OCR (A) Biology A-level

Topic 6.1: Genetics and evolution

Notes

Cellular control

Mutations are changes in the sequence of nucleotides in DNA molecules. Types of mutations include:

- **Insertion/deletion mutations** where one or more nucleotide pairs are inserted or deleted from the sequence. This type of mutation alters the sequence of nucleotides after the insertion/deletion point known as a **frameshift**.
- **Point mutation/substitution** occurs where one base pair is replaced by another.
- **A nonsense mutation** is one where translation is stopped early thus giving rise to a truncated polypeptide due to premature introduction of a stop codon.
- **A missense mutation** is a codon change which results in the production of a different amino acid, thus resulting in altered tertiary structure of the protein
- **A silent mutation** is a codon change which does not affect the amino acid sequence produced. Silent mutations are possible due to the **degenerate** nature of the genetic code.

Mutations can either have **neutral effects** where the mutation causes no change to the organism, for example in a case where the mutation occurs in a non-coding region of DNA or is a silent mutation, as described above. A mutation can also be neutral when a **change in tertiary structure of the protein has no effect** on the organism.

Some mutations are beneficial, for instance, humans developed trichromatic vision through a mutation. Harmful mutations include a mutation in the CFTR protein which causes **cystic fibrosis.**

Whether a mutation proves to be beneficial or detrimental to an organism will depend on the environment of the organism.

Controlling gene expression

Gene expression can be controlled **the transcriptional, post-transcriptional and post-transcriptional and post-translational levels**.

An example of **transcriptional control** is the **lac operon**, which is a length of DNA composed of structural genes and control sites which controls the **expression of beta-galactosidase responsible for hydrolysis of lactose in E.coli**. The operon consists of a **promoter region** which is the binding site for RNA polymerase to initiate transcription, **operator region** where the inhibitor binds and structural genes which give rise to 3 products, beta galactosidase, lactose permease and another enzyme. The inhibitor protein is coded for by a **regulator gene**, located outside the operon..

In a case where the **concentration of glucose is high** and the **concentration of lactose is low**, the transcription of the structural genes is **inhibited** due to **binding of the repressor to the operator region**. However, in a case where the **concentration of glucose is low** and **concentration of lactose is high**, lactose **binds to the repressor** thus **causing the shape of its DNA binding site to change**, therefore making it ineffective. This means that it can no longer bind to the operator region therefore RNA polymerase is able to bind to the promotor region and **transcription of the structural genes** takes place.

Gene expression can also be controlled by **transcription factors** which have the ability to switch genes on and off. They do so through interaction with the **promoter sequence** of DNA to **either initiate or inhibit transcription**.

Gene expression is controlled at **post-transcriptional** level by editing of the **primary mRNA** transcript, during which the non-coding regions called **introns** are removed, thus creating a **mature transcript** consisting only of protein-producing regions known as **exons**.

Gene expression can be controlled at the **post-translational** level. For example, proteins such as adrenaline can be activated with the help of **cyclic AMP**. This occurs when **adrenaline binds to a complementary receptor**, which activates the enzyme **adenylate cyclase** which **converts ATP to cyclic AMP** which **starts a cascade** of enzyme reactions within the cell, thus activating the protein.

Genetic control

Homeobox genes are involved in controlling the development of body plan of organisms thus aiding the development of a zygote to a complete organism. They **code for transcription factors** that bind to DNA to **regulate transcription** by **switching genes on and off** when they are required at particular stages of development, for instance during limb formation in humans. Homeobox genes are arranged in **hox clusters**.

Apoptosis

Apoptosis is a form of **programmed cell death** which can act as a mechanism to control the **development of body plans** such as fingers and toes. It is an integral part of tissue development in both plants and animals. It is an ordered, controlled series of biochemical events leading to cell death. It is the **opposite of necrosis**, which is cell death resulting from damage and release of hydrolytic enzymes. It is a means of **controlling the number of cells and ensuring that it remains constant to prevent cancer**.

During the process, enzymes break down the cytoskeleton of the cell, DNA and proteins. As the contents of the cell are broken down, the cell begins to shrink and break up. Subsequently, the **cell fragments are engulfed by phagocytes and destroyed**.

Patterns of inheritance

Phenotypic variation: There are two types of variation, **discontinuous and continuous**. Discontinuous variation is used to describe variation which can be assigned to a particular category, for instance **shoe size or blood type**. Whereas continuous variation is a type of variation where the differences between phenotypes are **quantitative, for instance height or weight**. Variation can be influenced by both **environmental factors** such as diet in animals and etiolation and chlorosis in plants and **genetic factors.**

Sexual reproduction and genetic variation

Meiosis is a form of cell division that gives rise to **genetic variation**. The main role of meiosis is **production of haploid gametes** as cells produced by meiosis have half the number of chromosomes. Meiosis produces genetically different cells, genetic variation is achieved through:

- **Crossing over of chromatids** where pairs of homologous chromosomes line up and exchange some of their genetic material
- **Independent assortment of chromosomes** – there are various combinations of chromosome arrangement

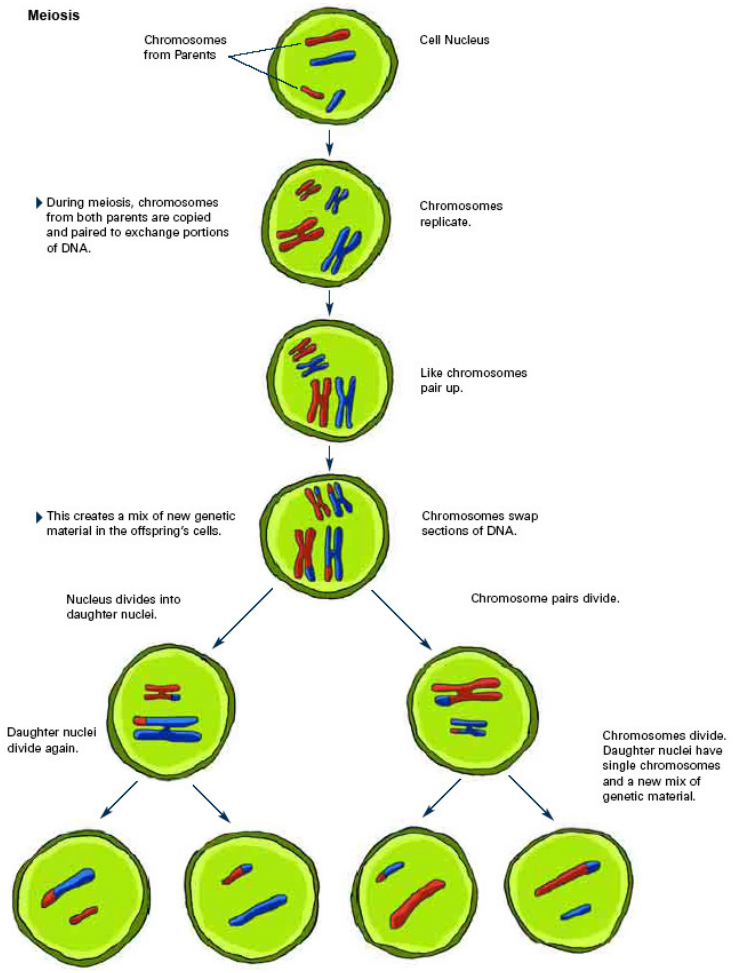

Figure 1 Biologypost

Key terms

- **Allele** – alternative form of a gene
- **Locus** – the specific **position** of a gene on a chromosome, the two alleles of a gene are found at the same loci on the chromosome pairs
- **Phenotype** – **observable characteristics** of an organism which are as a result of **genotype and environment**
- **Genotype** – the **alleles** present within cells of an organism, for a particular **trait or characteristic**
- **Dominant** – only a **single allele** is required for the characteristic to be expressed, that allele is always expressed in the phenotype

- **Recessive** – the characteristic is only expressed if there **is no dominant allele present**
- **Homozygous** – two **identical** alleles
- **Heterozygous** – two **different** alleles
- **Codominance** – **both** alleles contribute to the phenotype

Linkage is the phenomenon where genes for **different characteristics are located at different loci on the same chromosome and so are inherited together**.

Monogenic inheritance – when a phenotype or trait is controlled by a **single gene**. For instance, **cystic fibrosis** where the individuals with doubly recessive genootype are affected.

Dihybrid cross – inheritance of two genes

Sex linkage – expression of an allele **dependent on the gender of the individual** as the gene is located on a **sex chromosome**, for instance, males are more likely to inherit an X-chromosome linked condition because they only have a single copy of the X chromosome. An example of sex linkage is haemophilia which is a recessive condition (hh).

Autosomal linkage – genes which are **located on the same chromosome** (which is not a sex chromosome) and tend to be **expressed together** in the offspring

Codominance – when both alleles are expressed in a heterozygote, that is**, both alleles contribute towards the phenotype**. Examples include **blood type**.

Epistasis – the **interaction of different loci** on the gene, one gene locus affects the other gene locus. One gene loci can either mask or suppress the expression of another gene locus.

Recessive epistasis occurs when the **presence of a recessive allele prevents the expression of another allele at a second locus**. Recessive epistasis gives the ratio of **9:3:4**.

Dominant epistasis is when a **dominant allele at one locus completely masks the alleles at a second locus**. Dominant epistasis gives a ratio of **12:3:1.**

Chi-squared test

$$X^2 = \sum \frac{(\text{observed - expected})^2}{\text{expected}}$$

The **chi squared test** is a **statistical test** which can be used to establish whether the difference between observed and expected results is small enough to occur purely **due to chance.**

- It can be used if the **sample size is sufficiently large**, that is over 20. It can only be used for **discontinuous variation data** in the form of raw counts.
- The chi squared test can be used to determine whether the **null hypothesis** is correct or not. The null hypothesis is the assumption that there **is no difference between observed and expected results.**
- The value obtained is compared to the **critical value**, and in a case where the value obtained is less than the critical value, the null hypothesis is accepted as the difference due to chance is not significant

- Whereas in a case where the x^2 value is greater than critical value, the null hypothesis is rejected meaning that the difference between observed and expected results is not due to chance, as it is significant.

Hardy-Weinberg principle

The **Hardy-Weinberg Equation** can be used to **estimate the frequency of alleles in a population** and to see whether a **change in allele frequency is occurring in a population over time**.

p = the frequency of the **dominant** allele (represented by A)
q = the frequency of the **recessive** allele (represented by a)
For a population in genetic equilibrium:
$p + q = 1.0$ (The sum of the frequencies of both alleles is 100%.)
$(p + q)^2 = 1$ so $p^2 + 2pq + q^2 = 1$
The three terms of this binomial expansion indicate the frequencies of the three genotypes:
p^2 = frequency of AA (**homozygous dominant**)
2pq = frequency of Aa (**heterozygous**)
q^2 = frequency of aa (**homozygous recessive)**

Evolution and speciation

The **niche** of a species is **its role within the environment**. Species which share the same niche compete with each other and a better adapted species survive. The idea that better adapted species survive is the basis of **natural selection**.

Organisms are adapted to their environment in various ways:

- **Anatomical adaptations** are **physical adaptations**, either external or internal e.g. presence of loops of Henlé which allow desert mammals to produce concentrated urine and minimise water loss
- **Behavioural adaptations** are **changes in behaviour** which improve the organism's chance of survival e.g. mating calls
- **Physiological adaptations** are **processes inside an organism's body** that increase its chance of survival e.g. regulation of blood flow through the skin

Natural selection is the process in which **fitter individuals** who are betted adapted to the environment **survive and pass on the advantageous alleles to future generations**.
Evolution is the process by which the **frequency of alleles in a gene pool changes over time as a result of natural selection.**

Evolution via natural selection:

- There's a **variety of phenotypes** within a population
- An **environmental change occurs** and as a result of that the **selection pressure changes**
- Some individuals possess **advantageous alleles which give them a selective advantage and allow them to survive and reproduce**

- The **advantageous alleles are passed on to their offspring**
- Over time and **many generations**, **the frequency of alleles in a population changes** and this leads to evolution

Factors that can affect the evolution of a species:

- **Genetic drift** is a phenomenon where there is a small change in **allele frequency** which occurs as a result of the fact that **not all the individuals in a population reproduce**. This effect is amplified in **very small groups, isolated from the rest** of the population.

- **Genetic bottleneck** – rapid reduction in **population size** which has an effect on the **population size and genetic variation in future generations** such as that caused by a natural disaster.

- **Founder effect** – decrease in **genetic diversity** which occurs when the population **descends from a small number of ancestors**

Speciation is the process by which new species arise after a **population becomes separated** and **cannot interbreed**. For instance, **allopatric speciation** is caused by a **physical barrier**. As the two groups become separated and reproductively isolated as a result, the **gene flow is reduced**. Each group experiences a different selection pressure as the environment they live in is different. Over time, the frequency of alleles changes through **natural selection** and the two parts of the population **can no longer interbreed and become separate species**.

Another type of speciation is **sympatric speciation** where new species evolve from a **single ancestral species** when **inhabiting the same geographic region**, for example as a result of a **chromosomal error during cell division** which leads to **reproductive isolation**.

Artificial selection

Artificial selection is the process where **selection pressures are artificially created** by humans thus allowing the **breeding of the desired characteristics**.

An example of artificial selection is the **dairy cow**. The **milk yield** of each cow is **measured** and **recorded** to **identify the cows with the highest milk yields**. This enables the identification of the best quality bulls. The cows with the highest yields are given **hormones to increase their egg production**. The eggs are **fertilised in vitro** and subsequently implanted into **surrogate mothers.**

Another example of artificial selection is **bread wheat.** Modern bread wheat is **hexaploid**, with 42 chromosomes meaning the cells need to be bigger in order to accommodate for the large number of chromosomes. Domesticated wheat is a **hybrid** composed of three genomes **A^UA^U, BB and DD**. A^UA^U comes from wild wheat species whereas genome BB originates from wild emmer wheat.

Principles of DNA sequencing

DNA sequencing begins with the process of **mapping** where the existing information about the genome is used to **identify the locus of a particular gene within the genome**. The DNA is **fragmented** with the use of **restriction enzymes** and the fragments produced are inserted into **bacterial artificial chromosomes**. This step results in the formation of a **genomic DNA library.** The fragments obtained from the **bacterial cultures** are again broken down into smaller fragments with the use of **restriction enzymes and sequenced** with the use of the **chain-termination**. This technique was developed by Sanger and is based on **selective incorporation of chain terminating nucleotides** into a growing chain by **DNA polymerase during replication.**

It occurs as following:

- The DNA sample is divided into **four separate sequencing reactions** which contain all **four standard** nucleotides, DNA polymerase, primers required for replication and **modified nucleotides** which have been **fluorescently labelled** for ease of identification.
- When a modified nucleotide is incorporated into a growing chain, replication is **terminated**
- DNA fragments of **different lengths** are formed across the **reaction vessels**
- **High resolution electrophoresis** is used to **separate the fragments by size** – single base differences can be seen
- The fragments are **visualised under UV light**, thus enabling the **base sequence** to be **read from the bottom of the gel upwards**

The **rapid advancement** of techniques used in sequencing increased the speed of sequencing and allowed **whole genome sequencing**, that is. **high-throughput sequencing.**

Gene sequencing allows for **genome-wide comparisons** between i**ndividuals and between species**. Comparing genomes between species is significant as it allows **evolutionary relationship**s between species to be determined, and it is also beneficial to **medical research.** Comparing genomes of individuals enables differences to be identified which can then be used for **development of personalised medicine** tailored to a particular genome, as well as in studies of **human diseases.**

Apart from allowing genome-wide comparisons to be made, gene sequencing has allowed for the **sequences of amino acids in polypeptides to be predicted** and has allowed for the **development of synthetic biology.**

DNA profiling

DNA profiling is a forensic technique used to **identify individuals by characteristics of their DNA**. It can also be used to **determine genetic relationships** between organisms. Main techniques used in DNA profiling are:

- **Polymerase chain reaction known as PCR** which is used to amplify the DNA by making millions of identical copies of a given DNA sample. It occurs as following:

  1) A reaction mixture is set up by mixing the **DNA sample, primers, free nucleotides and DNA polymerase** which is the enzyme involved in creating new DNA strands.
  2) The mixture is then **heated to 95 degrees** to break the hydrogen bonds and to separate the two strands.
  3) The mixture is then **cooled to a temperature between 50-65 degrees** depending on the type of primers used, so that they can bind to the strands
  4) Temperature is increased to about **70 degrees** as this is the optimum temperature DNA polymerase works at.
  5) **DNA polymerase** creates a copy of the sample by **complementary base pairing using the free nucleotides**
  6) **This cycle is repeated around 30 times** and gives rise to an amount of DNA sufficient to create a DNA profile.

- **Gel electrophoresis** is a process used to **separate the DNA fragments and proteins according to their size using an electric current**. The diagram on the left demonstrates how the process is carried out.
- **The DNA fragments move through the gel due to their negative charge**

Figure S-2: Gel Electrophoresis

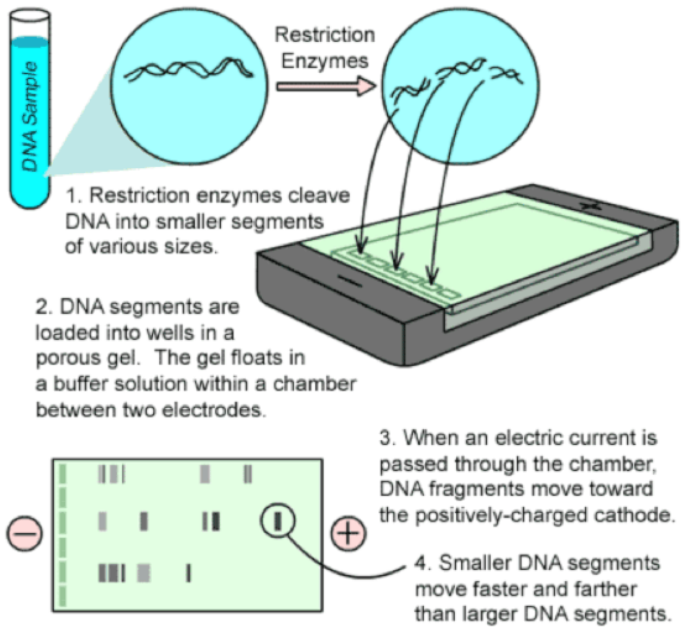

Figure 2 Living Environment course

Genetic engineering

Restriction enzymes cut DNA at specific base sequences and so are extremely useful in genetic engineering whereby genes from one species can be added to a different species. To carry the DNA into a host cell it must be placed into a **vector**. **Plasmids** are the most common vectors but **viruses** are also used.

Isolated DNA fragments can be placed in plasmids in a following way:

- Plasmid and gene are cut with the same restriction enzyme to create **complementary ends.** If sticky ends are missing, they can be added
- The fragments are **incubated with the plasmids.** If a plasmid takes up the insert, base pairing takes place between the complementary ends which are then **sealed with the use of DNA ligase which forms phosphodiester linkages**
- A **recombinant DNA** molecule is created

In the formation of **transgenic microorganisms**, **electroporation** is used to stimulate bacterial cells to take up plasmids. Electroporation facilitates the process by **increasing the permeability of bacterial membranes thus increasing the chance of success**. This is achieved via the use of **calcium salts and rapid temperature change from 0 to 40 degrees**. Bacteria which have successfully taken up a plasmid can be identified with the help of **marker genes**. For instance, if a plasmid contains an **antibiotic resistance gene**, the bacteria will be resistant to the antibiotic, and if grown on the media, **only the bacteria which have been successfully transformed will survive.**

Other types of vectors include **bacteriophages, liposomes and yeast artificial chromosome**.

Gene therapy

Gene therapy is the insertion of a normal allele into target cells to **replace a faulty allele**, such as the allele which causes a cystic fibrosis. There are two types of gene therapy: **somatic gene therapy** where the allele is introduced to the **target cells only** and **germ line gene therapy** where the allele is introduced to **embryonic cells**, thus meaning **every cell contains the normal allele**. **Somatic** gene therapy is a **short-term solution only** and needs to be repeated, whereas **germ-line** therapy is a **permanent solution which will be passed down to the offspring.**

There are many **ethical considerations** regarding genetic engineering. Benefits of genetic engineering include insect resistance in crops such as soya and genetically used animals used to produce pharmaceuticals. Some people object to genetic engineering due to the potential effect it might have on the environment, or because of the idea that genetically modified seeds would not be as easily available to poorer farmers.