

DNA and the Genome - Unit 1 Summary

1. The structure and replication of DNA

(a) Structure of DNA

- Nucleotides are composed of deoxyribose sugar, phosphate and a base. These nucleotides join together to form a sugar-phosphate backbone.
- Base pairing occurs between the two strands of DNA between adenine, thymine and guanine, cytosine. These base pairs bond by weak hydrogen bonds.
- The DNA helix is double stranded, and has an anti-parallel structure, with deoxyribose and phosphate at 3' and 5' ends of each strand.

(i) Organisation of DNA in different organisms

- Prokaryotes are a domain of organisms comprising the bacteria and cyanobacteria, characterized by the absence of a distinct, membrane-bound nucleus or membrane-bound organelles, and by DNA that is not organised into chromosomes.
- Eukaryotes are a domain of organisms having cells each with a distinct nucleus within which the genetic material is contained. Eukaryotes include fungi, plants, and animals.
- Circular chromosomal DNA and plasmids can be found in **prokaryotes**.
- Circular plasmids can be found in yeast
- Circular chromosome can be found in mitochondria and chloroplasts of **eukaryotes**.
- Linear chromosomes can be found in the nucleus of **eukaryotes**. This DNA is tightly packaged with associated proteins.

(b) Replication of DNA

- Prior to cell division DNA is replicated by DNA polymerase. It needs a primer to start replication.
- DNA is unwound and unzipped to form two template strands. This process occurs at several locations on the DNA molecule.
- DNA is replicated by DNA polymerase in only one direction; adding complementary nucleotides to the deoxyribose (3') end of a DNA strand.

This means that one strand is replicated continuously and the other is replicated in fragments.

- Fragments of DNA are joined together by ligase.

(i) Polymerase chain reaction (PCR) is the amplification of DNA (in vitro) using complementary primers for specific target sequences at the two ends of the region to be amplified.

- DNA is heated to separate strands then cooled for primer binding to target sequence.

- Heat-tolerant DNA polymerase then replicates the region of DNA.

- Repeated cycles of heating and cooling amplify this region of DNA. Primers are about as accurate as a 'needle in a haystack' and amplification by PCR is about achieving a 'haystack from the needle'.

2. Gene expression

The genetic code used in transcription and translation is found in all forms of life. The phenotype is determined by the proteins produced as the result of gene expression, influenced by intra- and extra-cellular environmental factors.

Only a fraction of the genes in a cell are expressed.

(a) Gene expression is controlled by the regulation of transcription and translation.

- mRNA is transcribed from DNA in the nucleus and translated into proteins by ribosomes in the cytoplasm.

i) Structure and functions of RNA.

- Single strand, replacement of thymine with uracil and deoxyribose with ribose compared to DNA.

- mRNA (messenger) carries a copy of the DNA code from the nucleus to the ribosome.

- **rRNA** and proteins form the ribosome.

- Each tRNA carries a specific amino acid.

(ii) Transcription of DNA into primary and mature RNA transcripts to include the role of RNA polymerase and complementary base pairing.

- RNA polymerase moves along DNA unwinding and unzipping the double helix and synthesising a primary transcript of RNA from RNA nucleotides by complementary base pairing.
- The introns of the primary transcript of mRNA are non-coding and are removed in RNA splicing.
- The exons are coding regions and are joined together to form mature transcript. This process is called RNA splicing.

(iii) Translation of mRNA into a polypeptide by tRNA at the ribosome.

- tRNA folds due to base pairing to form a triplet anticodon site and an attachment site for a specific amino acid.
- Triplet codons on mRNA and anticodons translate the genetic code into a sequence of amino acids.
- Start and stop codons exist.
- tRNA recognises the mRNA codon and joins it bringing along its corresponding amino acid, peptide bond form between amino acids and tRNA exits from the ribosome as the polypeptide is formed.

(iv) One gene, many proteins as a result of RNA splicing and post-translational modification.

- Different mRNA molecules are produced from the same primary transcript depending on which RNA segments are treated as exons and introns.
- Post translation protein structure modification by cutting and combining polypeptide chains or by adding phosphate or carbohydrate groups to the protein.

(v) Proteins are held in a three-dimensional shape.

- Proteins have a large variety of structures and shapes resulting in a wide range of functions
- Peptide bonds form to bind amino acids in a chain called a polypeptide, polypeptide chains can be folded to form the 3 dimensional shape of a protein, held together by hydrogen bonds and interactions between individual amino acids.
- Fibrous proteins include silk, actin, collagen and keratin.
- Globular proteins include hormones, antibodies, membrane proteins and enzymes.

(b) Cellular differentiation

The process by which a cell develops more specialised functions by expressing the genes characteristic for that type of cell.

- Differentiation into specialised cells from meristems in plants; embryonic and tissue (adult) stem cells in animals.
- Meristems are regions of unspecialised cells in plants that are capable of cell division.
- Stem cells are relatively unspecialised cells in animals that can continue to divide and can differentiate into specialised cells of one or more types. In the very early embryo, embryonic stem cells differentiate into all the cell types that make up the organism.

(i) Research and therapeutic uses of stem cells by reference to the repair of damaged or diseased organs or tissues.

- Stem cell research provides information on how cell processes such as cell growth, differentiation and gene regulation work.
- Stem cells can be used as model cells to study how diseases develop or for drug testing.
- Stem cells can be used in the repair of diseased or damaged tissue e.g. used in skin grafts, bone marrow transplants and cornea repair
- Sources of stem cells can include embryonic stem cells, tissue (adult) stem cells and attempts to reprogram specialised cells to embryonic state (induced pluripotent stem cells).
- Tissue (adult) stem cells replenish differentiated cells that need to be replaced and give rise to a more limited range of cell types. Once a cell becomes differentiated it only expresses the genes that produce the characteristic for that cell type.
- The ethical issues of stem cell use and the regulation of their use.
- Embryo cells must not be allowed to develop beyond 14 days, around the same time a blastocyst would be implanted in a uterus.
- Other ethical considerations include the use of induced pluripotent stem cells and the use of nuclear transfer techniques.

3. Genome

The genome of an organism is its hereditary information encoded in DNA. DNA sequences that code for protein are defined as genes. A genome is made up of genes and other DNA sequences that do not

code for proteins. Most of the eukaryotic genome consists of these non-coding sequences.

(a) The structure of the genome

- Coding and non-coding sequences include those that regulate transcription and those that are transcribed to RNA but are never translated.
- Non-translated forms of RNA include tRNA, rRNA (ribosomal) and RNA fragments.
- Some non-coding sequences have no known function.

(b) Mutations are changes in the genome that can result in no protein or an altered protein being expressed.

i) Single gene mutations involve the alteration of a DNA nucleotide sequence as a result of the substitution, insertion or deletion of nucleotides.

- Single nucleotide substitutions and inversion are point mutations and can be silent, neutral, missense, or nonsense.
- Sickle cell anaemia and lactose intolerance are examples of point mutations.
- Single-nucleotide substitutions include: missense, nonsense and splice-site mutations.
- Splice site mutations can alter post-translational processing.
- Nucleotide insertions or deletions result in frame-shift mutations or an expansion of a nucleotide sequence repeat.
- Regulatory sequence mutations can alter gene expression.

(ii) Chromosome structure mutations — duplication, deletion and translocation.

(iii) The importance of mutations and gene duplication in evolution e.g. evidence of formation of human chromosome 2 from fusion of two ancestral chromosomes and gene duplication leading to alpha and beta globins in haemoglobin.

(iv) Polyploidy - errors during the separation of chromosomes during cell division can result in cells with whole genome duplications. Polyploidy is important in the evolution of human food crops. Polyploidy examples include banana (triploid) and potato (tetraploid)

as well as swede, oil seed rape, wheat and strawberry. Polyploidy is very rare in animals.

(c) Evolution

- The changes in organisms over generations as a result of genomic variations.

i) Gene transfer.

- Vertical (inheritance) - from parent to offspring as a result of sexual or asexual reproduction.

- Horizontal - prokaryotes and viruses can exchange genetic material in this way (conjugation). This can result in rapid evolutionary change.

- Prokaryotes and viruses can also transfer sequences horizontally into the genomes of eukaryotes.

(ii) Selection.

- Natural selection is the non-random increase in frequency of DNA sequences that increase survival. A non-random reduction in frequency of deleterious DNA sequences. Sexual selection is an increase in successful reproduction. There is a difference in outcome as a result of stabilising, directional and disruptive selection.

(iii) Genetic drift. The random increase and decrease in frequency of sequences, particularly in small populations, as a result of neutral mutations and founder effects.

(iv) Speciation is the generation of new biological species by evolution.

- A species is a group of organisms capable of interbreeding and producing fertile offspring, and which does not normally breed with other groups.

- There are different definitions of the term species e.g. biological species concept and phylogenetic species concept.

- It is difficult to apply the species definition to asexually reproducing organisms.

- The importance of geographical barriers in allopatric speciation.

- The importance of behavioural or ecological barriers in sympatric speciation.

- Hybrid zones form in regions where the ranges of closely related species meet e.g. hooded crow and carrion crow zone in Scotland.

(d) Genomic sequencing

- The sequence of nucleotide bases can be determined for individual genes and entire genomes.
- Fluorescence tagging of nucleotides can be used to identify the base sequences.
- To compare sequence data, computer and statistical analyses (bioinformatics).

i) Evidence from phylogenetics and molecular clocks are used to determine the main sequence of events in evolution.

- This is the use of sequence data to study the evolutionary relatedness among groups of organisms. Sequence divergence is used to estimate time since lineages diverged. For example, comparison of sequences provides evidence for three main domains (bacteria, archaea and eukaryotes).
- Highly conserved DNA sequences are used for comparisons of distantly related organisms
- The use of sequence data and fossil evidence to determine the main sequence of events in evolution of life: cells, last universal ancestor, photosynthetic organisms, eukaryotes, multicellularity, animals, land plants, vertebrates.

(ii) Comparison of genomes from different species.

- Comparison of genomes reveals that many genes are highly conserved across different organisms.
- Many genomes have been sequenced, particularly of disease-causing organisms, pest species and species that are important model organisms for research.
- Comparison of human and chimp genomes reveals rapid change in genes for immune system and regulation of neural development over the last 6 million years.

(iii) Personal genomics and health - Pharmacogenetics.

- Analysis of an individual's genome may lead to personalised medicine through knowledge of the genetic component of risk of disease and likelihood of success of a particular treatment.
- Comparison of individual's genomes focuses on point mutations, repetitive sequence errors and blocks of duplication and deletion.
- There are difficulties in distinguishing between neutral and harmful mutations in both genes and regulatory sequences, and in understanding the complex nature of many diseases.