Essential idea: The structure of DNA allows efficient storage of genetic information.

- There is 2m of DNA in each human cell, however the most cells in the human body have a diameter of 10 μm. This DNA is divided in chromosomes and coiled around proteins called histones so that it can be efficiently stored in each cell's nucleus. The human genome project which has decoded the case sequence for the whole 2m of the human genome requires a data warehouse (pictured) to store the information electronically. This should give a good idea of just how efficient DNA is at storing information and why it needs to be so.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Guidance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2.6.U1</strong></td>
<td>The nucleic acids DNA and RNA are polymers of nucleotides.</td>
</tr>
<tr>
<td><strong>2.6.U2</strong></td>
<td>DNA differs from RNA in the number of strands present, the base composition and the type of pentose.</td>
</tr>
<tr>
<td><strong>2.6.U3</strong></td>
<td>DNA is a double helix made of two antiparallel strands of nucleotides linked by hydrogen bonding between complementary base pairs.</td>
</tr>
<tr>
<td><strong>2.6.A1</strong></td>
<td>Crick and Watson’s elucidation of the structure of DNA using model making.</td>
</tr>
<tr>
<td><strong>2.6.S1</strong></td>
<td>Drawing simple diagrams of the structure of single nucleotides of DNA and RNA, using circles, pentagons and rectangles to represent phosphates, pentoses and bases. <strong>In diagrams of DNA structure, the helical shape does not need to be shown, but the two strands should be shown antiparallel. Adenine should be shown paired with thymine and guanine with cytosine, but the relative lengths of the purine and pyrimidine bases do not need to be recalled, nor the numbers of hydrogen bonds between the base pairs.</strong></td>
</tr>
</tbody>
</table>
2.6.U1 The nucleic acids DNA and RNA are polymers of nucleotides.

Some key points about the structure of DNA:

DNA is a **double-helix**: it has **two strands** that twist around each other.

Each strand is made of single units called **nucleotides**.

It has a **sugar-phosphate backbone**.

Bases join the two strands by hydrogen bonds. These bases are **cytosine**, **guanine**, **adenine** and **thymine**.

**Complementary base pairing** is a key idea in genetics: **C pairs with G**, **T with A**.

Each strand of DNA can be **millions of base-pairs** in length and is coiled up to make **chromosomes**.

http://youtu.be/qy8dk5iS1f0
2.6.U1 The nucleic acids DNA and RNA are polymers of nucleotides.

**A nucleotide**: a single unit of a nucleic acid

Nucleic acids are very large molecules that are constructed by linking together nucleotides to form a polymer.

There are two types of nucleic acid: DNA and RNA.
2.6.U1 The nucleic acids DNA and RNA are polymers of nucleotides.

**A nucleotide**: a single unit of a nucleic acid

- five carbon atoms = a pentose sugar
- If the sugar is Deoxyribose the polymer is Deoxyribose Nucleic Acid (DNA)
- If the sugar Ribose the polymer is Ribose Nucleic Acid (RNA)

- acidic
- negatively charged
- contains nitrogen
- has one or two rings in it’s structure
What is the relevance of this coffee cup to this topic?
2.6.U1 The nucleic acids DNA and RNA are polymers of nucleotides.

There are **four nitrogenous bases** in DNA:

- Adenine (A)
  - ![Adenine](image)
- Guanine (G)
  - ![Guanine](image)
- Cytosine (C)
  - ![Cytosine](image)
- Thymine
  - ![Thymine](image)

RNA Shares the same bases except that Uracil (U) replaces Thymine.

*n.b. when talking about bases always use the full name on the first instance*
2.6.U1 The nucleic acids DNA and RNA are polymers of nucleotides.

A strand of nucleotides is joined by covalent bonds

The sequence of bases makes up the genetic code.

- Nucleotides are linked into a single strand via a condensation reaction.
- Bonds are formed between the phosphate of one nucleotide and the pentose sugar of the next.
- The phosphate group (attached to the 5'-C of the sugar) joins with the hydroxyl (OH) group attached to the 3'-C of the sugar.
- This results in a **phosphodiester bond** between the two nucleotides and the formation of a water molecule.
- Successive condensation reactions between nucleotides result in the formation of a long strand.
2.6. U2 DNA differs from RNA in the number of strands present, the base composition and the type of pentose.

<table>
<thead>
<tr>
<th></th>
<th>RNA</th>
<th>DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bases</strong></td>
<td>Adenine (A)</td>
<td>Adenine (A)</td>
</tr>
<tr>
<td></td>
<td>Guanine (G)</td>
<td>Guanine (G)</td>
</tr>
<tr>
<td></td>
<td><strong>Uracil (U)</strong></td>
<td><strong>Thymine (T)</strong></td>
</tr>
<tr>
<td></td>
<td>Cytosine (C)</td>
<td>Cytosine (C)</td>
</tr>
<tr>
<td><strong>Sugar</strong></td>
<td><img src="http://commons.wikimedia.org/wiki/File:RiboseAndDeoxy.gif" alt="Ribose" /></td>
<td><img src="http://commons.wikimedia.org/wiki/File:RiboseAndDeoxy.gif" alt="Deoxyribose" /></td>
</tr>
<tr>
<td><strong>Number of strands</strong></td>
<td><strong>Single stranded</strong>, and often, but not always, linear in shape</td>
<td><strong>Two anti-parallel, complementary strands</strong> form a double helix</td>
</tr>
</tbody>
</table>

http://commons.wikimedia.org/wiki/File:RiboseAndDeoxy.gif
DNA is a double strand of polynucleotides.

The sugar-phosphate backbone is on the outside, the bases are on the inside.

The strand is held together by hydrogen bonds between the bases.

A only pairs with T, G only pairs with C. This is called complementary base pairing.

The two strands run in opposite directions. This is called anti-parallel.

DNA twists into a double-helix, held by more hydrogen bonds.

http://science.nhmccd.edu/biol/bio1int.htm#dna
2.6. U3 DNA is a double helix made of two antiparallel strands of nucleotides linked by hydrogen bonding between complementary base pairs.
2.6. U3 DNA is a double helix made of two antiparallel strands of nucleotides linked by hydrogen bonding between complementary base pairs.

How is the double helix structure maintained?

- **Hydrogen bonds** hold adjacent sections together.
- **Hydrogen bonds** hold complementary base pairs together.
- Complementary base pairing ensures that mistakes are not made when copying or transcribing DNA.
- Sugar-phosphate backbone is hydrophilic, so it is positioned on the outside.
- Nitrogenous bases are very reactive, so they are protected on the inside.
- Polynucleotides are anti-parallel: they run in opposite directions.

http://science.nhmccd.edu/biol/biol1int.htm#dna
2.6. U3 DNA is a double helix made of two antiparallel strands of nucleotides linked by hydrogen bonding between complementary base pairs.

In Summary:

- Each polynucleotide chain (strand) consists of a chain of **nucleotides bonded covalently**.

- Two polynucleotide chains of DNA are held together by hydrogen bonds between **complementary base pairs**:
  - Adenine pairs with thymine (A=T) via two hydrogen bonds
  - Guanine pairs with cytosine (G=C) via three hydrogen bonds

- In order for bases to be facing each other and thus able to pair, the two strands must run in opposite directions (i.e. they are **anti-parallel**)

- As the polynucleotide chain lengthens, the atoms that make up the molecule will arrange themselves in an optimal energy configuration. This position of least resistance results in the double-stranded DNA twisting to form a **double helix** with approximately 10 - 15 bases per twist.
2.6.S1 Drawing simple diagrams of the structure of single nucleotides of DNA and RNA, using circles, pentagons and rectangles to represent phosphates, pentoses and bases.

Use this simple, but very effective You Tube video to learn how to draw the nucleotides making up a short section of a DNA molecule.

To make sure you have learn this skill you need to practice it repeatedly.

n.b. ideally label lines should be drawn with a ruler and should not have arrow heads.

http://youtu.be/kTH13ol8BSI
Whilst others worked using an experimental basis Watson and Crick used stick-and-ball models to test their ideas on the possible structure of DNA. Building models allowed them to visualize the molecule and to quickly see how well it fitted the available evidence.

It was not all easy going however. Their first model, a triple helix, was rejected for several reasons:

- The ratio of Adenine to Thymine was not 1:1 (as discovered by Chargaff)
- It required too much magnesium (identified by Franklin)

From their setbacks they realized:

- DNA must be a double helix.
- The relationship between the bases and base pairing
- The strands must be anti-parallel to allow base pairing to happen

Because of the visual nature of their work the second and the correct model quickly suggested:

- Possible mechanisms for replication
- Information was encoded in triplets of bases
2.6 A1 Crick and Watson’s elucidation of the structure of DNA using model making.

Watson and Crick gained Nobel prizes for their discovery. It should be remembered that their success was based on the evidence they gained from the work of others. In particular, the work of Rosalind Franklin and Maurice Wilkins, who were using X-ray diffraction was critical to their success.

Find out more about the discovery of DNA:

http://www.nobelprize.org/educational/medicine/dna_double_helix/readmore.html

http://youtu.be/sf0YXnAFBs8
2.7 DNA replication, transcription and translation

Essential Idea: Genetic information in DNA can be accurately copied and can be translated to make the proteins needed by the cell.

- The image shows an electron micrograph of a Polysome, i.e. multiple ribosomes simultaneous translating a molecule of mRNA. The central strand is the mRNA. The darker circular structures are the ribosomes and the side chains are the newly formed polypeptides.

http://bioknowledgy.weebly.com/

### Understandings

<table>
<thead>
<tr>
<th>Statement</th>
<th>Guidance</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.7.U1 The replication of DNA is semi-conservative and depends on complementary base pairing.</td>
<td></td>
</tr>
<tr>
<td>2.7.U2 Helicase unwinds the double helix and separates the two strands by breaking hydrogen bonds.</td>
<td></td>
</tr>
<tr>
<td>2.7.U3 DNA polymerase links nucleotides together to form a new strand, using the pre-existing strand as a template.</td>
<td>The different types of DNA polymerase do not need to be distinguished.</td>
</tr>
<tr>
<td>2.7.U4 Transcription is the synthesis of mRNA copied from the DNA base sequences by RNA polymerase.</td>
<td></td>
</tr>
<tr>
<td>2.7.U5 Translation is the synthesis of polypeptides on ribosomes.</td>
<td></td>
</tr>
<tr>
<td>2.7.U6 The amino acid sequence of polypeptides is determined by mRNA according to the genetic code.</td>
<td></td>
</tr>
<tr>
<td>2.7.U7 Codons of three bases on mRNA correspond to one amino acid in a polypeptide.</td>
<td></td>
</tr>
<tr>
<td>2.7.U8 Translation depends on complementary base pairing between codons on mRNA and anticodons on tRNA.</td>
<td></td>
</tr>
<tr>
<td>Statement</td>
<td>Guidance</td>
</tr>
<tr>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td>2.7.A1</td>
<td>Use of Taq DNA polymerase to produce multiple copies of DNA rapidly by the polymerase chain reaction (PCR).</td>
</tr>
<tr>
<td>2.7.A2</td>
<td>Production of human insulin in bacteria as an example of the universality of the genetic code allowing gene transfer between species.</td>
</tr>
<tr>
<td>2.7.S1</td>
<td>Use a table of the genetic code to deduce which codon(s) corresponds to which amino acid.</td>
</tr>
<tr>
<td>2.7.S2</td>
<td>Analysis of Meselson and Stahl’s results to obtain support for the theory of semi-conservative replication of DNA.</td>
</tr>
<tr>
<td>2.7.S3</td>
<td>Use a table of mRNA codons and their corresponding amino acids to deduce the sequence of amino acids coded by a short mRNA strand of known base sequence.</td>
</tr>
<tr>
<td>2.7.S4</td>
<td>Deducing the DNA base sequence for the mRNA strand.</td>
</tr>
</tbody>
</table>
2.7.U2 Helicase unwinds the double helix and separates the two strands by breaking hydrogen bonds.

Helicase
- The ‘ase’ ending indicates it is an enzyme
- This family of proteins varies, but are often formed from multiple polypeptides and doughnut in shape

2.7. U2 Helicase unwinds the double helix and separates the two strands by breaking hydrogen bonds.

1. DNA Helicase unwinds and unzips DNA

- Unwinds the DNA Helix
- Separates the two polynucleotide strands by breaking the hydrogen bonds between complementary base pairs
- ATP is needed by helicase to both move along the DNA molecule and to break the hydrogen bonds
- The two separated strands become parent/template strands for the replication process
2.7. U3 DNA polymerase links nucleotides together to form a new strand, using the pre-existing strand as a template.

DNA Polymerase
- The ‘ase’ ending indicates it is an enzyme
- This protein family consists of multiple polypeptides sub-units
- This is DNA polymerase from a human.
- The polymerisation reaction is a condensation reaction
2.7. U3 DNA polymerase links nucleotides together to form a new strand, using the pre-existing strand as a template.

2. **DNA Polymerase** creates **complementary strands**

- DNA polymerase always moves in a 5’ to 3’ direction.
- DNA polymerase catalyses the covalent phosphodiester bonds between sugars and phosphate groups.
- DNA Polymerase proof reads the complementary base pairing. Consequently mistakes are very infrequent occurring approx. once in every billion bases pairs.

Free nucleotides are collected by DNA polymerase and attached to the new strand by complementary base pairing.

- Free nucleotides are deoxynucleoside triphosphates.
- The extra phosphate groups carry energy which is used for formation of covalent bonds.
2.7. U3 DNA polymerase links nucleotides together to form a new strand, using the pre-existing strand as a template.

DNA replication moves in a 5′ to 3′ direction
- This means the 5′ end of the new strand

Free nucleotides in the nucleus (deoxynucleoside triphosphates)

• DNA polymerase always moves in a 5′ to 3′ direction
• DNA polymerase catalyses the covalent phosphodiester bonds between sugars and phosphate groups
2.7.U1 The replication of DNA is semi-conservative and depends on complementary base pairing.

1. Each of the nitrogenous bases can only pair with its partner (A=T and G=C) this is called **complementary base pairing**.

2. The two new strands formed will be identical to the original strand.
2.7.U1 The replication of DNA is semi-conservative and depends on complementary base pairing.

3. Each new strand contains one original and one new strand, therefore DNA Replication is said to be a **Semi-Conservative** Process.
**Polymerase Chain Reaction (PCR)**

- Typically used to copy a segment of DNA – not a whole genome
- Used to amplify small samples of DNA
- In order to use them for DNA profiling, recombination, species identification or other research.
- The process needs a thermal cycler, primers, free DNA nucleotides and DNA polymerase.

Learn the detail using the virtual lab and/or the animation:

http://learn.genetics.utah.edu/content/labs/pcr/

http://www.slideshare.net/gurustip/genetic-engineering-and-biotechnology-presentation
Review: 2.7.A1 Use of Taq DNA polymerase to produce multiple copies of DNA rapidly by the polymerase chain reaction (PCR).

After clicking on the myDNA link choose techniques and then amplifying to access the tutorials on the polymerase chain reaction (PCR).

Alternatively watch the McGraw-Hill tutorial.

http://www.dnai.org/b/index.html

http://highered.mcgraw-hill.com/olc/dl/120078/micro15.swf
To summarise:

PCR is a way of producing large quantities of a specific target sequence of DNA. It is useful when only a small amount of DNA is available for testing e.g. crime scene samples of blood, semen, tissue, hair, etc.

PCR occurs in a thermal cycler and involves a repeat procedure of 3 steps:
1. **Denaturation**: DNA sample is heated to separate it into two strands
2. **Annealing**: DNA primers attach to opposite ends of the target sequence
3. **Elongation**: A heat-tolerant DNA polymerase (Taq) copies the strands

- One cycle of PCR yields two identical copies of the DNA sequence
- A standard reaction of 30 cycles would yield 1,073,741,826 copies of DNA ($2^{30}$)
2.7. S2 Analysis of Meselson and Stahl’s results to obtain support for the theory of semi-conservative replication of DNA.

Before Meselson and Stahl’s work there were different proposed models for DNA replication. After their work only semi-conservative replication was found to be biologically significant.
2.7.2 Analysis of Meselson and Stahl's results to obtain support for the theory of semi-conservative replication of DNA.

**Question** Which model of DNA replication - conservative, dispersive, or semi-conservative - applies to *E. coli*?

**Experiment**

1. **15N medium**
   - Transfer to 14N medium and replicate
   - Spin

2. **Label 1**
   - After one round of replication, the DNA appeared as a single band intermediate between that expected for DNA with 15N and that expected for DNA with 14N.

3. **14N medium**
   - Replication in 14N medium
   - Spin

4. **Label 2**
   - After a second round of replication, DNA appeared as two bands, one in the position of hybrid DNA (half 15N and half 14N) and the other in the position of DNA that contained only 14N.

5. **14N medium**
   - Replication in 14N medium
   - Spin

6. **Label 3**
   - Samples taken after additional rounds of replication appeared as two bands, as in step 3.

**Results**

- **Light (14N)**
- **Heavy (15N)**

**Conclusion** DNA replication in *E. coli* is semi-conservative.

Learn about Meselson and Stahl’s work with DNA to discover the mechanism of semi-conservative replication:


Analysis of Meselson and Stahl’s results to obtain support for the theory of semi-conservative replication of DNA.

At the start of a Meselson and Stahl experiment (generation 0) a single band of DNA with a density of 1.730 g cm\(^{-3}\) was found. After 4 generations two bands were found, but the main band had a density of 1.700 g cm\(^{-3}\).

a. Explain why the density of the main band changed over four generations. (2)

b. After one generation one still only one DNA band appears, but the density has changed.
   i. Estimate the density of the band. (1)
   ii. Which (if any) mechanisms of DNA replication are falsified by this result? (1)
   iii. Explain why the identified mechanism(s) are falsified. (1)

c. Describe the results after two generations and which mechanisms and explain the identified mechanism(s) (if any) are falsified as a consequence. (3)

d. Describe and explain the result found by centrifuging a mixture of DNA from generation 0 and 2. (2)
2.7.S2 Analysis of Meselson and Stahl’s results to obtain support for the theory of semi-conservative replication of DNA.

At the start of a Meselson and Stahl experiment (generation 0) a single band of DNA with a density of 1.730 g cm\(^{-3}\) was found. After 4 generations two bands were found, but the main band had a density of 1.700 g cm\(^{-3}\).

a. Explain why the density of the main band changed over four generations. (2)

- \(N_{15}\) isotope has a greater mass than \(N_{14}\) isotope due to the extra neutron
- Generation 0 contained DNA with exclusively \(N_{15}\) isotopes (giving it a greater density)
- With each generation the proportion \(N_{14}\) isotope (from free nucleotides) increases as the mass of DNA doubles
- After four generations most strands contain only \(N_{14}\) isotope – the dominant band at a density of 1.700 g cm\(^{-3}\).
- \(N_{15}\) isotope remains, but is combined in strands with \(N_{14}\) isotope – a second band at a density between 1.730 and 1.700 g cm\(^{-3}\).
2.7.2 Analysis of Meselson and Stahl’s results to obtain support for the theory of semi-conservative replication of DNA.

At the start of a Meselson and Stahl experiment (generation 0) a single band of DNA with a density of 1.730 g cm\(^{-3}\) was found. After 4 generations two bands were found, but the main band had a density of 1.700 g cm\(^{-3}\).

b. After one generation only one DNA band appeared, but the density had changed.
   i. Estimate the density of the band. (1)
      • The band would contain equally amounts of N\(_{14}\) isotope and N\(_{15}\) isotope
      • Density of an all N\(_{15}\) isotope band is 1.730 g cm\(^{-3}\).
      • Density of an all N\(_{14}\) isotope band is 1.700 g cm\(^{-3}\).
      • Density of an the mixed isotope band is the average of the two:
        \[
        \frac{1.730 \text{ g cm}^{-3} + 1.700 \text{ g cm}^{-3}}{2} = 1.715 \text{ g cm}^{-3}
        \]
   ii. Which (if any) mechanisms of DNA replication are falsified by this result? (1)
       • conservative replication
   iii. Explain why the identified mechanism(s) are falsified. (1)
       • For conservative replication to be the case two bands should appear in all generations after generation 0
2.7.S2 Analysis of Meselson and Stahl’s results to obtain support for the theory of semi-conservative replication of DNA.

At the start of a Meselson and Stahl experiment (generation 0) a single band of DNA with a density of 1.730 g cm\(^{-3}\) was found. After 4 generations two bands were found, but the main band had a density of 1.700 g cm\(^{-3}\).

c. Describe the results after two generations and which mechanisms and explain the identified mechanism(s) (if any) are falsified as a consequence. (3)

- **2 bands:**
- **One band containing a mixture of \(N_{15}\) and \(N_{14}\) isotopes** – semi-conservative replication preserves the DNA strands containing \(N_{15}\) isotopes, but combines them with \(N_{14}\) nucleotides during replication.
- **One band containing all \(N_{14}\) isotopes** - during replication from generation 1 to generation 2. The new strands consisting of \(N_{14}\) isotopes are replicated using \(N_{14}\) nucleotides creating strands containing just \(N_{14}\) isotopes.
- **Dispersive replication is falsified** as this model would continue to produce a single band, containing proportionally less \(N_{15}\) isotope.
2.7.S2 Analysis of Meselson and Stahl’s results to obtain support for the theory of semi-conservative replication of DNA.

At the start of a Meselson and Stahl experiment (generation 0) a single band of DNA with a density of 1.730 g cm-3 was found. After 4 generations two bands were found, but the main band had a density of 1.700 g cm-3.

d. Describe and explain the result found by centrifuging a mixture of DNA from generation 0 and 2. (2)

- **3 bands:**
  - **One band from generation 0 containing all N\textsubscript{15} isotopes** – no replication has occurred.
  - **One band from generation 2 containing a mixture of N\textsubscript{15} and N\textsubscript{14} isotopes** – semi-conservative replication preserves the DNA strands containing N\textsubscript{15} isotopes, but combines them with N\textsubscript{14} nucleotides during replication.
  - **One band from generation 2 (all replicated DNA) containing all N\textsubscript{14} isotopes** - during replication from generation 1 to generation 2. The new strands consisting of N\textsubscript{14} isotopes are replicated using N\textsubscript{14} nucleotides creating strands containing just N\textsubscript{14} isotopes.
2.7.U4 Transcription is the synthesis of mRNA copied from the DNA base sequences by RNA polymerase.
2.7.U5 Translation is the synthesis of polypeptides on ribosomes.

Q - What is the purpose of transcription and translation?

A - These processes work together to create a polypeptide which in turns folds to become a protein. Proteins carry many essential functions in cells. For more detail review 2.4.U7 Living organisms synthesize many different proteins with a wide range of functions.

Catalysis  | Tensile strengthening  | Transport of nutrients and gases  | Cell adhesion
---|---|---|---
Muscle contraction  |  |  |  
Cytoskeletons  |  |  |  
Blood clotting  |  |  |  
Membrane transport  |  |  |  

Use the learn.genetics tutorial to discover one example:

http://learn.genetics.utah.edu/content/molecules/firefly/
2.7.U4 Transcription is the synthesis of mRNA copied from the DNA base sequences by RNA polymerase.

2.7.U5 Translation is the synthesis of polypeptides on ribosomes.

**TRANSCRIPTION**: In the nucleus, the cell's machinery copies the gene sequence into messenger RNA (mRNA), a molecule that is similar to DNA. Like DNA, mRNA has four nucleotide bases - but in mRNA, the base uracil (U) replaces thymine (T).

**TRANSLATION**: The protein-making machinery, called the ribosome, reads the mRNA sequence and translates it into the amino acid sequence of the protein. The ribosome starts at the sequence AUG, then reads three nucleotides at a time. Each three-nucleotide codon specifies a particular amino acid. The "stop" codons (UAA, UAG and UGA) tell the ribosome that the protein is complete.
Transcription is the synthesis of mRNA copied from the DNA base sequences by RNA polymerase.

Three main types of RNA are predominantly synthesised:

- **Messenger RNA (mRNA):** A transcript copy of a gene used to encode a polypeptide
- Transfer RNA (tRNA): A clover leaf shaped sequence that carries an amino acid
- Ribosomal RNA (rRNA): A primary component of ribosomes

We are focusing on mRNA

http://www.nature.com/scitable/topicpage/Translation-DNA-to-mRNA-to-Protein-393
The enzyme RNA polymerase binds to a site on the DNA at the start of a gene. The sequence of DNA that is transcribed into RNA is called a gene. RNA polymerase separates the DNA strands and synthesises a complementary RNA copy from the antisense DNA strand. It does this by covalently bonding ribonucleoside triphosphates that align opposite their exposed complementary partner (using the energy from the cleavage of the additional phosphate groups to join them together). Once the RNA sequence has been synthesised:
- RNA polymerase will detach from the DNA molecule
- RNA detaches from the DNA
- The double helix reforms
Transcription occurs in the nucleus (where the DNA is) and, once made, the mRNA moves to the cytoplasm (where translation can occur).
Translation is the process of protein synthesis in which the genetic information encoded in mRNA is translated into a sequence of amino acids in a polypeptide chain.

A ribosome is composed of two halves, a large and a small subunit. During translation, ribosomal subunits assemble together like a sandwich on the strand of mRNA:

- Each subunit is composed of RNA molecules and proteins
- The small subunit binds to the mRNA
- The large subunit has binding sites for tRNAs and also catalyzes peptide bonds between amino acids
2.7.U6 The amino acid sequence of polypeptides is determined by mRNA according to the genetic code.

Messenger RNA (mRNA): A transcript copy of a gene used to encode a polypeptide

- The length of mRNA molecules varies - the average length for mammals is approximately 2,200 nucleotides (this translates to approximately 730 amino acids in the average polypeptide)
- Only certain genes in a genome need to be expressed depending on:
  - Cell specialism
  - Environment
- Therefore not all genes (are transcribed) and translated
- If a cell needs to produce a lot of a certain protein (e.g. β cells in the pancreas specialize in secreting insulin to control blood sugar) then many copies of the required mRNA are created.
2.7.U7 Codons of three bases on mRNA correspond to one amino acid in a polypeptide.

The genetic code is the set of rules by which information encoded in mRNA sequences is converted into proteins (amino acid sequences) by living cells.

- **Codons are a triplet of bases** which encodes a particular amino acid.
- As there are four bases, there are 64 different codon combinations ($4 \times 4 \times 4 = 64$).
- The codons can translate for **20 amino acids**.
- Different codons can translate for the same amino acid (e.g. GAU and GAC both translate for Aspartate) therefore the genetic code is said to be **degenerate**.
- The order of the codons determines the amino acid sequence for a protein.
- The **coding region always starts with a START codon** (AUG) therefore the first amino acid in all polypeptides is Methionine.
- The **coding region of mRNA terminates with a STOP codon** - the STOP codon does not add an amino acid – instead it causes the release of the polypeptide.

Amino acids are carried by **transfer RNA (tRNA)**. The anti-codons on tRNA are complementary to the codons on mRNA.
2.7. U8 Translation depends on complementary base pairing between codons on mRNA and anticodons on tRNA.

**Key components** of translation that enable genetic code to synthesize polypeptides

- **tRNA molecules** have an anticodon of three bases that binds to a complementary codon on mRNA.
- **tRNA molecules** carry the amino acid corresponding to their codon.
- **mRNA** has a sequence of codons that specifies the amino acid sequence of the polypeptide.
- **Ribosomes:**
  - act as the binding site for mRNA and tRNA
  - catalyse the peptide bonds of the polypeptide

![Peptide synthesis diagram](https://upload.wikimedia.org/wikipedia/commons/0/0f/Peptide_syn.png)
2.7. U8 Translation depends on complementary base pairing between codons on mRNA and anticodons on tRNA.

An **outline of translation** and polypeptide synthesis

- **The large subunit binds to the small subunit of the ribosome.**
- **There are three binding sites on the large subunit of the ribosome, but only two can contain tRNA molecules at a time.**

The mRNA contains a series of codons (3 bases) each of which codes for an amino acid.

mRNA binds to the small subunit of the ribosome.

tRNA molecules contain anticodons which are complementary to the codons on the mRNA. tRNA molecules bind to a specific amino acid that corresponds to the anticodon.
Translation depends on complementary base pairing between codons on mRNA and anticodons on tRNA.

An **outline of translation** and polypeptide synthesis

A peptide bond is formed between the two amino acids (carried by the tRNAs)

- tRNAs with anticodons complementary to the codons bind (the bases are linked by the formation of hydrogen bonds)

The ribosome moves along the mRNA and presents codons in the first two binding sites

https://upload.wikimedia.org/wikipedia/commons/0/0f/Peptide_syn.png
Translation depends on complementary base pairing between codons on mRNA and anticodons on tRNA.

**An outline of translation** and polypeptide synthesis

The process (i.e. the last two steps) repeats forming a polypeptide.

Another tRNA carrying an amino acid binds to the first site and a second peptide bond is formed.

As the ribosome moves along mRNA a tRNA moves to the third binding site and detaches.
2.7.S1 Use a table of the genetic code to deduce which codon(s) corresponds to which amino acid.
2.7.S3 Use a table of mRNA codons and their corresponding amino acids to deduce the sequence of amino acids coded by a short mRNA strand of known base sequence.
2.7.S4 Deducing the DNA base sequence for the mRNA strand.

The diagram summarizes the process of protein synthesis. You should be able to use a section of genetic code, transcribe and translate it to deduce the polypeptide synthesized.
Practice transcribing and translating using the learn.genetics tutorial.

http://learn.genetics.utah.edu/content/molecules/transcribe/
n.b. You just have to be able to use the table. You do not have to memorize which codon translates to which amino acid.
1. Deduce the codon(s) that translate for Aspartate.

2. If mRNA contains the base sequence CUGACUAGGUCCGGA
   a. deduce the amino acid sequence of the polypeptide translated.

   b. deduce the base sequence of the DNA antisense strand from which the mRNA was transcribed.

3. If mRNA contains the base sequence ACUAUAC deduce the base sequence of the DNA sense strand.
1. Deduce the codon(s) that translate for Aspartate.

2. If mRNA contains the base sequence CUGACUAGGUCCGGA
   a. deduce the amino acid sequence of the polypeptide translated.

   b. deduce the base sequence of the DNA antisense strand from which the mRNA was transcribed.

3. If mRNA contains the base sequence ACUAAC deduce the base sequence of the DNA sense strand.
1. Deduce the codon(s) that translate for Aspartate.

   GAU, GAC

2. If mRNA contains the base sequence CUGACUAGGUCCGGA
   a. deduce the amino acid sequence of the polypeptide translated.
      Leucine + Threonine + Lysine + Arginine + Serine + Glycine

   b. deduce the base sequence of the DNA antisense strand from which the mRNA was transcribed.
      GACTGATCCAGGCCT (the antisense strand is complementary to the mRNA, but remember that uracil is replaced by thymine)

3. If mRNA contains the base sequence ACUAAC deduce the base sequence of the DNA sense strand.
   ACTAAC (the sense strand is the template for the mRNA the only change is that uracil is replaced by thymine)
2.7.S1 Use a table of the genetic code to deduce which codon(s) corresponds to which amino acid.

Transcribe this DNA strand into mRNA:

DNA: ACGTTACGGGATTACAGTCCCCAAACTAC

mRNA:
Transcribe this DNA strand into mRNA:

DNA: ACGTTACGGATTACAGTCCCCAAACTAC

mRNA: UGCAAUUGCCUAUUGUCAGGGGUUUGAUG

Don't forget: on mRNA, Uracil takes the place of Thymine

Uracil is complementary to Adenine
2.7. S1 Use a table of the genetic code to deduce which codon(s) corresponds to which amino acid.

Now translate the mRNA into a polypeptide:

DNA: ACGTTACGGATTACAGTCCCCAAACTAC

mRNA: UGCAAUGCCUAUAUGUCAGGGGUUUGAUG

Remember: the 'Met' codon is 'Start'

There are 20 amino acids and a Stop codon:

Phe, Leu, Ile, Met, Val, Ser, Pro, Thr, Ala, Tyr

His, Gln, Asn, Lys, Asp, Glu, Cys, Trp, Arg, Gly, Stop
2.7.S1, 2.7.S3, 2.7.S4

How many amino acids in the polypeptide?:

**DNA:**

ACGTACGGATTACAGTCCCAAACCTAC

**mRNA:**

UGCAAUUGCCUAAUUGUCAGGGGUUUGAUAG

**Met**  **Pro**  **Asn**  **Val**  **Arg**  **Val**  **Stop**

The 'Met' codon is always the first.

There are 20 amino acids and a Stop codon:

Phe  Leu  Ile  Met  Val  Ser  Pro  Thr  Ala  Tyr

His  Gln  Asn  Lys  Asp  Glu  Cys  Trp  Arg  Gly  Stop
2.7.S1 Use a table of the genetic code to deduce which codon(s) corresponds to which amino acid.

Things to remember about the genetic code:

DNA: ACGTTACGGA TTACAGTGCCCCAAACTAC

- this gene = 21 base pairs
- giving 7 codons

mRNA: UGCAAUGCCUAAUGUCAGGGGUUGAUG

- Met Pro Asn Val Arg Val
- 6 amino acids in the polypeptide
- plus stop

The 'Met' codon is always the first.

Number of amino acids = codons - 1
- or = (base pairs/3) - 1

There are 20 amino acids and a Stop codon:

- Phe Leu Ile Met Val Ser Pro Thr Ala Tyr
- His Gln Asn Lys Asp Glu Cys Trp Arg Gly Stop
Diabetes in some individuals is due to destruction of cells in the pancreas that secrete the hormone insulin. It can be treated by injecting insulin into the blood. Porcine and bovine insulin, extracted from the pancreases of pigs and cattle, have both been widely used. Porcine insulin has only one difference in amino acid sequence from human insulin and bovine insulin has three differences. Shark insulin, which has been used for treating diabetics in Japan, has seventeen differences.

Despite the differences in the amino acid sequence between animal and human insulin, they all bind to the human insulin receptor and cause lowering of blood glucose concentration. However, some diabetics develop an allergy to animal insulins, so it is preferable to use human insulin. In 1982 human insulin became commercially available for the first time. It was produced using genetically modified E. coli bacteria. Since then methods of production have been developed using yeast cells and more recently safflower plants.
2.7.A2 Production of human insulin in bacteria as an example of the universality of the genetic code allowing gene transfer between species.

- All living things use the **same bases** and the **same genetic code**.
- Each **codon** produces the **same amino acid** in transcription and translation, regardless of the species.
- So the sequence of amino acids in a **polypeptide remains unchanged**.
- Therefore, we can take genes from one species and insert them into the genome of another species.

---

**“The Genetic Code is Universal”**
Restriction enzymes ‘cut’ the desired gene from the genome.

E. coli bacteria contain small circles of DNA called plasmids.

These can be removed.

The same restriction enzyme cuts into the plasmid.

Because it is the same restriction enzyme the same bases are left exposed, creating ‘sticky ends’

Ligase joins the sticky ends, fixing the gene into the E. coli plasmid.

The recombinant plasmid is inserted into the host cell. It now expresses the new gene. An example of this is human insulin production.
7.1 DNA Structure and Replication

Essential idea: The structure of DNA is ideally suited to its function.

From reactive bases that bond easily with their complements to the nucleoside tri-phosphates that can with them the energy to bond and build a strand of DNA. The structure of DNA is ideally suited to replicating itself and storing information.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Guidance</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1.U1  Nucleosomes help to supercoil the DNA.</td>
<td></td>
</tr>
<tr>
<td>7.1.U2  DNA structure suggested a mechanism for DNA replication.</td>
<td></td>
</tr>
<tr>
<td>7.1.U3  DNA polymerases can only add nucleotides to the 3’ end of a primer.</td>
<td></td>
</tr>
<tr>
<td>7.1.U4  DNA replication is continuous on the leading strand and discontinuous on the lagging strand.</td>
<td>Details of DNA replication differ between prokaryotes and eukaryotes. Only the prokaryotic system is expected.</td>
</tr>
<tr>
<td>7.1.U5  DNA replication is carried out by a complex system of enzymes.</td>
<td>The proteins and enzymes involved in DNA replication should include helicase, DNA gyrase, single strand binding proteins, DNA primase and DNA polymerases I and III.</td>
</tr>
<tr>
<td>7.1.U6  Some regions of DNA do not code for proteins but have other important functions.</td>
<td>The regions of DNA that do not code for proteins should be limited to regulators of gene expression, introns, telomeres and genes for tRNAs.</td>
</tr>
<tr>
<td>7.1.A1  Rosalind Franklin’s and Maurice Wilkins’ investigation of DNA structure by X-ray diffraction.</td>
<td></td>
</tr>
<tr>
<td>7.1.A2  Use of nucleotides containing dideoxyribonucleic acid to stop DNA replication in preparation of samples for base sequencing.</td>
<td></td>
</tr>
<tr>
<td>7.1.A3  Tandem repeats are used in DNA profiling.</td>
<td></td>
</tr>
<tr>
<td>7.1.S1  Analysis of results of the Hershey and Chase experiment providing evidence that DNA is the genetic material.</td>
<td></td>
</tr>
<tr>
<td>7.1.S2  Utilization of molecular visualization software to analyse the association between protein and DNA within a nucleosome.</td>
<td></td>
</tr>
</tbody>
</table>
In the mid-twentieth century scientists were unsure as to whether proteins or chromosomes were the genetic material of cells.

Alfred Hershey and Martha Chase wanted to solve this problem by finding out if protein or DNA was the genetic material of viruses.

Viruses infect cells and transform them into virus-producing factories:
- Viruses inject their genetic material into cells.
- The non-genetic part of the virus remains outside the cell.
- Infected cells produce large numbers of the virus.
- The cell bursts releasing the copied virus.

Hershey and Chase chose to study the **T2 bacteriophage**, which infects the **E. Coli** bacterium (image of both above), because of its very simple structure consisting of just:
- **Protein** coat (capsid)
- **DNA** inside the coat
7.1. S1 Analysis of results of the Hershey and Chase experiment providing evidence that DNA is the genetic material.

Use one or more of the animations to learn about the Hersey-Chase experiment:

Watch and take notes using at least one animation. You will be expected to be able to analyse results from similar experiments.

http://nortonbooks.com/college/biology/animations/ch12a02.htm


https://smartsite.ucdavis.edu/access/content/user/00002950/bis10v/media/ch09/bacteriophage_studies.html

http://nortonbooks.com/college/biology/animations/ch12a02.htm
Amino acids containing **Radioactive isotopes** were used to label the virus:

- **Sulfur** ($^{35}$S) for the **Protein** coat (capsid)
- **Phosphorus** ($^{32}$P) for the **DNA**

The experiments combined T2 bacteriophage with *E. Coli* bacteria. At the end of the experiment a **centrifuge** was used to separate them:

- the smaller **virus** remained in the **supernatant** (liquid)
- the **bacteria** formed a **pellet**

Separate experiments with the two isotopes found that:

- **Sulfur** ($^{35}$S) remained in **supernatant**
- **Phosphorus** ($^{32}$P) was found in the **pellet**

Hershey and Chase **deduced** that **DNA** therefore was the **genetic material** used by viruses because **DNA** (labelled by $^{32}$P) was being **transferred** into the **bacteria**.
Nature of Science: Making careful observations—Rosalind Franklin’s X-ray diffraction provided crucial evidence that DNA is a double helix. (1.8)

**How was DNA discovered?**

Is the story of its discovery an example of:
- Cooperation?
- Internationalism?
- Sexism?
- Competition?

Discuss how important was Rosalind Franklin’s careful observation and interpretation of the photographic evidence was to Crick’s and Watson’s successful discovery of the structure of DNA?

[https://youtu.be/sf0YXnAFBs8](https://youtu.be/sf0YXnAFBs8)

What is x-ray diffraction?

When X-rays are directed at a material some is scattered by the material. This scattering is known as diffraction. For X-ray diffraction to work well the material ideally should be crystallised so that the repeating pattern causes diffraction to occur in a regular way. DNA cannot be crystallised but the molecules were arranged regularly enough for the technique to work.

Use the animation to understand how to interpret the Rosalind Franklin’s and Maurice Wilkins’ X-ray diffraction photographs of DNA.
Core Review: 2.6.U3 DNA is a double helix made of two antiparallel strands of nucleotides linked by hydrogen bonding between complementary base pairs.
Core Review: 2.6.U3 DNA is a double helix made of two antiparallel strands of nucleotides linked by hydrogen bonding between complementary base pairs.
DNA replication and mechanisms by which it can happen are **implied by complementary base pairing**. Outline the **evidence** that supports complementary base pairing:

- (X-ray diffraction showed that) the DNA helix is both tightly packed and regular* therefore pyrimidines need to be paired with purines.
- The electrical charges of adenine and thymine are compatible (and opposite) allowing two hydrogen bonds to form between them.
- The pairing of cytosine with guanine allows for three hydrogen bonds to form between them.

The logical **deduction** is that if an adenine (A) base occurs on one strand then opposite it the only possible base is thymine (T), vice-versa is true too. It also follows that if a cytosine (C) base occurs on one strand then opposite it the only possible base is guanine (G), vice-versa is true too.

* This evidence also showed that opposite bases need to be “upside down” in relation each other - the helix is anti-parallel however this is not important for complementary base pairing.
7.1.U1 Nucleosomes help to supercoil the DNA.

Eukaryotic DNA supercoiling is organised by nucleosomes

- Nucleosomes both protect DNA and allow it to be packaged, this in turn allows DNA to be supercoiled.
- Nucleosomes are formed by wrapping DNA around histone proteins.

n.b. Prokaryotic DNA is, like eukaryotic DNA, supercoiled, but differently: Prokaryotic DNA may be associated with proteins, but it is not organised by histones and is therefore sometimes referred as being ‘naked’.

7.1. U1 Nucleosomes help to supercoil the DNA.

Structure of a simplified nucleosome

The H1 histone binds DNA in such a way to form a structure called the 30 nm fibre (solenoid) that facilitates further packing.

[Image: http://commons.wikimedia.org/wiki/File:Nucleosome_organization.png]
Why does Eukaryotic DNA need to be supercoiled?

*Supercoiling is when a DNA strand has been wound back on itself multiple times to so that the molecule becomes compacted.*

The facts:

- The length of DNA in a human (eukaryotic) cell is approx. 2 m
- Eukaryotic chromosomes are 15 – 85 mm in length
- The nucleus in a eukaryotic cell has a diameter of approximately 10 μm
7.1.U1 Nucleosomes help to supercoil the DNA.

Why does Eukaryotic DNA need to be supercoiled?

- **essential** to **pack genetic material** into the nucleus
- to **organise** DNA to allow **cell division** to occur (most DNA supercoiling occurs at this time)
- to **control DNA expression** - supercoiled DNA cannot be transcribed
- allow **cells** to **specialise** by permanently supercoiling DNA (**heterochromatin**)
- **transcription** of active chromatin (**Euchromatin**) can be **promoted** or **inhibited** by the associated **histones**
7.1.S2 Utilization of molecular visualization software to analyse the association between protein and DNA within a nucleosome.

Use the **RCSB Protein Data Bank** to find out more about nucleosomes:

- Jmol visualisation: [http://www.rcsb.org/pdb/explore/jmol.do?structureId=1AOI&bionumber=1](http://www.rcsb.org/pdb/explore/jmol.do?structureId=1AOI&bionumber=1)

Use the Jmol visualisation to:

1. Identify the two copies of each histone protein. This can be done by locating the tail* of each protein.
2. Suggest how the positive charges help to form the nucleosome (with the negatively charged DNA molecule).

*Tails of the histone proteins are involved in regulating gene expression.*
7.1. U3 DNA polymerases can only add nucleotides to the 3’ end of a primer.

DNA Replication is initiated at many points in eukaryotes:

This makes DNA replication faster and more efficient.*

These points are known as origins of initiation and will have the same (or very similar) base-sequence.

Proteins called Origin Recognition Complexes will bind here and then DNA Helicase will be able to attach, to begin replication.

*In this topic prokaryote DNA replication is examined; prokaryotes have a single replication fork and a simpler mechanism of replication.

More information:
Core Review: 2.7.U2 Helicase unwinds the double helix and separates the two strands by breaking hydrogen bonds.

- Unwinds the DNA Helix
- Separates the two polynucleotide strands by breaking the hydrogen bonds between complementary base pairs
- ATP is needed by helicase to both move along the DNA molecule and to break the hydrogen bonds
- The two separated strands become parent-template strands for the replication process
Core Review: 2.7.U3 DNA polymerase links nucleotides together to form a new strand, using the pre-existing strand as a template.

- Free nucleotides are deoxynucleoside triphosphates
- The extra phosphate groups carry energy which is used for formation of covalent bonds

n.b. For AHL you need to distinguish between DNA polymerase III, as shown here, and DNA polymerase I, which is dealt with later.

DNA polymerase always moves in a 5’ to 3’ direction
DNA polymerase catalyses the covalent phosphodiester bonds between sugars and phosphate groups
DNA Polymerase proof reads the complementary base pairing. Consequently mistakes are very infrequent occurring approx. once in every billion bases pairs
7.1 U3 DNA polymerases can only add nucleotides to the 3’ end of a primer.

**RNA primers** provide an attachment and initiation point for DNA polymerase III.

RNA primers consist of a short sequence (generally about 10 base pairs) of **RNA nucleotides**.

[Diagram of DNA replication with RNA primers highlighted]
Core Review: 2.7.U3 DNA polymerase links nucleotides together to form a new strand, using the pre-existing strand as a template.

DNA replication moves in a 5' to 3' direction
- This means the 5' end of the new strand

Free nucleotides in the nucleus (deoxynucleoside triphosphates)

- DNA polymerase always moves in a 5' to 3' direction
- DNA polymerase catalyses the covalent phosphodiester bonds between sugars and phosphate groups
DNA polymerases can only add nucleotides to the 3' end of a primer. DNA polymerase III adds new nucleotides to the C3 hydroxyl group on the ribose/deoxyribose sugar such that the strand grows from the 3' end. RNA primers provide an attachment and initiation point for DNA polymerase III.

RNA primers consists of a short sequence (generally about 10 base pairs) of RNA nucleotides.

Therefore DNA polymerase III moves along the new strand in a 5' - 3' direction (3' - 5' direction on the template strand)

Core Review: 2.7.U1 The replication of DNA is semi-conservative and depends on complementary base pairing.

1. Each of the nitrogenous bases can only pair with its partner (A=T and G=C) this is called complementary base pairing.

2. The two new strands formed will be identical to the original strand.
Core Review: 2.7.U1 The replication of DNA is semi-conservative and depends on complementary base pairing.

3. Each new strand contains one original and one new strand, therefore DNA Replication is said to be a Semi-Conservative Process.
7.1. U5 DNA replication is carried out by a complex system of enzymes.

**Summary of the enzymes involved in DNA Replication**

DNA Gyrase *(aka* topoisomerase*) moves in advance of helicase and relieves strain and prevents supercoiling on the separated strands.

**DNA Helicase** unwinds and separates the double stranded DNA by breaking the hydrogen bonds between base pairs.

http://commons.wikimedia.org/wiki/File:DNA_replication_en.svg
7.1.U5 DNA replication is carried out by a complex system of enzymes.

**Summary of the enzymes involved in DNA Replication**

**DNA Ligase** joins the Okazaki fragments together to create a continuous strand.

**DNA Polymerase III** adds deoxynucleoside triphosphates (dNTPs) to the 3' end of the polynucleotide chain, synthesising in a 5' - 3' direction.

**RNA Primase** synthesises a short RNA primer on each template strand to provide an attachment and initiation point for DNA polymerase III.

**DNA Polymerase I** removes the RNA primers and replaces them with DNA.

7.1. U4 DNA replication is continuous on the leading strand and discontinuous on the lagging strand.

**Detailed summary of DNA Replication – part I**

- DNA replication occurs during (S phase of interphase, in preparation for cell division
- Helicase unwinds the double helix separating the strands of DNA
- It breaks the hydrogen bonds between the two strands
- Single stranded binding proteins keep the separated strands apart so that nucleotides can bind
- DNA gyrase moves in advance of helicase and relieves strain and prevents the DNA supercoiling again.
- Each strand of parent DNA is used as template for the synthesis of the new strands
- Synthesis always occurs in $5' \rightarrow 3'$ direction on each new strand
- Therefore synthesis is continuous on leading strand (in the same direction as helicase) and dis-continuous on lagging strand (away from helicase)
- This leads to the formation of Okazaki fragments on the lagging strand
7.1. DNA replication is continuous on the leading strand and discontinuous on the lagging strand.

**Detailed summary of DNA Replication – part II**

- To synthesise a new strand first an RNA primer is synthesized on the parent DNA using RNA primase.

- Next DNA polymerase III adds the nucleotides (to the 3’ end) added according to the complementary base pairing rules; adenine pairs with thymine and cytosine pairs with guanine; (names needed, letters alone not accepted).

- Nucleotides added are in the form of deoxynucleoside triphosphate. Two phosphate groups are released from each nucleotide and the energy is used to join the nucleotides into a growing DNA chain.

- DNA polymerase I then removes the RNA primers and replaces them with DNA.

- DNA ligase next joins Okazaki fragments on the lagging strand.

- Because each new DNA molecule contains both a parent and newly synthesised strand DNA replication is said to be semi-conservative.

[Image of DNA replication process]
Some regions of DNA do not code for proteins but have other important functions. The percentage varies greatly between organisms, but in all organisms there are regions of DNA that are not expressed as polypeptides. This non-coding DNA is still important to organisms for a variety of reasons.

data source: [http://www.nature.com/nrg/journal/v11/n8/fig_tab/nrg2814_T2.html](http://www.nature.com/nrg/journal/v11/n8/fig_tab/nrg2814_T2.html)
Some regions of DNA do not code for proteins but have other important functions.

**Genes**, the regions of DNA that code for polypeptides, contain both **intron** and **exon** DNA.

- **Introns** are ‘edited’ out of messenger RNA (mRNA)
- mRNA is translated by ribosomes into polypeptides
- Therefore **only exons code** for the **polypeptides**
Some regions of DNA do not code for proteins but have other important functions. Between genes exist non-coding regions of DNA. Although such DNA does not code for polypeptides it can affect transcription of mRNA.

**Promoters** sequences are attachment points for RNA polymerase adjacent to the gene.

Some of these regions act as binding sites for particular proteins, which in turn affect transcription of the nearby gene:

- **Enhancers** are sequences that increase the rate of transcription (when a protein is bound to it)
- **Silencers** inhibit transcription (when a protein is bound to it)
Some regions of DNA do not code for proteins but have other important functions.

The end of chromosomes contain highly repetitive DNA sequences. These regions are called Telomeres and they protect the DNA molecule from degradation during replication.

n.b. the electron micrograph image is falsely coloured to indicate the telomere regions.

http://med.stanford.edu/content/dam/sm-news/images/2015/01/telomeres.jpg
7.1.A2 Use of nucleotides containing dideoxyribonucleic acid to stop DNA replication in preparation of samples for base sequencing.

Use the video short introduction to understand the principles of Sanger’s technique for **DNA base sequencing** using dideoxyribonucleic acid.

7.1.A2 Use of nucleotides containing dideoxyribonucleic acid to stop DNA replication in preparation of samples for base sequencing.

Dideoxyribonucleic acid stops DNA replication when it is added to a new DNA strand.

Fluorescent dye markers are attached to dideoxyribonucleic acids so that the base present when replication stops can be identified. From this the base on the parent strand deduced.

DNA replication is carried out with dideoxyribonucleic acid mixed in with normal deoxyribonucleic acid.

A range of new strands of differing lengths are produced.

The length of strand and the terminal base are identified by sequencing machines.

http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/F/FluorDideoxySeq.gif
Tandem repeat sequences are short sequences of (non-coding) DNA, normally of length 2-5 base pairs, that are repeated numerous times in a head-tail manner.

- The same repeating sequence (GATA) is found in all participants.
- The number of times the sequence is repeated varies from 8 to 10 times.
7.1.A3 Tandem repeats are used in DNA profiling.

How tandem repeats are used in **DNA profiling**?

- The TRs vary greatly in terms of the different number of copies of the repeat element that can occur in a population.
- For maternal profiling it is usually mitochondrial DNA. For paternal profiling commonly the Y chromosome is used.
- Chromosomes occur in pairs (if not using the Y chromosome) and the tandem repeat on each may vary.

- Dyes markers (e.g. attached to dideoxyribonucleic acids) are attached to the tandem repeats during PCR (DNA Replication).
- Restriction enzymes can be used to cut DNA between the tandem repeats.
- Electrophoresis enables scientists therefore to calculate the length of the tandem repeat sequence of an individuals.
- If different tandem repeats at different loci are used then a unique profile, for an individual can be identified.

7.2 Transcription and gene expression

Essential idea: Information stored as a code in DNA is copied onto mRNA.

"The genetic code is frequently referred to as a blueprint because it contains the instructions a cell requires in order to sustain itself. We now know that there is more to these instructions than simply the sequence of letters in the nucleotide code, however. For example, vast amounts of evidence demonstrate that this code is the basis for the production of various molecules, including RNA and protein ... In transcription, a portion of the double-stranded DNA template gives rise to a single-stranded RNA molecule."

http://www.nature.com/scitable/topicpage/dna-transcription-426#

The image shows how DNA is used as a template to create portable molecules of genetic code, i.e. mRNA, that can leave the nucleus for translation in other regions of the cell.

https://bioknowledgy.weebly.com/

https://commons.wikimedia.org/wiki/File:Simple_transcription_elongation1.svg
### Understandings, Applications and Skills

<table>
<thead>
<tr>
<th>Statement</th>
<th>Guidance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>7.2.U1</strong> Transcription occurs in a 5’ to 3’ direction.</td>
<td>RNA polymerase adds the 5´ end of the free RNA nucleotide to the 3´ end of the growing mRNA molecule.</td>
</tr>
<tr>
<td><strong>7.2.U2</strong> Nucleosomes help to regulate transcription in eukaryotes.</td>
<td></td>
</tr>
<tr>
<td><strong>7.2.U3</strong> Eukaryotic cells modify mRNA after transcription.</td>
<td></td>
</tr>
<tr>
<td><strong>7.2.U4</strong> Splicing of mRNA increases the number of different proteins an organism can produce.</td>
<td></td>
</tr>
<tr>
<td><strong>7.2.U5</strong> Gene expression is regulated by proteins that bind to specific base sequences in DNA.</td>
<td></td>
</tr>
<tr>
<td><strong>7.2.U6</strong> The environment of a cell and of an organism has an impact on gene expression.</td>
<td></td>
</tr>
<tr>
<td><strong>7.2.A1</strong> The promoter as an example of non-coding DNA with a function.</td>
<td></td>
</tr>
<tr>
<td><strong>7.2.S1</strong> Analysis of changes in the DNA methylation patterns.</td>
<td></td>
</tr>
</tbody>
</table>
Between genes exist **non-coding regions of DNA**. Although such DNA does not code for polypeptides it **can affect transcription** of mRNA.

**Promoters** sequences are **attachment points** for **RNA polymerase** adjacent to the gene.

Some other regions act as binding sites for particular proteins, which in turn affect transcription of the nearby gene:

- **Enhancers** are sequences that **increase the rate of transcription** (when a protein is bound to it)
- **Silencers inhibit transcription** (when a protein is bound to it)
Non-coding regions have important functions, for example promoters:

The promoter is a DNA sequence located near a gene. It acts as the binding site for RNA polymerase.

The adjacent gene is transcribed, but the promoter region is not.

Operator is a region of DNA that can regulate transcription, typically inhibiting transcription (silencers are a type of operator).

RNA polymerase transcribes the gene into RNA, typically mRNA.

Edited from: http://commons.wikimedia.org/wiki/File:Lac_Operon.svg
Gene expression is regulated by proteins that bind to specific base sequences in DNA.

One well known example of the regulation of gene expression by proteins is the metabolism of lactose in *E. Coli* bacterium. The diagram below illustrates this example.

The repressor protein is bound to the operator preventing RNA Polymerase from transcription of the genes.

Operator is a region of DNA that can regulate transcription, typically inhibiting transcription, such as this silencer sequence.

The promoter is a DNA sequence located near a gene. It acts as the binding site for RNA polymerase.

Genes involved in the metabolism (breakdown) of lactose.

The consequence of the inhibition of the lactose metabolism is that the concentration of undigested lactose now increases in *E. Coli* ...

Edited from: http://commons.wikimedia.org/wiki/File:Lac_Operon.svg
7.2 U5 Gene expression is regulated by proteins that bind to specific base sequences in DNA.

One well known example of the regulation of gene expression by proteins is the metabolism of lactose in *E. Coli* bacterium. The diagram below illustrates this example.

Lactose binds to the repressor protein inhibiting it: the repressor can no longer bind to the operator.

Lactose molecules build up inside the *E. Coli*

RNA polymerase binds with the promoter, and express the genes (by transcribing them), which in turn synthesizes lactase

With the synthesis of lactase the lactose is broken down, as it’s concentration decreases the inhibition of the repressor molecules will decrease ‘silencing’ the gene again.
Gene expression is regulated by proteins that bind to specific base sequences in DNA. This regulation involves activator and repressor proteins interacting with DNA sequences known as enhancers and silencers, as well as the promoter region.

### Summary of common types of regulating proteins and associated sequences found in eukaryotes

<table>
<thead>
<tr>
<th>DNA Sequence</th>
<th>Binding protein</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enhancers</td>
<td>Activator</td>
<td>Activator proteins bind to enhancer sequences of DNA to greatly increase the rate of transcription of a gene.</td>
</tr>
<tr>
<td>Silencers</td>
<td>Repressor</td>
<td>Repressor proteins bind to non-coding regions of DNA to either block or reduce the transcription of a gene.</td>
</tr>
<tr>
<td>Promoter</td>
<td>RNA Polymerase</td>
<td>A region of DNA located close to a specific gene. Once bound to the sequence RNA polymerase transcribes the gene.</td>
</tr>
</tbody>
</table>
Review: 7.1.U1 Nucleosomes help to supercoil the DNA.

Eukaryotic DNA supercoiling is organised by nucleosomes

- Nucleosomes both protect DNA and allow it to be packaged, this in turn allows DNA to be supercoiled.
- Nucleosomes are formed by wrapping DNA around histone proteins.

n.b. Prokaryotic DNA is, like eukaryotic DNA, supercoiled, but differently: Prokaryotic DNA maybe associated with proteins, but it is not organised by histones and is therefore sometimes referred as being ‘naked’.

Review: 7.1.U1 Nucleosomes help to supercoil the DNA.

Structure of a simplified nucleosome

The H1 histone binds DNA in such a way to form a structure called the 30 nm fibre (solenoid) that facilitates further packing.

http://commons.wikimedia.org/wiki/File:Nucleosome_organization.png
7.2. U2 Nucleosomes help to regulate transcription in eukaryotes.

**Methylation** is the addition of methyl groups to **DNA**

**Acetylation** is the addition of Acetyl groups to **histones**

Edited from: [http://www.nature.com/neuro/journal/v13/n4/images/nn0410-405-F1.jpg](http://www.nature.com/neuro/journal/v13/n4/images/nn0410-405-F1.jpg)
Methylation is the addition of methyl groups to DNA

Methylation of DNA inhibits transcription

Processes that inhibit transcription bind the DNA more tightly to the histone making it less accessible to transcription factors (forming heterochromatin).

*Chromatin is a complex of DNA, protein and RNA. Tightly packed chromatin which cannot be transcribed is referred to as heterochromatin.
7.2. U2 Nucleosomes help to regulate transcription in eukaryotes.

**Acetylation** is the addition of Acetyl groups to **histones**

Acetylation promotes transcription

Processes that **promote transcription** bind the DNA more loosely to the **histone** making it **more accessible** to transcription factors (forming **euchromatin**).

![Diagram showing DNA and histones with acetylation and methylation marks](http://www.nature.com/neuro/journal/v13/n4/images/nn0410-405-F1.jpg)

**n.b.** **Methylation** of histones can also occur, this process **can both promote and inhibit** transcription.

*Chromatin is a complex of DNA, protein and RNA. Loosely packed chromatin which can be transcribed is referred to as euchromatin.*

Edited from: [http://www.nature.com/neuro/journal/v13/n4/images/nn0410-405-F1.jpg](http://www.nature.com/neuro/journal/v13/n4/images/nn0410-405-F1.jpg)
The environment of an organism impacts gene expression. For example human hair and skin colour are impacted by the exposure to sunlight and high temperatures. Similarly pigments in the fur of Himalayan rabbits (*Oryctolagus cuniculus*) are regulated by temperature.

**Gene C** controls fur pigmentation in Himalayan rabbits. The gene is **active** when environmental temperatures are **between 15 and 25°C**. At higher temperatures the gene is inactive.

In **low temperatures** Gene C becomes active in the rabbit's **colder extremities** (ears, nose, and feet) and produces a **black** pigment.

In the warm weather no pigment is produced and the fur is **white**.
Only a small number of genes are involved in determining body patterns during embryonic development.

Morphogens diffuse across the surfaces of cells from a concentrated source. Therefore different embryonic cells get different concentrations of morphogens.

This results in the activation and inhibition of different genes in different cells. This in turn controls how long your fingers should be, where your nose is on your face, and other specifics about body structure.
7.2. U2 Nucleosomes help to regulate transcription in eukaryotes.

Changes in the environment affect the cell metabolism, this in turn can directly or indirectly affect processes such as Acetylation & Methylation.

Methylation and acetylation mark the DNA to affect transcription. These markers are known as **epigenetic tags**.

Reprogramming scours the genome and erases the epigenetic tags to return the cells to a genetic "blank slate".

*The branch of genetics concerned with heritable changes not caused by DNA is called Epigenetics.*

For a new organism to grow it needs unmarked DNA that can develop into lots of different specialised cell types.

For a small number of genes, epigenetic tags make it through this process unchanged hence get passed from parent to offspring.

http://learn.genetics.utah.edu/content/epigenetics/inheritance/images/Reprogramming.jpg
7.2.S1 Analysis of changes in the DNA methylation patterns.

The images show a mapping of chromosomal regions with differential DNA methylation in monozygotic (identical) twins. The sample is of metaphase chromosomes (humans have 23 pairs).

- Hypomethylation (low levels of methylation) in one twin compared to the other.
- Hypermethylation (high levels of methylation) in one twin compared to the other.
- Similar levels of methylation in both twins.
- Hypomethylation (low levels of methylation) in one twin compared to the other.

The diagrams maps changes between the twins’ levels of methylation of DNA across the chromosomes.

Your task: compare the diagrams (not the graphs) then analyse the evidence and deduce conclusions.
7.2.S1 Analysis of changes in the DNA methylation patterns.

Comparison (evidence):
The variation in the levels of methylation between the twins increases with age.

Hypomethylation occurs more away from the end of the chromosomes.

Chromosome 3 shows the greatest variation in methylation at both ages.

Chromosome 17 changes the least with age.

The levels of both hypermethylation & hypomethylation increase with age.

http://www.pnas.org/content/102/30/10604/F3.expansion.html
7.2.S1 Analysis of changes in the DNA methylation patterns.

**Analysis and deductions:**

The variation in the levels of methylation between the twins increases with age.

The twins will have different experiences/environment stimuli which will in turn cause different levels of methylation on different chromosomes.

http://www.pnas.org/content/102/30/10604/F3.expansion.html
Analysis of changes in the DNA methylation patterns.

**Analysis and deductions:**

The levels of both hypermethylation & hypomethylation increase with age.

**Methylation inhibits transcription:** as the organism ages cells becomes more greatly specialised due to more inhibited and promoted DNA.

http://www.pnas.org/content/102/30/10604/F3.expansion.html
Analysis of changes in the DNA methylation patterns.

**Analysis and deductions:**

Chromosome 17 shows the greatest variation in methylation at both ages.

Chromosome 20 changes the least with age.

*Numbers, sizes and roles of genes vary between chromosomes. Therefore it is expected that some chromosomes vary more than others in their degree of methylation and how they are affected by time/environment.*

http://www.pnas.org/content/102/30/10604/F3.expansion.html
7.2.S1 Analysis of changes in the DNA methylation patterns.

**Analysis and deductions:**

Chromosome 17 shows the greatest variation in methylation at both ages.

Chromosome 20 changes the least with age.

**Numbers, sizes and roles of genes varies between chromosomes. Therefore it is expected that some chromosomes vary more than others in their degree of methylation and how they are affected by time/environment.**

http://www.pnas.org/content/102/30/10604/F3.expansion.html
Review: 2.7.U4 Transcription is the synthesis of mRNA copied from the DNA base sequences by RNA polymerase.

TRANSCRIPTION: In the nucleus, the cell's machinery copies the gene sequence into messenger RNA (mRNA), a molecule that is similar to DNA. Like DNA, mRNA has four nucleotide bases - but in mRNA, the base uracil (U) replaces thymine (T).

TRANSLATION: The protein-making machinery, called the ribosome, reads the mRNA sequence and translates it into the amino acid sequence of the protein. The ribosome starts at the sequence AUG, then reads three nucleotides at a time. Each three-nucleotide codon specifies a particular amino acid. The "stop" codons (UAA, UAG and UGA) tell the ribosome that the protein is complete.
Review: 2.7.U4 Transcription is the synthesis of mRNA copied from the DNA base sequences by RNA polymerase.

Transcription is the process by which an RNA sequence is produced from a DNA template:

Three main types of RNA are predominantly synthesised:

- **Messenger RNA (mRNA):** A transcript copy of a gene used to encode a polypeptide
- Transfer RNA (tRNA): A clover leaf shaped sequence that carries an amino acid
- Ribosomal RNA (rRNA): A primary component of ribosomes

We are focusing on mRNA

http://www.nature.com/scitable/topicpage/Translation-DNA-to-mRNA-to-Protein-393
The enzyme RNA polymerase binds to a site on the DNA at the start of a gene (The sequence of DNA that is transcribed into RNA is called a gene).

RNA polymerase separates the DNA strands and synthesises a complementary RNA copy from the antisense DNA strand.

**Transcription occurs in a 5’ to 3’ direction**: RNA polymerase adds the 5’ end of the free RNA nucleotide to the 3’ end of the growing mRNA molecule (RNA polymerase moves along the antisense strand in a 3’ to 5’ direction).

It does this by covalently bonding ribonucleoside triphosphates that align opposite their exposed complementary partner (using the energy from the cleavage of the additional phosphate groups to join them together).

Once the RNA sequence has been synthesised:
- RNA polymerase will detach from the DNA molecule
- RNA detaches from the DNA
- The double helix reforms

Transcription occurs in the nucleus (where the DNA is) and, once made, the mRNA moves to the cytoplasm (where translation can occur).

https://commons.wikimedia.org/wiki/File:DNA_transcription.jpg
Eukaryotic cells modify mRNA after transcription. Eukaryotic genes (unlike prokaryote) contain base sequences that are not translated into polypeptides.

**Exons** are coding sections of the gene. **Introns** are non-coding sections of the gene.

Exons are removed then broken down back into nucleotides ready for use. Introns are removed and then broken down back into nucleotides. The Spliceosome forms and causes the introns to form loops which allows the exons to be joined.

**Mature mRNA** contains only exons leaves the nucleus to be translated into polypeptides.

*The spliceosome is a complex assembled from small nuclear RNA (snRNA) and proteins.*
7.2.U4 Splicing of mRNA increases the number of different proteins an organism can produce.

The splicing process above can happen in different ways to the same gene. Particular exons (of a gene) may be included within or excluded from mature mRNA. Multiple proteins produced by a single gene. Each protein produced will vary in its biological function. An example of this is the IgM gene which produces different immunoglobulins (antibodies) to fight different pathogens.

https://commons.wikimedia.org/wiki/File:DNA_alternative_splicing.gif
7.3 Translation

Essential idea: Information transferred from DNA to mRNA is translated into an amino acid sequence.

"The genes in DNA encode protein molecules, which are the "workhorses" of the cell, carrying out all the functions necessary for life ... In the simplest sense, expressing a gene means manufacturing its corresponding protein"

http://www.nature.com/scitable/topicpage/translation-dna-to-mrna-to-protein-393

https://commons.wikimedia.org/wiki/File:Aminoacids_table.svg
<table>
<thead>
<tr>
<th>Statement</th>
<th>Guidance</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.3.U1 Initiation of translation involves assembly of the components that</td>
<td>Examples of start codons are not required. Names of the tRNA binding sites</td>
</tr>
<tr>
<td>carry out the process.</td>
<td>are expected as well as their roles.</td>
</tr>
<tr>
<td>7.3.U2 Synthesis of the polypeptide involves a repeated cycle of events.</td>
<td></td>
</tr>
<tr>
<td>7.3.U3 Disassembly of the components follows termination of translation.</td>
<td></td>
</tr>
<tr>
<td>7.3.U4 Free ribosomes synthesize proteins for use primarily within the cell.</td>
<td></td>
</tr>
<tr>
<td>7.3.U5 Bound ribosomes synthesize proteins primarily for secretion or for use in lysosomes.</td>
<td></td>
</tr>
<tr>
<td>7.3.U6 Translation can occur immediately after transcription in prokaryotes due to the absence of a nuclear membrane.</td>
<td></td>
</tr>
<tr>
<td>7.3.U7 The sequence and number of amino acids in the polypeptide is the primary structure.</td>
<td></td>
</tr>
<tr>
<td>7.3.U8 The secondary structure is the formation of alpha helices and beta pleated sheets stabilized by hydrogen bonding.</td>
<td></td>
</tr>
<tr>
<td>7.3.U9 The tertiary structure is the further folding of the polypeptide stabilized by interactions between R groups.</td>
<td>Polar and non-polar amino acids are relevant to the bonds formed between R groups.</td>
</tr>
<tr>
<td>7.3.U10 The quaternary structure exists in proteins with more than one polypeptide chain.</td>
<td>Quaternary structure may involve the binding of a prosthetic group to form a conjugated protein.</td>
</tr>
</tbody>
</table>
### Applications and Skills

<table>
<thead>
<tr>
<th>Statement</th>
<th>Guidance</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.3.A1</td>
<td>tRNA-activating enzymes illustrate enzyme–substrate specificity and the role of phosphorylation.</td>
</tr>
<tr>
<td>7.3.S1</td>
<td>Identification of polysomes in electron micrographs of prokaryotes and eukaryotes.</td>
</tr>
<tr>
<td>7.3.S2</td>
<td>The use of molecular visualization software to analyse the structure of eukaryotic ribosomes and a tRNA molecule.</td>
</tr>
</tbody>
</table>
Review: 2.7.U4 Transcription is the synthesis of mRNA copied from the DNA base sequences by RNA polymerase. AND 2.7.U5 Translation is the synthesis of polypeptides on ribosomes.

**TRANSCRIPTION:** In the nucleus, the cell's machinery copies the gene sequence into messenger RNA (mRNA), a molecule that is similar to DNA. Like DNA, mRNA has four nucleotide bases - but in mRNA, the base uracil (U) replaces thymine (T).

**TRANSLATION:** The protein-making machinery, called the ribosome, reads the mRNA sequence and translates it into the amino acid sequence of the protein. The ribosome starts at the sequence AUG, then reads three nucleotides at a time. Each three-nucleotide codon specifies a particular amino acid. The "stop" codons (UAA, UAG and UGA) tell the ribosome that the protein is complete.

http://learn.genetics.utah.edu/content/molecules/transcribe/
Translation is the process of protein synthesis in which the genetic information encoded in mRNA is translated into a sequence of amino acids in a polypeptide chain.

A ribosome is composed of two halves, a large and a small subunit. During translation, ribosomal subunits assemble together like a sandwich on the strand of mRNA:

- Each subunit is composed of RNA molecules and proteins
- The small subunit binds to the mRNA
- The large subunit has binding sites for tRNAs and also catalyzes peptide bonds between amino acids

http://www.nature.com/scitable/topicpage/ribosomes-transcription-and-translation-14120660
Review: 2.7.U5 Translation is the synthesis of polypeptides on ribosomes. (plus AHL detail)

To understand how ribosomes work learn how to draw, label and annotate their structure.

Use this slide and the previous one to construct your own diagram.

tRNA currently holding the growing peptide is located at this site.

tRNAs leave the ribosome at this site.

tRNA carrying amino acids bind first in this site.

http://old-www.hartnell.edu/tutorials/biology/images/ribosome_sm.jpg
7.3.S2 The use of molecular visualization software to analyse the structure of eukaryotic ribosomes and a tRNA molecule.

The structure of tRNA matches its function.

Function: to bring amino acids from the cytoplasm to the growing polypeptide and to attach them in the correct location.

Clover-leaf structure

single chain of RNA

7.3.S2 The use of molecular visualization software to analyse the structure of eukaryotic ribosomes and a tRNA molecule.

Use the RCSB Protein Bank to explore the following Jmol images:
- The large ribosomal subunit
- The Path of Messenger RNA Through the Ribosome
- A tRNA molecule

For each image:
- Identify the structures you know of from your diagrams
- Compare and contrast the usefulness of the jmol image and your diagrams

http://www.rcsb.org/pdb/education_discussion/educational_resources/tRNA_jmol.jsp
Nature of Science: Developments in scientific research follow improvements in computing—the use of computers has enabled scientists to make advances in bioinformatics applications such as locating genes within genomes and identifying conserved sequences. (3.7)

Without computers analysis of the molecular structure such as ribosomal and tRNA structure would not be possible. The calculations that now take seconds to complete and would take weeks and months to do manually.

Bioinformatics also relies on computers to large extent as it involves analysing large amounts of data stored in vast databases. Bioinformatics is an interdisciplinary field of science, bioinformatics combines computer science, statistics, mathematics, and engineering to study and process biological data, in particular genetics and genomics.
tRNA is activated by a tRNA activating enzyme

tRNA delivers amino acids to the growing polypeptide chain in translation.

It picks up new amino acids when activated by a specific tRNA activating enzyme. This uses ATP.

There are 20 of these enzymes, corresponding to the 20 amino acids, for which the tRNA molecule has the complementary anticodon.

The energy in the bond linking the tRNA molecule to the amino acid will be used in translation to form a peptide bond between adjacent amino acids.
The genetic code is the set of rules by which information encoded in mRNA sequences is converted into proteins (amino acid sequences) by living cells.

- **Codons are a triplet of bases** which encodes a particular amino acid.
- As there are four bases, there are 64 different codon combinations ($4 \times 4 \times 4 = 64$).
- The codons can translate for 20 amino acids.
- Different codons can translate for the same amino acid (e.g., GAU and GAC both translate for Aspartate) therefore the genetic code is said to be **degenerate**.
- The order of the codons determines the amino acid sequence for a protein.
- The **coding region always starts with a START codon** (AUG) therefore the first amino acid in all polypeptides is Methionine.
- The **coding region of mRNA terminates with a STOP codon** - the STOP codon does not add an amino acid – instead it causes the release of the polypeptide.

Amino acids are carried by transfer RNA (tRNA)
The anti-codons on tRNA are complementary to the codons on mRNA.
Review: 2.7.U8 Translation depends on complementary base pairing between codons on mRNA and anticodons on tRNA.

Key components of translation that enable genetic code to synthesize polypeptides

1. mRNA has a sequence of codons that specifies the amino acid sequence of the polypeptide.
2. tRNA molecules carry the amino acid corresponding to their codon.
3. tRNA molecules have an anticodon of three bases that binds to a complementary codon on mRNA.
4. Ribosomes:
   - act as the binding site for mRNA and tRNA
   - catalyse the peptide bonds of the polypeptide
7.3.U1 Initiation of translation involves assembly of the components that carry out the process. AND 7.3.U2 Synthesis of the polypeptide involves a repeated cycle of events. AND 7.3.U3 Disassembly of the components follows termination of translation.

The process of translation can be summarised by the following steps:

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>mRNAbinds to the small subunit of the ribosome.</td>
</tr>
<tr>
<td>2.</td>
<td>The small subunit of the ribosome moves along the mRNA molecule in a 5' - 3' direction until it reaches a start codon (AUG)</td>
</tr>
<tr>
<td>3.</td>
<td>A molecule of tRNA (with it’s amino acid, Methionine attached) complementary to the start codon (UAC) binds to the P site of the ribosome</td>
</tr>
<tr>
<td>4.</td>
<td>The large subunit of the ribosome binds to the tRNA and small subunit</td>
</tr>
<tr>
<td>5.</td>
<td>A second tRNA (with amino acid attached) complementary to the second codon on the mRNA then binds to the A site of the ribosome</td>
</tr>
<tr>
<td>6.</td>
<td>The amino acid carried by the tRNA in the P site is transferred to the amino acid in the A site as a consequence of the ribosome catalyzing a new peptide bond (condensation reaction). The growing polypeptide increases in length.</td>
</tr>
</tbody>
</table>
| 7.   | The ribosome moves one codon along the mRNA (in a 5’ – 3’ direction):  
  • The tRNA in the P site is moved to the E site and then released  
  • The tRNA in the A site is moved into P site |
| 8.   | Another tRNA binds, complementary to the next codon on the mRNA, binds to the A site. |
| 9.   | Steps 6, 7, and 8 are repeated until a stop codon is reached. |
| 10.  | When a stop codon is reached translation is stopped:  
  • a release factor attaches to the A site  
  • the polypeptide chain is released  
  • the ribosome complex dissembles ready for reuse translating another mRNA molecule |
7.3.U1 Initiation of translation involves assembly of the components that carry out the process.

**Initiation:**

1. mRNA binds to the small subunit of the ribosome.
2. The small subunit of the ribosome moves along the mRNA molecule in a 5' - 3' direction until it reaches a start codon (AUG).
3. A molecule of tRNA (with its amino acid, Methionine attached) complementary to the start codon (UAC) binds to the P site of the ribosome.
   
   **Note:** A, P, E notation, though not necessary makes the process of translation easier to explain.

   n.b. The A, P, E notation, though not necessary makes the process of translation easier to explain.

4. The large subunit of the ribosome binds to the tRNA and small subunit.

http://www.hartnell.edu/tutorials/biology/translation.html
7.3. U2 Synthesis of the polypeptide involves a repeated cycle of events.

**Elongation:**

5. A second tRNA (with amino acid attached) complementary to the second codon on the mRNA then binds to the A site of the ribosome.

5. The amino acid carried by the tRNA in the P site is transferred to the amino acid in the A site as a consequence of the ribosome catalyzing a new peptide bond (condensation reaction). The growing polypeptide increases in length.

7. The ribosome moves one codon along the mRNA (in a 5’ – 3’ direction):
   - The tRNA in the P site is moved to the E site and then released
   - The tRNA in the A site is moved into P site

8. Another tRNA binds, complementary to the next codon on the mRNA, binds to the A site.

7. Steps 6, 7, and 8 are repeated until a stop codon is reached.
7.3.U3 Disassembly of the components follows termination of translation.

**Termination:**

10. When a stop codon is reached translation is stopped:
- a release factor attaches to the A site
- the polypeptide chain is released
- the ribosome complex dissembles ready for reuse translating another mRNA molecule
Translation can occur immediately after transcription in prokaryotes due to the absence of a nuclear membrane.

In prokaryotes ribosomes can be adjacent to the chromosomes whereas in eukaryotes the mRNA needs to be relocated from the nucleus to the cytoplasm (through the nuclear membrane).

In eukaryotes the mRNA is modified, e.g. spliced, (here labelled processing) after transcription before translation, this does not occur in prokaryotes.
A polysome is a structure that consists of multiple ribosomes attached to a single mRNA. Multiple ribosomes translating mRNA simultaneously enables the cell to quickly create many copies of the required polypeptide.
In prokaryotes the chromosome may have numerous polysomes attached directly to it.

Features to take note of:
- The arrow indicates the chromosome
- Polysomes show as the side chains. The dark ball-like ribosomes on each polysome side chain are clearly visible

Electron micrograph showing growing polysomes on a chromosome from E. coli.

*n.b. polysomes in eukaryotes occur separately in the cytoplasm or on the ER. This is because Ribosomes attach to the mRNA as it is being translated.*
Review: 2.7.U4 Transcription is the synthesis of mRNA copied from the DNA base sequences by RNA polymerase. AND 2.7.U5 Translation is the synthesis of polypeptides on ribosomes.

Q - What is the purpose of transcription and translation?
A - These processes work together to create a polypeptide which in turn folds to become a protein. Proteins carry many essential functions in cells. For more detail review 2.4.U7 Living organisms synthesize many different proteins with a wide range of functions.

Catalysis | Tensile strengthening | Transport of nutrients and gases | Cell adhesion
--- | --- | --- | ---
Muscle contraction | Heartbeats | Hormones
Cytoskeletons | Muscle contraction | Receptors
Blood clotting | Blood clotting | Packing of DNA
Membrane transport | Cell adhesion | Immunity

Use the learn.genetics tutorial to discover one example:

Http://learn.genetics.utah.edu/content/molecules/firefly/
Review: 2.4.U6 The amino acid sequence determines the three-dimensional conformation of a protein.

Polypeptides are chains of amino acids joined by peptide bonds:

![Peptide bond diagram](image)

There are 20 different amino acids. These can be combined in any order. Each amino acid has unique properties:
- some are polar (hydrophilic)
- some are non-polar (hydrophobic)
- some are positively or negatively charged
- some contain sulphur

The properties of the amino acids determine how a polypeptide folds up into a protein.
The amino acid sequence determines the three-dimensional conformation of a protein. Proteins are commonly described as either being fibrous or globular in nature. Fibrous proteins have structural roles whereas globular proteins are functional (active in a cell’s metabolism).

<table>
<thead>
<tr>
<th>Properties</th>
<th>Fibrous Protein</th>
<th>Globular Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shape</td>
<td>Long and narrow</td>
<td>Rounded / spherical</td>
</tr>
<tr>
<td>Role</td>
<td>Structural (strength and support)</td>
<td>Functional (catalytic, transport, etc.)</td>
</tr>
<tr>
<td>Solubility</td>
<td>(Generally) insoluble in water</td>
<td>(Generally) soluble in water</td>
</tr>
<tr>
<td>Sequence</td>
<td>Repetitive amino acid sequence</td>
<td>Irregular amino acid sequence</td>
</tr>
<tr>
<td>Stability</td>
<td>Less sensitive to changes in heat, pH, etc.</td>
<td>More sensitive to changes in heat, pH, etc.</td>
</tr>
<tr>
<td>Examples</td>
<td>Collagen, myosin, fibrin, actin, keratin, elastin</td>
<td>Catalase, haemoglobin, insulin, immunoglobulin</td>
</tr>
</tbody>
</table>

In globular proteins the hydrophobic R groups are folded into the core of the molecule, away from the surrounding water molecules, this makes them soluble. In fibrous proteins the hydrophobic R groups are exposed and therefore the molecule is insoluble.
There are four levels of protein structure. Which level a protein conforms to is determined by its amino acid sequence.

**Primary Structure**
- The order / sequence of the amino acids of which the protein is composed
- Formed by covalent peptide bonds between adjacent amino acids
- Controls all subsequent levels of structure

**Secondary Structure**
- The chains of amino acids fold or turn upon themselves
- Held together by hydrogen bonds between (non-adjacent) amine (N-H) and carboxylic (C-O) groups
- H-bonds provide a level of structural stability
- **Fibrous proteins**

**Tertiary Structure**
- The polypeptide folds and coils to form a complex 3D shape
- Caused by interactions between R groups (H-bonds, disulphide bridges, ionic bonds and hydrophilic / hydrophobic interactions)
- Tertiary structure may be important for the function (e.g. specificity of active site in enzymes)
- **Globular proteins**

**Quaternary Structure**
- The interaction between multiple polypeptides or prosthetic groups
- A prosthetic group is an inorganic compound involved in a protein (e.g. the heme group in haemoglobin)
- **Fibrous and Globular proteins**

Research types of proteins.