

## Glossary of less known terms used in this course

- **Adaptors:** a short chemically synthesised single or double-strand DNA (oligonucleotide) used in some sequencing methodologies that allow the addition of a DNA barcode or other oligonucleotides downstream of an unknown or amplified DNA strand.
- **Amplicon:** a DNA fragment that has been amplified using polymerase chain reaction (PCR) or a method that results in the generation of multiple copies of that fragment.
- **A-tailing:** an enzymatic reaction to add a sequence of adenines at the 3'-terminus of a DNA fragment for sequencing purposes.
- **Barcoding:** addition of a tag of known DNA sequence (barcode, also called index) to an amplified DNA strand that permits sequencing multiple samples in parallel, and stratifying sample data informatically post/during sequencing.
- **cDNA:** complementary DNA, a DNA molecule synthesised from an RNA molecule.
- **Consensus sequence:** a sequence of DNA, RNA or protein generated from a set of aligned sequences. It is calculated based on the most frequent nucleic acids or residues in each position of the alignment.
- **COVID-19:** coronavirus disease 2019.
- **DNA:** deoxyribonucleic acid, which is an information molecule forming the "base code" for a living organism.
- **Enzyme:** a protein able to catalyse, i.e. accelerate chemical reactions.
- **Library:** in the sequencing context is a pool of DNA fragments attached to adaptors and/or other oligonucleotides used during sequencing preparation procedures.
- **ONT:** Oxford Nanopore Technologies, a company for sequencing technology. Their sequencing platform is also referred to as ONT.

- **PCR:** Polymerase chain reaction (PCR) is a laboratory technique for rapidly making (amplifying) millions to billions of copies of a given section of DNA, which may then be analysed further.
- **Primer:** a short chain of oligonucleotides used to target and 'prime' the initiation of DNA replication.
- **Protein translation:** is the process of synthesising proteins from a transcript.
- **NGS:** next-generation sequencing, a high throughput sequencing methodology.
- **Nucleotides:** the individual subunits that form DNA and RNA molecules.
- **RNA:** ribonucleic acid, an information molecule, can be the "base code" for viruses.
- **SARS-CoV-2:** Severe Acute Respiratory Syndrome Coronavirus 2.
- **SNP:** single-nucleotide polymorphism is a nucleotide substitution in a specific position of the genome. It is identified when compared to a reference sequence.
- **Transcription:** is the process of copying a segment of DNA into RNA.
- **Tagmentation:** is the initial step in library preparation where unfragmented DNA is cleaved and tagged for analysis.
- **Trimming:** in bioinformatics, it is the process of removing the ends of the reads, leaving only a region of high-quality bases and those relating to the sample, not the sequencing chemistry or DNA barcoding.
- **Viral replication:** is the mechanism in how viruses propagate during the infection cycle, it is the process of virus multiplication.
- **Viral genome variant:** a virus that has one or more mutations in its genome.
- **Whole-genome sequencing (WGS):** is the method of determining the entirety, or almost the entirety, of an organism's genomic DNA sequence all at once. This entails sequencing all of an organism's chromosomal DNA, and DNA found in mitochondria and, in plants, chloroplasts.