Manipulating Genomes

Methods

PCR: for amplifying DNA.

1. Load thermal cycler machine with:
    - target DNA: sequence to be amplified
    - primers: short sequences that bind to the start of the target sequence
    - free nucleotides: for building new DNA strands
    - *taq* polymerase: thermally stable DNA polymerase
2. Heat to 95 °C to separate target DNA strands.
3. Cool to 55 °C to allow primers to bind.
4. Heat again to 72 °C to allow polymerase to catalyse formation of new complementary DNA strand.
5. Repeat 2-5.

Gel electrophoresis: for separating DNA fragments or proteins by size.

1. Collect sample of DNA.
2. Amplify using PCR.
3. Cut with restriction enzymes to produce fragments of different length.
4. Load samples into wells in gel and apply an electric current.
5. DNA fragments move towards the positive electrode.
6. Smaller fragments move faster so will travel further in a set period of time.
7. Transfer fragments to paper / a membrane.
8. Add DNA probes that are complementary to the DNA and that have a radioactive or fluorescent tag: the probes will hybridise with the DNA.
9. Use X-ray or UV to show the location of the bands of DNA.

DNA sequencing: for determining the base sequence of a piece of DNA.

Sanger sequencing:

1. Carry out PCR to amplify target DNA (that will be sequenced).
2. Load four separate test tubes with:
    - target DNA
    - free nucleotides
    - either A, C, T or G terminator nucleotides with radioactive / fluorescent tags
    - DNA polymerase
    - DNA primers
3. Allow DNA replication to occur in each of the tubes (very similar to PCR): this will result in four tubes that contain multiple DNA fragments of different length, all of which end in the same base, i.e. tube 1 fragments will all end with A, tube 2 fragments will all end with C etc.
4. Put the contents of each of the four tubes into four different wells in a gel and run gel electrophoresis.
5. This will separate the fragments by size, allowing the position of each base in the sequence to be determined, i.e. if the shortest fragment ends with A, then we know that A is the first base in the sequence, then the next shortest fragment ends with C, so we know that C is the next base in the sequence, and so on.

High-throughput sequencing is a set of sequencing technologies that can read millions of different fragments in parallel, producing very large amounts of sequence data rapidly and at much lower cost than traditional Sanger sequencing.

Genetic engineering: inserting a gene from one organism into the genome of another.

1. Identify the gene of interest using DNA sequencing data, e.g. from a computerised database.
2. Cut the DNA from the genome using restriction enzymes.
3. Amplify the sequence using PCR.
4. Insert the sequence into a vector, e.g. a bacterial plasmid.
    - Use the same restriction enzyme to cut the plasmid to produce complementary sticky ends.
    - Use DNA ligase to join the new DNA to the cut plasmid; this forms a recombinant plasmid.
5. Mix recombinant plasmids with bacterial cells and encourage the cells to take up the plasmids by electroporation; cells containing new DNA are said to be transformed.
    - Electroporation: a brief electric pulse that temporarily makes holes in the plasma membrane.
6. Select cells containing recombinant DNA using marker genes, e.g. genes that code for fluorescent proteins will result in transformed cells that glow green.
7. Culture the recombinant cells and purify their new protein products, e.g. insulin.

Applications

DNA profiling: for determining the length of DNA fragments in the genome of an individual

1. Collect a DNA sample, e.g. from a crime scene or a patient blood sample.
2. Run PCR to amplify the DNA.
3. Use gel electrophoresis to separate the fragments (see method above).
4. Analyse the resulting DNA fingerprints, e.g.:
    - in forensics look for suspect / victim DNA banding patterns that match the crime scene
    - in disease risk analysis look for particular banding patterns that indicate the presence of a mutated allele
    - in paternity testing aim to match the 50% of bands in a child's DNA (that do not match with the mother) with the bands in a potential father's DNA
    - in classification / evolutionary research look for similarities in banding patterns between different species

Genome-wide comparisons: DNA sequencing data from different organisms can be uploaded to computer databases; the use of computer tools to store and analyse sequence data is known as bioinformatics.

This allows for comparison of sequences, e.g. for:

- investigation of the relationship between genotype and phenotype, i.e. which DNA sequences correspond with which characteristics
- studying the genomes of infectious organisms so that their spread can be tracked, e.g. the development of COVID19 tests
- comparison of the DNA sequences of different groups of organisms to determine evolutionary relationship

Storing DNA data in computer databases also allows for the generation of computer models and simulations, e.g. simulating how changing a DNA sequence would affect chemical pathways in a cell.

Synthetic biology: designing and building new biological parts, systems and organisms.

- DNA sequencing allows scientists to determine the exact base sequence of genes and how they translate to proteins.
- Knowing the precise sequence of a gene allows scientists to design synthetic sequences, e.g. PCR primers, or genes to insert into organisms during genetic engineering that will give rise to specific proteins.