AQA A2 Biology Unit 5

Contents

Specification		2
Human Nervous system	Nerve Cells	4
	The Nerve Impulse	6
	Synapses	10
	Receptors	14
	Muscle	18
	Animal Responses	25
	Control of Heart Rate	28
	The Hormone System	30
	Homeostasis	33
	Temperature Homeostasis	34
	Blood Glucose Homeostasis	38
	Control of Mammalian Oestrus	42
	Plant Responses	44
Molecular Genetics	The Genetic Code	47
	Transcription	50
	Translation	52
	Gene Mutations	54
	Stem Cells	58
	Control of Gene Expression	62
	Biotechnology	66
	PCR	69
	DNA sequencing	72
	Southern Blot	75
	Genetic Fingerprinting	77
	In vivo cloning	80
	Genetically Modified Organisms	85
	Gene Therapy	89
	Genetic Screening and Counselling	92

Neil C Millar (nmillar@ntlworld.co.uk)
Head of Biology, Heckmondwike Grammar School
High Street, Heckmondwike, WF16 0AH
Jan 2012

Biology Unit 5 Specification

Control Systems

Organisms increase their chance of survival by responding to changes in their environment.

The Nerve Impulse

The structure of a myelinated motor neurone. The establishment of a resting potential in terms of differential membrane permeability, electrochemical gradients and the movement of sodium and potassium ions. Changes in membrane permeability lead to depolarisation and the generation of an action potential. The all-or-nothing principle. The passage of an action potential along non-myelinated and myelinated axons, resulting in nerve impulses. The nature and importance of the refractory period in producing discrete impulses. Factors affecting the speed of conductance: myelination and saltatory conduction; axon diameter; temperature.

Synapses

The detailed structure of a synapse and of a neuromuscular junction. The sequence of events involved in transmission across a cholinergic synapse and across a neuromuscular junction. Explain unidirectionality, temporal and spatial summation and inhibition. Predict and explain the effects of specific drugs on a synapse (recall of the names and mode of action of individual drugs will not be required).

Receptors

Receptors only respond to specific stimuli. The creation of a generator potential on stimulation.

- The basic structure of a Pacinian corpuscle as an example of a receptor. Stimulation of the Pacinian corpuscle membrane produces deformation of stretch-mediated sodium channels leading to the establishment of a generator potential.
- Differences in sensitivity and visual acuity as explained by differences in the distribution of rods and cones and the connections they make in the optic nerve.

Muscle

The sliding filament theory of muscle contraction. Gross and microscopic structure of skeletal muscle. The ultrastructure of a myofibril. The roles of actin, myosin, calcium ions and ATP in myofibril contraction. The role of ATP and phosphocreatine in providing the energy supply during muscle contraction. The structure, location and general properties of slow and fast skeletal muscle fibres

Animal Responses

A simple reflex arc involving three neurones. The importance of simple reflexes in avoiding damage to the body. Taxes and kineses as simple responses that can maintain a mobile organism in a favourable environment. Investigate the effect of external stimuli on taxes and kineses in suitable organisms.

Control of Heart Rate

The role of receptors, the autonomic nervous system and effectors in controlling heart rate.

Hormones

Nerve cells pass electrical impulses along their length. They stimulate their target cells by secreting chemical neurotransmitters directly on to them. This results in rapid, short-lived and localised responses. Mammalian hormones are substances that stimulate their target cells via the blood system. This results in slow, long-lasting and widespread responses. Histamine and prostaglandins are local chemical mediators released by some mammalian cells that affect only cells in their immediate vicinity.

Homeostasis

Homeostasis in mammals involves physiological control systems that maintain the internal environment within restricted limits.

Negative and Positive feedback

- Negative feedback restores systems to their original level. The possession of separate mechanisms involving negative feedback controls departures in different directions from the original state, giving a greater degree of control.
- Positive feedback results in greater departures from the original levels. Positive feedback is often associated with a breakdown of control systems, e.g. in temperature control.

Interpret diagrammatic representations of negative and positive feedback.

Temperature Homeostasis

The importance of maintaining a constant core temperature and constant blood pH in relation to enzyme activity. The contrasting mechanisms of temperature control in an ectothermic reptile and an endothermic mammal. Mechanisms involved in heat production, conservation and loss. The role of the hypothalamus and the autonomic nervous system in maintaining a constant body temperature in a mammal.

Blood Glucose Homeostasis

The factors that influence blood glucose concentration. The importance of maintaining a constant blood glucose concentration in terms of energy transfer and water potential of blood. The role of the liver in glycogenesis and gluconeogenesis. The role of insulin and glucagon in controlling the uptake of glucose by cells and in activating enzymes involved in the interconversion of glucose and glycogen. The second messenger model of adrenaline and glucagon action. The effect of adrenaline on glycogen breakdown and synthesis.Types I and II diabetes and control by insulin and manipulation of the diet.

Control of Mammalian Oestrus

The mammalian oestrous cycle is controlled by FSH, LH, progesterone and oestrogen. The secretion of FSH, LH, progesterone and oestrogen is controlled by interacting negative and positive feedback loops. Interpret graphs showing the blood concentrations of FSH, LH, progesterone and oestrogen during a given oestrous cycle.

Plant Responses

Tropisms as responses to directional stimuli that can maintain the roots and shoots of flowering plants in a favourable environment. In flowering plants, specific growth factors diffuse from growing regions to other tissues. They regulate growth in response to directional stimuli. The role of indoleacetic acid (IAA) in controlling tropisms in flowering plants.

Genetics

The Genetic Code

The genetic code as base triplets in mRNA which code for specific amino acids. The genetic code is universal, non-overlapping and degenerate. The structure of molecules of messenger RNA (mRNA) and transfer RNA (tRNA). Candidates should be able to compare the structure and composition of DNA, mRNA and tRNA

Protein Synthesis

- Transcription as the production of mRNA from DNA. The role of RNA polymerase. The splicing of pre-mRNA to form mRNA in eukaryotic cells.
- Translation as the production of polypeptides from the sequence of codons carried by mRNA. The role of ribosomes and tRNA.

Show understanding of how the base sequences of nucleic acids relate to the amino acid sequence of polypeptides, when provided with suitable data. Interpret data from experimental work investigating the role of nucleic acids. Recall of specific codons and the amino acids for which they code, and of specific experiments, will not be tested.

Gene Mutations

Gene mutations might arise during DNA replication. The deletion and substitution of bases. Gene mutations occur spontaneously. The mutation rate is increased by mutagenic agents. Some mutations result in a different amino acid sequence in the encoded polypeptide. Due to the degenerate nature of the genetic code, not all mutations result in a change to the amino acid sequence of the encoded polypeptide. Evaluate the effect on diagnosis and treatment of disorders caused by hereditary mutations and those caused by acquired mutations.

Mutations and Cancer

The rate of cell division is controlled by proto-oncogenes that stimulate cell division and tumour suppressor genes that slow cell division. A mutated proto-oncogene, called an oncogene, stimulates cells to divide too quickly. A mutated tumour suppressor gene is inactivated, allowing the rate of cell division to increase. Interpret information relating to the use of oncogenes and tumour suppressor genes in the prevention, treatment and cure of cancer.

Stem Cells

Totipotent cells are cells that can mature into any body cell. During development, totipotent cells translate only part of their DNA, resulting in cell specialisation.

- In mature animals only a few totipotent cells, called stem cells, remain. These can be used in treating some genetic disorders. Evaluate the use of stem cells in treating human disorders.
- In mature plants, many cells remain totipotent. They have the ability to develop in vitro into whole plants or into plant organs when given the correct conditions. Interpret data relating to tissue culture of plants from samples of totipotent cells

Regulation of Gene Expression

- Transcription of target genes is stimulated only when specific transcriptional factors move from the cytoplasm into the nucleus. The effect of oestrogen on gene transcription.
- Small interfering RNA (siRNA) as a short, double-strand of RNA that interferes with the expression of a specific gene. Interpret data provided from investigations into gene expression.

Genetic Engineering Techniques

- The use of restriction endonucleases to cut DNA at specific, palindromic recognition sequences. The importance of "sticky ends".
- conversion of mRNA to cDNA, using reverse transcriptase
- Interpret data showing the results of gel electrophoresis to separate DNA fragments.
- The base sequence of a gene can be determined by restriction mapping and DNA sequencing.
- The use of the polymerase chain reaction (PCR) in cloning DNA fragments.
- The use of labelled DNA probes and DNA hybridisation to locate specific genes. Understand the principles of these methods. Be aware that methods are continuously updated and automated.
- The technique of genetic fingerprinting in analysing DNA fragments that have been cloned by PCR, and its use in determining genetic relationships and in determining the genetic variability within a population. Explain the biological principles that underpin genetic fingerprinting techniques. An organism's genome contains many repetitive, non-coding base sequences. The probability of two individuals having the same repetitive sequences is very low. Explain why scientists might use genetic fingerprints, in the fields of forensic science, medical diagnosis, animal and plant breeding.
- The use of ligases to insert DNA fragments into vectors, which are then transferred into host cells.
- The identification and growth of transformed host cells to clone the desired DNA fragments.

The relative advantages of in vivo and in vitro cloning.

Genetically Modified Organisms

The use of recombinant DNA technology to produce transformed organisms that benefit humans. Interpret information relating to the use of recombinant DNA technology. Evaluate the ethical, moral and social issues associated with the use of recombinant technology in agriculture, in industry and in medicine. Balance the humanitarian aspects of recombinant DNA technology with the opposition from environmentalists and anti-globalisation activists.

Gene Therapy

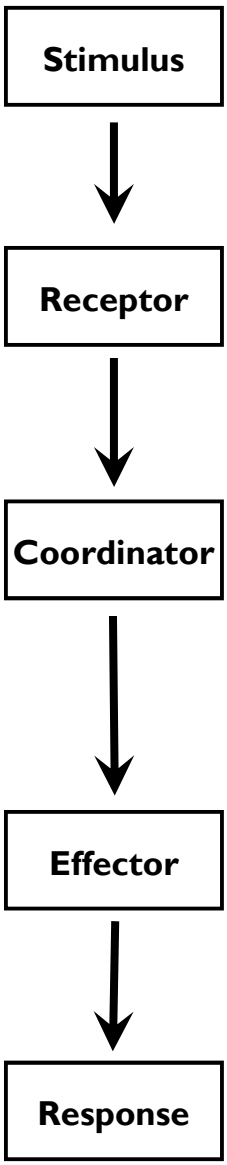
The use of gene therapy to supplement defective genes. Be able to evaluate the effectiveness of gene therapy.

Genetic Screening

Many human diseases result from mutated genes or from genes that are useful in one context but not in another, e.g. sickle cell anaemia. DNA sequencing and PCR are used to produce DNA probes that can be used to screen patients for clinically important genes. The use of this information in genetic counselling, e.g., for parents who are both carriers of defective genes and, in the case of oncogenes, in deciding the best course of treatment for cancers.

The Human Nervous System

Humans, like all living organisms, can respond to changes in the environment and so increase survival. Humans have two control systems to do this: the nervous system and the endocrine (hormonal) system. We'll look at the endocrine system later, but first we'll look at the nervous system. The human nervous system controls everything from breathing and standing upright, to memory and intelligence. It has three parts: detecting stimuli; coordinating; and effecting a response:



Stimuli are changes in the external or internal environment, such as light waves, pressure or blood sugar. Humans can detect at least nine external stimuli and dozens of internal stimuli, so the commonly-held believe that humans have just five senses is obviously very wide of the mark!

Receptor cells detect stimuli. Receptor cells are often part of sense organs, such as the ear, eye or skin. Receptor cells all have special receptor proteins on their cell membranes that actually do the sensing, so "receptor" can confusingly mean a protein, a cell or a group of cells.

The coordinator is the name given to the network of interneurones connecting the sensory and motor systems. It can be as simple as a single interneurone in a reflex arc, or as complicated as the human brain. Its job is to receive impulses from sensory neurones and transmit impulses to motor neurones.

Effectors are the cells that effect a response. In humans there are just two kinds: muscles and glands. Muscles include skeletal muscles, smooth muscles and cardiac muscle, and they cause all movements in humans, such as walking, talking, breathing, swallowing, peristalsis, vasodilation and giving birth. Glands can be exocrine – secreting liquids to the outside (such as tears, sweat, mucus, enzymes or milk); or endocrine – secreting hormones into the bloodstream.

Responses aid survival. They include movement of all kinds, secretions from glands and all behaviours such as stalking prey, communicating and reproducing.

We're going to be looking at each of these stages in turn, but first we'll look at the cells that comprise the nervous system.

Nerve Cells

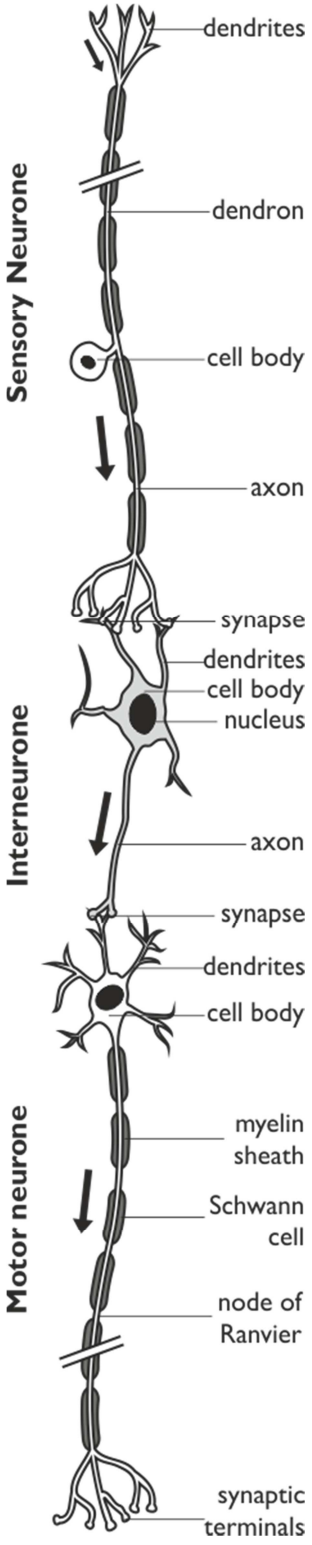
The nervous system composed of nerve cells, or neurones. A neurone has a cell body with extensions leading off it. Several dendrons carry nerve impulses towards the cell body, while a single long axon carries the nerve impulse away from the cell body. Axons and dendrons are only 10µm in diameter but can be up to 4m in length in a large animal (a piece of spaghetti the same shape would be 400m long)! A nerve is a discrete bundle of several thousand neurone axons.

Nerve impulses are passed from the axon of one neurone to the dendron of another at a synapse. Numerous dendrites provide a large surface area for connecting with other neurones.

Most neurones also have many companion cells called Schwann cells, which are wrapped around the axon many times in a spiral to form a thick lipid layer called the myelin sheath. The myelin sheath provides physical protection and electrical insulation for the axon. There are gaps in the sheath, called nodes of Ranvier, which we'll examine later.

Humans have three types of neurone:

- Sensory neurones have long dendrons and transmit nerve impulses from sensory receptors all over the body to the central nervous system.
- Effector neurones (also called motor neurones) have long axons and transmit nerve impulses from the central nervous system to effectors (muscles and glands) all over the body.
- Interneurones (also called connector neurones or relay neurones) are much smaller cells, with many interconnections. They comprise the central nervous system. 99.9% of all neurones are interneurones.

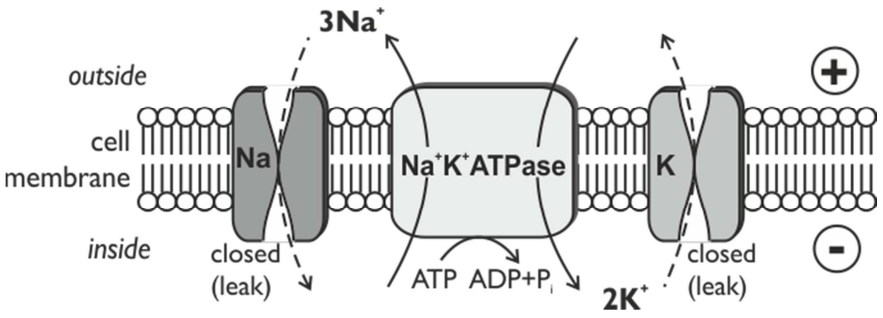


The Nerve Impulse

Neurones transmit simple on/off signals called impulses (never talk about nerve signals or messages). These impulses are due to events in the cell membrane, so to understand the nerve impulse we need to revise some properties of cell membranes.

The Membrane Potential

All animal cell membranes contain a protein pump called the Na+K+ATPase. This uses the energy from ATP splitting to simultaneously pump 3 sodium ions out of the cell and 2 potassium ions in. If this was to continue unchecked there would be no sodium or potassium ions left to pump, but there are also sodium and potassium ion channels in the membrane. These channels are normally closed, but even when closed, they "leak", allowing sodium ions to leak in and potassium ions to leak out, down their respective concentration gradients.



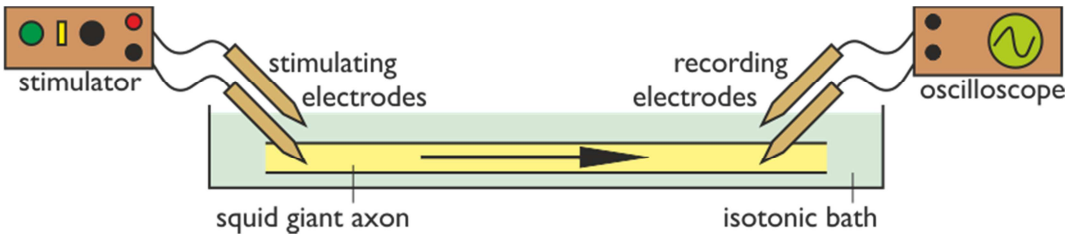
The combination of the Na^+K^+ATPase pump and the leak channels cause a stable imbalance of Na^+ and K^+ ions across the membrane. This imbalance causes a potential difference across all animal cell membranes, called the membrane potential. The membrane potential is always negative inside the cell, and varies in size from –20 to –200mV in different cells and species. The Na^+K^+ATPase is thought to have evolved as an osmoregulator to keep the internal water potential high and so stop water entering animal cells and bursting them. Plant cells don't need this pump as they have strong cells walls to prevent bursting (which is why plants never evolved a nervous system).

The Action Potential

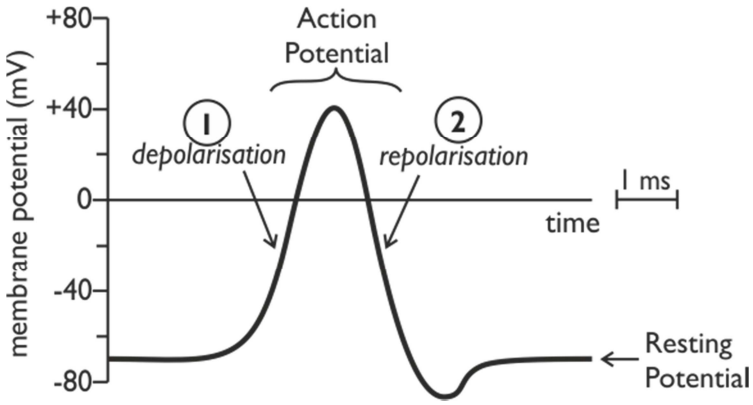
In nerve and muscle cells the membranes are electrically excitable, which means that they can change their membrane potential, and this is the basis of the nerve impulse. The sodium and potassium channels in these cells are voltage gated, which means that they can open and close depending on the size of the voltage across the membrane.

The nature of the nerve impulse was discovered by Hodgkin, Huxley and Katz in Plymouth in the 1940s, for which work they received a Nobel Prize in 1963. They used squid giant neurones, whose axons are almost 1 mm in diameter (compared to 10 µm normally), big enough to insert wire electrodes so that they could measure the potential difference across the cell membrane. In a typical experiment they would apply

an electrical pulse at one end of an axon and measure the voltage changes at the other end, using an oscilloscope:

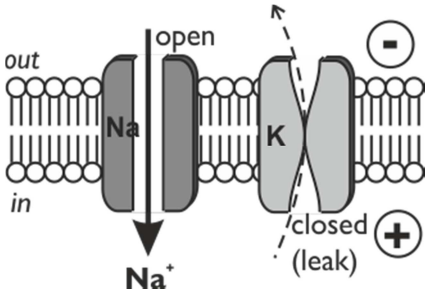


The normal membrane potential of these nerve cells is –70mV (inside the axon), and since this potential can change in nerve cells it is called the resting potential. When a stimulating pulse was applied a brief reversal of the membrane potential, lasting about a millisecond, was recorded. This brief reversal of the membrane potential is actually the nerve impulse, and is also called the action potential:

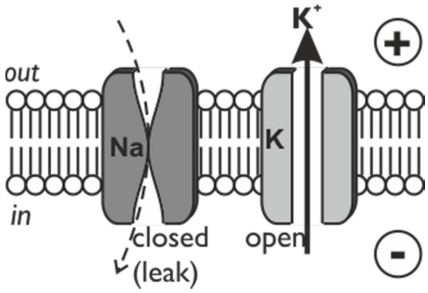


The action potential has 2 phases called depolarisation and repolarisation.

1. Depolarisation. The sodium channels open for 0.5ms, causing sodium ions to diffuse **in** down their gradient, and making the inside of the cell more positive. This is a depolarisation because the normal voltage polarity (negative inside) is reversed (becomes positive inside).



2. Repolarisation. The potassium channels open for 0.5ms, causing potassium ions to diffuse **out** down their concentration gradient, making the inside more negative again. This is a repolarisation because it restores the original polarity.



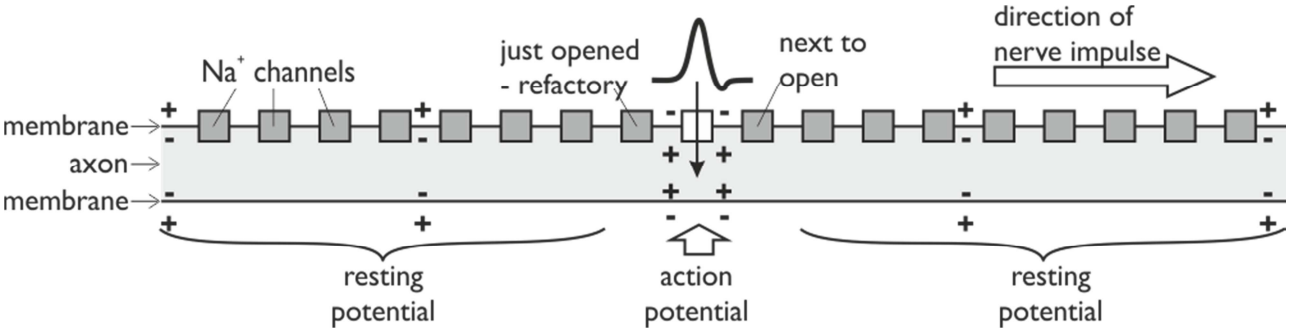
Since both channels are voltage-gated, they are triggered to open by changes in the membrane potential itself. The sodium channel opens at–30mV and the potassium channel opens at 0V. The Na^+K^+ATPase pump runs continuously, restoring the resting concentrations of sodium and potassium ions.

How do Nerve Impulses Start?

In the squid experiments the action potential was initiated by the stimulating electrodes. In living cells they are started by receptor cells. These all contain special receptor proteins that sense the stimulus. The receptor proteins are sodium channels that are not voltage-gated, but instead are gated by the appropriate stimulus (directly or indirectly). For example chemical-gated sodium channels in tongue taste receptor cells open when a certain chemical in food binds to them; mechanically-gated ion channels in the hair cells of the inner ear open when they are distorted by sound vibrations; and so on. In each case the correct stimulus causes the sodium channel to open; which causes sodium ions to diffuse into the cell; which causes a depolarisation of the membrane potential, which affects the voltage-gated sodium channels nearby and starts an action potential.

How are Nerve Impulses Propagated?

Once an action potential has started it is moved (propagated) along an axon automatically. The local reversal of the membrane potential is detected by the surrounding voltage-gated ion channels, which open when the potential changes enough.

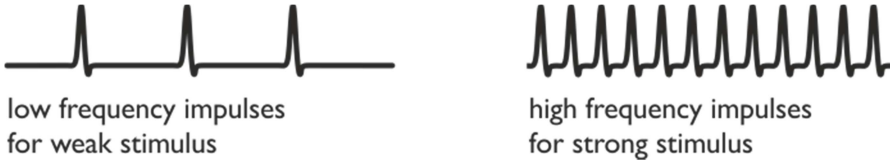


The ion channels have two other features that help the nerve impulse work effectively:

- After an ion channel has opened, it needs a "rest period" before it can open again. This is called the refractory period, and lasts about 2ms. This means that, although the action potential affects all other ion channels nearby, the upstream ion channels cannot open again since they are in their refractory period, so only the downstream channels open, causing the action potential to move one way along the axon.
- The ion channels are either open or closed; there is no half-way position. This means that the action potential always reaches +40mV as it moves along an axon, and it is never attenuated (reduced) by long axons. In other word the action potential is all-or-nothing.

How can Nerve Impulses convey strength?

How do impulses convey the strength of the stimulus? Since nerve impulses are all-or-nothing, they cannot vary in size. Instead, the strength of stimulus is indicated by the frequency of nerve impulses. A weak stimulus (such as dim light, a quiet sound or gentle pressure) will cause a low frequency of nerve impulses along a sensory neurone (around 10Hz). A strong stimulus (such as a bright light, a loud sound or strong pressure) will cause a high frequency of nerve impulses along a sensory neurone (up to 100Hz).



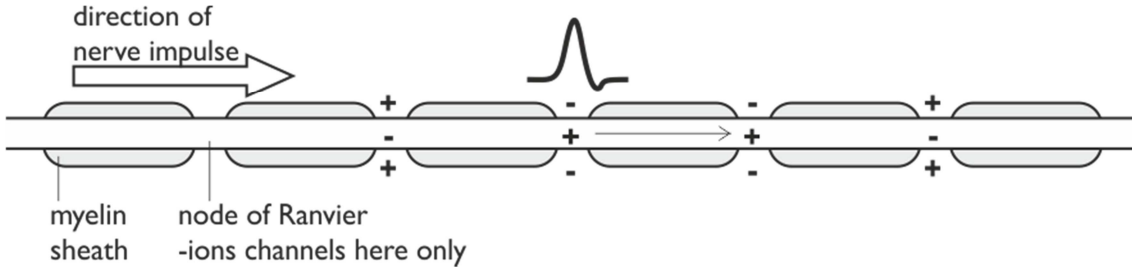
low frequency impulses for weak stimulus

high frequency impulses for strong stimulus

How Fast are Nerve Impulses?

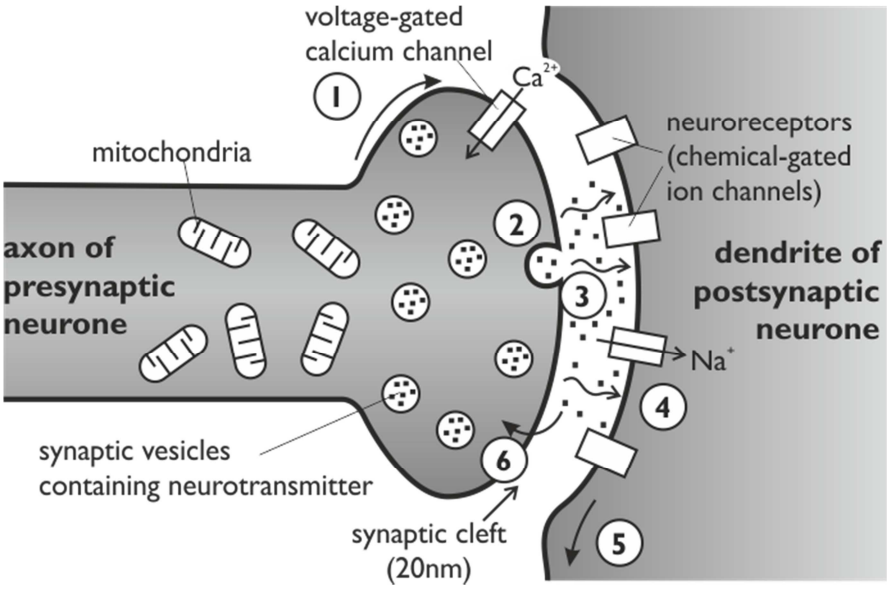
Action potentials can travel along axons at speeds of 0.1-100 ms^{-1}. This means that nerve impulses can get from one part of a body to another in a few milliseconds, which allows for fast responses to stimuli. (Impulses are much slower than electrical currents in wires, which travel at close to the speed of light, $3x10^8$ ms^{-1}.) The speed is affected by 3 factors:

- **Temperature**. The higher the temperature, the faster the speed. So homeothermic (warm-blooded) animals have faster responses than poikilothermic (cold-blooded) ones.
- **Axon diameter**. The larger the diameter, the faster the speed. So marine invertebrates, which live at temperatures close to 0°C, have developed thick axons to speed up their responses. This explains why squid have their giant axons.
- **Myelin sheath**. Only vertebrates have a myelin sheath surrounding their neurones. The voltage-gated ion channels are found only at the nodes of Ranvier, and between the nodes the myelin sheath acts as a good electrical insulator. The action potential can therefore jump large distances from node to node (1 mm), a process that is called saltatory propagation. This increases the speed of propagation dramatically, so while nerve impulses in unmyelinated neurones have a maximum speed of around 1 ms^{-1}, in myelinated neurones they travel at 100 ms^{-1}.



Synapses

The junction between two neurones is called a synapse. An action potential cannot cross the gap between the neurones (called the synaptic cleft), and instead the nerve impulse is carried by chemicals called neurotransmitters. These chemicals are made by the cell that is sending the impulse (the pre-synaptic neurone) and stored in synaptic vesicles at the end of the axon. The cell that is receiving the nerve impulse (the post-synaptic neurone) has chemical-gated ion channels in its membrane, called neuroreceptors. These have specific binding sites for the neurotransmitters.



1. At the end of the pre-synaptic neurone there are voltage-gated calcium channels. When an action potential reaches the synapse these channels open, causing calcium ions to diffuse into the cell down their concentration gradient.
2. These calcium ions cause the synaptic vesicles to fuse with the cell membrane, releasing their contents (the neurotransmitter chemicals) by exocytosis.
3. The neurotransmitters diffuse across the synaptic cleft.
4. The neurotransmitter binds to the neuroreceptors in the post-synaptic membrane, causing the ion channels to open. In the example shown these are sodium channels, so sodium ions diffuse in down their gradient.
5. This causes a depolarisation of the post-synaptic cell membrane, called the post-synaptic potential (PSP), which may initiate an action potential.
6. The neurotransmitter must be removed from the synaptic cleft to stop the synapse being permanently on. This can be achieved by breaking down the neurotransmitter by a specific enzyme in the synaptic cleft (e.g. the enzyme cholinesterase breaks down the neurotransmitter acetylcholine). The breakdown products are absorbed by the pre-synaptic neurone by endocytosis and used to re-synthesise more neurotransmitter, using energy from the mitochondria. Alternatively the neurotransmitter may be absorbed intact by the pre-synaptic neurone using active transport.

Different Types of Synapse

The human nervous system uses a number of different neurotransmitters and neuroreceptors, and they don't all work in the same way. We can group synapses into 5 types:

1. Excitatory Ion-channel Synapses.

These synapses have neuroreceptors that are sodium (Na^+) channels. When the channels open, positive ions diffuse in, causing a local **depolarisation** called an excitatory postsynaptic potential (EPSP) and making an action potential **more** likely. This was the kind of synapse described on the previous page. Typical neurotransmitters in these synapses are acetylcholine, glutamate or aspartate.

2. Inhibitory Ion-channel Synapses.

These synapses have neuroreceptors that are chloride (Cl^-) channels. When the channels open, negative ions diffuse in causing a local **hyperpolarisation** called an inhibitory postsynaptic potential (IPSP) and making an action potential **less** likely. So with these synapses an impulse in one neurone can inhibit an impulse in the next. Typical neurotransmitters in these synapses are glycine or GABA.

3. Non-channel Synapses.

These synapses have neuroreceptors that are not channels at all, but instead are membrane-bound enzymes. When activated by the neurotransmitter, they catalyse the production of a "messenger chemical" (e.g. Ca^{2+}) inside the cell, which in turn can affect many aspects of the cell's metabolism. In particular they can alter the number and sensitivity of the ion channel receptors in the same cell. These synapses are involved in slow and long-lasting responses like learning and memory. Typical neurotransmitters are adrenaline, noradrenaline (NB adrenaline is called epinephrine in America), dopamine, serotonin, endorphin, angiotensin, and acetylcholine.

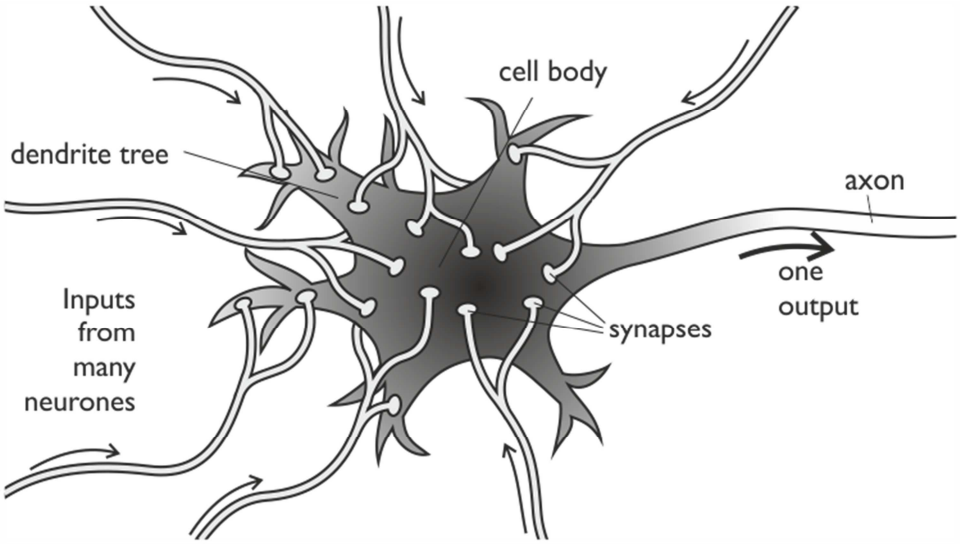
4. Neuromuscular Junctions.

These are the synapses formed between effector neurones and muscle cells. They always use the neurotransmitter acetylcholine, and are always excitatory. We shall look at these when we do muscles. Effector neurones also form specialised synapses with secretory cells.

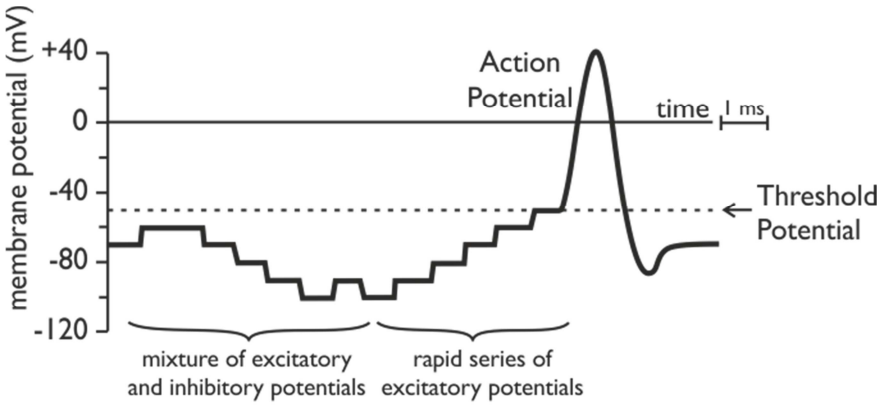
5. Electrical Synapses.

In these synapses the membranes of the two cells actually touch, and they share proteins. This allows the action potential to pass directly from one membrane to the next without using a neurotransmitter. They are very fast, but are quite rare, found only in the heart and the eye.

Synaptic Integration



One neurone can have hundreds (or even thousands) of synapses on its cell body and dendrites. This arrangement is called axon convergence, and it means the post-synaptic neurone has many inputs but only one output, through its axon. Some of these synapses will be excitatory and some will be inhibitory, and each synapse will produce its own local voltage change, called a postsynaptic potential (PSP). The excitatory and inhibitory PSPs from all that cell's synapses sum together to form a grand postsynaptic potential (GPSP) in the neurone's membrane. Only if this GPSP is above a threshold potential will an action potential be initiated in the axon. This process is called synaptic integration or summation.



- Spatial summation is the summing of PSPs from different synapses over the cell body and dendrite tree
- Temporal summation is the summing of a sequence of PSPs at one synapse over a brief period of time.

Summation is the basis of the processing power in the nervous system. Neurones (especially interneurones) are a bit like logic gates in a computer, where the output depends on the state of one or more inputs. By connecting enough logic gates together you can make a computer, and by connecting enough neurones together to can make a nervous system, including a human brain.

Drugs and Synapses

Almost all drugs taken by humans (medicinal and recreational) affect the nervous system, especially synapses. Drugs can affect synapses in various ways, shown in this table:

Drug action	Effect	Examples
1. Mimic a neurotransmitter	stimulate a synapse	levodopa
2. Stimulate the release of a neurotransmitter	stimulate a synapse	cocaine, caffeine
3. Open a neuroreceptor channel	stimulate a synapse	alcohol, marijuana, salbutamol
4. Block a neuroreceptor channel	Inhibit a synapse	atropine, curare, opioids
5. Inhibit the breakdown enzyme	stimulate a synapse	DDT

Drugs that stimulate a synapse are called agonists, and those that inhibit a synapse are called antagonists. By designing drugs to affect specific synapses, drugs can be targeted at different parts of the nervous system. The following examples show how some common drugs work. You do not need to learn any of this, but you should be able to understand how they work.

- Caffeine, theophylline, amphetamines, ecstasy (MDMA) and cocaine all promote the release of neurotransmitter in excitatory synapses in the part of the brain concerned with wakefulness, so are stimulants.
- Alcohol, benzodiazepines (e.g. mogadon, valium, librium), barbiturates, and marijuana all activate the inhibitory neuroreceptors in the same part of the brain, so are tranquillisers.
- The narcotics or opioid group of drugs, which include morphine, codeine, opium, methadone and diamorphine (heroin), all block opiate receptors, blocking transmission of pain signals in the brain and spinal chord. The brain's natural endorphins appear to have a similar action.
- Parkinson's disease (shaking of head and limbs) is caused by lack of the neurotransmitter dopamine in the midbrain. The balance can be restored with levodopa, which mimics dopamine.
- Curare and α-bungarotoxin (both snake venoms) block the acetylcholine receptors in neuromuscular junctions and so relax skeletal muscle.
- Nerve gas and organophosphate insecticides (DDT) inhibit acetylcholinesterase, so acetylcholine receptors in neuromuscular junctions are always active, causing muscle spasms and death.
- Tetrodotoxin (from the Japanese puffer fish) blocks voltage-gated sodium channels, while tetraethylamonium blocks the voltage-gated potassium channel. Both are powerful nerve poisons.

Receptors

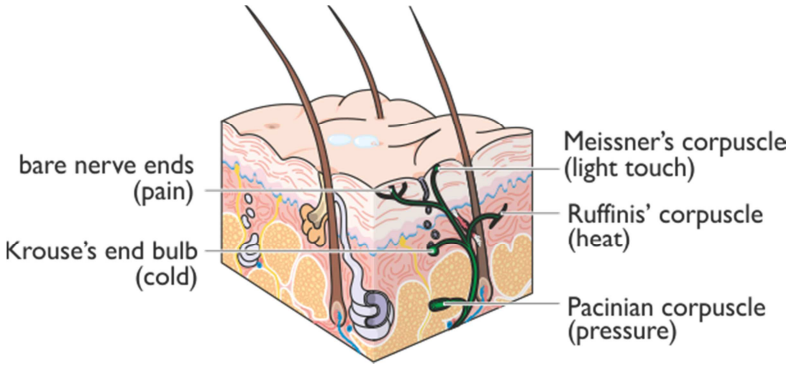
Receptor cells detect stimuli. Humans have over 20 different kinds of receptors. They can be classified as

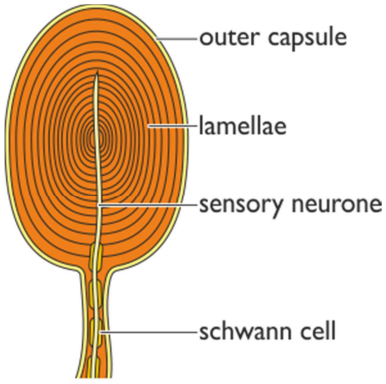
- photoreceptors– detecting light and other kinds of electromagnetic radiation
- mechanoreceptors– detecting movements, pressures, tension, gravity and sound waves
- chemoreceptors – detecting specific chemicals such as glucose, H^+ or pheromones
- thermoreceptors – detecting hot and cold temperatures
- Other animals have electroreceptors and magnetoreceptors.

In some receptors the receptor cell is the sensory neurone itself, while in others, there is a separate receptor cell that synapses with a sensory neurone. Receptor cells are often part of sense organs, such as the ear, eye or skin. Receptor cells all have special receptor proteins on their cell membranes that actually do the sensing, so "receptor" can confusingly mean a protein, a cell or a group of cells. We'll look at pressure receptors and light receptors in more detail.

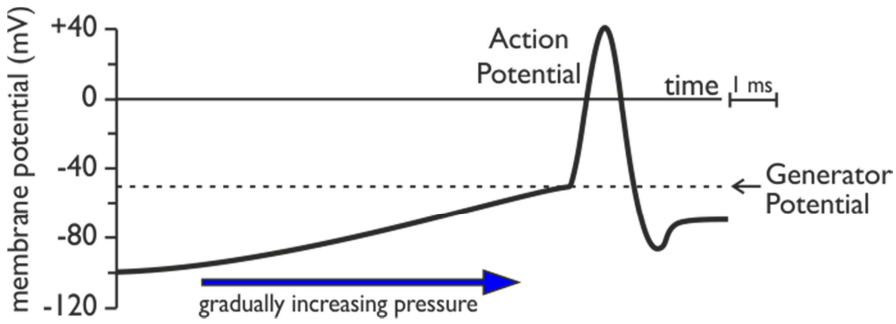
Receptors in the Skin

The skin is a major sense organ, containing at least six different types of receptors, detecting pressure, temperature and pain. Some are shown here.



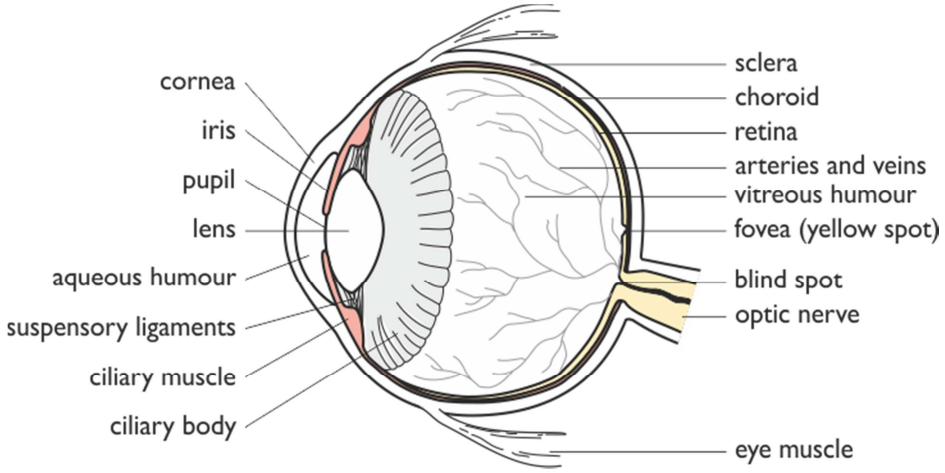


Pacinian corpuscles are mechanoreceptors found in the skin and in joints. They detect strong pressure and vibrations. Pacinian corpuscles look like microscopic onion bulbs, about 1mm long. Each corpuscle consists of a sensory neurone surrounded by a capsule of 20-60 layers of flattened Schwann cells and fluid, called lamellae. Pacinian corpuscles are situated deep in the skin, so are only sensitive to intense pressure, not light touch. Pressure distorts the neurone cell membrane, and opens mechanically-gated sodium channels. This allows sodium ions to diffuse in, causing a local depolarisation, called the generator (or receptor) potential. The stronger the pressure, the greater the generator potential, until it reaches a threshold, when an action potential is triggered.

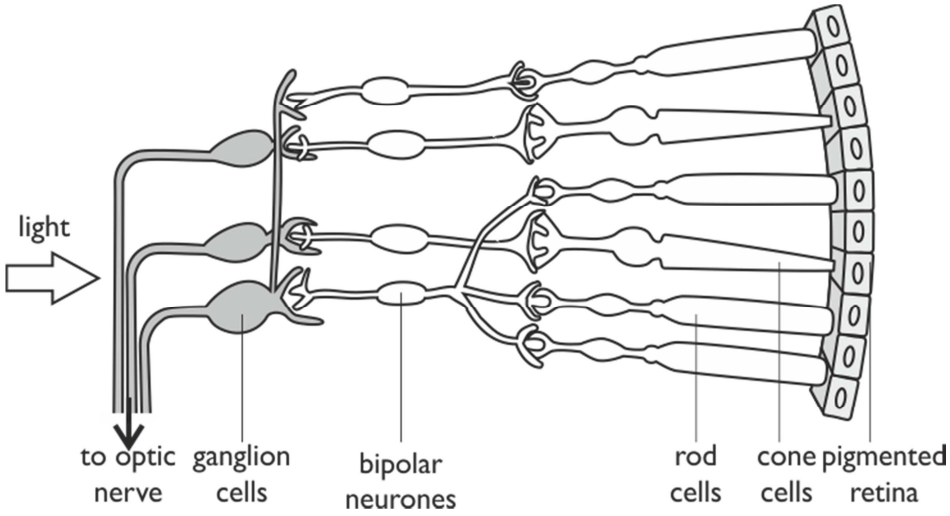


Receptors in the Eye

The eye is a complex sense organ, not just detecting light, but regulating its intensity and focussing it to form sharp images. The structure of the eye is shown here.



The actual detection of light is carried out by photoreceptor cells in the retina. The structure of the retina is shown in more detail in this diagram:

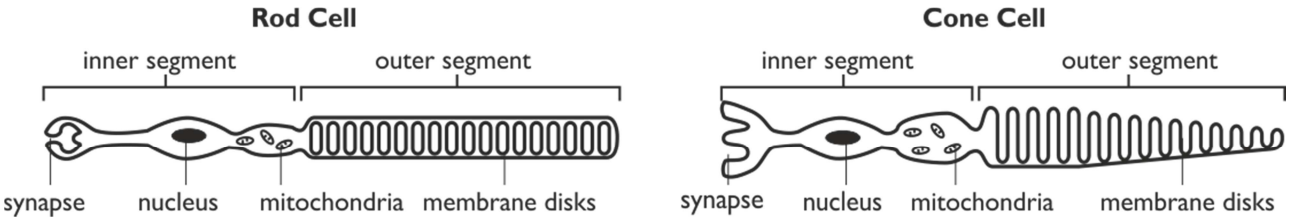


There are two kinds of photoreceptor cells in human eyes: rods and cones, and we shall look at the difference between these shortly. These rods and cones form synapses with special interneurones called bipolar neurones, which in turn synapse with sensory neurones called ganglion cells. The axons of these ganglion cells cover the inner surface of the retina and eventually form the optic nerve (containing about a million axons) that leads to the brain. Each cone cell is connected to one bipolar neurone, while rod cells are connected in groups of up to 100 to a single bipolar neurone. This linking together is called retinal convergence.

A surprising feature of the retina is that it is back-to-front (inverted). The photoreceptor cells are at the back of the retina, and the light has to pass through several layers of neurones to reach them. This is due to the evolutionary history of the eye, but in fact doesn't matter very much as the neurones are small and transparent. However, it does mean that the sensory neurones must pass through the retina where converge at the optic nerve, causing the blind spot.

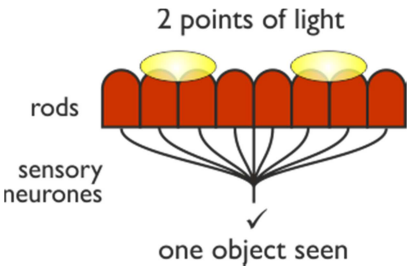
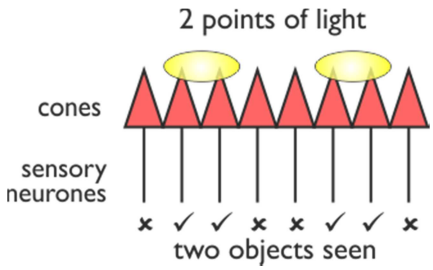
Rods and Cones

The photoreceptor cells have the following structures:



The detection of light is carried out on the membrane disks in the outer segments. These membranes contain thousands of molecules of rhodopsin, the photoreceptor protein. When illuminated, rhodopsin molecules change shape and can bind to sodium channels in the receptor cell membranes. This binding opens the sodium channels, allowing sodium ions to diffuse in, causing a local depolarisation. When enough sodium channels are open the depolarisation reaches a threshold, and an action potential is triggered in the rod or cone cell. This action potential is passed to the bipolar neurones and then to the ganglion cells (sensory neurones) in the retina.

The rods and cones serve two different functions as shown in this table:

	Rods	Cones
Shape	Outer segment is rod shaped	Outer segment is cone shaped
Density	10^9 cells per eye, distributed throughout the retina, so used for peripheral vision.	10^6 cells per eye, found mainly in the fovea, so can only detect images in centre of retina.
Colour	Only 1 type, so only monochromatic vision.	3 types (red green and blue), so are responsible for colour vision.
Connections	Many rods connected to one bipolar cell, giving retinal convergence.	Each cone connected to one bipolar cell, so no convergence.
Sensitivity (ability to detect low light intensity)	High sensitivity due to high concentration of rhodopsin, and to retinal convergence – one photon per rod will sum to cause an action potential. Used for night vision.	Low sensitivity due to lower concentration of rhodopsin and to no convergence – one photon per cone not enough to cause an action potential. Need bright light, so work best in the day.
Acuity (ability to resolve fine detail)	Poor acuity due to low density in periphery of retina, and retinal convergence (i.e. rods are not good at resolving fine detail).	Good acuity due high density in fovea and 1:1 connections with interneurones (i.e. cones are used for resolving fine detail such as reading).
		

Although there are far more rods than cones, we use cones most of the time because they have better acuity and can resolve colours. Since the cones are almost all found in the fovea, a 1mm² region of the retina directly opposite the lens, we constantly move our eyes so that images are always focused on the fovea. You can only read one word of a book at a time, but your eyes move so quickly that it appears that you can see much more. The more densely-packed the cone cells, the better the visual acuity. In the fovea of human eyes there are 160 000 cones per mm^2, while hawks have 1 million cones per mm^2, so they really do have far better acuity.

Muscle

ENGINE FOR SALE

Powerful ($100 Wkg^{-1}$)
Large Force ($200 kNm^{-2}$)
Very Efficient (>50%)
Silent Operation
Non-Polluting
Doesn't Overheat (38°C)
Uses a Variety of Fuels
Lasts a Lifetime
Good to Eat
£10-00 per kg at your Supermarket

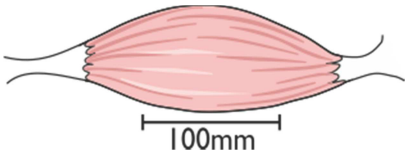
Muscle is indeed a remarkable tissue. In engineering terms it superior to any motor we have been able to invent, and it is responsible for almost all movements in animals.

There are three types of muscle:

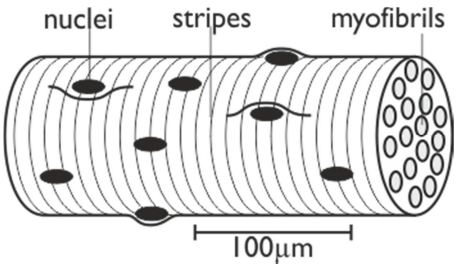
- Skeletal muscle (striated, voluntary)
  This is always attached to the skeleton, and is under voluntary control via the motor neurones of the somatic nervous system. It is the most abundant and best understood type of muscle. It can be subdivided into red (slow) muscle and white (fast) muscle (see pxx).
- Cardiac Muscle
  This is special type of red skeletal muscle. It looks and works much like skeletal muscle, but is not attached to skeleton, and is not under voluntary control.
- Smooth Muscle
  This is found in internal body organs such as the wall of the gut, the uterus, blood arteries and arterioles, the iris, ciliary body and glandular ducts. It is under involuntary control via the autonomic nervous system or hormones. Smooth muscle usually forms a ring, which tightens when it contracts, so doesn't need a skeleton to pull against.

There are also many examples of non-muscle motility, such as cilia (in the trachea and oviducts) and flagella (in sperm). These movements use different "motor proteins" from muscle, though they work in similar ways. Unless mentioned otherwise, the rest of this section is about skeletal muscle.

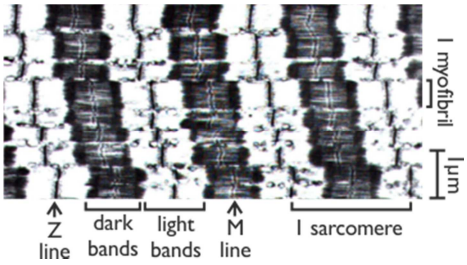
Muscle Structure



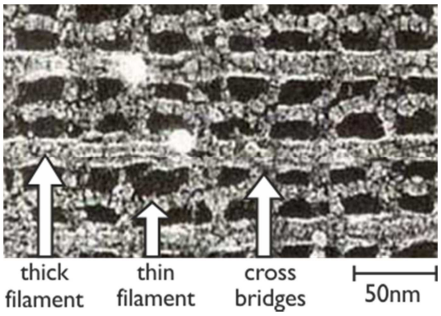
A single muscle (such as the biceps) contains around 1000 muscle fibres running the whole length of the muscle and joined together at the tendons.



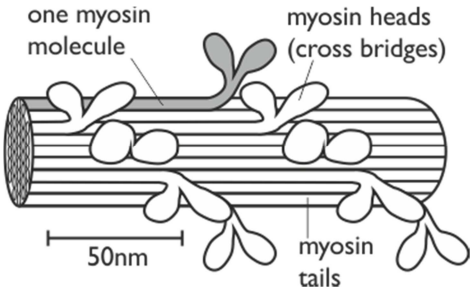
Each muscle fibre is actually a single muscle cell about 100µm in diameter and a few cm long. These giant cells have many nuclei, as they were formed from the fusion of many smaller cells. Their cytoplasm is packed full of myofibrils, bundles of protein filaments that cause contraction. They also have many mitochondria to provide ATP for contraction.



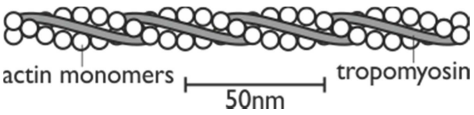
The electron microscope shows that each myofibril is made up of repeating dark and light bands. In the middle of the dark band is a line called the M line and in the middle of the light band is a line called the Z line. The repeating unit from one Z line to the next is called a sarcomere.



A very high resolution electron micrograph shows that each myofibril is made of parallel filaments. There are two kinds of alternating filaments, called the thick and thin filaments. These two filaments are linked at intervals by blobs called cross-bridges, which actually stick out from the thick filaments.

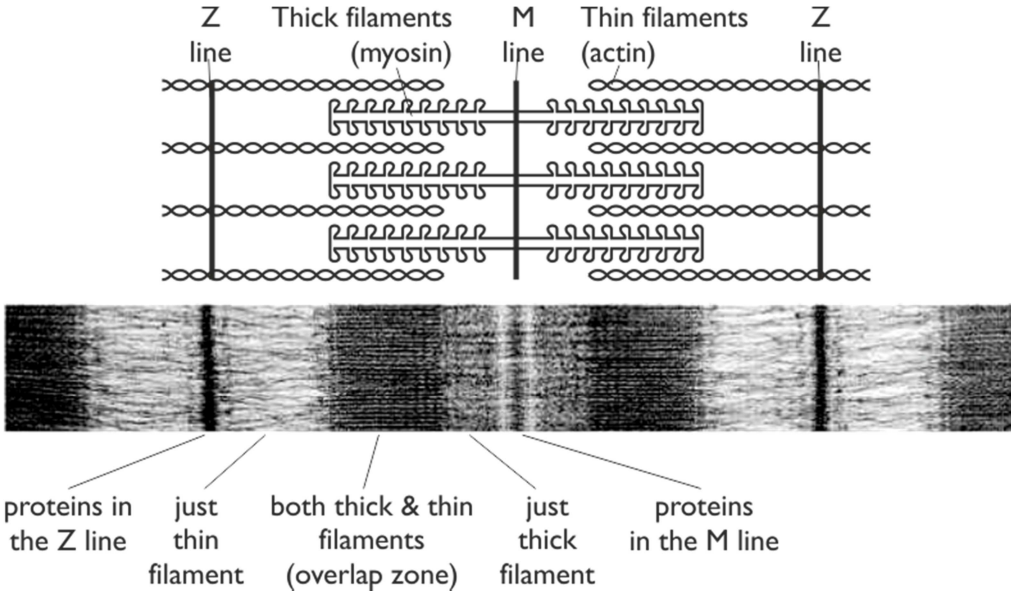


The thick filament is made of a protein called myosin. A myosin molecule is shaped a bit like a golf club, but with two heads. Many of these molecules stick together to form the thick filament, with the "handles" lying together to form the backbone and the "heads" sticking out in all directions to form the cross-bridges.



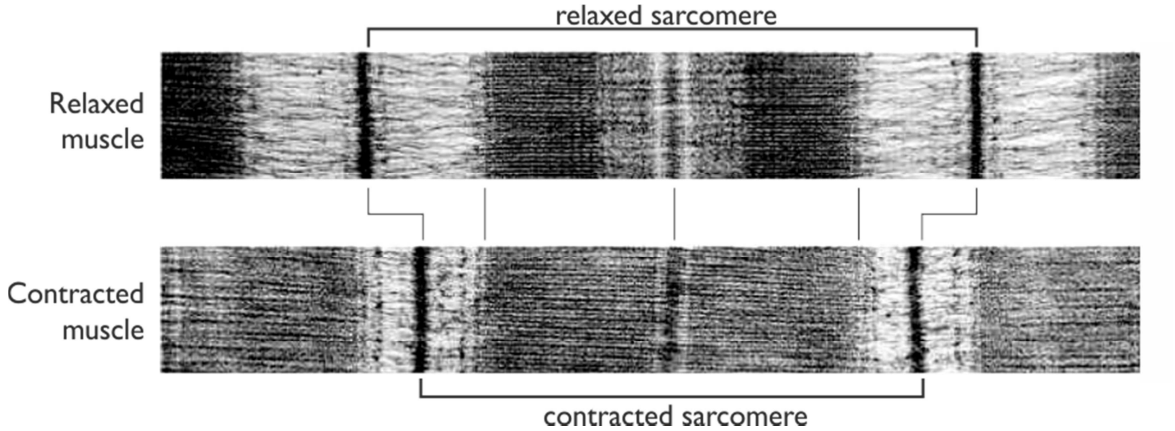
The thin filament is made of a protein called actin. Actin is a globular molecule, but it polymerises to form a long double helix chain. The thin filament also contains troponin and tropomyosin, two proteins involved in the control of muscle contraction.

The thick and thin filaments are arranged in a precise lattice to form a sarcomere. The thick filaments are joined together at the M line, and the thin filaments are joined together at the Z line, but the two kinds of filaments are not permanently joined to each other. The position of the filaments in the sarcomere explains the banding pattern seen by the electron microscope:



Mechanism of Muscle Contraction – the Sliding Filament Theory

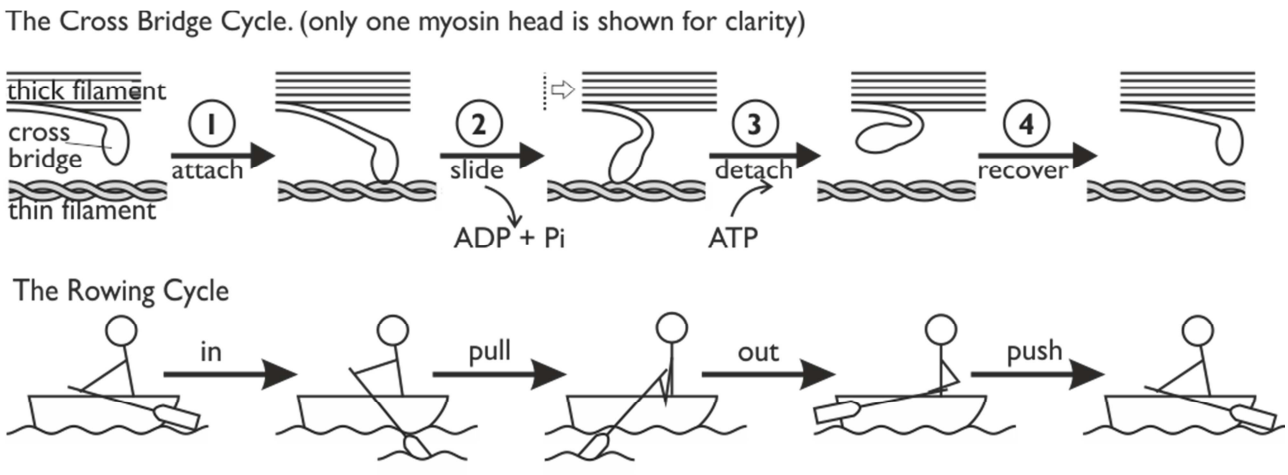
Knowing the structure of the sarcomere enables us to understand what happens when a muscle contracts. The mechanism of muscle contraction can be deduced by comparing electron micrographs of relaxed and contracted muscle:



These show that each sarcomere gets shorter when the muscle contracts, so the whole muscle gets shorter. But the dark band, which represents the thick filament, does not change in length. This shows that the filaments don't contract themselves, but instead they must slide past each other. This sliding filament theory was first proposed by Huxley and Hanson in 1954, and has been confirmed by many experiments since.

The Cross-Bridge Cycle

What makes the filaments slide past each other? Energy is provided by the hydrolysis of ATP, and the ATPase enzyme that does this splitting is located in the myosin cross-bridge head. These cross-bridges can also attach to actin, so they are able to cause the filament sliding by "walking" along the thin filament. This cross-bridge walking is called the cross-bridge cycle, and it has four steps. One step actually causes the sliding, while the other three simply reset the cross-bridge back to its starting state. It is analogous to the four steps involved in rowing a boat:



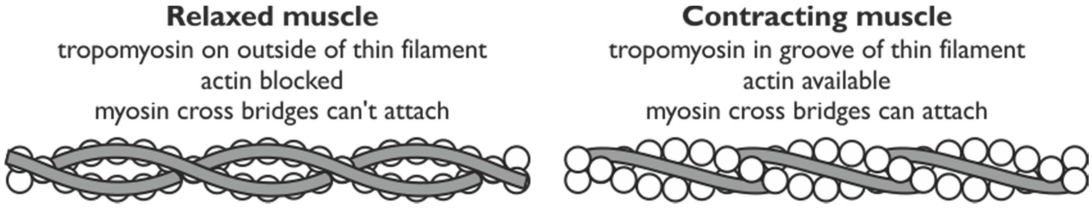
1. The cross-bridge swings out from the thick filament and attaches to the thin filament.
2. The cross-bridge changes shape and rotates through 45°, causing the filaments to slide. The energy from ATP splitting is used for this "power stroke" step, and the products (ADP + Pi) are released.
3. A new ATP molecule binds to myosin and the cross-bridge detaches from the thin filament.
4. The cross-bridge changes back to its original shape, while detached (so as not to push the filaments back again). It is now ready to start a new cycle, but further along the thin filament.

One ATP molecule is split by each cross-bridge in each cycle, which takes a few milliseconds. During a contraction, thousands of cross-bridges in each sarcomere go through this cycle thousands of times, like a millipede running along the ground. Fortunately the cross-bridges are all out of sync, so there are always many cross-bridges attached at any time to maintain the force.

ATP has two roles in the crossbridge cycle: to provide the energy for the power stroke and to detach the crossbridge heads. ATP is therefore still needed in a relaxed muscle and the myosin ATPase still hydrolyses ATP, but only very slowly. When a muscle starts contracting the ATPase rate increases dramatically to provide enough energy for the contraction. After death, all ATP in muscle cells is depleted, so cross-bridges cannot detach from the thin filament. With all cross-bridges attached the muscle becomes stiff – rigor mortis.

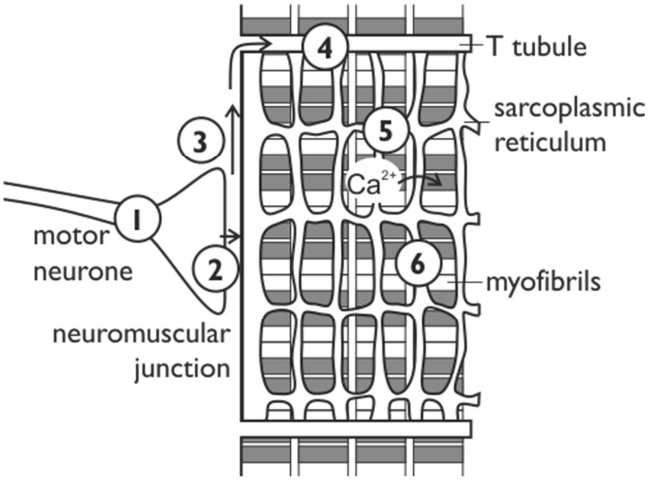
Control of Muscle Contraction

How is the cross-bridge cycle switched off in a relaxed muscle? This is where the regulatory proteins on the thin filament, troponin and tropomyosin, are involved. Tropomyosin is a long thin molecule, and it can change its position on the thin filament. In a relaxed muscle is it on the outside of the filament, covering the actin molecules so that myosin cross-bridges can't attach. This is why relaxed muscle is compliant: there are no connections between the thick and thin filaments. In a contracting muscle the tropomyosin has moved into the groove of the double helix, revealing the actin molecules and allowing the cross-bridges to attach.



Contraction of skeletal muscle is initiated by a nerve impulse, and we can now look at the sequence of events from impulse to contraction (sometimes called excitation-contraction coupling).

1. An action potential arrives at the end of a motor neurone, at the neuromuscular junction.
2. This causes the release of the neurotransmitter acetylcholine.
3 This initiates an action potential in the muscle cell membrane.
4. This action potential is carried quickly throughout the large muscle cell by invaginations in the cell membrane called T-tubules.

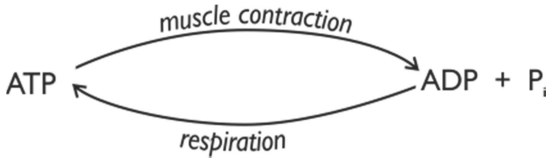


5. The action potential causes the sarcoplasmic reticulum (a large vesicle) to release its store of calcium ions into the myofibrils.
6. The calcium ions bind to troponin on the thin filament, which changes shape, moving tropomyosin into the groove in the process.
7. Myosin cross-bridges can now attach to actin, so the cross-bridge cycle can take place. Cross-bridges keep cycling, and muscles keep shortening or producing force, so long as calcium ions are present.

Relaxation is the reverse of these steps. This process may seem complicated, but it allows for very fast responses so that we can escape from predators and play the piano.

Energy for Muscle Contraction

In animals more energy (in the form of ATP) is used for muscle contraction than for any other process.



A muscle cell has only enough ATP for about three seconds of contraction, so the ATP supply has to be constantly replenished. Muscle cells have three systems for making ATP:

1. The Aerobic System

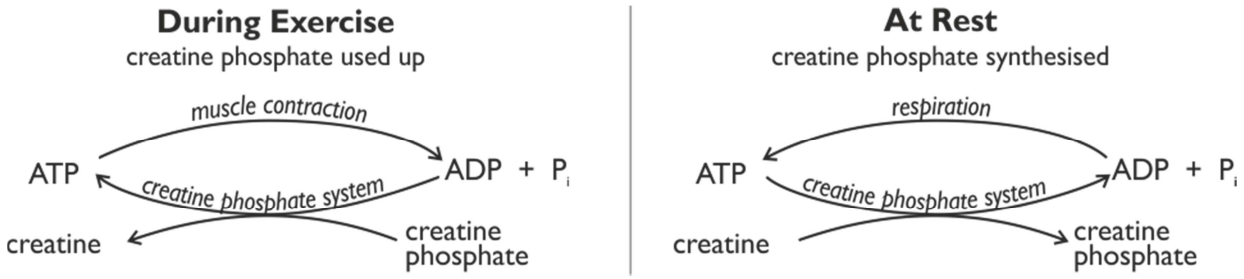
Most of the time, when muscles are resting or moderately active, muscles use aerobic respiration of glucose from the glycogen stores in the liver to make ATP. The aerobic system provides an almost unlimited amount of energy, but contraction is fairly slow, as it is limited by how quickly oxygen and glucose can be provided by the blood. This is why you can't run a marathon at the same speed as a sprint. The aerobic system is the only process to respire fats (triglycerides), so aerobic exercise is the best way to lose body fat.

2. The Anaerobic System (or Glycogen-Lactate System)

If the rate of muscle contraction increases then ATP starts to be used faster than it can be made by aerobic respiration. In this case muscles switch to anaerobic respiration of local muscle glycogen. This is quick, since nothing is provided by the blood, but the lactate produced causes muscle fatigue. The glycogen store in the muscle provides enough ATP to last for about 90s of muscle contraction. The anaerobic system is used for short-distance races like 400m, and when running for a bus.

3. The Creatine Phosphate System

For maximum speed of muscle contraction even the anaerobic system isn't fast enough, so muscles make ATP in a very fast, one-step reaction from creatine phosphate. Creatine phosphate is a short-term energy store in muscle cells, and there is about ten times more creatine phosphate than ATP. It is made from ATP while the muscle is relaxed and can very quickly be used to make ATP when the muscle is contracting. This allows about 10 seconds of fast muscle contraction, enough for short bursts of intense activity such as a 100 metre sprint or running up stairs.

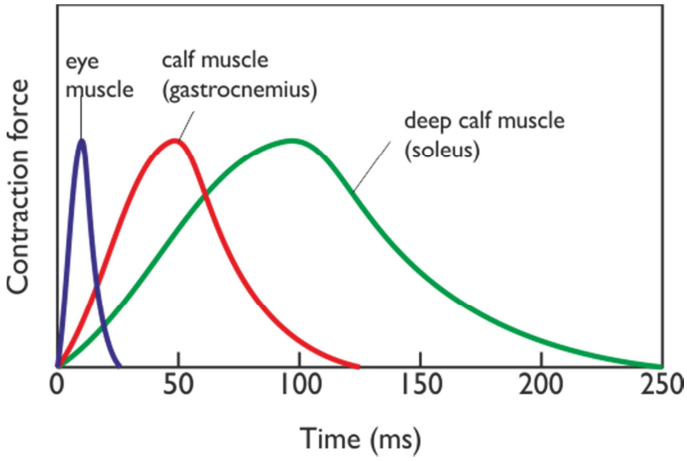


Slow and Fast Muscles

Skeletal muscle cells can be classified as slow-twitch (type I) or fast-twitch (type II) muscles on the basis of their type of respiration and speed of contraction.

Slow-twitch muscle	Fast-twitch muscle
Adapted for aerobic respiration.	Adapted for anaerobic respiration.
Can continue contracting for long periods.	Can only sustain short bursts of activity.
Slow contraction speed, limited by rate of oxygen supply.	Fast contraction speed, not limited by blood supply.
No lactate produced, so not susceptible to muscle fatigue.	Lactate production leads to low pH and muscle fatigue (reduced force and pain).
Contain many mitochondria and the protein myoglobin, which give these cells a red colour. Myoglobin is similar to haemoglobin, and is used as an oxygen store in these muscles, helping to provide the oxygen needed for aerobic respiration.	Contain a lot of glycogen, but few mitochondria and little myoglobin, so cells are a white colour. Mitochondria are not needed for anaerobic respiration.
Found in heart, some leg and back muscles (red meat).	Found in finger muscles, arm muscles, birds' breast muscle and frogs legs (white meat).

The speed of contraction of different muscles is illustrated in this graph showing the muscle force produced over time following stimulation of the muscle. The muscles that move the eye contain mostly fast-twitch cells; the soleus muscle contains mostly slow-twitch cells; while the gastrocnemius muscle contains both types.



Animal Responses

Now we'll look at three simple responses of animals to stimuli. In each case these responses are involuntary responses that aid survival.

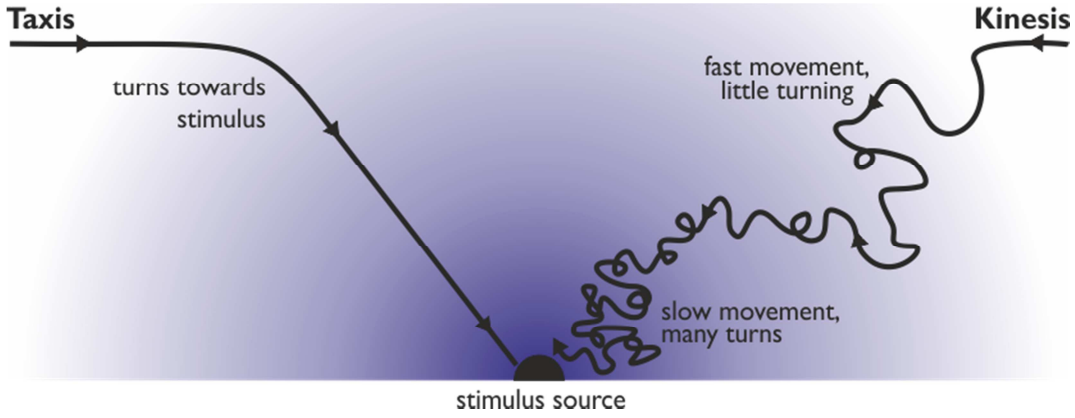
Taxes

A taxis (plural taxes) is a directional response to a directional stimulus. Taxes are common in invertebrates, and even motile bacteria and protoctists show taxes. The taxis response can be positive (towards the stimulus) or negative (away from the stimulus). Common stimuli include:

- **light** (phototaxis), e.g. fly larvae (maggots) use negative phototaxis to move away from light to avoid exposure and desiccation. Adult flies use positive phototaxis to fly towards light to warm up.
- **gravity** (geotaxis), e.g. earthworms show positive geotaxis to burrow downwards underground.
- **chemicals** (chemotaxis) e.g. male moths show positive chemotaxis when flying towards a pheromone.
- **movement** (rheotaxis) e.g. moths use positive rheotaxis to fly into the wind and salmon use positive rheotaxis to swim upstream.

Kineses

A kinesis (plural kineses) is a response to a changing stimulus by changing the amount of activity. Sometimes the speed of movement varies with the intensity of the stimulus (orthokinesis) and sometimes the rate of turning depends on the intensity of the stimulus (klinokinesis). The response is not directional, so kinesis is a suitable response when the stimulus isn't particularly localised. The end result is to keep the animal in a favourable environment. For example woodlice, who breathe using gills, use kinesis to stay in damp environments, where they won't dry out. In dry environments woodlice move quickly and don't turn much, which increases their chance of moving out of that area. If, by chance, they find themselves in a humid environment, they slow down and increase their rate of turning. This tends to keep them in the humid area.

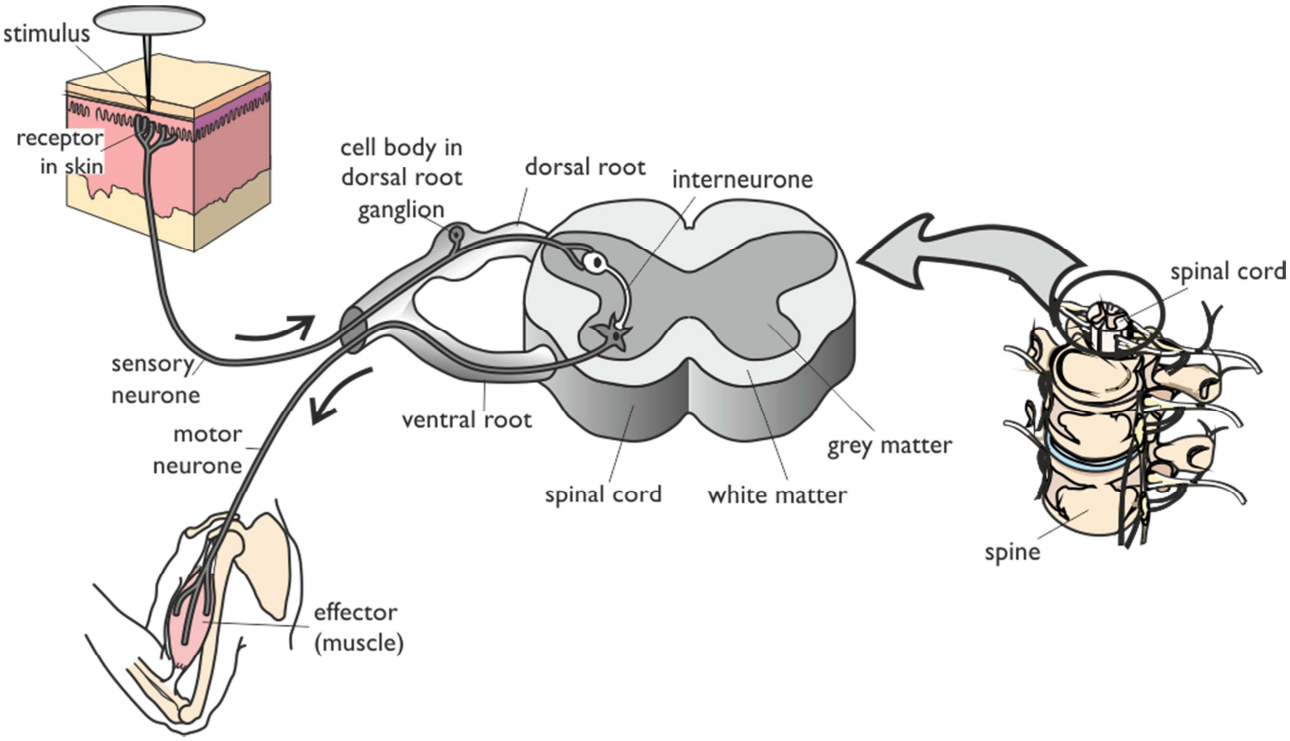


Reflexes

A reflex is a specific response to a specific stimulus. For example when an earthworm detects vibrations in the ground it escapes by quickly withdrawing into its burrow. This reflex response protects the earthworm from predators. All animals have reflexes (even humans) and they are essential for survival. The five key features of a reflex are:

- Involves few neurones. Usually 3, but can be more.
- Immediate. Reflex arcs are very fast, since they involve few neurones, so few synapses.
- Involuntary. No choice or thought is involved.
- Invariable. A given reflex response to a specific stimulus is always exactly the same.
- Innate. Reflex responses are genetically-programmed, not learned.

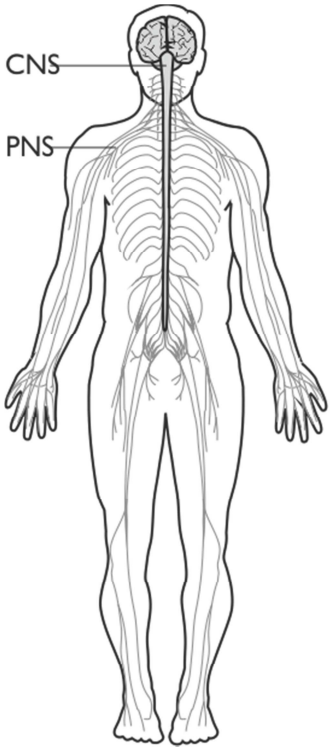
Humans have reflex responses too, which protect the body from damage. For example withdrawing your hand from a sharp object is a reflex and you respond before your brain has registered the pain. Only three neurones are involved. A sensory neurone carries an impulse from a skin pain receptor to an interneurone in the spinal cord, and a motor neurone carries the nerve impulse from the interneurone to a muscle, which responds by contracting.



The interneurone in the spinal cord will usually also synapse with other interneurones, which transmit impulses to other muscles and the brain, and these allow slower, secondary responses in addition to the reflex. Other human reflexes are the knee-jerk reflex (which protects the knee joint from damage) and the pupil reflex (which protects the retina from damage by intense light).

The Organisation of the Human Nervous System

The organisation of the human nervous system is shown in this diagram:



Human Nervous System

Central Nervous System (CNS)
Brain and spinal cord
Interneurones

Peripheral Nervous System (PNS)
Everything else
Sensory & effector neurones

Sensory Neurones
Input from receptors

Effector neurones
Output to muscles & glands

Somatic Nervous System
Voluntary responses
Output to skeletal muscles

Autonomic Nervous System
Involuntary responses
Output to smooth muscles & glands

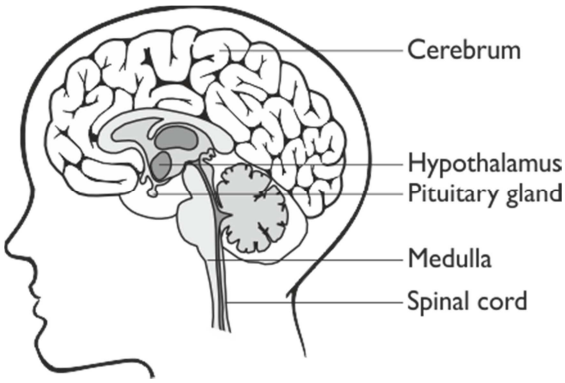
Sympathetic Neurones
"fight or flight" responses
Neurotransmitter: noradrenaline
"Adrenergic System"

Parasympathetic Neurones
relaxing responses
Neurotrasmitter: acetylcholine
"Cholinergic System"

It is easy to forget that much of the human nervous system is concerned with routine, involuntary jobs, such as homeostasis, digestion, posture, breathing, etc. These are the jobs of the autonomic nervous system. Its functions are split into two divisions, with anatomically-distinct neurones. Most body organs are innervated by two separate sets of effector neurones; one from the sympathetic system and one from the parasympathetic system. These neurones have opposite (or antagonistic) effects. In general the sympathetic system stimulates the "fight or flight" responses to threatening situations, while the parasympathetic system relaxes the body. For example, as we shall see, sympathetic neurones speed up heart rate, while parasympathetic neurones slow it down.

Certain key parts of the brain are involved in involuntary functions, and are connected to the autonomic nervous system:

- The medulla controls heart rate
- The hypothalamus controls homeostasis
- The pituitary gland secretes LH and FSH

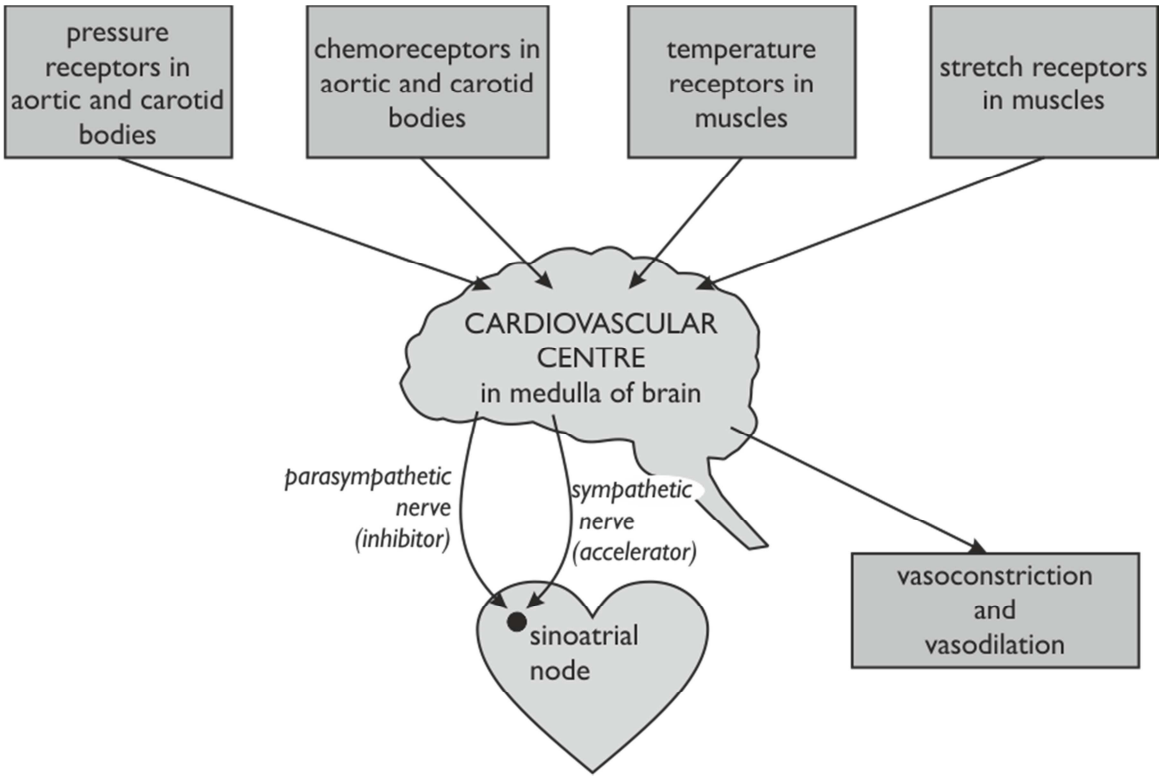


Control of Heart Rate

In unit 1 we learnt that the heart is myogenic – each heart beat is initiated by the sinoatrial node (SAN) in the heart itself, not by nerve impulses from the CNS. But the heart rate and the stroke volume can both be controlled by the CNS to alter the cardiac output – the amount of blood flowing in a given time:

Cardiac output = heart rate x stroke volume

Control of the heart rate is an involuntary reflex, and like many involuntary processes (such as breathing, coughing and sneezing) it is controlled by a region of the brain called the medulla. The medulla and its nerves are part of the autonomic nervous system. The part of the medulla that controls the heart is called the cardiovascular centre. It receives inputs from various receptors around the body and sends output through two nerves to the sino-atrial node in the heart.



How does the cardiovascular centre control the heart?

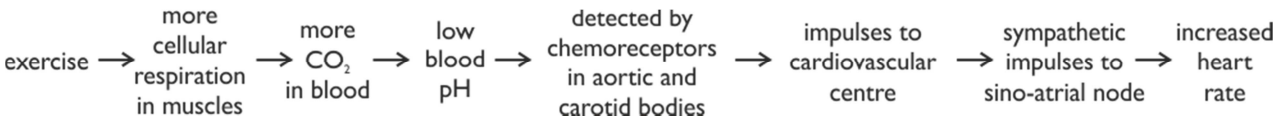
The cardiovascular centre can control both the heart rate and the stroke volume. There are two separate nerves from the cardiovascular centre to the sino-atrial node: the sympathetic nerve to speed up the heart rate and the parasympathetic nerve to slow it down. The cardiovascular centre can also change the stroke volume by controlling blood pressure. It can increase the stroke volume by sending nerve impulses to the arterioles to cause vasoconstriction, which increases blood pressure so more blood fills the heart at diastole. Alternatively it can decrease the stroke volume by causing vasodilation and reducing the blood pressure.

How does the cardiovascular centre respond to exercise?

When the muscles are active they respire more quickly and cause several changes to the blood, such as decreased oxygen concentration, increased carbon dioxide concentration, decreased pH (since the carbon dioxide dissolves to form carbonic acid) and increased temperature. All of these changes are detected by various receptor cells around the body, but the pH changes are the most sensitive and therefore the most important. These pH changes are detected by chemoreceptors found in:

- The walls of the aorta (the aortic body), monitoring the blood as it leaves the heart
- The walls of the carotid arteries (the carotid bodies), monitoring the blood to the head and brain
- The medulla, monitoring the tissue fluid in the brain

The chemoreceptors send nerve impulses to the cardiovascular centre indicating that more respiration is taking place, and the cardiovascular centre responds by increasing the heart rate.



A similar job is performed by temperature receptors and stretch receptors in the muscles, which also detect increased muscle activity.

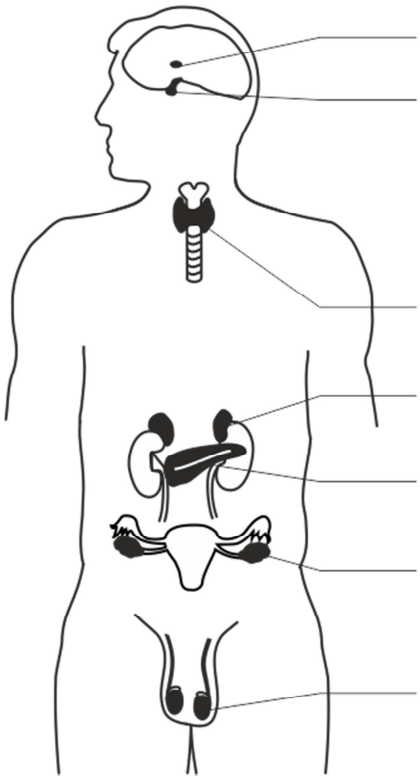
The Hormone System

Humans have two complementary control systems that they can use to respond to their environment: the nervous system and the endocrine (hormonal) system. We've looked at the nervous system, so now we'll now look briefly at the hormone system.

Hormones are secreted by glands into the blood stream. There are two kinds of glands:

- Exocrine glands secrete solutions to the outside, or to body cavities, usually through ducts (tubes). e.g. sweat glands, tear glands, mammary glands, digestive glands.
- Endocrine glands do not have ducts but secrete chemicals directly into the tissue fluid, whence they diffuse into the blood stream. e.g. thyroid gland, pituitary gland, adrenal gland. The hormone-secreting glands are all endocrine glands.

This table shows some of the main endocrine glands and their hormones. The hormones marked with a * are ones that we shall look at in detail later.



Gland	Hormone	Target organ	Function
Pineal gland	melatonin	many	biological clock
Pituitary gland	FSH*	ovaries	menstrual cycle
	LH*	ovaries	menstrual cycle
	ADH	kidneys	water homeostasis
	growth hormone	many	stimulates cell division
	oxytocin	uterus	birth contractions
	prolactin	mammary glands	milk production
Thyroid gland	thyroxine*	liver	metabolic rate
Adrenal glands	adrenaline*	many	fight or flight
	cortisol	many	anti-stress
Pancreas	insulin*	liver	glucose homeostasis
	glucagon*	liver	glucose homeostasis
Ovaries	oestrogen*	uterus	menstrual cycle
	progesterone*	uterus	menstrual cycle
Testes	testosterone	many	male characteristics

Once a hormone has been secreted by its gland, it diffuses into the blood stream and is carried all round the body to all organs. However, it only affects certain target organs, which can respond to it. These target organs have specific receptor molecules in their cells to which the hormone binds. These receptor molecules are proteins, and they form specific hormone-receptor complexes, very much like enzyme-substrate complexes. Cells without the specific receptor will just ignore a hormone. The hormone-

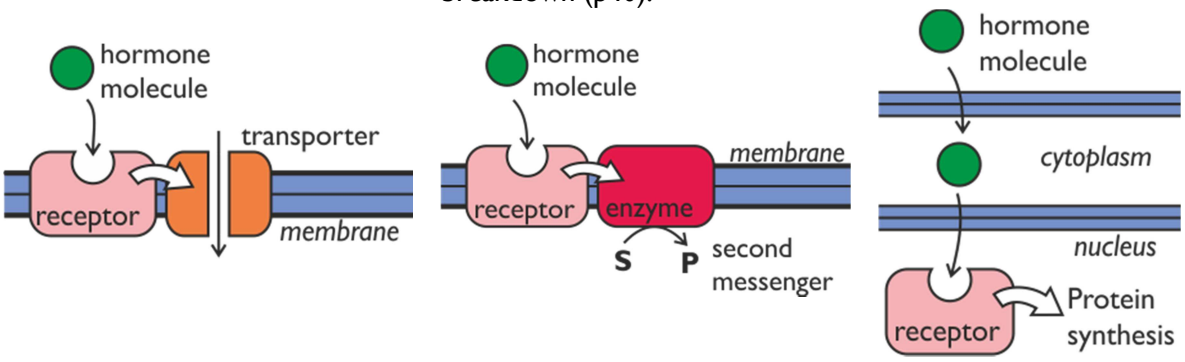
receptor complex can affect almost any aspect of a cell's function, including metabolism, transport, protein synthesis, cell division or cell death.

There are three different ways in which a hormone can affect cell function:

① Some hormones affect the permeability of the cell membrane. They bind to a receptor on the membrane, which then activates a transporter, so substances can enter or leave the cell. For example insulin stimulates glucose uptake (p40).

② Some hormones release a "second messenger" inside the cell. They bind to a receptor on the membrane, which then activates an enzyme in the membrane, which catalyses the production of a chemical in the cytoplasm, which affects various aspects of the cell. For example adrenaline stimulates glycogen breakdown (p40).

③ The steroid hormones are lipid-soluble so can easily pass through membranes by lipid diffusion. They diffuse to the nucleus, where they bind to a receptor, which is a transcription factor that activates protein synthesis. For example oestrogen stimulates growth in the uterus (p63).



So in most cases, the hormone does not enter the cell. The effect of a hormone is determined not by the hormone itself, but by the receptor in the target cell. Thus the same hormone can have different effects in different target cells.

Comparison of Nervous and Hormone Systems

Nervous System	Hormone System
Transmitted by specific neurone cells	Transmitted by the circulatory system
Effect localised by neurone anatomy	Effect localised by target cell receptor proteins
Fast-acting (ms–s)	Slow-acting (minutes–days)
Short-lived response	Long-lived response

The two systems work closely together: endocrine glands are usually controlled by the nervous system, and a response to a stimulus often involves both systems.

Paracrine Signalling

Paracrine signalling is communication between close cells using chemicals called local chemical mediators. These chemicals are released by cells into the surrounding tissue fluid, but not into the blood. Thus they only have a local effect on the cells surrounding their release, in contrast to hormones. Like hormones and neurotransmitters they bind to receptors on the surface of the target cells to cause an effect. Two such local mediators are prostaglandins and histamine, which control the inflammatory response.

- **Prostaglandins** are lipid molecules that are produced by cells in almost every tissue in the body, targeting local smooth muscle cells and endothelial (lining cells). There are over a dozen different prostaglandins known, causing a number of effects, especially the inflammatory response to injury and infection (unit 1). They cause vasodilation by stimulating smooth muscle cells in the walls of local arterioles to relax, so increasing the flow of blood to the area (so the area turns red). They also stimulate the blood clotting process (so wounds are sealed) and they stimulate the pain neurones. Anti-inflammatory drugs, like aspirin and ibuprofen, reduce inflammation and associated pain by inhibiting a key enzyme in the synthesis of prostaglandins.

- **Histamine** is also involved in the inflammatory response, including the inflammation following stings from insects or nettles. Histamine is made from the amino acid histidine, and is stored in granules in mast cells, which are found in connective tissue, especially in the skin. When stimulated by the immune system or by injury, the mast cells release the histamine into the surrounding tissue fluid. There it stimulates vasodilation of nearby arterioles and loosening of nearby capillary walls, causing the capillaries to leak, so plasma and leukocytes can reach the injury. Histamine also causes bronchoconstriction in the airways. In some cases too much histamine is released, resulting in an extreme inflammatory response called an allergic reaction. Antihistamine drugs inhibit the release of histamine and so are used to counter allergic reactions.

- **Neurotransmitters** are also local chemical mediators, since these chemicals only have a local effect at synapses and are not released into the blood.

Homeostasis

Homeostasis literally means "standing still" and it refers to the process of keeping the internal body environment in a steady state. The importance of this cannot be over-stressed, and a great deal of the hormone system and autonomic nervous system is dedicated to homeostasis.

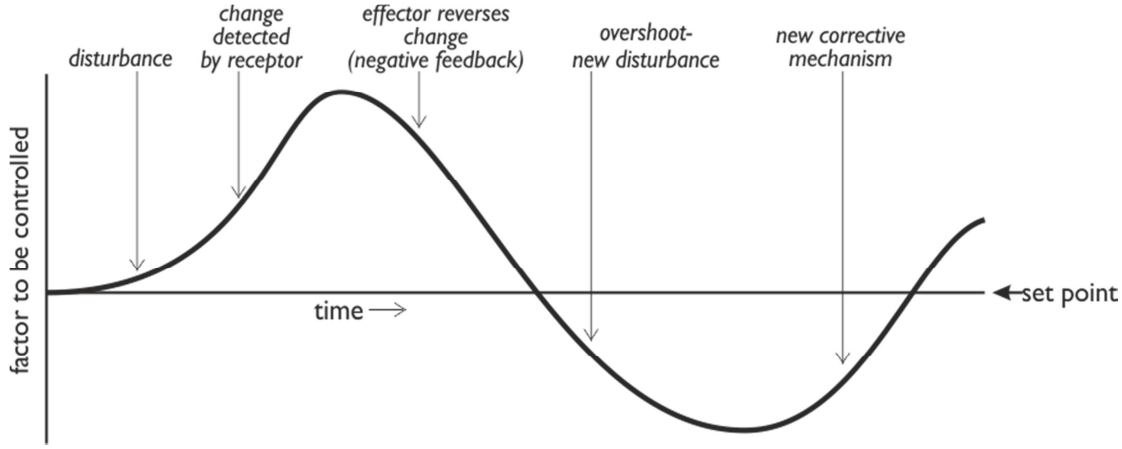
Factors that are controlled include

- body temperature – to keep enzymes working near their optimum temperature and stop them denaturing
- blood pH – to keep enzymes working near their optimum pH
- blood glucose concentration – to ensure there is enough glucose available for cellular respiration, but not enough to lower the blood water potential and dehydrate cells
- blood water potential – to prevent loss or gain of water from cells by osmosis

We shall look at two examples of homeostasis in detail: temperature and blood glucose.

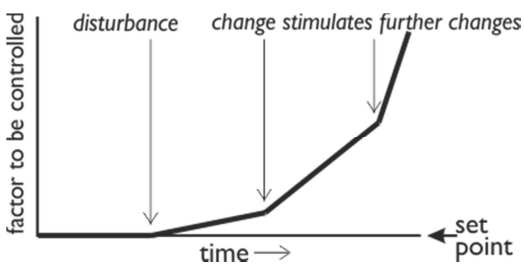
Negative and positive feedback

All homeostatic mechanisms use negative feedback to maintain a constant value (called the set point). Negative feedback means that whenever a change occurs in a system, the change automatically causes a corrective mechanism to start, which reverses the original change and brings the system back to normal. It also means that the bigger then change the bigger the corrective mechanism. Negative feedback applies to electronic circuits and central heating systems as well as to biological systems.



So in a system controlled by negative feedback the level is never maintained perfectly, but constantly oscillates about the set point. An efficient homeostatic system minimises the size of the oscillations.

Positive feedback occurs when the change stimulates a further change in the same direction. Positive feedback is potentially dangerous and is usually associated with a breakdown in the normal control mechanism (e.g. hyperthermia p37).



Temperature Homeostasis (Thermoregulation)

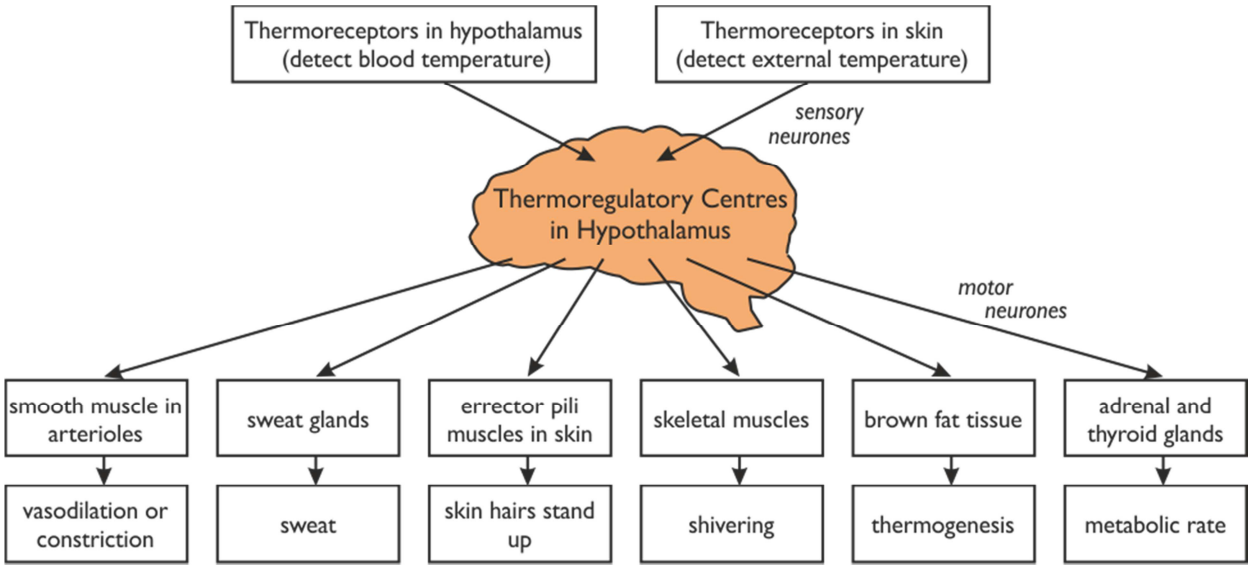
One of the most important examples of homeostasis is the regulation of body temperature. Body temperature needs to be regulated to keep enzymes working close to their optimum temperature and to prevent them from denaturing. There are basically two ways of doing this: mammals and birds can generate their own heat and are called endotherms; while all other animals rely on gaining heat from their surroundings, so are called ectotherms.

Temperature Homeostasis in Endotherms

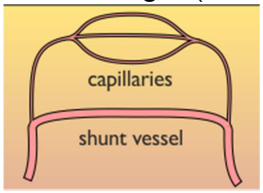
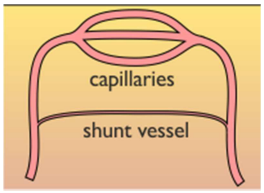
Endotherms (mammals and birds) can generate heat internally, have thermal insulation, and can usually maintain a remarkably constant body temperature. Humans and most other mammals maintain a set point of 37.5 ± 0.5 °C, while birds usually have set points of 40-42°C. Endotherms are sometimes called warm-blooded animals, but this term isn't very scientific, partly because endotherms can get quite cold (e.g. during hibernation). Endotherms don't keep their whole body at the same temperature: they maintain a constant core temperature; while allowing the peripheral temperature to be colder, especially the surface, which is in contact with the surroundings.

In humans temperature homeostasis is controlled by two thermoregulatory centres in the hypothalamus. The thermoregulatory centres receive input from two sets of thermoreceptors: receptors in the hypothalamus itself monitor the temperature of the blood as it passes through the brain (the core temperature), and receptors in the skin monitor the external temperature. Both pieces of information are needed so that the body can make appropriate adjustments.

- One of the thermoregulatory centres – the heat loss centre – is activated when the core temperature rises. It sends impulses to several different effectors in the body to reduce the core temperature.
- The other thermoregulatory centre – the heat gain centre – is activated when the core temperature falls. It sends impulses to several different effectors in the body to increase the core temperature.



The thermoregulatory centre is part of the autonomic nervous system, so the various responses are all involuntary. The exact responses to high and low temperatures are described in the table below:

Effector	**Response to Low Temperature (controlled by the heat gain centre)**	**Response to High Temperature (controlled by the heat loss centre)**
Smooth muscles in peripheral arterioles in the skin.	Muscles contract causing vasoconstriction. Less heat is carried from the core to the surface of the body, maintaining core temperature. Extremities can turn blue and feel cold and can even be damaged (frostbite). 	Muscles relax causing vasodilation. More heat is carried from the core to the surface, where it is lost by convection and radiation. Skin turns red. 
Sweat glands	No sweat produced.	Glands secrete sweat onto surface of skin, where it evaporates. This is an endothermic process, and since water has a high latent heat of evaporation, it takes a lot of heat from the body. Some hairy mammals pant instead of sweating.
Skeletal muscles	Muscles contract and relax repeatedly and involuntarily, generating heat by friction and from metabolic reactions.	No shivering.
Erector pili muscles in skin (attached to skin hairs)	Muscles contract, raising skin hairs and trapping an insulating layer of still, warm air next to the skin. Not very effective in humans, just causing "goosebumps".	Muscles relax, lowering the skin hairs and allowing air to circulate over the skin, encouraging convection and evaporation.
Brown fat tissue	Non-shivering thermogenesis. Increased respiration in brown fat tissue – specialised fat tissue packed with "uncoupled" mitochondria that respire triglycerides to produce heat but no ATP. Especially important in babies and hibernating mammals.	Decreased respiration in brown fat tissue.
Adrenal and thyroid glands	Glands secrete adrenaline and thyroxine respectively, which increase the metabolic rate in different tissues, especially the liver, so generating heat.	Glands stop releasing adrenaline and thyroxine, so metabolic rate slows.
Behaviour	Curling up, huddling, finding shelter, putting on more clothes, etc.	Basking, stretching out, finding shade, swimming, removing clothes, etc.

Note that

- some of the responses to low temperature actively generate heat (thermogenesis), while others just conserve heat.
- Similarly some of the responses to warm temperatures actively cool the body down, while others just reduce heat production or transfer heat to the surface.

The body thus has a range of responses available, depending on the internal and external temperatures.

Mammals can alter their set point in special circumstances:

- Fever. White blood cells release chemicals called pyrogens as part of the inflammatory response to infection. These pyrogens raise the set point of the thermoregulatory centre causing the whole body temperature to increase by 2-3 °C. This helps to kill bacteria and explains why patients shiver even though they are hot.
- Hibernation. Some mammals release hormones that reduce their set point to around 5°C while they hibernate. This drastically reduces their metabolic rate and so conserves their food reserves.
- Torpor. Bats and hummingbirds reduce their set point every day while they are inactive. They have a high surface area: volume ratio, so this reduces heat loss.

Failure of temperature homeostasis

- Hypothermia occurs when heat loss exceeds heat generation, due to prolonged exposure to cold temperatures. As the core temperature decreases the metabolic rate also decreases, leading to less thermogenesis. If the core temperature drops below 32°C shivering stops so the core temperature drops even further. If the core temperature falls below 30°C hypothermia is usually fatal.
- Hyperthermia occurs when heat gain exceeds heat loss, usually due to prolonged exposure to high temperatures. This situation is often associated with dehydration, which reduces sweating, the only effective way to cool down. A rise in core temperature increases the metabolic rate, fuelling a further increase in temperature. If the core temperature rises above 40°C hyperthermia is usually fatal.

Both hypothermia and hyperthermia are examples of the dangers of positive feedback.

The advantage of being endothermic is that animals can survive in a wide range of environmental temperatures, and so can colonise almost any habitat, and remain active at night and in cold weather. This gives endothermic predators an obvious advantage over ectothermic prey. The disadvantage is that it requires a lot of energy, so endotherms need to eat far more than ectoderms.

Temperature Control in Ectotherms

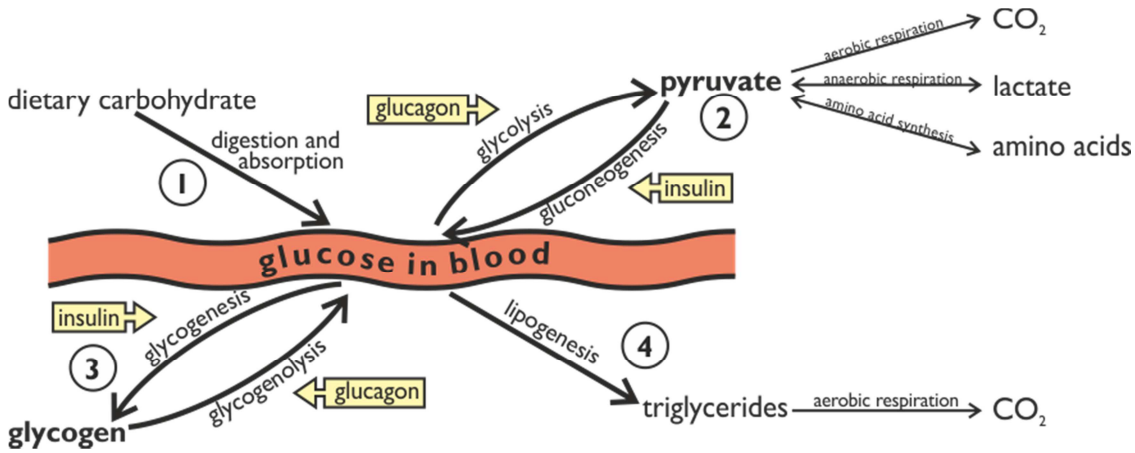
Ectotherms (all animals except mammals and birds) rely on external heat sources to warm up; do not have thermal insulation; and their body temperature varies with the environmental temperature. Ectotherms are sometimes called cold-blooded animals, but this term isn't very scientific, because ectoderms can get very warm. Reptiles, such as lizards, iguanas and crocodiles are classic ectotherms. They cannot warm up by shivering because, if their temperature is low, they cannot respire fast enough to make ATP for rapid muscle contraction. Instead, reptiles regulate their body temperature by thermoregulatory behaviour e.g.

- Iguanas start every day by basking on rocks in the sun until their metabolic rate is fast enough for them to become active.
- Lizards lie down on warm ground to gain heat, and raise themselves off the ground if they get too hot.
- To prevent overheating in the midday sun lizards take shelter under rocks or vegetation.
- Some lizards can adjust the amount of heat they gain by changing their angle to the sun. Turning their backs to the sun presents the maximum surface area, while pointing towards the sun presents the minimum surface area.
- Crocodiles can move between the land and the water during the day to maintain a constant temperature.
- At night, lizards shelter in burrows, which provide insulation to reduce heat loss (and hide them from predators).

The advantage of being ectothermic is that animals use far less energy than endoderms. At rest the metabolic rate of a reptile is only 10% of that of a mammal of similar size. At night, when the core temperature of ectotherms drops with the temperature of the surroundings, their metabolic rate drops still further. This means ectothermic animals need to eat far less than endotherms, and can often survive for weeks without eating. The disadvantages are that, at certain times of the day, ectotherms can only move slowly and have slow reactions. This makes them easy prey and poor predators.

Blood Glucose Homeostasis

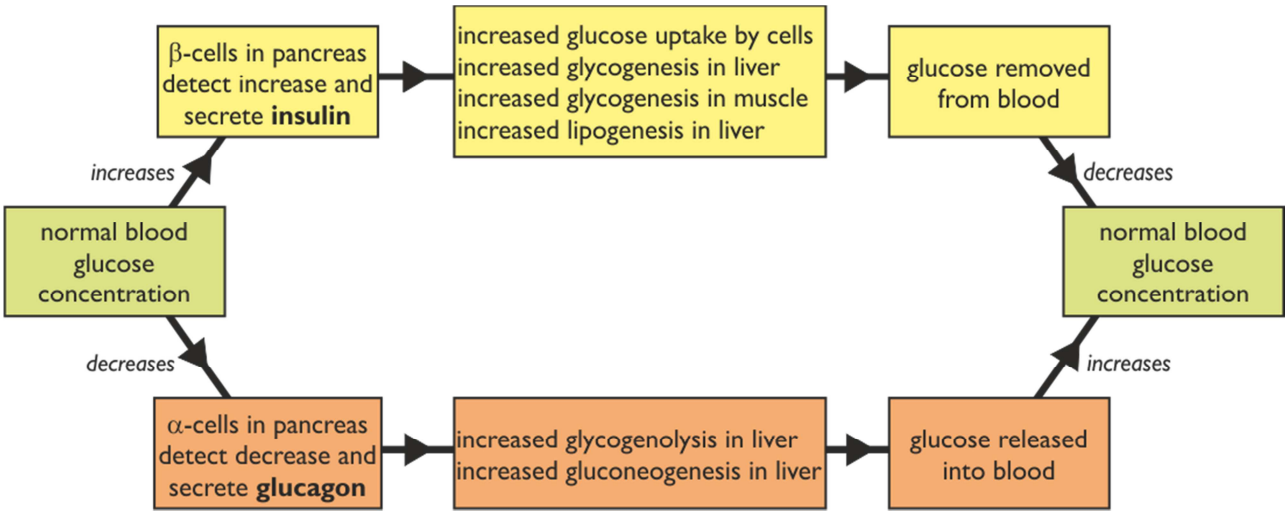
Glucose is the transport carbohydrate in animals, and its concentration in the blood affects every cell in the body. The brain in particular can only respire glucose (not lipids) but it doesn't store glycogen. Very low concentrations of glucose (hypoglycaemia) will cause brain cells to die and very high concentrations (hyperglycaemia) will lower the blood water potential and kill cells by dehydration. The concentration of glucose in the blood is therefore strictly controlled within the range 80-100 mg 100cm^{-3}. This diagram shows the main sources and fates of blood glucose.



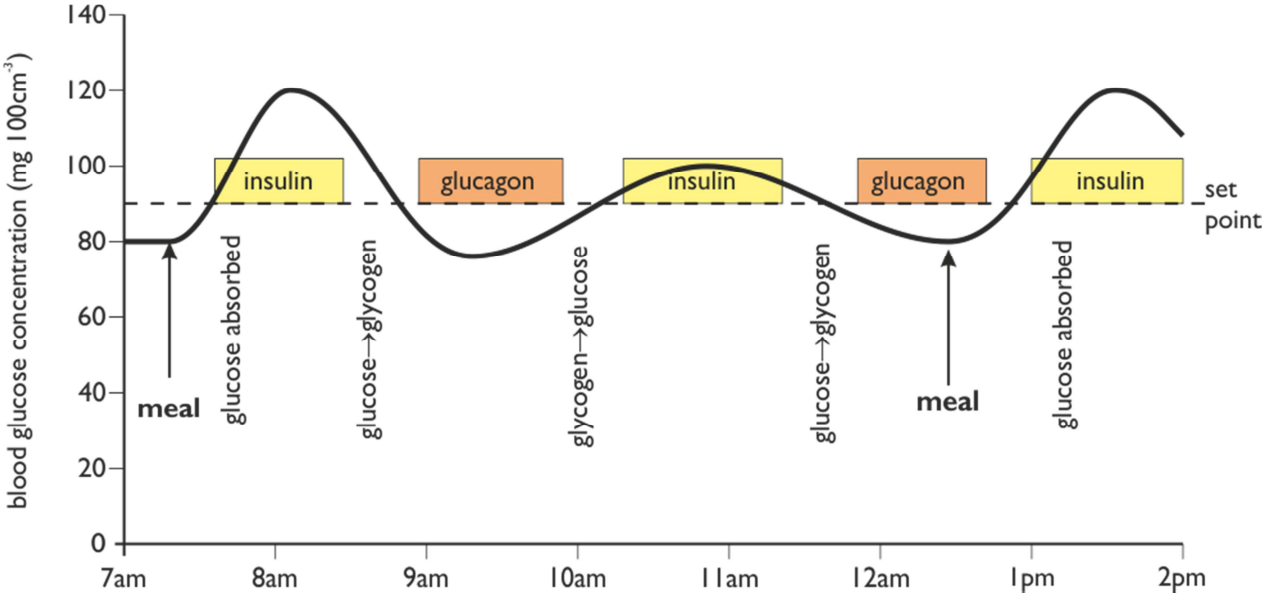
1. The main source of blood glucose is the digestion and absorption of dietary carbohydrate (mostly starch). Blood from the intestine goes first through the hepatic portal vein to the liver, where glucose homeostasis is controlled, before being carried to the rest of the body.
2. Glucose is taken up and used for respiration by all cells. It can also be used to synthesise amino acids, nucleotides, etc. Occasionally glucose can be re-synthesised from pyruvate and amino acids, e.g. when dietary glucose is scarce (gluconeogenesis).
3. For storage glucose can be converted to the polysaccharide glycogen in liver and muscle cells (glycogenesis). This is reversible, and when glucose is needed the glycogen can be broken down again (glycogenolysis).
4. Excess glucose can also be converted to triglycerides in the liver (lipogenesis) then transported as lipoproteins to adipose tissue for storage. This process is irreversible: triglycerides can be used in aerobic respiration but cannot be used to make glucose.

Control of blood glucose concentration

Blood glucose concentration is unusual in that it is not controlled by the CNS, but by the pancreas, which is both an exocrine and an endocrine organ. Regions of the pancreas, called the islets of Langerhans, serve as both glucose receptors and as endocrine cells, releasing hormones to effect the control of glucose. There are two kinds of islet cells, called α and β cells. Both cells have glucose receptors, but the α cells detect low glucose concentrations and respond by secreting glucagon, while the β cells detect high glucose concentrations and respond by secreting insulin. These two hormones are antagonistic, which means that they have opposite effects on blood glucose:

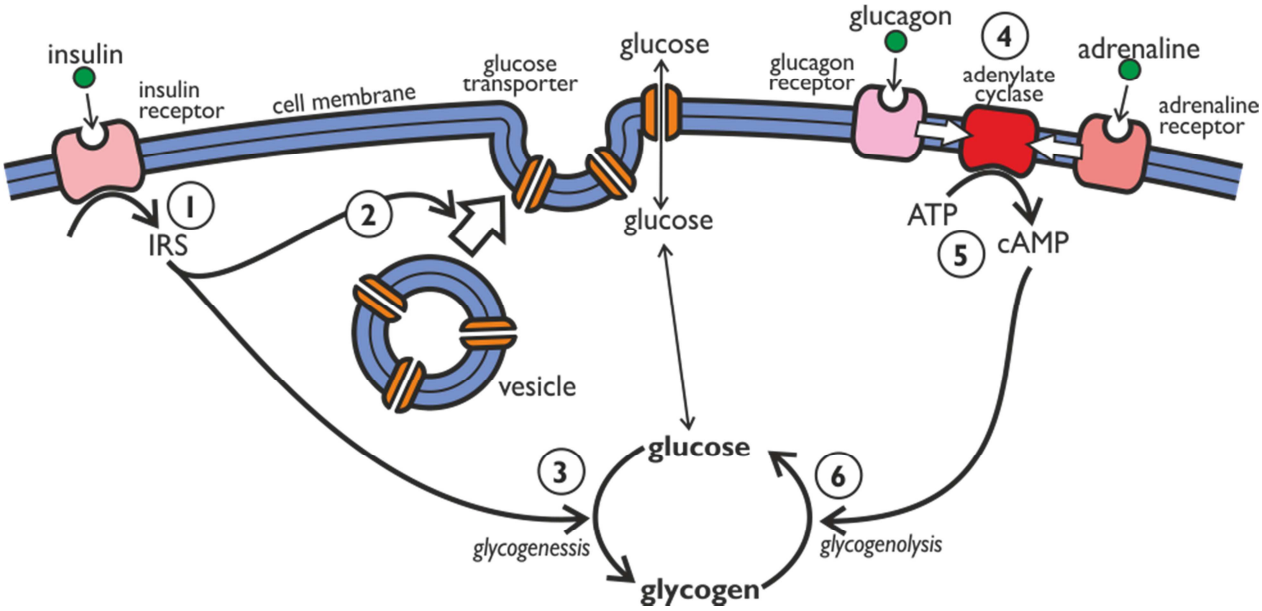


After a meal, glucose is absorbed from the gut into the bloodstream, increasing the blood glucose concentration. This increase is detected by the pancreas, which secretes insulin from its β cells in response. Insulin causes glucose to be taken up by the liver and converted to glycogen. This reduces blood glucose, which is detected by the pancreas, which stops secreting insulin. If the glucose level falls too far, the pancreas detects this and releases glucagon from its α cells. Glucagon causes the liver to break down some of its glycogen store to glucose, which diffuses into the blood. This increases blood glucose, which causes the pancreas to stop producing glucagon. These negative feedback loops continue all day, as shown in this graph:



The mechanism of glucose homeostasis hormone action

This diagram summarises how the hormones insulin and glucagon exert their effects. It also includes the hormone adrenaline, which is the "fight or flight" hormone released by the adrenal glands. Amongst many other effects, adrenaline also stimulates the release of glucose from the liver, to provide more energy for muscle contraction.



1. Insulin molecules bind to an insulin receptor protein in the cell membrane. This binding activates an enzyme active site on the inner surface of the same membrane protein, which catalyses a reaction activating a molecule called insulin receptor substrate (IRS). IRS is the second messenger for insulin.
2. IRS has two separate effects in the cell. Firstly it increases the rate of glucose uptake by recruiting more glucose transporters to the cell membrane. These receptors are stored in cytoplasmic vesicles, and IRS causes these vesicles to fuse with the cell membrane.
3. Secondly, IRS actives the enzyme glycogen synthase, which synthesises glycogen from cytoplasmic glucose. In fact there are several steps here: IRS activates one enzyme, which activates another, which activates another, which activates glycogen synthase. This cascade amplifies the effect, so each molecule of insulin can activate thousands of molecules of glycogen synthase.
4. If present in the blood, glucagon and adrenaline each bind to their specific membrane receptor proteins.
5. These hormone-receptor complexes now activate a membrane-bound enzyme called adenylate cyclase, which catalyses the conversion of ATP to cAMP (cyclic adenosine monophosphate), which is the second messenger for these hormones.
6. cAMP activates the enzyme glycogen phosphorylase, which catalyses the breakdown of glycogen to glucose. This process involves another multi-step cascade amplification.

These hormones demonstrate two advantages of the second messenger system: one hormone (insulin in this case) can have two completely different effects in a cell; while two different hormones (glucagon and adrenaline in this case) can have the same effect in a cell. Each hormone can have different effect in different cell types too.

Diabetes Mellitus

Diabetes is a disease caused by a failure of glucose homeostasis. There are two forms of the disease.

- In insulin-dependent diabetes (also known as IDDM, type I or early-onset diabetes) there is a severe insulin deficiency due to autoimmune killing of β cells (possibly due to a virus). This type usually appears in childhood.
- In non insulin-dependent diabetes (also known as NIDDM, type II or late-onset diabetes) insulin is produced, but the insulin receptors in the target cells don't work, so insulin has no effect. This type tends to appear in overweight people at around age 40, and it accounts for 90% of diabetes cases in the industrialised world.

In both cases there is a very high blood glucose concentration after a meal, so the kidney can't reabsorb it all back into the blood, so much of the glucose is excreted in excess urine (diabetes mellitus means "sweet fountain"). This leads to the symptoms of diabetes:

- high thirst due to osmosis of water from cells to the blood, which has a low water potential.
- copious urine production due to excess water in blood.
- poor vision due to osmotic loss of water from the eye lens.
- tiredness due to loss of glucose in urine and poor uptake of glucose by liver and muscle cells, so no glycogen stores.
- muscle wasting due to gluconeogenesis caused by increased glucagon.

Until the discovery of insulin in 1922 by Banting and Best, diabetes was an untreatable, fatal disease. Today diabetes can be treated by injections with insulin or by careful diet. Insulin can be extracted from the pancreas tissue of cattle and pigs, or it can be made in fermenters by genetically-engineered bacteria (pxx). Treatment is improving all the time, with the development of fast-acting and slow-acting insulin preparations; simpler injection pens; oral insulin preparations (insulin, a protein hormone, is normally digested in the intestine); portable insulin infusion pumps; and even the possibility of islets of Langerhans transplants.

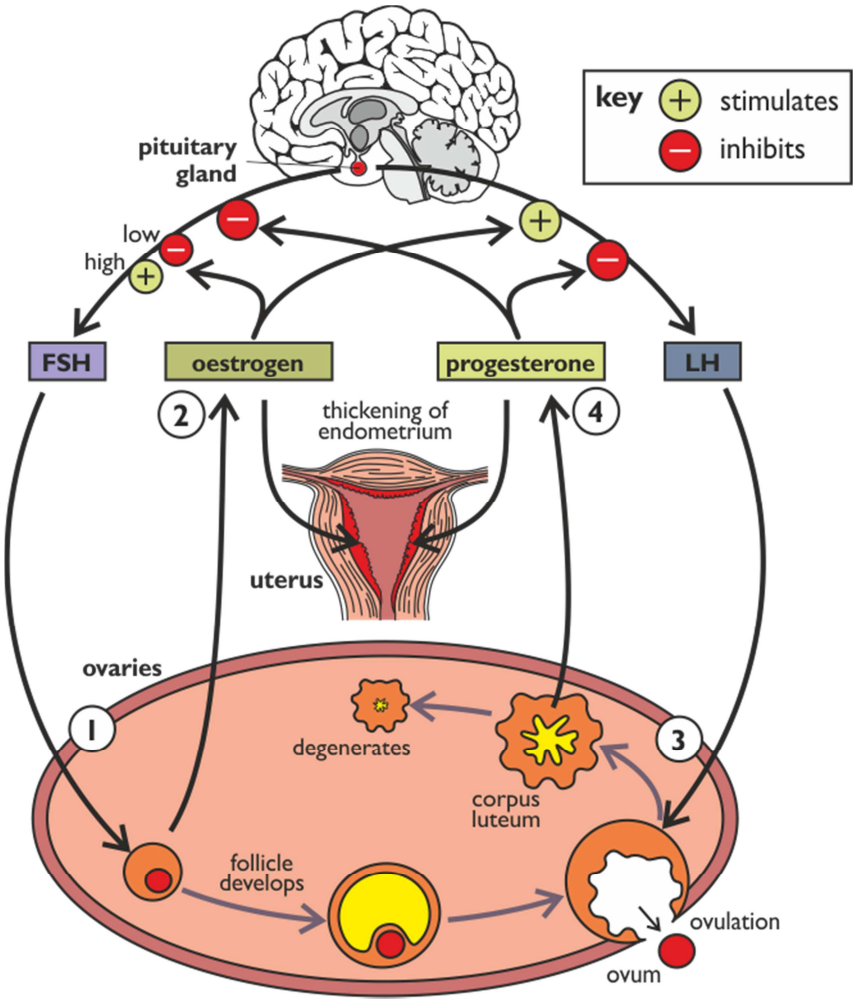
Control of the Mammalian Oestrous Cycle

Female mammals produce eggs and become receptive to mating in a regular cycle, called the oestrus cycle. The oestrus cycle can vary in length from five days in mice to one year for red deer. Oestrus itself refers to the period of the cycle when the female is sexually receptive, or "on heat". For example dogs usually come into heat twice a year. In humans the oestrus cycle is also known as the menstrual cycle because it is one month long (from the Latin *mens*, month). Humans (and other primates) are unusual in that the females are sexually receptive throughout the year, and the uterine lining (the endometrium) is shed each month through the vagina in the process of menstruation (the "period").

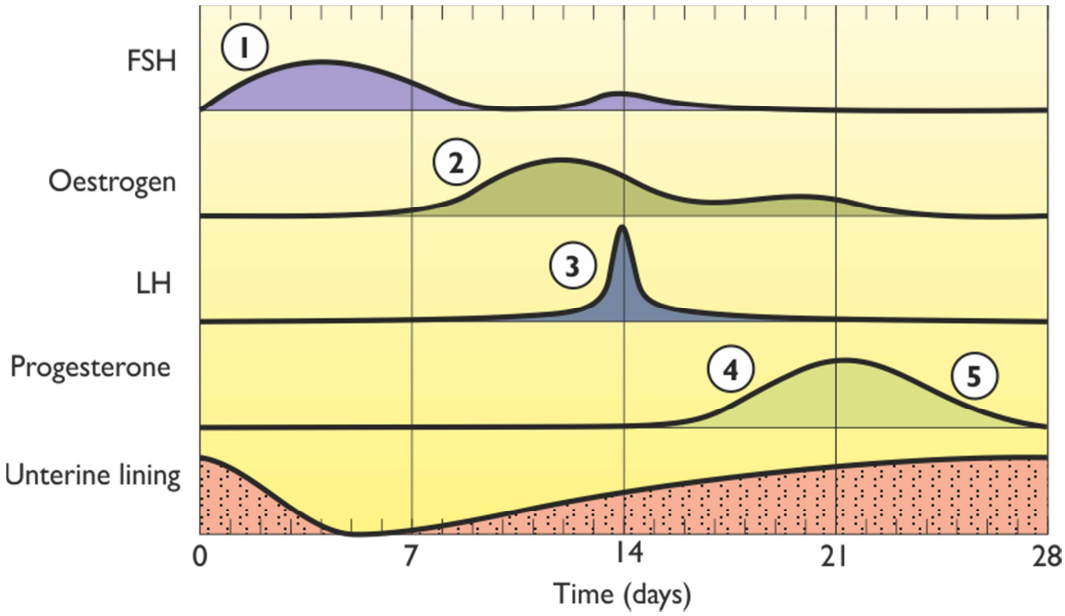
The menstrual cycle in humans is controlled by four hormones secreted by two glands.

- The pituitary gland, below the hypothalamus in the brain, secretes the hormones follicle stimulating hormone (FSH) and luteinising hormone (LH), which target the ovaries.
- The ovaries are endocrine organs as well as creating and releasing ova. They secrete the hormones oestrogen and progesterone, which target the pituitary gland and the uterus.

The effects of these four hormones are shown in this diagram.



Each hormone affects the release of other hormones by negative and positive feedback loops, so each hormone is produced in sequence. This chart shows how the concentration of each hormone in the blood changes throughout a 28-day cycle.



1. FSH is secreted by the pituitary glands, and stimulates the development of a Graafian follicle in one of the ovaries. This follicle contains a single ovum cell surrounded by other cells.
2. The follicle secretes oestrogen, which is a transcription factor, stimulating protein synthesis (p64). It causes the uterus to rebuild the endometrium wall that has been shed during menstruation. Oestrogen also affects the pituitary gland, initially inhibiting the release of FSH. However, as the follicle gradually develops, the concentration of oestrogen in the blood rises, and it starts to stimulate the release of FSH and LH by the pituitary gland. FSH and LH are protein hotmones.
3. The sudden surge of LH at about day 14 causes the fully developed follicle to burst, release the ovum into the oviduct – ovulation. LH also stimulates the follicle to develop into a body called the corpus luteum, which secretes progesterone.
4. Progesterone stimulates the uterus to complete the development of the endometrium wall, which is now ready to receive an embryo. Progesterone also inhibits the release of LH and FSH by the pituitary gland, which in turn stops the release of oestrogen and progesterone by the ovaries.
5. The corpus luteum degenerates over the next 10 days, , so less progesterone is secreted. When the concentration of progesterone drops low enough, menstruation is triggered. The inhibition of the pituitary gland is also removed, so FSH starts to be released and the cycle starts again.

Pregnancy

The ovum is fertilised as it travels down the oviduct and the embryo implants in the endometrium at about 7 days. This is when pregnancy begins. The developing embryo secretes a hormone called human chorionic gonadotropin (hCG). hCG works like LH, so it stops the corpus luteum degenerating and so progesterone continues to be produced and there is no menstruation. Progesterone also stops the pituitary releasing FSH, so no more ova are matured during pregnancy. Pregnancy test kits test for the presence of small amounts of hCG in the urine.

Plant Responses

Plants don't have nervous or endocrine systems, but they can sense and respond to stimuli. For example plants can grow towards the light, their roots can grow towards water and they can flower at certain times of the year. Many of these responses are directional growth responses, called tropisms. Tropisms can be positive (growing towards the stimulus) or negative (growing away from the stimulus) and occur in response to a variety of stimuli:

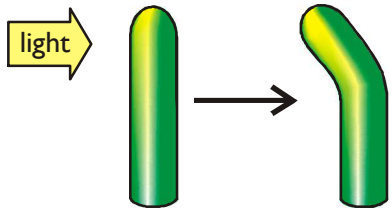
Name of tropism	Stimulus	Examples
Phototropism	Light	Young shots (positive) roots (negative)
Gravitropism	Gravity	Roots (positive)
Hydrotropism	Water	Roots (positive)
Chemotropism	Chemicals	Pollen tube (positive)
Thigmotropism	Touch	Vines, bindweed (positive)

Phototropisms have been studied in detail by many scientists over the last 150 years, and gradually some details of the mechanisms have been revealed.

Darwin's Experiments

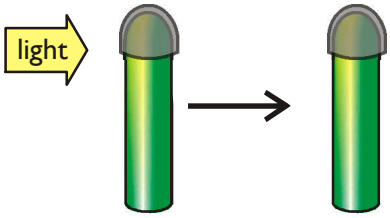
Some of the earliest experiments on phototropism were carried out by Charles Darwin and his son Francis in the 1880s. They observed the effects of directional light on young grass seedlings. When grass seeds germinate, the leaves are initially protected by a cylindrical sheath called a coleoptile. These coleoptiles are ideal for studying phototropism.

Experiment 1



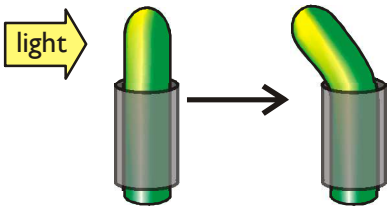
This shows that young shoots bend towards directional light.

Experiment 2



If an opaque foil cap is placed over the tip, bending does not occur.

Experiment 3



If an opaque foil collar is placed below the tip, bending does occur as usual.

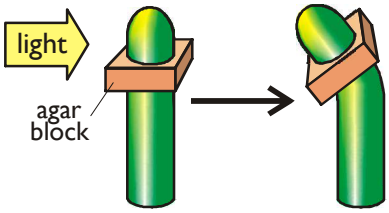
From these experiments the Darwins concluded that light is detected only at the shoot tip, and an "influence" was transmitted from the tip down the shoot to cause bending further down. They had no idea what this "influence" was.

Went's Experiments

In the 1920s the Dutch scientist Frits Went carried out a series of elegant experiments on oat seedling coleoptiles that showed that the "influence" was a chemical. He knew that seedlings with their tips cut off would not grow, while seedlings with an intact tip would. So he cut off the tips off growing seedlings,

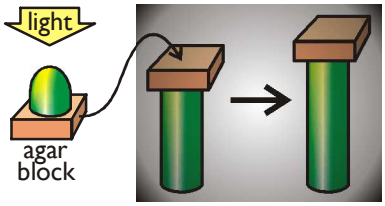
placed them on small blocks of agar for two hours, and then placed the agar blocks on top of cut seedlings in the dark. Agar jelly allows chemicals to diffuse through but contains no living cells, and a control experiment using agar blocks that had not been in contact with shoot tips did not promote growth.

Experiment 4



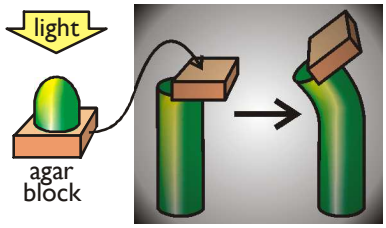
Coleoptiles still show phototropism if they are cut below the tip and a block of agar is inserted between the tip and the rest of the shoot. This shows that the "influence" can diffuse through agar – it must be a chemical.

Experiment 5



The agar block stimulated growth in the cut coleoptile. A control experiment showed that this wasn't the agar itself, so a chemical must have diffused from the tip into the agar, and then from the agar into the cut coleoptile.

Experiment 6



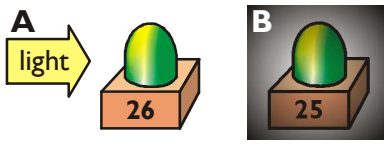
When the agar block was placed asymmetrically on the cut shoot, the shoot bent away from the agar block. This showed that bending was due to growth on one side of the shoot.

Went concluded that a chemical substance had diffused from the shoot tip into the agar, and that this substance stimulated growth further down the shoot. He called this substance auxin.

Brigg's Experiments

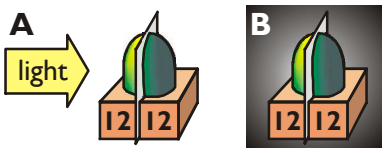
In the 1960s Winslow Briggs used Went's method (experiment 6) to assay the amount of auxin in plant material. He found that the greater the amount of auxin, the greater the bending. In the experiments below the numbers refer to the angle of bending and therefore to the amount of auxin.

Experiment 7



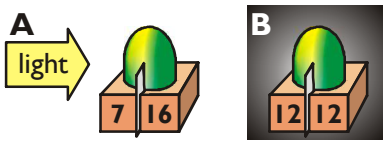
This shows that the total amount of auxin produced by the shoot tip is the same in the light or dark. It suggests that auxin moves from one side to the other rather than being made or destroyed. Note that in 7A the auxin can diffuse throughout the agar, so this doesn't tell us where in the tip the auxin is coming from.

Experiment 8



In this case the shoot tips and the agar blocks are split by thin sheets of mica, which stop diffusion. Again this shows that the total amount of auxin is the same in the dark or light. Result 8A is consistent with previous experiments that had shown that the phototropic response is removed if the tip is split by a thin sheet, presumably because the sheet prevents movement of auxin.

Experiment 9



In this experiment the mica sheet only splits the agar and the lower half of the tip. The total amount of auxin is the same in both cases, but the uneven distribution of auxin in 9A suggests that auxin is moving from the light to the dark side. This uneven distribution of auxin explains why shoots bend towards the light.

These discoveries are summarised in the Cholodny-Went theory, which states that auxin is synthesized in the coleoptile tip; asymmetric illumination is detected by the coleoptile tip and this causes auxin to move into the darker side; auxin diffuses down the coleoptile; and the higher auxin concentration on the darker side causes the coleoptile to bend toward the light source.

Although there is a lot of evidence supporting this theory, it is by no means certain and some recent studies using radioactive tracers have found no difference in auxin concentration on the dark and light sides of a shoot. An alternative mechanism is that auxin is present on both sides but is somehow inhibited on the light side, so there is little growth.

How does auxin work?

In 1934 auxin was identified as a compound called indoleacetic acid, or IAA. It was the first of a group of substances controlling plant growth responses called plant growth regulators, or PGRs. PGRs are a bit like animal hormones, but the term hormones is not used for plants because PGRs are not made in glands and do not travel in blood. IAA is hydrophobic so it can diffuse through cell membranes and so move around the plant.

IAA stimulates growth by:

1. Binding to a receptor protein in the target cell membranes and activating a proton pump.
2. This pump pumps protons (hydrogen ions) from the cytoplasm of these cells to their cell walls.
3. The resulting decrease in pH activates an enzyme that breaks the bonds between cellulose microfibrils.
4. This loosens the cell wall and so allows the cell to elongate under the internal turgor pressure.

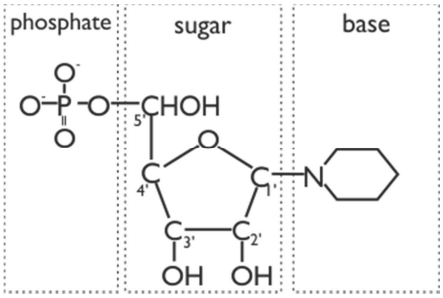
Genetics

In unit 2 we learnt about the structure and replication of DNA. Now we shall learn more about how DNA controls characteristics by controlling protein synthesis. This process is called gene expression and is summarised in the "central dogma" of genetics:

As this diagram shows, expression is split into two parts – transcription and translation, and it involves another nucleic acid – RNA (ribonucleic acid). RNA is a nucleic acid like DNA, but with 4 differences:

- RNA is made of ribose nucleotides instead of deoxyribose nucleotides
- RNA has the base uracil instead of thymine (so the four bases in RNA are A, U, C and G).
- RNA is single stranded (though it can fold into 3-dimentional structures)
- RNA is shorter than DNA

This shows a ribose nucleotide, found in RNA. Note the OH group on C3 of the sugar.



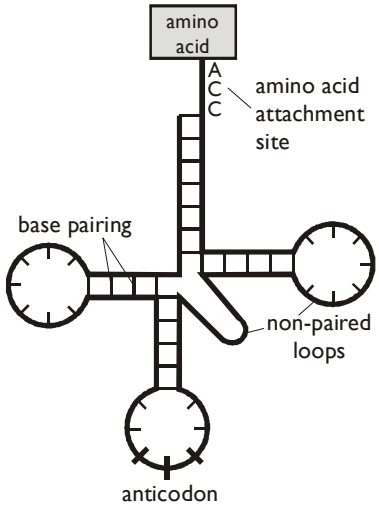
There are three kinds of RNA, with three different jobs:

Messenger RNA (mRNA)

mRNA carries the "message" that codes for a particular protein from the nucleus (where the DNA master copy is) to the cytoplasm (where proteins are synthesised). It is single stranded and just long enough to contain one gene only (about 1000 nucleotides). It has a short lifetime and is degraded soon after it is used.

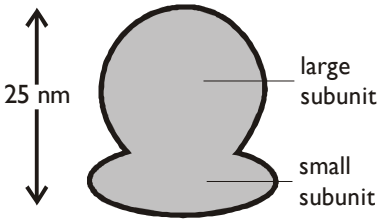
Transfer RNA (tRNA)

tRNA is an "adapter" that matches amino acids to their codon. tRNA is only about 80 nucleotides long, and it folds up by complementary base pairing to form a looped clover-leaf structure. At one end of the molecule there is always the base sequence ACC, where the amino acid binds. On the middle loop there is a triplet nucleotide sequence called the anticodon. There are 64 different tRNA molecules, each with a different anticodon sequence complementary to the 64 different codons. The amino acids are attached to their tRNA molecule by specific aminoacyl tRNA synthase enzymes. These are highly specific, so that each amino acid is attached to a tRNA adapter with the appropriate anticodon.



Ribosomal RNA (rRNA)

rRNA together with proteins forms ribosomes, which are the site of mRNA translation and protein synthesis. Ribosomes have two subunits, small and large, and are assembled in the nucleolus of the nucleus and exported into the cytoplasm. rRNA is coded for by numerous genes in many different chromosomes. Ribosomes free in the cytoplasm make proteins for use in the cell, while those attached to the RER make proteins for export.



The Genetic Code

In unit 2 we learnt that the sequence of bases on DNA codes for the sequence of amino acids in proteins. The sequence is read in groups of three bases called codons, where each codon codes for one amino acid. The meaning of each of the 64 codons is called the genetic code.

The Genetic Code (mRNA codons)

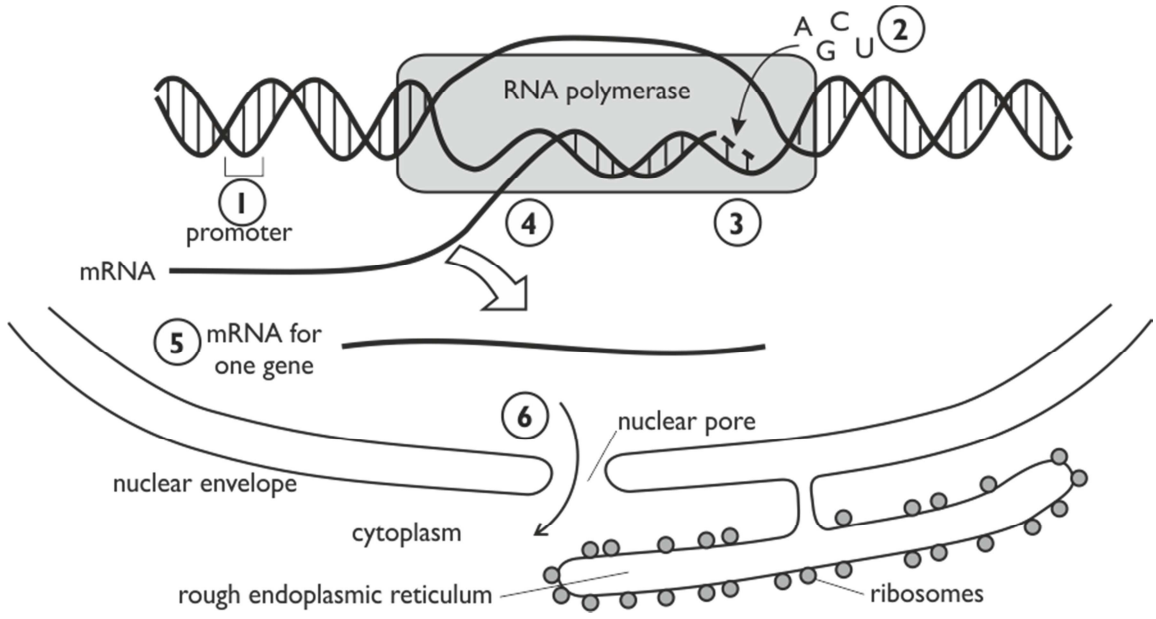
UUU phe UUC phe UUA leu UUG leu	CUU leu CUC leu CUA leu CUG leu	AUU ile AUC ile AUA ile AUG start/met	GUU val GUC val GUA val GUG val
UCU ser UCC ser UCA ser UCG ser	CCU pro CCC pro CCA pro CCG pro	ACU thr ACC thr ACA thr ACG thr	GCU ala GCC ala GCA ala GCG ala
UAU tyr UAC tyr UAA stop UAG stop	CAU his CAC his CAA gln CAG gln	AAU asn AAC asn AAA lys AAG lys	GAU asp GAC asp GAA glu GAG glu
UGU cys UGC cys UGA stop UGG trp	CGU arg CGC arg CGA arg CGG arg	AGU ser AGC ser AGA arg AGG arg	GGU gly GGC gly GGA gly GGG gly

There are several interesting points from this code:

- The code is degenerate, i.e. there is often more than one codon for an amino acid.
- The degeneracy is on the third base of the codon, which is therefore less important than the others.
- One codon means "start" i.e. the start of the gene sequence. It is AUG, which also codes for methionine. Thus all proteins start with methionine (although it may be removed later). AUG in the middle of a gene simply codes for methionine.
- Three codons mean "stop" i.e. the end of the gene sequence. They do not code for amino acids.
- The code is read from the 5' to 3' end of the mRNA, and the protein is made from the N to C terminus ends.

Transcription - RNA Synthesis

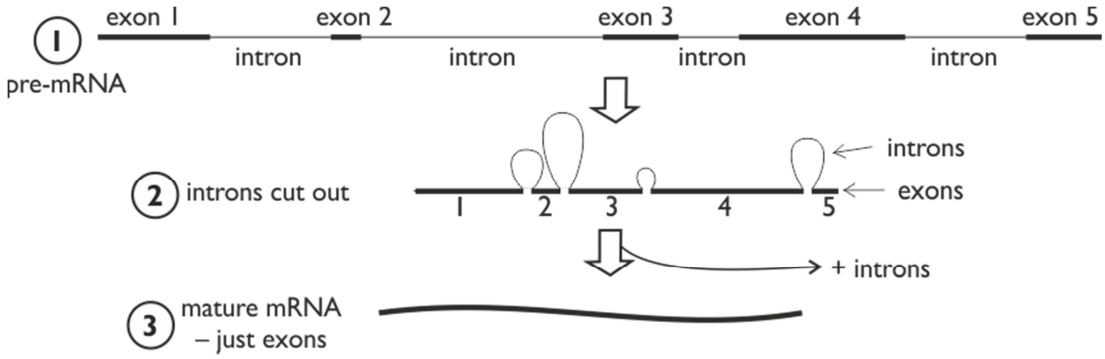
DNA never leaves the nucleus, but proteins are synthesised in the cytoplasm, so a copy of each gene is made to carry the "message" from the nucleus to the cytoplasm. This copy is mRNA, and the process of copying is called transcription.



1. The start of each gene on DNA is marked by a special sequence of bases called the promoter.
2. The RNA molecule is built up from the four ribose nucleotides (A, C, G and U) in the nucleoplasm. The ribose nucleotides attach themselves to the bases on the DNA by complementary base pairing, just as in DNA replication. However, only one strand of RNA is made. The DNA strand that is copied is called the template strand (or antisense strand). The other strand is a complementary copy, called the non-template strand (or sense strand).
3. The new nucleotides are joined to each other by strong covalent phosphodiester bonds by the enzyme RNA polymerase.
4. Only about 8 base pairs remain attached at a time, since the mRNA molecule peels off from the DNA as it is made. A winding enzyme rewinds the DNA.
5. At the end of the gene the transcription stops, so the mRNA molecule is just the length of the gene.
6. The mRNA diffuses out of the nucleus through a nuclear pore into the cytoplasm. There, it attaches to ribosomes for translation. It usually doesn't have far to go to find a ribosome, as many are attached to the rough endoplasmic reticulum, which is contiguous with the nuclear envelope.

Post-Transcriptional Modification

In unit 2 we learnt that Eukaryotic DNA contains non-coding sequences as well as the coding sequences. The non-coding sequences within genes are called introns, while the coding sequences are called exons. The introns need to be removed before the mRNA can be translated into protein. This removal is called post-transcriptional modification.

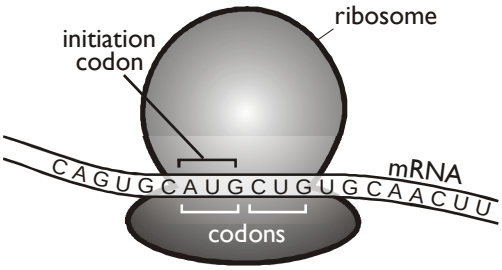


1. The initial mRNA that is transcribed is called the primary transcript or pre-mRNA. Pre-mRNA is an exact copy of the gene on the DNA, so it contains exons and introns.
2. The introns in the mRNA are cut out and the exons are joined together by enzymes in a process called splicing. Some of this splicing is done by the RNA intron itself, acting as an RNA enzyme. The recent discovery of these RNA enzymes, or ribozymes, illustrates what a diverse and important molecule RNA is. Other splicing is performed by RNA/protein complexes called snurps.
3. The result is a shorter mature RNA containing only exons. The introns are broken down.

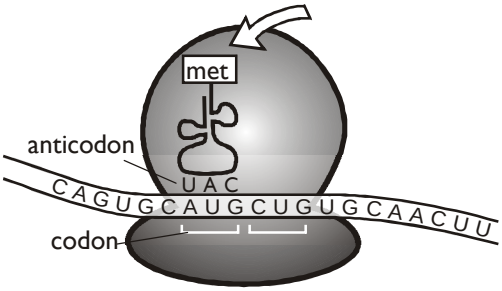
Prokaryotic DNA does not have introns, so mRNA processing is not needed.

Translation - Protein Synthesis

1. A ribosome attaches to the mRNA at an initiation codon (AUG). The ribosome encloses two codons.

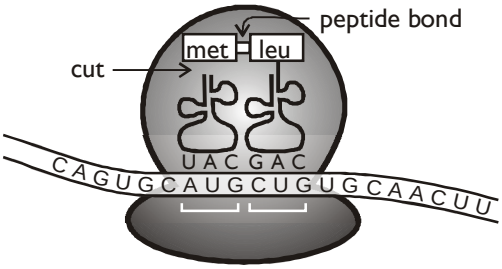


2. The first tRNA molecule with an amino acid attached (met-tRNA) diffuses to the ribosome. Its anticodon attaches to the first mRNA codon by complementary base pairing.

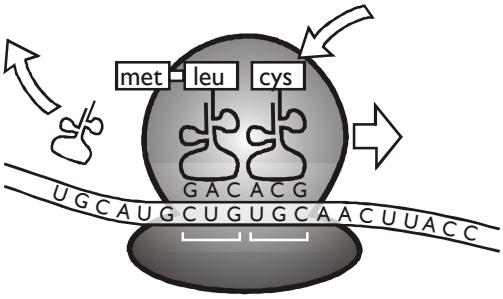


3. The next amino acid-tRNA attaches to the adjacent mRNA codon (CUG, leu in this case) by complementary base pairing.

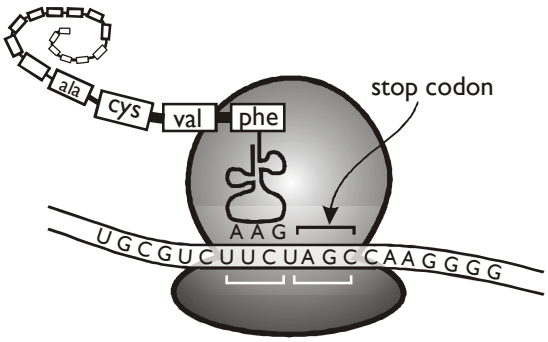
4. The bond between the amino acid and the tRNA is cut and a peptide bond is formed between the two amino acids. These operations are catalysed by enzymes in the ribosome called ribozymes.



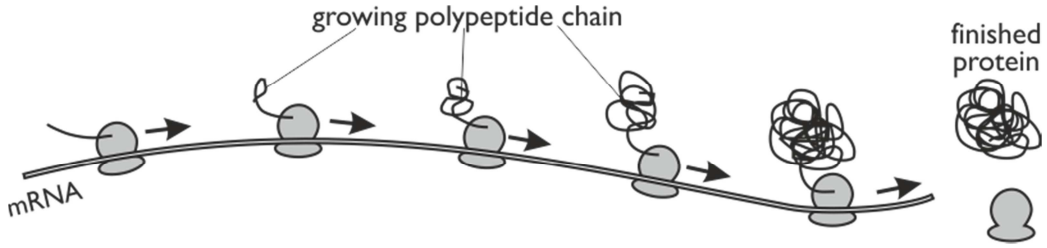
5. The ribosome moves along one codon so that a new amino acid-tRNA can attach. The free tRNA molecule leaves to collect another amino acid. The cycle repeats from step 3.



6. The polypeptide chain elongates one amino acid at a time, and peels away from the ribosome, folding up into a protein as it goes. This continues for hundreds of amino acids until a stop codon is reached, when the ribosome falls apart, releasing the finished protein.



A single piece of mRNA can be translated by many ribosomes simultaneously, so many protein molecules can be made from one mRNA molecule. A group of ribosomes all attached to one piece of mRNA is called a polyribosome, or a polysome.



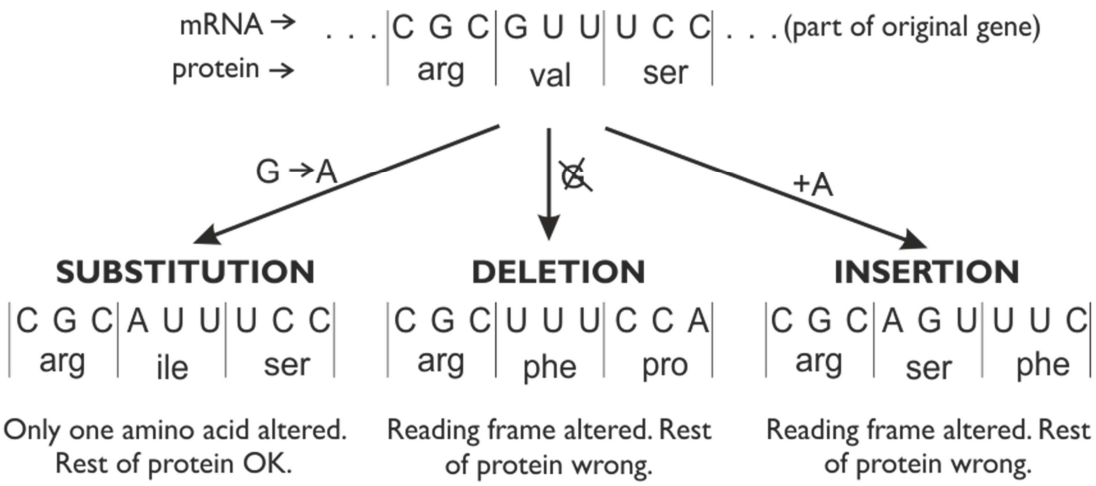
Post-Translational Modification

In eukaryotes, proteins often need to be altered before they become fully functional. Because this happens after translation, it is called post-translational modification. Modifications are carried out by other enzymes and include: chain cutting, adding methyl or phosphate groups to amino acids, adding sugars (to make glycoproteins) or lipids (to make lipoproteins).

Gene Mutations

Mutations are changes in genes, which are passed on to daughter cells. DNA is a very stable molecule, and it doesn't suddenly change without reason, but bases can change when DNA is being replicated. Normally replication is extremely accurate, and there are even error-checking procedures in place to ensure accuracy, but very occasionally mistakes do occur (such as a T-C base pair). So a mutation is a base-pairing error during DNA replication.

There are three kinds of gene mutation, shown in this diagram:



- Substitution mutations only affect one amino acid, so tend to have less severe effects. In fact if the substitution is on the third base of a codon it may have no effect at all, because the third base often doesn't affect the amino acid coded for (e.g. all codons beginning with CC code for proline). These are called silent mutations. However, if a mutation leads to a premature stop codon the protein will be incomplete and certainly non-functional. This is called a nonsense mutation.
- Deletion and insertion mutations have more serious effects because they are frame shift mutations i.e. they change the codon reading frame even though they don't change the actual sequence of bases. So all amino acids "downstream" of the mutation are wrong, and the protein is completely wrong and non-functional. However, the effect of a deletion can be cancelled out by a near-by insertion, or by two more deletions, because these will restore the reading frame. A similar argument holds for a substitution.

Mutation Rates and Mutagens

Mutations are normally very rare, which is why members of a species all look alike and can interbreed. However the rate of mutation is increased by chemicals or by radiation. These are called mutagenic agents or mutagens, and include:

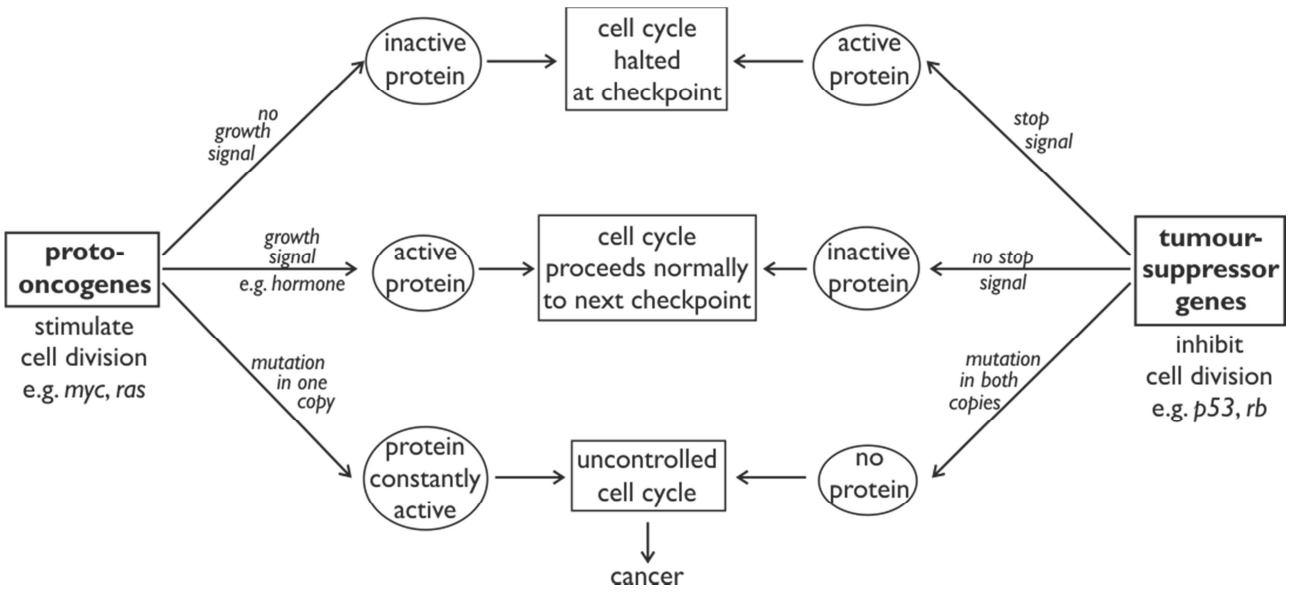
- High-energy, ionising radiation such as x-rays, ultraviolet rays, α, β, or γ rays from radioactive sources. This ionises the bases so that they don't form the correct base pairs. Note that low-energy radiation (such as visible light, microwaves and radio waves) doesn't have enough energy to affect DNA and so is harmless.
- Intercalating chemicals such as mustard gas (used as a weapon in warfare), which bind to DNA separating the two strands.
- Chemicals that react with the DNA bases such as benzene, nitrous acid, and tar in cigarette smoke.
- Viruses. Some viruses can change the base sequence in DNA causing genetic disease and cancer.

During the Earth's early history there were far more of these mutagens than there are now, so the mutation rate would have been much higher than now, leading to a greater diversity of life. Some of these mutagens are used today in research, to kill microbes or in warfare. They are often carcinogens since a common result of a mutation is cancer.

Mutations and Cancer

In unit 2 we learnt that mutations in some genes can cause cancer. Two of these genes are proto-oncogenes and tumour suppressor genes. More than 50% of cancers involve mutations in these genes, sometimes just a single-base substitution. These genes encode proteins that control the checkpoints in the cell cycle (see unit 2) and so are part of the normal tight control of cell division in organisms. Despite their names, these are not "cancer genes" – their normal job is to control growth by regulating when cells divide. Only when they mutate and fail in their job do cancers occur.

- **Proto-oncogenes** encode proteins that stimulate cell division when activated by growth signals, such as hormones. The activated proteins allow cells to pass the checkpoints in the cell cycle and so proceed to divide. A mutation in a proto-oncogene turns it into an oncogene, which makes too much active protein all the time, regardless of the growth signals. The cell divides unchecked, which can lead to cancer. A mutation in one copy of the gene is enough to cause cell proliferation; in other words the mutation is dominant.
- Tumour-suppressor genes encode proteins that inhibit cell division when activated by stop signals, such as damaged DNA. The activated proteins halt the cell cycle at the checkpoints and so stop division. Mutated tumour-suppressor genes stop making the protein so the cell can pass the checkpoints, regardless of stop signals. Both copies of the gene need to be mutated to cause cell proliferation; in other words the mutation is dominant.



The cell cycle contains several checkpoints and, for cancer to occur, all the checkpoints must be broken. So cancer cannot be caused by a single mutation, but instead by many different mutations accumulating. This explains why cancers are more likely to occur in older patients as there has been more time for cell cycle mutations to accumulate. Most adults will have some cells with mutated oncogenes or tumour-suppressor genes, but these won't necessarily go on to become cancerous.

Blank Page

Stem Cells

In unit 2 we saw that most body cells only express a few of their genes, so can only carry out a limited set of functions. These cells are specialised (or differentiated). In animals specialised cells are irreversibly differentiated, so they and their daughter cells will keep their specialisation and cannot become any other kind of cell. However, there are a few cells that remain undifferentiated and so can become any type of cell. These are called stem cells.

Stem cells possess two key properties:

- Stem cells are potent – they have the potential to differentiate into specialized cell types.
- Stem cells are immortal – they can divide indefinitely.

Because of these two properties stem cells can be grown in the lab (in culture) and used for research and medicine.

Where do they come from?

- **Embryo stem cells** are grown *in vitro* from human embryos called blastocysts (i.e. five day old embryos containing around 150 cells, before implantation). These embryos are created by in vitro fertilisation (IVF) to help infertile couples reproduce, but any "spare" embryos, no longer needed for reproduction, can be used to create stem cells, with the informed consent of the donor couple. These stem cells are totipotent, so can be differentiated into any cell type for clinical use. However, there remains a debate about whether it is ethical to use human embryos for this purpose. Embryonic stem cell research had been banned across Europe by the European parliament, although the UK government does allow such research in the UK.
- **Adult stem cells** are extracted from certain tissues of the body. It is thought that most organs and tissues maintain a small number of undifferentiated stem cells, which the body uses to replace and repair damaged tissue. Tissues where stem cells have been found include the brain, bone marrow, blood vessels, muscle, skin, heart, gut and liver. These stem cells are multipotent, which means they can differentiate into their own family of cells (e.g. blood cells, muscle cells), but not others. The use of these cells has no ethical issues, and there are also no problems of rejection if the stem cells are taken from the patient's own body. However, they are difficult to find and difficult to grow in culture.
- **Induced pluripotent stem cells** (iPSCs) are normal, specialised adult cells that have been genetically reprogrammed to become undifferentiated, pluripotent stem cells. iPSCs are a very new development, still at the research stage, but are likely to solve the problems of both adult and embryo stem cells, making the use of embryo stem cells obsolete.

Stem Cells in Medicine

The discovery of stem cells in the 1950s immediately led to suggestions that they could be used clinically. The idea is to transplant tissue grown from stem cells into a patient, where it would grow and replace damaged tissue. Some potential examples of these cell-based therapies are shown in the table:

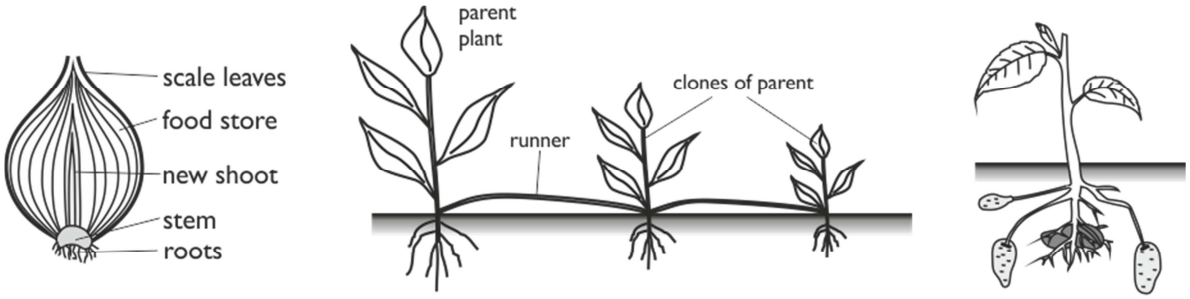
Type of cell	Disease that could be treated
Myocytes	Myocardial infarction
Pancreatic β cells	Type I diabetes
Skeletal muscle cells	Muscular dystrophy
Blood cells	Leukaemia
Nerve cells	Parkinson's disease, multiple sclerosis, strokes, paralysis due to spinal injury
Skin cells	Burns
Bone cells	Osteoporosis
Cartilage cells	Osteoarthritis
Retina cells	Macular degeneration

Since the tissues could be grown from the patient's own adult stem cells, there is no risk of rejection, unlike conventional transplants. Blood cell-forming (hematopoietic) stem cells from bone marrow are already being used successfully to treat leukaemia (cancer of white blood cells); and heart disease and diabetes have be treated in mice. In addition, human stem cells grown in culture are being used to test the effects of new drugs, without harming humans or animals.

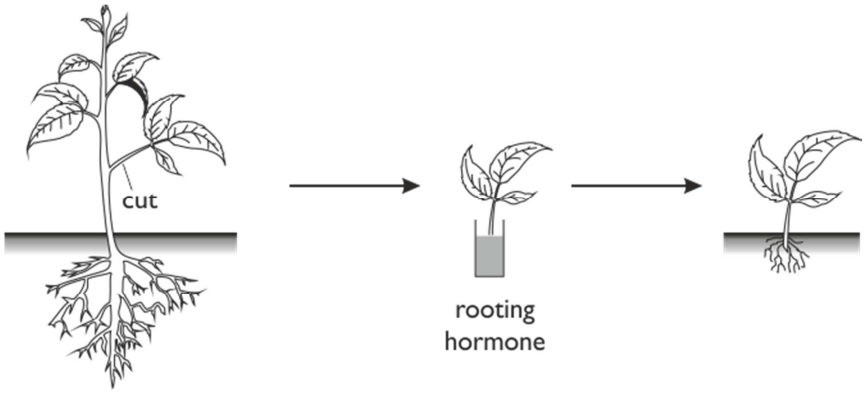
Stem Cells in Plants

Plants, unlike animals, grow only from meristem tissue in their shoots and roots. This meristem tissue contains totipotent stem cells that have the ability to develop into any other plant tissue, including whole plants. Meristem cells have small vacuoles, undeveloped chloroplasts called proplastids and thin cell walls. In most plants differentiated cells are unable to divide, so division and growth only takes place in the meristems. These meristematic totipotent plant cells are used naturally and artificially:

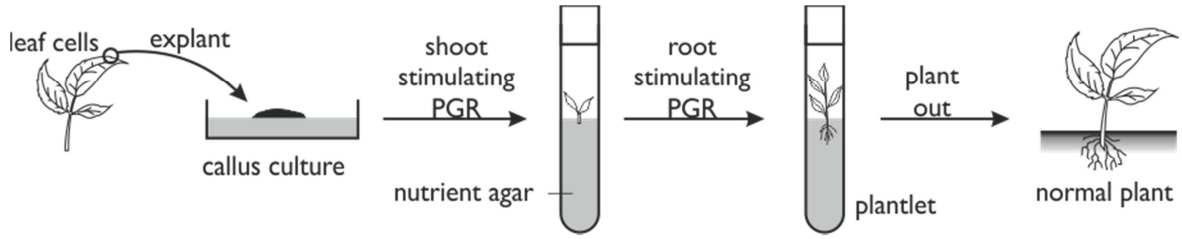
- **Vegetative Propagation.** This is the natural method of asexual reproduction used by plants, where a new plant grows from meristem tissue in a vegetative (i.e. not reproductive) part of the plant. There are numerous forms of vegetative reproduction, including: bulbs (e.g. onion, daffodil); corms (e.g. crocus, gladiolus); rhizomes (e.g. iris, couch grass); stolons (e.g. blackberry, bramble); runners (e.g. strawberry, creeping buttercup); tubers (e.g. potato, dahlia); tap roots (e.g. carrot, turnip) and tillers (e.g. grasses). Many of these methods are also perenating organs, which means they contain a food store and are used for survival over winter as well as for asexual reproduction. Since vegetative reproduction relies entirely on mitosis, all offspring are clones of the parent.



- **Cuttings.** This is an old artificial method of cloning plants. Parts of a plant stem (or even leaves) are cut off and simply replanted in wet soil. Each cutting produces roots and grows into a complete new plant, so the original plant can be cloned many times. The cut stems contain meristematic cells that start dividing and re-differentiating to become the various root tissues. This division and differentiation is helped if the cuttings are dipped in "rooting hormone" (a plant growth regulator, such as IAA). Many flowering plants, such as geraniums, African violet and chrysanthemums are reproduced commercially by cuttings.



- **Tissue Culture (or micropropagation).** This is a more modern, and very efficient, way of cloning plants. Small samples of plant tissue, called an explant, can be grown on agar plates in the laboratory in much the same way that bacteria can be grown. Any plant tissue can be used for tissue culture, even differentiated, non-meristematic tissue (e.g. from a leaf). The plant tissue can be separated into individual cells, each of which can grow into a mass of undifferentiated cells called a callus. If the correct plant growth regulators (PGRs, p47) are added these cells can develop into whole plantlets, which can eventually be planted outside, where they will grow into normal-sized plants. Conditions must be kept sterile to prevent infection by microbes.

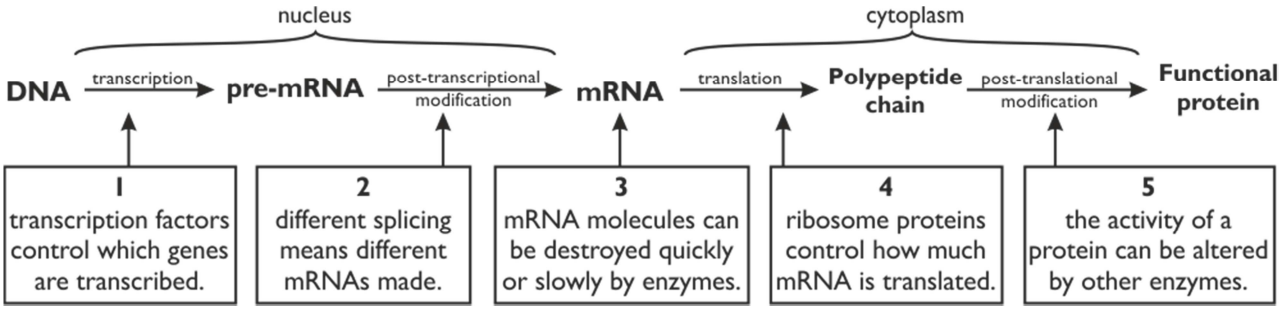


With micropropagation thousands of clones of a particularly good plant can be made quickly and in a small space. Micropropagation is used on a large scale for fruit trees, ornamental plants and plantation crops such as oil palm, date palm, sugar cane and banana, that cannot be asexually propagated by other means. Because whole plants can be grown from single cells, this technique could be combined with genetic engineering to grow genetically modified plants from a single genetically-modified cell.

All these methods of reproduction using plant stem cells result in clones. This is useful for exactly reproducing plants with desired characteristics, but the initial variation to create those desired characteristics can only be produced using sexual reproduction. Producing plants sexually from seeds is also cheaper, easier and very low tech.

Control of Gene Expression

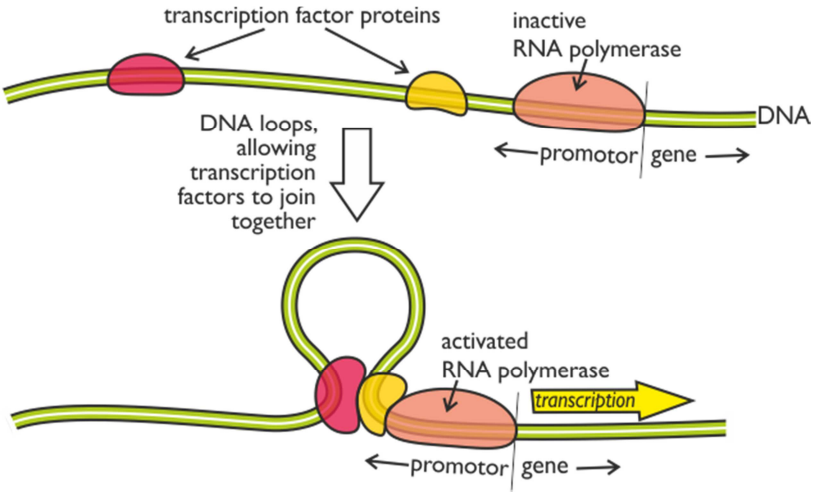
We've seen how genes are expressed through transcription and translation into proteins, which give cells their functions and properties. But cells don't express all their genes all the time. Gene expression can be switched on or off by other genes (e.g. during embryo development), by stimuli (e.g. light, injury, nutrients), or by hormones (e.g. growth hormone, oestrogen). There are five control points along the gene expression pathway:



We'll look at step 1 (transcription factors) and step 3 (destruction of mRNA) in more detail.

Control of Transcription by Transcription Factors

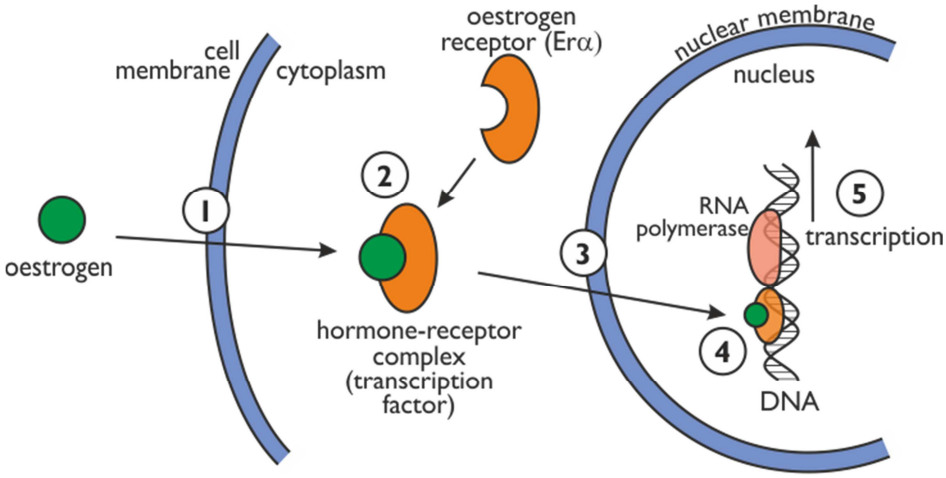
Transcription is the most important control, since it is the earliest and most efficient point, so mRNA is not made if it's not needed. We've seen that, for transcription to happen, the enzyme RNA polymerase must bind to the DNA molecule just upstream of the gene, at a region called the promoter. However, RNA polymerase binds only weakly at first and it needs the assistance of a number of other DNA-binding proteins, called transcription factors, before it can start transcribing the gene. The transcription factors also bind to the promoter region, upstream of the DNA polymerase. Each transcription factor protein has a specific binding site that binds to a particular DNA sequence in the promoter. Once bound to DNA, the different transcription factors bind together as the DNA molecule bends and loops. The combined transcription factors activate RNA polymerase, which can now move along the DNA molecule, transcribing the gene.



This example shows two transcription factors, but in practice there can be over a dozen different proteins involved. Some promote transcription, while others inhibit it, so transcription factors provide a very flexible method of control. Since transcription factors are proteins, they are synthesised in the cytoplasm and transported into the nucleus to bind to DNA. And of course their own production is controlled by transcription factors!

Oestrogen

Steroid hormones control protein synthesis (p22), and they do this via transcription factors. For example oestrogen is secreted by the ovaries and stimulates synthesis of many proteins in target tissues such as the uterus, hypothalamus and breasts.



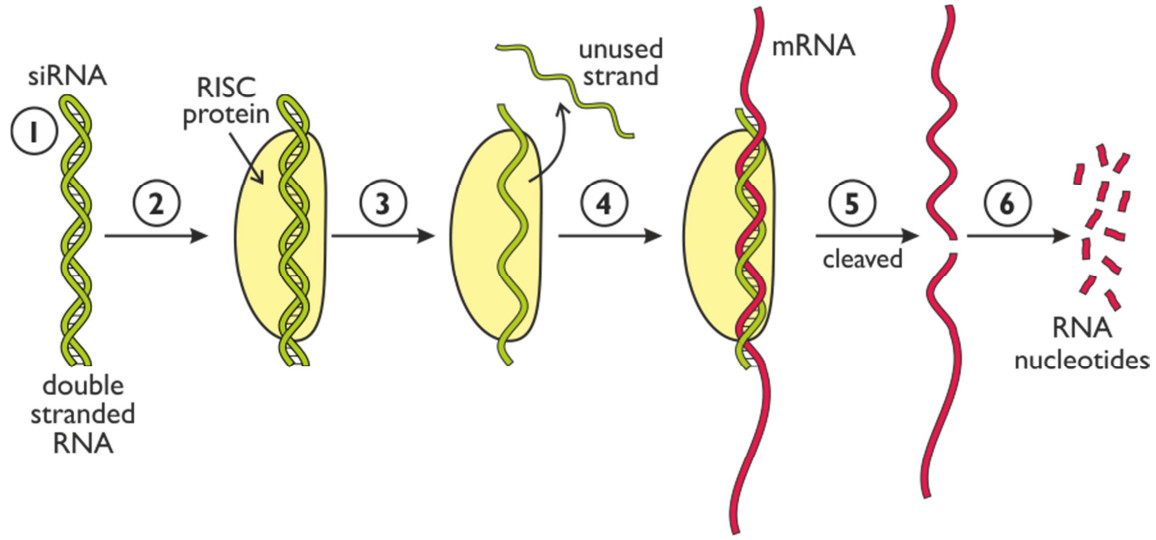
1. Oestrogen, like all steroids hormones, is a lipid, so crosses the cell membrane by lipid diffusion and enters the cytoplasm.
2. Oestrogen binds to its receptor protein (called ERα) in the cytoplasm to form a hormone-receptor complex. This complex is now a transcription factor.
3. The active transcription factor diffuses into the nucleus through a nuclear pore.
4. In the nucleus the transcription factor binds to a specific base sequence on a DNA promoter, upstream of RNA polymerase.
5. This binding stimulates RNA polymerase to transcribe genes and so stimulates protein synthesis.

Control of mRNA by RNA Interference

mRNA molecules have a fairly short lifetime – they are degraded by enzymes after they have been used for translation. The quicker the mRNA is broken down, the less protein can be made, and this is another point where gene expression can be controlled. This control is carried out by special RNA molecules called small interfering RNA (siRNA) and the process is called RNA interference (RNAi).

1. siRNA is a short double-stranded RNA molecule, about 20 base pairs long. It is made by special regulatory genes. These genes are transcribed as normal to make single-stranded RNA, which then folds back on itself by complementary base pairing to make a hairpin-like double-stranded molecule.

2. In the cytoplasm siRNA binds to a protein called the RNA-induced silencing complex (RISC).
3. RISC breaks the double-stranded siRNA into its separate strands. One strand remains attached to the RISC protein, while the other strand is discarded.



4. The RISC-RNA complex now binds to mRNA molecules in the cytoplasm by complementary base pairing. Any mRNA molecules with a base sequence complementary to the 20-base siRNA sequence will bind.
5. This binding causes RISC to cut the mRNA molecule in two.
6. This cleaved mRNA can no longer be used in translation, and is broken down by nuclease enzymes.

RNA interference seems to be mainly used by cells in defence against viral attack. Viruses enter a host cell and use the host cell's ribosomes to translate viral mRNA and so make more viral proteins. In defence, host cells make siRNA with complementary sequences to viral mRNA, so preventing synthesis of viral proteins.

Applications of RNA interference

Since the discovery of RNA interference in 1998, scientists quickly realised that they could use siRNA to "silence" genes. A 20-base pair siRNA molecule can be synthesised in the laboratory with a sequence that is complementary to part of the gene that is to be silenced. This siRNA is then introduced into cells to observe the effect of silencing that gene. Applications include:

- Identifying the function of genes in cells or simple organisms by silencing the gene and observing the functional effects.
- Genetically-modifying crop plants by silencing undesirable genes e.g. those that make toxins or allergens. In the Flavr Savr tomato a gene causing ripening was silenced, causing the tomatoes to stay fresh for longer.
- Treating human patients to silence essential genes in cancer cells, or in antiviral therapies e.g. silencing the gene for the receptor protein used by HIV.

Biotechnology

The United Nations Convention on Biological Diversity defines biotechnology as:

"Any technological application that uses biological systems, living organisms, or derivatives thereof, to make or modify products or processes for specific use."

In practice biotechnology usually refers to the application of molecular (DNA) biology in the laboratory. Biotechnology has applications in:

- **research** e.g. human genome project
- **medicine** e.g. genetically-engineered drugs, gene therapy
- **agriculture** e.g. improving crops
- **industry** e.g. manufacturing enzymes, biosensors

A particular aspect of biotechnology is genetic engineering, which means altering the genes in a living organism to produce a Genetically Modified Organism (GMO) with a new genotype. Genetic engineering can include inserting a foreign gene from one species into another (forming a transgenic organism); altering an existing gene so that its product is changed or changing gene expression.

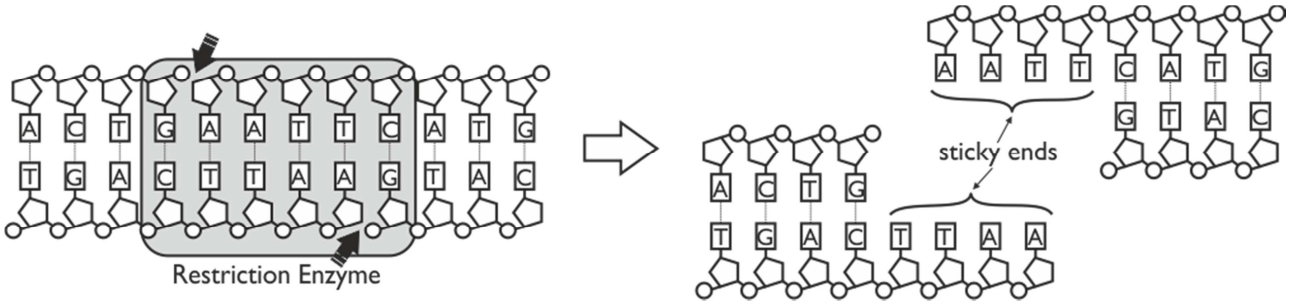
Techniques of Biotechnology

Modern biotechnology is possible due to the development of techniques from the 1960s onwards, which arose from our greater understanding of DNA and how it functions, following the discovery of its structure by Watson and Crick in 1953. This table lists the techniques that we shall look at in detail.

	Technique	Purpose	Type
1	Restriction Enzymes	To cut DNA at specific points, making small fragments	manipulating DNA
2	DNA Ligase	To join DNA fragments together	
3	Reverse transcriptase	To make DNA from mRNA	
4	PCR	To amplify very small samples of DNA	Analysing DNA
5	DNA Sequencing	To read the base sequence of a length of DNA	
6	Restriction Mapping	To make a map of a length of DNA	
7	Electrophoresis	To separate fragments of DNA	
8	Southern Blot	To look for specific sequences in DNA	
9	Genetic Fingerprinting	To compare different peoples' DNA	
10	Vectors	To carry DNA into cells and ensure replication	Cloning DNA
11	Transformation	To deliver a gene into a living cell	
12	Marker Genes	To identify cells that have been transformed	

1. Restriction Enzymes

These are enzymes that cut DNA at specific sites. They are properly called restriction endonucleases because they cut phosphodiester bonds in the middle of the polynucleotide chain. Some restriction enzymes cut straight across both chains, forming blunt ends, but most enzymes make a staggered cut in the two strands, forming sticky ends.

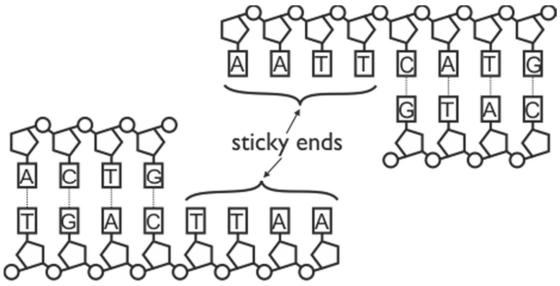


The cut ends are "sticky" because they have short stretches of single-stranded DNA with complementary sequences. These sticky ends will stick (or anneal) to another sticky end by complementary base pairing (i.e. with weak hydrogen bonds), but only if the sticky ends have both been cut with the same restriction enzyme so that they have complementary sequences. Restriction enzymes have highly specific active sites, and will only cut DNA at specific base sequences, 4-8 base pairs long, called recognition sequences. Recognition sequences are usually palindromic, which means that the sequence and its complement are the same but reversed (e.g. GAATTC has the complement CTTAAG).

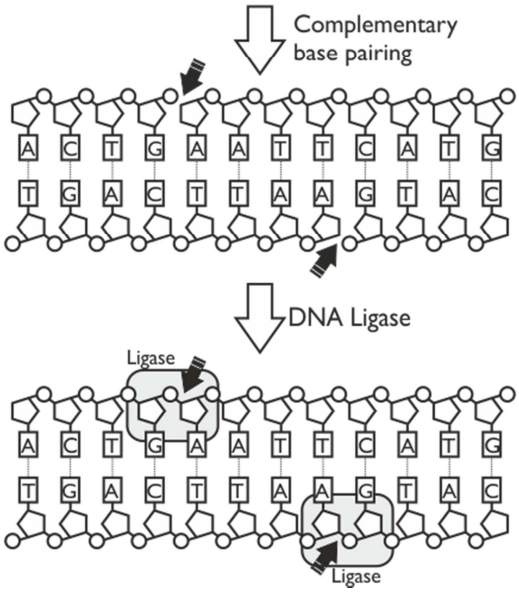
Restriction enzymes are produced naturally by bacteria as a defence against viruses (they "restrict" viral growth), but they are enormously useful in genetic engineering for cutting DNA at precise places ("molecular scissors"). Short lengths of DNA cut out by restriction enzymes are called restriction fragments. There are thousands of different restriction enzymes known, with over a hundred different recognition sequences. Restriction enzymes are named after the bacteria species they came from, so EcoR1 is from *E. coli* strain R, and HindIII is from *Haemophilis influenzae*.

2. DNA Ligase

This enzyme repairs broken DNA by joining two nucleotides in a DNA strand. Ligase is therefore a bit like DNA polymerase. It is commonly used in genetic engineering to do the reverse of a restriction enzyme, i.e. to join together complementary restriction fragments.



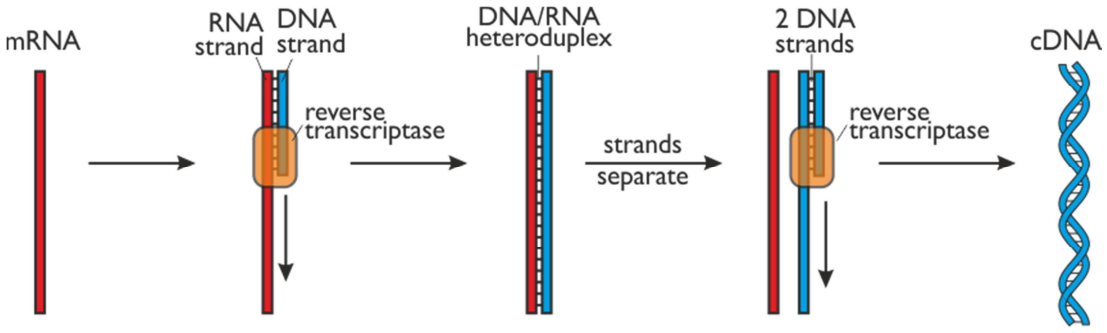
Two restriction fragments can anneal if they have complementary sticky ends, but only by weak hydrogen bonds, which can quite easily be broken, say by gentle heating. The backbone is still incomplete.



DNA ligase completes the DNA backbone by forming covalent phosphodiester bonds. Restriction enzymes and DNA ligase can therefore be used together to join lengths of DNA from different sources.

3. Reverse Transcriptase

The enzyme reverse transcriptase does the reverse of transcription: it synthesises DNA from an RNA template (so it is an RNA-dependent DNA polymerase enzyme). Reverse transcriptase is produced naturally by a group of viruses called the retroviruses (which include HIV), and it helps them to invade cells. In biotechnology reverse transcriptase is used to make an "artificial gene", called complementary DNA (cDNA), from an mRNA template as shown in this diagram:



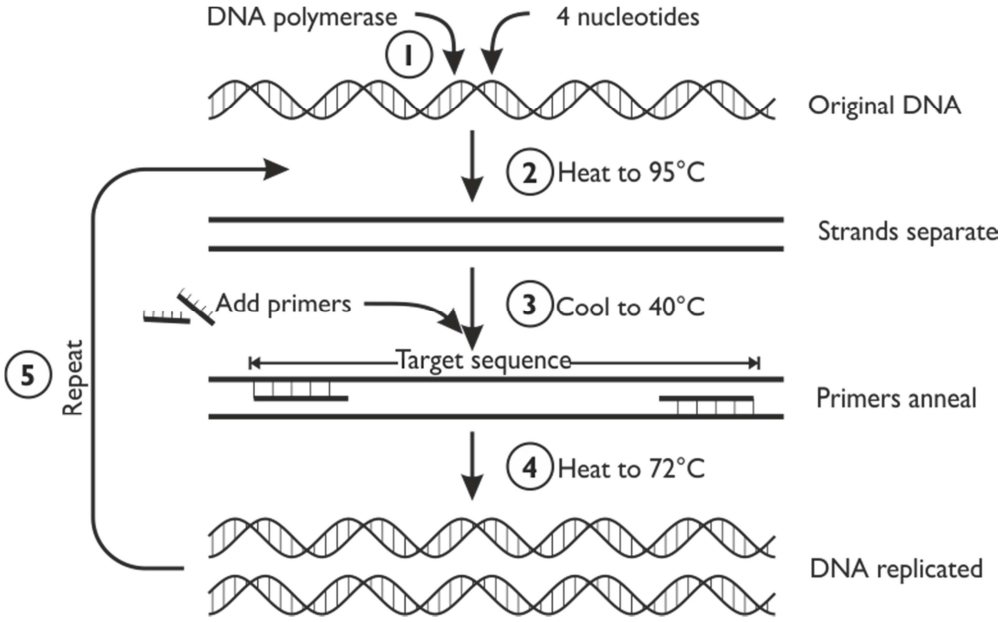
Mature mRNA (without introns) is extracted from cells and mixed with reverse transcriptase and DNA nucleotides. A new strand of DNA is synthesised, complementary to the mRNA strand, forming a double-stranded DNA/RNA "heteroduplex" molecule. The two strands of this molecule are then separated and reverse transcriptase now synthesises a second DNA stand, complementary to the first. The result is a normal double-stranded DNA molecule called cDNA. Note that the cDNA molecule is much shorter than the original gene in the organism's DNA (typically <50% the size), since the cDNA doesn't have introns. cDNA is therefore an "artificial gene".

Reverse transcriptase has several uses in biotechnology:

- It makes genes without introns. Eukaryotic genes with many introns are often too big to be incorporated into a bacterial plasmid, and bacteria are unable to splice out the introns anyway. The artificial cDNA gene is made from mRNA that already has the introns spliced out of it, so it can be expressed in bacteria.
- It makes a stable copy of a gene, since DNA is less readily broken down by enzymes than RNA.
- It makes genes easier to find. There are some 20 000 genes in the human genome, and finding the DNA fragment containing one gene out of this many is a very difficult task. However a given cell only expresses a few genes, so only makes a few different kinds of mRNA molecule. For example the β cells of the pancreas make insulin, so make lots of mRNA molecules coding for insulin. This mRNA can be isolated from these cells and used to make cDNA of the insulin gene.

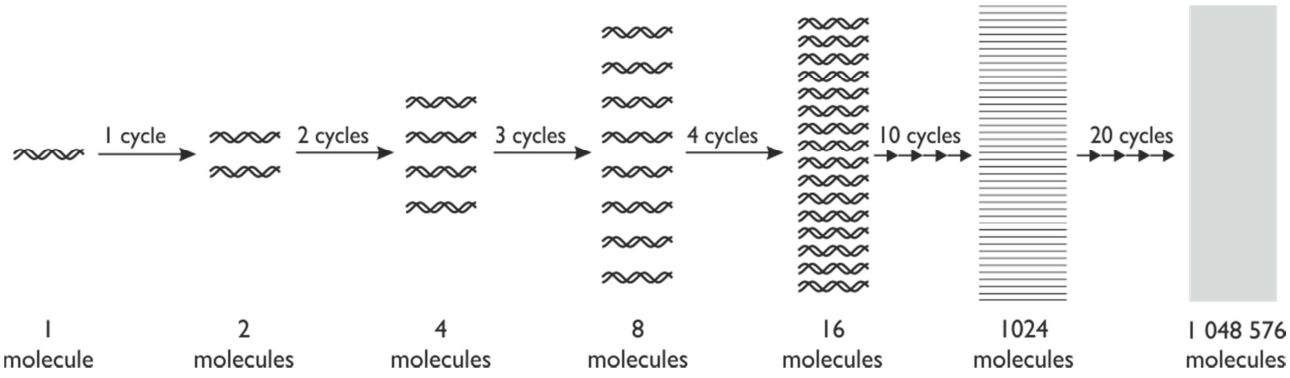
4. Polymerase Chain Reaction (PCR)

The polymerase chain reaction is a technique used to copy (or amplify) DNA samples as small as a single molecule. It was developed in 1983 by Kary Mullis, for which discovery he won a Nobel Prize in 1993. PCR is simply DNA replication in a test tube. If a length of DNA is mixed with the four nucleotides (A, T, C and G) and the enzyme DNA polymerase in a test tube, then the DNA will be replicated many times. The details are shown in this diagram:



1. Start with a sample of the DNA to be amplified, and add the four nucleotides and the enzyme DNA polymerase.
2. Heat to 95°C for two minutes to breaks the hydrogen bonds between the base pairs and separate the two strands of DNA. Normally (*in vivo*) the DNA double helix would be separated by an enzyme.
3. Add primers to the mixture and cool to 40°C. Primers are short lengths of single-stranded DNA (about 20 bp long) that anneal (i.e. form complementary base pairs) to complementary sequences on the two DNA strands forming short lengths of double-stranded DNA. The DNA is cooled to 40°C to allow the hydrogen bonds to form. There are two reasons for making short lengths of double-stranded DNA:
   - The enzyme DNA polymerase requires some existing double stranded DNA to get it started.
   - Only the DNA between the primer sequences is replicated, so by choosing appropriate primers you can ensure that only a specific target sequence is copied.
4. The DNA polymerase enzyme can now build new stands alongside each old strand to make double-stranded DNA. Each new nucleotide binds to the old strand by complementary base pairing and is joined to the growing chain by a phosphodiester bond. The enzyme used in PCR is derived from the thermophilic bacterium *Thermus aquaticus*, which grows naturally in hot springs at a temperature of 90°C, so it is not denatured by the high temperatures in step 2. Its optimum temperature is about 72°C, so the mixture is heated to this temperature for a few minutes to allow replication to take place as quickly as possible.

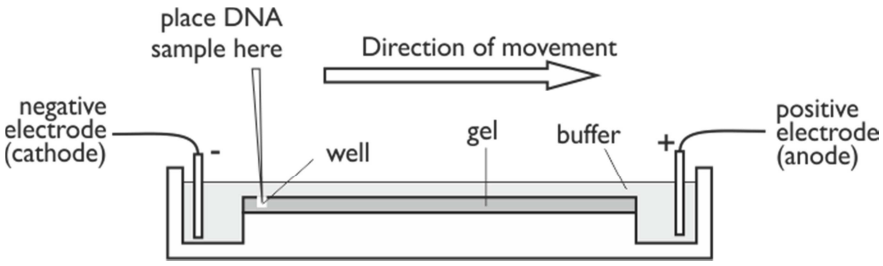
5. Each original DNA molecule has now been replicated to form two molecules. The cycle is repeated from step 2 and each time the number of DNA molecules doubles. This is why it is called a chain reaction, since the number of molecules increases exponentially, like an explosive chain reaction. After n cycles, there is an amplification factor of 2^n. Typically PCR is run for 20-30 cycles.



PCR can be completely automated, so in a few hours a tiny sample of DNA can be amplified millions of times with little effort. The product can be used for further studies, such as cloning, electrophoresis, or gene probes. Because PCR can use such small samples it can be used in forensic medicine (with DNA taken from samples of blood, hair or semen), and can even be used to copy DNA from mummified human bodies, extinct woolly mammoths, or from an insect that's been encased in amber since the Jurassic period. One problem of PCR is having a pure enough sample of DNA to start with. Any contaminant DNA will also be amplified, and this can cause problems, for example in court cases.

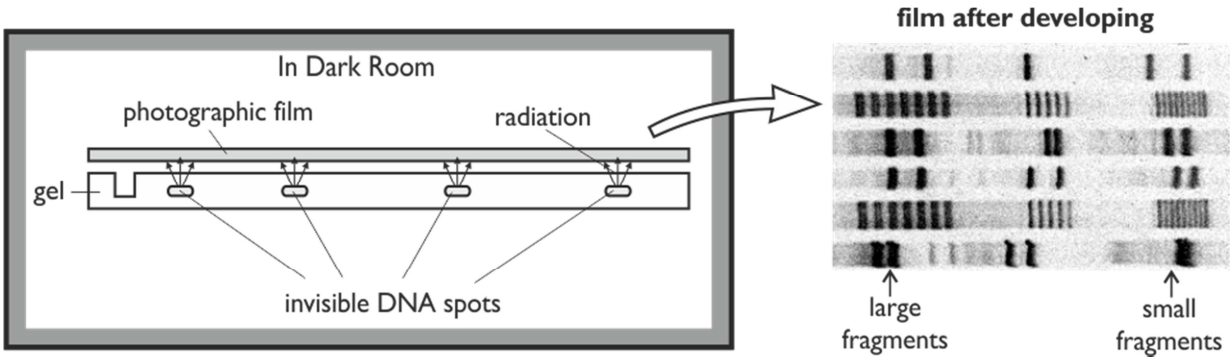
5. Electrophoresis

This is a form of chromatography used to separate different pieces of DNA on the basis of their length. It might typically be used to separate restriction fragments. The DNA samples are placed into wells at one end of a thin slab of gel made of agarose or polyacrylamide, and covered in a buffer solution. An electric current is passed through the gel. Each nucleotide in a molecule of DNA contains a negatively-charged phosphate group, so DNA is attracted to the anode (the positive electrode). The molecules have to diffuse through the gel, and longer lengths of DNA are retarded by the gel so move more slowly than shorter lengths. So the smaller the length of the DNA molecule, the further down the gel it will move in a given time. At the end of the run the current is turned off.



Unfortunately the DNA on the gel cannot be seen, so it must be visualised. There are two common methods for doing this:

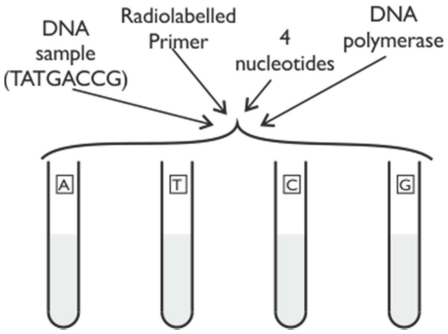
- The DNA can be stained with a coloured chemical such as azure A (which stains the DNA bands blue), or a fluorescent molecule such as ethidium bromide (which emits coloured light when the finished gel is illuminated with invisible ultraviolet light).
- The DNA samples at the beginning can be radiolabelled with a radioactive isotope such as ^{32}P, then visualised using autoradiography. Ordinary photographic film (sometimes called X-ray film) is placed on top of the finished gel in the dark for a few hours, and the radiation from any radioactive DNA on the gel exposes the film. When the film is developed the position of the DNA shows up as dark bands on the film. This method is extremely sensitive.



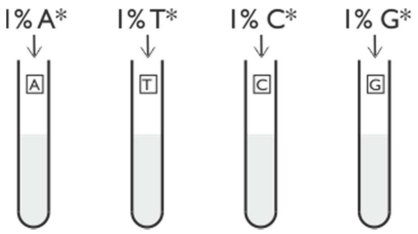
6. DNA Sequencing

This means reading the base sequence of a length of DNA. DNA sequencing is based on a beautifully elegant technique developed by Fred Sanger in Cambridge in 1975, and now called Sanger Sequencing.

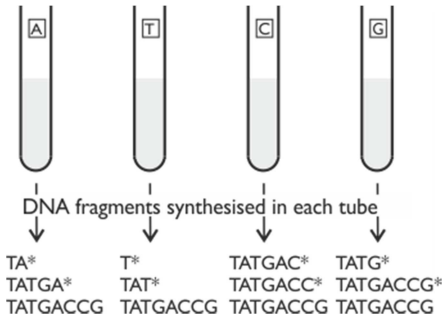
1. Label 4 test tubes A, T, C and G. Into each test tube add: a sample of the DNA to be sequenced (containing many millions of individual molecules); a radioactive primer (so the DNA can be visualised later on the gel); the four DNA nucleotides and the enzyme DNA polymerase.



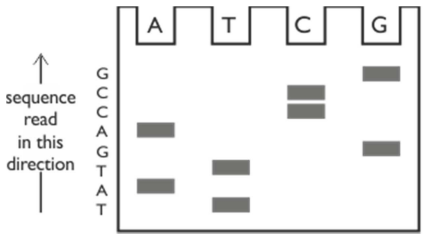
2. In each test tube add a small amount of a special modified dideoxy nucleotide that cannot form a phosphodiester bond and so stops further synthesis of DNA. Tube A has dideoxy A (A*), tube T has dideoxy T (T*), tube C has dideoxy C (C*) and tube G has dideoxy G (G*). The dideoxy nucleotides are present at about 1% of the concentration of the normal nucleotides.



3. Let the DNA polymerase synthesise many copies of the DNA sample. From time to time at random a dideoxy nucleotide will be added to the growing chain and synthesis of that chain will then stop. A range of DNA molecules will be synthesised ranging from full length to very short. The important point is that in tube A, all the fragments will stop at an A nucleotide. In tube T, all the fragments will stop at a T nucleotide, and so on.

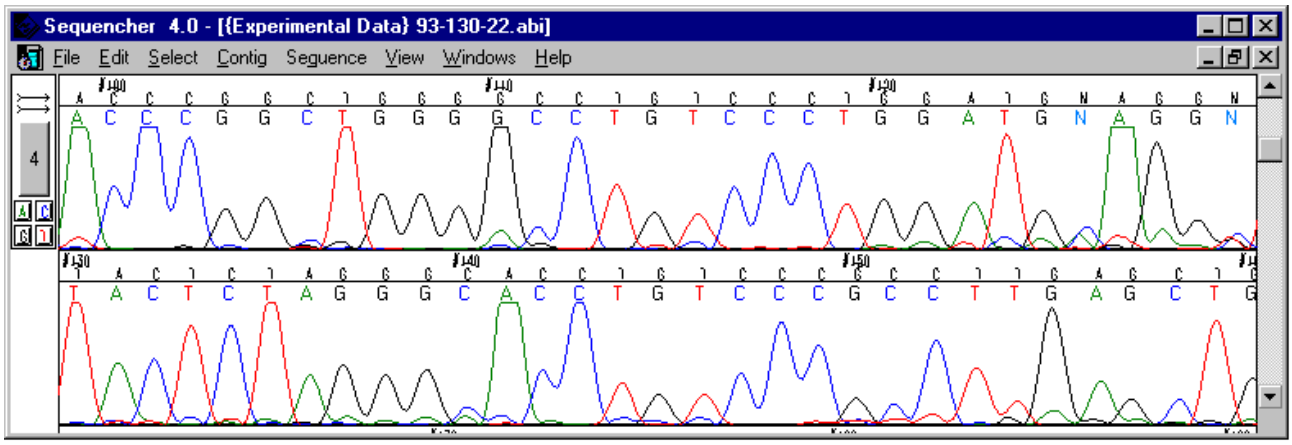


4. The contents of the four tubes are now run side by side on an electrophoresis gel, and the DNA bands are visualised by autoradiography. Since the fragments are now sorted by length the sequence can simply be read off the gel starting with the smallest fragment (just one nucleotide) at the bottom and reading upwards.



There is now a modified version of the Sanger method called cycle sequencing, which can be completely automated. The primers are not radiolabelled, but instead the four dideoxy nucleotides are fluorescently labelled, each with a different colour (A* is green, T* is red, C* is blue and G* is yellow). The polymerisation reaction is done in a single tube, using PCR-like cycles to speed up the process. The

resulting mixture is separated using capillary electrophoresis, which gives good separation in a single narrow tube gel. The gel is read by a laser beam and the sequence of colours is converted to a DNA sequence by computer program (like the screenshot below). This technique can sequence an amazing 12 000 bases per minute.



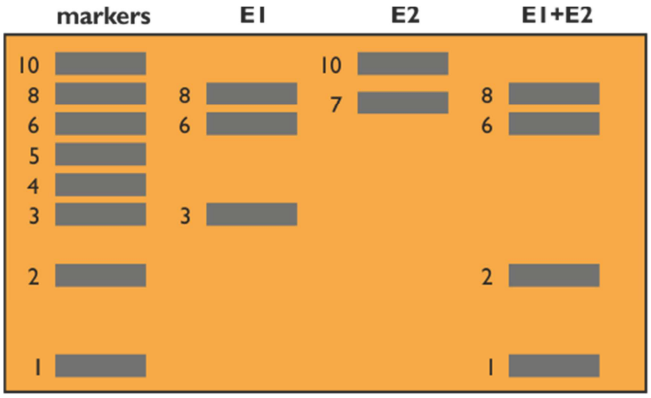
Many thousands of genes have been sequenced using these methods and the entire genomes of several organisms have also been sequenced. A huge project to sequence the complete 3-billion base sequence of the human genome was recently completed. This information is giving us unprecedented knowledge about ourselves, and is likely to lead to dramatic medical and scientific advances. Once a gene sequence is known the amino acid sequence of the protein that the DNA codes for can also be determined, using the genetic code table. The sequence can also be compared with DNA sequences from other individuals and even other species to work out relationships between individuals or species.

These genome sequences represent vast amount of data that must be analysed and compared to existing sequences. Powerful computers, huge databases and intelligent search programs are being developed to deal with this data, which has led to a whole new branch of biology called bioinformatics, and a new way of doing biology without touching a living thing: *in silico*.

7. Restriction Mapping

A restriction map is simply a diagram of a piece of DNA marked with the locations of sites where it is cut by restriction enzymes.

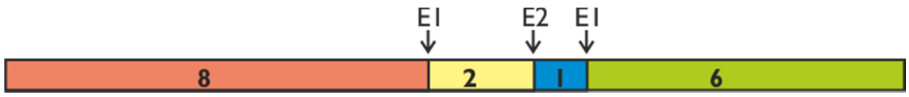
Usually a piece of DNA is cut with two different restriction enzymes, both on their own, and together. This gives three different mixtures of restriction fragments, which are run on an electrophoresis gel (labelled E1, E2 and E1+E2 on the gel below). The first lane on this gel contains a "DNA ladder" – a mixture of DNA fragments of known sizes – which is used to calibrate the gel. By comparison with the ladder bands, the length of each restriction fragment can be measured (marked in kilobases, kb on this diagram).



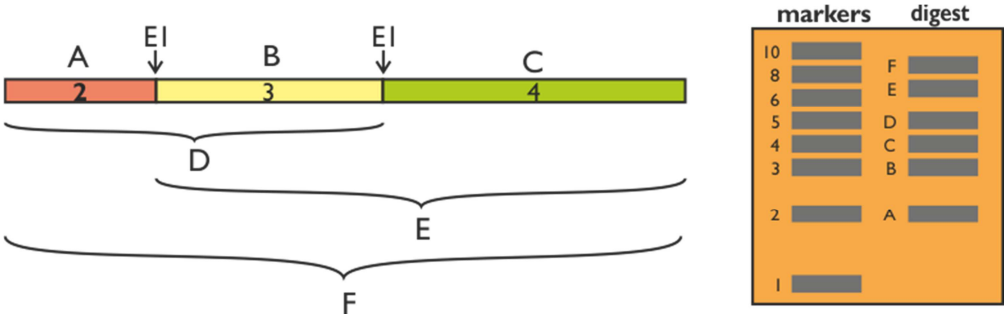
Conclusions:

- There are 2 recognition sites for E1
- There is 1 recognition sites for E2
- The total length is 17kb

From this information alone we have to deduce the restriction map. By a process of logic, we can construct the following restriction map to account for the banding pattern above.



Restriction maps can also be made using one restriction enzyme under conditions when digestion is incomplete. Some DNA molecules will be cut at all recognition sites, some at just one, and some won't be cut at all. This partial digest thus gives a range of fragments from which a restriction map can be deduced.



Restriction maps are useful in themselves, as they can be used to choose restriction enzymes to generate known-sized fragments. But they are even more useful as aids to DNA sequencing. A long piece of DNA can be cut into shorter restriction fragments, which are easy to sequence. A computer then uses the restriction map to "stitch together" the individual sequences in the correct order and orientation, to generate the complete sequence. The human genome was sequenced in this way.

8. Southern Blot

A Southern blot is used to detect a specific target sequence in samples of DNA. DNA samples separated in an electrophoresis gel can't be manipulated or stored, since the electrophoresis gel is fragile, and the separated DNA fragments continue to diffuse through it, blurring the bands. So, in a blot, the DNA is transferred to a stronger substrate, where it can be fixed. The target sequence is then detected using a DNA probe.

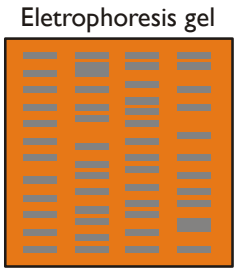
A DNA probe is simply a short length of single-stranded DNA (100-1000 nucleotides long) with a label attached. There are two common types of label used:

- A radioactively-labelled probe (synthesised using the isotope ^{32}P) can be visualised using autoradiography.
- A fluorescently-labelled probe will emit visible light when illuminated with invisible ultraviolet light. Probes can be made to fluoresce with different colours.

If a probe is added to a mixture of different pieces of single-stranded DNA (e.g. restriction fragments) it will anneal to any fragments of DNA containing the complementary sequence forming regions of double-stranded DNA. This double-stranded DNA is a mixture or hybrid (since it contains DNA from two different sources), so the DNA formed is called hybrid DNA, and the process is called hybridisation. The hybrid DNA fragments will now be labelled and will stand out from the rest of the DNA.

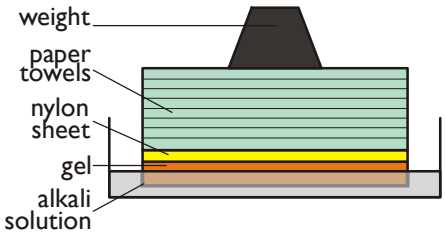
The Southern blot method is:

1. DNA is extracted from the source cells (e.g. different patients or different species) and amplified by PCR to make enough DNA for the hybridisation. The DNA samples are then digested by a restriction enzyme into many small fragments, and the fragments separated on an electrophoresis gel.

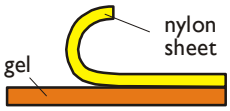


2. The gel is placed in an alkali solution, which breaks the hydrogen bonds between the DNA bases causing the stands to separate.

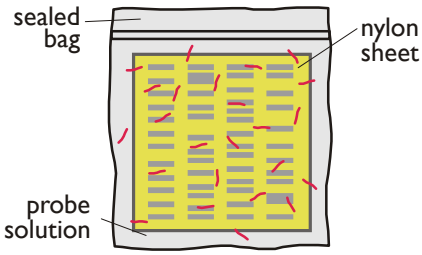
3. A thin sheet of nylon or nitrocellulose is placed on top of the gel and covered with a stack of paper towels, weighted down to make good contact. The alkali solution is drawn up through the gel to the paper towels by capillary action, bringing the DNA with it. The negatively-charged DNA sticks to the positively-charged nylon membrane.



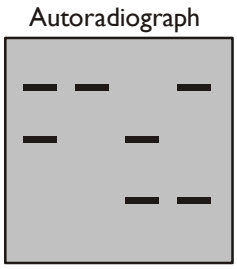
4. After a few hours the DNA fragments are all transferred to the nylon sheet (though of course they can't be seen). The nylon sheet is separated from the gel, and treated with UV light to fix the DNA molecules.



5. The nylon sheet is placed in a plastic bag containing a solution of labelled probes and mixed thoroughly. The probes will anneal by complementary base-pairing to DNA fragments in the nylon membrane that have a complementary sequence, forming hybrid DNA molecules stuck to the nylon sheet (but again they can't be seen).



6. The location of the hybrid DNA can be visualised by different methods, depending on the label used. If the probes were radioactive then they can be visualised by autoradiography, and the probes show up as bands on photographic film. If the probes were fluorescent then they can be visualised as bands of light when illuminated by ultraviolet light.



The Southern blot has many uses:

- To identify restriction fragments containing a particular gene out of the thousands of restriction fragments formed from an entire genome.
- To identify genes in one species that are similar to those of another species. Genes are remarkably similar in sequence from one species to another, so for example a gene probe for a mouse gene will probably anneal with the same gene from a human. This has aided the identification of human genes.
- To identify the short DNA sequences used in DNA fingerprinting.
- To screen for genetic diseases, such as Huntington's disease and cystic fibrosis (see p91).

The Southern blot was invented by Edwin Southern at Edinburgh University in 1975 (as a biological joke, similar blotting techniques using RNA or protein are called northern and western blots respectively).

9. Genetic Fingerprinting

Genetic fingerprinting is used to distinguish between DNA samples from different people. Although 99.9% of human DNA is the same in every person, there is enough variation in the remaining 0.1% (over 100 000 base pairs) to be used to distinguish one individual from another. The differences are found in non-coding DNA (see unit 2), since these regions can accumulate mutations without affecting function, so will show considerable variation between individuals of the same species.

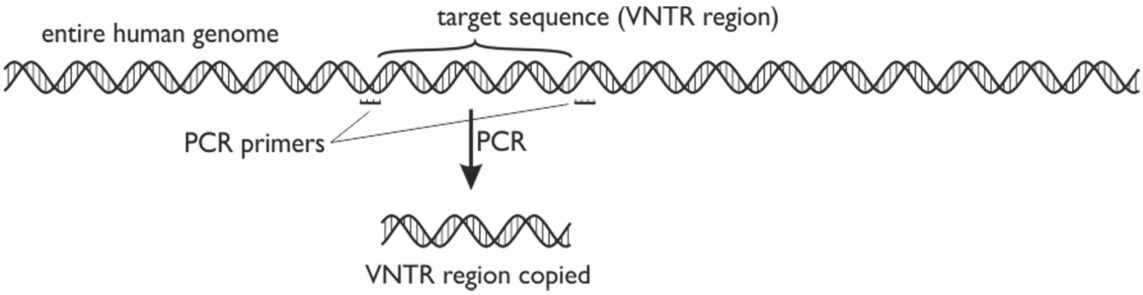
I principle a specific region of non-coding DNA from individuals' DNA samples could be sequenced and the sequences compared. The more differences in the bases, the more distantly related the individuals are. However, sequencing is expensive and slow (although new methods are making it quicker and cheaper all the time). Fortunately there is a quicker way to compare non-coding DNA using the cheap and established techniques of restriction enzymes and electrophoresis. It so happens that non-coding DNA is mainly made up of long repetitive sequences. There are different kinds of repetitive sequences:

- Satellite DNA – any DNA region with repetitive sequences. Satellite DNA gets its name because it forms a separate (or satellite) band when separated by density centrifugation.
- Tandem repeat – when a simple sequence of bases in a DNA molecule is repeated without a break e.g. ATATATAT (a 2-base repeat) or CCTAAGCCTAAGCCTAAG (a 6-base repeat).
- STR (Short Tandem Repeat) or Microsatellite – a tandem repeat with a sequence of less than10 bases.
- Minisatellite – a tandem repeat with a sequence of 10–60 bases.
- VNTR (Variable Number Tandem Repeat) – a region of DNA containing a particular tandem repeat. While everyone has the same repeating sequence in this region, different individuals have different numbers of repeats.
- Polymorphism – a range of different phenotypes of a characteristic or genotypes of a gene within a population. In molecular biology it means a range of different DNA sequences at a particular locus.
- RFLP (Restriction Fragment Length Polymorphism, pronounced "riff-lip") – when restriction fragments of the same region of DNA have different lengths in different people.

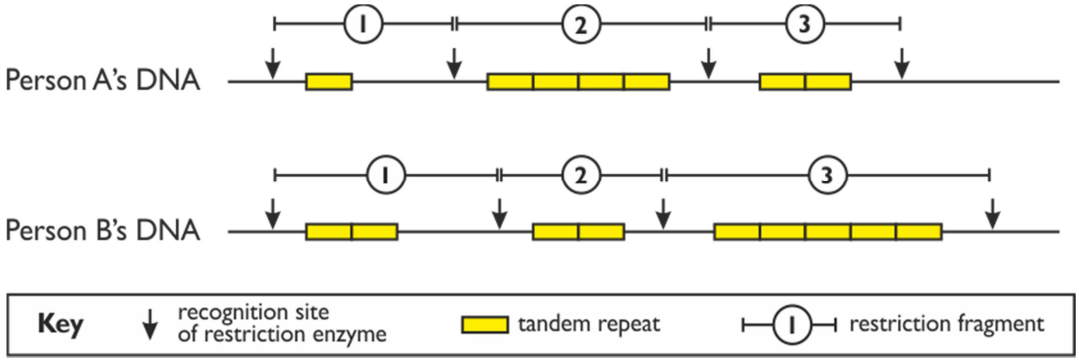
Genetic Fingerprinting was invented by Sir Alec Jeffreys at Leicester University in 1984. Today there are many different methods of genetic fingerprinting, but Jeffreys' original method was based on Restriction Fragment Length Polymorphism (RFLPs):

Method

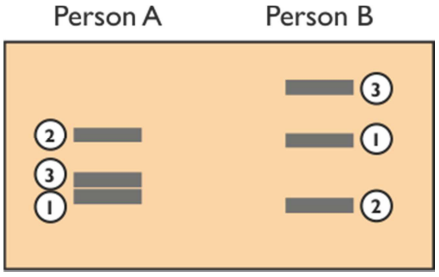
1. Samples of DNA to be compared (from different individuals) are amplified by PCR. The PCR primers are designed to target and copy a region of the human DNA containing VNTRs.



2. The amplified DNA of the VNTR region is digested by a restriction enzyme. The same recognition sequences should be present in the different samples, so the same number of fragments should be produced, but the fragments will be different lengths depending on the number of tandem repeats in each restriction fragments.



3. The digested DNA is run on an electrophoresis gel to separate out the fragments by size, and the fragments are visualised using a Southern blot. The pattern of bands is different for the two samples because of the restriction fragment length polymorphism i.e. the "same" restriction fragments have different lengths due to different numbers of tandem repeats.



Samples of the same DNA (or from identical twins) give the same banding pattern, but samples from different people give different banding patterns.

This original fingerprinting method takes a few days to carry out and can produce results that are difficult to interpret. Newer methods of genetic fingerprinting have been developed that offer improved speed, sensitivity, discrimination and reliability, but the same principles are used. In the current technique, called DNA profiling and used by the police in most countries of the world, 10-13 different regions containing variable numbers of short tandem repeats (STRs) are targeted and amplified by PCR and then sorted and sized using an automated capillary sequencer. This avoids the need for restriction enzymes or Southern blot and means the process can be automated. The size data can also be stored on computers, forming National DNA Databases. It is always possible that two unrelated people might by chance have the same genetic fingerprint, but using DNA profiling it is calculated that the chance of this is less than 1 in 10^9.

DNA fingerprinting is used

- in forensic science, to match DNA samples collected from a crime (e.g. from sperm, blood, hair, skin) with that of suspects. Normally suspects are arrested because of some other link to the crime, but with the development of a national DNA database, that doesn't have to be the case. Convictions based on DNA evidence can be difficult due to problems of contamination and poor laboratory technique.
- to determine family relationships, usually between a child and a suspect father (paternity testing). Since children inherit half their DNA from each parent, they inherit half their DNA fingerprint fragments from each parent, so 50% of the bands should match their mother's bands and 50% match their father's bands.
- to prevent undesirable inbreeding during breeding programs in farms and zoos. It can also identify plants or animals that carry a particular desirable (or undesirable) allele.
- to determine relationships between ancient humans (using DNA extracted from archaeological remains) and modern humans.
- to establish evolutionary relationships between different species.
- to measure genetic diversity within a population.

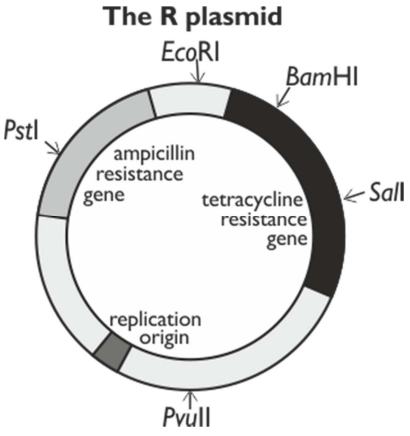
10. Vectors

Now we turn to genetically-modifying living cells. The next three techniques are concerned with transferring genes (DNA) from the test tube into cells, so that the genes can be replicated and expressed in those cells. To do this we first need a vector.

In genetic engineering a vector is a length of DNA that carries the gene we want into a host cell. A vector is needed because a length of DNA containing a gene on its own won't actually do anything inside a host cell. Since it is not part of the cell's normal genome it won't be replicated when the cell divides, it won't be expressed, and in fact it will probably be broken down pretty quickly. A vector gets round these problems by having these properties:

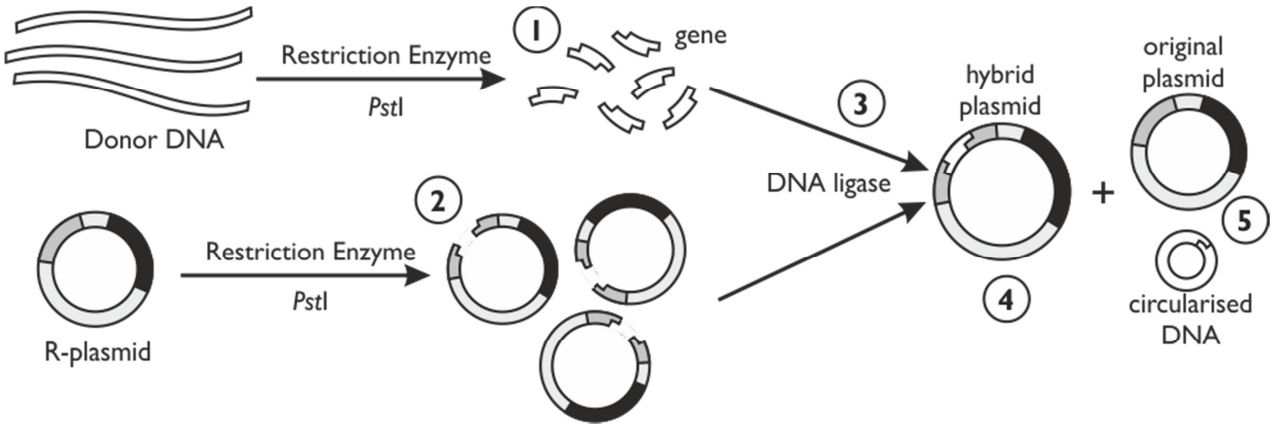
- It is big enough to hold the gene we want (plus a few others), but not too big.
- It is circular (or more accurately a closed loop), so that it is less likely to be broken down (particularly in prokaryotic cells where DNA is always circular).
- It contains control sequences, such as a replication origin and a transcription promoter, so that the gene will be replicated, expressed, or incorporated into the cell's normal genome.
- It contains marker genes, so that cells containing the vector can be identified.

Plasmids are the most common kind of vector, so we shall look at how they are used in some detail. Plasmids are short circular bits of DNA found naturally in bacterial cells. A typical plasmid contains 3-5 genes and there are usually around 10 copies of a plasmid in a bacterial cell. Plasmids are copied separately from the main bacterial DNA when the cell divides, so the plasmid genes are passed on to all daughter cells. They are also used naturally for exchange of genes between bacterial cells, so bacterial cells will readily take up a plasmid. Because they are so small, plasmids are easy to handle in a test tube, and foreign genes can quite easily be incorporated into them using restriction enzymes and DNA ligase.



One of the most common plasmids used is the R-plasmid. This plasmid contains a replication origin, several recognition sequences for different restriction enzymes (with names like *PstI* and *EcoRI*), and two marker genes, which in this case confer resistance to antibiotics. The R plasmid gets its name from these resistance genes.

The diagram below shows how a gene can be incorporated into a plasmid using restriction and ligase enzymes.



1. A restriction enzyme (*Pst1* here) is used to cut the gene from the donor DNA, with sticky ends.
2. The same restriction enzyme cuts the plasmid in the middle of one of the marker genes (we'll see why this is useful later).
3. The gene and plasmid are mixed in a test tube and they anneal because they were cut with the same restriction enzyme and have the same sticky ends.
4. The fragments are joined covalently by DNA ligase to form a hybrid vector (in other words a mixture or hybrid of bacterial and foreign DNA).
5. Several other products are also formed: some plasmids will simply re-anneal with themselves to re-form the original plasmid, and some DNA fragments will join together to form chains or circles. These different products cannot easily be separated, but it doesn't matter, as the marker genes can be used later to identify the correct hybrid vector.

This technique takes place entirely in test tubes; there are no cells involved. So the next step is to insert our modified DNA (the hybrid vector) into a living cell.

10. Transformation

Transformation means inserting new DNA (usually a plasmid) into a living cell (called a host cell), which is thus genetically modified, or transformed. A transformed cell can replicate and express the genes in the new DNA. DNA is a large molecule that does not readily cross cell membranes, so the membranes must be made permeable in some way. There are different ways of doing this depending on the type of host cell.

- **Heat Shock**. Bacterial and animal cells in culture can be made to take up DNA from their surroundings by raising the temperature suddenly raised by about 40°C.
- **Electroporation**. The most efficient method of delivering genes to bacterial cells is to use a high-voltage pulse, which temporarily disrupts the membrane and allows the plasmid to enter the cell.
- **Micro-Injection**. To transform individual cells, such as fertilised animal egg cells, the DNA is injected directly into the nucleus using an incredibly fine micro-pipette.
- **Gene Gun**. Tiny gold particles coated with DNA can be fired at plant cells using a compressed air gun. The particles can penetrate the tough cell wall and deliver the DNA to the nucleus.
- **Plant Tumours**. Plant cells are infected with a transformed bacterium, which inserts its plasmid into the plant cells' chromosomal DNA. Whole new plants are grown from these cells by micropropagation.
- **Liposomes**. Human cells *in vivo* can be transformed by DNA encased in liposomes, which fuse with the cell membrane, delivering the DNA into the cell.
- **Viruses**. Human cells *in vivo* can be infected by genetically-engineered viruses, which deliver the DNA into host cells. The viruses must first be made it safe, so they can't cause disease.

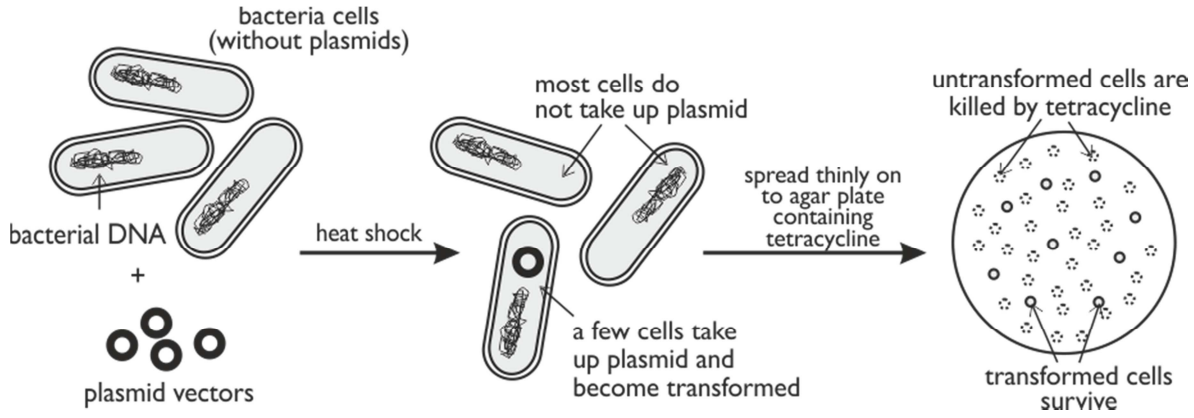
Most of these transformation techniques have a very low success rate (<<1%), so we need to be able to identify those few cells that have taken up the foreign DNA and been transformed. This is where the plasmid's marker genes are used.

11. Marker Genes

Marker genes (or reporter genes) are used to find which cells have actually taken up the hybrid vector. Following transformation, there are (at least) these four possible outcomes:

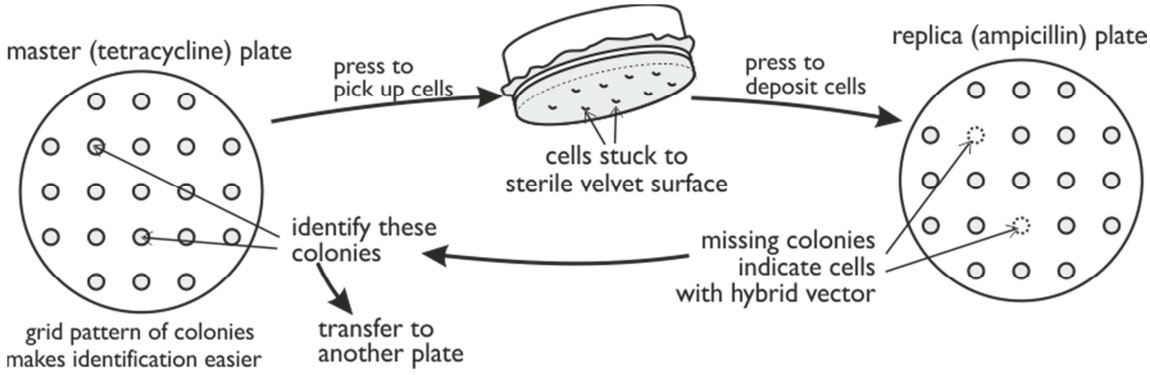
Vectors contain two different marker genes, which are needed to identify the required cells.

The first marker gene distinguishes between cells that have taken up a plasmid from those that haven't. The plasmid contains a gene for resistance to an antibiotic such as tetracycline, so bacterial cells taking up this plasmid can make this gene product and so are resistant to this antibiotic. The cells are therefore grown on a medium containing tetracycline, which kills all the normal untransformed cells (>99.99%). Only the few transformed cells will survive, and these can then be grown and cloned on another plate.



The second marker gene distinguishes between cells that have taken up the hybrid plasmid from those that have taken up the original plasmid. The trick here is that the foreign DNA is inserted <u>inside</u> the second marker gene, so cells with the hybrid plasmid cannot make that gene product. Different genes are used for this second marker:

- The marker gene can be a gene for resistance to another antibiotic, such as ampicillin. In this case, cells with the hybrid vector are not resistant to ampicillin. Since this means killing the cells we want, the ampicillin test is done on a replica plate, where the colonies are transferred to a second agar plate using a cloth stamp. Colonies that grow on the first (tetracycline) plate but not on the replica (ampicillin) plate are the ones we want.



- The marker gene can be a gene for the enzyme β-galactosidase (lactase). This enzyme turns a white substrate in the agar plate into a blue product. So colonies of cells with the original plasmid turn blue, while those with the hybrid plasmid remain white, and can easily be identified.
- The marker gene can be a gene for green fluorescent protein (GFP). Colonies of cells with the original plasmid fluoresce green in UV light, while those with the hybrid plasmid do not fluoresce, and can easily be identified.

Gene Cloning

Gene cloning simply means make multiple copies of a piece of DNA. It is a necessary step in just about any aspect of molecular biotechnology, such as genetic engineering, genome sequencing or genetic fingerprinting. There are two different ways to clone DNA:

- *In vitro* gene cloning uses PCR to clone DNA in the test tube.
- *In vivo* gene cloning uses restriction enzymes, vectors, DNA ligase, transformation of bacterial cells, marker genes and growth cultures to clone DNA inside bacterial cells.

The two techniques have different advantages and disadvantages:

***in vitro* cloning (using PCR)**	***in vivo* cloning (using living cells)**
Newer technique (1983)	Older technique (1977)
Simple, automated technique, which can be completed in a few hours	Complex, multi-step process, needing several days to complete
Very sensitive, can clone a single molecule	Large amounts of original DNA needed
Can use DNA from different kinds of source, including degraded DNA from crime scenes or archaeological sources	Needs intact, pure DNA
Clones DNA molecules up to 1kbp long	Clones DNA molecules up to 2Mbp long
High error rate, since no error-correction	Low error rate due to cellular error-correcting mechanisms
DNA is made in the test tube, so cannot be expressed directly	DNA is made in cells, so can be expressed easily

Genetically Modified Organisms

We have looked at some of the many techniques used in biotechnology. We'll now turn to some applications of these techniques. The applications involve altering the genes in a living organism to produce a Genetically Modified Organism (GMO) with a new genotype. The GMO is designed to benefit human in some way. If a foreign gene in copied from one species into another the GMO is called a transgenic organism, but remember that not all GMOs are transgenic: the genetic modification might just alter an existing gene so that its product is changed or change its gene expression. We'll consider the applications in three groups.

- Gene Products — Using genetically modified organisms (usually microbes) to produce chemicals (usually proteins) for medical or industrial applications.
- New Phenotypes — Using gene technology to alter the characteristics of organisms (usually farm animals or crops).
- Gene Therapy — Using gene technology on humans to treat a disease.

Gene Products

The biggest and most successful kind of genetic engineering is the production of gene products. These products are of medical, agricultural or commercial value to humans. This table shows a few of the examples of genetically engineered products that are already available.

Product	Use	Host Organism
Insulin	human hormone used to treat diabetes	bacteria, yeast
HGH	human growth hormone, used to treat dwarfism	bacteria
Encephalin	human hormone	plants
BST	bovine growth hormone, used to increase milk yield of cows	bacteria
Factor VIII	human blood clotting factor, used to treat haemophiliacs	pigs
Anti-thrombin	anti-blood clotting agent used in surgery	goats
Collagen	used in reconstructive surgery	plants
Vaccines	hepatitis B antigen, for vaccination	yeast, plants
Antibodies	research and clinical use	goats, plants
AAT	enzyme inhibitor used to treat cystic fibrosis and emphysema	sheep, yeast
α-glucosidase	enzyme used to treat Pompe's disease	rabbits
DNAse	enzyme used to treat CF	bacteria
rennin	enzyme used in manufacture of cheese	bacteria /yeast
cellulase	enzyme used in paper production	bacteria
PHB	biodegradable plastic	plants

The products are mostly proteins, which are produced directly when a gene is expressed, but they can also be non-protein products produced by genetically-engineered enzymes. The basic idea is to transfer a gene (often human) to another host organism (usually a microbe) so that it will make the gene product quickly, cheaply and ethically. It is also possible to make "designer proteins" by altering gene sequences, but while this is a useful research tool, there are no commercial applications yet.

Since the end-product is just a chemical, in principle any kind of organism could be used to produce it. By far the most common group of host organisms used to make gene products are the bacteria, since they can be grown quickly and the product can be purified from their cells. Unfortunately bacteria cannot always make human proteins, and so fungi, animals and plants have also been used to make gene products. This table shows some of the advantages and disadvantages of using different organisms for the production of genetically-engineered gene products.

Type of organism	Advantages	Disadvantages
Prokaryotes (Bacteria)	No nucleus so DNA easy to modify; have plasmids; small genome; genetics well understood; asexual so can be cloned; small and fast growing; easy to grow commercially in fermenters; will use cheap carbohydrate; few ethical problems.	Can't splice introns; no post-translational modification; small gene size.
Eukaryotes	Can splice introns; can do post-translational modifications; can accept large genes.	Do not have plasmids (except yeast); often diploid so two copies of genes may need to be inserted; control of expression not well understood.
Fungi	Asexual so can be cloned; haploid, so only one copy needed; can be grown in vats.	Can't always make animals' gene products.
Plants	Photosynthetic so don't need much feeding; can be cloned from single cells; products can be stored in seeds or secreted from roots or in sap.	Cell walls difficult to penetrate by vector; slow growing; multicellular.
Animals (pharming)	Most likely to be able to make human proteins; products can be secreted in milk or urine.	Multicellular; slow growing; expensive to produce, ethical issues.

New Phenotypes

This means altering the characteristics of organisms by genetic engineering. The organisms are generally commercially-important crops or farm animals, and the object is to improve their quality in some way. This can be seen as a high-tech version of selective breeding, which has been used by humans to alter and improve their crops and animals for at least 10 000 years. This table gives an idea of what is being done.

Organism	Modification
long life tomatoes	There are two well-known projects, both affecting the gene for the enzyme polygalactourinase (PG), a pectinase that softens fruits as they ripen. Tomatoes that make less PG ripen more slowly and retain more flavour. The American "Flavr Savr" tomato used antisense technology to silence the gene, while the British Zeneca tomato disrupted the gene. Both were successful and were on sale for a few years, but neither is produced any more.
Insect-resistant crops	Genes for various powerful protein toxins have been transferred from the bacterium Bacillus thuringiensis to crop plants including maize, rice and potatoes. These Bt toxins are thousands of times more powerful than chemical insecticides, and since they are built-in to the crops, insecticide spraying (which is non-specific and damages the environment) is unnecessary.
virus-resistant crops	Gene for virus coat protein has been cloned and inserted into tobacco, potato and tomato plants. The coat protein seems to "immunise" the plants, which are much more resistant to viral attack.
herbicide resistant crops	The gene for resistance to the herbicide BASTA has been transferred from Streptomyces bacteria to tomato, potato, corn, and wheat plants, making them resistant to BASTA. Fields can safely be sprayed with this herbicide, which will kill all weeds, but not the crops. However, this means continued use of agrochemicals, so is controversial.
pest-resistant legumes	The gene for an enzyme that synthesises a chemical toxic to weevils has been transferred from Bacillus bacteria to The Rhizobium bacteria that live in the root nodules of legume plants. These root nodules are now resistant to attack by weevils.
Nitrogen-fixing crops	This is a huge project, which aims to transfer the 15-or-so genes required for nitrogen fixation from the nitrogen-fixing bacteria Rhizobium into cereals and other crop plants. These crops would then be able to fix their own atmospheric nitrogen and would not need any fertiliser. However, the process is extremely complex, and the project is nowhere near success.
crop improvement	Proteins in some crop plants, including wheat, are often deficient in essential amino acids (which is why vegetarians have to watch their diet so carefully), so the protein genes are being altered to improve their composition for human consumption.
mastitis-resistant cattle	The gene for the enzyme lactoferrin, which helps to resists the infection that causes the udder disease mastitis, has been introduced to Herman – the first transgenic bull. Herman's offspring inherit this gene, do not get mastitis and so produce more milk.
tick-resistant sheep	The gene for the enzyme chitinase, which kills ticks by digesting their exoskeletons, has bee transferred from plants to sheep. These sheep should be immune to tick parasites, and may not need sheep dip.
Fast-growing fish	A number of fish species, including salmon, trout and carp, have been given a gene from another fish (the ocean pout) which activates the fish's own growth hormone gene so that they grow larger and more quickly. Salmon grow to 30 times their normal mass at 10 times the normal rate.
environment cleaning microbes	Genes for enzymes that digest many different hydrocarbons found in crude oil have been transferred to Pseudomonas bacteria so that they can clean up oil spills.

Evaluating Biotechnology

The whole point of creating genetically-modified organisms is to benefit humans, and the benefits are usually fairly obvious, but nevertheless there has been some vocal opposition to GMOs. Opposition is often based on ethical, moral or social grounds, such as harm to animals or the environment, though there can also be more practical issues, such as distrust of large corporations.

Benefits

- Medicines and drugs can be produced safely in large quantities from microbes rather than from slaughtered animals. These medicines benefit humans and can spare animal suffering as well.
- Agricultural productivity can be improved while using less pesticides or fertilisers, so helping the environment. GM crops can grow on previously unsuitable soil or in previously unsuitable climates.
- GM crops can improve the nutrition and health of millions of people by improving the nutritional quality of their staple crops.

Risks

- **Risks to the modified organism.** Genetic modification of an organism may have unforeseen genetic effects on that organism and its offspring. These genetic effects could include metabolic diseases or cancer, and would be particularly important in vertebrate animals, which have a nervous system and so are capable of suffering. The research process may also harm animals.
- **Transfer to other organisms.** Genes transferred into GMOs could be transferred again into other organisms, by natural accidents. These natural accidents could include horizontal gene transmission in bacteria, cross-species pollination in plants, and viral transfer. This could result in a weed being resistant to a herbicide, or a pathogenic bacterium being resistant to an antibiotic. To avoid transfer via cross-pollination, genes can now be inserted into chloroplast DNA, which is not found in pollen.
- **Risks to the ecosystems.** A GMO may have an unforeseen effect on its food web, affecting other organisms. Many ecosystems are often delicately balanced, and a GMO could change that balance.
- **Risk to biodiversity.** GMOs may continue to reduce the genetic biodiversity already occurring due to selective breeding.
- **Risks to human societies.** There could be unexpected and complicated social and economic consequences from using GMOs. For example if GM bananas could be grown in temperate countries, that would be disastrous for the economies of those Caribbean countries who rely on banana exports.
- **Risks to local farmers.** Developing GMOs is expensive, and the ownership of the technology remains with the large multi-national corporations. This means the benefits may not be available to farmers in third world countries who need it most.

Gene Therapy

The idea of gene therapy is to genetically alter humans in order to treat a disease. This could represent the first opportunity to cure hitherto incurable diseases. Note that this is quite different from using genetically-engineered microbes to produce a drug, vaccine or hormone to treat a disease by conventional means. Gene therapy means altering the genotype of a tissue or even a whole human. The most promising targets of gene therapy are single gene disorders ("genetic diseases") like Huntington's disease, cystic fibrosis and muscular dystrophy, since these could in principle be treated by replacing just a single gene.

Methods of Gene Therapy

The techniques of gene therapy are inherently more difficult than those used to make other GMOs because living humans are involved and, above all, the technique must not cause further harm. The correct gene (the therapeutic DNA) needs to be introduced into the cells of a living person. Once there, the therapeutic DNA needs to be expressed by the cells, making a protein and affecting the cells' function. Some of the most common methods are:

- **Liposomes**. The therapeutic DNA is encased in a lipid vesicle called a liposome. The liposome membrane then fuses with the cell membrane, delivering the therapeutic DNA into the cell.
- **Viruses**. Normal viral infection depends on the virus delivering its own DNA into host cells, where it can be expressed to make new virus particles. So genetically-modified viruses can be used to deliver human genes by the same method. The virus must first be genetically engineered to make it safe, so that it can't reproduce itself or make toxins.
- **Stem cells**. In some cases stem cells can be removed from the patient (e.g. from bone marrow), genetically modified *in vitro* with the therapeutic DNA, then the stem cells injected back into the patient. This method is safer and avoids immune rejection, but only works for some tissues.

Note that gene therapy doesn't alter or replace the existing mutated gene, which will still continue to make the non-functional protein. But in addition, the new gene will make working protein, which will allow the modified cells to function normally.

Examples of Gene Therapy

- **Cystic Fibrosis**. Cystic fibrosis (CF) is the most common genetic disease in the UK, affecting about 1 in 2500. CF is caused by a mutation in the gene for the protein CFTR, which is a chloride ion channel, and it results too much sticky mucus being produced by epithelial cells. This sticky mucus causes breathlessness, lung infections, malnutrition and other symptoms. CF is always fatal, though life expectancy has increased from 1 year to about 20 years due to modern treatments. If about 10% of the epithelial cells could be genetically modified with the correct ion-channel gene, this would allow enough chloride ion transport to relieve the symptoms of the disease. In clinical trials liposomes and viruses are

delivered using an aerosol inhaler, like those used by asthmatics, but as yet no therapy has been shown to be successful.

- **SCID**. Severe Combined immunodeficiency Disease (SCID) is a rare genetic disease that affects the immune system. It is caused by a mutation in the gene for the enzyme adenosine deaminase (ADA). Without this enzyme white blood cells cannot be made, so sufferers have almost no effective immune system and would quickly contract a fatal infection unless they spend their lives in sterile isolation. For gene therapy bone marrow cells from the patient are transfected with the ADA gene *in vitro*, then injected into the patient. This procedure has successfully cured a number of patients so far, though the treatment has to be repeated every few years.
- **Retinal disease**. Gene therapy has been used to treat an inherited eye disease caused by mutation of the gene RPE65. This mutation stops rhodopsin working, resulting in almost complete blindness. A patient was treated with a virus containing the correct RPE65 gene and their vision improved with no apparent side effects.
- **Cancer**. White blood cells have been genetically modified to produce tumour necrosis factor (TNF), a protein that kills cancer cells, making these cells more effective against tumours. Genes could also be targeted directly at cancer cells, causing them to die, or to revert to normal cell division.

Problems of Gene Therapy

- Most gene therapy attempted so far has had only a short-lived effect. Problems with integrating the therapeutic DNA into the host cell, and of replicating new DNA when the host cell divides, have meant that patients have to repeat the gene therapy treatment at intervals.
- Therapeutic DNA and modified host cells are recognised as non-self by the immune system and so destroyed in a primary immune response. Subsequent repeated treatments stimulate a greater secondary immune response, which can be harmful to the patient.
- There is a chance that the therapeutic DNA is integrated in the host genome in the middle of another gene, for example in a tumour suppressor gene. The gene therapy could therefore induce cancer. This occurred in a clinical trial of SCID, when 3 out of 20 patients developed leukaemia.
- Viruses are the most successful vectors in gene therapy but they also present a variety of potential problems to the patient including toxicity, immune and inflammatory responses and recovery of pathogenicity. In 1999 a patient in a gene therapy trial died from a massive immune response to the virus being used for gene therapy.
- Many common genetic disorders, such as heart disease, high blood pressure, Alzheimer's disease, arthritis and diabetes, are caused by the combined effects of many genes. These multigene disorders are probably impossible to treat effectively using gene therapy.

Types of Gene Therapy

It is important to appreciate the different between somatic cell therapy and germ-line therapy.

- **Somatic cell therapy** means genetically altering specific body (or somatic) cells, such as trachea epithelial cells, bone marrow cells, pancreas cells, or whatever, in order to treat the disease. This therapy may treat the disease in the patient, but any genetic changes will **not** be passed on the offspring of the patient.
- **Germ-line therapy** means genetically altering those cells (sperm cells, sperm precursor cells, ova, ova precursor cells, zygotes or early embryos) that will pass their genes down the "germ-line" to future generations. Alterations to any of these cells will affect every cell in the resulting human, and in all his or her descendants.

Germ-line therapy would be highly effective, but is also potentially dangerous (since the long-term effects of genetic alterations are not known), unethical (since it could easily lead to eugenics) and immoral (since it could involve altering and destroying human embryos). It is currently illegal in the UK and most other countries, and current research is focussing on somatic cell therapy only. All gene therapy trials in the UK must be approved by the Gene Therapy Advisory Committee (GTAC), a government body that reviews the medical and ethical grounds for a trial. Germ-line modification is allowed with animals, and indeed is the basis for producing GMOs.

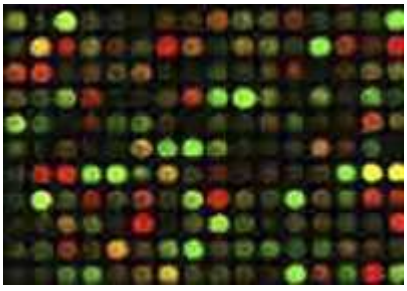
Genetic Screening and Counselling

There are around 25 000 genes in the human genome, and scientists have now identified and sequenced thousands of human genes. In many cases they also know exactly how the base sequence is different in the different versions (or alleles) of the genes that cause genetic diseases, susceptibility to disease or other clinically-important symptoms. This knowledge can be used to test, or screen, someone for the presence of these clinically-important alleles. The information gained from a screen can be used to select appropriate treatment at an early stage, even before symptoms appear. Even if a genetic disease is incurable, the screen may help in family planning and care of the patient.

Genetic screening is based on DNA hybridisation, and is similar to the Southern blot technique. Short lengths of DNA, 100-1000 nucleotides long, are synthesised with base sequences found only in the mutant alleles. These short lengths of DNA are called gene probes or hybridisation probes. Tiny quantities of these single-stranded gene probes are covalently fixed to a glass or silicon slide in a grid pattern. This slide is called a DNA microarray (or gene chip) and a single microarray a couple of centimetres long can hold 10 000 different gene probes.

The genetic screen procedure is:

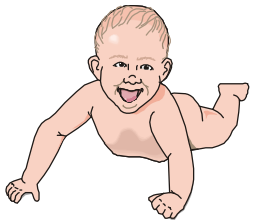
1. A few cells are taken from a patient by biopsy.
2. DNA is extracted from the biopsy and if necessary amplified by PCR.
3. The DNA is labelled with a fluorescent chemical.
4. The DNA is denatured (i.e. separated into single strands) by heating or strong alkali.
5. The single-stranded DNA is added to the microarray and mixed for a few hours. DNA that has a complementary sequence to any of the probes fixed to the microarray will hybridise to the probes by complementary base pairing.
6. The microarray is washed with a buffer, washing away any loose fluorescent DNA that is not hybridised to the probes on the chip.
7. If the microarray is illuminated with ultra-violet light, any spots where the subject's DNA has hybridised to a probe will fluoresce and can be seen, while the spots where the probes haven't hybridised will remain dark. In practice the microarray is "read" with a laser, which illuminates each spot in turn, recording the amount of fluorescence in each case on a computer.



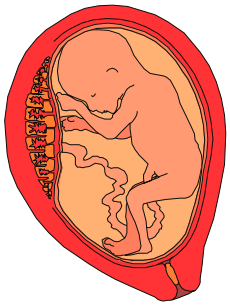
8. The locations of the fluorescent, "positive", spots are matched with the name of that probe's allele, giving a detailed genetic profile of the patient.

Genetic screens can be carried out at different times for different reasons:

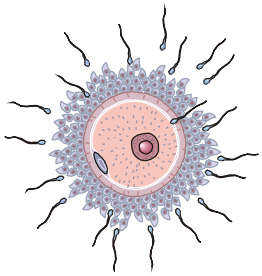
Adult screening is carried out to test adults for alleles of late-onset diseases (such as Huntington's or breast cancer) before any symptoms appear. Adult screening is also used test for carriers of a genetic disease (such as cystic fibrosis). The sample is usually taken by a buccal smear – a scraping of cells from the inside of the cheek.



Newborn screening is carried out immediately after birth, when a blood sample is taken with a lancet from the baby's heel. DNA is extracted from the baby's cells and tested for alleles for conditions such as phenylketonuria and congenital hypothyroidism. These conditions benefit from being treated early in life.



Prenatal screening is carried out on fetal cells before birth. It is offered to parents when there is a risk that the fetus might have a serious genetic disability. Amniocentesis is used to collect fetal skin cells in the amniotic fluid between 14 to 20 weeks of gestation, and the DNA from these cells is tested for alleles for conditions such as Tay Sachs disease, sickle cell anaemia, thalassemia, cystic fibrosis, and fragile x syndrome. If the tests are positive the parents have the option of terminating the pregnancy or preparing themselves to care for a disabled child.



Pre-implantation screening (or preimplantation genetic diagnosis, PGD) is performed on human embryos created by *in vitro* fertilization. One or two cells are carefully removed from an 8-cell embryo and the DNA is tested. This procedure doesn't harm the rest of the embryo, which continues to develop normally. Only embryos without genetic problems are used for implantation, so this avoids the need for terminations associated with prenatal screening.

Genetic Counselling

The results of genetic screening can be complex and distressing, both for subjects and their families. So genetic screening is often accompanied by genetic counselling. A genetic counsellor can explain the meaning of the test results, discuss the implications for the patient and their family, and advise on the next course of action. For example:

- Sickle-cell anaemia is single-gene disorder, caused by a recessive mutation the haemoglobin gene (see module 4). About 25% of the population are carriers of the sickle-cell allele (i.e. heterozygous). A couple who are both carriers have a 25% risk that their children may suffer from sickle-cell anaemia. So

before starting a family a couple with a family history of sickle-cell anaemia may decide to be tested for the sickle-cell allele to see if they are both carriers. If they are the genetic counsellor could discuss the options of pre-natal or pre-implantation screening.

- Huntington's disease is a single-gene disorder caused by a dominant mutation in the huntingtin gene. Huntington's disease causes muscle spasms and death, but there are no symptoms until middle age, and there is no cure. Someone with a history of Huntington's disease in the family may choose to be tested for the mutated allele. Since this is a dominant allele, a positive test result means that they will develop the disease in their forties, and there is a 50% chance that they will pass the disease on to any children they may have. The test can also indicate the likely severity of the disease. Genetic counselling and support is very important in this case since the impact on the individual and their family are so severe. 95% of individuals at risk choose not to undergo genetic testing, since there is no cure. The uncertainly of not knowing whether they will develop Huntington's disease must be weighed against the stress of knowing that they eventually will, or even the "survivor guilt" of knowing that they don't have the disease.
- Cancers can be caused by mutations in many different genes, such as oncogenes and tumour-suppressor genes (p57). Gene probes have been made for hundreds of these cancer-related alleles, so they can be tested in genetic screens. Screening and counselling can help to choose the best treatment for a particular patient. For example the drug herceptin is only effective at treating a certain type of breast cancer, caused by mutation in the HER2 gene.

There can be other issues with genetic screening. Mistakes in procedures or interpretation can result in false-positive or false-negative results. Results of genetic tests must remain confidential since they may reveal information about the subject's family and their future health prospects. This could lead to genetic discrimination, where people are denied jobs or insurance because of their genes. In the USA it is illegal for a company to discriminate for employment or health insurance on the basis of genetic test results.