SCoPE

Scientists' Collaborative Project with Educators

Exploring Epigenetics

Student sheets: The Genetic Code

Task

Your task is to work from Sanger-sequenced DNA, to determine the mRNA sequence and then the amino acid sequence of the polypeptide chain.

Afterwards, using the single letter designations for the amino acids, you will find a message. Research the meaning of this message in preparation for learning about gene regulation.

Learning objectives

In order to complete this task you need to recall that:

- Sanger sequencing can be used to determine DNA sequence
- During transcription, a section of a DNA strand is transcribed into messenger RNA (mRNA)
- During translation the mRNA is then translated into a polypeptide chain formed of amino acids
- A DNA (deoxyribonucleic acid) nucleotide consists of a deoxyribose sugar, a phosphate group and one of the nitrogen-containing organic bases: adenine (A), cytosine (C), guanine (G) or thymine (T)
- An RNA (ribonucleic acid) nucleotide consists of a ribose sugar, a phosphate group and one of the nitrogen-containing organic bases: adenine (A), cytosine (C), guanine (G) or uracil (U)
- In DNA, base pairing of nitrogen-containing organic bases always occurs so that Adenine pairs with Thymine (Uracil in RNA), and Guanine pairs with Cytosine
- The order of bases on DNA is called the genetic code and consists of triplets of bases, or codons, each of which codes for a particular amino acid
- The genetic code is degenerate, so more than one triplet codes for the same amino acid
- The genetic code contains start and stop codons
- An mRNA molecule is a long single strand created during transcription in which the base sequence is complementary to DNA. Each set of three bases, called a codon, match a triplet on the DNA. Transcription ends when a stop codon is reached.
- During translation at the ribosome, mRNA codons attach to a tRNA by complementary base pairing and the amino acid attached to the tRNA molecule joins to the extending polypeptide chain. Translation ends when a stop codon is reached.

SCoPE

Instructions

The process of working from Sanger sequence data to determine the amino acid sequence of the polypeptide chain is illustrated in Diagram 1. This is the process that you are going to follow. Each of the techniques, processes and steps you need to undertake are explained in more detail on the following pages and you can complete the exercise on the student sheets provided on pages 6 and 7.

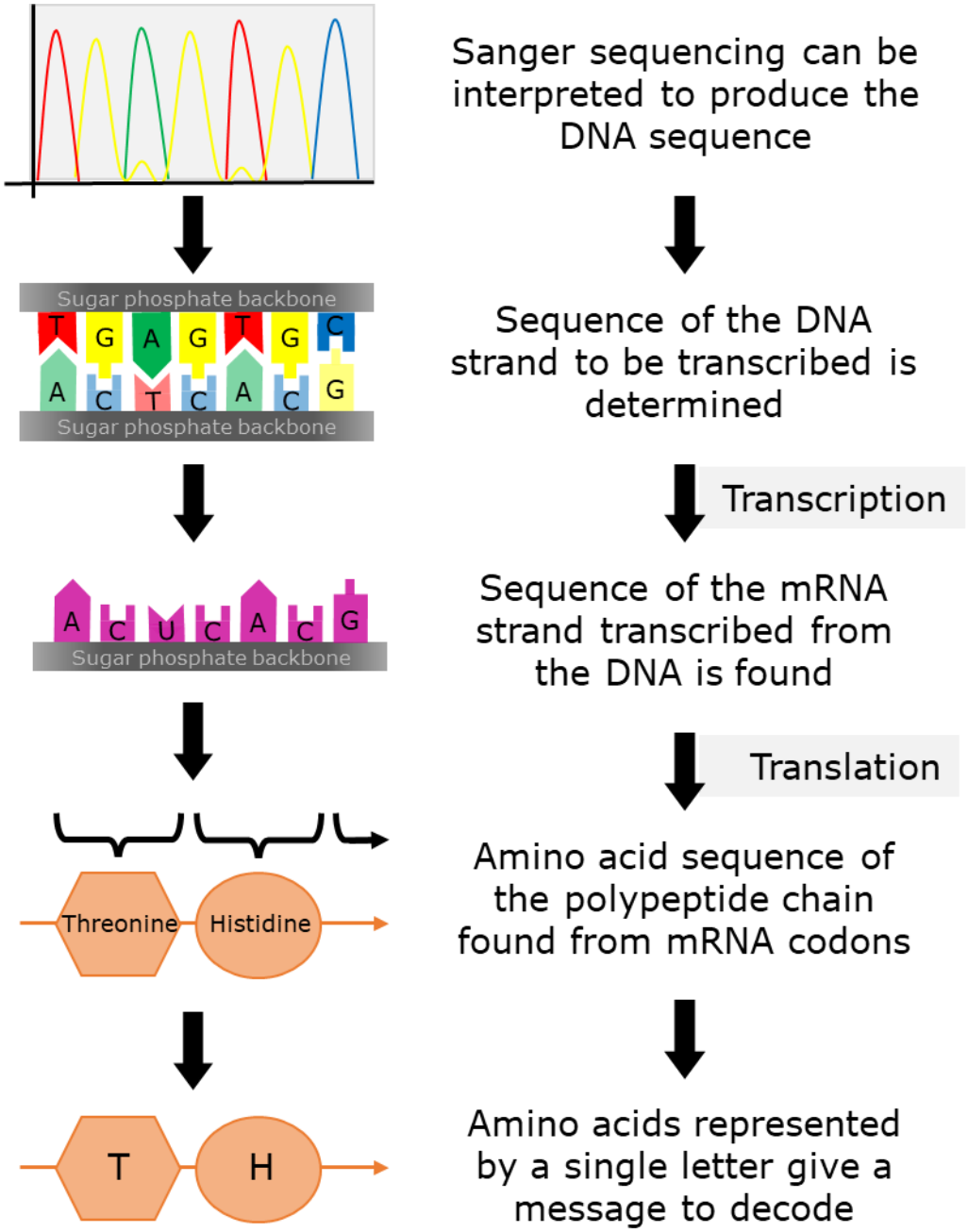


Diagram 1. The process of working from Sanger sequence data to determine the amino acid sequence of a polypeptide chain.

A. Sanger sequencing data

SCoPE

Sanger sequencing was named after Frederick Sanger, a British scientist who won 2 Nobel Prizes. Sanger sequencing can be used to determine the sequence of nitrogen containing organic bases (adenosine, cytosine, guanine and thymine) in DNA. To do this:

1. The DNA to be sequenced is split into 4 reactions. Within each reaction the DNA is replicated multiple times, using 4 standard nucleotides and a modified version of one nucleotide that is fluorescently labelled and prevents DNA polymerase continuing replication:
    - Modified adenosine is associated with a green dye
    - Modified cytosine with blue dye
    - Modified guanine with yellow dye
    - Modified thymine with red dye
2. This means that in each of the 4 reactions when modified nucleotides are incorporated, replication stops, producing fluorescently labelled fragments of different lengths.
3. These fragments of different length are separated by size (using electrophoresis) and the colour of the fluorescent dye at each position is used to determine the DNA sequence.

B. Complete the DNA sequence

On pages 6-7, standard fluorescent dye colours have been used. Write the DNA sequence into the boxes underneath where it says 'B. Complete the DNA sequence'.

C. Complete the mRNA sequence

An mRNA molecule is a long, single strand created during transcription. DNA is the template for the mRNA molecule. RNA nucleotides form complementary base pairs with the DNA. Each set of three bases in the mRNA, called a codon, matches a triplet on the DNA. Once complementary base pairing occurs, the enzyme RNA polymerase catalyses the formation of phosphodiester bonds between nucleotides to form the RNA polymer.

In mRNA the sequence AUG is the start codon where transcription begins. Sequences UAA, UAG and UGA all act as stop codons, ending transcription.

Transcription occurs in the nucleus. Completed, mature mRNA (formed by splicing in eukaryotic cells) moves out of the nucleus and attaches to a ribosome in the cytoplasm where translation occurs.

On pages 6-7, use the DNA sequence you completed in part B to find the mRNA sequence. Write the mRNA sequence into the boxes underneath where it says 'C. Complete the mRNA sequence'.

D. Complete the amino acid sequence

A sequence of amino acids, joined together by peptide bonds, forms a polypeptide chain during translation. The genetic code is degenerate, so more than one triplet codes for the same amino acid. Diagram 2 below shows the amino acids that different mRNA codons encode.

During translation, mRNA attaches to a ribosome where tRNA (transfer RNA) molecules are involved in forming the polypeptide. tRNAs are small, single-stranded molecules, made up of around 80 nucleotides folded into a clover leaf shape. At one end of the tRNA molecule there is an anticodon and at the other end is an extended region where corresponding amino acids bind. At the ribosome, the anticodon of a tRNA attaches itself to mRNA by complementary base pairing. The amino acid attached to the other end of the tRNA molecule joins to the growing polypeptide chain, then the tRNA molecule detaches from the amino acid.

Once completed, the polypeptide chain folds to form a protein.

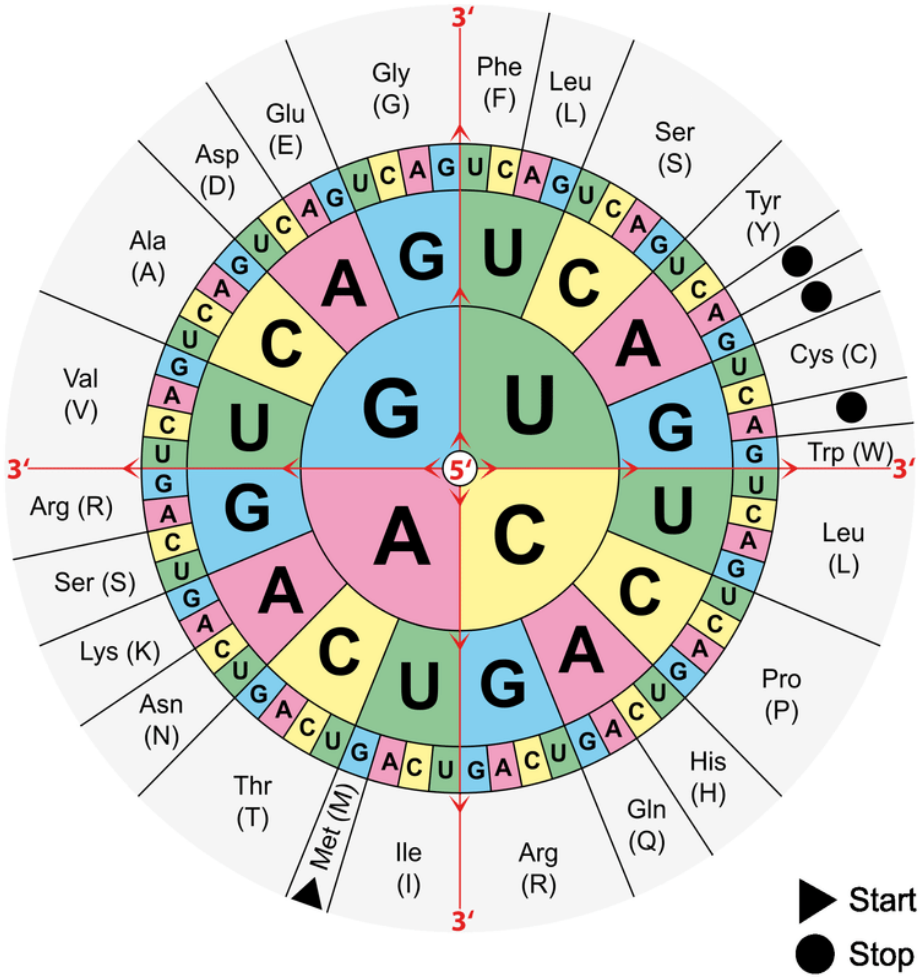


Diagram 2. The amino acids encoded by different mRNA codons.

On pages 6-7, use the information about the amino acids encoded by different mRNA codons (from Diagram 2), to complete the amino acid sequence. Write the amino acid sequence into the boxes underneath where it says 'D. Complete the amino acid sequence'.

E. Convert the amino acids to single letters

Each of the amino acids is also given a single letter code, to prevent the need to write out their names in long-hand.

On pages 6-7, use the information about the amino acids' single letter codes (from Diagram 2), to complete the amino acid sequence. Write the single letter amino acid sequence into the boxes underneath where it says 'E. Convert the amino acids to single letters'.

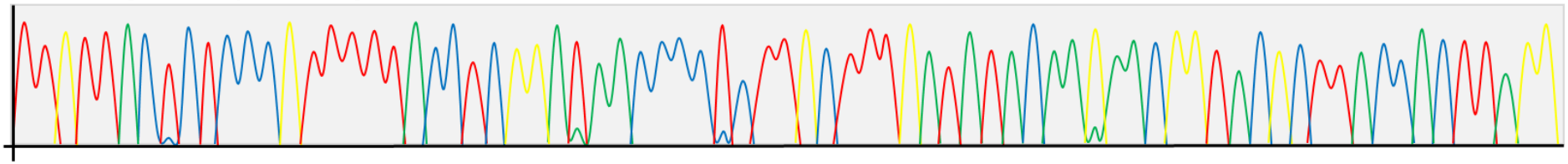
F. Research the meaning of the message

Finally, after writing in the single letter designations for the amino acids in part E, you should be able to read a message between the start and stop codons. Research the meaning of this message in preparation for learning about gene regulation.

SCoPE

Sheet 1

A. Sanger sequencing data



B. Complete the DNA sequence

C. Complete the mRNA sequence

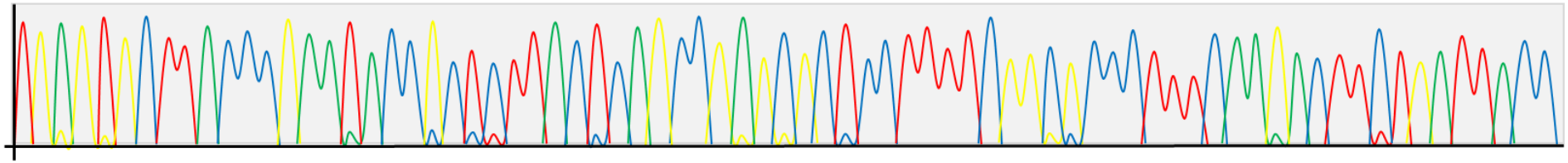
D. Complete the amino acid sequence

E. Convert the amino acids to single letters

SCoPE

Sheet 2

A. Sanger sequencing data



B. Complete the DNA sequence

C. Complete the mRNA sequence

D. Complete the amino acid sequence

E. Convert the amino acids to single letters