Polymerase chain reaction and its application in diagnosis of COVID-19

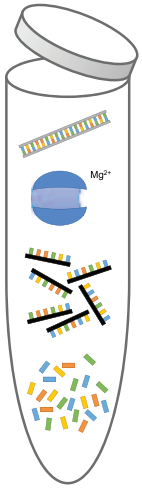
The polymerase chain reaction (PCR) is a method that is used in many laboratories to make millions (or even billions) of copies of a piece of DNA. The reaction is carried out "*in vitro*", meaning outside of a living organism. The amplified DNA can be used in forensics to help solve crimes, determine paternity (DNA fingerprinting) and diagnose genetic disorders.

It is also used extensively in microbiology to diagnose infectious diseases such as COVID-19 and in research to help understand how microbes evolve, function and are transmitted.

Amplification of DNA by PCR

The method uses **primers**, which are pieces of single stranded DNA that are complementary to a specific target sequence of the DNA strand that is to be amplified. Components that are required for the PCR reaction are shown in Figure 1. The reaction involves three steps (Figure 2):

- DNA is heated to break hydrogen bonds between the base pairs and separate the strands **(Step 1. Denaturing)**.
- The temperature is then decreased to allow primers to bind (anneal) to the complementary sequence of target DNA at the 3′ end of each strand **(Step 2. Annealing)**.
- The temperature is then increased to allow a heat-tolerant **DNA polymerase** to add nucleotides to the DNA strands at the 3′ end of each bound primer **(Step 3. Extension)**.



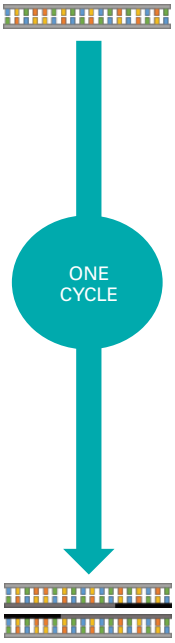
DNA from sample containing region to be amplified

DNA polymerase in buffer with co-factor

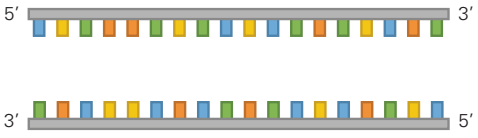
Two types of primer

4 types of nucleotides

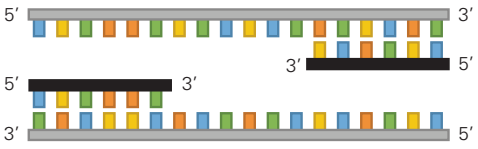
Figure 1. PCR components



STEP 1. DENATURING
DNA is heated to separate the strands.
Short time at 92-98°C



STEP 2. ANNEALING
Primers bind to target DNA.
Short time at 50-65°C



STEP 3. EXTENSION
DNA polymerase replicates a target section of DNA by binding nucleotides to DNA strands.
Longer time at 70-80°C

Original DNA in sample
Newly synthesised DNA

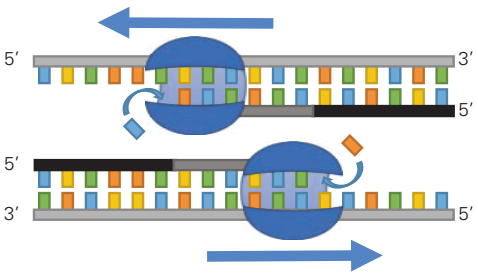


Figure 2. Steps in one cycle of PCR

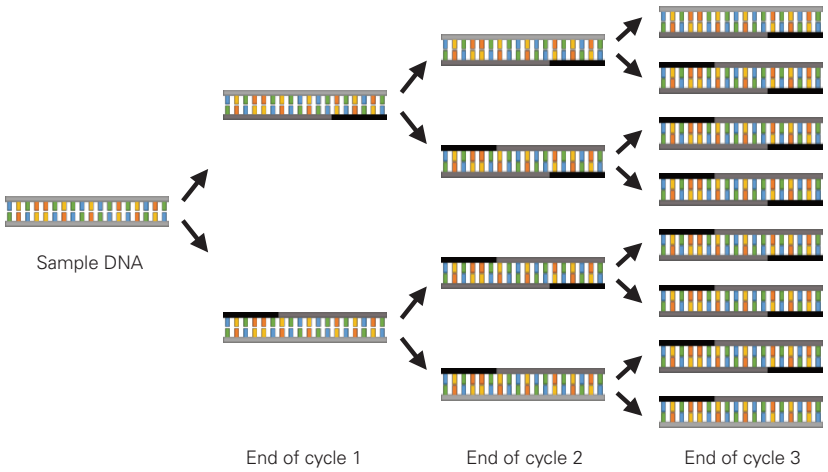


Figure 3. Graphical representation of PCR amplification

This process replicates a target section of DNA. The cycle is repeated 20-30 times and **the amount of DNA present doubles with every cycle** (Table 1 and Figure 3).

Each cycle produces two identical molecules for each target sequence in the sample. This gives an exponentially increasing quantity of specific DNA molecules.

In most laboratories a **thermal cycler** is used, which is a piece of computerised heating apparatus that can control the temperature changes required for PCR (Figure 4).

Visualisation of amplified DNA by gel electrophoresis

Amplified DNA is often visualised using **gel electrophoresis**. This method is used to separate electrically charged molecules so the quality and size of DNA fragments can be checked.

The process involves passing a current though an agarose gel. Samples are loaded into small wells at the negative end and DNA molecules travel through the gel to the positive end (Figure 5).

Molecules are separated according to size with the smallest molecules travelling more quickly and therefore the farthest in a given period of time. The DNA can be stained for visualisation. DNA fragments that are the same size, such as those generated during PCR, appear as a single band on the gel.

DNA standards are also run on the gel. These are DNA fragments of known size (molecular weight) so the size of DNA fragments within the sample can be checked. A **positive control** is used to ensure that the reaction worked and a **negative control** is used to check that the reagents are not contaminated.

Number of PCR cycles (n)	Number of copies of template DNA $(2)^n$
0	1
1	2
2	4
3	8
4	16
5	32
10	1,024
20	1,048,576
30	1,073,741,824

Table 1: DNA copy number per number of cycles

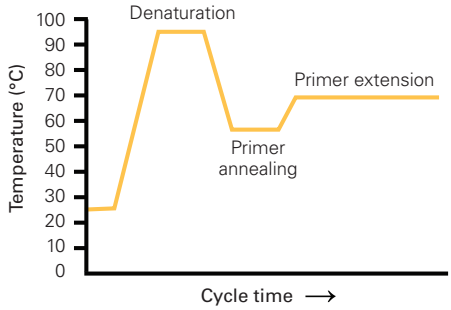
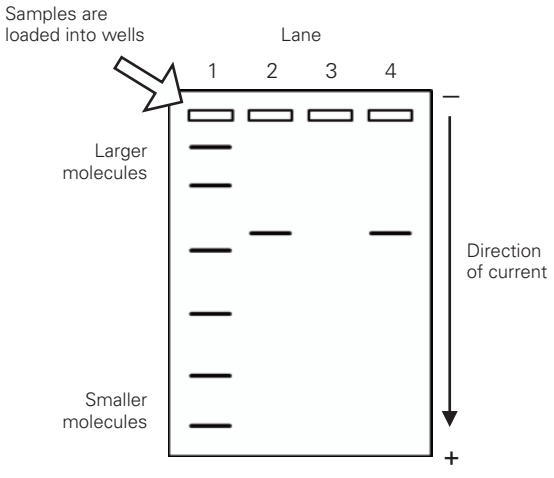


Figure 4. Different thermal steps of PCR



1. DNA standards.
2. Positive control.
3. Negative control.
4. PCR positive sample.

Figure 5. Gel electrophoresis

Quantitative PCR

Amplified DNA can also be monitored using fluorescent dyes which bind to DNA. These can be added onto primers or included in the PCR reagent mix. Once the dye is bound to DNA it emits a signal that can be measured using specialised equipment.

The intensity of the signal is proportional to the amount of DNA, allowing the DNA to be quantified. this method is called **quantitative PCR (qPCR)** or real-time PCR, because the amount of DNA can be measured over time.

An example of the data generated during a qPCR reaction is shown in Figure 6. Initially, the DNA is generated exponentially, then the curve flattens at the top because the number of new copies made decreases as reaction components are depleted and become limiting.

This method is often used in diagnostic laboratories because the results of the PCR can be obtained without the need for gel electrophoresis, and therefore a diagnosis can be reached more rapidly. It is also possible to perform multiple reactions within the same tube, using different fluorescent dyes. This means you can test for a number of infectious agents at the same time.

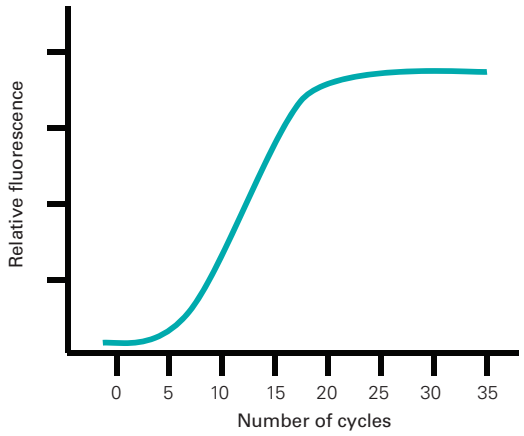


Figure 6. Example of data generated during a qPCR experiment

Diagnosing COVID-19 infection

PCR is used extensively in diagnosing infectious diseases. Nucleic acids are extracted from patient samples and regions that are specific to the infectious agents are amplified by PCR. This generates sufficient quantities of DNA to detect and confirm that the patient is infected.

SARS-CoV-2 (**s**evere **a**cute **r**espiratory **s**yndrome **co**rona**v**irus **2**) is the virus that causes COVID-19 (Figure 7). SARS-CoV-2 is a coronavirus, which is a group of viruses that can infect humans and animals. In humans they usually cause respiratory infections but they can cause other diseases including diarrhoea in cattle, pigs, chickens cats and dogs.

Soon after the identification of COVID-19 in late 2019, scientists began working on diagnostic tests to detect the virus. Because SARS-CoV-2 is a newly emerged virus, a number of steps were needed before a PCR diagnostic test could be developed (Figure 8). This included confirming SARS-CoV-2 as the cause of COVID-19 and sequencing of the virus genome.

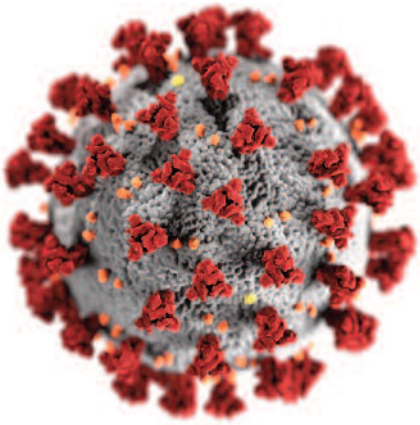


Figure 7. Illustration of SARS-CoV-2

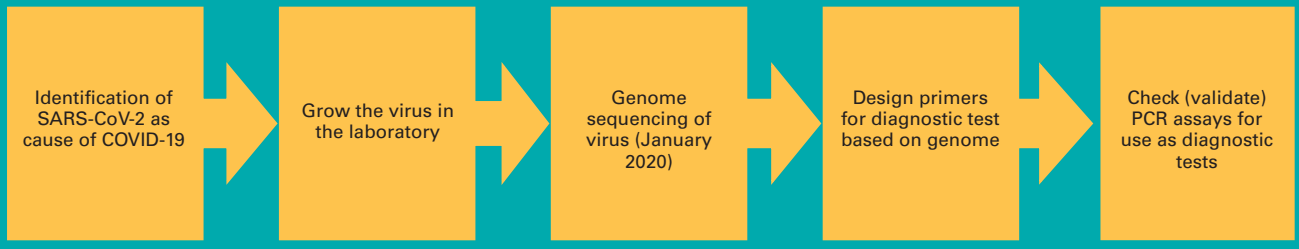


Figure 8. Steps required to develop diagnostic PCR test for SARS-CoV-2

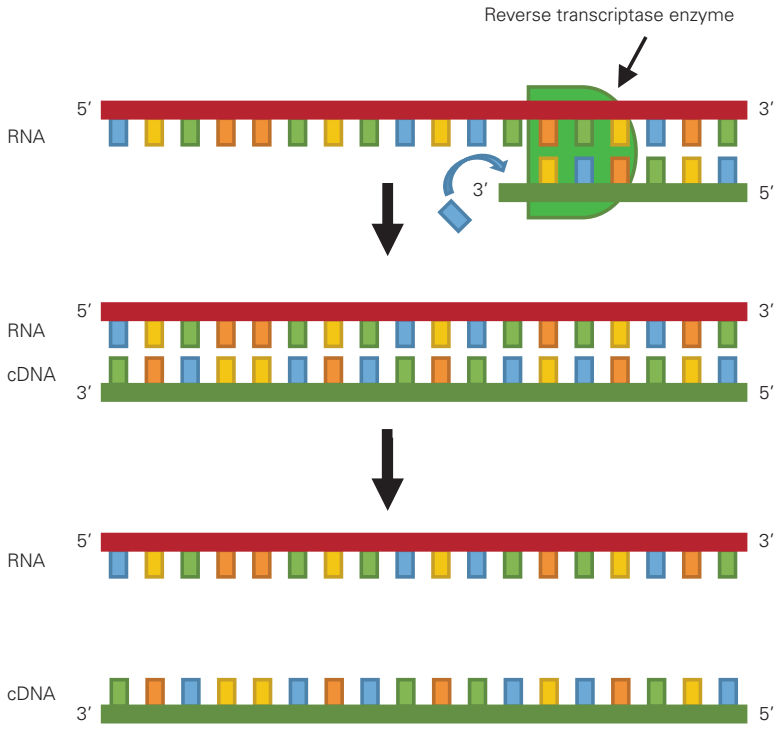
Reverse transcription PCR

SARS-CoV-2 is an "RNA virus" which means it carries its genetic information as RNA rather than DNA. Only DNA can be amplified using PCR. Because SARS-CoV-2 contains RNA, an extra step is needed in the laboratory before PCR, to convert the viral RNA into DNA.

The conversion of RNA to DNA is carried out by an enzyme called **reverse transcriptase** because the process is the reverse of normal gene transcription where DNA is transcribed to messenger RNA (mRNA). DNA that is synthesised in this way is called complementary DNA (cDNA). This type of PCR is called **reverse transcription PCR (Figure 9)**.

A number of PCR tests have been developed to detect SARS-CoV-2 since the virus was sequenced in January 2020. These tests use a variety of different viral RNA targets, including regions within the gene encoding the surface spike proteins and the RdRp gene that codes for an RNA-dependent RNA polymerase.

During an infection, the virus uses the infected host's cellular machinery to make more copies of itself by replicating viral RNA, expressing the virus proteins and assembling the virus particle. The RdRp gene codes for a viral enzyme that enables replication of viral RNA.



Nucleotides are added by the reverse transcriptase enzyme.

A strand of cDNA is synthesised, complementary to the RNA.

During the first denaturation step of PCR, the strands are separated and then the single cDNA strand is amplified as usual.

Figure 9. Reverse transcription of RNA

Designing a diagnostic test for COVID-19: considerations

During the course of a disease outbreak such as COVID-19, large numbers of people must be tested to see if they currently have the infection. This is to ensure they are diagnosed and then treated properly, to ensure measures are taken to stop the infection spreading and to understand more about how the disease is transmitted in a population.

In the case of a newly emerged disease such as this, new diagnostics and procedures must be developed quickly. This is not an easy task and there are a number of considerations that are needed to make sure data are collated accurately:

The PCR test:

Is the test specific enough?

Primers must not match nucleotide sequences in other viruses, such as the common cold virus. This ensures that the likelihood of "false positive" results is low and that a positive result is due to the presence of SARS-CoV-2. When selecting primers, scientists consult large databases of DNA and RNA sequence data to check this.

Is the test sensitive enough?

Some PCR assays work better than others. Scientists try to **optimise** the assays to ensure that even small amounts of viral RNA can be detected. If the assay isn't sensitive enough a patient might be misdiagnosed as being un-infected, also known as a false negative.

Will the PCR keep working if the virus evolves?

Viral genomes often change over time due to mutations. Some regions are more likely to change than others. **The region of viral RNA that has been selected for PCR analysis must be stable and unlikely to change**. If the region of DNA that the primers bind to changes, the PCR might not work.

Sampling:

Which sample is most appropriate?

For respiratory-transmitted infections such as COVID-19, a nose and/or throat swab or a sputum sample is currently used, depending on the stage of the infection.

Labelling and documentation

It is crucial that the sample is labelled properly so the patient gets the correct result. The results must also be stored safely to ensure patient confidentiality.

Samples are often labelled with bar codes which are unique for each patient. This reduces the amount of paperwork for each sample and allows more rapid processing.

The laboratory:

Transport stability and safety

Swabs taken from patients must be stored safely and transferred to secure diagnostic laboratories. RNA degrades quickly but samples can be stored in a liquid called Viral Transport Medium. This preserves the RNA.

Extraction of RNA

Viral RNA must be extracted (isolated) from the sample. This ensures the conditions for PCR are optimal, increasing the concentration of RNA and removing substances that might stop the reaction from working. Special laboratory procedures have been developed to do this efficiently and safely.

Logistics

In the case of an outbreak such as SARS-CoV-2, extremely large numbers of tests must be carried out. This means that the production of any **reagents** and **consumables** such as primers, Viral Transport Medium, DNA polymerase and swabs must be able to occur quickly and materials are distributed to diagnostic laboratories safely and consistently.

Staff must be appropriately trained and equipment and personal protective equipment (PPE) must be available.

For the patient:

What will this test tell us?

It is very important that we understand what the test is telling us and that we interpret the results correctly.
A **negative PCR result** might mean that the patient is not infected but it might also mean that the patient is infected but has not yet started to produce the virus in quantities that can be detected.

The PCR test is useful because it can tell us whether or not a person **is carrying** the virus and is likely to infect others, before the patient has the chance to generate antibodies.
It cannot however tell us if the person **has had** the virus once they recover. For this, we would need a test to detect circulating antibodies which are made by the body in response to infection and can be detected around 2 weeks after initial infection.

How long will it take?

Currently, all samples for SARS-CoV-2 testing must be analysed within a laboratory with special equipment and it can take 3 days to get a result.

Scientists are currently working on tests that will detect the virus immediately where the sample has been taken, for example at the bedside of a patient. This is called a "point-of-care" or "near-patient" test.

For the population:

The PCR test also plays an important role in contact tracing, which is a method that is used in public health to prevent spread of the disease in populations.

People who develop symptoms are tested using the PCR test, and for positive results public health teams identify anyone the infected person has been in contact with so that they can self-isolate and prevent the virus from spreading further.

This is important when lockdown measures are relaxed.

Glossary

Agarose	A polysaccharide that is used in gel electrophoresis. When powdered agarose is dissolved in water then cooled, a gel is created with pores that are large enough for molecules such as DNA and proteins to travel through.
Annealing	The second step in PCR where the temperature of the sample is decreased to 50-65°C. This allows primers to bind (anneal) to template DNA.
Antibody	Antibodies are Y-shaped proteins that form part of the immune response. During an infection, antibodies are generated that uniquely bind to specific structures on the infectious agent. Detection of these specific antibodies is evidence that the body has been exposed to that infectious agent. Also known as immunoglobulin.
Assay	In biology this is an experiment or test that is carried out to assess the presence, amount or activity of a substance within a sample.
Co-factor	In biochemistry a co-factor is a chemical, often a metal ion, that is required for an enzyme's activity. The co-factor for the DNA polymerase enzyme is a magnesium ion (Mg^{2+}), which is added to the buffer as magnesium chloride.
Complementary DNA	Strand of DNA that is generated from RNA, by reverse transcription. Also known as cDNA.
Consumable	In science laboratories this refers to equipment that is used up rapidly and cannot be re-used e.g. swabs, gloves, tubes, plastic pipettes.
Contact tracing	A method used in public health to prevent spread of certain infections within populations. People who have been exposed to an infected person are traced and in the case of COVID-19, asked to self-isolate.
Coronavirus	Family of viruses that infect humans and animals. A coronavirus carries RNA as its genetic material, which is enclosed in a spherical membrane. Spike proteins project from the surface like a crown, giving the virus its name ("corōna" is Latin for "crown or garland").
COVID-19	Infectious disease caused by the SARS-CoV-2 virus.
Denaturing	The first step in PCR where DNA is heated to a high temperature (92-98°C). This breaks hydrogen bonds between the strands so they separate into single strands.
Diagnostic test	In medicine this is a procedure that is used to establish whether or not a patient has a specific condition or disease.
DNA	DNA is short for deoxyribonucleic acid. DNA is made from nucleotides, and carries the genetic information for all eukaryotes, bacteria and archaea and some viruses.
DNA fingerprinting	Laboratory technique used to analyse patterns of DNA sequences. It can be used in forensics to provide a link between evidence at a crime scene and a suspect or to establish how genetically similar two individuals are (e.g. to determine paternity).
DNA polymerase	An enzyme that synthesises DNA molecules from nucleotides. During PCR a heat-tolerant DNA polymerase adds nucleotides to the 3' end of a primer that is bound to template DNA.
DNA sequencing	Laboratory technique used to determine the exact sequence of nucleotides in a DNA molecule.
Extension	The third step in PCR where the temperature is raised to 70-80°C so the DNA polymerase can add nucleotides to DNA strands at the 3' end of each bound primer.
False positive	In diagnostics this refers to a positive test result from an uninfected individual, possibly as a result of interfering substances in the patient sample or contamination.

False negative	In diagnostics this refers to a negative test result from an infected individual, possibly as a result of an inappropriate sample being used, problems with the reagents or because the sample has been taken when the virus is not present at sufficient levels for detection.
Fluorescent dye	Fluorescent dyes can absorb and re-emit light at a different frequency. These dyes are often used in biology to detect or visualise organisms, cells, DNA and proteins.
Forensics	Scientific methods that are applied to solve crimes.
Gel electrophoresis	Technique used to separate DNA fragments according to size. DNA fragments that are separated can be visualised on the gel using chemicals that bind the DNA.
Genome	The complete set of an organism's genetic material.
Hydrogen bond	A link between an electronegative atom and a hydrogen atom which is bound to another electronegative atom. These bonds form between the complementary base pairs of two DNA strands. They can be broken easily and during PCR, this is achieved by heating the sample to a high temperature.
in vitro	Describes studies that are carried out outside a living organism, often in a test tube. From the Latin term meaning "within the glass".
Microbiology	The study of organisms that are too small to be seen by the naked eye. This includes bacteria, archaea, viruses, fungi, prions, protozoa and algae. These organisms are known as microbes.
Messenger RNA	Messenger RNA (mRNA) is a single-stranded RNA molecule that is complementary to one DNA strand of a gene. mRNA is translated into an amino acid chain within the ribosome.
Mutation	A change that occurs in an organism's genetic code resulting in an altered DNA or RNA sequence. This can be a substitution, insertion or deletion.
Negative control	In a diagnostic test this is used to confirm that the reagents are not contaminated and that the test is specific. A negative control is expected to give a negative result.
Nucleic acid	A biopolymer comprised of nucleotides. DNA and RNA are nucleic acids.
Nucleotide	The building blocks of DNA and RNA. Each nucleotide consists of either deoxyribose or ribose sugar, which is joined to a phosphate group and a base. The 4 bases for DNA are thymine, cytosine, adenine and guanine. The 4 bases for RNA are uracil, cytosine, adenine and guanine.
Optimise	For diagnostics this refers to the process of altering protocols to ensure the test is as specific and sensitive as possible.
Paternity	The fact of being a father.
Point-of-care test	Diagnostic test that is performed at the time and place of patient care, rather than within a diagnostic laboratory. It often involves a hand-held device. Also known as a "near-patient" or "bedside" test.
Polymerase chain reaction	Method used to make millions of copies of a specific sequence of DNA.
Positive control	In a diagnostic test this is used to validate the test and ensure that good quality results can be obtained from the experimental protocol. A positive control is expected to give a positive result.
Primers	Short pieces of single-stranded DNA that bind to a target sequence of template nucleic acid within a sample.
Public health	Organised measures to prevent disease, promote health and prolong life within the community.

Quantitative PCR	Quantitative PCR (qPCR) is a technique where the amplification of DNA by PCR is monitored over time. Also known as real-time PCR.
RdRp gene	This gene is found in the SARS-CoV-2 virus and is used as a target for diagnostic PCR tests. The RdRp gene codes for an enzyme called an RNA-dependent RNA polymerase.
Reagent	Substance or mixture for use in chemical or biological analysis or a chemical reaction. For PCR, reagents include DNA polymerase, four types of nucleotide, primers and a buffer.
Respiratory infection	An infection that affects the sinuses, nose, throat, airways or lungs.
Reverse transcriptase	An enzyme that synthesises complementary DNA from an RNA template.
Reverse transcription	The process by which a reverse transcriptase enzyme generates a complementary strand of DNA to an RNA molecule.
RNA polymerase	Enzyme that transcribes DNA into mRNA, synthesising mRNA from RNA nucleotides by complementary base pairing.
SARS-CoV-2	The virus that causes COVID-19.
Self-isolating	People who are self-isolating remain indoors and avoid contact with other people to prevent spreading infectious disease amongst the population.

This resource was created as support material for Higher Biology: DNA and the Genome.

Version 1, August 2020