) Hershey and Chase
(c) Morgan and Sturtevant
(d) Beadle and Tatum

51 Which one of the following pairs is correctly matched with regard to the codon and the amino acid coded by it?
(a) UUA - valine (b) AAA - lysine
(c) AUG - cysteine (d) CCC - alanine

52 One gene-one enzyme hypothesis was postulated by
(a) R - Franklin (b) Hershey and Chase
(c) A Garrod (d) Beadle and Tatum

53 One gene-one enzyme relationship was established for the first time in
(a) *Neurospora crassa*
(b) *Salmonella typhimurium*
(c) *Escherichia coli*
(d) *Diplococcus pneumoniae*

54 Process of protein synthesis in a cell is called
(a) transcription (b) translation
(c) transduction (d) translocation

55 Many ribosomes may associate with a single *m*RNA to form multiple copies of a polypeptide simultaneously. Such strings of ribosomes are termed as
➜ NEET 2018, 2016
(a) plastidome (b) polyhedral bodies
(c) polysome (d) nucleosome

56 A DNA strand is directly involved in the synthesis of all the following except
(a) *t*RNA molecule (b) *m*RNA molecule
(c) another DNA strand (d) protein synthesis

57 In an animal cell, protein synthesis takes place
(a) only on the ribosomes protein in the cytosol
(b) only on the ribosomes attached to nuclear envelope and endoplasmic reticulum
(c) on ribosomes present in the nucleolus as well as in cytoplasm
(d) on ribosomes present in the cytosol as well as in the mitochondria

58 In translocation on ribosome, the amino acid shifts from
(a) A site to P-site (b) P-site to A-site
(c) *t*RNA to A-site (d) *t*RNA to P-site

59 Amino acid sequence in protein synthesis is decided by the sequence of
(a) *t*RNA (b) *m*RNA
(c) *c*DNA (d) *r*RNA

60 Which one of the following hydrolyses internal phosphodiester bonds in a polynucleotide chain?
➜ CBSE-AIPMT 2005
(a) Lipase (b) Exonuclease
(c) Endonuclease (d) Protease

61 The equivalent of a structural gene is ➜ NEET-II 2016
(a) muton (b) cistron
(c) operon (d) recon

62 Gene and cistron words are sometimes used synonymously because
(a) one cistron contains many genes
(b) one gene contains many cistrons
(c) one gene contains one cistron
(d) one gene contains no cistron

63 Select the correct match. ➝ NEET 2018

(a) Matthew Meselson and F Stahl — *Pisum sativum*
(b) Alfred Hershey and Martha Chase — TMV
(c) Alec Jeffreys — *Streptococcus pneumoniae*
(d) Francois Jacob and Jacques Monod — *Lac* operon

64 All of the following are parts of an operon except ➝ NEET 2018

(a) an enhancer
(b) structural genes
(c) an operator
(d) a promoter

65 Which of the following is required as inducer (s) for the expression of *lac* operon? ➝ NEET-I 2016

(a) galactose
(b) lactose
(c) lactose and galactose
(d) glucose

66 Which enzyme/s will be produced in a cell in which there is a non-sense mutation in the *lac* Y-gene?

(a) β-galactosidase
(b) Lactose permease
(c) Transacetylase
(d) Lactose permease and transacetylase

67 Human genome project was coordinated by which country?

(a) Japan (b) Europe
(c) China (d) USA

68 Which of the following is not required for any of the techniques of DNA fingerprinting available at present? ➝ NEET-I 2016

(a) Zinc finger analysis
(b) Restriction enzymes
(c) DNA - DNA hybridisation
(d) Polymerase Chain reaction

69 DNA fingerprinting refers to

(a) molecular analysis or profiles of DNA samples
(b) analysis of DNA samples using imprinting device
(c) techniques used for molecular analysis of different specimens of DNA
(d) techniques used for identification of fingerprints of individuals

70 Match the following columns.

	Column I		Column II
A.	DNA synthesis	1.	Manufactures ribosomes
B.	Crick	2.	Ribosomes
C.	Nucleolus	3.	DNA polymerase
D.	Translation	4.	Central dogma
		5.	Mitochondria

Codes

	A	B	C	D
(a)	2	1	3	5
(b)	5	2	4	1
(c)	3	4	1	2
(d)	1	2	3	4

71 Select the two statements out of the four (I-IV) given below about *lac* operon.

I. Glucose or galactose may bind with the repressor and inactivate it.
II. In the absence of lactose, the repressor binds with the operator region.
III. The Z-gene codes for permease.
IV. This was elucidated by Francois Jacob and Jacques Monod.

The correct statements are ➝ CBSE-AIPMT 2010

(a) I and III (b) I and IV (c) II and IV (d) I and II

72 **Assertion** *Lac* operon is seen in *E. coli.*

Reason *E. coli* lacks a definite nucleus.

(a) Both Assertion and Reason are true and Reason is the correct explanation of Assertion
(b) Both Assertion and Reason are true, but Reason is not the correct explanation of Assertion
(c) Assertion is true, but Reason is false
(d) Both Assertion and Reason are false

DAY PRACTICE SESSION 2

PROGRESSIVE QUESTIONS EXERCISE

1 Wobble hypothesis establishes
(a) peptide chain formation
(b) initiation of peptide chain
(c) termination of peptide chain
(d) economy in *t*RNA molecules

2 Genes that are involved inturning on or off the transcription of a set of structural genes are called
(a) polymorphic genes
(b) operator genes
(c) reductant genes
(d) regulatory genes

3 Artificial gene for alanine *t*RNA with 77 base pairs was first synthesised by
(a) Khorana *et. al* (1968)
(b) Nirenberg and Mathaei (1967)
(c) Beadle and Tatum (1966)
(d) Watson and Crick (1958)

4 Cistron was called as functional gene in 1955 by
(a) Bateson (b) Waldeyer
(c) Benzer (d) Balling

5 Genetic code was discovered by
(a) Holley and Ochoa
(b) Watson and Crick
(c) Nirenberg and Matthaei
(d) Holley, Nirenberg and Khorana

6 Polypeptide chain is initiated by
(a) leucine (b) lysine
(c) methionine (d) glycine

7 The transforming substance in *Pneumococcus* as found out by Avery, MacLeod and McCarty was
(a) *m*RNA (b) DNA
(c) protein (d) polysaccharide

8 A DNA with unequal nitrogen bases would most probably be
(a) single-stranded (b) double-stranded
(c) triple-stranded (d) four-stranded

9 Which one of the following pairs of codons is correctly matched with their function or the signal for the particular amino acid?
(a) GUU, GCU — Alanine
(b) UAG, UGA — Stop
(c) AUG, ACG — Start / methionine
(d) UUA, UCA — Leucine

10 A molecule that can act as a genetic material must fulfil the traits given below, except ➜ NEET-II 2016
(a) it should be able to express itself in the form of 'Mendelian characters'
(b) It should be able to generate its replica
(c) It should be unstable structurally and chemically
(d) It should provide the scope for slow changes that are required for evolution

11 *E. coli* about to replicate was placed in a medium containing radioactive thymidine for five minutes. Then it was made to replicate in a normal medium. Which of the following observations will be correct?
(a) Both the strands of DNA will be radioactive
(b) One strand will be radioactive
(c) Each strand will be half radioactive
(d) None of the strands will be radioactive

12 *E. coli* cells with a mutated Z-gene of the *lac* operon cannot grow in medium containing only lactose as the source of energy because
(a) in the presence of glucose, *E. coli* cells do not utilise lactose
(b) they cannot transport lactose from the medium into the cells
(c) the *lac* operon is constitutively active in these cells
(d) they cannot synthesise functional β-galactosidase

13 The Okazaki fragments in DNA chain growth
(a) result in transcription
(b) polymerise in the 3′ to 5′ direction and forms replication fork
(c) prove semiconservative nature of DNA replication
(d) polymerise in the 5′ to 3′ direction and explain 3′ to 5′ DNA replication

14 Identify the correct order of organisation of genetic material from largest to smallest.
(a) Chromosome, gene, genome, nucleotide
(b) Genome, chromosome, nucleotide, gene
(c) Genome, chromosome, gene, nucleotide
(d) Chromosome, genome, nucleotide, gene

15 The basis for DNA fingerprinting is
(a) occurrence of Restriction Fragment Length Polymorphism (RFLP)
(b) phenotypic differences between individuals
(c) availability of cloned DNA
(d) knowledge of human karyotype

16 Protein synthesis in an animal cell takes place
(a) only in the cytoplasm
(b) in the nucleolus as well as in the cytoplasm
(c) in the cytoplasm as well as in mitochondria
(d) only on ribosomes attached to a nucleus

17 During translation initiation in prokaryotes, a GTP molecule is needed in
(a) association of 30S *m*RNA with formyl-met-*t*RNA
(b) association of 50S subunit of ribosome with initiation complex
(c) formation of formyl-met-*t*RNA
(d) binding of 30S subunit of ribosome with *m*RNA

18 Khorana first deciphered the triplet codons of
(a) serine and isoleucine
(b) threonine and histidine
(c) tyrosine and tryptophan
(d) phenylalanine and methionine

19 If there are 999 bases in an RNA that codes for a protein with 333 amino acids and the base at position 901 is deleted such that the length of the RNA becomes 998 bases, how many codons will be altered?
(a) 1 (b) 11
(c) 33 (d) 333

20 Match the following columns.

	Column I		**Column II**
A.	Transcription	1.	Adenine
B.	*t*RNA	2.	Cytosine
C.	Purine	3.	DNA fingerprinting
D.	Pyrimidine	4.	RNA polymerase
		5.	Anticodon loop

Codes

	A	B	C	D		A	B	C	D
(a)	4	5	1	2	(b)	1	2	3	4
(c)	2	1	3	5	(d)	5	2	4	3

21 **Assertion** Replication and transcription occur in the nucleus but translation occurs in the cytoplasm.
Reason *m*RNA is transferred from the nucleus into the cytoplasm, where ribosomes and amino acids are available for protein synthesis.
(a) If both Assertion and Reason are true and Reason is the correct explanation of Assertion
(b) If both Assertion and Reason are true but Reason is not the correct explanation of Assertion
(c) If Assertion is true but Reason is false
(d) If both Assertion and Reason are false

ANSWERS

SESSION 1

1 (c)	**2** (b)	**3** (c)	**4** (c)	**5** (a)	**6** (d)	**7** (c)	**8** (d)	**9** (c)	**10** (b)
11 (c)	**12** (c)	**13** (b)	**14** (c)	**15** (b)	**16** (c)	**17** (b)	**18** (a)	**19** (b)	**20** (a)
21 (c)	**22** (d)	**23** (b)	**24** (a)	**25** (c)	**26** (b)	**27** (c)	**28** (a)	**29** (c)	**30** (d)
31 (c)	**32** (a)	**33** (a)	**34** (c)	**35** (a)	**36** (b)	**37** (c)	**38** (a)	**39** (b)	**40** (a)
41 (d)	**42** (c)	**43** (b)	**44** (a)	**45** (d)	**46** (b)	**47** (a)	**48** (d)	**49** (a)	**50** (a)
51 (b)	**52** (d)	**53** (a)	**54** (b)	**55** (c)	**56** (d)	**57** (d)	**58** (b)	**59** (b)	**60** (c)
61 (b)	**62** (c)	**63** (d)	**64** (a)	**65** (b)	**66** (a)	**67** (d)	**68** (a)	**69** (a)	**70** (c)
71 (c)	**72** (b)								

SESSION 2

1 (d)	**2** (a)	**3** (a)	**4** (c)	**5** (c)	**6** (c)	**7** (b)	**8** (a)	**9** (b)	**10** (d)
11 (b)	**12** (c)	**13** (d)	**14** (c)	**15** (a)	**16** (c)	**17** (a)	**18** (b)	**19** (c)	**20** (a)
21 (a)									