## CUET (UG) Biology Notes: Molecular Basis of Inheritance

### 1. Structure and Packaging of DNA

DNA is a long polymer of deoxyribonucleotides. The length of DNA is defined by the number of base pairs (bp).

**The Double Helix Model (Watson & Crick, 1953)**

- Backbone: Formed by sugar-phosphate. Bases project inwards.
- Base Pairing: Adenine pairs with Thymine (2 hydrogen bonds); Guanine pairs with Cytosine (3 hydrogen bonds). Purines always pair with pyrimidines, ensuring uniform distance between the strands.
- Dimensions: The pitch of the helix is 3.4 nm, and there are roughly 10 bp per turn. The distance between consecutive base pairs is 0.34 nm.

**Packaging in Eukaryotes**

- Nucleosome: Negatively charged DNA is wrapped around a positively charged histone octamer (rich in basic amino acids lysine and arginine) to form a nucleosome. A typical nucleosome contains 200 bp of DNA helix.
- Chromatin: Nucleosomes constitute the repeating unit of a structure in the nucleus called chromatin (appears as "beads-on-string").
- Euchromatin vs. Heterochromatin:
    - Euchromatin: Loosely packed, lightly stained, transcriptionally active.
    - Heterochromatin: Densely packed, darkly stained, transcriptionally inactive.

### 2. The Search for Genetic Material

| Experiment | Key Figures | Organism Used | Conclusion / Outcome |
|---|---|---|---|
| Transforming Principle (1928) | Frederick Griffith | Streptococcus pneumoniae | Found that the non-virulent R-strain was "transformed" into the virulent S-strain by absorbing some heat-stable material from the dead S-strain. |
| Biochemical Nature (1944) | Avery, MacLeod, McCarty | Purified biochemicals from S-strain | Proved that DNA is the transforming principle. Only DNAse (DNA-digesting enzyme) inhibited transformation; proteases and RNAse did not. |
| The Blender Experiment (1952) | Hershey & Chase | Bacteriophage (T2) & E. coli | Unequivocally proved DNA is the genetic material. Used radioactive isotopes: $^{32}\text{P}$ to label DNA and $^{35}\text{S}$ to label protein coats. Only $^{32}\text{P}$ entered the bacterial cells. |

### 3. DNA Replication

Replication is semi-conservative (one old strand is conserved, one new strand is synthesized).

- Meselson & Stahl Experiment (1958): Proved semi-conservative replication using E. coli grown in heavy nitrogen ($^{15}\text{N}$) and then shifted to normal nitrogen ($^{14}\text{N}$). They used CsCl density gradient centrifugation to separate the DNA densities.

**The Machinery and Enzymes**

- DNA-dependent DNA Polymerase: The main enzyme. Highly efficient and highly accurate. It only polymerizes in the 5' $\rightarrow$ 3' direction.
- Continuous vs. Discontinuous Synthesis:

  - On the template strand with 3' $\rightarrow$ 5' polarity, replication is continuous (Leading strand).
  - On the template strand with 5' $\rightarrow$ 3' polarity, replication is discontinuous (Lagging strand), forming Okazaki fragments.
- DNA Ligase: Joins the discontinuously synthesized Okazaki fragments together.

---

## 4. Transcription & The RNA World

Transcription is the process of copying genetic information from one strand of DNA into RNA.

### The Transcription Unit

1. Promoter: Located at the 5'-end (upstream) of the structural gene. It provides the binding site for RNA polymerase.
2. Structural Gene: The segment of DNA flanked by the promoter and terminator.
3. Terminator: Located at the 3'-end (downstream) and defines the end of transcription.

### Types of RNA

- mRNA (Messenger): Provides the template for protein synthesis.
- tRNA (Transfer): Brings amino acids and reads the genetic code. Has an anticodon loop and an amino acid acceptor end. (Clover-leaf 2D structure, inverted L-shaped 3D structure).
- rRNA (Ribosomal): Plays structural and catalytic roles during translation.

### Process in Eukaryotes (Post-Transcriptional Modifications)

Eukaryotic structural genes are split (contain coding exons and non-coding introns). The primary transcript (hnRNA) must undergo processing:

1. Splicing: Introns are removed, and exons are joined together.
2. Capping: An unusual nucleotide (methyl guanosine triphosphate) is added to the 5'-end.
3. Tailing: Adenylate residues (poly-A tail, 200-300) are added at the 3'-end.

---

## 5. Genetic Code and Translation

The genetic code determines the sequence of amino acids during protein synthesis.

- Features: Triplet code (61 codons code for amino acids, 3 are stop codons: UAA, UAG, UGA). It is universal, unambiguous (one codon = one amino acid), and degenerate (some amino acids are coded by more than one codon).
- Start Codon: AUG (codes for Methionine and acts as an initiator).

### Translation Process

Translation is the polymerization of amino acids to form a polypeptide.

1. Charging of tRNA: Amino acids are activated in the presence of ATP and linked to their cognate tRNA (aminoacylation).
2. Ribosome: The cellular factory for protein synthesis. In bacteria, the 23S rRNA in the large subunit acts as a catalyst (ribozyme) for peptide bond formation.

---

## 6. Regulation of Gene Expression: The Lac Operon

An operon is a polycistronic structural gene regulated by a common promoter and regulatory genes (found in prokaryotes). The lac operon was elucidated by Jacob and Monod.

- Structure: Consists of one regulatory gene ($i$ gene) and three structural genes ($z$, $y$, and $a$).
    - $i$ gene: Codes for the repressor of the lac operon.
    - $z$ gene: Codes for $\beta$-galactosidase (cleaves lactose into galactose and glucose).
    - $y$ gene: Codes for permease (increases cellular permeability to lactose).
    - $a$ gene: Codes for a transacetylase.
- Mechanism: It is strictly regulated by the substrate lactose (the inducer).
    - Absence of Inducer: Repressor binds to the operator region, preventing RNA polymerase from transcribing the operon (Switched OFF).
    - Presence of Inducer: Lactose binds to the repressor, inactivating it. RNA polymerase gains access to the promoter, and transcription proceeds (Switched ON).

## 7. Human Genome Project (HGP) & DNA Fingerprinting

### Human Genome Project (1990 - 2003)

A mega project with the goal of sequencing the entire human genome.

- Salient Features:
    - The human genome contains 3164.7 million base pairs.
    - The total number of genes is estimated at 30,000.
    - Almost 99.9% of nucleotide bases are exactly the same in all people.
    - The largest known human gene is dystrophin (2.4 million bases).
    - Chromosome 1 has the most genes (2968), and the Y chromosome has the fewest (231).
    - Scientists identified about 1.4 million locations where single-base DNA differences (SNPs - Single Nucleotide Polymorphisms) occur in humans.

### DNA Fingerprinting (Alec Jeffreys)

A technique to identify differences in the DNA of individuals based on highly repetitive sequences that do not code for proteins.

- Principle: Relies on Polymorphism in DNA sequences.
- Probe: Uses Variable Number of Tandem Repeats (VNTR), which belong to a class of satellite DNA called mini-satellites.
- Process: Includes DNA isolation, digestion by restriction endonucleases, separation by gel electrophoresis, Southern blotting (transferring to synthetic membranes), hybridization using a radioactive VNTR probe, and autoradiography.