DNA and Polypeptide Synthesis

DNA in prokaryotes and eukaryotes
The genetic code is universal- the same nucleotide base-pairing code is used in all living organisms, both prokaryotes and eukaryotes, to instruct protein synthesis.

Prokaryotic DNA

DNA of prokaryotes and eukaryotes is chemically the same, but differs in how it is structurally arranged into chromosomes and packaged inside cells.

Location and structure
- Prokaryotic cells have a simpler structure
- They contain a single chromosome in the form of a circular strand of DNA
- Chromosome has no membrane and floats in the cytoplasm, in a dense region known as the nucleoid.
- The circular, double-stranded prokaryotic DNA is not a helix, but two circles of single-stranded DNA twisted around each other like two pieces of string, each joining at its own ends.

Non-chromosomal DNA
- Prokaryotic cells may have one or more small rings of non-chromosomal DNA, called plasmids, floating separately in the cytoplasm. The genes on these plasmids code for features that are not essential to the survival of the cell, but often provide bacteria with a selective advantage, such as resistance to antibiotics
- Plasmids replicate independently of the chromosome.

Packaging of prokaryotic DNA
- DNA is much larger than the cell it often is in.
- DNA is isolated intact, the circular DNA is supercoiled and forms loops around protein/s to form a nucleoid.
- The dense protein is sometimes referred to as the scaffold.

**Eukaryotic DNA**
- Located in a membrane-bound nucleus within the cell.
- DNA is separated into chromosomes, and each chromosome is larger and more complex than the chromosome in a prokaryotic cell.
- There is a large proportion of non-coding DNA in most eukaryotic cells. This is called introns.
- Coding sequences in DNA are called exons.

![Diagram of eukaryotic cell and DNA structure](image)

**Packaging of eukaryotic DNA**
- DNA of eukaryotes is linear rather than circular and winds around proteins (histones) tightly, but does not supercoil.
- It coils in a way that forms nucleosomes - bead-like structures made up of long sequences of DNA wrapped around eight histone protein cores.

**Non-nuclear DNA in eukaryotes**
- Mitochondria and chloroplasts are organelles in eukaryotic cells that contain their own DNA. This is inherited independently of nuclear DNA.
- mtDNA can be used to trace maternal inheritance, since all mitochondria in a female lineage possess identical mtDNA.
The use of mtDNA is of advantage, because mitochondria:
- Are inherited only from the mother, which allows tracing of a direct genetic line
- Don’t combine paternal and maternal genes
- Occur in large numbers in every cell, so they’re easy to access and sample
- Evolve very quickly because mtDNA doesn’t repair enzymes, so mutations often arise during replication.
- Nuclear DNA is present inside the nucleus of each of our cells, and has about 3 billion base pairs and around 200000 protein-coding genes
- The mitochondrial genome exists outside the cell nucleus, and has 13 protein-coding genes, 24 genes coding for RNA and about 16500 base pairs.
- mtDNA has a higher rate of mutation than nuclear DNA, making it easier to identify differences between closely related individuals.
- mtDNA is used to study evolutionary relatedness, construct evolutionary trees, investigate family relatedness and identify people in forensic science.
Polypeptide synthesis

Polypeptides are chains of amino acids and polypeptides join to form proteins. A polypeptide is a molecule made up of a chain of many amino acids joined by peptide bonds. There are about 20 different amino acids that can be linked together in a linear sequence, to form chains up to 300 amino acids in length.

One or more polypeptides twist and join together into a particular three dimensional shape, forming proteins in cells. The sequence and arrangement of amino acids determines the configuration of the protein, and any change in the amino acid sequence may result in a change in the protein’s shape and its ability to carry out its function.

Gene concept- genes in a cell contain all the information for the synthesis and functioning of cellular components. When end product has been made the gene has been ‘expressed’. In specialised cells, only certain genes are expressed in each cell type.

Process of polypeptide synthesis

- DNA does not leave the nucleus since the molecules are too large to pass through the pores in the nuclear membrane.
Hence, an intermediate molecule called messenger RNA (mRNA) is created and this carries a transcribed copy of the instructions from the nucleus to the ribosomes in the cytoplasm.

Ribosomes translate the message carried by mRNA into a cell product such as protein.

Types of RNA

- Messenger RNA (mRNA)
  - Single-stranded and not twisted into a helix
  - Much shorter than DNA (few thousand bases long)
  - Found in both nucleus and cytoplasm
  - Serves as an intermediate molecule, carrying information from DNA in the nucleus to ribosomes in the cytoplasm

- Transfer RNA (tRNA)
  - Molecules occur in cytoplasm
  - Molecule is 75 nucleotides long and twisted into clover leaf
  - At one end - three unpaired bases, called an anticodon → these attach the tRNA to its complementary bases (codon) on the mRNA strand.
- Other end of the tRNA is able to bind with an amino acid temporarily
- Each tRNA molecule will only attach to one particular amino acid
- Specific sequence of three bases at anticodon end determines which amino acid will be made

- Ribosomal RNA (rRNA)
  - Forms structural part of ribosomes
  - Made in the nucleolus of the cell

**Transcription**

Transcriptions occurs when RNA polymerase (an enzyme) binds to a section of DNA and begins building a chain of RNA nucleotides to form a complementary strand of RNA.

1. RNA polymerase binds to a part of the DNA called the promoter and the DNA ‘unzips’ - the hydrogen bonds between the two strands break and the strands separate. This happens in only the part of the DNA that contains the gene to be used. Only one strand of DNA contains the genetic information to make a protein; it is called the non-coding strand or sense strand; the other strand is called the coding strand (it has the same code as the mRNA being made) or antisense strand.
2. The sense strand acts as a template and RNA nucleotides are assembled, forming complementary single-stranded mRNA molecule (DNA is transcribed into mRNA). Sequence of nucleotide bases on mRNA molecule is the same as DNA coding strand, except for U instead of T.
3. mRNA moves out nucleus and into cytoplasm, where it encounters some of the millions of ribosomes in the cell. Usually one mRNA molecule is read by a large number of ribosomes, so multiple chains of the same polypeptide product are produced from one mRNA template molecule.

**Translation**

4. Translation occurs when ribosomes move along the mRNA molecule and attach tRNA molecules to mRNA by temporarily pairing the bases of the tRNA anticodons with their complementary triplets of bases (codons) on the mRNA.
5. Amino acids from the tail end of each tRNA are linked to one another by an enzyme to form a polypeptide chain. Each amino acid is then spliced off its tRNA carrier.
6. tRNAs move away from the mRNA, leaving the growing chain of amino acids, and move back into the cytoplasm where they can pick up another amino acid and be reused.
7. The polypeptide chain is then processed in the cell; it may be joined to one or more other polypeptides and be further processed folded into the correct shape for its functioning, forming the final protein end product.
8. The mRNA is broken down to its individual nucleotides, which can be reused.

**RNA Processing**

In eukaryotic cells, mRNA that is transcribed from DNA is termed pre-mRNA; further editing of this RNA takes place in the nucleus before it acts as a template for translation into a polypeptide. Pre-mRNA contains coding sequences of nucleotides, called exons, which will be translated into amino acid chains (exons are expressed as proteins).

After transcription, mRNA is edited/spliced and introns are removed by a complex molecule called a spliceosome. Result of this splicing is formation of a mature mRNA molecule, which moves from nucleus to cytosol for translation by ribosomes. Instructions for splicing the mRNA are found within the introns - they code for their own removal.

Splicing is the second step in gene regulation and serves an important purpose in complex organisms. A strand of mRNA produced from one gene is not always spliced in the same way. Alternative ways of splicing mRNA give rise to different versions of the same protein.

Eukaryotes have a large amount of non-coding DNA. It was originally deemed as useless but advanced technology has helped scientists discover that its removal disrupts development of individuals in the embryonic state. When some parts of this non-coding DNA are expressed, chromatin is less tightly packed.

Codons in DNA and its transcribed sequence in mRNA determined the amino acid sequence in proteins that are synthesised. There are 20 amino acids made during the translation on the ribosome. The codon AUG is the start codon, signifying where translation should be initiated. There are three stop codons: UAA, UAG, and UGA. There is more than one codon that codes for each amino acid, creating some flexibility for errors in the genetic code.
Functions and importance of polypeptide synthesis: Genes to Proteins

A gene is considered the smallest unit of heredity, and each gene is a portion of DNA with a specific sequence of bases that encodes for a particular trait that can be passed from parent to offspring. A locus is the position of a gene on a chromosome. Genes influence particular characteristics (phenotypes). A chromosome can therefore be described as a linear sequence of genes. The total amount of genetic material that an organism has in each of its cells is called a genome. Alleles are different forms of the same gene.

A gene is a sequence of nucleotides that codes for any molecular cell product, such as a polypeptide chain.

In eukaryotes, mRNA may be processed after transcription (introns are spliced out) before it passes into the cytoplasm and binds with ribosomes to direct the formation of a polypeptide. mRNA is translated, with the help of tRNA in ribosomes, into polypeptide chains.

Homologous chromosomes carry genes for the same traits in the same position (locus), but these genes may have alternative forms (alleles). Specific genes are associated with the specific traits and are expressed in specialised cells.

How genes and the environment affect phenotypic expression

The genotype is the set of genes in an organism’s DNA that is responsible for a particular trait. Genotype can be determined by biological testing. The phenotype is the physical expression, or characteristics, of that trait and can be determined by observation.
Gene expression and phenotype

Stem cells are unspecialised cells that are capable of dividing and becoming specialised tissue. Embryonic stem cells are capable of dividing and giving rise to any type of tissue within an organism, and so they are termed pluripotent. Stem cells also occur in some adult tissue but there are not pluripotent, as they are only able to give rise to cells of one tissue type. Stem cells undergo asymmetric division - they give rise to two daughter cells, one of which will continue to divide and another that will follow the path of differentiation.

Genes must produce the proteins that regulate their own expression. The accurate synthesis of proteins according to DNA instructions is therefore of ultimate importance in assembling amino acids in the correct order in each polypeptide, as this gives rise to the three-dimensional structure of each protein.

Regulation of genes for phenotypic expression

Gene regulation by modifying DNA for transcription

Gene expression can be understood in terms of the switching on and off of genes, as needed.

- Initiation of transcription seems to be determined by how densely DNA is bound to histones and this in turn affects gene expression.
- Epigenetics explores how chemical modifications to either the histones or the chemical structure of DNA may affect the density of DNA binding and, in turn, affect gene expression. Cancer was the first human disease to be linked to epigenetics.

Gene regulation during the transcription process

Elongation of RNA and the termination of transcription are other steps in the process of transcribing DNA to mRNA where regulation of gene expression may occur. The regulation of gene expression in prokaryotes differs from that of eukaryotes, because prokaryotes do not have introns. Bacterial gene regulation has been studied in detail by looking at the LAC operon. The LAC operon relies on the fact that some bacteria use lactose for sustenance, but most bacteria use glucose. If no glucose is present, bacteria turn on a gene that allows them to turn lactose into glucose.

Summary

- Gene expression is the switching on and off of genes to make the required proteins and other end products in particular cell types.
- Phenotypic expression is the result of gene expression - the structure, physiology and behaviour of an individual as a result of genes that have been expressed.
- Some variations within a population are due to the influence of the environment, rather than having a genetic (DNA sequence) basis.
Identical twins have identical genotypes, and therefore any phenotypic differences can be attributed to environmental influences.

- An example of variation brought about by the environment is the difference in colour of flowers in hydrangeas (dependent on pH of the soil).
- Chemical modifications of DNA that do not involve a change in the sequence of nucleotides are termed ‘epigenetic’ modifications. They may be the mechanism by which some environmental factors bring about variation. The result is a change in phenotypic without a change in genotypic.
- Epigenetic changes show links to disease, including cancers and metabolic diseases.

The structure and function of proteins in living things

**Structure of Proteins**
Proteins are made up of one or more long chains of nitrogen-containing amino acids. Each chain is called a polypeptide. Each protein within an organism is folded into a particular shape that is crucial to its functioning. Proteins bind with other molecules to carry out their functions, and so their shape and chemical properties (such as electrical charge and attraction to water) allow the proteins to function in a particular way.

**Chemical structure**
Proteins contain chemical elements carbon, hydrogen, oxygen, nitrogen, and sometimes sulfur. These elements combine to form amino acids, which are the building blocks of proteins. There are about 20 amino acids; they can be put together in a chain of up to 300 amino acids. The amino acids in each linear sequence of polypeptide chains are held together by chemical bonds known as peptide bonds.

One or more polypeptides can be twisted together into a particular shape, resulting in the overall structure of a protein. The sequence and arrangement of the amino acids determines the type of protein, in the same way that sequences of the letters of the alphabet can be used to make words and then sentences.

**Physical structure**
The physical structure of proteins can be described at four levels’ primary, secondary, tertiary, and quaternary.
Types of proteins in cells

**Fibrous proteins** form structural components of cells and tissues; together with water, they form the basic structure of proplasm (the cytoskeleton). They are long and insoluble in water, and also present in ligaments and tendons. Keratin is a fibrous protein in hair and nails.

**Globular proteins** are usually spherical in shape and are compact and soluble in water. They are often transport proteins, such as hemoglobin in the blood. Immunoglobulins (antibodies), hormones and enzymes are other examples of globular proteins.

**Summary**
- Proteins are macromolecules (polymers) made up of one or more polypeptide chains that are formed by sequences of amino acids linked by peptide bonds.
- They have a primary structure (sequence of amino acids in the polypeptide chains), secondary structure (polypeptide chains arranged in a helix or pleats), tertiary structure
(3D shape formed by folding the secondary structure) and quaternary structure (two or more polypeptide chains combined).

- The sequence of amino acids in a protein is responsible for the overall bonding, folding and 3D shape of a protein, which determines its ability to function. A change in one or more amino acids in the sequence has the potential to alter the entire structure of a protein.
- There are three main types of protein: fibrous, globular, and conjugated.

Functions of proteins
The biological properties of proteins depend on the interactions of the proteins with each other and with other molecules. For example, enzymes bind with substrates to catalyse reactions, regulatory proteins control DNA replication and turn on genes, antibodies bind with pathogens, and hormones bind with receptors on target cells.

Structural proteins- support and movement
- Structural proteins are often fibrous and stringy, and found in connective tissues such as skin, cartilage, bone, tendons, and ligaments. Structural proteins also make up shells in invertebrates, and are in hair, nails, and hooves in vertebrates.
- Some, such as tubulin (in microtubules) are responsible for forming the cytoskeleton, which maintains the shape of cells. Contractile proteins also occur in cells, allowing movement to occur.

Enzymes - control of biochemical reactions
- Enzymes are protein molecules involved in all biochemical aspects of cellular metabolism. They catalyse reactions such as those in the chemical respiration pathway and in digestion.
- The shape of an enzyme’s active site determines its binding specificity and therefore its ability to function.
- Enzymes are particularly important in gene functioning, replicating, repairing, and transcribing DNA to make new proteins.

Proteins for cell communication, cell signalling, and biological recognition
- Cells communicate by means of chemical signals. Some proteins embedded in membranes form channels to carry substances that are essential for cell functioning between cells and the environment. Some proteins, such as hormones and neurotransmitters, act as chemical messengers between cells and communicate messages to a cell about the environment around it and trigger responses in the target cell’s functioning.
Antibodies are defence proteins that may be attached to cell membranes or floating freely in body fluids. They also recognise foreign invaders by the different proteins on the surfaces of their cells.

Transport and storage proteins
- Some proteins bind to and carry or store chemicals in the body. These are termed ligand-binding proteins and must be able to bind easily with the chemical and also release it when and where it is needed.
- E.g. hemoglobin- loses its affinity for oxygen in areas where there is a low oxygen concentration, such as in active tissues.
- Some proteins store chemicals for use by the organism - for example, ferritin stores iron, while albumin (in egg whites) and casein (in milk) store amino acids.

Sensory proteins - responding to stimuli
- Some proteins change their shape or biochemical activity in response to stimuli. For example, opsins are proteins in the retina of the eye that detect light.

Summary
- Proteins are grouped according to their structure and functioning to make it easier to remember and compare them
- Structural proteins form the structural and functional part of cell membranes. Some proteins in the cell membrane also function in communication and regulate the movement of substances across the cell membrane.
- Some proteins, such as myosin and actin in muscle cells, are responsible for cell motility, joining together to form filaments that allow movement.
- Storage and transport proteins bind and store or carry other molecules in cells. For example, histones package DNA in a compact form, the protein ferritin stores iron and haemoglobin carries oxygen.
- Some proteins regulate metabolic functioning - these include enzymes (chemical catalysts) and chemical messengers such as hormones and neurotransmitters.
- Proteins involved in cell recognition include antibodies, which defend against disease.

Technology, big data, and the investigation of protein structure and function
- Proteomics is the large-scale study of sets of proteins produced in an organism or biological system to find out how the proteins in a variety of cell types work together.
- X-ray crystallography was first used over a century ago.
- Synchrotron (a particle accelerator) emits electromagnetic radiation where charged particles can be accelerated to move at speeds close to the speed of light.
- Bioinformatics - field of combined computer science and genetics.
- Development of gene banks as well