

What is a gene?

A gene is a section of DNA that contains the coded information for making polypeptides and functional RNA.

The coded info is in the form a specific sequence of bases along the DNA molecule.

Polypeptides make up proteins and so genes determine the proteins of an organism. (enzymes are proteins) as enzymes control chemical reactions they are responsible for an organisms development and activities.

Genes (along with other environmental factors) determine the nature and development of all organisms.

A gene is a section of DNA located at a particular position, called a locus on a DNA molecule.

The gene is a base sequence of DNA that codes for:

- The amino acid sequence of a polypeptide
- Or a functional RNA, including ribosomal RNA and transfer RNA's

The Genetic Code:

Scientists suggested there must be a **minimum of 3 bases that coded for each amino acid**. Their reasoning was:

1. only 20 different amino acids regularly occur in proteins
2. Each amino acid must have its own code of bases on the DNA
3. Only four different bases (A,G,C,T) are present in DNA
4. If each base coded for a different amino acid, only four different amino acids could be coded for
5. Using a pair of bases, 16 different codes are possible (still inadequate)
6. Three bases produces 64 different codes, more than enough to satisfy the requirements of 20 different amino acids.

As the code has 3 bases for each amino acid, each one is called a triplet.

Features of the code:

- A few amino acids are coded for by only a single triplet.
- The remaining amino acids are coded for by between 2 and 6 triplets each.
- The code is known as the 'degenerate code' because amino acids are coded for by more than one triplet.
- A triplet is always read in one particular direction along the DNA strand.
- The start of a DNA sequence that codes for a polypeptide is always the same triplet. This codes for the amino acid methionine. If this first molecule doesn't form part of the final polypeptide its later removed.
- Three triplets do not code for any amino acid. These are called 'stop codes' and mark the end of a polypeptide chain.
- The code is 'non-overlapping' = each base in the sequence is only read once. Thus six bases numbered 123456 are read as triplets 123 and 456 rather than as triplets: 123 234 345 456.
- The code is 'universal', with a few minor exceptions, each triplet codes for the same amino acid in all organisms. This is indirect evidence for evolution.

Much of the DNA in eukaryotes doesn't code for polypeptides. For example, **between genes there are non-coding sequences made up of multiple repeats of base sequences. Even within genes, only certain sequences code for amino acids. These coding sequences are called EXONS. Within the gene these exons are separated by further non-coding sequences called INTRONS.**