MASTER BIOLOGY

Biology A-level

Topic 8: The control of gene expression

For more notes, tutors who are qualified doctors and medical admissions webinars visit
www.mastermedicine.co.uk

For more notes, tutors who are qualified doctors and medical admissions webinars visit
www.mastermedicine.co.uk

Mutations

Mutations are changes in the sequence of nucleotide bases in DNA. The different types of mutations are:

- Insertion/deletion mutations - where one or more nucleotide pairs are inserted or deleted from the sequence. This type of mutation alters the sequence of nucleotides after the insertion/deletion point, meaning they code for different amino acids. This is known as a **frameshift**
- Duplication - one or more bases are repeated and produce a frameshift
- Inversion - a group of bases become separated from the DNA sequence and then rejoin at the same position but in the reverse order. This therefore affects the amino acid that is produced
- Translocation - a group of bases become separated from the DNA sequence on one chromosome and are inserted into the DNA sequence on another chromosome. This can often lead to significant effects on the phenotype

Cause of mutations

Gene mutations occur randomly during DNA replication. Whilst they can happen without reason, certain factors can make these changes more likely. These factors can be **mutagenic** agents which can affect the replication of DNA. Apart from spontaneous errors, the two most common **mutagens** are:

1. Chemical mutagens - alcohol, benzene and substances in materials such as **asbestos** and **tar** which is found in tobacco
2. Ionising radiation - alpha and beta radiation are very **ionising** (and therefore damaging to DNA). UV light is also a form of radiation, as well as X-rays. Both of these can also cause damage to DNA.

Mutations can be **silent** as the mutation in the triplet of nucleotide bases may not actually code for a different amino acid. Another way in which it may be silent is if the mutation occurs in a region of the DNA that is **non-coding**. A mutation can also be silent even when the mutation causes the amino acid to change. This is because the change in **tertiary structure** of the protein might not actually affect its function.

Stem cells

Stem cells are **undifferentiated** cells which can keep dividing to give rise to other cell types. The types of stem cells you should be aware of are:

Pluripotent cells are able to give rise to many types of specialised cells apart from embryonic cells. They may also be used to repair some types of tissue, although research on this is early

Totipotent cells which can give rise to all types of specialised cells, including embryonic cells

Totipotent stem cells that are able to differentiate into any type of cell found in body and into extra embryonic cells such as those in the placenta. These cells are found in the embryo at an early stage called the **blastomere** . These stem cells are sometimes called **embryonic stem cells**. They have been used in more recent research to repair areas of damage.

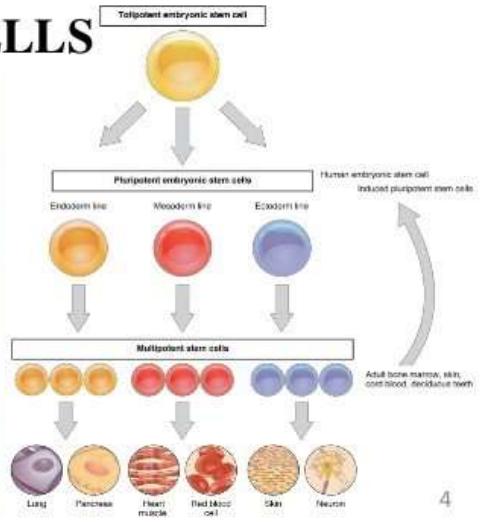
Totipotent cells in the embryo are initially unspecialised however when they become specialised they differentiate to form tissues which make up the foetus. This is because certain genes are switched on and off, irreversibly, which causes them to differentiate into a unipotent cell.

Multipotent cells can differentiate into various other cells types but are much more limited than pluripotent and totipotent cells. Once a cell is multipotent, it only has a limited number of cells it can be become. An example of this is multipotent stem cells in the bone marrow, which can only become types of blood cell. There are also some multipotent cells found in the umbilical cord, which have been the subject of lots of research in the last decade.

Unipotent cells are specialised cells. These can only differentiate into one type of cell. Scientists have found a way to create **pluripotent cells** from unipotent cells. These pluripotent cells are called induced pluripotent cells or iPCs.

TYPES OF STEM CELLS

Type of stem cell	What it can be	Examples
Totipotent cells	Each cell can develop into a new individual	Cells of embryo of 1-3 days
Pluripotent cells	Each cell can form any cell type (over 200)	Cells of blastocyst 5-14 days
Multipotent cells	Cells differentiate and can form a number of tissue types.	Fetal tissue, cord blood, adult cells



Regulation of transcription and translation

Transcription Factors

A transcription factor is a protein which can control and alter the rate of transcription of certain genes in DNA. It will do this by binding to a certain area on the DNA molecule called **promoter regions.**

An example of how transcription factors help to control gene expression lies within **oestrogen**

Oestrogen is a **steroid hormone** which has the ability to alter transcription by **modifying transcription factors.** The action of oestrogen in controlling transcription is described below:

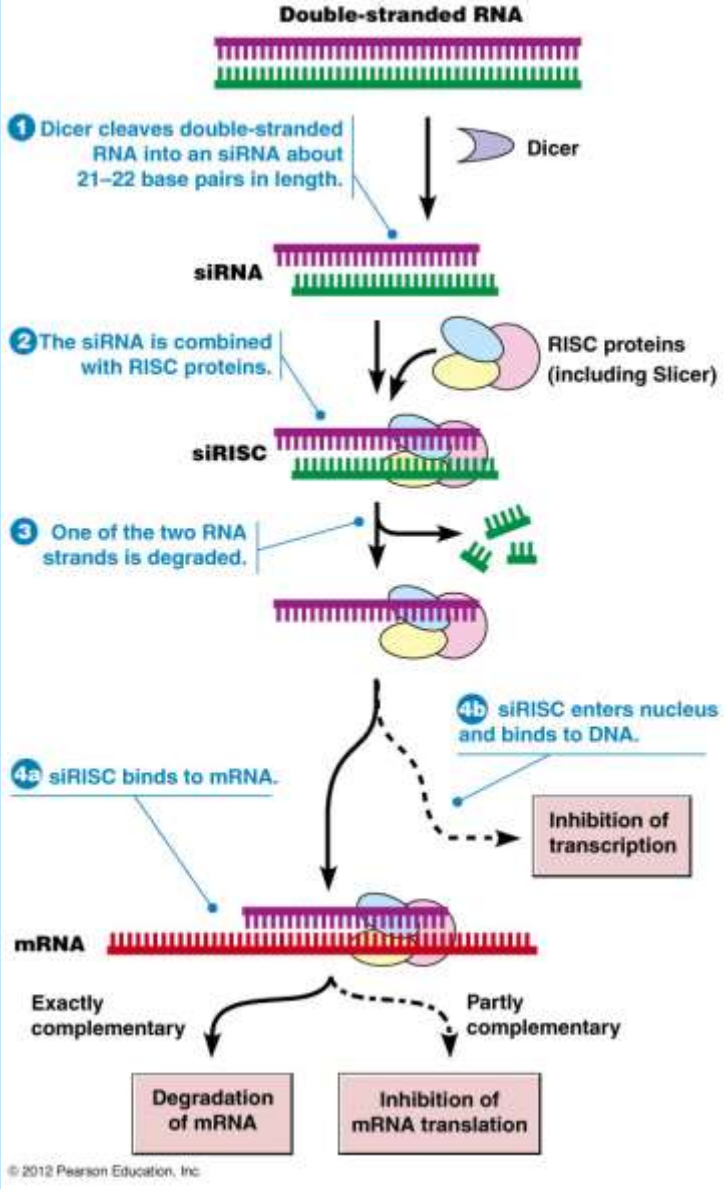
1. Oestrogen enters the cell by diffusing directly through the membrane. It does this as it is **fat (lipid) soluble**.
2. Once in the cell, oestrogen binds to its specific **receptor** within the cell (ER).
3. This receptor is attached to a **transcription factor**. The binding of oestrogen to its receptor causes a conformational change in the **tertiary structure** of the transcription factor
4. The area of the transcription factor that usually binds to the promoter region **changes shape** and it **becomes complementary** to the promoter region on DNA
5. The transcription factor can therefore enter the nucleus through a nuclear pore and bind to the promoter region on DNA. This causes the transcription of genes

RNA Interference

Normally, when we switch on transcription of a gene, it produces copious amounts of mRNA in the cytoplasm ready for **translation**. If we switch off the transcription process by interfering with the **transcription factor and/or promoter region**, we will still have mRNA in the cell which will continue to be translated, therefore allowing protein synthesis to continue long after the gene has been switched off. To prevent this from happening, both prokaryotes and eukaryotes have developed a way of **degrading** mRNA in the cytoplasm when it is no longer needed. This is called small interfering RNA.

Small interfering RNA (siRNA), also called **silencing** RNA, is used for short term switching off of genes. The way it does this can be found below:

1. Double stranded RNA in the cell is first **cleaved** by an enzyme called **dicer**. This creates shorter sections of double stranded RNA which are about 21 base pairs long
2. This **shorter section** of double stranded RNA then combines with a large protein complex called **RISC**
3. The RISC complex **degrades** one of the strands of the RNA, creating one siRNA molecule which is still attached to the RISC complex. We call this **siRISC**. Note that until the second strand of RNA is degraded, the RNA is **NOT siRNA**.
4. siRISC binds to mRNA via hydrogen bonds between complementary base pairs. In other words, the siRNA must have a **complementary base sequence** to that of the mRNA it wants to degrade
5. If the sequence is entirely complementary, the binding of siRISC will target the mRNA for **degradation**
6. If it is **partially** complementary, it will **inhibit** the translation of the mRNA
7. siRISC can also enter the nucleus and inhibit transcription of the gene as well, although this is not its main role



Epigenetic Changes

Epigenetics involves **heritable** changes in gene function, without changes to the base sequence of DNA. It shows that **environmental factors** can make changes to the function of genes which can be inherited.

- **DNA methylation** is a process by which **methyl** groups are added to **cytosine** in DNA. Methylation **suppresses** gene transcription. The change is **permanent** and prevents the cell from converting back into a stem cell or a different cell type. This prevents binding of **transcription factors** to DNA and stimulates **decreased acetylation** of histones.

- DNA **acetylation** also changes DNA structure. DNA is wound around proteins called **histones**. Acetylation of histones allows the DNA to unravel loosely from the histones, allowing transcription factors to bind. **Deacetylation** of histones causes DNA to bind more tightly to this histone. When this happens, transcriptional factors can no longer access the DNA, so the gene is switched off.

Cancer

Cancer arises due to a mutation in DNA which leads to uncontrolled cell division. Usually, cell division is controlled and regulated by various pathways which allow it to be stopped when not needed. In cancer, uncontrolled cell division leads to a tumour.

There are two types of tumours. **Benign** tumours do not tend to cause much harm apart from light mechanical damage caused by pressing against blood vessels or other cells. Benign tumours grow slowly and do not spread, whereas **malignant** tumours grow rapidly and can spread to other areas of the body via **metastasis**. Malignant tumours difficult to treat and generally if they have spread (metastasised), they are much more difficult to cure.

Cancer Definitions:

- Proto-oncogenes – healthy genes which usually stimulate normal cells to divide. They allow checkpoints in the cell cycle to be passed in a controlled manner. When they are **proto-oncogenes,** they are healthy and will not cause cancer. However, these are genes that if mutated, will become **oncogenes**, and cause cancer.

- Oncogenes - these are mutated proto-oncogenes and are permanently switched on. This allows cell division to continue unchecked and uncontrolled. It does this by permanently activating a **cell surface receptor** or growth factor.

- Tumour suppressor genes – these are healthy genes which control cell division. They check for damage in the DNA and stop cell division by stopping the cell cycle if damage is found. They can also cause cells to 'commit suicide' via programmed cell death. This is called **apoptosis.** If tumour suppressor genes are switched off, the cell cycle becomes unregulated
    - As mentioned above, methylation activates transcription of genes. The abnormal **methylation** of BRCA1 or BRCA2, tumour suppressor genes, commonly leads to breast and/or ovarian cancer
    - Abnormal **methylation** of tumour suppressor genes and **demethylation** and **acetylation** of oncogenes causes uncontrolled division

- Increased **oestrogen** concentrations can cause breast cancer. This is because oestrogen binds to its receptor on the transcription factor, activating the genes promoting cell division (oncogenes), leading to tumour formation

Gene Sequencing projects

Genome - the specific haploid set of chromosomes a species has
Proteome - All of the proteins that the genome can code for.

In recent years, we have sequenced the genome of many organisms. This means we have found the **exact sequence** of the **nucleotide bases** in every single chromosome each species has. The most ambitious was the **human genome project.**

The **Human Genome Project** is an international scientific project which has successfully determined the sequence of bases of the human genome. Potential applications include:

- Genetic **screening** of individuals with symptoms for mutated sequences causing disease
- Screening potential parents to see if they are **carriers** of disease
- **Pre-implantation** screening of embryos before they are implanted in the womb in in vitro fertilisation
- Screening for disorders such as Huntington's disease before the symptoms appear

However, there are many ethical concerns regarding the Human Genome Project, such as potential discrimination against those with certain genetic sequences. Another potential downfall may involve the misuse and ownership of the genetic information.

Determining the **genome** of simpler organisms allows us to work out the sequence of bases that code for the proteins in that organism. This could be used to work out **potential antigens** for use in vaccine production.

In more complex organisms, there is a lot more **non-coding DNA** (introns) than in simpler ones. Because of this, it is difficult to know which sections of the DNA code for proteins and therefore work out the sequence of nucleotides that codes for each proteins. The **genome** cannot easily be translated into the **proteome**. In addition, due to selective gene expression, not all proteins will be found in every cell in the body as each cell has a different function.

Gene sequencing allows for comparisons of **genomes** between individuals and between species. Comparing genomes between species allows us to learn about the evolutionary relationships between them and work out how closely related they are. We can use this for medical research too.

Comparing genomes of human individuals enables differences to be identified. This can be used for a new area of research known as "**personalised medicine**", where treatment drugs and doses are tailored to each person's particular genome. Human comparisons between an individual who has a **heritable disease** and someone who is healthy allow us to identify which genes are mutated in order to cause the disease.

Recombinant DNA technology

Recombinant DNA technology involves many ways of manipulating DNA to create **DNA Fragments**, these processes are all detailed below.

Reverse Transcriptase

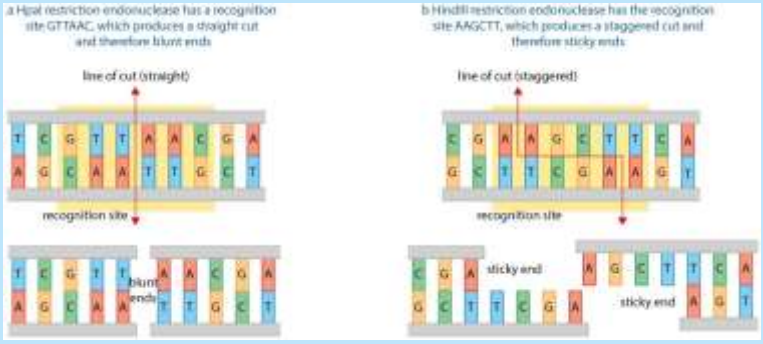
The enzyme **reverse transcriptase** is an enzyme found in some viruses and bacteria which can create **double stranded DNA** from a single strand of RNA. This allows us to create DNA of a certain gene simply by finding the mRNA in a cell for a certain protein.

We can find a cell that produces the desired protein, and extract the mRNA. We can use reverse transcriptase to produce DNA from this RNA. We can then create double stranded DNA from this using **DNA polymerase.**

Restriction Endonucleases

Restriction endonucleases are **enzymes** from bacteria, that cut DNA at specific sequences. The most useful restriction endonucleases are those that make **staggered cuts**, as they leave unpaired nucleotide bases at the ends which are available for making hydrogen bonds with **complementary bases.** As a result, we call these 'sticky ends'

Sticky ends are important because if the same restriction endonuclease is used to cut two DNA fragments, then the **ends will be complementary**. This allows then to attach together.



In-vivo gene cloning

If a DNA fragment was placed in a cell, it would be digested by enzymes. As a result, we must use a **vector** is used to **insert DNA of a useful gene** into the DNA already within the nucleus of cells. The most common vector we use, is the **plasmids** from bacterial cells.

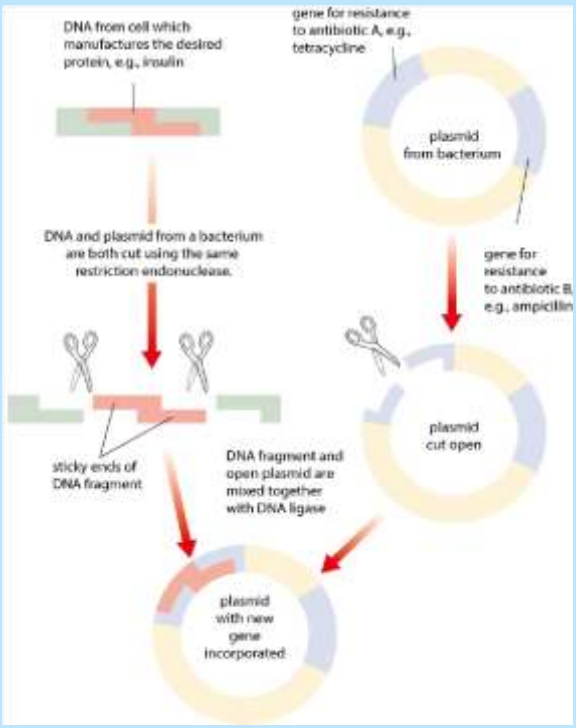
Isolated DNA fragments can be placed in plasmids in a following way:

1. The gene that we want to insert in a certain cell will be cut out of its DNA strand using a **restriction endonuclease**
2. A marker is inserted onto this DNA, such as a **fluorescent protein** which will be visible under a microscope
3. A plasmid from a bacterial cell is cut with the same restriction endonuclease. The plasmid is known as the **vector**
4. This means that the cut ends of the plasmid and the cut ends of the gene are complementary to each other. These are known as **'sticky ends'**
5. The fragments are incubated with the plasmids. If a plasmid takes up the insert, base pairing takes place between the complementary ends are joined with the use **of DNA ligase** which makes **phosphodiester** bonds between the two ends
6. Some plasmids will take up the DNA. These are known as **recombinant plasmids**
7. The plasmids are then inserted into bacteria through **electroporation**, which is essentially using an electric shock on the bacteria.
8. Not all bacteria will take up the recombinant plasmids. In order to check recombinant plasmid **has been taken up** by the bacteria, we check to see if the cells are labelled – in

this instance, the bacteria will be **fluorescent under a microscope** as they are now producing the fluorescent protein

Note: another way of working out if the bacteria has taken up a plasmid with recombinant DNA (ie. with our desired gene), is to insert the desired gene within an **existing gene for antibiotic** resistance, for instance resistance to tetracycline. Those which contain the recombinant DNA plasmid will be killed by tetracycline as we have disrupted the gene for antibiotic resistance. Those that are killed are therefore identified via **replica plating**.

Enzyme markers can also be used



Gene markers

There are different types of gene markers, these are **antibiotic restraint genes**, **fluorescent** markers and **enzyme markers**. These genes are incorporated into the plasmid so that the bacteria who have the plasmid can be separated from the bacteria that do not.

Gene therapy

A mechanism by which genetic diseases are treated or cured by masking the effect of a faulty allele and replacing it with a healthy allele

SCID

The treatment of this disease (**Severe Combined Immunodeficiency**) involves the scientific use of viruses to insert healthy alleles into their genome.

The mutation which causes **SCID** is a mutation of the gene coding for an enzyme adenosine deaminase (**ADA**). This means that T lymphocytes in the immune system do not work properly as they do not divide when they meet a pathogen. This means other cells, such as B lymphocytes, are not activated and therefore the immune system does not fight infection as well. We can treat this via injections of the actual enzyme, bone marrow transplants and finally, gene therapy.

This means we can insert the healthy allele for ADA into the **T lymphocyte**. Here is the step by step of how:

1. Isolating the healthy ADA allele using genetic markers, and cutting it out via restriction endonucleases
2. We then insert this gene into **retroviruses.** Retroviruses are able to insert their own RNA into the cells they infect, as DNA
3. We can therefore give a deactivated version of a retrovirus with **recombinant DNA** and therefore produce the healthy enzyme, to the T cells
4. We infect the person's T cells with the retroviruses, and it incorporates its own healthy allele into the T cell's DNA
5. This means that now, the T cells **can produce healthy ADA**

Ethical worries

- May be a slippery slope towards designer babies
- May impact the functioning of other healthy genes instead

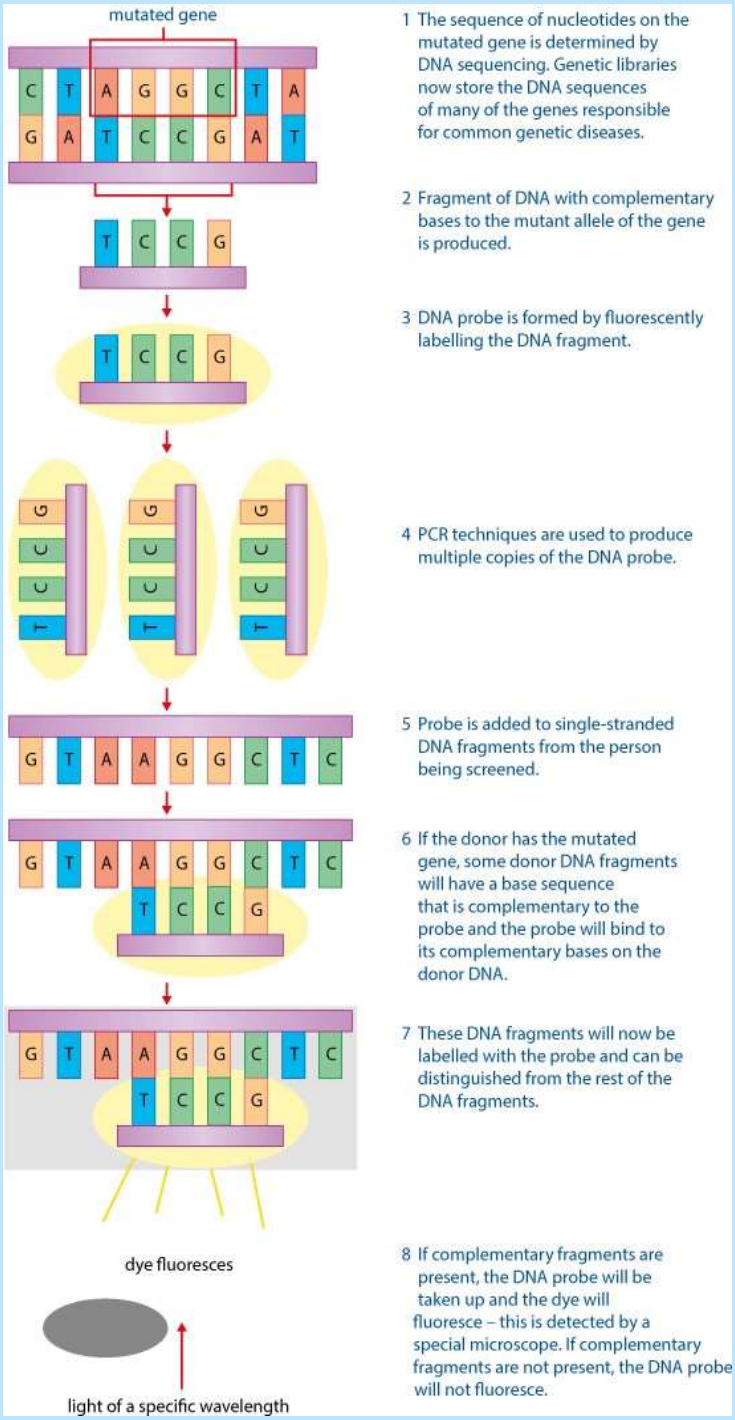
Gene technologies

Polymerase Chain Reaction

Polymerase chain reaction known as PCR is used to amplify DNA by making millions of copies of a given DNA sample. This is because often, a DNA sample given is far too small to sequence or genetically fingerprint (see section on genetic fingerprinting).

1. A reaction mixture is set up by mixing the **DNA sample, primers, free nucleotides** and **DNA polymerase** which is the enzyme involved in creating new DNA strands.
2. The mixture is then heated to **95 degrees** to **break the hydrogen bonds**. This separates the DNA into two strands, each with nucleotide bases on each strand now exposed
3. The mixture is then **cooled** to a temperature of between **50-65 degrees** so free nucleotides can bind (**anneal**) to their **complementary** nucleotide bases on the exposed DNA strands via hydrogen bonding
4. Temperature is increased to 70 degrees as this is the optimum temperature of DNA polymerase (AKA **Taq Polymerase**). This enzyme is found in hot springs.
5. DNA polymerase creates a copy of the sample by complementary base pairing using the free nucleotides.
6. The result is two double DNA strands that were identical to the first one

7. This cycle is **repeated** around 30 times and gives rise to an amount of DNA sufficient to create a **DNA profile**.



In-vivo and in-vitro cloning

In vitro – this does not require a living organism. It is this type of gene cloning we see in PCR. This is fast, automated and reliable once conditions are established

- Pros - This does not require living cells
- Cons - Contamination and errors. Also can be expensive

In-vivo - gene cloning by placing recombinant plasmids into bacteria and allowing them to replicate

- Pros – Accurate. Actually produces the gene that the DNA codes for.
- Cons - Time consuming, requires monitoring of cell growth and optimum conditions for bacterial survival and reproduction, requires living cells

DNA probes

A DNA probes are a short, single stranded DNA molecule designed to complementary to a sequence that we want to detect. We can create one or two DNA probes and then create millions of copies (amplification) using **Polymerase Chain Reaction**. These probes usually have a marker on them which allows them to be visible in some way. Common types include **radioactive probes** or **fluorescent probes**.

Genetic fingerprinting

Genetic fingerprinting is a technique that can detect differences and differences in DNA. We can use it to see how closely related two individuals are, or use it in forensics to identify if a suspect's DNA is the same as that found at the crime scene.

It uses **variable number tandem repeats (VNTRs)** which are very short sequences of bases which repeat over and over again. They are found in over 1000 different positions on the human genome

Whilst every single humans VNTRs are at the same loci, their **lengths differ between individuals.** In other words, one person may only have 3 repeats, but another may have 300 in the same position. As a result, this makes their DNA profile **unique.**

This is a common A Level question, so here is a step by step of each step. The first step seems pointless, but there are many mark schemes on which you are given a mark for it!

1. Collect DNA by **swabbing of the cheeks** for cheek cells, or can use teeth, bones or hairs too at a crime scene
2. Amplify the DNA using **Polymerase Chain Reaction (PCR)**
3. The DNA is cut into smaller pieces using **restriction endonucleases**. They will all cut the DNA at the same sequence. The specific sequence is known as a **recognition site**
4. The restriction endonucleases will cut parts of the DNA where there is **NOT a VNTR.** They will leave the VNTRs intact
5. Therefore, if a person has **many repeats in their VNTRs**, it is going to create a fragment which is much larger than a person who has barely any repeats
6. People will therefore have different fragment lengths
7. The fragments are **labelled with a radioactive probe** through hybridisation
8. The fragments will be **separated out via gel electrophoresis** (see section on gel electrophoresis if you don't know what this is)
9. The smallest fragments travel the furthest and the longest fragments travel the least along the gel

10. An X-ray film is placed over the top of the gel tray. Because the **fragments are radioactively labelled**, they will expose the X ray film and leave a **band**
11. Because everyone has different length fragments, they will have **bands in different places.** We call this their **genetic fingerprint**
12. The more similar DNA is to each other, the more similar the placement of the bands in the gel. Therefore, people who are related will have similar genetic fingerprints

We can compare multiple samples of DNA to others in order to identify if individuals are related, or to see if someone's DNA is the same as that found at a crime scene

Gel Electrophoresis

Gel electrophoresis is a process used to separate the DNA fragments and proteins according to their size using an electric current.

1. The DNA samples are all placed wells at one end of a tray of gel
2. There is an **anode** (negatively charged) at the same end as the wells and a **cathode** (positively charged) at the other end of the gel
3. Electric current is run between the anode and cathode
4. Because DNA is negatively charged, it will start to move in the direction of the positively charged cathode at the other end
5. Smaller fragments can move faster and will therefore move **further**. Larger fragments of DNA are the opposite
6. When the current is turned off, there will be **bands in the gel** which are exposed onto an X-Ray film. These bands will be at different places depending on whose DNA it is.
7. The position of the bands is called the **genetic fingerprint (see also the section on genetic fingerprinting)**