USAID FROM THE AMERICAN PEOPLE

RISE
Reaching Impact, Saturation, and Epidemic Control

## Reaching Impact, Saturation, and Epidemic Control (RISE)

# Learning Resource Package for Genome Sequencing of SARS-CoV-2 using Oxford Nanopore Sequencing Platform

**Guidelines on sequencing methodologies, sequencing strategies, and setting up of SARS-CoV-2 genome sequencing laboratories using Oxford Nanopore Sequencing Platform**

**May 2023**

## Executive Summary

The first case of novel coronavirus disease 2019 (COVID-19) were reported by the World Health Organization (WHO) in December 2019. Since then, there have been almost 625 million reported cases of COVID-19, and more than 6 million deaths due to this disease. COVID-19 is caused by a virus known as SARS-CoV-2. The COVID-19 pandemic is one of the most severe public health catastrophes the world has seen.

In the early stages of the pandemic there were few tools at our disposal to tackle the spread of the virus. The only control measures were non-pharmaceutical interventions (NPIs), such as social distancing, wearing of face masks and, most extreme of all, lockdowns. While such measures can reduce the transmission of COVID-19, they also have detrimental effects on mental health, children's education, and a country's economy.

We now have a wider range of tools at our disposal. There has been some progress in treatment of COVID-19 with the availability of new and repurposed drugs. Importantly, several effective vaccines have been developed, which indicates that there may be light at the end of the pandemic tunnel. However, despite the medical and technological advances that have been made, their rollout globally has been uneven, and there are considerable issues in terms of access, especially to vaccines. Therefore, diagnostics remain a key weapon in the fight against COVID-19.

Sequencing is a process used to decipher and interpret the genetic makeup of a biological organism; Next Generation Sequencing (NGS) are the available high-throughput, rapid, and scalable sequencing technologies used to determine the order of nucleotides present in DNA or RNA sequences of complete genomes or their parts. Applying NGS techniques enable rapid identification of unknown pathogens, discovery of genetic variations, and molecular understanding of disease-causing pathogens, to inform the development and utility of tests, treatments, and vaccines. An increasingly critical application of sequencing is genomic surveillance, which uses sequenced data from outbreak causing pathogens to identify them, and to understand how pathogens are introduced and spread through a population.

NGS-based diagnostic tests for COVID-19 became available from June 2020 and are not widely used. As the COVID-19 pandemic progresses, both knowledge of the disease and virus, and strategies for managing infection rate and reducing transmission, are evolving.

Indian SARS-CoV-2 Genomics Consortium (INSACOG) was established to expand Whole Genome Sequencing of SARS-CoV-2 across the nation, aiding our understanding of how the virus spreads and evolves. The Consortium initially started with a network of ten regional genome sequencing laboratories spread across the country and has now expanded to 58 (as on 30-Apr-2023) additional INSACOG Genome Sequencing Laboratories (IGSLs), that are mapped to most States and UTs to facilitate a smooth processing of all available positive samples.

The Learning Resource Package presented here summarizes the current knowledge in the evolution of NGS technology, protocols, and best practices. It begins with the training presentations for genome sequencing methods and procedures, bioinformatics analysis and interpretation. There are trainers' scripts and manual, standard operating procedures for few of the protocols the training evaluation questionnaire used in the workshops.

**Intended use:**

This LRP document is a comprehensive resource for all laboratory technicians, scientists, microbiologists and pathologists involved in the diagnostic laboratory ecosystem.

**How to use:**

This LRP document can be used for training laboratory personnel working on SARS-CoV-2 genome sequencing and bioinformatics analysis. The presentations can be used for cascade training and the SOPs and the worksheets can be used as templates in the laboratory..

**Intended Audience(s)**

This LRP is intended for members of committees appointed for developing or updating a SARS-CoV-2 sequencing facility, and other relevant stakeholders who influence SARS-CoV-2 sequencing related activities, such as laboratory personnel, laboratory managers, administrators, and program managers working in SARS-CoV-2 diagnostics.

Submitted by:

FIND, in collaboration with: JHPIEGO

## Acknowledgments

Reaching Impact, Saturation, and Epidemic Control (RISE) is a 5-year global project funded by the U.S. Agency for International Development (USAID). RISE works with countries to achieve a shared vision of attaining and maintaining epidemic control, with stronger local partners capable of managing and achieving results through sustainable, self-reliant, and resilient health systems by 2024.

We are grateful to the team at FIND, the global alliance for diagnostics, for insights derived from its vast experience in laboratory management. Further, this volume would not have been possible without the support of JHPIEGO, a non-profit organization for international health affiliated with Johns Hopkins University.

This document was made possible with support from the United States Agency for International Development (USAID) funded RISE program, under the terms of the cooperative agreement 7200AA19CA00003. The contents are the responsibility of the RISE program and do not necessarily reflect the views of USAID or the United States Government.

CONTRIBUTORS:

### FIND, the global alliance for diagnostics

Dr Sanjay Sarin, Vice President, Access
Dr Sarabjit Chadha, Regional Technical Director-India & South-East Asia
Dr Anita Suresh, Deputy Director of Genomics & Sequencing, FIND Geneva

### Authors

Mr Nithyanandan T, Senior Technical Associate
Mr. Subbaiah TB, External Technical Associate
Mr. Jithin S Sunny, External Bioinformatics Associate
Ms. Infanta Saleth Teresa Eden, External Bioinformatics Associate

### Reviewers *(in alphabetical order)*

Dr Anita Desai, Formerly Professor & Head of Neurovirology, National Institute of Mental Health and Neurosciences (NIMHANS)
Dr Lakshmi Soundararajan, Senior Microbiologist
Dr Liliana Rutaihwa, Scientist, FIND Geneva
Dr Marco Marklewitz, Scientist, FIND Geneva
Dr Nandhini Palani, Project Manager
Dr Swapna Uplekar, Senior Scientific Officer, FIND Geneva
Ms Preetishirin Katapur, Deputy Project Manager

### JHPIEGO, John Hopkins University Affiliates (India)

Dr Vineet Kumar, Chief of Party, RISE
Dr Nochiketa Mohanty, Deputy Chief of Party, RISE
Dr Prashant Singh, Senior Advisor Laboratory Strengthening, RISE

## List of Abbreviations

| | |
|---|---|
| ABI | Applied Biosystem |
| APHL | Association of Public Health Laboratories |
| APS | Adenosine 5´ Phosphosulfate |
| BAM | Binary Alignment Map |
| BGI | Beijing Genomics Institute |
| BLAST | Basic Local Alignment Search Tool |
| bp | base pairs |
| CCS | Circular Consensus Sequence |
| CDC | Centers for Disease Control and Prevention |
| cDNA | complementary DNA |
| CLR | Continuous Long Read |
| COVID-19 | coronavirus disease 2019 |
| cPAL | Combinatorial Probe Anchor Ligation |
| CT | Cycle Threshold |
| CTC | Configuration Test Cell |
| ddNTPs | di-deoxyribonucleotide triphosphates |
| DNA | Deoxyribonucleic acid |
| DNA-Seq | DNA Sequencing |
| DNB | DNA Nanoball |
| dNTPs | deoxyribonucleotide triphosphates |
| DOE | Department of Energy |
| dsDNA | A double stranded DNA |
| ELSI | Ethical, Legal and Social Issues |
| Gb | Gigabyte |
| GFF | General Feature Format |
| GISA | ID Global Initiative on Sharing All Influenza Data |
| GTF | Gene Transfer Format |
| HGP | Human Genome Project |
| HLA | Human Leukocyte Antigen |

| | |
|---|---|
| HMW | High Molecular Weight |
| HPC | High Performance Computing |
| HUGO | Human Genome Organization |
| INDEL | Insertion/Deletion |
| INSACOG | Indian SARS-CoV-2 Genomic Consortium |
| ITS | Internal Transcribed Spacer |
| IVD | In vitro diagnostics |
| kb | Kilobases |
| LED | Light Emitting Diode |
| mb | Megabases |
| mRNA | messenger RNA |
| NaCl | Sodium Chloride |
| NCBI | National Center for Biotechnology Information |
| NGS | Next Generation Sequencing |
| NHGRI | National Human Genome Research Institute |
| NIH | National Institute of Health |
| NMP | Nucleotide Mono Phosphate |
| ONT | Oxford Nanopore Technology |
| PacBio | Pacific Biosciences |
| PCR | Polymerase Chain Reaction |
| PE | Paired end |
| PEG | Polyethylene Glycol |
| PPi | Inorganic Pyrophosphate |
| QC | Quality Control |
| QMS | Quality Management System |
| qPCR | Quantitative Polymerase Chain Reaction |
| RNA | Ribonucleic acid |
| RNA-Seq | RNA Sequencing |
| rRNA | ribosomal RNA |
| SARS-CoV-2 | Severe Acute Respiratory Syndrome Coronavirus 2 |
| SBS | Sequencing by Synthesis |

| | |
|---|---|
| SE | Single end |
| SMRT | Single Molecule Real Time |
| SNP | Single Nucleotide Polymorphism |
| SNV | Single Nucleotide Variant |
| SPRI | Solid-Phase Reversible Immobilization |
| SV | Structural Variant |
| TS | Targeted Sequencing |
| UV | Ultraviolet |
| VCF | Variant Call Format |
| VOC | Variant of Concern |
| VOI | Variant of Interest |
| WES | Whole Exome Sequencing |
| WGBS | Whole Genome Bisulfite Sequencing |
| WGS | Whole Genome Sequencing |
| WHO | World Health Organisation |
| WTS | Whole Transcriptome Sequencing |
| ZMW | Zero-Mode Waveguide |

# Table of Contents

# Chapter 1: Presentation Module

**This chapter provides training presentations on SARS-CoV-2 genome sequencing for the master trainers covering the following topics.**

- Introduction to Next-Generation Sequencing
- Next-Generation Sequencing - Workflow
- Introduction to Nanopore – Its Principle & Application
- SARS-CoV-2 Sequencing Workflow
- Nanopore Sequencing Instrument Setup
- SARS-CoV-2 Bioinformatics Analysis
- Tool Kit - A comprehensive manual for Next-Generation Sequencing with a focus on SARS-CoV-2.

# 1. INTRODUCTION

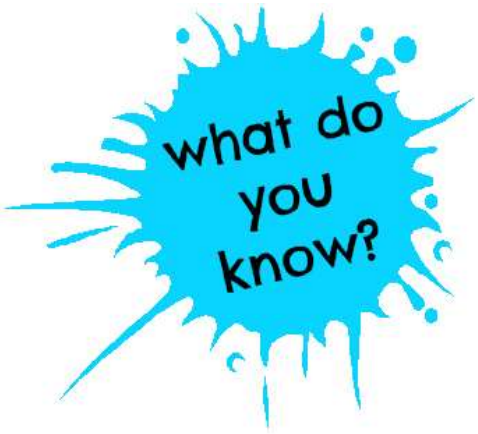
# TO

# NEXT-GENERATION SEQUENCING

## Concepts Covered

- What is NGS?
- DNA Sequencing Evolution
- Sequencing Methods
- Comparison of NGS Technologies
- Various Sequencing Platforms
- Traditional Sequencing vs NGS Sequencing
- Next-Generation Sequencing Workflow
- NGS Library Construction
- Types of Libraries
- Applications of NGS Sequencing
- Important Terminologies

## Next-Generation Sequencing (NGS)




## What is NGS?

Next-generation sequencing (NGS), high-throughput sequencing, or **massively parallel sequencing** are related terms that describe the DNA sequencing technology which has revolutionized the biological research. Characterized by **ultra-high throughput, scalability**, and speed, NGS **enables researchers to sequence a human genome within a single day,** which Sanger sequencing took over a decade to accomplish the final human genome draft.

NGS represents diverse modern high-throughput sequencing technologies, including Illumina sequencing, Ion Torrent sequencing, PACBIO sequencing and Nanopore Sequencing.

NGS platforms sequence numerous short reads in a single stroke. To do this, the input sample should first be cleaved into small fragments. The lengths of these sections depend on the sequencing instrument type and configuration used.

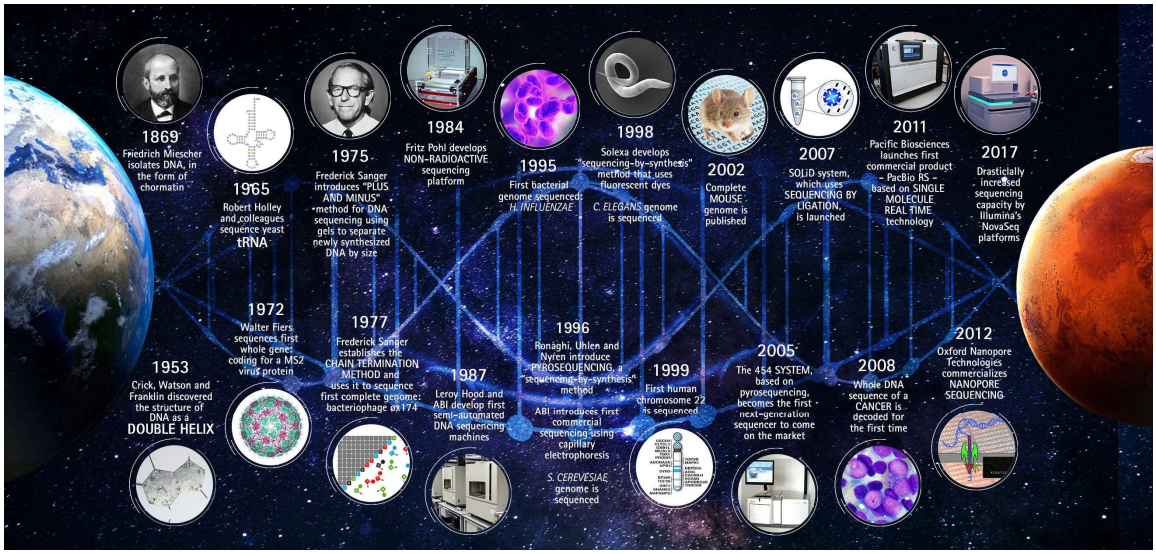
Adapted from Behjati S, Tarpey P S. What is next generation sequencing?. Archives of Disease in Childhood-Education and Practice, 2013, 98(6): 236-238

## "The Human Genome Project"

The Human Genome Project (HGP) was one of the great feats of exploration in history. The discovery was led by an international team of researchers looking to sequence and map all of the genes -- together known as the genome -- of members of our species, Homo sapiens. Beginning on October 1, 1990 and completed in April 2003, the HGP gave us the ability, for the first time, to read nature's complete genetic blueprint for building a human being.

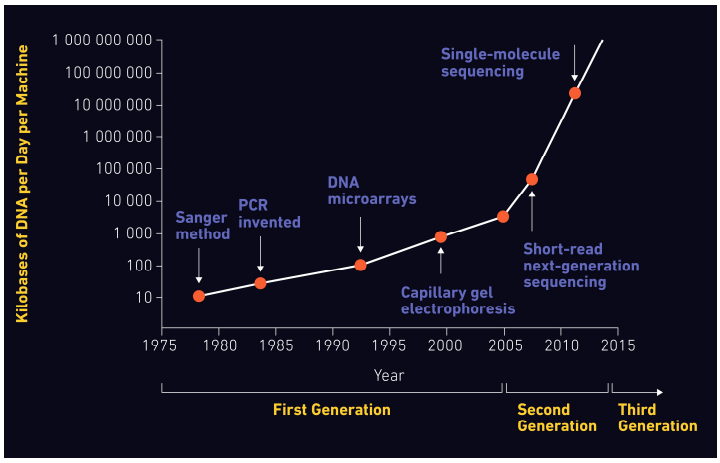
Courtesy: National Human Genome Research Institute

## DNA Sequencing Evolution



## Sequencing Methods

- First Generation Sequencing
  - Maxam-Gilbert Method
  - Sanger Sequencing
- Second Generation Sequencing
  - Roche 454 Sequencing
  - SOLiD
  - Ion Torrent
  - Illumina
- Third Generation Sequencing
  - PacBio-SMRT Sequencing
  - Oxford Nanopore Sequencing

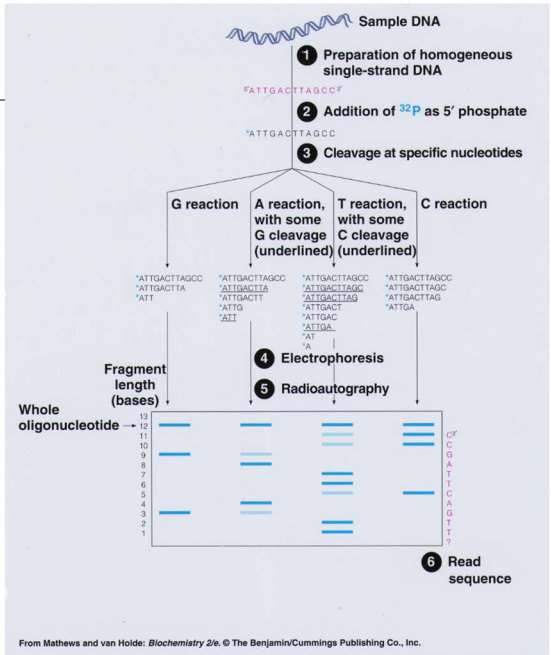



The evolution of sequencing methodologies

Image Credit: https://medicaltrend.org/2021/03/16/overview-of-next-generation-sequencing-technology/

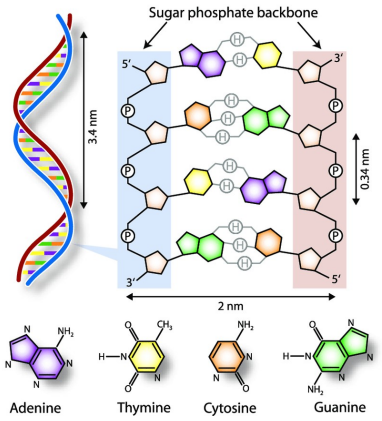
## Maxam Gilbert Method

- Maxam–Gilbert sequencing is a method of DNA sequencing developed by Allan Maxam and Walter Gilbert in 1976–1977.
- This method is based on nucleobase-specific partial chemical modification of DNA and subsequent cleavage of the DNA backbone at sites adjacent to the modified nucleotides.



## Sanger Sequencing – Amplicon PCR

- The DNA sequence of interest is used as a template for a special type of PCR called chain-termination PCR. Chain-termination PCR works just like standard PCR, but with one major difference: the addition of modified nucleotides (dNTPs) called dideoxyribonucleotides (ddNTPs).
- In the extension step of standard PCR, DNA polymerase adds dNTPs to a growing DNA strand by catalyzing the formation of a phosphodiester bond between the free 3'-OH group of the last nucleotide and the 5'-phosphate of the next.




Schematic representation of double-stranded DNA

Adapted from *Gauthier, Michel. (2007)*

## Sanger Sequencing – Cycle Sequencing PCR

- In chain-termination PCR, low ratio of chain-terminating ddNTPs are mixed with the normal dNTPs in the PCR reaction. ddNTPs lack the 3'-OH group required for phosphodiester bond formation; therefore, when DNA polymerase incorporates a fluorescently labelled ddNTP at random, extension ceases. The result of chain-termination PCR is millions to billions of oligonucleotide copies of the DNA sequence of interest, terminated at a random lengths (n) by 5'-ddNTPs.

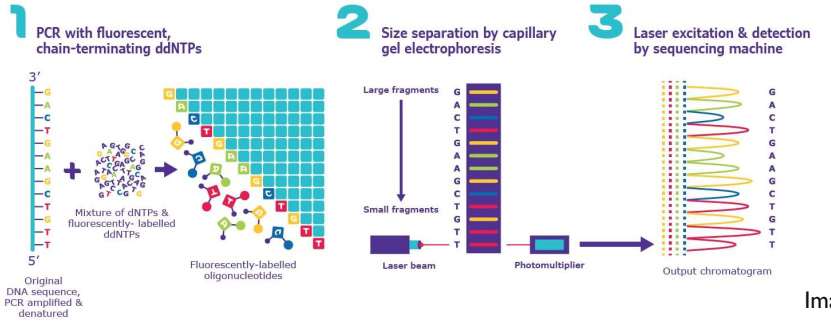
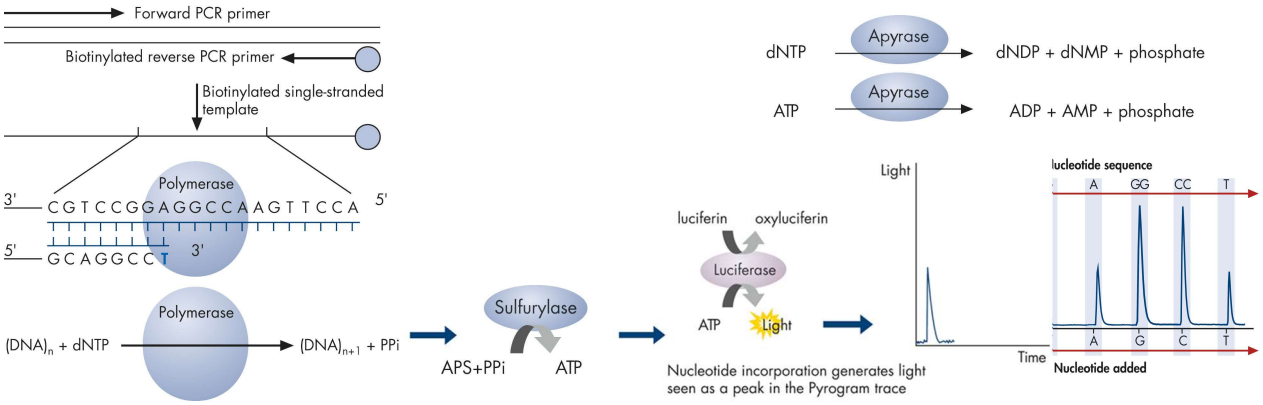



Image credit: www.sigmaaldrich.com

## Pyrosequencing Technology

The pyrosequencing technology is based on the sequencing-by-synthesis principle, where the detection of signal is achieved in real time by detecting the nucleotide incorporated by a DNA polymerase.




*Adapted from* www.Qiagen.com

## "Second" Generation Sequencing

After the original sanger methods using radioactivity and slab gels, the need for higher throughput sequencing of large genomes at lower cost triggered the development of many second-generation or "Nextgen" technologies using a variety of creative methodologies in addition to automated sanger sequencing.

## 2nd Gen NGS – The future in 1990's

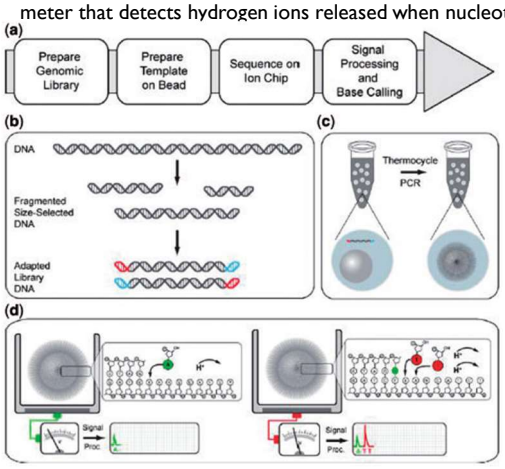
Second generation sequencing methods can be grouped into two major categories: sequencing by hybridization and sequencing by synthesis (SBS).

SBS methods are a further development of Sanger sequencing, without the dideoxy terminators, in combination with repeated cycles of synthesis, imaging, and methods to incorporate additional nucleotides in the growing chain. At first glance, these new methods may seem expensive, but the reactions are run in parallel often at nanoliter, picoliter, or zeptoliter volumes in small chambers, and thus the cost per base pair sequenced is nominal.

A note about costs: Costs for sequencing encompass many variables, some of which are often left out of commonly presented estimates of "cost per base". Costs usually do not include labor and the bioinformatics pipeline at the end of the process. Nevertheless, goals such as the "$1,000 human genome" or reducing the "cost per base" are gold standards to be met by the sequencing technology and research community.

## Ion Torrent Sequencing

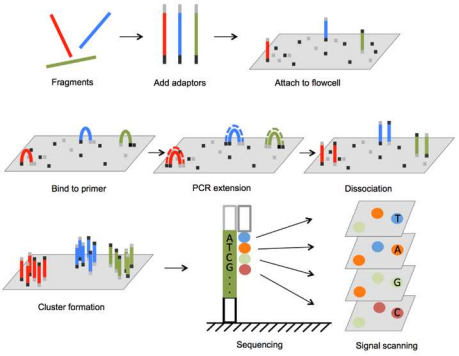
Ion Torrent uses Ion Semiconductor Sequencing Technology. The Ion Torrent chip is an ultrasensitive pH meter that detects hydrogen ions released when nucleotides are incorporated during DNA synthesis.




https://www.youtube.com/watch?v=zBPKj0mMcDg

## Illumina Sequencing

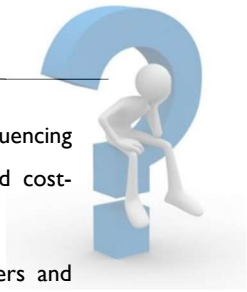
The Illumina next-generation sequencing (NGS) method is based on sequencing-by-synthesis (SBS), and reversible dye-terminators that enable the identification of single bases as they are introduced into DNA strands.




Adapted from http://en.biomarker.com.cn/platforms/illumina

https://www.youtube.com/watch?v=fCd6B5HRaZ8

## 2nd Generation(2G) NGS Limitations



2G NGS technologies in general offer several advantages over alternative sequencing techniques, including the ability to generate sequencing reads in a fast, sensitive and cost-effective manner.

However, there are also disadvantages, including poor interpretation of homopolymers and incorporation of incorrect dNTPs by polymerases, resulting in sequencing errors.

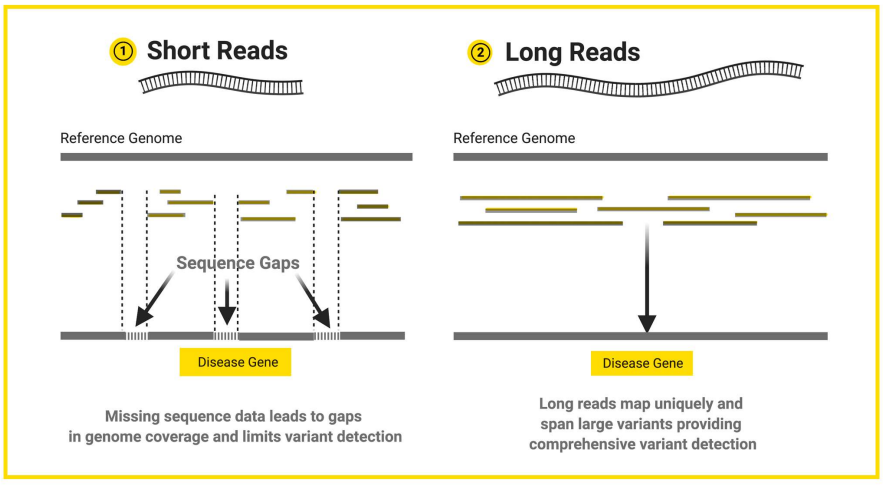
The short read lengths also create the need for deeper sequencing coverage to enable accurate contig and final genome assembly.

The major disadvantage of all 2G NGS techniques is the need for PCR amplification prior to sequencing. This is associated with PCR bias during library preparation (sequence GC-content, fragment length and false diversity) and analysis (base errors or favoring certain sequences over others).

## "Third" Generation Sequencing
## (Large Fragment Single Molecule)

In contrast to second generation sequencing methods, third generation sequencing methods aim to sequence long DNA (and RNA) molecules. This is fueled in part by cost per reaction and in part by the desire to obtain as much primary sequence read information as possible to circumvent sequence context issues such repeated DNA elements.

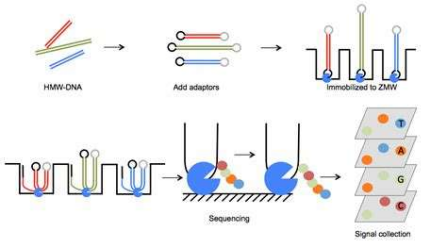
## Short Reads vs Long Reads




*Adapted from* Sarah Sharman (2021). Piecing together the genome: the long and short of it all. https://www.hudsonalpha.org/piecing-together-the-genome-the-long-and-short-of-it-all/

## Pacbio SMRT Sequencing

- PacBio long-read sequencing enabled by SMRT (single molecule real time) Sequencing technology requires no PCR amplification and the read length is 100 times longer than that of NGS.
- PacBio SMRT Sequencing uses the innovation of ZMW (Zero-Mode Waveguide) to distinguish the ideal fluorescent signal from the strong fluorescent backgrounds caused by unincorporated free-floating nucleotides.
- The binding of a DNA polymerase and the template DNA strand is anchored to the bottom glass surface of a ZMW.
- Laser light travels through the bottom surface of a ZMW and not completely penetrates it, since the ZMW dimensions are smaller than the wavelength of the light.
- Therefore, it allows selective excitation and identification of light emitted from nucleotides recruited for base elongation.




https://www.youtube.com/watch?v=_ID8JyAbwEo

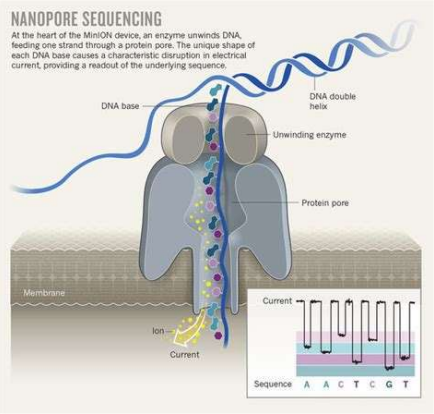
## Nanopore Sequencing

Determine the sequence of DNA fragments by passing DNA through a protein (or other) pore in a membrane.

Nanopore is a nano-scale hole. In nanopore sequencers, an ionic current passes through the nanopores.

When the DNA strand passes through the nanopore, the sequencer measures the change in current.

This change is used to identify the bases in the strand with the help of different electrochemical structures of the different bases.




https://www.youtube.com/watch?v=RcP85JHLmnI

## Various Sequencing Platforms

Illumina Miseq

Illumina NextSeq

Illumina iSeq 100

Illumina Hiseq X

Illumina Novaseq 6000

Ion Torrent PGM

Ion Torrent S5

PacBio

Nanopore Minion

Nanopore Gridion

Image Credit: Illumina Inc, Thermo Scientific, Inc, Pacbio Inc, Nanopore Technologies

## Comparison of NGS Platforms

| | Illumina | | | | | | Thermofisher | | | PacBio | | | ONT | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **System Platform** | iSeq | MiniSeq | MiSeq | NextSeq 550 | NextSeq 1000 & 2000 | NovaSeq 6000 | GeneStudio S5 | Genexus | Ion PGM Dx | Sequel | Sequel II | Sequel IIe | Flongle | MinION | GridION | PromethION |
| **Sequencing principle** | Sequence by synthesis | | | | | | | | | PacBio single molecule sequencing | | | Nanopore single molecule sequencing | | | |
| **Detection** | Fluorescent | | | | | | Ion | | | Fluorescent | | | Electrical conductivity | | | |
| **Read length** | 2 x 150 bp | | 2 x 300 bp | 2 x 150 bp | 2 x 150 bp | 2 x 250 bp | 600 bp | 400 bp | 200 bp | 300 kb | | | Longest read so far > 4 Mb | | | |
| **Output data per run** | 1.2 Gb | 7.5 Gb | 15 Gb | 120 Gb | 330 Gb | 3000 Gb | 15 Gb | 24 Gb | 1 Gb | 75 Gb | 600 Gb | | 1-2 Gb | 10-30-50 Gb | | 100-200-300 Gb |
| **Sequencing run time** | 9.5-19 hr | 5-24 hr | 4-56 hr | 11-29 hr | 11-48 hr | 13-44 hr | 4.5-21.5 hr | 14-31 hr | 4.4 hr | Up to 20 hr | Up to 30 hr | | 1 min-16 hr | 1 min - 72 hr | | |
| **Accuracy** | Q30 ≥ 75% | | | | | | ≥99% | ≥99% | ≥99% | HIFI reads >99% accuracy | | | R9 accuracy: >98.3%, R10 accuracy: >97.5%, new chemistry accuracy: >99%, Consensus R9.4.1: Current best Q45 (>99.99%), R10: current best Q50 (99.999%) | | | |
| **Equipment cost (USD)** | $19,900 | $49,500 | $99,000 | $275,000 | $335,000 | On request | | | | $525,000 | | | $1,460 | - | $69,955 | 24 flow cells: $335,455 48 flow cells: $530,000 |
| **Advantages** | Highest NGS accuracy, wide application, high throughput | | | | | | Short run times, low cost, no fluorescent labelling, no optics | | | Long reads resolve ambiguities, no DNA amplification required, fast turnaround time | | | Very long reads resolve ambiguities, no DNA amplification required, low cost, simple preparation, fast sequencing, portability | | | |
| **Disadvantages** | Complex library prep, long run time, short read lengths, poor resolution of structural variants and repetitive regions | | | | | | Higher error rate in homopolymer regions | | | Expensive sequencing equipment, diffult installation | | | Higher error rates, dynamic improvement | | | |

*Note*. Adapted from Hu, T., Chitnis, N., Monos, D., & Dinh, A. (2021). Next-generation sequencing technologies: An overview. Hum Immunol, 82(11), 801-811. https://doi.org/10.1016/j.humimm.2021.02.012

## Sequencing Methods (Recap)

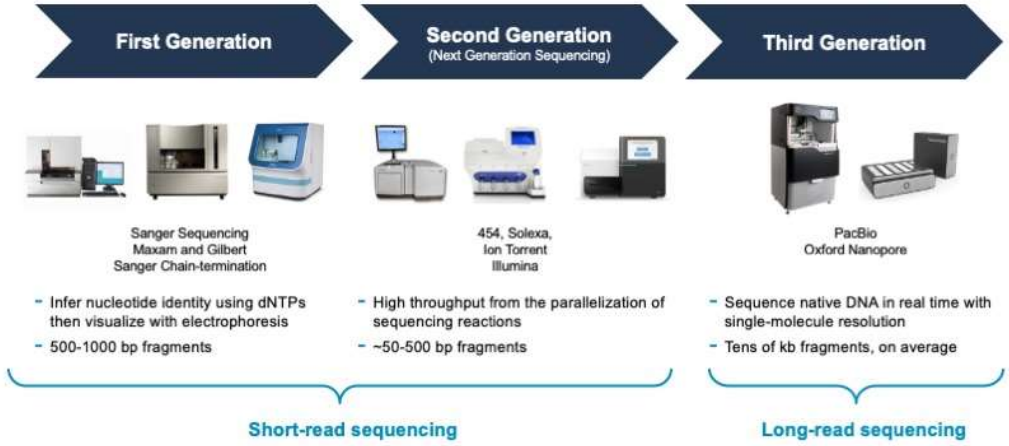
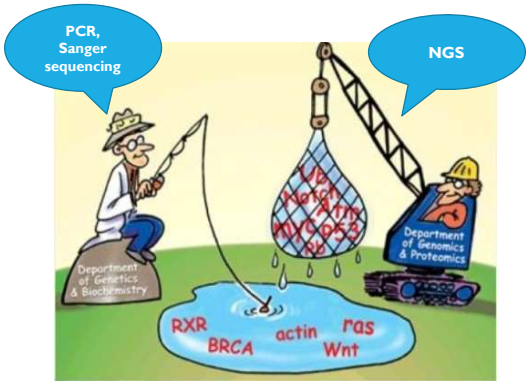



*Image adapted from* https://www.pacb.com/blog/the-evolution-of-dna-sequencing-tools/

## Traditional Sequencing vs Next-Generation Sequencing



## Key differences between NGS and Sanger sequencing

- In principle, the concepts behind Sanger vs. next-generation sequencing (NGS) technologies are similar. In both NGS and Sanger sequencing (also known as dideoxy or capillary electrophoresis sequencing), DNA polymerase adds fluorescent nucleotides one by one onto a growing DNA template strand. Each incorporated nucleotide is identified by its fluorescent tag.

- The critical difference between Sanger sequencing and NGS is sequencing volume. While the Sanger method only sequences a single DNA fragment at a time, NGS is massively parallel, sequencing millions of fragments simultaneously per run. This process translates into sequencing hundreds to thousands of genes at one time. NGS also offers greater discovery power to detect novel or rare variants with deep sequencing.

*Adapted from* https://sapac.illumina.com/science/technology/next-generation-sequencing/ngs-vs-sanger-sequencing.html

## Sanger, Real Time PCR, NGS Comparison of Technologies

| Sanger | Real time PCR | Targeted Sequencing - NGS |
|---|---|---|
| Low Sensitivity (>20%) | High sensitivity upto 0.1% | High sensitivity (>1%) |
| High input DNA | Low input DNA | Low input DNA |
| Low discovery power and resolution | High discovery power and resolution | High discovery power and resolution |
| Low scalability (< 20 genes) | Low scalability, can detect only one marker at a time, multiplexing is limited to 3 or 4 markers per reaction | High Scalability >20 genes |
| Not as cost effective for multiple genes. However good to go with single gene testing. | Cost effective with high sensitivity and short turn around time for quick decision making in treatment management. Addition of multiple markers adds cost to the testing which is a limitation. | Highly cost effective for multiple genes and multiple mutation screening |
| Detection of tumor heterogeneity and novel mutations is limited by sensitivity | Limitations: Cannot detect novel mutations | Highly sensitivity in detection of tumor heterogeneity and novel mutations – limited with longer turn around time as compared to Realtime PCR for detection of specific set of mutations, although highly cost effective |

## Applications of NGS Sequencing: Unlimited Opportunities

Whole-Genome Sequencing

Exome Sequencing

De novo Sequencing

Targeted Sequencing

Total RNA & mRNA Sequencing

Targeted RNA Sequencing

Small RNA & Non coding RNA Sequencing

Methylation Sequencing

ChIP Sequencing

## Question & Answer Session

1) What is the principle behind Sanger Sequencing method?

a) chemical degradation method

b) dideoxy chain termination method

c) emulsion PCR method

2) How original sanger sequencing method was improvised in the 2nd Generation?

a) by parallelly doing multiple slab gels

b) by employing more lab technicians

c) by changing conventional method to automated method

3) Which one was the first commercial sequencer?

a) Ion Torrent

b) Illumina

c) SOLiD

d) None of the above

## Question & Answer Session

4) What is the abbreviation for HGP?

a) Human genetics & proteins

b) Human genome proportion

c) Human genome project

5) How Illumina works?

a) Sequencing by Synthesis

b) Sequencing using Ion Semiconductor Chip

c) Sequencing using ZMW

6) What are the improvements in 3rd Generation Technology?

a) Fragment length

b) Cost per base

c) High quality bases

d) More number of bases sequenced per second

## References


1. Behjati S, Tarpey P S. What is next generation sequencing?. Archives of Disease in Childhood-Education and Practice, 2013, 98(6): 236-238
2. A Journey Through The History Of DNA Sequencing. https://the-dna-universe.com/2020/11/02/a-journey-through-the-history-of-dna-sequencing/
3. Slatko, B. E., Gardner, A. F., & Ausubel, F. M. (2018). Overview of Next-Generation Sequencing Technologies. Current protocols in molecular biology, 122(1), e59. https://doi.org/10.1002/cpmb.59
4. Gauthier, Michel. (2007). Simulation of polymer translocation through small channels: A molecular dynamics study and a new Monte Carlo approach.
5. Sanger Sequencing Steps & Method. https://www.sigmaaldrich.com/IN/en/technical-documents/protocol/genomics/sequencing/sanger-sequencing, last accessed 30-Apr-2023.
6. Pyrosequencing, QIAGEN. https://www.qiagen.com/us/home/resources/technologies/pyrosequencing-resource-center/#:~:text=Pyrosequencing%20-%20QIAGEN%20Pyrosequencing%20Pyrosequencing%20%E2%80%94%20the%20unique,enables%20rapid%20and%20accurate%20quantification%20of%20sequence%20variation

## References


7. Slatko BE, Gardner AF, Ausubel FM. Overview of Next-Generation Sequencing Technologies. Curr Protoc Mol Biol. 2018 Apr;122(1):e59. doi: 10.1002/cpmb.59. PMID: 29851291; PMCID: PMC6020069.
8. Golan, David & Medvedev, Paul. (2013). Using state machines to model the Ion Torrent sequencing process and to improve read error rates. Bioinformatics (Oxford, England). 29. i344-i351. 10.1093/bioinformatics/btt212.
9. Athina Gkazi, PhD (2021). An Overview of Next-Generation Sequencing. https://www.technologynetworks.com/genomics/articles/an-overview-of-next-generation-sequencing-346532
10. Sarah Sharman (2021). Piecing together the genome: the long and short of it all. https://www.hudsonalpha.org/piecing-together-the-genome-the-long-and-short-of-it-all/
11. Overview of PacBio SMRT sequencing: principles, workflow, and applications. https://www.cd-genomics.com/pacbio-smrt-system-single-molecule-real-time-sequencing.html#:~:text=PacBio%20SMRT%20sequencing%20requires%20no%20PCR%20amplification%2C%20can,read%20length%20is%208-15kb%20and%20up%20to%2040-70kb.
12. Untergasser, Gerold & Bucher, Philipp & Dresch, Philipp. (2019). Metagenomics Profiling of Tumours Using 16S-rRNA Amplicon Based Next Generation Sequencing.


## References


13. Hu, T., Chitnis, N., Monos, D., & Dinh, A. (2021). Next-generation sequencing technologies: An overview. Hum Immunol, 82(11), 801-811. https://doi.org/10.1016/j.humimm.2021.02.012
14. Differences Between NGS and Sanger Sequencing. https://sapac.illumina.com/science/technology/next-generation-sequencing/ngs-vs-sanger-sequencing.html
15. Akalin, A. Computational Genomics with R. (2020). https://compgenomr.github.io/book/fasta-and-fastq-formats.html

# 2. Next-Generation Sequencing – Workflow

## Concepts Covered

**Next-Generation Sequencing Workflow**

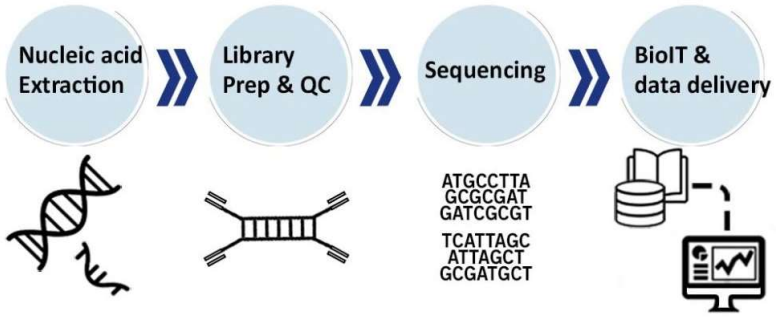
**NGS Library Construction**

**Types of Libraries**

**Applications of NGS Sequencing**

**Important Terminologies in NGS**

## Next-Generation Sequencing Workflow




## NGS Library Construction

The workflow for library construction involves the following steps:

- Determine the quality of genomic DNA (gDNA) using fluorometric assay
- Shear gDNA using a mechanical/enzymatic/chemical method
- Repair DNA damage and ends of fragmented DNA
- Adapter ligation using adapters
- Purify & Size select (optional) adapter ligated library.
- PCR enrichment & submission to a sequencer

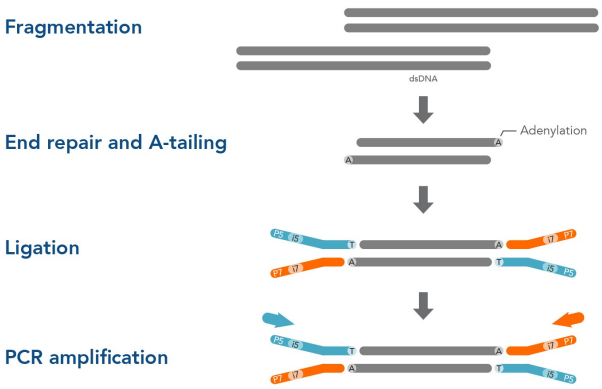



Image Credit: Illumina, Inc

## Types of Libraries

**Genome Libraries**

- Whole Genome
- WG Bisulphite

**Exome Libraries**

- Whole Exome
- Targeted Exome

**Amplicon Libraries**

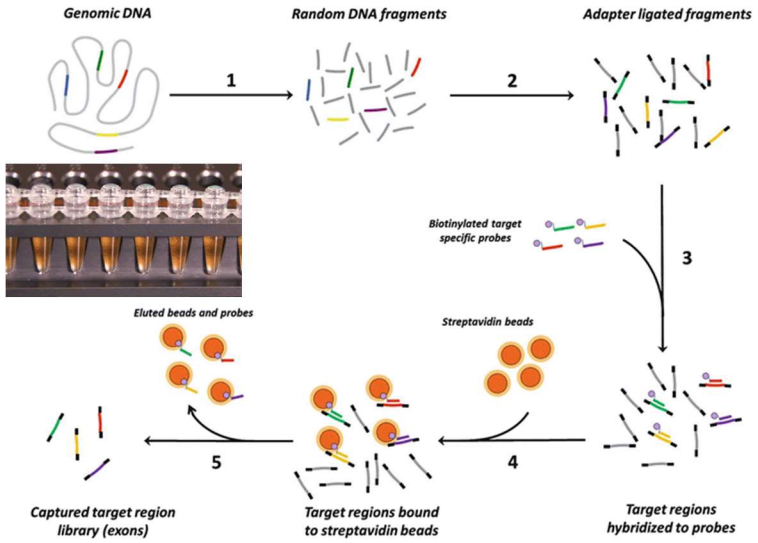
**cDNA Libraries**

- Whole Transcriptome
- mRNA
- Small RNA

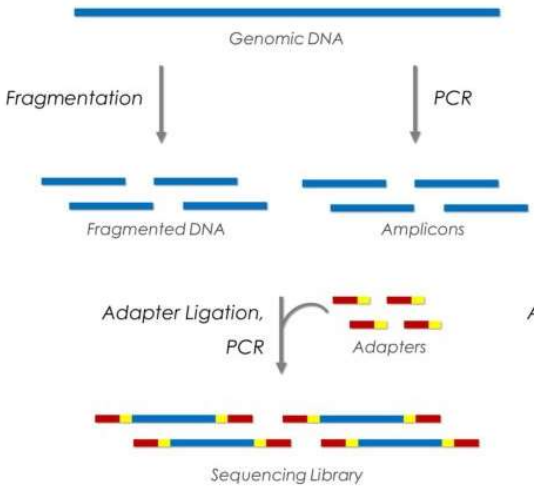
**Metagenome libraries**

- V3-V4
- ITS

## Exome Library - Capture / Hybridization based

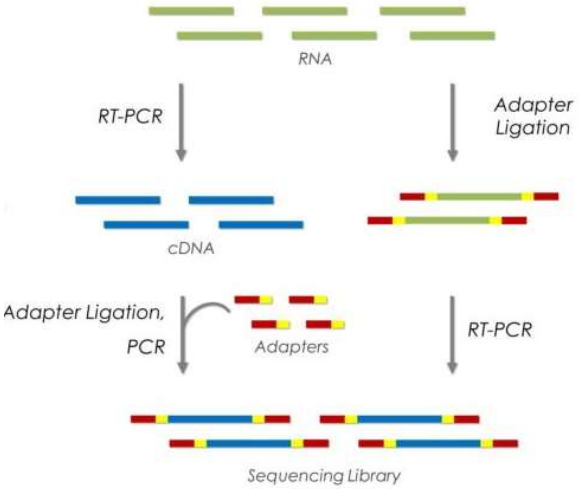


## Amplicon Library – PCR based




*Adapted from* https://www.biocompare.com/Molecular-Biology/9187-Next-Generation-Sequencing/

## cDNA Libraries



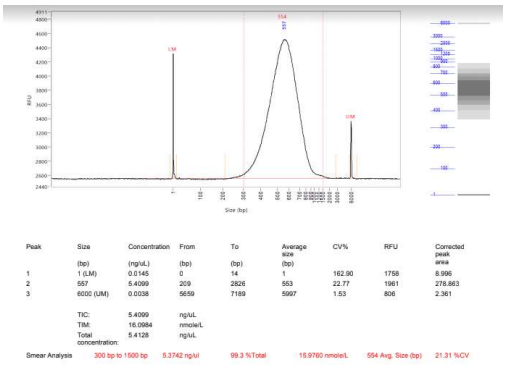

*Adapted from* https://www.biocompare.com/Molecular-Biology/9187-Next-Generation-Sequencing/

## Final Library QC

1. **Quantity check by Qubit Fluorometer**

   Fluorometers detect fluorescent dyes specifically bound to the target molecule. They can distinguish dsDNA from ssDNA or intact from degraded RNA, even in extremely small amounts or in the presence of contaminants.

2. **Quality check by capillary electrophoresis** for fragment size determination.



| Peak | Size (bp) | Concentration (ng/uL) | From (bp) | To (bp) | Average size (bp) | CV% | RFU | Corrected peak area |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 (LM) | 0.0145 | 0 | 14 | 1 | 162.90 | 1758 | 8.996 |
| 2 | 557 | 5.4099 | 209 | 2826 | 553 | 22.77 | 1961 | 278.863 |
| 3 | 6000 (UM) | 0.0038 | 5659 | 7189 | 5997 | 1.53 | 806 | 2.361 |
| | TIC: | 5.4099 | ng/uL | | | | | |
| | TIM: | 16.0984 | nmole/L | | | | | |
| | Total concentration: | 5.4128 | ng/uL | | | | | |
| Smear Analysis | 300 bp to 1500 bp | 5.3742 ng/ul | 99.3 %Total | 15.9760 nmole/L | 554 Avg. Size (bp) | 21.31 %CV | | |

Library Profile
(output from fragment analyzer instrument)

## Important Terminologies in NGS

## Important Terminologies

**Library**

**Adapter**

**Barcoding/Indexing**

**SPRI Beads**

**Flow cell**

**Base Call files & Raw Data**

**Phred Quality Score**

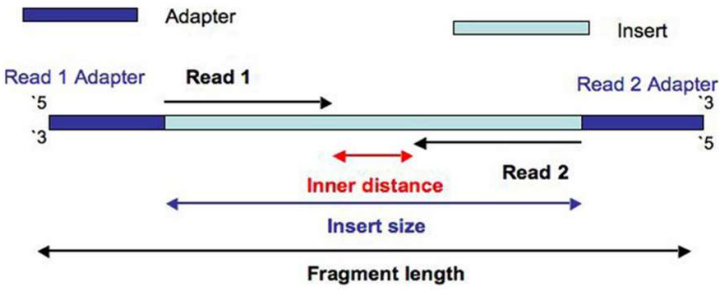
**Depth & Coverage**

**Percent on-target reads**

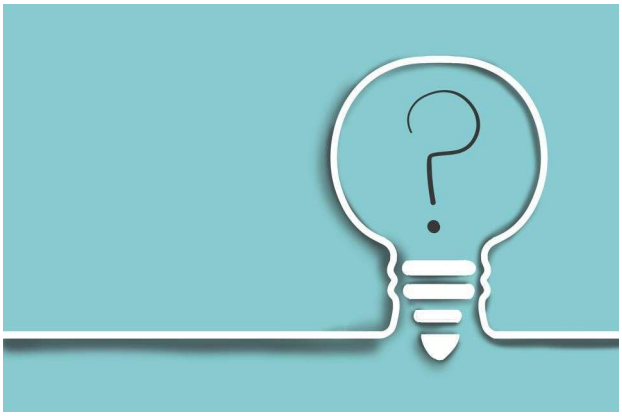
**Amplicon dropouts**

## Library

Library is a population of DNA fragments of defined lengths with defined oligomer sequences at both ends to be compatible with the applied sequencing technique.

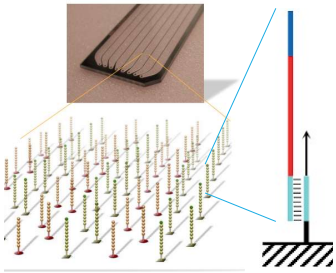


## How do you sequence random short fragments?



## Adapters

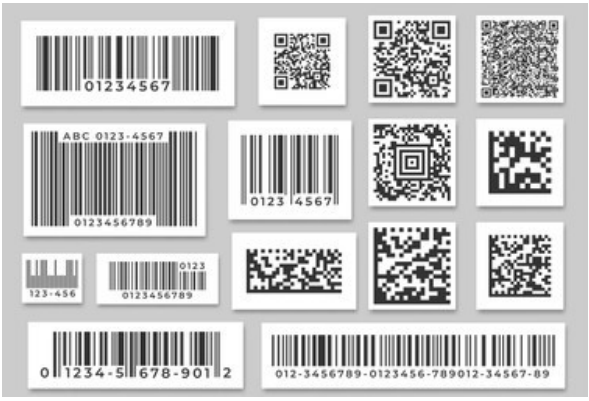
- An adapter or adaptor is a short, chemically synthesized, single-stranded or double-stranded oligonucleotide that can be ligated to the ends of other DNA or RNA fragments of interest.
- Double stranded adapters can be synthesized to have blunt ends to both terminals or to have sticky end at one end and blunt end at the other.
- Two adapters could base pair to each other to form dimers.



Flow cell coated with a lawn of Oligos complementary to library adapters


Image Credit: Illumina, Inc

## What is the purpose of Barcodes?




## Barcoding/ Indexing of libraries - Sample Multiplexing

Sample multiplexing, also known as multiplex sequencing, allows large numbers of libraries to be pooled and sequenced simultaneously during a single run on sequencing instruments.

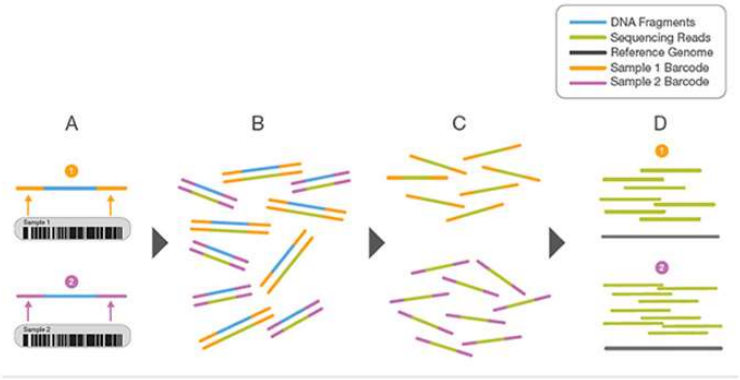
Sample multiplexing is useful when targeting specific genomic regions or working with smaller genomes.

Pooling samples exponentially increases the number of samples analyzed in a single run, without drastically increasing cost or time.

With multiplex sequencing, individual "barcode" sequences are added to each DNA fragment during next-generation sequencing (NGS) library preparation so that each read can be identified and sorted before the final data analysis.

These barcodes, or index adapters, can follow one of two major indexing strategies depending on your library prep kit and application.

## Multiplex Sequencing Highlights




A. Two representative DNA fragments from two unique samples, each attached to a specific barcode sequence that identifies the sample from which it originated.

B. Libraries for each sample are pooled and sequenced in parallel. Each new read contains both the fragment sequence and its sample-identifying barcode.

C. Barcode sequences are used to de-multiplex, or differentiate reads from each sample.

D. Each set of reads is aligned to the reference sequence.

- Fast high-throughput strategy: Large sample numbers can be simultaneously sequenced during a single experiment
- Cost-effective method: Sample pooling improves productivity by reducing time and reagent use
- Simplified analysis: Automatic sample identification with "barcodes" using data analysis software

Image Credit: Illumina, Inc

## What are the ways to purify the DNA or Amplified product?

## SPRI Beads

Solid Phase Reversible Immobilization beads were developed at the Whitehead Institute (DeAngelis et al 1995) for purification of PCR amplified colonies in the DNA sequencing group.

SPRI beads are paramagnetic (magnetic only in a magnetic field) and this prevents them from clumping and falling out of solution. Each bead is made of polystyrene surrounded by a layer of magnetite, which is coated with carboxyl molecules.

It is these that reversibly bind DNA in the presence of the "crowding agent" polyethylene glycol (PEG) and salt (20% PEG, 2.5M NaCl is the magic mix).

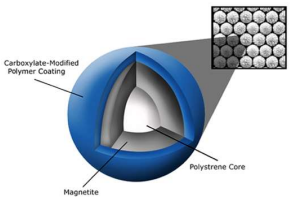


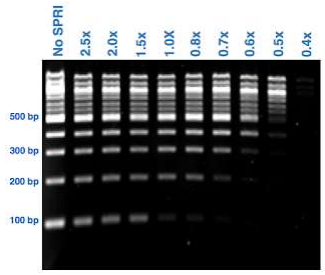

Image Credit: Beckman Coulter, Inc

## SPRI Beads

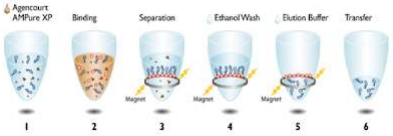
PEG causes the negatively-charged DNA to bind with the carboxyl groups on the bead surface. As the immobilization is dependent on the concentration of PEG and salt in the reaction, the volumetric ratio of beads to DNA is critical.

SPRI is great for low concentration DNA cleanup that is why it is used in so many kits.

The reagents are easy to handle, and a user can process 96 samples very easily in a standard plate. Alternatively, the protocol can be easily automated, and tens or hundreds of plates can be run on a robot in a working day. The binding capacity of SPRI beads is huge. 1ul of AmpureXP will bind over 3µg DNA.




SPRI size selection from Broad "boot camp"




typical SPRI protocol

Image Credit: Beckman Coulter, Inc

## Where does the Sequencing Happens?

## Flowcell

An Illumina flow cell is a hollow glass slide with one or more channels ("lanes"), coated with oligonucleotides which are complementary to the sequencing adapters so that single-stranded, adapter-ligated DNA fragments can attach through hybridization.

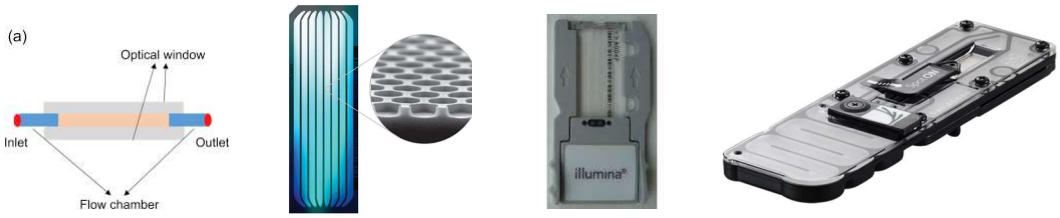


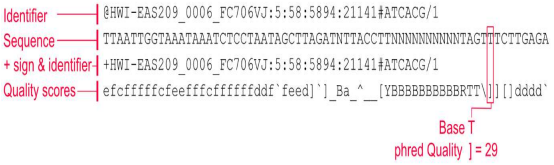

Image Credit: Illumina Inc & Oxford Nanopore Technologies

## Base Call files & Raw Data

The raw output of all sequencing machines varies by format. For Illumina, it is the . bcl format whereas for Nanopore, FAST5 file. These files are named after, and represent base calls per cycle, which is a binary file that contains both the base call and the quality of that base call.

Base call files are further processed, demultiplexed using barcode sequence assigned at the time of library preparation (Sample sheet) and the sample wise data is stored in the FASTQ file format.

**FASTQ** files have four lines per sequence.
**Line 1** begins with a '@' character and is followed by a sequence identifier and an optional description.
**Line 2** is the sequence.
**Line 3** begins with a '+' character and is optionally followed again by the same sequence identifier and description.
**Line 4** encodes the quality values for the sequence, which corresponds to a quality score.




*Note*. Reprinted from Akalin, A. Computational Genomics with R. (2020). https://compgenomr.github.io/book/fasta-and-fastq-formats.html License: CC BY-NC-SA 4.0

## Phred Quality Score

A next-generation sequencing experiment consists of a series of discrete steps that uniquely contribute to the overall quality of a data set.

Sequencing quality metrics can provide important information about the accuracy of each step in this process, including library preparation, base calling, read alignment, and variant calling.

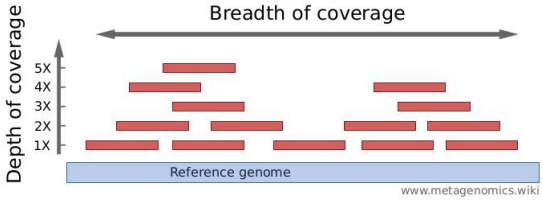
Base calling accuracy, measured by the Phred quality score (Q score), is the most common metric used to assess the accuracy of a sequencing platform. It indicates the probability that a given base is called incorrectly by the sequencer.

| Phred Quality Score | Probability of Incorrect Base Call | Base Call Accuracy |
|---|---|---|
| 10 | 1 in 10 | 90% |
| 20 | 1 in 100 | 99% |
| 30 | 1 in 1,000 | 99.9% |
| 40 | 1 in 10,000 | 99.99% |
| 50 | 1 in 100,000 | 99.999% |

Quality Scores & Base Calling Accuracy

## Depth & Coverage

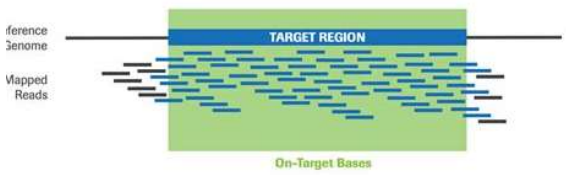
- Depth of Coverage - Refers to the number of times a nucleotide is read during sequencing. A greater depth of coverage can increase confidence in the final results. Deep coverage aids in differentiating sequencing errors from single nucleotide polymorphisms.
- Breadth of coverage refers to the number of times sequenced nucleotide bases covered the target genome. For example, if genome size is 100 Mbp and you have sequenced 5 M reads of 100 bp size, then sequencing coverage at genome level would be 5X.




## Percent on-target reads

The On-target rate is commonly used terminology to measure sensitivity (% target bases in sequence reads) and specificity (% of sequence reads on target site) in targeted enrichment NGS methods such as whole exome sequencing (WES).

On-target rate refers to how many nucleotide bases or reads are covered at the target site. On-target rate is determined by percent on-target bases (percentages of nucleotide bases mapped to the target region of the genome) and percent on-target reads (percentages of sequencing reads that covers the target region of the genome).



## Lineage, Variant & Mutation

- Lineage: A lineage is a group of closely related Organism with a common ancestor. Example: SARS-CoV-2 has many lineages; all cause COVID-19.
- Variant: A variant is a viral genome (genetic code) that may contain one or more mutations. In some cases, a group of variants with similar genetic changes, such as a lineage or group of lineages, may be designated by public health organizations as a variant of concern (VOC) or a variant of interest (VOI) due to shared attributes and characteristics that may require public health action.
- Mutation: A mutation is defined as any change in a DNA sequence away from normal. This implies there is a normal allele that is prevalent in the population and that the mutation changes this to a rare and abnormal variant.

## Polymorphism

- In contrast, a polymorphism is a DNA sequence variation that is common in the population. In this case no single allele is regarded as the standard sequence. Instead there are two or more equally acceptable alternatives. The arbitrary cut-off point between a mutation and a polymorphism is 1 percent. That is, to be classed as a polymorphism, the least common allele must have a frequency of 1 per cent or more in the population. If the frequency is lower that this, the allele is regarded as a mutation.
- Polymorphic sequence variants usually do not cause overt debilitating diseases. Many are found outside of genes and are completely neutral in effect. Others may be found within genes, but may influence characteristics such as height and hair color rather than characteristics of medical importance.

## Amplicon Dropout

The failure of amplification of one or few of the target regions which may be due to either sequence independent factors or sequence variations. The presence of single nucleotide variants (SNVs) in the forward and/or reverse oligo primer binding sites may lead to the complete or partial "drop" out of amplification during the PCR process.

## References


1. Illumina, Inc. https://www.illumina.com/
2. Beckman Coulter, Inc. https://www.beckmancoulter.com/
3. Akalin, A. Computational Genomics with R. (2020). https://compgenomr.github.io/book/fasta-and-fastq-formats.html

# 3. Introduction to Nanopore - Its Principle & Application

## Concepts Covered

Nanopore Sequencing Technology

Nanopore Sensing & Sequencing Events

Features of the technology

Advantages

Applications of Nanopore

Few Scientific Reports

Potential Disadvantage and Limitations

List of Nanopore Sequencers

Flow cell format

What's inside the box

## Why Nanopore ? - A Game Changer

- Lab-in-a-suitcase: A portable (size of a stapler) sequencing instrument that can be deployed to remote and resource-limited locations.
- Analysis of anything, by anyone, anywhere.
- Ease of set up, real-time analysis and control over time-to-results, make a difference.
- No capital investment required
- Generates Ultra long reads

Image Credit: Oxford Nanopore Technologies

## Nanopore Sequencing Technology

- In nature the protein nanopores function as gateways between the two systems.
- Carefully engineered protein nanopores by mutating key residues in the barrel of the pore.

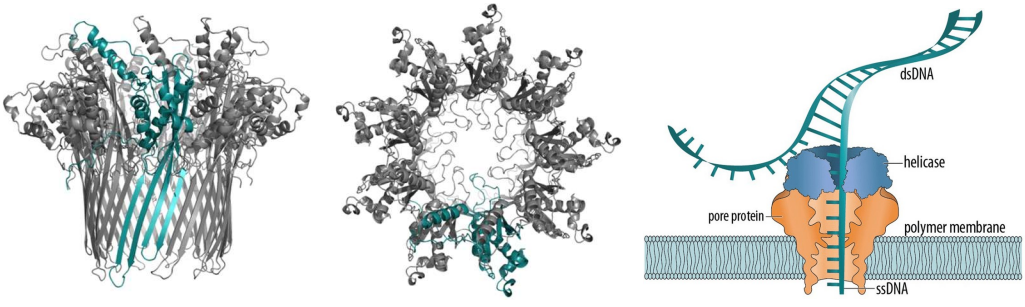
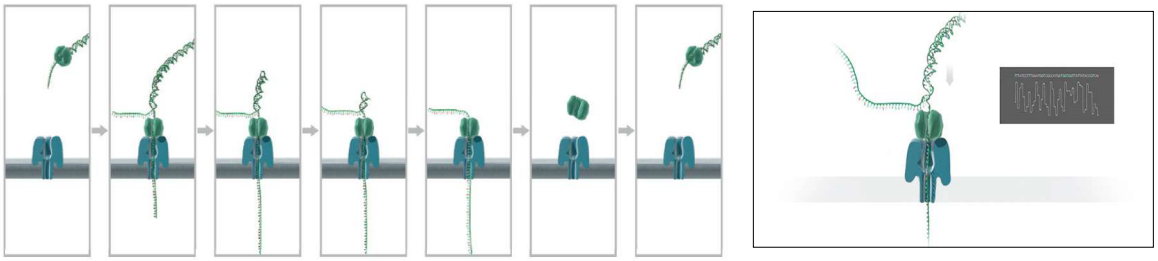



Image Credit: Oxford Nanopore Technologies

## Sequencing Events

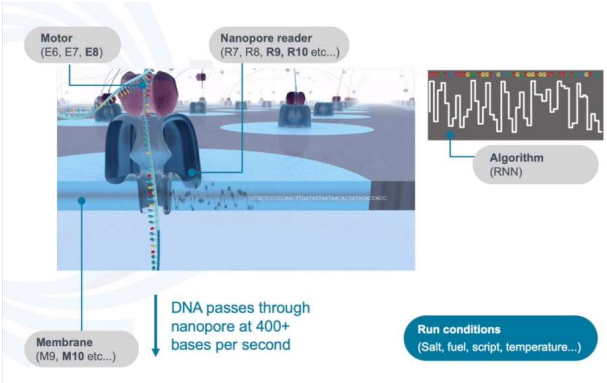


**Sequencing starts from the 5' end of the leader adapter.**
**The motor protein unwinds the dsDNA allowing single-stranded DNA to pass through the pore.**

Image Credit: Oxford Nanopore Technologies

## Nanopore DNA/RNA sequencing

DNA/RNA strand passes through the pore → signal interpreted into sequence data



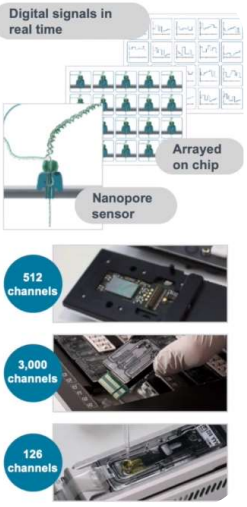




Image Credit: Oxford Nanopore Technologies

## Nanopore Sensing

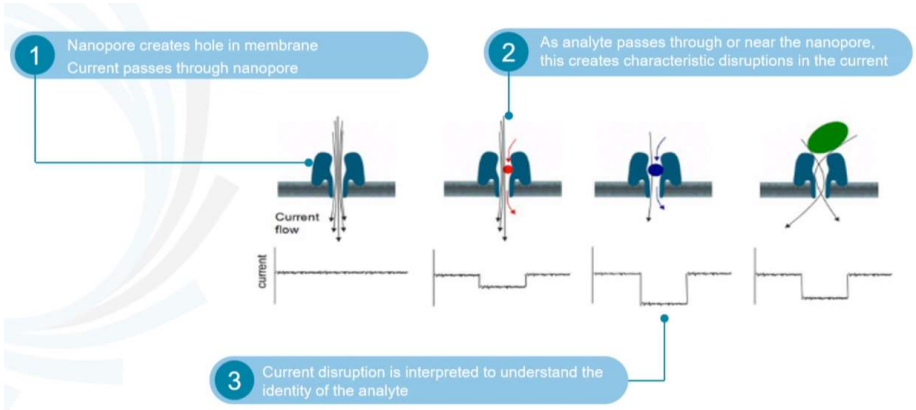
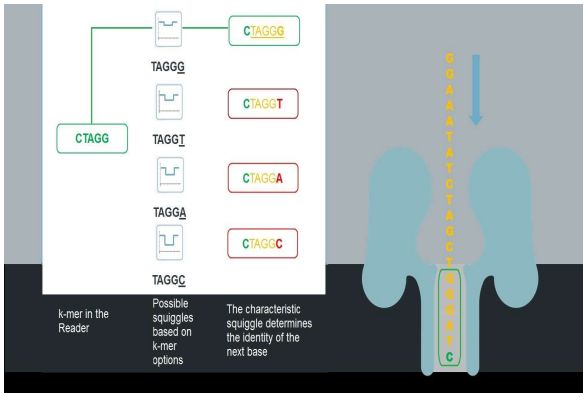



Image Credit: Oxford Nanopore Technologies

## Base Identification




**The nanopore houses the Reader that captures the signal contributed by the combination of nucleotides as the DNA moves through the pore.**

Image Credit: Oxford Nanopore Technologies

## Library preparation and multiplexing

**For DNA based applications**

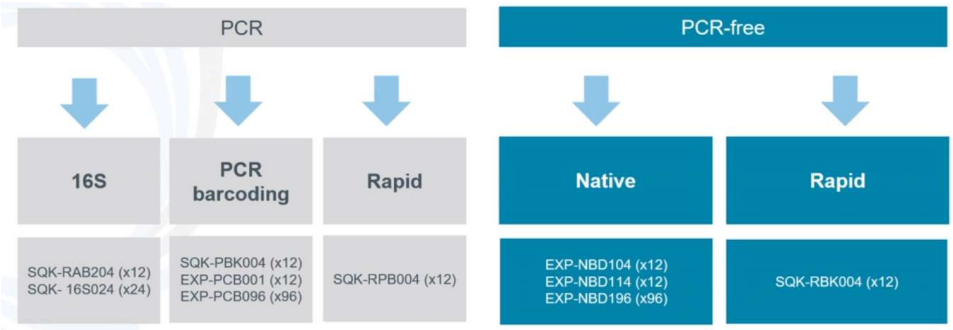



Image Credit: Oxford Nanopore Technologies

## Features of the Technology

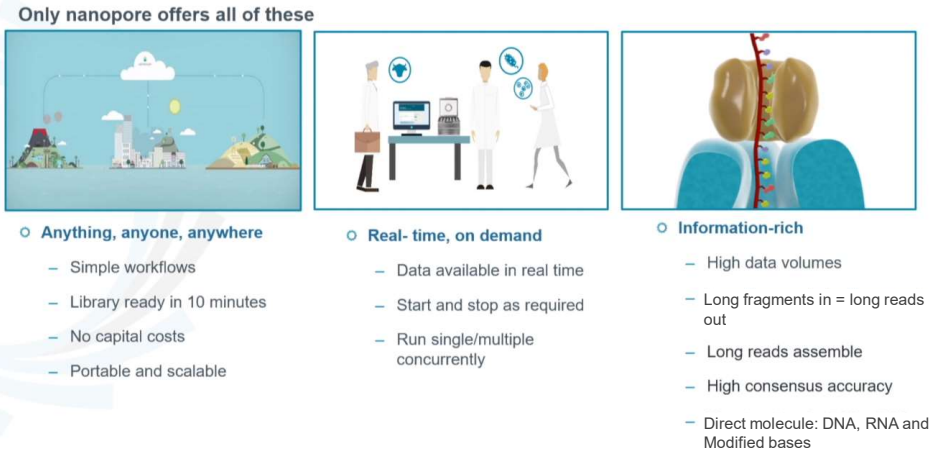
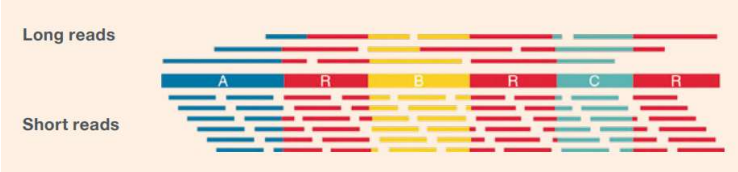



Image Credit: Oxford Nanopore Technologies

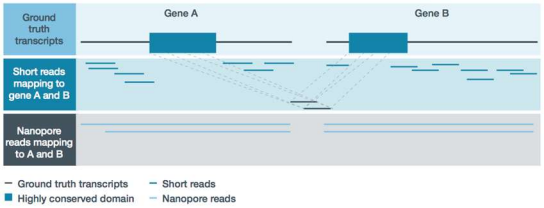
## Advantages of Long read Nanopore technology

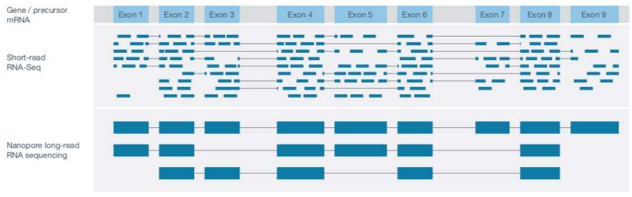
- ✓ offers simplified and less ambiguous genome assembly with the best assembly quality.
- ✓ easier assembly and the ability to span repetitive genomic regions
- ✓ Identification of large structural variation
- ✓ Generating complete, closed microbial genomes
- ✓ Identifying and quantifying individual microorganisms from mixed populations



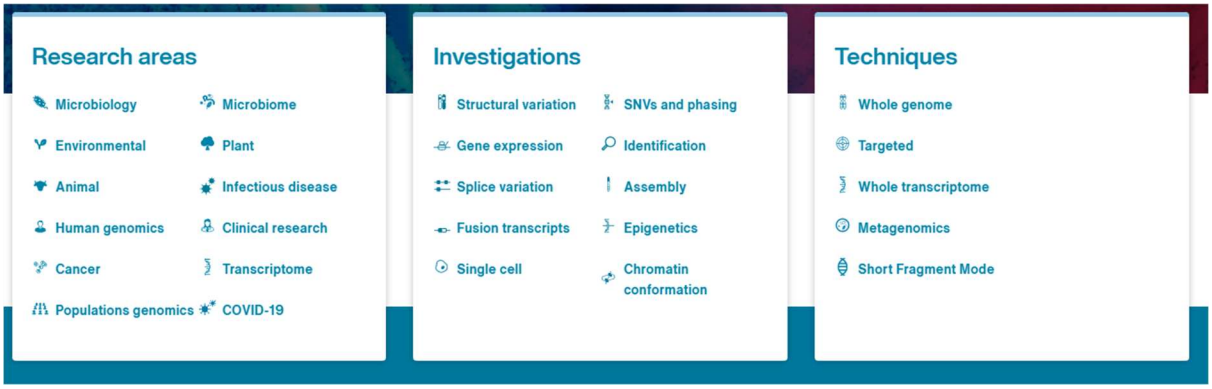

## Advantages in RNA Sequencing

- Nanopore sequencing for Transcriptome analysis
  a) Obtaining and analyzing full-length transcripts
  b) Accurate identification of multiple isoforms per gene
  c) Problematic isoform identification and quantification
- Long nanopore reads, can span full-length transcripts simplifying their identification.



## Nanopore Applications




## Nanopore Technology for array of applications

**DNA based**

1. De novo genome assembly
2. Structural variation detection
3. Comparative genome analyses
4. Metagenomic profiling.
5. Multiplexed amplicon sequencing
6. Long read and short read hybrid assembly
7. Ultra long read sequencing
8. 16s rRNA/ITS sequencing
9. Methylation analysis
10. Cell line characterization
11. HLA typing
12. Mitogenome sequencing

**RNA based**

1. Full length Transcriptome sequencing
2. Direct RNA sequencing
3. Identify splice variants and Isoforms
4. Fusion specific information
5. Sequence RNA viruses: SARS-CoV-2

## Nanopore Applications using MinION and GridION

**Genome size: 0.2 to 10 Mb**

**Smaller Genomes**
1. Bacteria
2. Bacteriophage
3. Virus

**Genome size: 50-500 Mb**

**Mid-sized Genomes:**
1. Fungus
2. Algae
3. Pomegranate
4. Rice
5. Teak
6. Drosophila
7. Fig Wasp

**Genome size: >500 Mb**

**Large Genomes:**
1. Shrimp
2. Mustard
3. Buffalo
4. Human
5. Bird
6. Elephant
7. Fish
8. Peacock


## Several Reports on Viral detection using Nanopore

SCIENTIFIC REPORTS

OPEN

Nanopore sequencing as a revolutionary diagnostic tool for porcine viral enteric disease complexes identifies porcine kobuvirus as an important enteric virus

PLOS ONE

RESEARCH ARTICLE

Early MinION™ nanopore single-molecule sequencing technology enables the characterization of hepatitis B virus genetic complexity in clinical samples

718 Biomed Environ Sci, 2017; 30(10): 718-726

Original Article

Rapid and Accurate Sequencing of Enterovirus Genomes Using MinION Nanopore Sequencer*

frontiers in Microbiology

TECHNOLOGY REPORT

MinION nanopore sequencing of an influenza genome

Genome Medicine

METHOD Open Access

Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis

Contents lists available at ScienceDirect

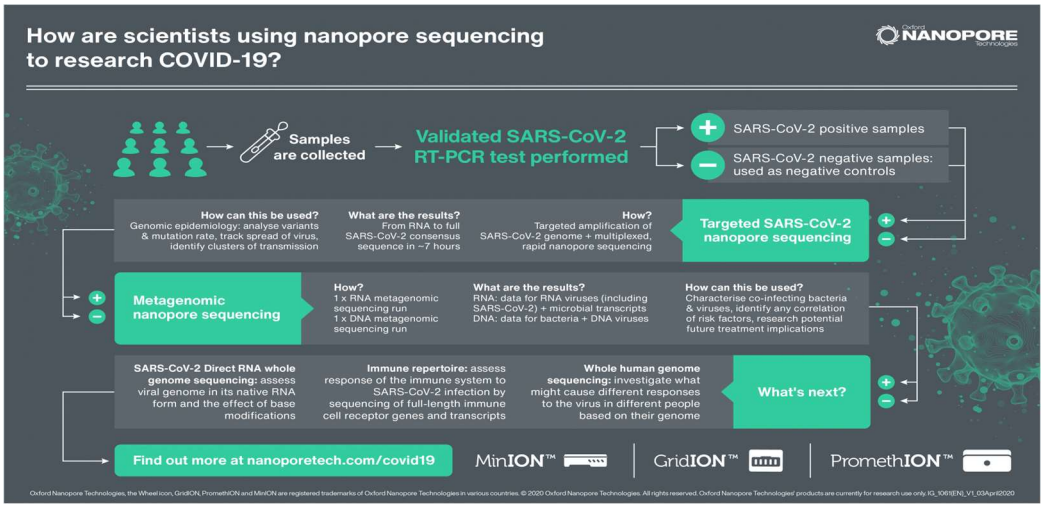
Journal of Virological Methods

journal homepage: www.elsevier.com/locate/jviromet

ELSEVIER

Metagenomic arbovirus detection using MinION nanopore sequencing

## How are scientists using Nanopore technology for SARS-CoV-2 sequencing?




## Potential Disadvantage and Limitations of Nanopore Sequencing

- Pore clogging
- Signal-to-noise ratio
- Error rate
- Read accuracy
- Cost of sequencing

Note: These potential disadvantages or limitations should be taken into account when considering Nanopore sequencing as a DNA sequencing technology. However, the benefits of long reads, real-time monitoring, and rapid sequencing offered by Nanopore sequencing still make it a promising technology for various applications.

Adapted from Branton, D., Deamer, D., Marziali, A. et al. The potential and challenges of nanopore sequencing. Nat Biotechnol 26, 1146–1153 (2008).

## List of Nanopore Sequencer

MinION

- Light Weight portable device for biological device in the palm of your hand.
- Pair with your laptop (MinION) or everything you need in one device (Mk1C)
- Generates up to 50 Gb in 72 hours

GridION

- Self-contained & easily deployable, benchtop device.
- Run upto 5 independently controllable MinION or Flongle Flow cells for multiple users and experiments.
- Enjoy the scalability & flexibility benefits of larger devices without specific infrastructure requirements.
- Generate upto 250 Gb of data basecalled in real time for immediate analysis.

PromethION

- Fully integrated devices, to high throughput solutions offering 24 or 48 flow cell positions.
- Easily deploy capacity for 2 promethION flow cells.
- Easily accommodate multiple devices in a single lab

Image Credit: Oxford Nanopore Technologies

## Flow cell formats

**MinION**

- The MinION and GridION Flow Cell contains up to 512 nanopore channels for sequencing DNA or RNA in real time.
- Available as R9.4.1 or our latest chemistry R10.4

**Flongle**

- Flongle is an adapter for MinION or GridION that enables direct, real-time DNA or RNA sequencing on smaller, single-use flow cells.
- Costing just $90 per flow cell, Flongle is the quickest, most accessible solution for smaller tests and experiments.

**PromethION**

- The PromethION Flow Cell contains up to 2675 nanopore channels for sequencing DNA or RNA in real time.
- Available as packs, contains 4 PromethION flow cells
- Available as R9.4.1 or our latest chemistry R10.4

Image Credit: Oxford Nanopore Technologies

## What's in your Box?


Image Credit: Oxford Nanopore Technologies

## The components of the MinION Mk1B


Image Credit: Oxford Nanopore Technologies

**LED lights**

There are four LED lights on each side of the device. The LEDs on the left side of the device display flow cell detected/undetected, Firmware loaded/not loaded, Communication signal with the MinKNOW software.

**Lid**

The lid can be flipped open to access the flow cell housing.

**Clip**

The clip holds the flow cell or Configuration Test Cell securely in place.

**Configuration Test Cell (CTC)**

The CTC is used during the hardware check to ensure that the communication between the device and the flow cell is working correctly.

**USB 3.0 port**

The USB 3.0 port is for connecting the MinION to the host computer for powering the device and data transfer.

## Sequencing Software: MinKNOW

- All Oxford Nanopore devices use MinKNOW™ as the primary software.
- The MinKNOW software carries out several core tasks: data acquisition, real-time analysis and feedback, basecalling, data streaming, controlling the device, and ensuring that the platform chemistry is performing correctly to run the samples.
- MinKNOW takes the raw data and converts it into reads by recognition of the distinctive change in current that occurs when a DNA strand enters and leaves the pore.
- MinKNOW then basecalls the reads and writes out the data into .fast5 or FASTQ files.


MinKNOW icon located on the desktop

## References

1. Oxford Nanopore Technologies. https://nanoporetech.com/
2. Branton, D., Deamer, D., Marziali, A. et al. The potential and challenges of nanopore sequencing. Nat Biotechnol 26, 1146–1153 (2008).

# 4. SARS-CoV-2 Sequencing Workflow

## Concepts Covered

**Why Genome Surveillance is Essential?**

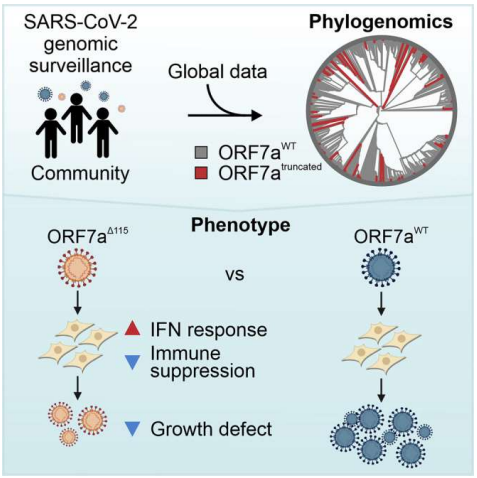
**About SARS-CoV-2 Virus & its Genome Structure**

**SARS-CoV-2 Library Preparation Methods & Approach**

**Nanopore Instrument Setup**

**Flowcell Preparation & Sequencing Run Setup**

## Why is Genome Surveillance Essential for Infectious Diseases




**Nemudryl et al 2021**

Slide reproduced from one of the Premas-FMR Workshop talk

- Vaccine Development
- Preparation of the Medical infrastructure
- Public Awareness
- Building up Resources
- Decision making

INSACOG
INDIAN SARS-COV-2 GENOMICS CONSORTIUM

COG-UK
COVID-19 GENOMICS UK CONSORTIUM

World Health Organization

CDC
Centers for Disease Control and Prevention

NATIONAL CENTRE FOR DISEASE CONTROL DELHI

## VOI & VOCs

**Variant of Interest**

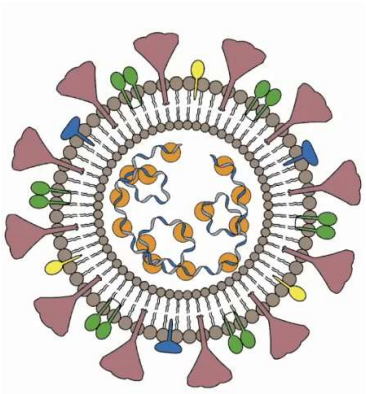
A variant with specific genetic markers that have been associated with changes to receptor binding, reduced neutralization by antibodies generated against previous infection or vaccination, reduced efficacy of treatments, potential diagnostic impact, or predicted increase in transmissibility or disease severity.

**Variant of Concern**

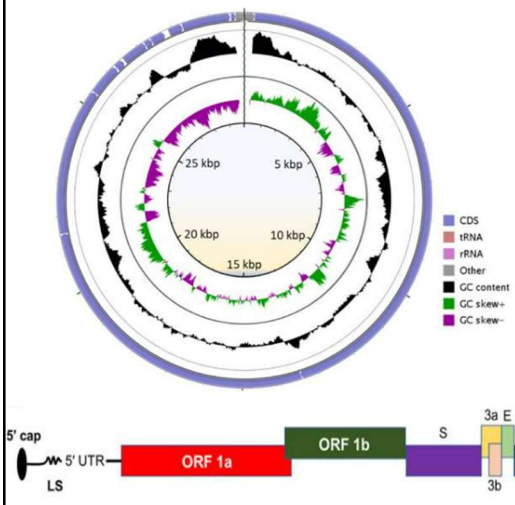
A variant for which there is evidence of an increase in transmissibility, more severe disease (for example, increased hospitalizations or deaths), significant reduction in neutralization by antibodies generated during previous infection or vaccination, reduced effectiveness of treatments or vaccines, or diagnostic detection failures.

## About SARS-CoV-2 Virus

- **S**evere **A**cute **R**espiratory **S**yndrome **C**oronavirus 2
- Causative agent of **COVID-19**
- Single stranded positive sense RNA virus of the beta coronavirus genus
- Genome length is ~30 kb.



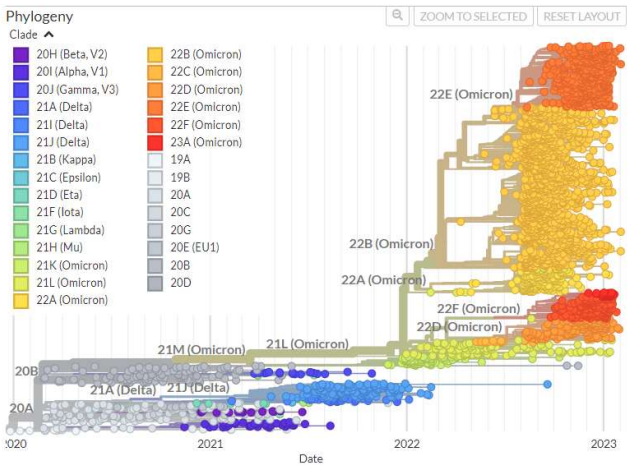
## hCoV-19 Genome Structure




| Protein | Description |
|---|---|
| Nsp1 | Nsp1 is the N-terminal product of the viral replicase |
| Nsp2 | Nsp2 is a replicase product essential for proofreading viral replication |
| Nsp3 | Nsp3 is a papain-like proteinase contains several domains. |
| Nsp4 | A membrane-spanning protein contains transmembrane domain 2 (TM2) |
| Nsp5 | 3C-like proteinase and main proteinase |
| Nsp6 | Putative transmembrane domain |
| Nsp7 | Nsp7 is an RNA-dependent RNA polymerase |
| Nsp8 | Multimeric RNA polymerase; replicase |
| Nsp9 | A single-stranded RNA-binding viral protein |
| Nsp10 | Growth-factor-like protein contains two zinc-binding motifs |
| Nsp12 | RNA-dependent RNA polymerase |
| Nsp13 | Zinc-binding domain |
| Nsp14 | Proofreading Exoribonuclease domain |
| Nsp15 | EndoRNAse |
| Nsp16 | 2'-O-ribose methyltransferase |
| Nsp11 | Made of 13 amino acids |

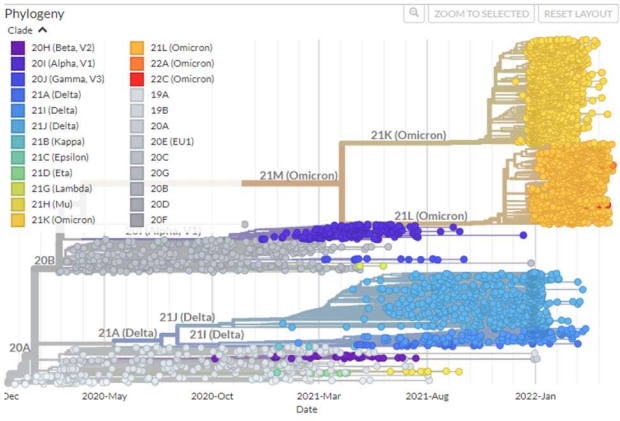
## Why to sequence SARS-CoV-2 Genome?

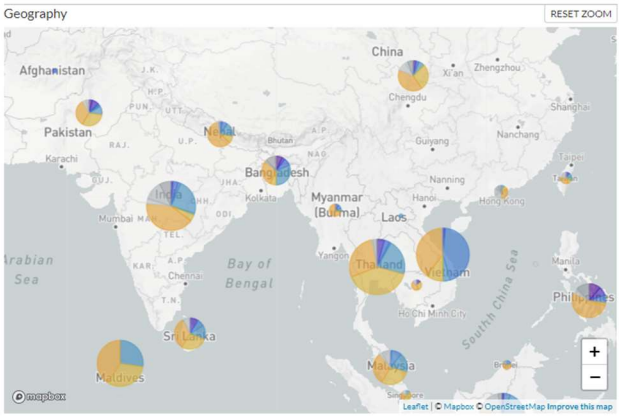
- **Viral epidemiology during an out break**
  - Monitoring spread
  - Help inform intervention methods
  - Monitoring the impact of intervention methods
- **Vaccine efficacy monitoring**
  - Tracking prevalence and distribution of key genes involved in vaccine development and deployment (i e spike protein)
  - Determine mutation rates to inform likelihood of vaccine longevity




## Genomic epidemiology of SARS-CoV-2 with Asia-focused subsampling

Several variants of SARS-CoV-2:
Alpha (B.1.1.7), Beta (B.1.351), Gamma (P1, Brazil), Delta (B.1.617.2) and Omicron (B.1.1.529).



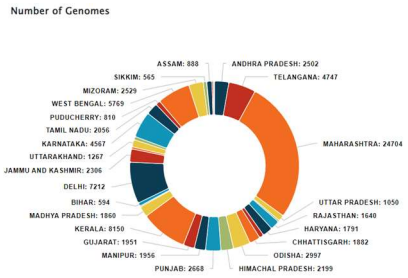
## A brief overview of online resources

**GISAID**

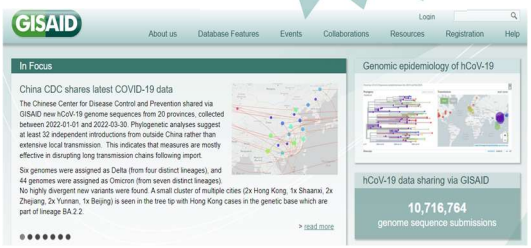
- The GISAID Initiative encourages the quick exchange of data from all influenza viruses as well as the COVID-19 coronavirus. This includes genetic sequence and related clinical and epidemiological data associated with human viruses, and geographical as well as species-specific data associated with avian and other animal viruses.

**INSACOG**

- The Indian SARS-CoV-2 Genomics Consortium (INSACOG), jointly initiated by the Union Health Ministry of Health, and Department of Biotechnology (DBT) with Council for Scientific & Industrial Research (CSIR) and Indian Council of Medical Research (ICMR), is a consortium of 58 laboratories to monitor the genomic variations in the SARS-CoV-2.




Webpage Information





## Methodology

## SARS-CoV-2 Sequencing – Available Methods

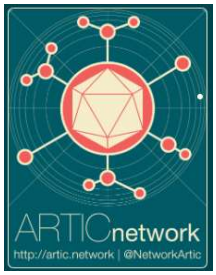
There are two methods available for whole-genome nanopore sequencing of SARS-CoV-2: **Midnight** and **ARTIC Classic**.

- **Midnight** is a simple, rapid method of sequencing SARS-CoV-2 genomes at a low cost per sample. The SARS-CoV-2 genome is amplified in ~1,200 bp overlapping segments, making it more resilient to drop-out caused by mutations in the viral genome.
- **ARTIC Classic** was the first SARS-CoV-2 nanopore sequencing protocol to be utilized and has been used by scientists around the world. In this method, the SARS-CoV-2 genome is amplified in ~400 bp fragments. This shorter length may help improve coverage for RNA samples that are likely to be degraded - for example, due to freeze-thaw cycles or storage at temperatures above -80°C.

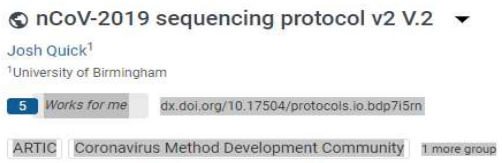
Both methods employ a PCR tiling approach in which the viral genome is amplified in overlapping sections, maximizing coverage across the full genome.

## What is an ARTIC Network?

**An independent network of researchers from:**




- University of Edinburgh
- University of Birmingham
- University of Cambridge
- KU Leuven
- University of Oxford
- Fred Hutchinson Cancer Research Center
- University of California Los Angeles




- Pioneered " an end-to-end system for processing samples from viral outbreaks to generate real-time epidemiological information that is interpretable and actionable by public health bodies.
- Responsible for protocol and bioinformatics pipelines and recommended by Oxford Nanopore
- Artic network is platform-independent and will use any technology available.

## Which approach to use?

1. Classic PCR tiling of SARS-CoV-2 virus
2. PCR tiling of SARS-CoV-2 virus with Native Barcoding Expansion 96 (EXP-NBD196)
3. Eco PCR tiling of SARS-CoV-2 virus with native barcoding (EXP-NBD104, EXP-NBD114, and EXP-NBD196)
4. PCR tiling of SARS-CoV-2 virus with rapid barcoding (SQK-RBK110.96)

| | ARTIC Classic | Midnight |
|---|---|---|
| Experience level required | ●●●○ | ●●○○ |
| Third-party reagent usage | ●●●○ | ●○○○ |
| Amplicon length generated | 400 bp | 1200 bp |
| Normalisation step included | Yes | No |
| Batching by Ct value recommended | No | Yes |
| Library prep method | Ligation | Rapid |
| Turnaround time of workflow | ●●●● | ●●○○ |

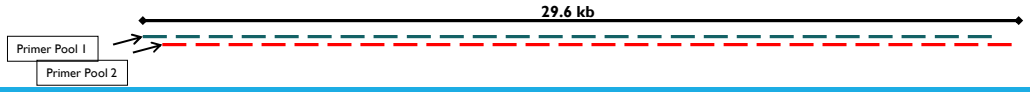
## Which approach to use?

**What does the protocol do??**

- Allows to generate a consensus sequence of the SARS-CoV-2 genome from RNA obtained from clinical samples
- Identifies sequence variants
- Allows decentralized data sharing for global sequence analysis of the SARS-CoV-2 genome

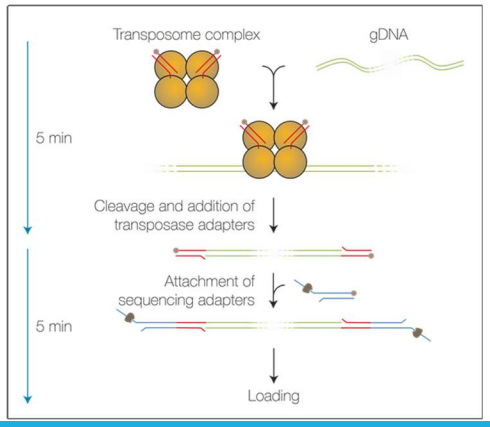
**How does it do that?**

- Generates 400 bp tiled amplicons from reverse transcribed viral RNA
- Bioinformatics pipelines from the ARTIC network use an alignment based approach to generate SNP and indel calls and subsequently a consensus sequence

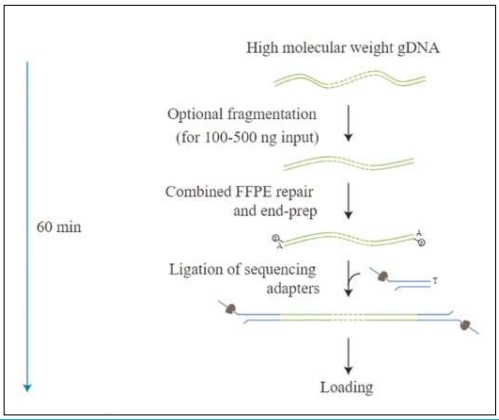


## ONT Library preparation method and multiplexing

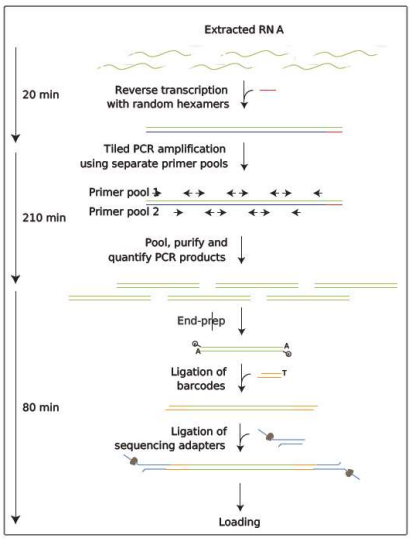
For minimal preparation time:
Rapid kit with Transposase



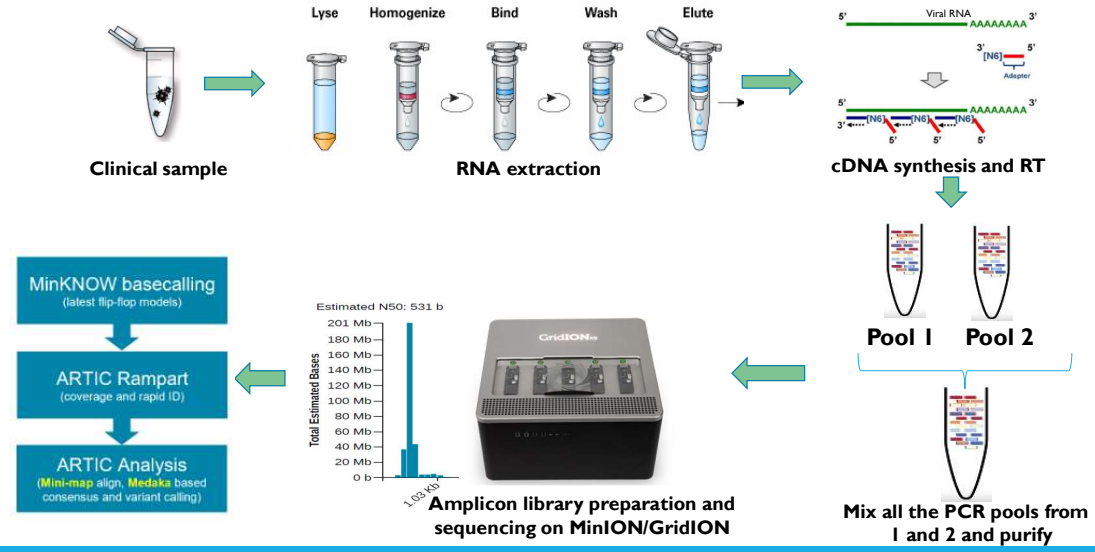

For maximum throughput:
Native Ligation kit




## Steps in the ARTIC protocol



## ARTIC Protocol - Workflow overview

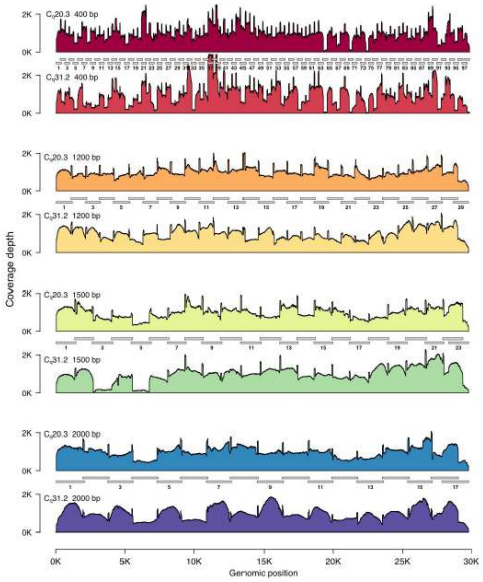



## Why not longer amplicons….? 1200, 1500 and 2000bp

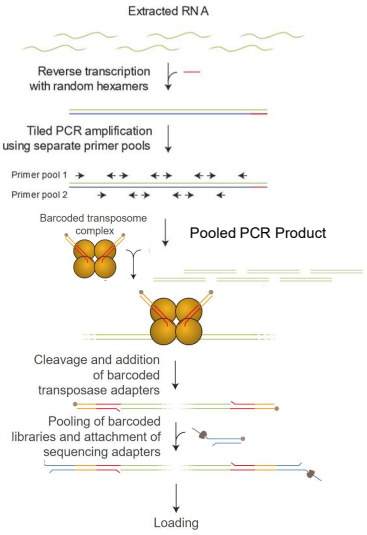
Rapid and Inexpensive Whole-Genome Sequencing of SARS-CoV2 using 1200 bp Tiled Amplicons and Oxford Nanopore Rapid Barcoding

Authors
Nikki E. Freed*, Markéta Vlková, Muhammad B. Faisal, Olin K. Silander*
Address for all authors:
School of Natural and Computational Sciences, Massey University, Auckland, New Zealand

- Longer amplicons with Rapid Barcode library kit allows for faster, simpler, and less expensive SARS-CoV-2 genome sequencing.
- Amplicon size of 1200 bp exhibit lower levels of variation in coverage compared to other commonly used primer sets.
- This method reduces the time by more than half compared to the more standard ligation-based library preparation method at considerably lower costs.



## Steps in the Midnight Artic protocol




## Overview of Rapid barcoding using midnight primers

Extracted RNA

cDNA synthesis

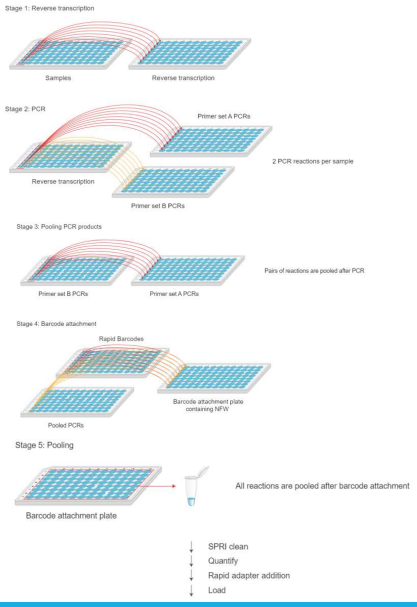
Multiplex PCR

Rapid barcoding

Sequencing

Analysis

7 hours



# Nanopore Instrument Setup

## Setting up the MinION Mk1B

**Insert the Configuration Test Cell (CTC) into the MinION Mk1B instrument.**

- ✓ Clip the CTC into place in the MinION Mk1B and connect to the host computer.
- ✓ Gently press down on the CTC; there will be a slight click as the CTC clips into place.
- ✓ Close the MinION Mk1B lid.

The Flongle CTC tests the MinION Mk1B and Flongle adapter are working correctly.

## Setting up the MinION Mk1B for Flongle

**Place the Flongle adapter into the MinION Mk1B**

- ✓ The adapter should sit evenly and flat on the MinION Mk1B. This ensures the flow cell assembly is flat during the next stage.

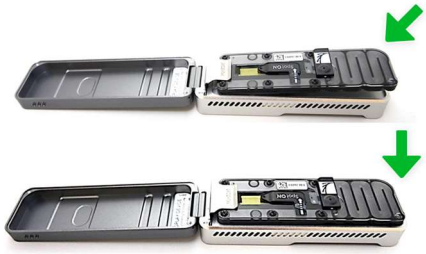
**Place the CTC into the Flongle adapter and press the flow cell down until you hear a click.**

- ✓ The CTC should sit evenly and flat inside the adapter
- ✓ Close the MinION Mk1B lid.

## Setting up the MinION Mk1B for Flow cell

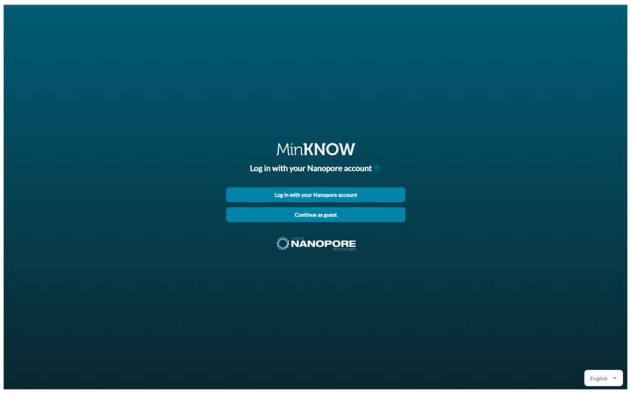
**Open the MinION Mk1B lid and insert the MinION Flow Cell**

- ✓ Insert the flow cell in the MinION Mk1B/GridION by sliding the flow cell under the clip. Firmly press down on the flow cell to ensure good thermal and electrical contact.

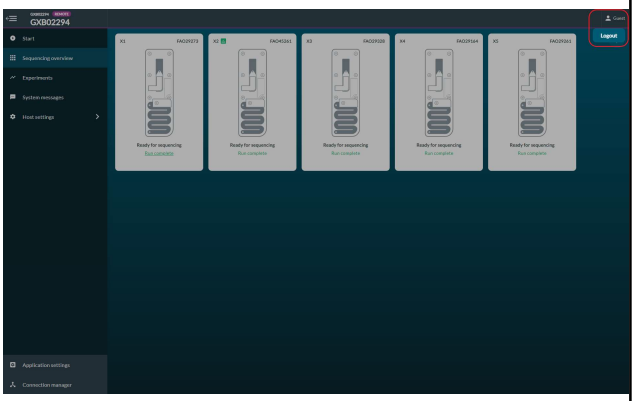


## Connecting to Device

To log in, you must be connected to the internet.




To logout of the MinKNOW software




Select the button in the top right corner, labelled with your initials or 'Guest' and click 'Logout'.

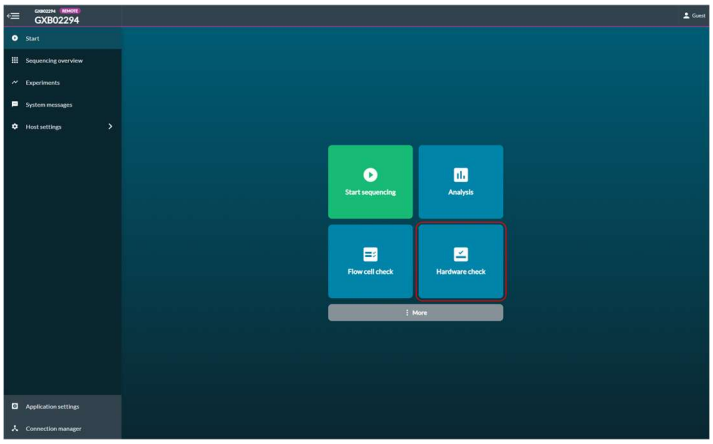
## Hardware Check

A hardware check must be performed on all new devices or when software has been upgraded. This uses the CTC, which comes pre-inserted into your device in place of flow cells.

Note: If using a Flongle flow cell, ONT recommends regularly checking the Flongle adapter by inserting an **empty** adapter and completing a hardware check. For checking the device or flow cell position, please use a MinION CTC, even if a Flongle flow cell will be used for sequencing.

## Hardware Check – Step 1 & 2

**Step 1. Insert a Configuration Test Cell (CTC) into the device.**



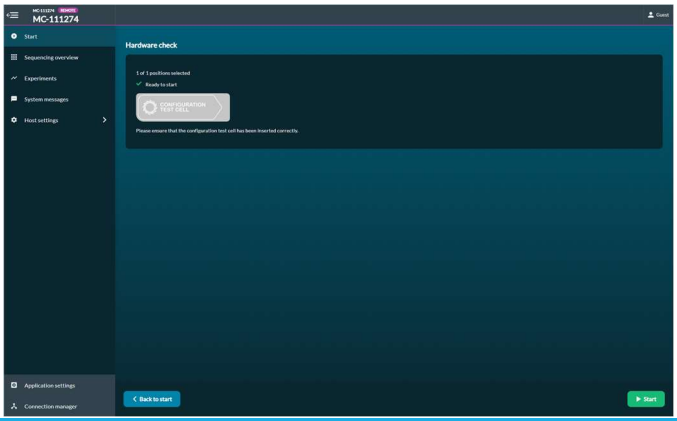

**Step 2: Navigate to the start homepage and select 'Hardware Check' option**

## Hardware Check – Step 3

**Step 3: Select 'Start' for the check to begin.**

Ensure the position selected is ready to start, as indicated on the GUI.

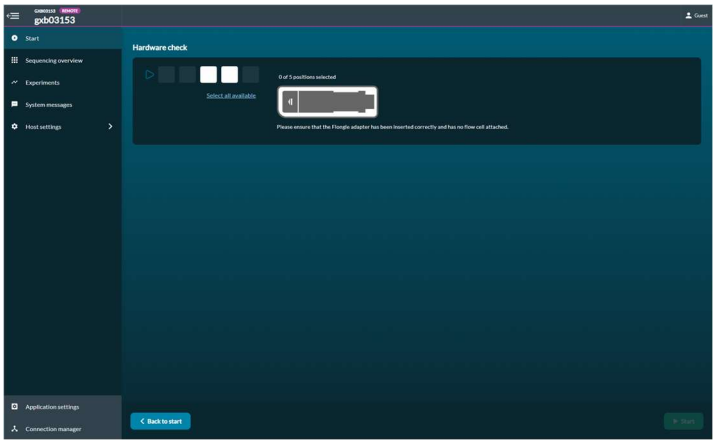
Using MinION flow cells:



## Hardware Check – Step 4

Using Flongle flow cells:

Note: To check the Flongle adapter, insert the EMPTY adapter. To check device or position, insert only a MinION CTC.



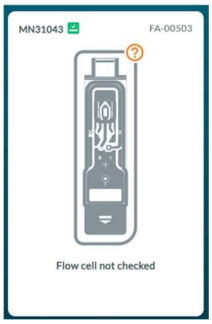

## Hardware Check Status

- You will be automatically navigated to the Sequencing Overview page.
- A loading bar will be displayed under the flow cell during the checks.

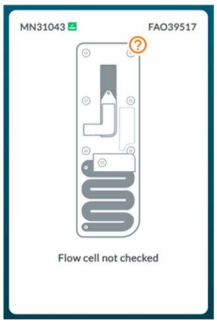
The hardware check will complete after approximately one minute.

✓ Hardware check pass is indicated by a green check icon.

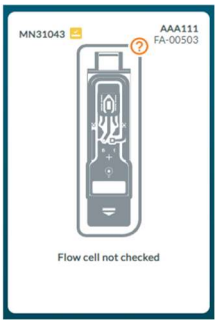
✓ An orange check icon is a fail.




Flongle hardware check pass




MinION hardware check pass




Flongle hardware check fail

## Flow Cell Check

A flow cell check must be carried out before loading a DNA or RNA library to assess the number of pores available.

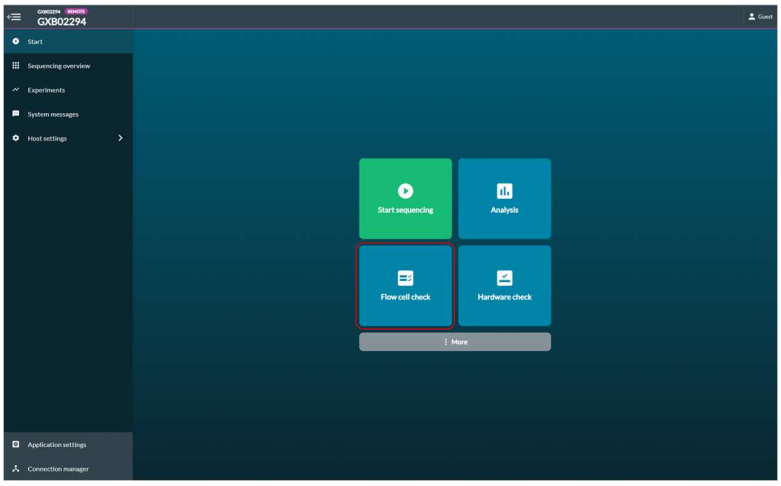
**Purpose of the flow cell check**

A Flow Cell Check is used to assess the number of nanopores that are available in the flow cell. It should be run when you are ready to use the flow cell, and within 3 months of purchase (except for Flongle flow cells, which currently have a shelf life of four weeks).

| Flow cell | Minimum number of active pores covered by warranty |
|---|---|
| Flongle Flow Cell (FLO-FLG001) | 50 |
| Flongle Flow Cell (FLO-FLGOP1) | 30 |
| MinION/GridION Flow Cell | 800 |
| PromethION Flow Cell | 5000 |

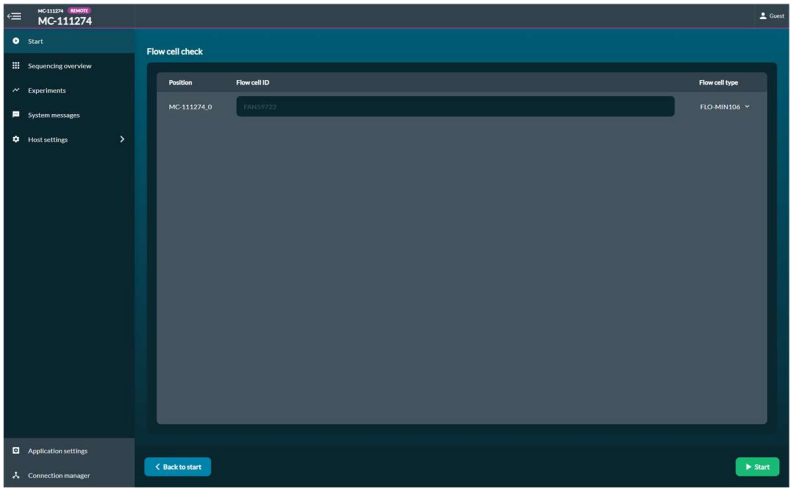
## Flow Cell Check – Step 1

**Navigate to the Start page and select 'Flow Cell Check' to open the flow cell check page.**



## Flow Cell Check – Step 2

**The flow cell type will be recognized by the device. Click 'Start' to begin.**



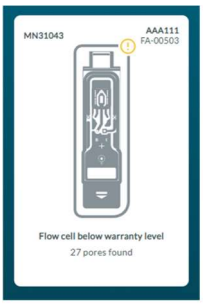

## Flow Cell Check Status

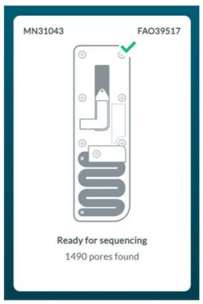
You will be automatically navigated to the Sequencing Overview page. A loading bar will be displayed under the flow cell during the check.

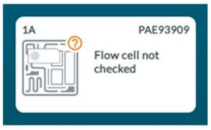
**Flow cell health indicators**

The quality of the flow cell will be shown as one of the three outcomes:

- ✓ Yellow exclamation mark (Flongle flow cell): The number of sequencing pores is below warranty.
- ✓ Green tick (MinION flow cell): The number of sequencing pores is above warranty.
- ✓ Question mark (PromethION flow cell): A flow cell check has not been run on the flow cell during this MinKNOW session.




Note: The indicator of quality will only remain visible during a MinKNOW session. Once the MinKNOW session has ended, the status of the flow cell will be erased.

## Flow Cell Replacement & Storage Recommendation

Oxford Nanopore Technologies will replace any flow cell that falls below the warranty number of active pores within three months of purchase, provided the result is reported within two days of performing the flow cell check and the storage recommendations have been followed.

| Flow cell | Storage instructions |
|---|---|
| Flongle Flow Cells | 2–8°C for four weeks |
| MinION/GridION Flow Cells | room temperature for one month, or at 2–8°C for 12 weeks |
| PromethION Flow Cells | 2–8°C for 12 weeks |

# Setting Up the Experiment

## Experiment Page: Introduction

- The experiments page displays summary information for all sequencing flow cells and device checks carried out on the device.
- Previous runs can be viewed here until MinKNOW service is restarted (e.g. after a device reboot).
- From this page, the user is able to control specific runs and identify real-time information, including flow cell health and reads, giving users real-time feedback for sequencing flow cells.

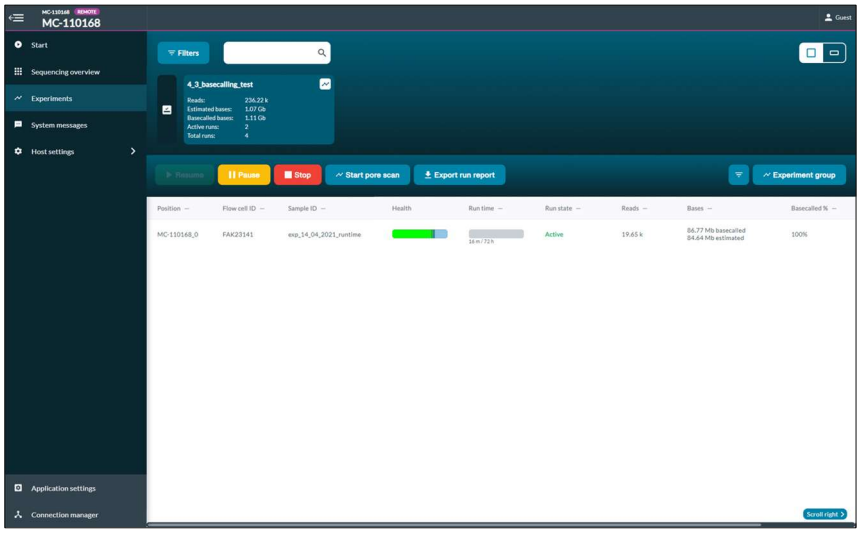
**Run statistics:** The total number of reads, estimated and basecalled bases across an experiment, and number or active and total runs.

**Run time:** The duration of the experiment.

**Run state:** The current state of the sequencing run; 'Active', 'Basecalling', 'Complete', 'Stopped with error'.

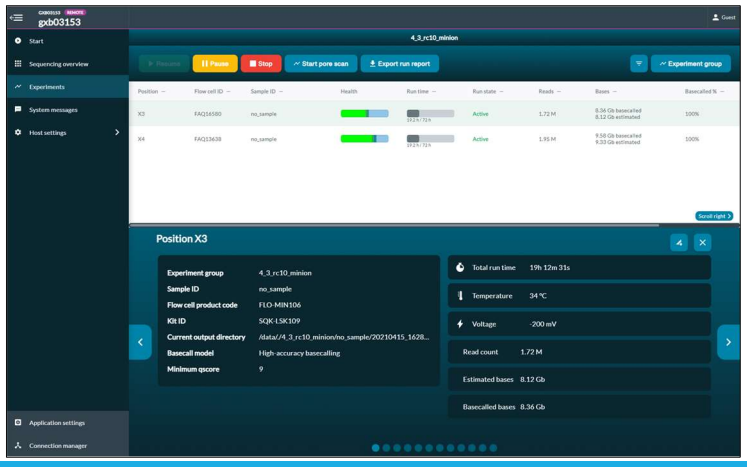
**Health:** The current flow cell health.

## Experiment Page of MinION/Flongle Flow Cell



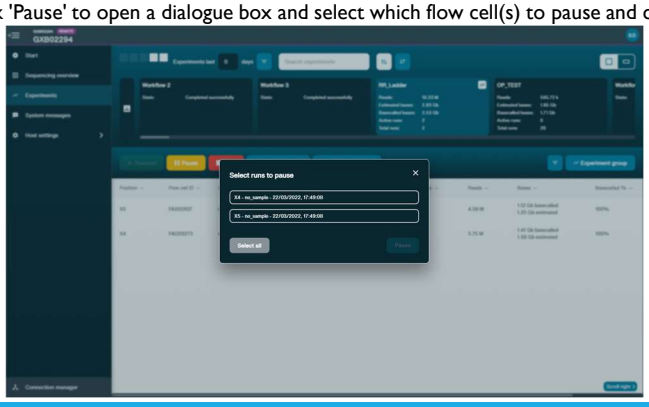
## Experiment Page : MinION/Flongle Runs

For more status information of a specific run, select the run to open the quick view, including current temperature and voltage.




## Experiment Page : Run Pausing

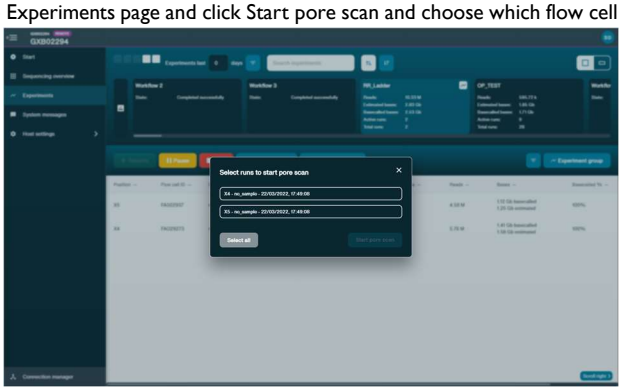
- Pausing works by dropping the voltage potential over the membrane to 0 mV to maintain a safe environment to add components, such as fuel, more DNA/RNA library or nucleases for a flow cell wash.
- Data acquisition will continue during this period, as this prevents sequencing data from being lost.
- To pause, click 'Pause' to open a dialogue box and select which flow cell(s) to pause and click 'Pause'.



## Experiment Page: Triggering a pore scan

The pore scan is used to assess the quality of the four wells in each channel to select the best performing pores. A new pore scan can be triggered every time a sequencing experiment is resumed after a pause (e.g. for a flow cell wash), or if the number of sequencing pores has significantly dropped during an experiment.

Navigate to the Experiments page and click Start pore scan and choose which flow cell to scan.




## Flow cell Preparation

## Flow cell Priming & Loading

**Mix and Vortex(Priming Mix)**

| Reagent | Volume |
|---|---|
| Flush Buffer(FB) | 1.17mL |
| Flush Tether (FLT) | 30mL |

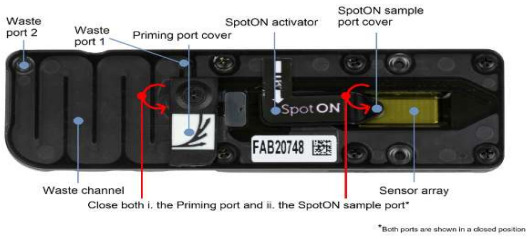
**Loading onto Flow Cell**


- Open port to 90° angle.
- After opening the priming port, check for a small air bubble under the cover.
- Draw back 20 – 30 uL to remove any bubbles.

## Flow cell Priming & Loading

Priming with Flush Buffer (F.B):

- Set pipette at 900uL
- Load 800 µl of the priming mix into the flow cell via the priming port, avoiding the introduction of air bubbles. Incubate for 5 minutes.
- Load 200 µl of the priming mix into the flow cell via the priming port (not on the SpotON sample port), avoiding the introduction of air bubbles.
- Mix the prepared library gently by pipetting up and down just prior to loading.
- Add 75 µl of sample to the flow cell via the SpotON sample port in a dropwise fashion.
- Gently replace the SpotON sample port cover, making sure the bung enters the SpotON port, close the priming port and replace the MinION Mk1B lid.



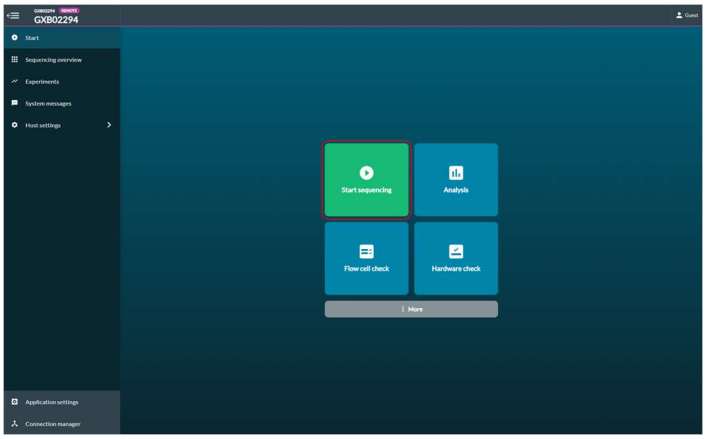
# Sequencing Run Setup

## Introduction to basecalling in MinKNOW

- For MinION Mk1B and Flongle (on MinION Mk1B), the MinKNOW software presents an option to basecall reads on the local computer. The basecalling is carried out live (or later), as the read files are generated during a sequencing experiment.
- Basecalling results are displayed in real-time in the MinKNOW GUI, and data is written out in the FASTQ or .fast5 file format.

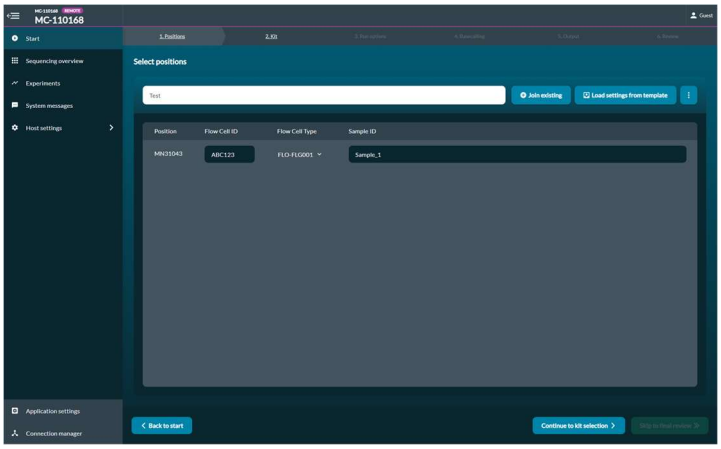
## Starting a sequencing run

Select the 'Start Sequencing' option on the Start homepage to choose the running parameters for your experiment




## Step 1: Starting a sequencing run

Type in the experiment name, sample ID and choose flow cell type from the dropdown menu. Fill in flow cell ID in the correct format: ABC123

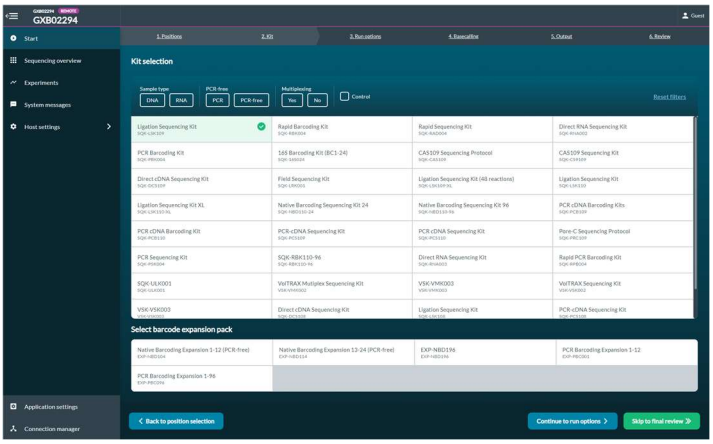




For Flongle, flow cell ID is not automatically assigned in MinKNOW, can be found at the back of the flongle.

## Step 2: Starting a sequencing run

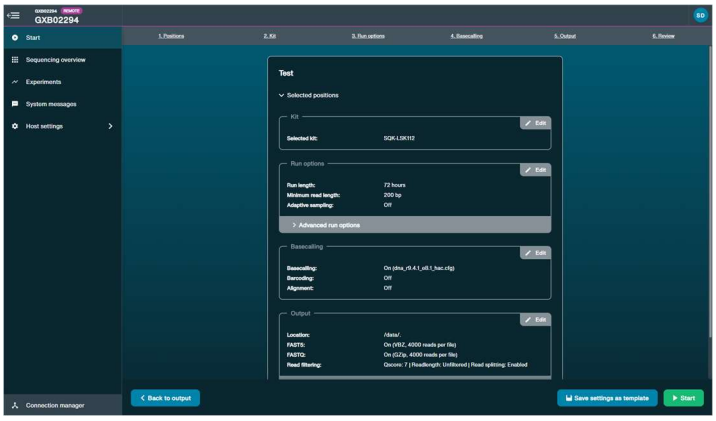
- ✓ Select Continue to Kit Selection to move to the next page. Select Sample Type: DNA & PCR Free: PCR & Multiplexing : Yes
- ✓ Choose the right kit used for the library preparation & Click Skip to Run review




## Step 3: Run Selection Review & Initiation

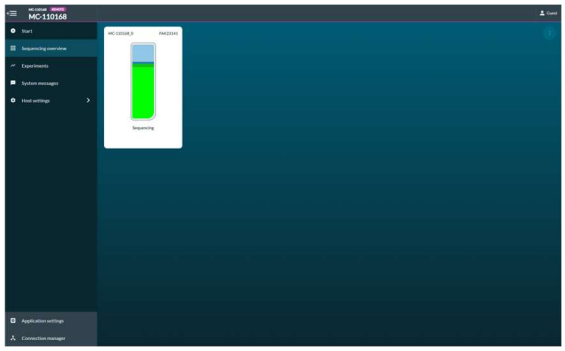
Click 'Start' to run the experiment.

The Review page is an overview of all run options selected.



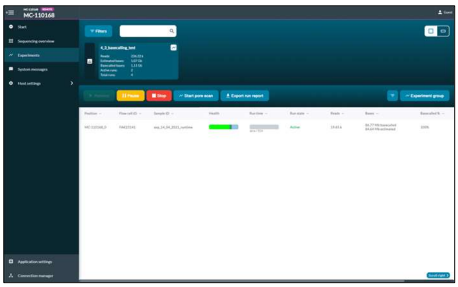
## Step 4: Sequencing Run Progress

- Users will be automatically navigated to the Sequencing Overview when sequencing starts.
- From here, users can see a progression bar below the flow cell to show the progression of the sequencing script.
- Flow cell health will be displayed after the first pore scan.

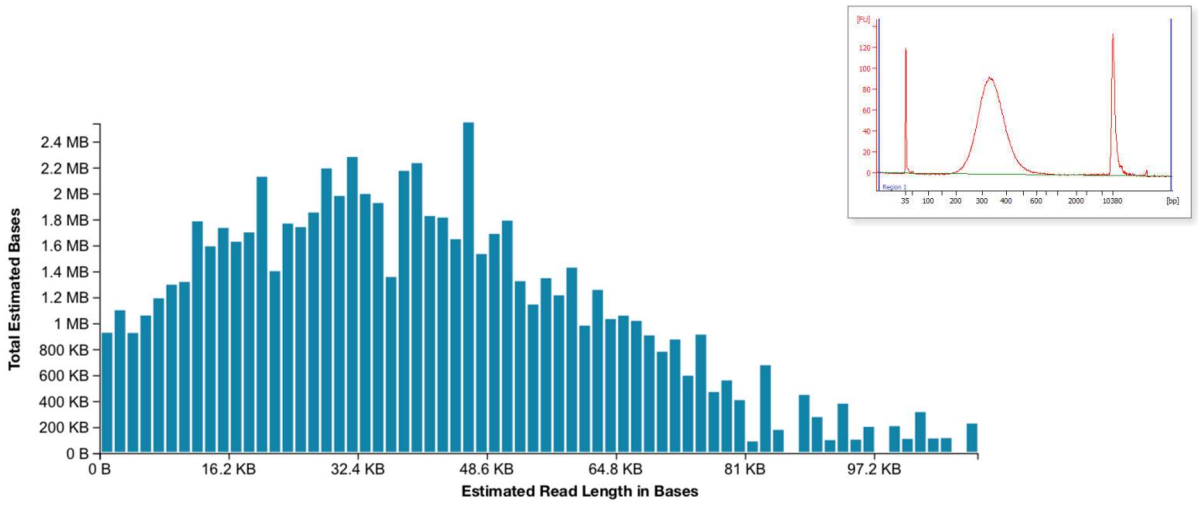



## Step 5: Flow cell active pores review

- Select the flow cell to open the quick view and check the number of active pores. The first pore scan should report a similar number of active pores (within 10-15%) to that reported in the flow cell check.
- If there is a significant reduction in active pores in the first pore scan, restart MinKNOW.
- If the numbers are still significantly different, close down the host computer and reboot.
- When the numbers are similar to those reported at the end of the flow cell check, restart the experiment. There is no need to load any additional library after the restart.



## A screenshot from Nanopore sequencing run




## Step 6: Sequencing run completion report

- Once sequencing is complete, generate a run report.
- A run report containing information about the sequencing run and performance graphs can be generated by clicking Export run report.
- For MinION Mk1B running on Windows or Mac, the run reports need to be manually generated. Click Export run report and select which experiment to export to html.
- A duty time CSV file is also generated for every run.
- The report and CSV files are saved to the same folder as the .fast5 and .fastq read files e.g. :\data\experiment\sample_ID\ for MinION Mk1B.

## Step 6: Flow Cell Refueling (optional)

Refuelling is the replenishment of motor fuel in the sequencing experiment through the addition of Flush Buffer (FB) from the Flow Cell Priming Kit (EXP-FLP002).

The translocation speed graph in MinKNOW can indicate whether it is necessary to top up fuel.

Volumes and processes of refuelling are specific to the flow cell type.

**When to refuel**

If the DNA translocation speed drops below 300 bases per second, you may start to see a reduction in quality of data reflected in the Qscore.


## Step 7: Flow cell washing

- Remove 1000 µl of buffer from 3rd port keeping the priming port and spot on port closed.
- De-gas the priming port.
- Flush the priming port with 400 µl mix. Incubate for 1 hour.

| Reagent | Volume |
|---|---|
| Wash Diluent | 398µl |
| Wash Mix | 2µl |

## Step 8: Flow cell Storage

- Thaw one tube of Storage Buffer (S) at room temperature
- Mix contents thoroughly by pipetting and spin down briefly.
- De-gas the priming port.
- Slowly add 500 µl of Storage Buffer (S) through the priming port of the flow cell
- Close the priming port.
- Using a P1000, remove all fluid from the waste channel through Waste port 1. As both the priming port and SpotON sample port are closed, no fluid should leave the sensor array area.
- Store the flow cell at 2 – 8°C until the next use (within the expiry of the flow cell).

Note: It is vital that the flow cell priming port and SpotON sample port are closed to prevent air from being drawn across the sensor array area, which would lead to a significant loss of sequencing channels.

## What's Next: What do I do with the Data

AAAGTTGGTATTAACGGATTCGGTCGCATTGGCCGTCTGGTGACCCGTGCTGCTTTCTTGACCAAGA
TGGAGATCGTGGCCATCAATGACCCATTCATTGACCTTGATTACATGGTTTACATGTTCCAGTACGA
CACCCATGGAAAGTACAAGGGTGAGGTTAAGGCAGAAGGCGGCAAACTGGTCATTGATGGTCATGCA
ACAGTCTATAGCGAGAGGGACCCAGCCAACATTAAGTGGGGTGATGCAGGTGCTACTTATGTTGTGG
CTACTGGTGTCTTCACTACTATTGAGAAGGCTTCTGCTCACATTAAGGGTGGTGCAAAGAGAGTCAT
CTCTGCCCCAAGTGCAGATGCCCCCATGTTTGTCATGGGTGTCAACCATGAGAAATATGACAACTCT
ACA
TTG
CTC
AAGGCTGTAGGCAAAGTAATTCCTGAGCTCAATGGCAAGCTTACTGGTATGGCCTTCCGTGTCCCCA
CCAATGTCTCTGTTGTGGATCTGACAGTCCGTCTTGAGAAACCTGCCAAGTATGATGAGATCAAGAA
CGTCAAGGCTGCAGCTGATGGGCCCATGAAAGGAATTCTGGGATACACGGAGCACCAGGTTGTGTCC
GACTTCAATGGGGATTGCCGTTCATCCATCTTTGACGCTGGTGCTGGTATTGCTCTCAACGATCACT
TCAAGCTGGTCACATGGTATGACAATGAGTTCGGTTACAGCAACCGTGTATGTGACCTGATGGCACA
GGCCTCCAAGGAGTAGATGTGACCCCTTTGCTGTTTCTTTTTTTTGATACGCGACCATTCTCCCATC
TTGAATGTTTGCACCACGTGCCTGGAAGGAAATTACATGCTTAAATTGAAGACCAATATTATTTTTA
ACTCTGTTCTGTTTCGTGTGTGAGGTTAAAAATAAATGTTGACTTCAAAGGCTTTTCTGTCTGTTAA

**How to analyze the sequence data….?????**

## Bioinformatics for Rescue

## References


1. PCR tiling of SARS-CoV-2 virus - rapid barcoding (SQK-RBK110.96). https://community.nanoporetech.com/docs/prepare/library_prep_protocols/pcr-tiling-of-sars-cov-2-virus-with-rapid-barcoding-sqk-rbk110/v/pctr_9125_v110_revh_24mar2021, requires ONT community access.
2. PCR tiling of SARS-CoV-2 virus - classic protocol (SQK-LSK109 with EXP-NBD104, EXP-NBD114 or EXP-NBD196. https://community.nanoporetech.com/docs/prepare/library_prep_protocols/classic-pcr-tiling-SARS-CoV-2/v/ptc_9096_v109_revx_06feb2020, requires ONT community access).
3. Nikki Freed, Olin Silander 2021. SARS-CoV2 genome sequencing protocol (1200bp amplicon "midnight" primer set, using Nanopore Rapid kit). protocols.io https://dx.doi.org/10.17504/protocols.io.bwyppfvnVersion created by Nikki Freed
4. SARS-CoV-2 ARTIC Network. https://artic.network/ncov-2019
5. MinKNOW Software, Oxford Nanopore Technologies. https://nanoporetech.com/


# 5. SARS-CoV-2 Bioinformatics analysis

## Concepts Covered

**Steps in bioinformatics analysis**

**Output Data Formats**

**Adapter Removal**

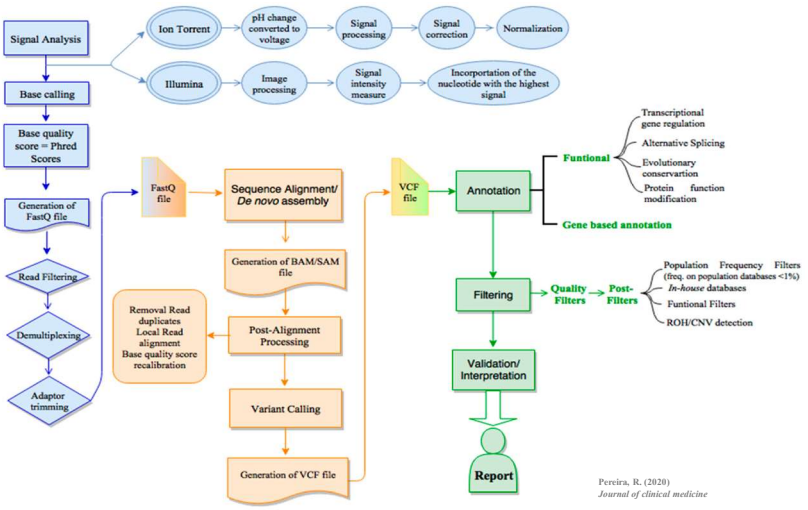
**Steps in Genome Assembly**

**Variant Discovery**

**InterARTIC Workflow**

**Lineage Identification & Visualization**

## Bioinfo Analysis – Multi Step Process



## Output Data Formats

### Fast5 Reads

| | | | |
|---|---|---|---|
| GXB01242_20180928_FAJ03745_GA20000_... | 9/28/2018 4:03 PM | FAST5 File | 39 KB |
| GXB01242_20180928_FAJ03745_GA20000_... | 9/28/2018 4:01 PM | FAST5 File | 35 KB |
| GXB01242_20180928_FAJ03745_GA20000_... | 9/28/2018 4:01 PM | FAST5 File | 44 KB |
| GXB01242_20180928_FAJ03745_GA20000_... | 9/28/2018 4:03 PM | FAST5 File | 36 KB |
| GXB01242_20180928_FAJ03745_GA20000_... | 11/15/2018 5:01 PM | FAST5 File | 48 KB |
| GXB01242_20180928_FAJ03745_GA20000_... | 11/15/2018 5:01 PM | FAST5 File | 58 KB |
| GXB01242_20180928_FAJ03745_GA20000_... | 9/28/2018 4:00 PM | FAST5 File | 41 KB |
| GXB01242_20180928_FAJ03745_GA20000_... | 9/28/2018 4:01 PM | FAST5 File | 58 KB |
| GXB01242_20180928_FAJ03745_GA20000_... | 11/15/2018 5:01 PM | FAST5 File | 35 KB |
| GXB01242_20180928_FAJ03745_GA20000_... | 11/15/2018 5:01 PM | FAST5 File | 32 KB |
| GXB01242_20180928_FAJ03745_GA20000_... | 9/28/2018 4:01 PM | FAST5 File | 31 KB |
| GXB01242_20180928_FAJ03745_GA20000_... | 11/15/2018 5:01 PM | FAST5 File | 33 KB |

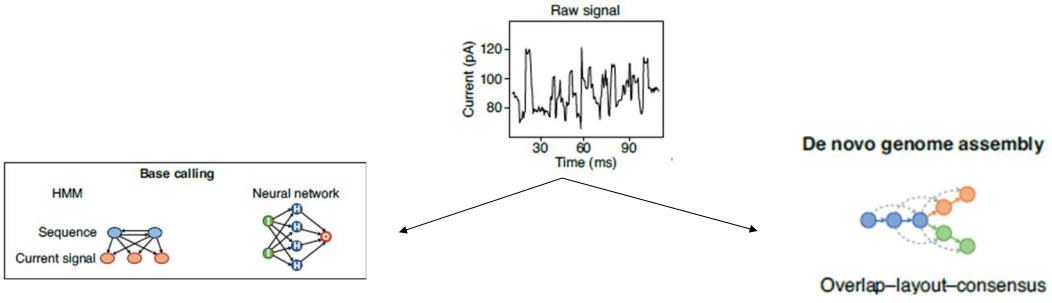
### Fasta Reads

```
>ATCAAACATTGCTTCGTTCAGTTACGTATTGCTAAGGTTAAAGTTCATTCCCACGGTAACACCAGCACCT
ATCAGCACCAACAGAAGGATATTGACTTGCCTGTCGCCCTATCTTCGGCAATATCAGCACCAACAGAAGAT
TCAGCACCAACAGAAAGATATTGACTTGCCTGTATCGCTCCTTCGGCAATATCAGTACCCAACAGAAGATA
ATATCGCTCAACAGAAGGATATTGTCTTGCCGTCAACTCTATCTTCGGCAATATCGCGCCAACAGAAAGAT
GGCAATATCAGCACCAACAGAGGATATTGTCTTGCCTGATAAACTCTATCTTCAGCAATATTAGCACCAAC
CTCTATCTTCAGCAATATCAGCACCAACAGAAAGATATTGACTCGCCTGTCAACTCTATCTTCGGCAAATA
AATAAGCTCTTCAGCAATATCAGCACCAACGAAGATGTTGTCTTGCCTGTCACTCTATCTTTCAGCATTAT
TGCCTGCCACTCTATCTTCAGCAATATCAGCACCAACAGAAGATATTTGTCTTGCCTGCCGCTCTATCTTC
GAACAACCTTCTGACCGATACTGATATTAACGAAAATAAACGACTGAAAATGAACTCTTTCTGTTAATGCC
AACCAATATCACTCTTGTTATTATGATATGCTGAAAATGAAGTGACGGCATCTTTCTGTTGATGCTATTTT
```

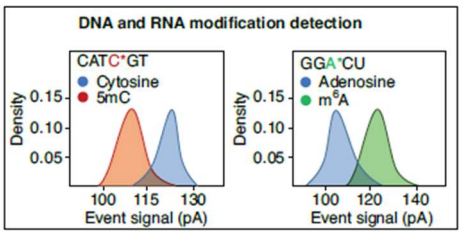
### Fastq Reads

```
barcode=barcode07
TTGTGTACTTCGTTCAGTTACGTATTGCTAAGGTTAAAAGGATTCATTCACGGCCAACACCAGCACCTGGCAAAGACAATATCTTTCTGTTGGTACTGATAATGCTGAGAGAATAAAGCGACGGGCAAACCAATATCTTTCTGTTGGTGC
TGATATTGCTGAAGATAGAGCGACAGGCAAGACAATATCTTTCTGTTGGTGCTGATATTGCTGAAGATGAGGCGACAGGCGAACAACCATCTTTTCATGTAGGTGCTGATGTTGCTGAAGATAGAGCGACAGAGCAAGACAATATCTTTC
TGTTGGTGCTGATATTGCTGAGGATAAAGCGACGGGCAAGACAATATCTTTCTGTTGGTGCAGCGGAATGTTGCTGCGAGAAGATGAGAGAGCGACGAAGGAAGGCAAGTTTTGAGGCGAGCGGTCAAACACTTGCAAGCAGACGCCGAA
GATATGAGCGACAGGCAAGTTTGAGGCGAGCTAAACGCAAAGCAGACGCCGAATGACTGACGGCAGAGTTTTGCAGGCGAGCGGTCGAAACTTTGCCTGCATCAATATCGGCGATCTGCTTGGTGTTTGACCGTCGCTATCACTCTTCAG
CAATATCGCTTTCCAACGAAAATCAATGTTTAACCTGTCATCTAATATTTCGGCAATATCGCTGCTGAAAAATACGCTGCTGTCATCAATATTCGGCAATATCGGCACCGGCTGAAAGACTGTTGACTCGCCTGTCAATACTATCTTCGG
CAATATCAGCACCAACTATTCTGTTGACAGTGCTGATGTATTATTTGAAGATGGGGCGACGGGCAGAGTCCCGTGATTGCCTTGTACGGTATTTGCTTTGCGCGCGCGCGCAGAAGAATGGAAGCGACGCGGAAGGCAGAGTCATTGTAG
CCGCTATCGCCATTGTTACCGTTTGGGGTTTTACGATATTATTAGAAGGAAAAATCCCTTCTCATAAAATATCTTC+$#")$(&-165/100+,3..-211060017922423:7,)**8,3-50-&(+
+,/103336+,,--;2362.:0./.3.-.11060-(&-+--.**-)'()&,*,-/-(&((-)-54-.67:2331.-,)13747716621;,:-0')%$%&**/7110228*,/%%*)+.572./024236/.
+,,40:9/06,6/06244110-1/./8729/+'%&((,-18,,4511651*.1%((-.07.%1,/./0.-0+*+++6;5640/2,-.-+((*453.,'*+61,-.(0+4//2:102-3/4..++*)+,/,8,*(*,031,32
-)/1.5,8657:589,;,,-55:6480501*('`$%%`&&``#*&%%%##$'%))&))%(&&')*++*$$'(&-(-(-*-41098-+,/4537796;9:260-&(+'#),,-/1672123,1-,-5'')(//348')0/
+/..761,455.(&&).2*/*,&''&,''/)3'(%'%&%(%$%%$(%&&'%(233+&$$(/*)(,*0426$(*'(*/((`&$#$&('(&%$''*.,*(&(+''*'/),,,1-,-'%$$##$$"&,&)(+-22($('%&(*43,)((',
(,&%%+,''%%&%&%+,&+''&&&*(&++*,0*$,.*4.:/00,-15''&$'*+.79-*'%&')&$&&%(&'()()&%')(%$%%&(+60'(`&+,*-*$%,.)((&`,-;;-,&#%&%()&&&%$'*++,-((S&(+(+,0.&
$((+(,$%&$&)(())+')*%*,)('()*'+&+&#%$#'''--).,*(9990:97/,,--+-#$##$(+%'&''#$"$#$%&%)&$%)(%&*)'%%&''&('),*&%'('%#$%&)*)%(.'')&&%$%%&$'%&&'%/,*)*)
&'#&%#'%#&(#%$#$+'-+*%$$$%&''-*%($%$#$%&&%-%&%$$$%$$##$#%$'-*#&#$$$##@3a636cff-853b-4f74-a5af-bba045657dc8
```

## Raw Signal to Basecalling

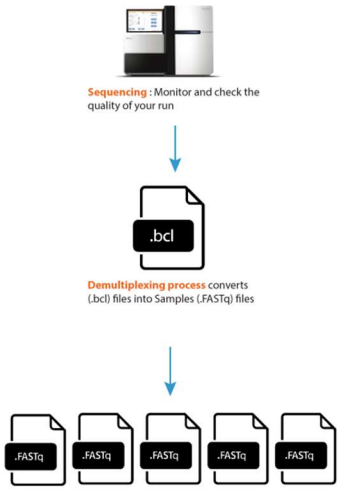



**Base calling, quality test**
**DNA/RNA modification detection**



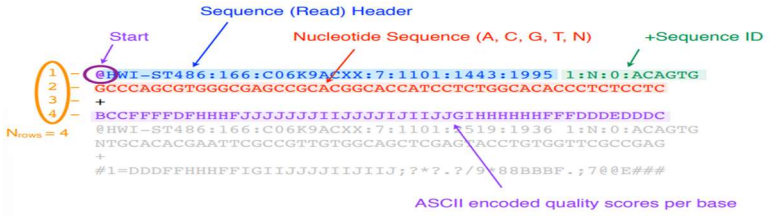

**Genome assembly**
**Haplotyping**
**Structural variation**

## Raw Data Output of NGS



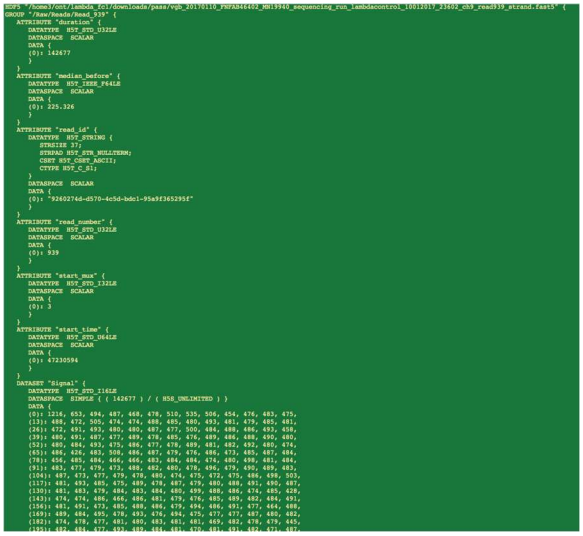

**Fastq format**

1. This format consists of four lines for each read.
2. The first line starts with an '@' character, followed by the sequencer identifier name given to the read by the sequencer.
3. The second line is the sequence of the read itself, the third line is simply a '+' symbol, acting as a spacer, and the fourth and final line is the Phred quality scores of the bases in the second line




## FAST5 format from Oxford Nanopore

The Oxford Nanopore (ONT) FAST5 format/HDF5, is a highly flexible data model, library, and file format for storing and handling data. It can store an infinite number of different data kinds.

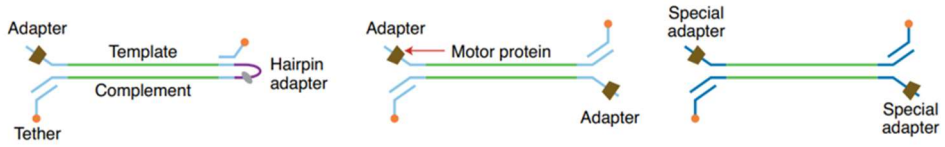


## Adapter removal

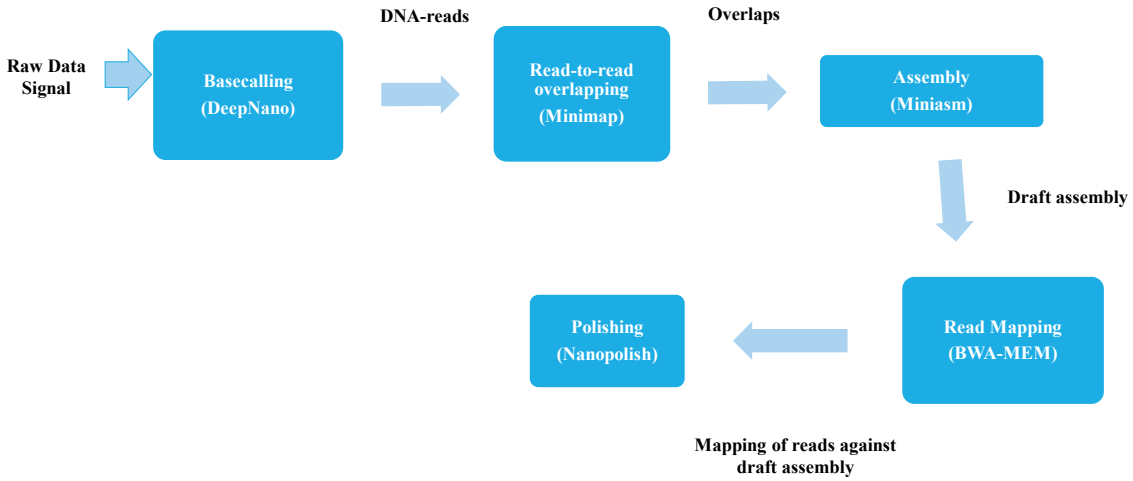
Most 3rd generation sequencing systems have linker, barcode, or adaptor sequences at the start or end of the read, similar to 2nd generation sequencing platforms.

Tools:

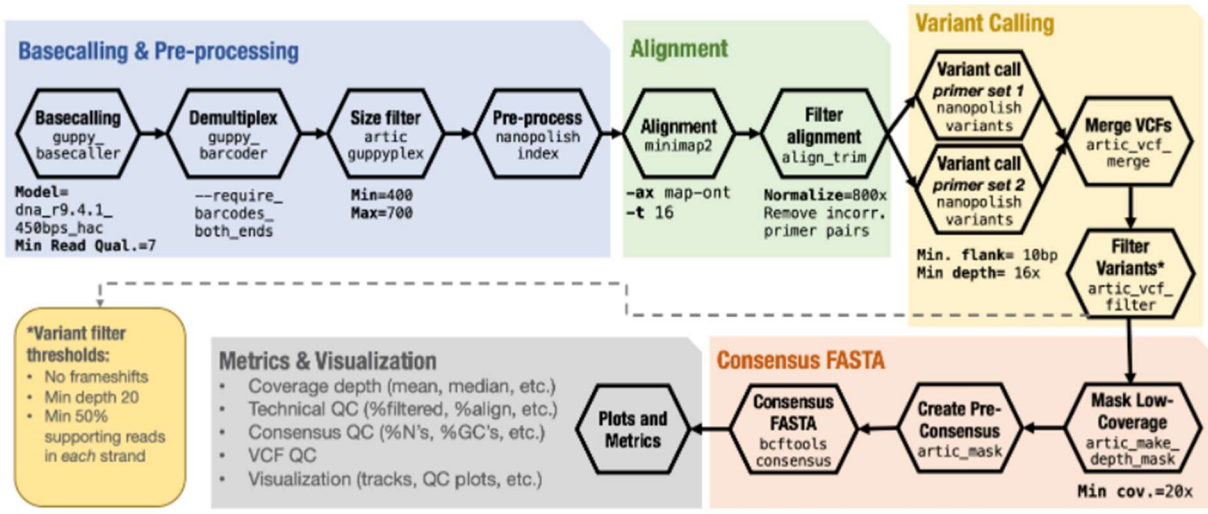
1. ONT's base-caller Guppy can trim adapters
2. Porechop



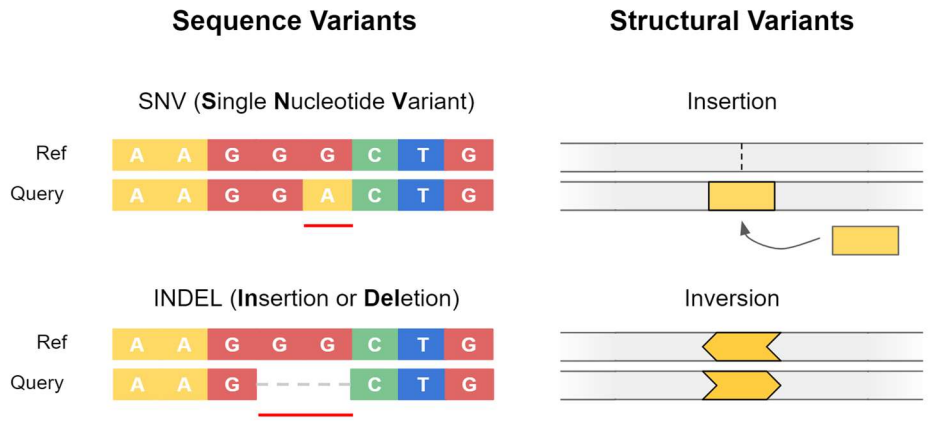

## SARS CoV-2 (2019) Genome assembly steps



## Overview of variant discovery




## Different types of variants



## Variant description

Variants/mutations/polymorphisms in DNA, RNA, and protein sequences are described using a specialized language; the HGVS nomenclature.

All variants are described in relation to a reference:

NM_004006.2:c.4375C>T

NC_000023.11:g.32389644G>A

a description of the variant

All variants are described in relation to a reference

## Genomic reference sequences

**Accession number** directly in front of the version number gives the number of the chromosome:

1-22, 23 for the X-chromosome and 24 for the Y-chromosome.

For NC_000023.10 "23" so a reference sequence of human **chromosome X**. ".10" is the version number.

NC_000023.9:g.32317682G>A : Chromosome

NG_012232.1:g.954966C>T: Gene or Genomic region

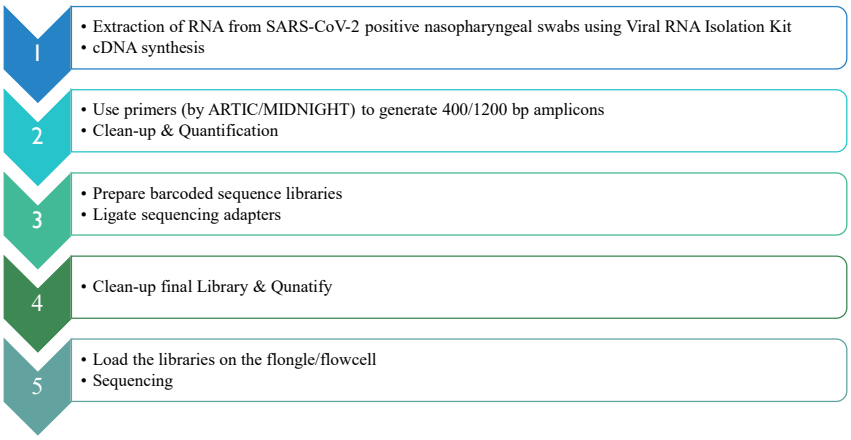
LRG_199:g.954966C>T: Gene or Genomic region

NM_004006.2:c.4375C>T

NP_003997.1:p.Arg1459*

# InterARTIC workflow

## Wet Bench Laboratory Workflow (recap)



## Overview of Bioinfo Analysis for SARS-CoV-2 Genome

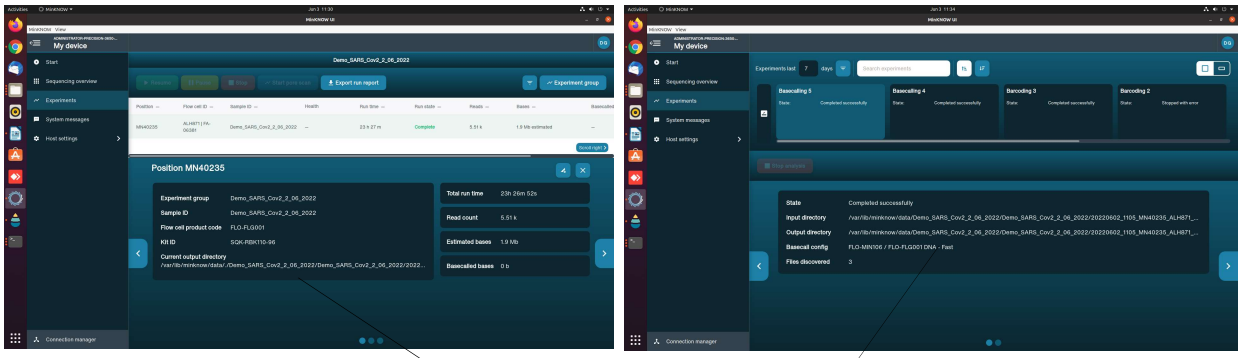
1. Demuliplexing using Minknow
2. FastQ Output from Minknow
3. Initialise InterARTIC
4. FastA output from InterARTIC
5. Save Consensus.fasta & variants table
6. Initialise Nextclade/Pangolin

## Basic file navigation and I/O functions

a) To know which directory you are in, you can use the **"pwd"**
b) Use **cat** to create a new .txt file
c) Use the **"ls"** command to know what files are in the directory you are in.
d) Use the **"cd"** command to go to a directory.
e) Use the **mkdir** command when you need to create a folder or a directory
f) Use the **rm** command to delete files and directories.
g) Use the **cp** command to copy files through the command line.
h) Use the **mv** command to move files through the command line.
i) The **locate** command is used to locate a file in a Linux system
j) Use **zip** to compress files into a zip archive, and **unzip** to extract files from a zip archive.

## MinKnow

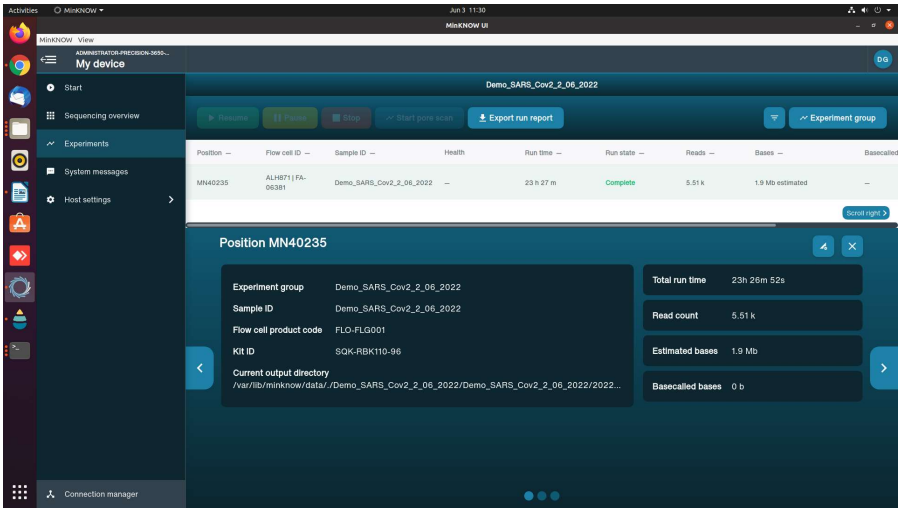
**Output page from which data is to be extracted**



The output directory index which has to be referred for the following files:

1. fast5
2. fastq

## Pre analysis QC

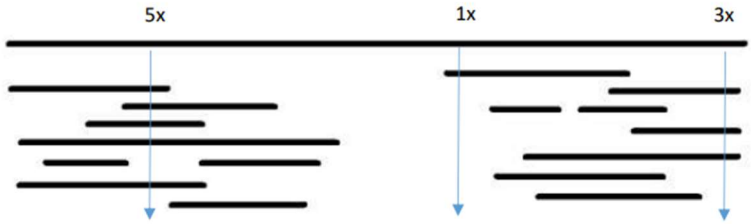



Following values should be noted prior to extracting data after MinKnow analysis:

1. Read count
2. Estimated bases
3. Basecalled bases

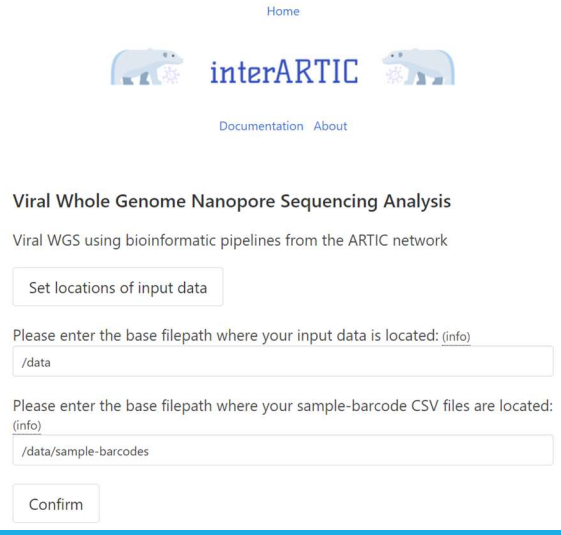
## PCR Amplicon Coverage

**Depth:** This is the amount of times a base within a genome has been sequenced. The greater the depth, the greater the confidence in the identity of the sequenced base.

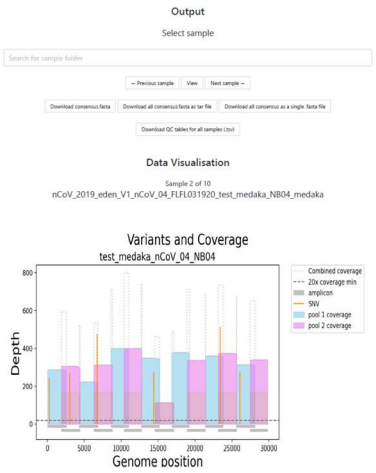



**Coverage:** this is the percentage of the whole genome that has been sequenced.

## InterARTIC usage



## Bioinformatics Result/QC Output




| | CHROM | POS | REF | ALT | QUAL | FILTER | DEPTH |
|---|---|---|---|---|---|---|---|
| 0 | MN908947.3 | 4540 | C | T | 500.0 | PASS | 153 |
| 1 | MN908947.3 | 8782 | C | T | 500.0 | dp | 612 |
| 2 | MN908947.3 | 9477 | T | A | 500.0 | PASS | 338 |
| 3 | MN908947.3 | 14805 | C | T | 500.0 | PASS | 344 |
| 4 | MN908947.3 | 14805 | C | T | 500.0 | PASS | 342 |
| 5 | MN908947.3 | 25979 | G | T | 500.0 | PASS | 215 |
| 6 | MN908947.3 | 28144 | T | C | 500.0 | PASS | 239 |
| 7 | MN908947.3 | 28311 | C | T | 500.0 | PASS | 216 |
| 8 | MN908947.3 | 28657 | C | T | 500.0 | PASS | 259 |
| 9 | MN908947.3 | 28863 | C | T | 500.0 | PASS | 244 |

## Analysis QC

The following values should be noted for future study

1. QUAL
2. FILTER
3. DEPTH

| | CHROM | POS | REF | ALT | QUAL | FILTER | DEPTH |
|---|---|---|---|---|---|---|---|
| 0 | MN908947.3 | 4540 | C | T | 500.0 | PASS | 153 |
| 1 | MN908947.3 | 8782 | C | T | 500.0 | dp | 612 |
| 2 | MN908947.3 | 9477 | T | A | 500.0 | PASS | 338 |
| 3 | MN908947.3 | 14805 | C | T | 500.0 | PASS | 344 |
| 4 | MN908947.3 | 14805 | C | T | 500.0 | PASS | 342 |
| 5 | MN908947.3 | 25979 | G | T | 500.0 | PASS | 215 |
| 6 | MN908947.3 | 28144 | T | C | 500.0 | PASS | 239 |
| 7 | MN908947.3 | 28311 | C | T | 500.0 | PASS | 216 |
| 8 | MN908947.3 | 28657 | C | T | 500.0 | PASS | 259 |
| 9 | MN908947.3 | 28863 | C | T | 500.0 | PASS | 244 |

## QC contd..

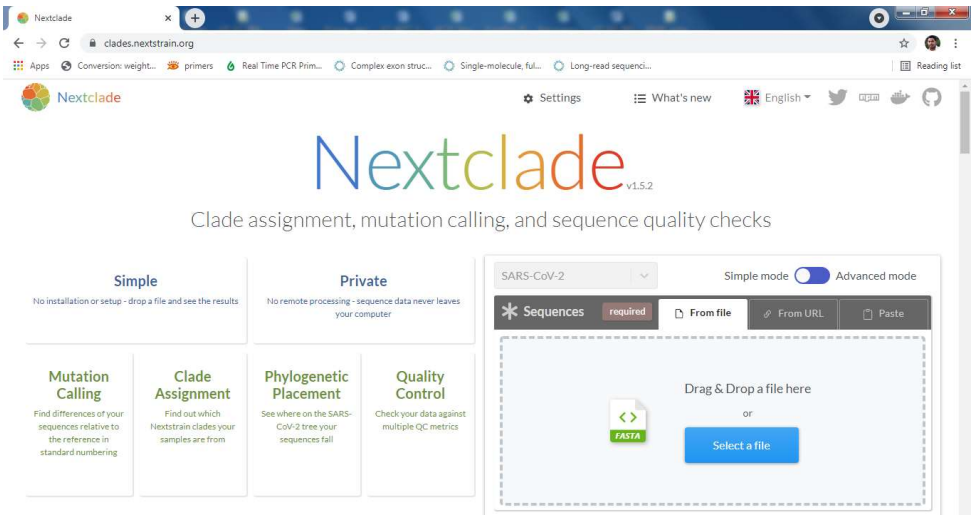
Amplicon wise result

The following values should be noted and recorded for further use:

1. Pass read count
2. Read quality mean
3. Amplicon's mean coverage

| | Metric | Value |
|---|---|---|
| 0 | Sample | nCoV_2019_midnight_V1_NcoV_02_run_03_06_2022-1_03-06-2022_medaka_RB02_medaka |
| 1 | Pass read count | 11823 |
| 2 | Read length mean | 647 |
| 3 | Read length stdev | 261.1 |
| 4 | Read quality mean | 18.69 |
| 5 | Read quality stdev | 2.07 |
| 6 | Total median coverage | 129 |
| 7 | Total mean coverage | 116 |
| 8 | Amplicon 1 mean coverage | 0 |
| 9 | Amplicon 2 mean coverage | 85 |
| 10 | Amplicon 3 mean coverage | 3 |

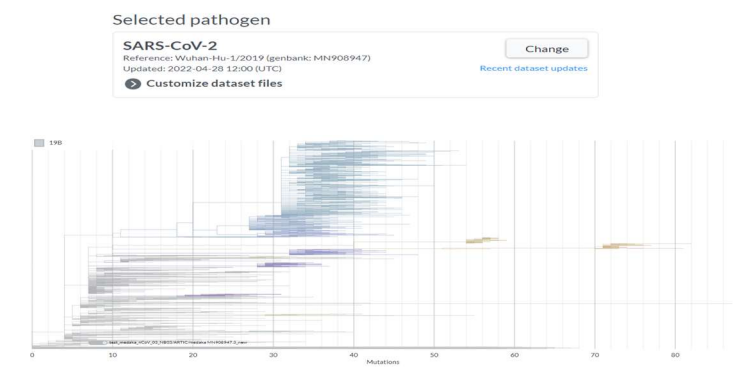
## Lineage Identification Tool: Nextclade



## Visualisations & Lineage Identification using Nextclade

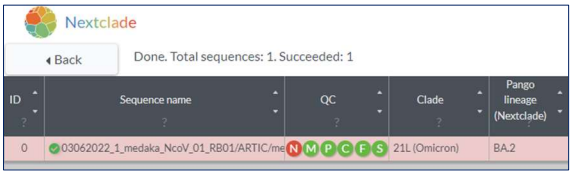
An online tool to convert consensus viral assembled fasta sequence into relevant annotations.

1. Phylogenetic tree
2. Mutations
3. Amino acids substitutions



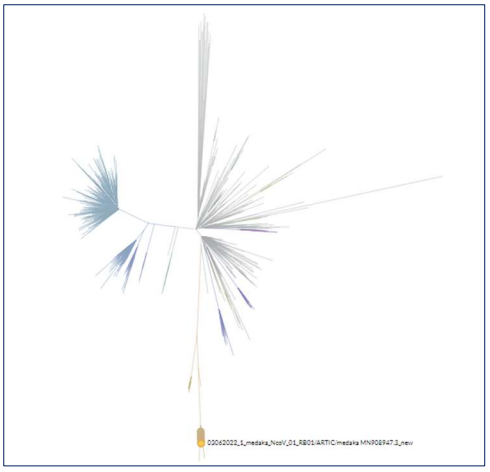

## Nextclade Output

1. Lineage

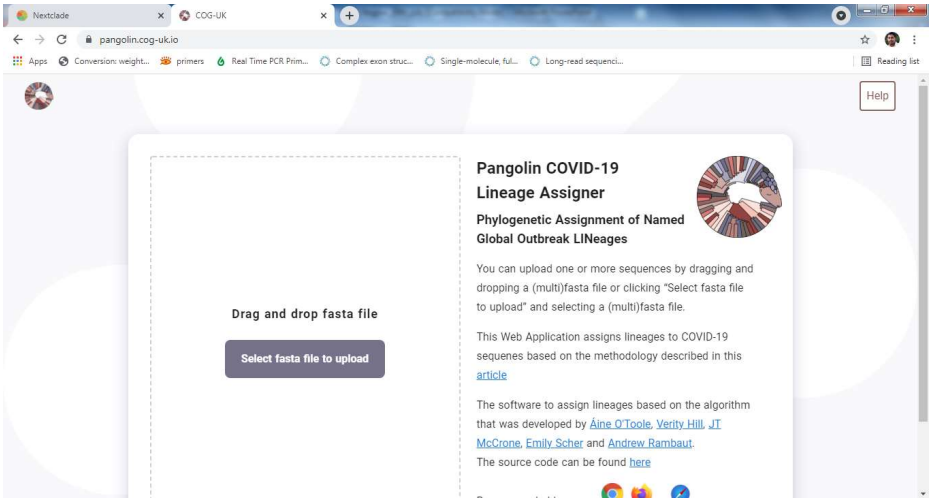



2. Variant details

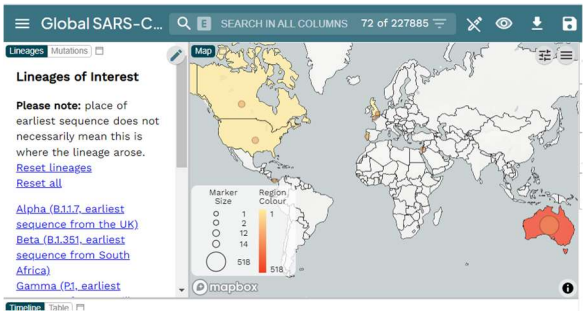

3. Phylogenetic tree



## Lineage Identification Tool: Pangolin




## Visualisation contd..

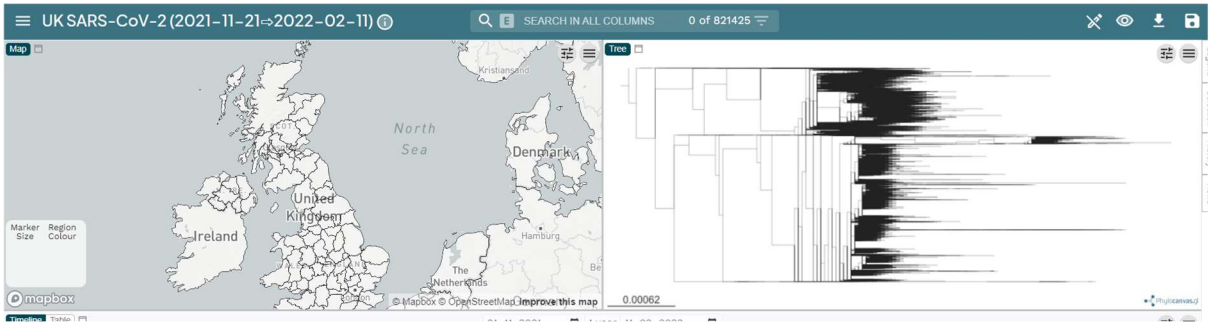




**Genome visualiser Pango**

**This tool also can provide us with**

1. **Lineage information**
2. **Lineage tracking**

## Lineage tracking using Pango




## References


1. Pereira, R., Oliveira, J., & Sousa, M. (2020). Bioinformatics and computational tools for next-generation sequencing analysis in clinical genetics. Journal of clinical medicine, 9(1), 132. https://doi.org/10.3390/jcm9010132
2. MinKNOW Post-run basecalling. https://community.nanoporetech.com/docs/prepare/library_prep_protocols/experiment-companion-minknow/v/mke_1013_v1_revcq_11apr2016/post-run-basecalling, requires ONT community access.
3. Ferguson JM, Gamaarachchi H, Nguyen T, Gollon A, Tong S, Aquilina-Reid C, Bowen-James R, Deveson IW. InterARTIC: an interactive web application for whole-genome nanopore sequencing analysis of SARS-CoV-2 and other viruses. Bioinformatics. 2022 Feb 7;38(5):1443-1446. doi: 10.1093/bioinformatics/btab846. PMID: 34908106; PMCID: PMC8826086 (https://github.com/Psy-Fer/interARTIC/).
4. Aksamentov, I., Roemer, C., Hodcroft, E. B., & Neher, R. A., (2021). Nextclade: clade assignment, mutation calling and quality control for viral genomes. Journal of Open Source Software, 6(67), 3773, https://doi.org/10.21105/joss.03773 (https://clades.nextstrain.org).
5. Pangolin COVID-19 Lineage Assigner (https://pangolin.cog-uk.io/).
6. Oxford Nanopore Technologies. https://nanoporetech.com/

## 6. Tool Kit

### A comprehensive manual for Next-Generation Sequencing with a focus on SARS-CoV-2

## Concepts covered in the Tool kit

- Introduction to Next-Generation Sequencing
- Comparison of NGS platforms
- NGS Workflow
- Disease outbreak and COVID-19 pandemic
- Bioinformatic analysis
- Setting up an NGS lab

## Tool Kit Chapters Details On…

- Current knowledge of NGS technology development, protocols, and best practices.
- Provides overview of the human genome project, followed by an in-depth look at various sequencing platforms, evolution, and their comparisons, along with a complete NGS workflow process.
- Contains chapters on COVID-19 genome surveillance, sequencing protocols using Nanopore technology, and an overview of the available bioinformatic tools developed specifically for biologists, besides the myriad conventional tools.
- Provides details on setting up of an NGS lab and the associated quality control procedures required in the laboratory.

## How to Access Tool Kit

A pdf version of the tool kit can be downloaded from the link provided below.

https://www.finddx.org/wp-content/uploads/2022/10/20221031_rep_india_toolkit_next_gen_sequencing_FV_EN.pdf

# Chapter 2: Annexures For the Master Trainers

**This chapter provides details on the trainer's script for the master trainers covering the following topics.**

- Demonstration and hands-on training on the use of laboratory cabinet(s) in the COVID-19 molecular diagnostics laboratory
- Use of Micropipettes
- Demonstration and hands-on training on the steps performed in the library preparation, sequencing, and bioinformatics analysis of the SARS-CoV-2 genome in a COVID-19 molecular diagnostics laboratory.
- Trainer's manual.

**Title:** Demonstration and hands-on training on the use of laboratory cabinet(s)** in the COVID-19 molecular diagnostics laboratory.

**Purpose:** This document describes demonstrations and hands-on exercises for safe work practices inside the cabinets. The key steps performed inside the cabinet during cDNA synthesis, SARS-COV-2 PCR, library preparation, and sequencing are demonstrated and practiced.

**Trainee Learning Objectives:**

- To operate and use the cabinet(s) correctly.
- To carry out the procedure inside the cabinet safely.
- To learn the correct procedure for handling the reagents and proper disposal of waste arising from procedures.

**Reference documents:**

- SOP for NGS library preparation & sequencing.
- SOP for cabinet use & maintenance.
- Equipment & material: Refer to the checklist provided in the SOP for library preparation & sequencing of SARS-CoV-2.

**Equipment**

- Certified Biosafety Cabinet Class II Type A2 or Laminar flow hood or PCR cabinet (whichever is available).
- Vortex mixer.
- -20°C/ -80°C (with free space for reagent and intermediate product storage).
- 4°C Refrigerator.
- Tabletop refrigerated centrifuge.
- Microfuge or Mini Spinner.

**Consumables**

- 0.5-10, 10-100, 20-200 & 100-1000µl pipettes & its compatible tips.
- 8-well PCR strips or 96-well plate.
- 1.5ml Microcentrifuge tubes.
- Biohazard bags.
- Markers.
- Absorbent liner.
- Twist tag (to tie the bags).

**Personal Protecting Equipment**

- Coverall/ Gowns.
- Gloves.
- Goggles..
- Shoe covers
- Head covers.

**Disinfectants**

- 70% Ethanol.
- 1% Sodium Hypochlorite stock solution.

**Updated Logbooks**

- Laboratory Cabinet use logbook.
- Maintenance logbook.
- Centrifuge use logbook.

**Mock Specimens**

- 80% Glycerol, Water, 80% Alcohol.

**Items to be provided to the trainee beforehand**

- SOP for Cabinet use.
- SOP for NGS library preparation & sequencing.

**Instructions to the trainer:**

- Please read the SOP and trainer script beforehand.
- Prepare for demonstration by collecting all the required material.
- Ensure you have the space and time for the number of trainees in your demonstration.
- Provide the SOP to the trainee and ensure its review by the trainee before the start of the exercise.
- Provide safety guidelines in case of adverse events such as chemical spills etc,. and ensure that are read and understood by trainees.
- Ensure that trainees are aware of biomedical waste management procedures followed.

| **Demonstration and hands-on training on the use of laboratory cabinet(s) for SARS-CoV-2 Sequencing workflow** | | |
|---|---|---|
| **Training Steps** | **Trainers Task** | **Trainer's observations** |
| Turning on the laboratory cabinet | The trainer shows the correct order of turning on the laboratory cabinet:<br>• UV Sterilization of chamber before use.<br>• Correct sash level.<br>• Turn on the blower after UV sterilization for 10 mins (if applicable for the cabinet).<br>• Magnahelic gauge reading and comparing with certification values.<br>• Entering the details in the logbook. | |
| Getting ready to work | **Task: setting up the work area**<br>• Assigns a trainee to set up the work area. Observes, whether the following practices followed:<br>o Checklist consulted.<br>o Disinfection carried out as per the protocol.<br>o Solid and liquid biohazard waste collection.<br>o Clean to dirty work area division achieved.<br>• Comment on right vs wrong practice. | |
| **Working inside the laboratory cabinet –**<br>Aliquoting or transfer of reagents and sample to microcentrifuge tubes or PCR strips or PCR plate | **Task: Aliquoting and transfer of reagents and sample into the reaction tube or PCR plate**<br>• Trainer demonstrates the right practice to aliquot or transfer at least with 4-5 samples at several steps in the procedure (mock).<br>• Trainer should assign a minimum of two samples per trainee to demonstrate aliquoting, transfer and pipette mixing procedure.<br>• Focus on pipetting techniques to avoid cross-contamination | |

| | | |
|---|---|---|
| | and aerosol generation.<br>• Focus on discarding the waste solution and used pipette tips.<br>• Comment on right vs wrong practice. | |
| **Working inside the laboratory cabinet –**<br>Cleaning the BSC post work | **Task: Cleaning the work area after work completion**<br>• Trainer assigns trainees to perform:<br>• Cleaning of the work area.<br>• Removing solid and liquid waste from the cabinet.<br>• Final surface cleaning.<br>• Transfer of microcentrifuge and PCR strips or PCR plate to the freezer/ refrigerator.<br>• UV Sterilization of chamber after use.<br>• Discuss right vs wrong practice | |
| **Working inside the laboratory cabinet –**<br>Transfer of microcentrifuge tubes and PCR strips to refrigerator or freezer | **Task: Storing the reagents and intermediate products and shut down of Cabinet.**<br>• Demonstrates correct storage procedure.<br>• Cleaning up the laboratory before leaving.<br>• Turning off the cabinet.<br>• Documenting lab activity.<br>• Exiting the Lab. | |

** Cabinets could refer to Biological Safety Cabinet or Laminar flow or PCR cabinet whichever is available in the laboratory for use.

**Reference:**

1) For detailed use of Biosafety cabinet and maintenance, please refer to the tool kit – "A comprehensive manual on SARS-CoV-2 diagnostics" from the link: https://www.finddx.org/wp-content/uploads/2022/08/20220819_usaid_covid_toolkit_FV_EN.pdf

**Title:** Use of Micropipettes

**Purpose:** This document describes demonstrations and hands-on exercises on the correct use of micropipettes in a PCR-based diagnostics laboratory.

**Equipment & material:**

**Micropipettes:**

- Certified micropipettes: (at least 2 sets for 10 trainees)
- 1-20 µl
- 200 µl
- 1000 µl
- Multichannel pipettes

**Other consumables:**

- Tube racks
- PCR tubes/96 well plates
- Plate covers
- Plate sealers
- 1.5 ml centrifuge tubes
- Tips for all the micropipettes
- Troughs for multichannel plates
- 96 well RNA extraction plate
- Precision balance (accuracy up to 10 mg)
- Paper boats
- Distilled water aliquots
- Notebook
- Marker
- Pens

**Disinfectants:**

- 1% sodium hypochlorite
- 70% Ethanol

**Mock Reagents**

- Nuclease-free water
- 80% glycerol
- Dye solution

**Reference material for the trainer:**

1. Micro pipetting:
    - https://www.youtube.com/watch?v=VEkfBStZSNc (reverse pipetting)
    - https://www.youtube.com/watch?v=1Y0U9jf5ZbI (reverse pipetting multichannel)
    - https://www.youtube.com/watch?v=QGX490kuKjg
    - https://www.youtube.com/watch?v=uEy_NGDfo_8&t=195s
2. Sealing the thermocycler plate:
    - https://www.youtube.com/watch?v=25OOn6W5gU0&t=10s

**Instructions to the trainer:**

- Please read the script beforehand.
- Prepare for the demonstration by collecting all the required material beforehand and setting up thedemonstration area.
- Practice your training to get an idea on time requirements and preparation gaps.
- Ensure you have the space and time for the number of trainees in your demonstration.

| **Demonstration and hands on exercises on the use of micropipettes in PCR based diagnostic laboratory** | | |
|---|---|---|
| **Training Steps** | **Trainer's Task** | **Trainer's observations** |
| Introduction & Learning objectives | • Describe the importance of correct pipetting techniques and their maintenance, especially calibration.<br>• Inform about learning objectives of the training. | |
| Micropipette volume range – Calibration & Volume setting | • Show types of micropipettes, describe their volume range.<br>• Show the certification label on the pipettes (or the calibration record).<br>• Show how to correctly hold the pipette.<br>• Show how to correctly set the volume and attach tips for each micropipette. | |
| Aspiration and dispensing liquids correctly | • Show how to aspirate the liquids.<br>• Angle of aspiration.<br>• How much to dip the tip in the liquid (immersion depth) during aspiration?<br>• How to do reverse pipetting?<br>• How to dispense?<br>• How to eject the tip into the wastebin? | |
| Measuring consistency of repeat pipetting | • Show the accuracy of repeat measurement by carrying out repeat pipetting of a dye solution on a piece of paraffin/wax paper. (Use 5-100µl volume for this exercise).<br>• Visual inspection of the size of the dye - drops are carried out for repeated volumes to show the consistency of repeated pipetting.<br>• Ask the trainee to test their pipetting skills in this way. | |
| Measuring pipetting accuracy | • Show the accuracy of measurement by carrying out repeat pipetting of distilled water on a paper boat. (Use 10-20 ul volume for this exercise).<br>• Repeated pipette volumes are noted on the paper and compared later for accuracy.<br>• Ask the trainee to test their pipetting skills in this way, once they return to their labs | |

| | | |
|---|---|---|
| Using Multichannel pipettes | • Show proper way of inserting tips.<br>• Using troughs for dispensing molecular-grade reagents.<br>• Reverse pipetting.<br>• Dispensing liquids to 96 well plates, avoiding bubbles.<br>• Common issues with multichannel pipettes. | |
| Pipetting PCR reaction mixes in 96well plates | • Show how to load the PCR reaction in a 96 well plate.<br>• How to seal the plate?<br>• Show how to use PCR tubes.<br>• How to cap them?<br>• Show how to spin the 96 well plates/tubes.<br>• Look for air bubbles and how to get rid of them. | |

**Title:** Demonstration and hands-on training on the steps performed in the library preparation, sequencing, and bioinformatics analysis of the SARS-CoV-2 genome in a COVID-19 molecular diagnostics laboratory.

**Purpose:** This document describes the various steps that are followed to prepare library preparation, sequencing, and bioinformatics analysis of the SARS-CoV-2 genome using Oxford Nanopore Sequencer.

**Trainee Learning Objectives:**

- To carry out NGS library preparation, sequencing, bioinformatics analysis and, interpretation of the results.
- To learn the corrective actions in case of QC failures.

**Duration of Training:**

Presentation –2 hrs and 30 minutes

Hands-on Training – 3 days

**Reference documents:**

- SOP for NGS library preparation and sequencing.
- SOP for SARS-CoV-2 bioinformatics analysis.
- Equipment & material: Refer to the checklist provided in SOP for library preparation & sequencing of SARS-CoV-2.

**Equipment**

- Certified Biosafety Cabinet Class II Type A2/Laminar Hood/PCR Cabinet
- Vortex mixer
- -70°C/ -80°C (with free space for sample storage)
- -20°C Freezers
- 2 - 8°C Refrigerator
- Tabletop refrigerated centrifuge
- Spinner
- Pipettes -1000µl, 200µl, 100µl, 10µl
- 1.5ml Centrifuge tubes
- Real-Time PCR/PCR Instrument

**Consumables**

- 0.5-10, 10-100, 20-200 & 100-1000µl pipettes & its compatible tips.

- 8-well PCR strips or 96-well plate.
- 1.5ml Microcentrifuge tubes.
- Biohazard bags
- Markers
- Twist tag (to tie the bags)
- Kits for NGS library preparation and sequencing

**Personal Protecting Equipment**

- Coverall/ Gowns
- Gloves
- Goggles
- Shoe covers
- Headcovers

**Disinfectants**

- 70% Ethanol
- 1% Sodium Hypochlorite stock solution

**Updated Logbooks**

- Laboratory Cabinet use logbook
- Maintenance logbook
- Centrifuge use logbook
- Instrument use logbook
- Worksheets for sample entries

**Mock Specimens**

- Nuclease-free water
- 80% glycerol
- 80% Ethanol

**Items to be provided to the trainee beforehand**

- SOP for NGS library preparation & sequencing.
- SOP for Bioinformatics analysis.

**Instructions to the trainer:**

- Please read the SOPs, QA plan, and trainer script beforehand.

- Prepare for demonstration by collecting all the required material.
- Ensure you have the space and time for the number of trainees in your demonstration.
- Provide the SOP to the trainee and ensure its review by the trainee before the start of the exercise.
- Provide safety guidelines in case of adverse events such as chemical spills etc. and ensure that are read and understood by trainees.
- Ensure that trainees are aware of biomedical waste management procedures followed.

| Demonstration and hands-on training on the library preparation, sequencing, and bioinformatics analysis of SARS-CoV-2 Genome | | |
|---|---|---|
| **Training Steps** | **Trainer's Task** | **Trainer's observations** |
| **Getting ready the laboratory to work** | **Task: Setting up the work area**<br>o Assigns a trainee to set up the workarea. Observe, whether the following practices are followed:<br>o Checklist entry and cross-checked<br>o Ensure the reagents are stored at an appropriate temperature as per the manufacturer's instruction<br>o Check the reagent's date and expiry.<br>o Check and arrange the RNA samples and cross-check with the laboratory numbers.<br>o Ensures the known samples are aliquoted and in the required amount for the assay.<br>o Proper labeling of the vials and entering the respective worksheets.<br>o Ensures proper biosafety practices including PPE are in place.<br>o The instruments and the pipettes are calibrated, and a proper check is done before starting the assay.<br>Entering the details in the logbook comments on the right vs wrong practice. | |
| **Working on the cDNA synthesis, PCR Setup, library Preparation & sequencing** | **Task: How to carry out library preparation from Viral RNA Sample**<br>• Trainer demonstrates how the cDNA synthesis, PCR setup, library preparation and sequencing is carried out stepwise emphasizing on the protocol with at least 3 samples and | |

| | | |
|---|---|---|
| | one control (mock).<br>• Trainer ensures that the cDNA synthesis, PCR setup, library preparation, and sequencing are carried out in the rooms designated for master mix preparation, pre-PCR & post PCR activity.<br>• Trainer should emphasize the importance of proper handling of pipettes and how to avoid cross contaminations.<br>• Trainer emphasizes on proper labeling and storage of the intermediate products in the library preparation & final library for sequencing.<br>• Trainer demonstrates the importance of cross-check signatures during each and every step.<br>• Trainer also ensures to comment on right vs wrong practice. | |
| **Working on the setting up of sequencing run** | **Task: How to prime the flow cell and loading of the library**<br>• Focus on pipetting techniques to avoid introduction of air bubbles and damaging integrity of pore membrane.<br>• Trainer ensures that the reagents and sequencing mix are well mixed before loading.<br>• Focus on the setting up of MinkNOW application and initiate the sequencing run.<br>• Focus on the returning of samples, intermediate products, and master mix reagents back to the proper temperature storage.<br>• Emphasize on the cleaning of the work area and pipettes after work completion.<br>• Demonstrate the quality control indicators. | |

| | | |
|---|---|---|
| | • Comments on right vs wrong practice. | |
| **Working on the Bioinformatics Analysis & Interpretation** | **Task: Launching bioinformatics tools & interpretation of results**<br>• Trainer assigns trainees to perform:<br>• Checking the QC metrics and overall performance of the sequencing run.<br>• Ensures the data is stored and maintained in a dedicated location.<br>• Ensures the analysis tools are launched properly without errors.<br>• Ensures all the results are cross verified.<br>• Interpret the results of the samples as per the SOP.<br>• Ensure the known control samples results are as expected.<br>• Document and release the reports if the controls are fine and known control samples results are as expected.<br>• Discuss right vs wrong practice | |
| **QC Testing** | **Task: QC Testing**<br>• Ensures the QC testing is done periodically with the designated reference laboratory as per the state guidelines.<br>• Ensure the proper aliquoting, storage, and transport of the samples to the designated laboratory.<br>• Ensures proper documentation of the ILQC tests | |
| **Working on the Troubleshooting Aspects** | **Task: Troubleshooting of Assay**<br>• Trainer assigns 2 different scenarios per trainee to demonstrate the same.<br>• Trainer to focus on the root cause analysis and instruct the action to be taken. | |

1. **General Laboratory Training****
   - **i)** Lab work Protocols
   - **ii)** Guidelines on Use of PPE
   - **iii)** Guidelines for Usage of Cabinets, Pipettes and All Instruments
   - **iv)** Biomedical Waste Management

2. **Laboratory Testing Protocol Training**
   - **i) Sample processing for SARS-CoV-2 genome sequencing**

     ***cDNA Synthesis, PCR, Library preparation & sequencing***

     (1) Checking of Reagents and Consumables as per the Checklist

     (2) Storage of RNA at the appropriate temperature

     (3) Processing of Extracted RNA for cDNA

     (4) Setting up of PCR using cDNA as input

     (5) Preparation of libraries using PCR amplicons

     (6) Priming and loading of flow cell

     (7) Setting up of the Nanopore instrument

     (8) Storing of reagents and extracted RNA after the experiment

     (9) Launching of bioinformatics tools & initiating analysis

     (10) Interpretation of results

     (11) Quality and Quantity check at required stages

     (12) Troubleshooting

     (13) Documentation

3. **Quality Control Procedures Training****

   (1) Personnel Responsibilities

   (2) Proficiency Testing

   (3) Interlaboratory Comparison Testing

   (4) Instrument Maintenance and Calibrations

   (5) Competency Assessment

   (6) Training for Laboratory technicians at regular intervals

****References**

For Si. No (1) & (3) training materials, please refer to the **tool kit** – "A comprehensive manual on SARS-CoV-2 diagnostics" from the link: https://www.finddx.org/wp-content/uploads/2022/08/20220819_usaid_covid_toolkit_FV_EN.pdf

# Chapter 3: Standard Operating Procedures

**This chapter provides details on standard operating procedures followed in the SARS-CoV-2 genome sequencing laboratory for the master trainers covering the following topics.**

- Midnight amplicon sequencing using rapid barcode protocol for SARS-CoV-2 Genome
- ARTIC amplicon sequencing using native barcoding protocol for SARS-CoV-2 Genome
- Whole Genome Analysis of SARS-CoV-2 using InterARTIC pipeline
- Flushing, Reloading and Storing of Mk1B Flow cell
- Annexure I: Fluorometric quantification of dsDNA
- Annexure II: Flongle/Flow cell QC Check
- Annexure III: Consumables & Reagent requirements for SARS-CoV-2 sequencing workflow
- Annexure IV: Recommended storage conditions of the reagents/consumables used in the process
- Annexure V: Sample processing sheet template

# SOP-001

# Midnight amplicon sequencing using rapid barcode protocol for SARS-CoV-2 Genome

| | LABORATORY STANDARD OPERATING PROCEDURES | | | |
|---|---|---|---|---|
| Country and Lab Demo Site | India | **DOCUMENT NUMBER** | **VERSION** | **EFFECTIVE DATE** |
| | | SOP-001 | 2.0 | |
| **PROCEDURE** | **Midnight amplicon sequencing using rapid barcode protocol for SARS-CoV-2 Genome** | | | |

| | | | | |
|---|---|---|---|---|
| Approved By | Name, Title | | Signature | Date |
| | | | | |
| | | | | |
| SOP Annual Review | Name, Title | | Signature | Date |
| | | | | |
| | | | | |
| | | | | |
| Revision History | Version # [0.0] | Revision Date [dd/mm/yy] | Description (notes) | |
| | | | | |
| | | | | |
| | | | | |
| Distributed Copies to | Name (or location) | # of copies | Name (or location) | # of copies |
| | | | | |
| | | | | |
| | | | | |

***I acknowledge that I have read, understand and agree to follow this SOP.***

| Name (print) | Signature | Designation | Date |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

# Table of Contents

## Objective

To perform whole genome sequencing of SARS-CoV-2 using Nanopore (Mk1b) instrument.

## Clinical Significance / Intended Use

It is essential to characterize the virus that is spreading rapidly & to estimate a particular variant's prevalence in a population to investigate the virus outbreaks through genome surveillance.

## Library Preparation Procedure

### Materials Required

1. Flow Cell (R9.4.1) FLO-MIN106D or Flongle FLO-FLG001
2. Flow Cell Wash Kit (EXP-WSH004)
3. Rapid Barcoding Kit 96 (SQK-RBK110.96)
4. Input RNA (< 25 / 25 Ct value)

### Consumables Required

1. LunaScript RT SuperMix Kit (NEB, E3010L)
2. Q5 Hot Start High-Fidelity 2X Master Mix (NEB, M0494L)
3. COVID ARCTIC WGS Primer Pools A& B (300 rxns)
4. PurBeads 60ml
5. Nuclease-free water
6. Freshly prepared 80% ethanol in nuclease-free water
7. Qubit dsDNA HS Assay Kit (ThermoFisher, Q32851)
8. 1.5 ml/2 ml Eppendorf tubes
9. 96 well PCR plates with (semi skirted) with heat seals

### Equipment required

1. Magnetic separator, suitable for 1.5 ml Eppendorf tubes
2. 96 well - Centrifuge
3. Microfuge
4. Vortex mixer
5. Thermal cycler (conventional/ real time instrument)
6. Multichannel pipettes suitable for dispensing 0.5–10 µl, 2–20 µl and 20–200 µl,
7. P1000 pipette
8. P200 pipette
9. P100 pipette
10. P20 pipette
11. P10 pipette
12. Filter tips (for the required range of pipettes as mentioned above)

13. Ice bucket with ice
14. Timer
15. Qubit fluorometer (or equivalent for QC check)
16. Eppendorf 5424 centrifuge (or equivalent)
17. PCR hood with UV sterilizer (optional but recommended to reduce cross-contamination)
18. PCR-Cooler (Eppendorf)

### Reverse Transcription PCR

Note: Keep the RNA sample on ice as much as possible to prevent nucleolytic degradation, which may affect sensitivity.

**Step 1:** In a clean pre-PCR hood, using a single or a multichannel pipette, add 2 µl of LunaScript RT SuperMix to a fresh 96-well plate (RT Plate).

**Step 2:** To each well containing LunaScript reagent of the RT plate, add 8 µl of RNA sample and gently mix by pipetting.

**Step 3:** Seal the RT plate and spin down. Return the plate to ice.

**Step 4:** Preheat the thermal cycler to 25°C.

**Step 5:** Incubate the samples in the thermal cycler using the following program:

| Step | Temperature | Time | Cycles |
|---|---|---|---|
| Primer annealing | 25°C | 2 min | 1 |
| cDNA synthesis | 55°C | 10 min | 1 |
| Heat inactivation | 95°C | 1 min | 1 |
| Hold | 4°C | ∞ | |

While the reverse transcription reaction is running, prepare the primer pool master mix as described in the next section.

### Midnight SARS-CoV-2 PCR

These primers are designed to generate 1200 bp amplicons that overlap by approximately 20 bp.

**Step 1:** In the pre-PCR hood, prepare the following master mixes in 1.5 ml/ 2 ml eppendorf tubes and mix thoroughly as follows:

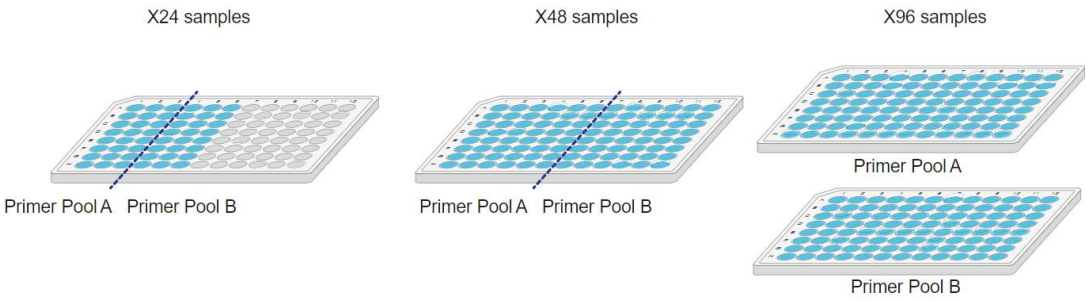
Volume per sample:

| Reagent | Pool A | Pool B |
|---|---|---|
| Nuclease free water | 2.25 µl | 2.25 µl |
| Primer pool A | 1.5 µl | - |
| Primer pool B | - | 1.5 µl |
| Q5 Hot Start HF 2x Master Mix | 6.25 µl | 6.25 µl |
| **Total** | **10 µl** | **10 µl** |

**Step 2:** Using a single or a multichannel pipette, aliquot 10 µl of Pool A and Pool B into a clean 96-well plate(s) as follows:

| Plate location | X24 samples | X48 samples | X96 samples |
|---|---|---|---|
| Columns | Pool A: 1-3<br>Pool B: 4-6 | Pool A: 1-6<br>Pool B: 7-12 | Pool A: 1-12<br>Pool B: 1-12 |

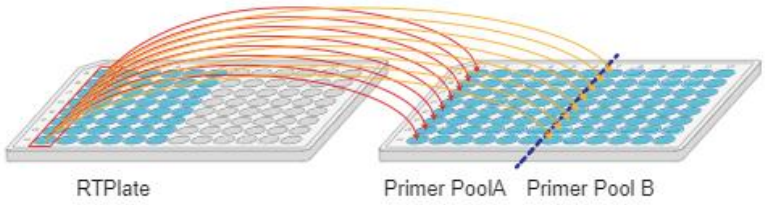
Note: For x96 samples, Pool A is a separate plate to Pool B.




**Step 3:** Using a multichannel pipette, transfer 2.5 µl of each RT reaction from the RT plate to the corresponding well for both Pool A and Pool B of the PCR plate(s). Mix by pipetting the contents of each well up and down.

- Carry forward the negative control from the reverse transcription reaction to monitor cross-contamination events.
- There should be two PCR reactions per sample.

Example for X48 samples:




**Step 4:** Seal the plate(s) and spin down briefly.

**Step 5:** Incubate using the following program, with the heated lid set to 105°C:

| Step | Temperature | Time | Cycles |
|---|---|---|---|
| Initial denaturation | 98°C | 30 sec | 1 |
| Denaturation<br>Annealing<br>Extension | 98°C<br>61°C<br>65°C | 15 sec<br>2 min<br>3 min | 32** |
| Hold | 4°C | ∞ | |

** For Ct Value 20/<20 – Keep 25 cycles and Ct value 20-30 – keep 30 cycles.

Note: If this is not feasible, laboratories can adopt 30 cycles uniformly Ct values of the number of positive samples selected for sequencing may fall within a range of 15- 30. Normalising the samples should be done when concentration is measured by Qubit before pooling the PCR products.

When PCR reaches 20-30 cycles, assemble the Rapid Barcode reaction plate as described in the next section.

Optional: Check the quality of the amplification by running an agarose gel.

### Addition of rapid barcodes

**Step 1:** Spin down the Rapid Barcode Plate and PCR reactions prior to opening to collect material in the bottom of the wells.

**Step 2:** Using a single or a multichannel pipette, add 2.5 µl of nuclease-free water to the wells in a clean 96-well plate (Barcode Attachment Plate).

**Step 3:** Using a multichannel pipette, transfer 2.5 µl of each well of PCR Pool A & the corresponding well of PCR Pool B (total volume of pooled PCR product should be 5 µl) to the barcode attachment plate and mix by pipetting.

**Step 4:** Using a multichannel pipette, transfer 2.5 µl from the Rapid Barcode Plate to the corresponding well of the Barcode Attachment Plate, taking care not to cross-contaminate different wells. Mix by pipetting.

**Step 5:** Seal the Barcode Attachment Plate and spin down (Total Volume - 10 µl).

**Step 6:** Incubate the plate in a thermal cycler at 30°C for 2 minutes and then at 80°C for 2 minutes.

### Pooling samples and clean-up

Note: Equilibrate Purbeads at room temperature for atleast 30 minutes prior to use.

**Step 1:** Briefly spin down the Barcode Attachment Plate to collect the liquid at the bottom of the wells prior to opening.

**Step 2:** Pool the barcoded samples in a 1.5 ml Eppendorf tube (expected to have about ~10 µl per sample).

Example:

| | X24 samples | X48 samples | X96 samples |
|---|---|---|---|
| Total volume | ~240 µl | ~480 µl | ~960 µl |

**Step 3:** Resuspend the SPRI beads/PurBeads by vortexing vigorously.

**Step 4:** To the entire pooled barcoded sample, add an equal volume of resuspended SPRI beads and mix by flicking the tube.

Example:

| | x24 samples | x48 samples | x96 samples |
|---|---|---|---|
| Volume of SPRI to add | 240 µl | 480 µl | 960 µl |

**Step 6:** Incubate for 5 minutes at room temperature (pulse flick the tube if the beads are settled at the bottom of the tube).

**Step 7:** Prepare at least 1 - 3 ml of fresh 80% ethanol in nuclease-free water.

**Step 8:** Spin down the sample and pellet on a magnetic separation rack. Keep the tube on the magnet, and pipette off the supernatant.

**Step 9:** Keep the tube on the magnet and wash the beads with 1.5 ml of freshly prepared 80% ethanol without disturbing the pellet. Remove the ethanol using a pipette and discard.

**Step 10:** Repeat the previous step.

**Step 11:** Briefly spin down and place the tube back on the magnet. Pipette off any residual ethanol. Allow to dry for 30 seconds, but do not dry the pellet to the point of cracking.

**Step 12:** Remove the tube from the magnetic rack and resuspend the pellet by pipetting in 15 - 30 µl Elution Buffer (EB). Incubate for 10 minutes at room temperature.

**Step 13:** Pellet the beads on a magnet until the eluate is clear and colourless.

**Step 14:** Remove and retain 15 - 30 µl of eluate containing the DNA library into a clean 1.5 ml Eppendorf tube.

**Step 15:** Quantify barcode ligated product concentration by using the Qubit dsDNA HS Assay Kit.

**Step 16:** Take 600–800 ng of library (for flow cell) and make up the volume to 11 µl with EB (Take ~150 ng of library, if flongle is used).

**Step 17:** Add 1 µl of Rapid Adapter F (RAP F) to 11 µl of barcoded DNA (If the number of samples are less (n<10) RAP F should be reduced to 0.5 µl).

**Step 18:** Incubate at room temperature for 5 minutes.

The prepared library is used for loading into the MinION flow cell. Store the library on ice until ready to load.

## Flow cell Preparation

### Introduction to the flow cell check

**Check your Flow cell:**

Oxford Nanopore Technologies highly recommend that you check the number of pores in your flow cell prior to starting a sequencing experiment. This should be done within three months of purchasing for MinION flow cells, or within four weeks of purchasing for Flongle flow cells. ONT will replace any flow cell with fewer than the number of pores in the table below, when the result is reported within two days of performing the flow cell check, and when the storage recommendations have been followed.

| Flow cell | Minimum number of active pores covered by warranty |
|---|---|
| Flongle Flow Cell | 50 |
| MinION/GridION Flow Cell | 800 |

### Shipping condition

Flow cells and kits are shipped together at 2–8°C. The packaging is designed to protect the flow cells from freezing. Sequencing kits are stable at room temperature for up to seven days, so users should not be concerned about their products should they arrive at ambient temperature.

### Storage and stability

We recommend to store flow cells unopened, as follows:

| Flow cell | Storage instructions |
|---|---|
| Flongle Flow Cells | 2–8°C for four weeks |
| MinION/GridION Flow Cells | lab temperature for one month, or at 2–8°C for 12 weeks |

### Flongle flow cell check

Consumables: Flongle device - flow cell and adapter

Equipment: MinION with a host computer connected to the Internet with MinKNOW installed

In this step, you will use the MinKNOW software to check the number of pores in your flow cell. This has to be done as soon as reasonably possible after receiving your Flongle flow cell. Flow cells can then be stored until required. If there is a delay between receiving and using the flow cells, another check should be performed before library preparation begins.

**How it works:**

Flow cells are shipped with a QC DNA molecule present in the buffer. This molecule produces a distinctive nanopore signal. The MinKNOW software uses this signal to validate the integrity of the nanopore array before use and provides the user with an estimate of the number of simultaneously available channels for the experiment.

**Step 1:** Place the Flongle adapter into the MinION.

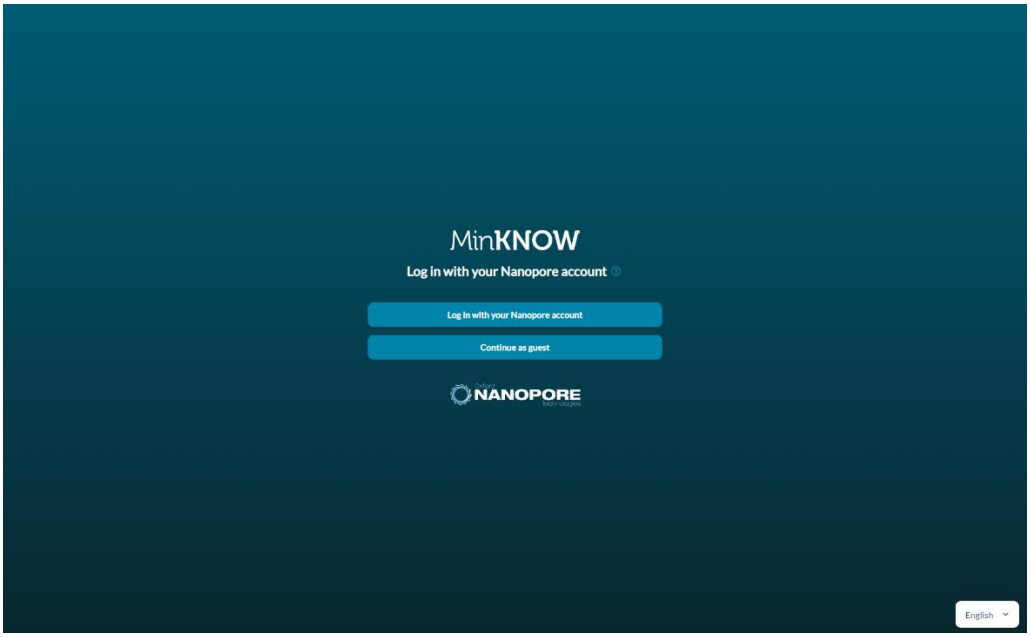
- The adapter should sit evenly and flat on the MinION Mk1B or GridION platform. This ensures the flow cell assembly is flat during the next stage.
- The adapter needs to be plugged into your device, and the device should be plugged in and powered on before inserting the Flongle flow cell.

**Step 2:** Place the flow cell into the Flongle adapter, and press the flow cell down until you hear a click.

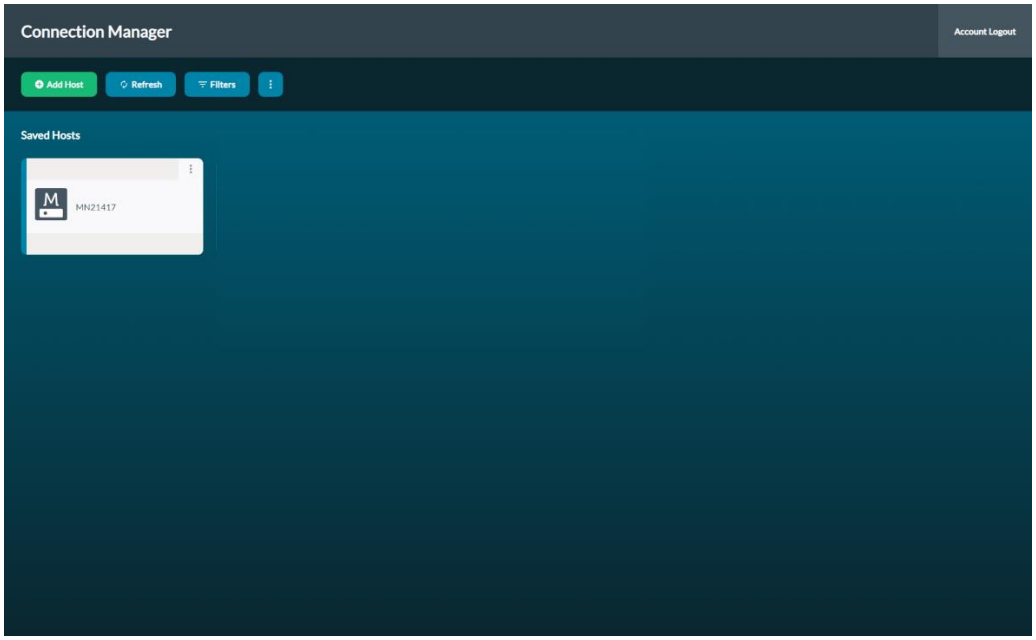
- The flow cell should sit evenly and flat inside the adapter, to avoid any bubbles forming inside the fluidic compartments.

**Step 3:** Double-click the **MinKNOW icon** located on the desktop to open the MinKNOW GUI.

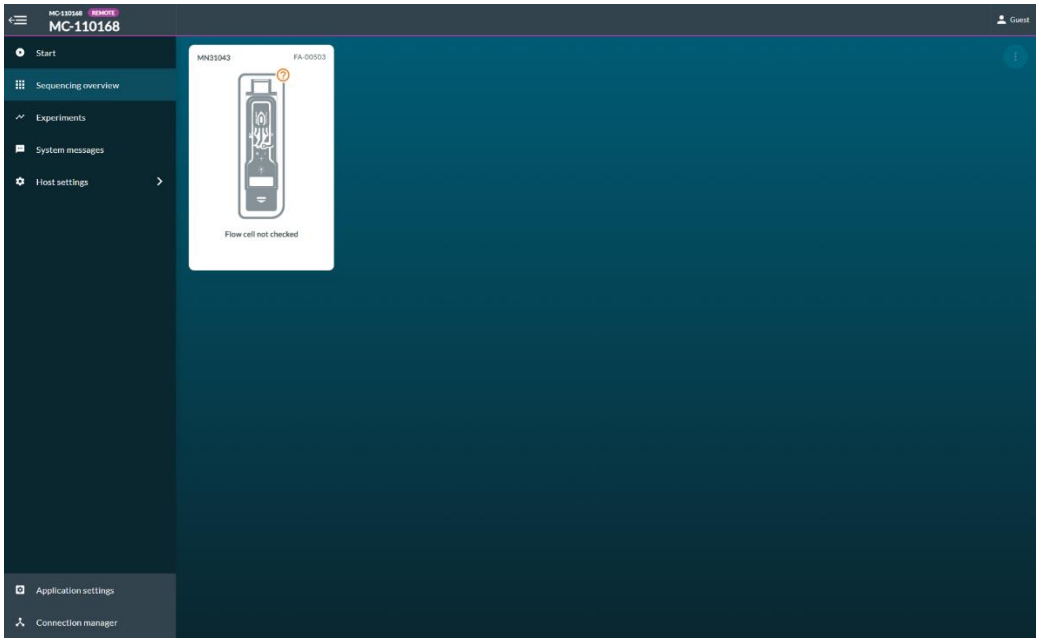

- To log in, you must be connected to the internet.



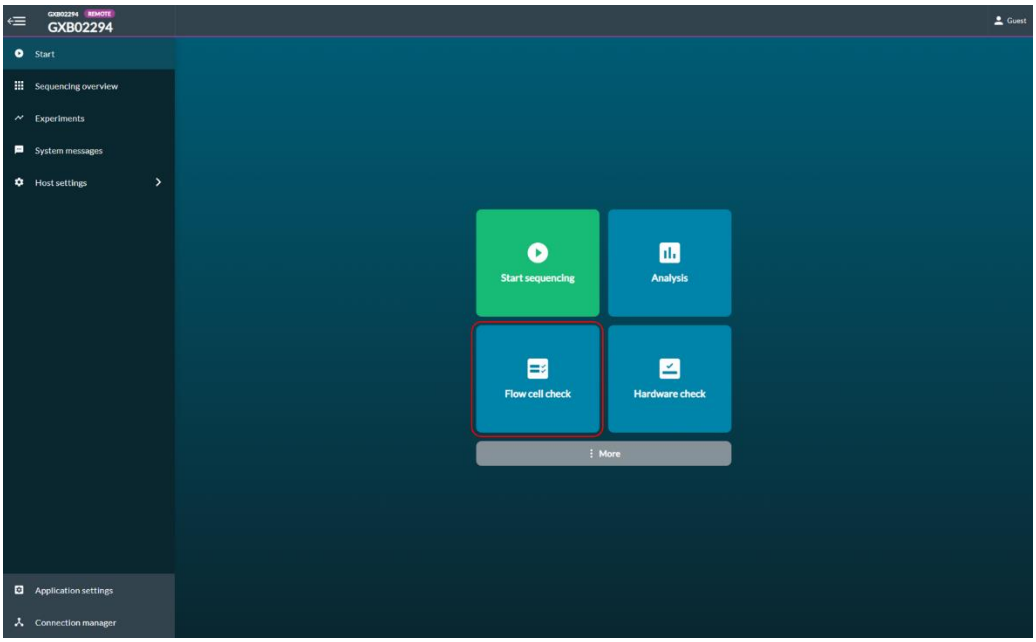

**Step 4:** Select the sequencing device connected to the computer.




- The Sequencing Overview should show the flow cell has not had any checks carried out.

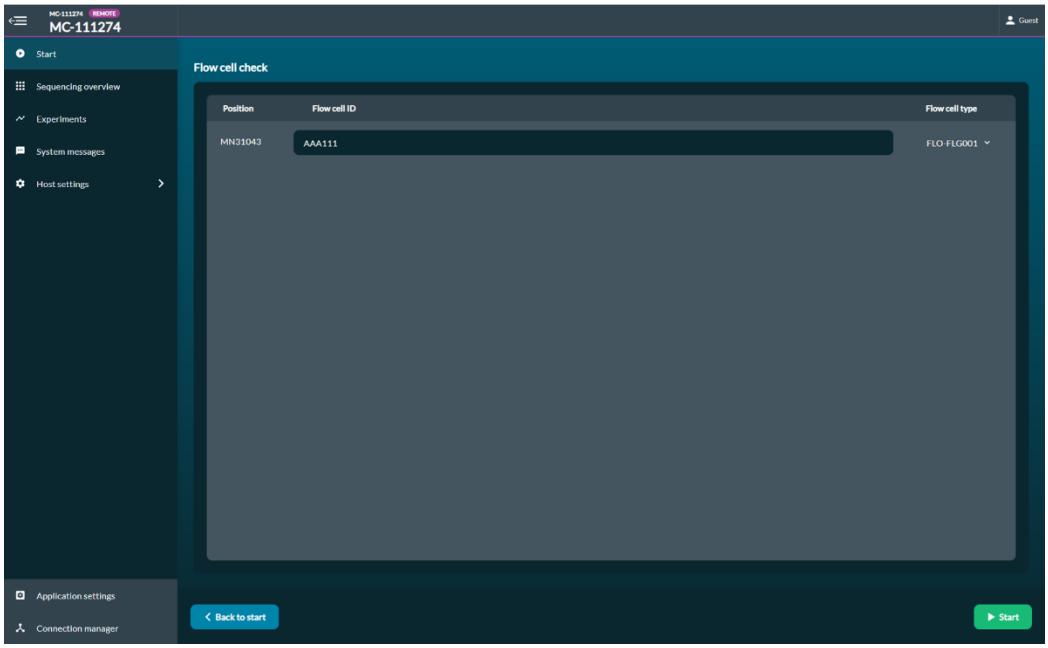



**Step 5:** Navigate to the start homepage and select **'Flow Cell Check**'.




**Step 6:** Choose the flow cell type from the drop down menu and fill in the flow cell ID.

Select **FLO-FLG001**.




- For Flongle, flow cell ID is not automatically assigned in MinKNOW. It is imperative the flow cell ID is entered in the correct format: ABC123 i.e. ([A-Z] x 3 [0-9] x 3).

    1. The ID is case sensitive with no spaces.
    2. The MinKNOW software will not allow you to proceed until the flow cell ID has been entered correctly.
    3. The Flongle flow cell ID is shown in the blue blox below:


**Step 7:** Select '**Start**' to begin the flow cell check.

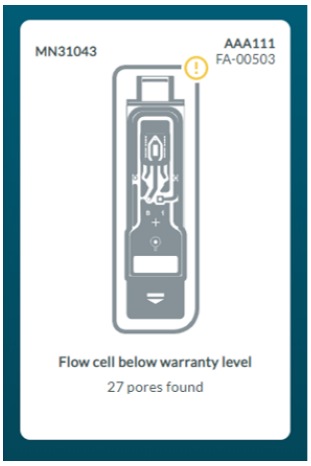
- You will be automatically navigated to the Sequencing Overview page.
- A loading bar will be displayed under the flow cell during the checks.
- The flow cell check should take a few minutes.

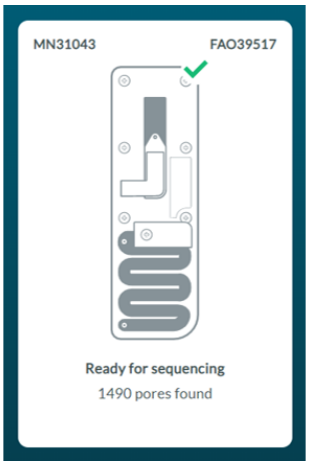
**Flow cell health indicators:**

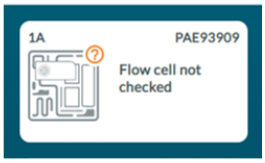
The quality of the flow cell will be shown as one of the three outcomes:

- Orange exclamation mark (Flongle flow cell) The number of sequencing pores is below warranty.
- Green tick (MinION flow cell) The number of sequencing pores is above warranty.

- Question mark (PromethION flow cell) A Flow Cell Check has not been run on the flow cell during this MinKNOW session.



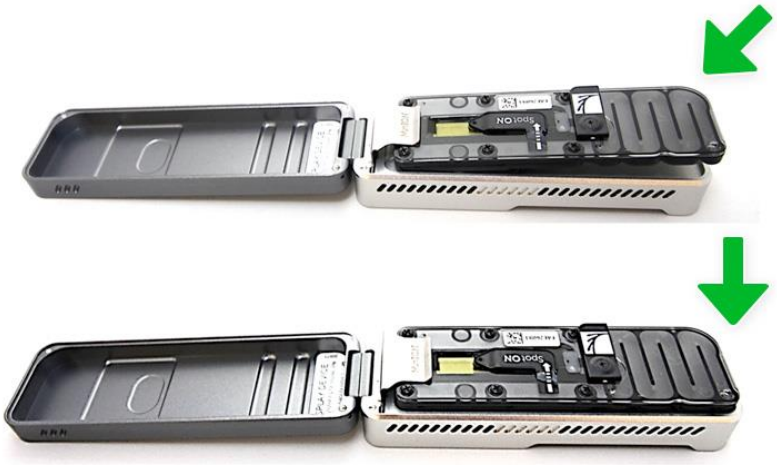

### MinION Mk1B flow cell check:

Consumables: SpotON Flow Cell

Equipment: MinION Mk1B with host computer connected to the Internet with MinKNOW installed

**Step1:** Open the MinION Mk1B/GridION lid and insert the MinION Flow Cell

- Insert the flow cell in the MinION Mk1B by sliding the flow cell under the clip. Firmly press down on the flow cell to ensure good thermal and electrical contact.

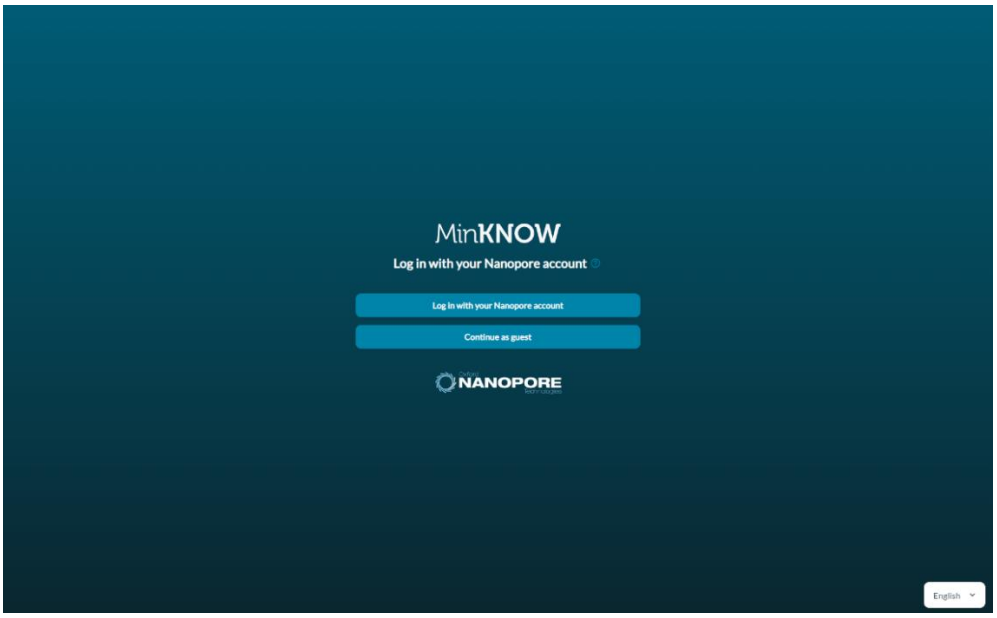


**Step 2:** Connect the assembled MinION Mk1B and flow cell to the host computer.

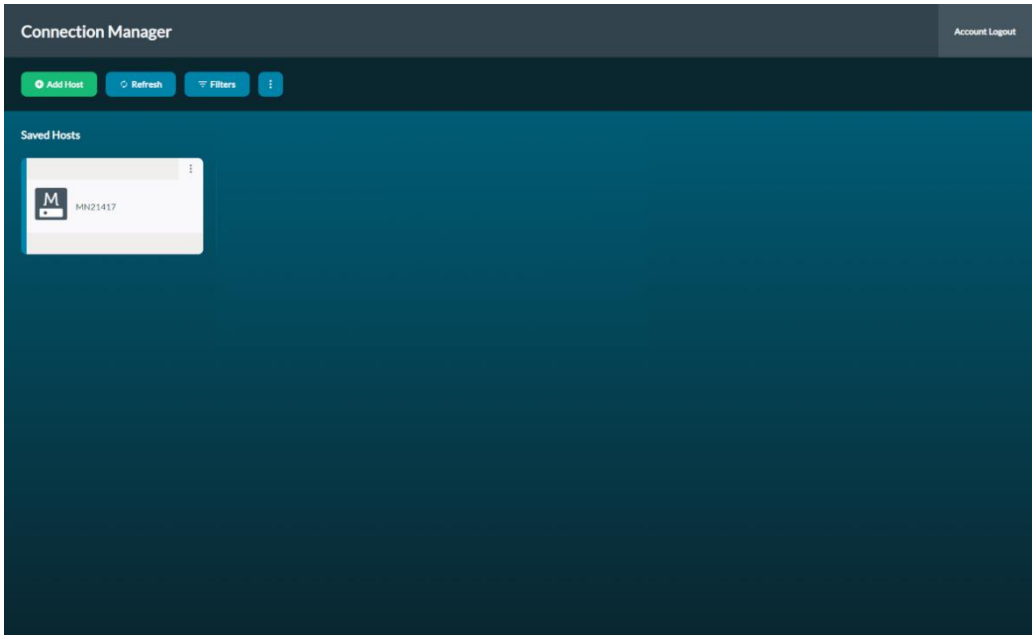
- Once successfully plugged in, you will see a light and hear the fan

**Step 3:** Double-click the **MinKNOW icon** located on the desktop to open the MinKNOW GUI.

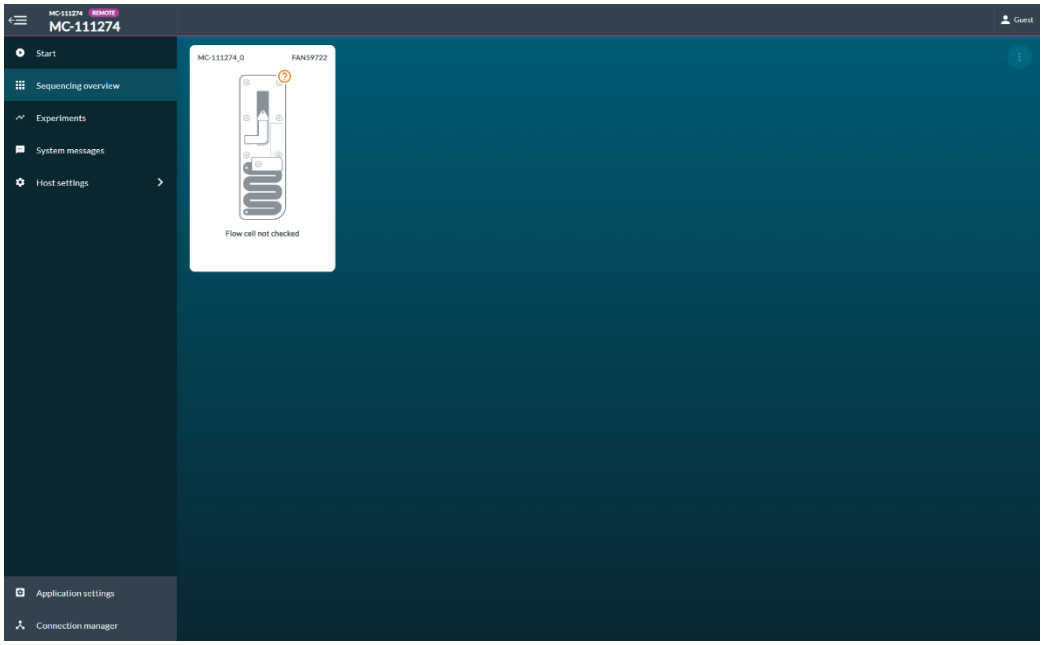

➢ To log in, you must be connected to the internet.



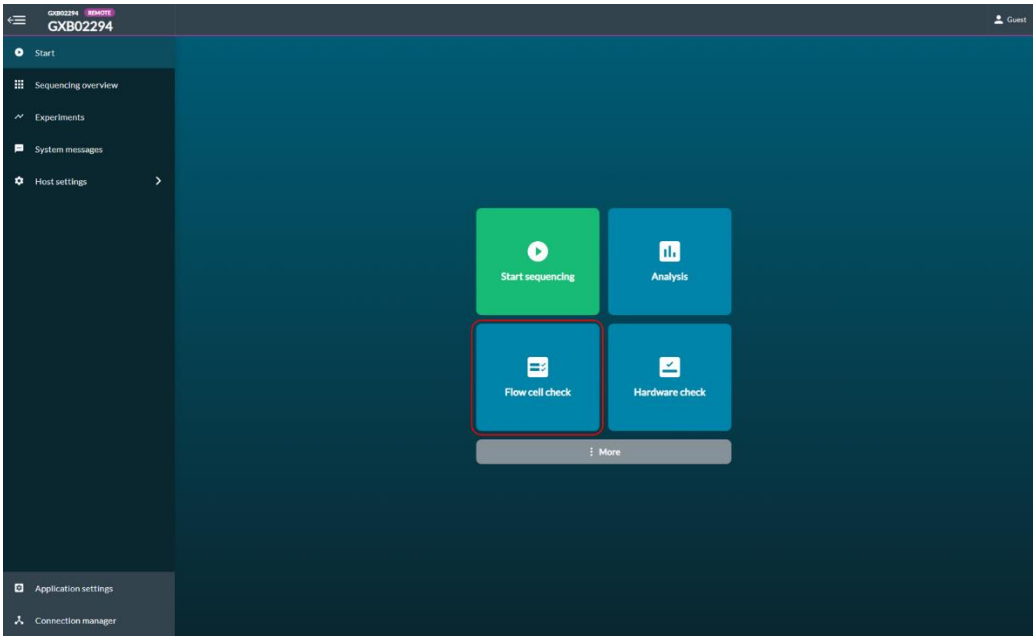

**Step 4:** Select the sequencing device connected to the computer.




- The Sequencing Overview should show the flow cell has not had any checks carried out.

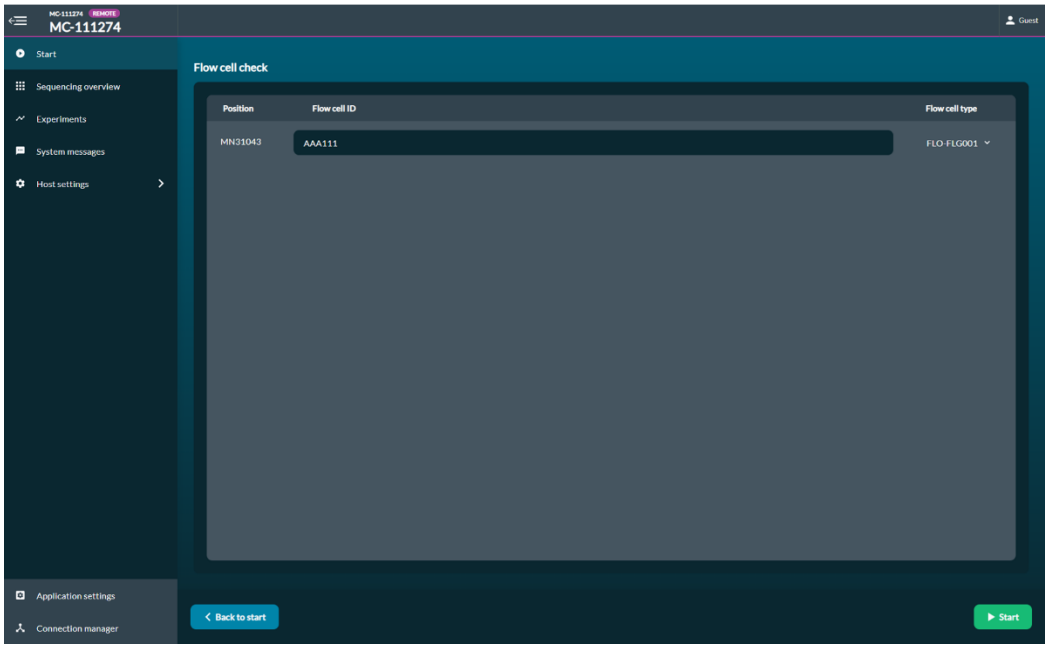



**Step 5:** Navigate to the start homepage and select '**Flow Cell Check**'.




**Step 6:** Choose the flow cell type from the drop-down menu.

Note: If you are using flow cells from your Starter Pack, please select **FLO-MIN106**.



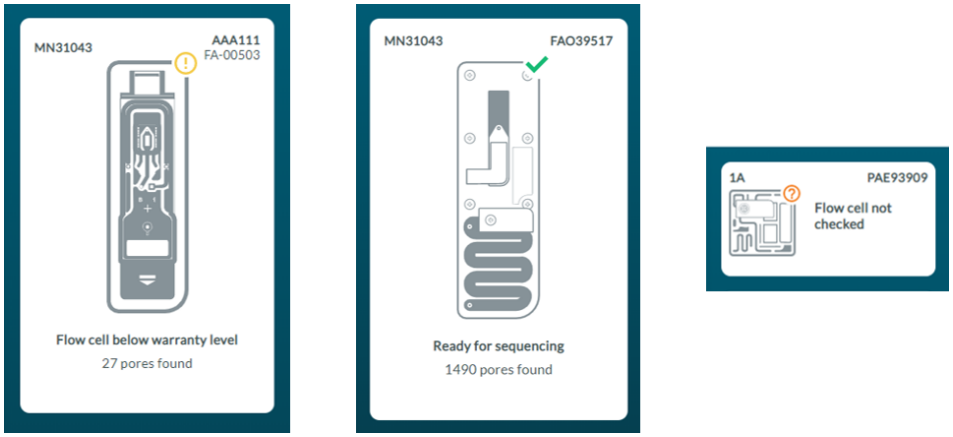

**Step 7:** Select '**Start**' to begin the flow cell check.

- You will be automatically navigated to the Sequencing Overview page.
- A loading bar will be displayed under the flow cell during the checks.
- The flow cell check should take a few minutes.

**Flow cell health indicators:**

The quality of the flow cell will be shown as one of the three outcomes:

- Yellow exclamation mark (Flongle flow cell) The number of sequencing pores is below warranty.
- Green tick (MinION flow cell) The number of sequencing pores is above warranty.
- Question mark (PromethION flow cell) A Flow Cell Check has not been run on the flow cell during this MinKNOW session.



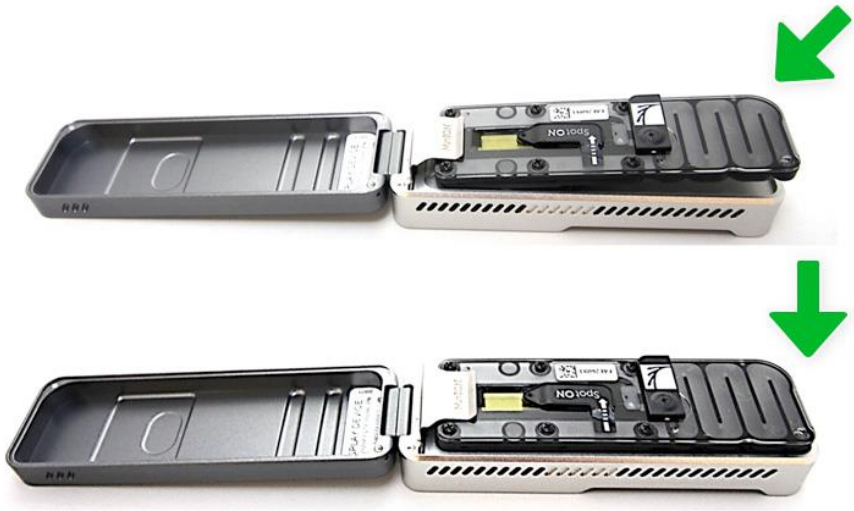

## Priming and loading the SpotON Flow Cell

**Step 1:** Thaw the Sequencing Buffer II (SBII), Loading Beads II (LBII) or Loading Solution (LS, if using), Flush Tether (FLT) and Flush Buffer (FB) at room temperature.

**Step 2:** Mix the Sequencing Buffer II (SBII), Flush Buffer (FB), Flush Tether (FLT) and Loading Solution (LS, if using) tubes by vortexing. Spin down the SBII and FLT at room temperature.

**Step 3**: Open the MinION Mk1B lid and slide the flow cell under the clip.

- Press down firmly on the flow cell to ensure correct thermal and electrical contact.



**Step 4:** Slide the **priming port** cover clockwise to open the priming port.


**How to prime and load the SpotON Flow Cell?**

Refer to the video first & then proceed with the instructions below.

**https://player.vimeo.com/video/186403871**

**Priming and loading:**

- The library is loaded dropwise without putting the pipette tip firmly into the port.
- Take care to avoid introducing any air during pipetting.

Take care when drawing back buffer from the flow cell. Do not remove more than 20-30 µl, and make sure that the array of pores are covered by buffer at all times. Introducing air bubbles into the array can irreversibly damage pores.

**Step 5 (Degassing):** After opening the priming port, check for a small air bubble under the cover. Draw back a small volume (20- 30 µl) to remove any bubbles (maximum 30 µl):

- Set a P1000 pipette to 200 µl
- Insert the tip into the priming port
- Turn the wheel until the dial shows 220-230 µl, or until you can see a small volume of buffer entering the pipette tip.

Note: Visually check that there is continuous buffer from the priming port across the sensor array.

**Step 6:** To prepare the **flow cell priming mix**, add 30 µl of thawed and mixed Flush Tether (FLT) directly to the tube of 1.17 ml of thawed and mixed Flush Buffer (FB), and mix by vortexing at room temperature (18 - 23 °C).

**Step 7:** Load 800 µl of the priming mix into the flow cell via the priming port, avoiding the introduction of air bubbles. Wait for 5 minutes. During this time, prepare the library for loading by following the steps below.

**Step 8:** Thoroughly mix the contents of the Loading Beads II (LBII) by pipetting.

The Loading Beads II (LBII) tube contains a suspension of beads. These beads settle very quickly. It is vital that they are mixed immediately before use.

**Step 9:** In a new tube, prepare the library for loading as follows:

| Reagent | Volume |
|---|---|
| Sequencing Buffer II (SBII) | 37.5 µl |
| Loading Beads II (LBII) mixed immediately before use, or Loading Solution (LS), if using | 25.5 µl |
| DNA library | 12 µl |
| **Total** | **75 µl** |

**Step 10:** Complete the flow cell priming:

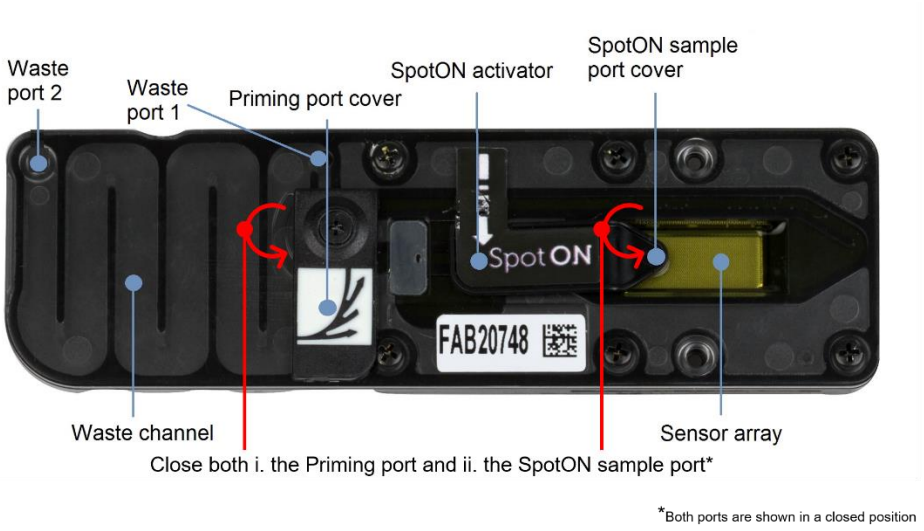
- Gently lift the SpotON sample port cover to make the SpotON sample port accessible.
- Load 200 µl of the priming mix into the flow cell via the priming port (not the SpotON sample port), avoiding the introduction of air bubbles.

Note: Load the library as soon as possible after this step.

**Step 11:** Mix the prepared library gently by pipetting up and down just prior to loading.

**Step 12:** Add 75 µl of sample to the flow cell via the SpotON sample port in a dropwise fashion. Ensure each drop flows into the port before adding the next, avoiding the introduction of air bubbles.

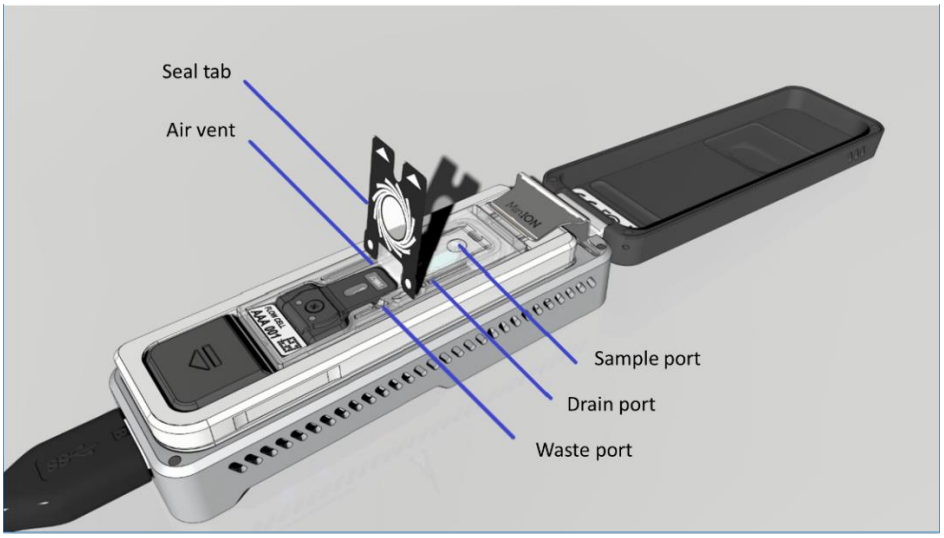
**Step 13:** Gently replace the SpotON sample port cover, making sure the bung enters the SpotON port, close the priming port and replace the MinION Mk1B lid.




**Loading the Flongle flow cell** (Skip to next section, if flongle is not used)

Do NOT touch the reverse side of the Flongle flow cell array or the contact pads on the Flongle adapter. ALWAYS wear gloves when handling Flongle flow cells and adapters to avoid damage to the flow cell or adapter.

**The diagram below shows the components of the Flongle flow cell:**




The seal tab, air vent, waste channel, drain port and sample port are visible here. The sample port, drain port and air vent only become accessible once the seal tab is peeled back.

**Step 1:** Thaw the Sequencing Buffer II (SBII), Loading Beads II (LBII) and Flush Buffer (FB) from the Flongle Sequencing Expansion and Flush Tether (FLT) from the Ligation Sequencing Kit (SQK-LSK110) at room temperature.

**Step 2:** Mix the Sequencing Buffer II (SBII), Flush Buffer (FB) and Flush Tether (FLT) tubes by vortexing and spin down at room temperature.

**Step 3:** In a fresh 1.5 ml Eppendorf DNA LoBind tube, mix 117 µl of Flush Buffer (FB) with 3 µl of Flush Tether (FLT) and mix by pipetting.

**Step 4:** Place the Flongle adapter into the MinION device.

- The adapter should sit evenly and flat on the MinION Mk1B or GridION platform. This ensures the flow cell assembly is flat during the next stage.

**Step 5:** Place the flow cell into the Flongle adapter, and press the flow cell down until you hear a click.

- The flow cell should sit evenly and flat inside the adapter, to avoid any bubbles forming inside the fluidic compartments.

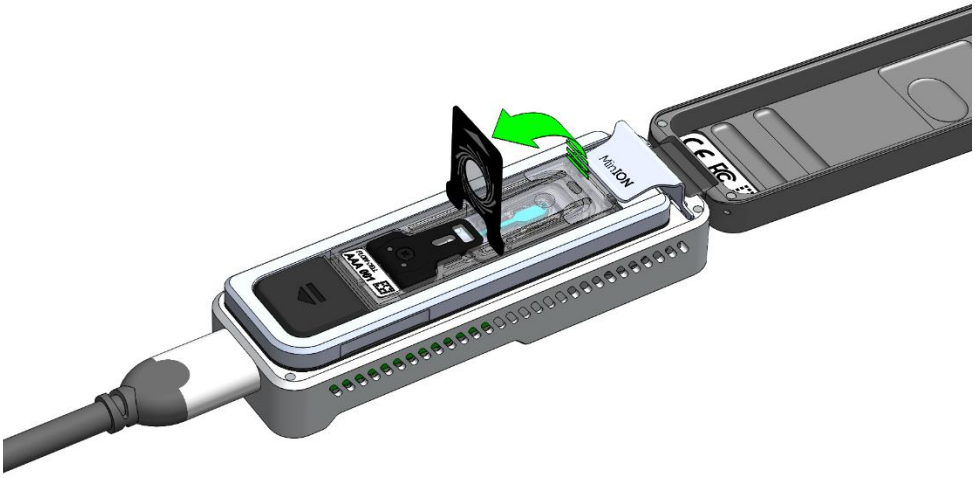
**How to prime and load a Flongle flow cell:**

Refer to the video & proceed further with the instruction below.

https://vimeo.com/297106442

**Step 6:** Peel back the seal tab from the Flongle flow cell, up to a point where the sample port is exposed, as follows:

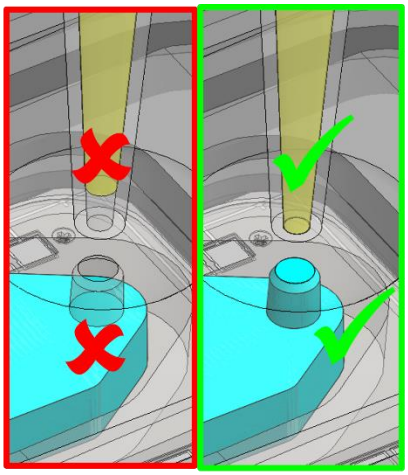
- **Lift up the seal tab:**

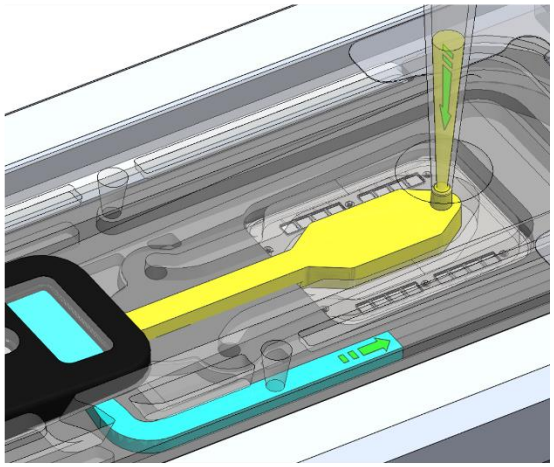



- **Pull the seal tab to open access to the sample port:**

- **Hold the seal tab open by using adhesive on the tab to stick to the MinION Mk 1B lid:**

**Step 7:** To prime your flow cell with the mix of Flush Buffer (FB) and Flush Tether (FLT) that was prepared earlier, ensure that there is no air gap in the sample port or the pipette tip. Place the P200 pipette tip inside the sample port and slowly dispense the priming fluid into the Flongle flow cell. To avoid flushing the flow cell too vigorously, load the priming mix by twisting the pipette plunger down.

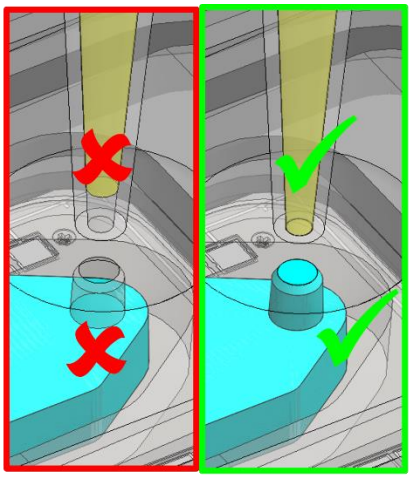


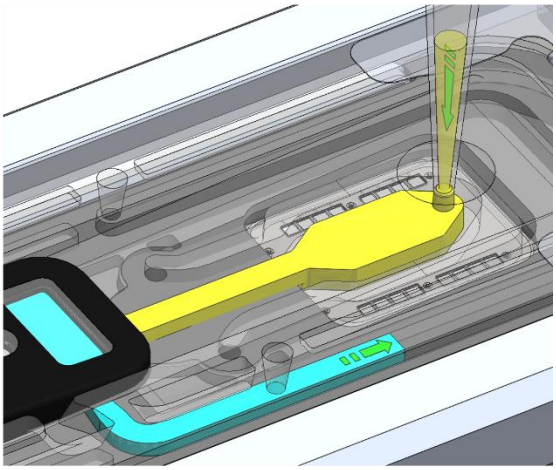


**Step 8:** Vortex the vial of Loading Beads II (LBII). Note that the beads settle quickly, so immediately prepare the Sequencing Mix in a fresh 1.5 ml Eppendorf DNA LoBind tube for loading the Flongle, as follows:

| **Reagents** | **Volume** |
|---|---|
| Sequencing Buffer II (SBII) | 15 µl |
| Loading Beads II (LBII) mixed immediately before use, or Loading Solution (LS), if using. LS can be used instead of LBII when preparing libraries with the Ligation Sequencing Kit (SQK-LSK110) | 10 µl |
| DNA library | 5 µl |
| **Total** | **30 µl** |

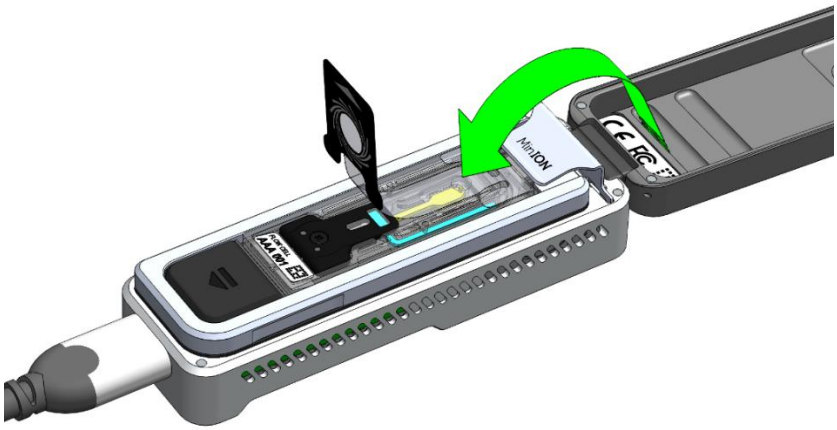
**Step 9:** To add the Sequencing Mix to the flow cell, ensure that there is no air gap in the sample port or the pipette tip. Place the P200 tip inside the sample port and slowly dispense the Sequencing Mix into the flow cell by twisting the pipette plunger down.



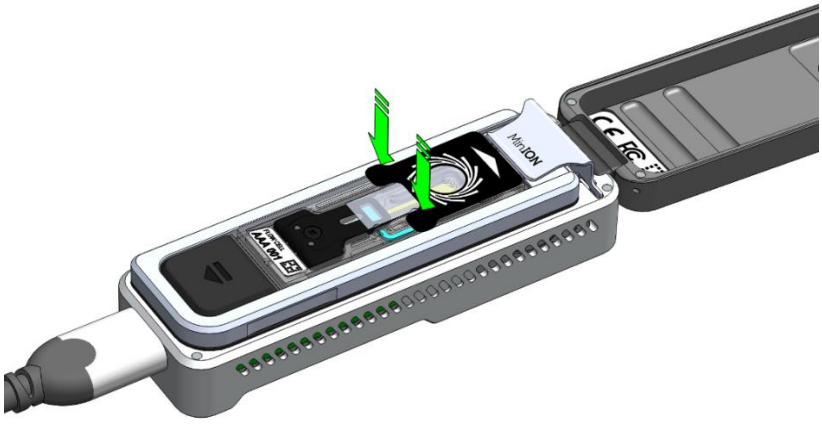


**Step 10:** Seal the Flongle flow cell using the adhesive on the seal tab, as follows:

- Stick the transparent adhesive tape to the sample port.




- Replace the top (Wheel icon section) of the seal tab to its original position.



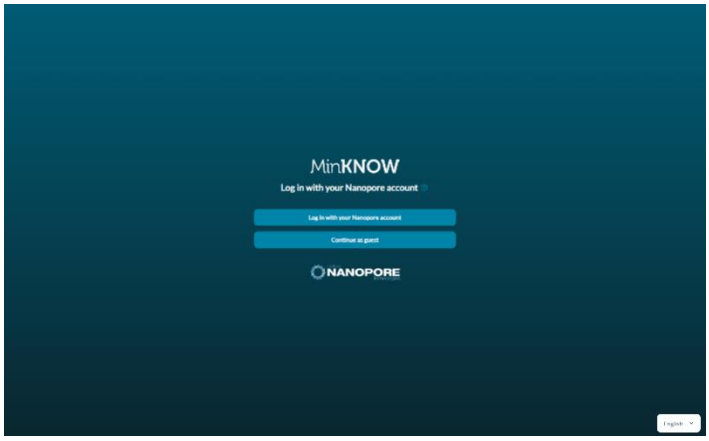

**Step 11:** Replace the sequencing platform lid.

## Sequencing Run Setup

### Starting a sequencing run on MinION Mk1B

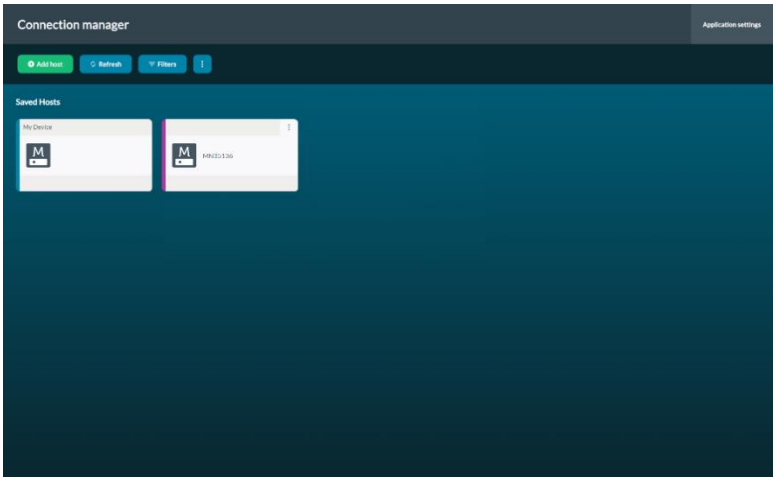
**Step 1:** Double-click the **MinKNOW icon** located on the desktop to open the MinKNOW GUI.


**Step 2:** To log in, you must be connected to the internet.

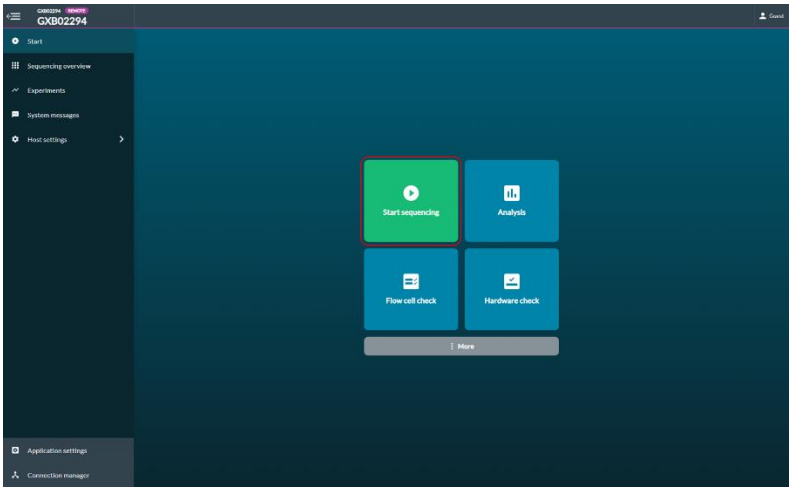



Step 3: If your MinION Mk1B was disconnected from the computer, plug it back in.

Step 4: Select the sequencing device connected to the computer.



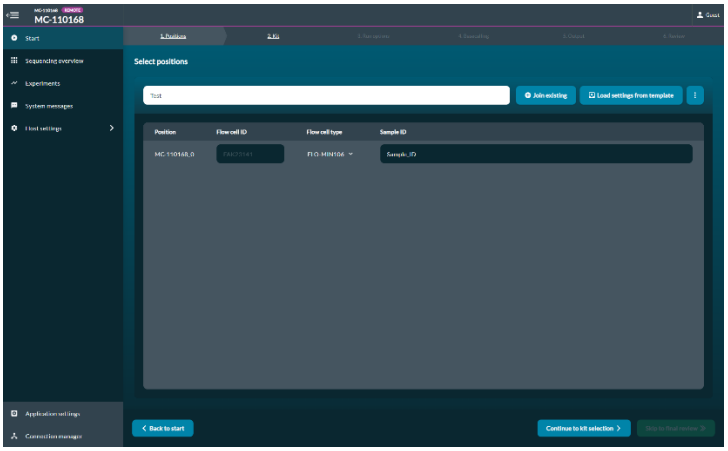

**Step 5:** Select the **'Start Sequencing'** option on the Start homepage to choose the running parameters for your experiment.




Step 6: Type in the experiment name, sample ID and choose flow cell type from the drop-down menu.

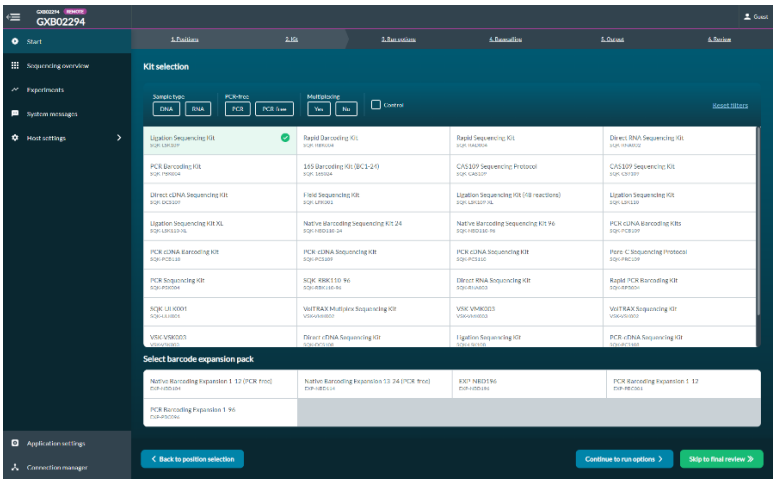
Note: If sample ID is not filled in, there will be no sample ID in the folder structure.

Step 7: Select **Continue** to Kit Selection to move to the next page.



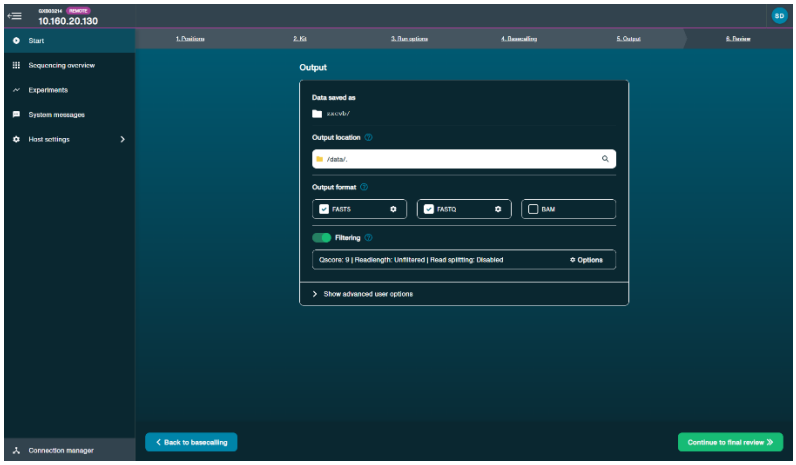

**Step 8:** Select the kit used from the **Kit selection menu**.

The filter options may be used to find the kit used. For example, for running SARS-CoV-2 protocol, select SQK-RBK110-96



**Step 9:** Select **continue to run** options, click further to **continue to basecalling** & click again to **continue to output**.

**Step 10:** Select the **output data location**, an alternative location can be selected to which processed reads will be written out by using Output Location.

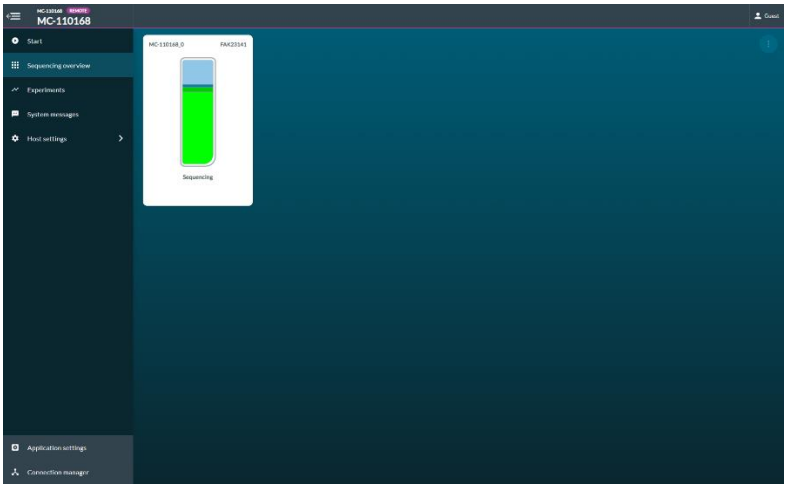



**Step 11:** Select **continue to final review** to proceed

**Step 12:** Click **'Start'** to run the experiment.

Step 13: Users will be automatically navigated to the Sequencing Overview when sequencing starts.

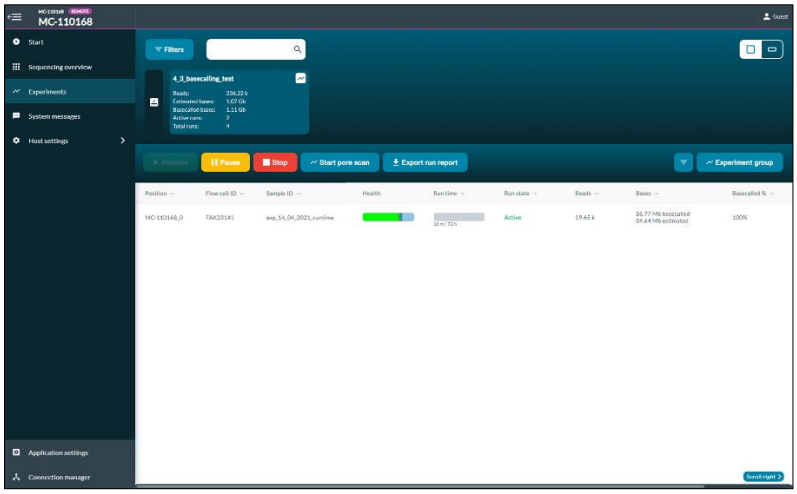
Flow cell health will be displayed after the first pore scan.




**Step 14:** Select the flow cell to open the quick view to check the number of active pores reported in the pore scan are similar (within 10-15%) to those reported at the end of the Flow Cell Check.

- If there is a significant reduction in the numbers, restart MinKNOW.
- If the numbers are still significantly different, close down the host computer and reboot.
- When the numbers are similar to those reported at the end of the Flow Cell Check, restart the experiment. There is no need to load any additional library after restart.

**Step 15:** To stop or pause the experiment, select 'Stop' or ‘Pause’ on the experiments page.




**Step 16:** Once sequencing is complete, generate a run report.

A run report containing information about the sequencing run and performance graphs can be generated by clicking Export run report.

## MUX scan

As the sequencing protocol starts, a MUX scan begins before the sequencing. This multiplex scan allows MinKNOW to prioritise the order in which the nanopores are used, maximising the data output in the initial stages of the run. There are four groups of active pores, and group 1 are used in the first eight hours. So for the Lambda Control Experiment, only group 1 pores are used but for longer sequencing runs the other groups are used as required.

If Active Channel Selection is enabled during the run, the software instantly switches to a new channel in the group if a channel is in the “Saturated” or “Multiple” state, or after ~5 minutes if a channel is “Recovering”.

## MinKNOW fallback folder

If writing out reads to your specified location fails (e.g. if the disk becomes full, or the connection is disrupted), the experiment will continue, and the reads will be written into a fallback folder:

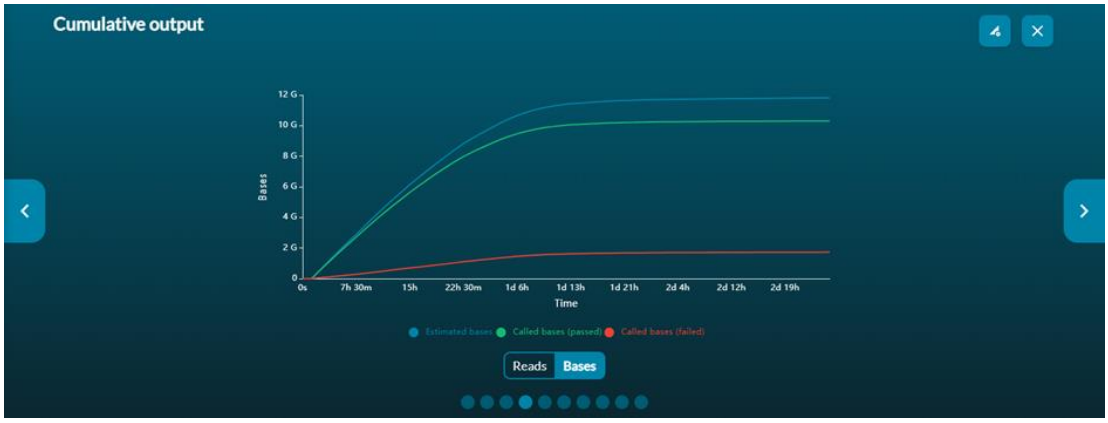
**Linux:**

.fast5 files: /var/lib/MinKNOW/data/fallback

FASTQ files: /var/lib/MinKNOW/data/fastq_fallback

## Ending the Sequencing Experiment

Step 1: When the cumulative output graph indicates that no reads are collected as seen by plateauing of the graph, a sequencing can be stopped before the set run time and can be proceeded for basecalling or further analysis.




Step 2: After your sequencing experiment is complete, if you would like to reuse the flow cell, please follow the Wash Kit instructions and store the washed flow cell at 2-8°C, or ONT recommends you to wash the flow cell as soon as possible after you stop the run. However, if this is not possible, leave the flow cell on the device and wash it the next day.

## Checks and monitoring

During the sequencing experiment, you can check various flow cell health and performance parameters that are shown in the MinKNOW GUI. Each parameter is described in more detail in this section.

Flow cell health:

During a sequencing experiment, the MinKNOW Sequencing Overview page shows a flow cell icon with coloured bars. The bars represent the combined health of all pores in a flow cell, and indicate how well the flow cell is performing. The colours are:

- Light green: sequencing
- Dark green: open pore
- Dark blue: pore recovering
- Light blue: pore inactive

This information is identical to the last bar of the pore activity plot (described later).

**Flongle/MinION flow cell health diagram:**

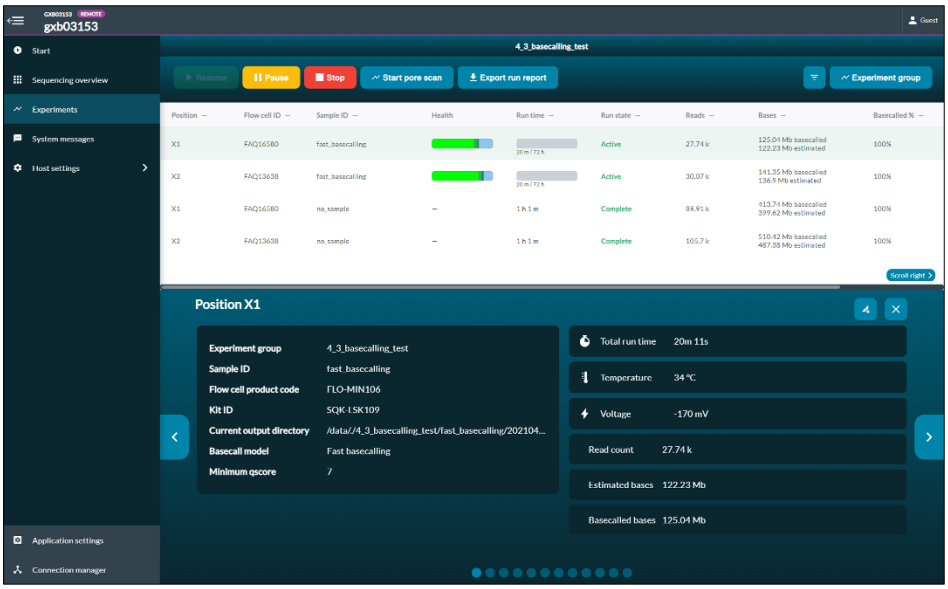
Note: The below image illustrates flow cell health on a GridION. A MinION device will only have one flow cell.


## Experiment summary information

The Experiments page displays summary information for all sequencing flow cells and device checks carried out on the device.

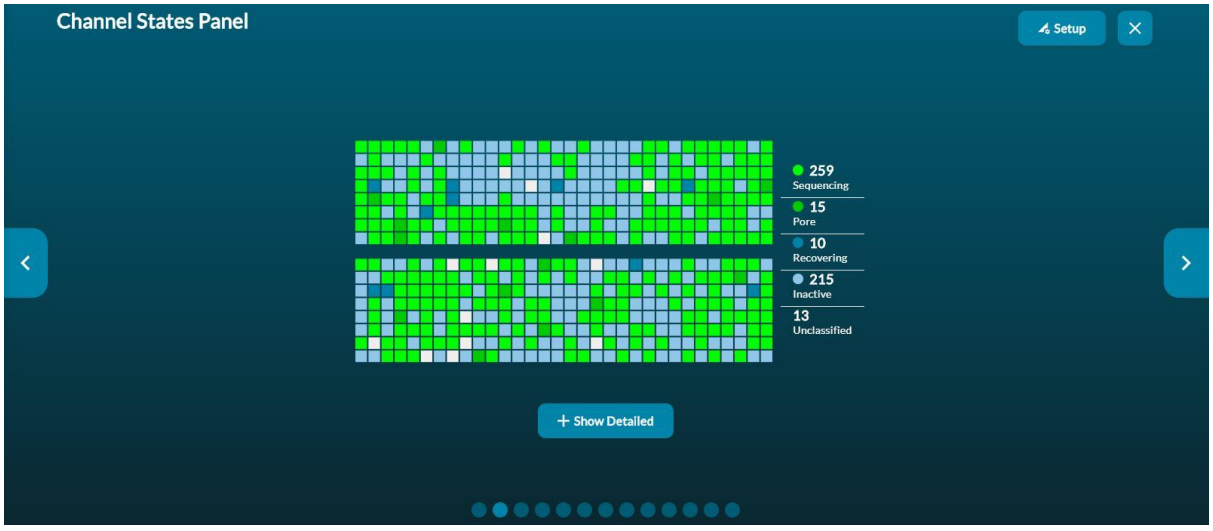
From this page, the user is able to control specific runs and identify real-time information including flow cell health and reads, giving users real-time feedback for sequencing flow cells.

- Run statistics: The total number of reads and bases produced across the experiment
- Basecall statistics: There are two values for basecalled reads:
    1. Basecalled reads as a percentage of the total reads produced across the experiment. This gives an indication as to the size of the queue for reads to be basecalled
    2. Total number of reads basecalled across the experiment.
- Run time: The duration of the experiment
- Temperature: The heatsink temperature of the selected position, which should reach 34°C (or 36°C on Flongle) for sequencing
- Voltage: The applied potential of the position at that point in time.




## Pore occupancy

- A good library will be indicated by a higher proportion of light green channels in Sequencing than are in Pore. The combination of **Sequencing** and **Pore** indicates the number of active pores at any point in time. A low proportion of Sequencing channels will reduce the throughput of the run.
- **Recovering** indicates channels that may become available for sequencing again. A high proportion of this may indicate additional clean up steps are required during your library preparation.
- **Inactive** indicates channels that are no longer available for sequencing. A high proportion of these as soon as the run begins may indicate an osmotic imbalance.
- **Unclassified** are channels that have not yet been assigned one of the above classifications.




Clicking on the Show Detailed button reveals a more detailed array of channel states:

**Strand:** the channel has strand

**Adapter:** the pore is sequencing the unligated sequencing adapter only. Reads will initially be classified as adapter until the DNA/RNA strand starts translocating through the pore and MinKNOW™ is able to reclassify the read

**Single pore:** the channel appears to show a single pore. Available for sequencing

**Unavailable:** the channel appears to show a pore that is currently unavailable for sequencing

**Active feedback:** the channel is reversing the current flow to eject the analyte

**No pore from scan:** the Mux scan has not detected a pore in the well

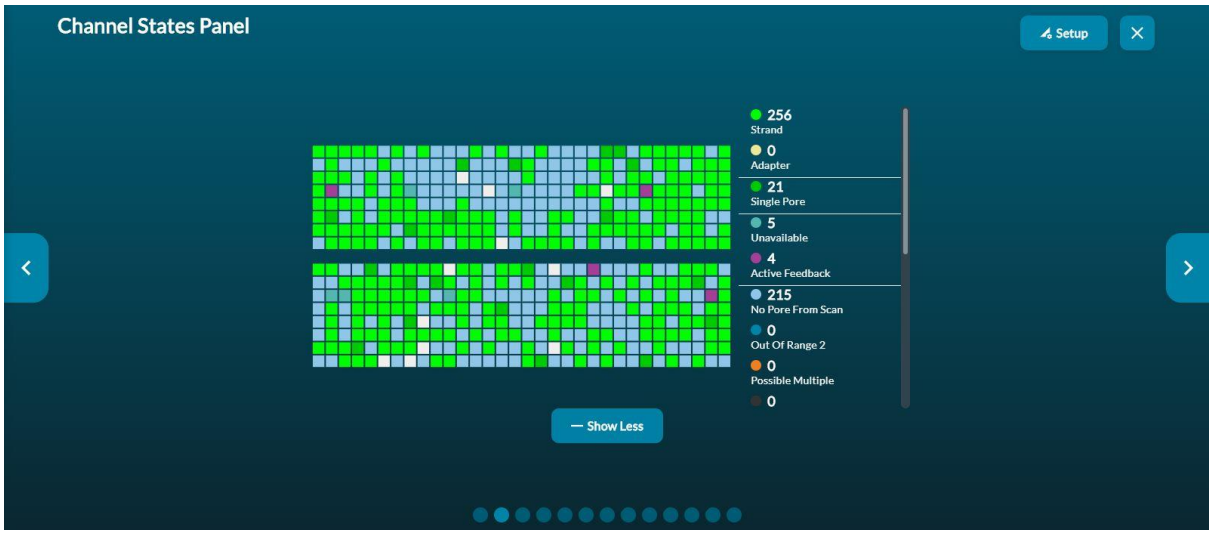
**Out of range 2:** current level is between 10 and 9999 pA. Currently unavailable for sequencing

**Possible multiple:** the channel appears to show more than one pore. Unavailable for sequencing

**Saturated:** the channel has switched off due to current levels exceeding hardware limitations

**Out of range 1:** current level is between -5 and -9999 pA. Currently unavailable for sequencing

**Zero:** Current level is between -5 and 10 pA. Currently unavailable for sequencing

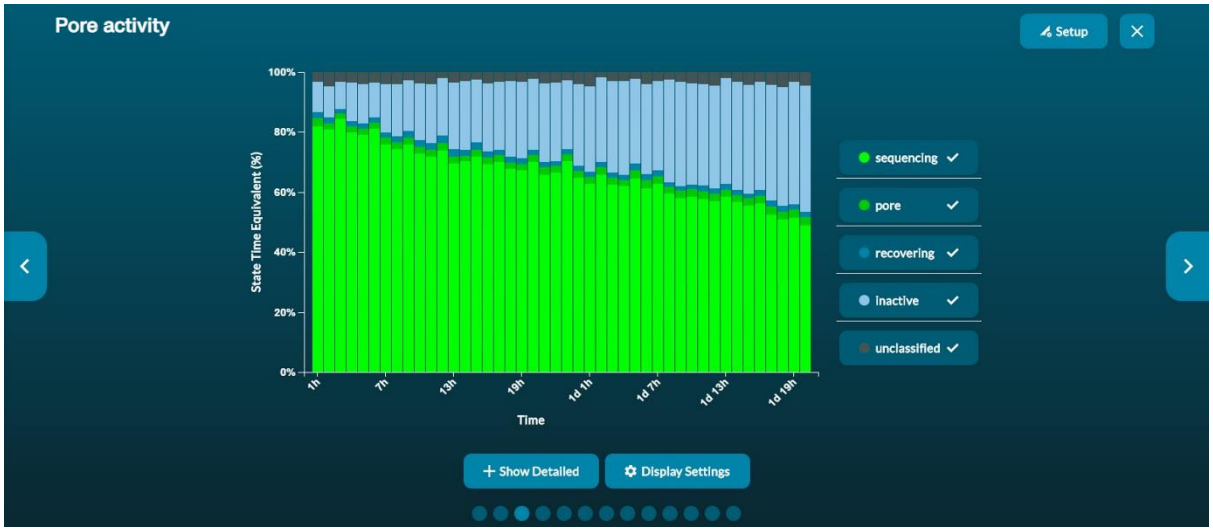



## Pore activity plots

The pore activity plot summarizes the channel states over time.

Each bar shows the sum of all channel activity in a particular amount of time. This time bucket defaults to 1 minute, and scales to 5 minutes automatically after reaching 48 buckets. However, bucket size can be adjusted in the "Bucket size" box in Display Settings.

The graph populates over time and can be used as a way to assess the quality of your sequencing experiment, and make an early decision whether to continue with the experiment or to stop the run.




## Read length histogram

The cumulative histogram shows reads compared to bases. Use the options below to choose the axis legends:

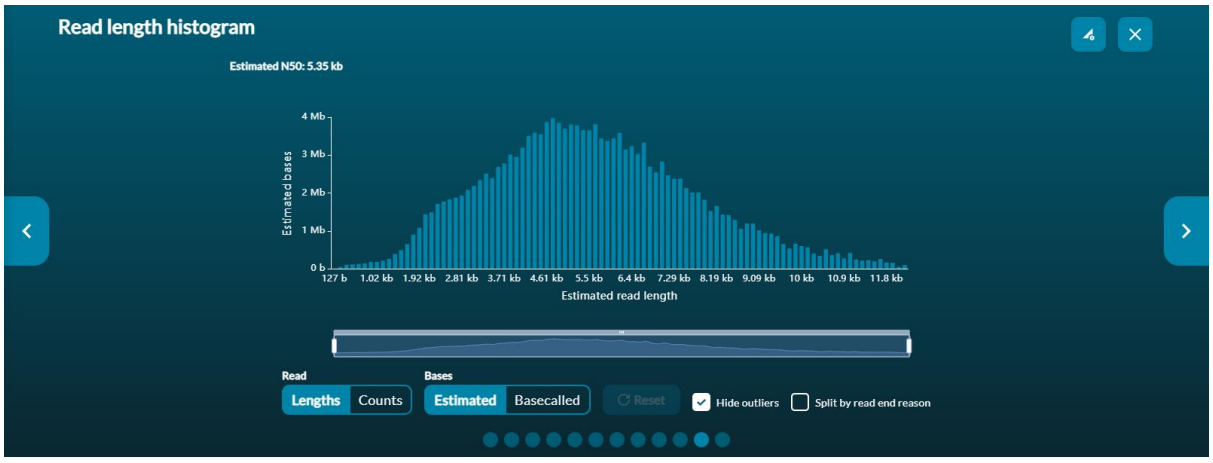
Y-axis: Estimated bases or basecalled bases

X-axis: Read length or read counts

Read count - this shows the number of reads vs read length. This enables the user to understand how the read lengths vary in number and size.

Read length - this shows the total number of bases vs the read length.

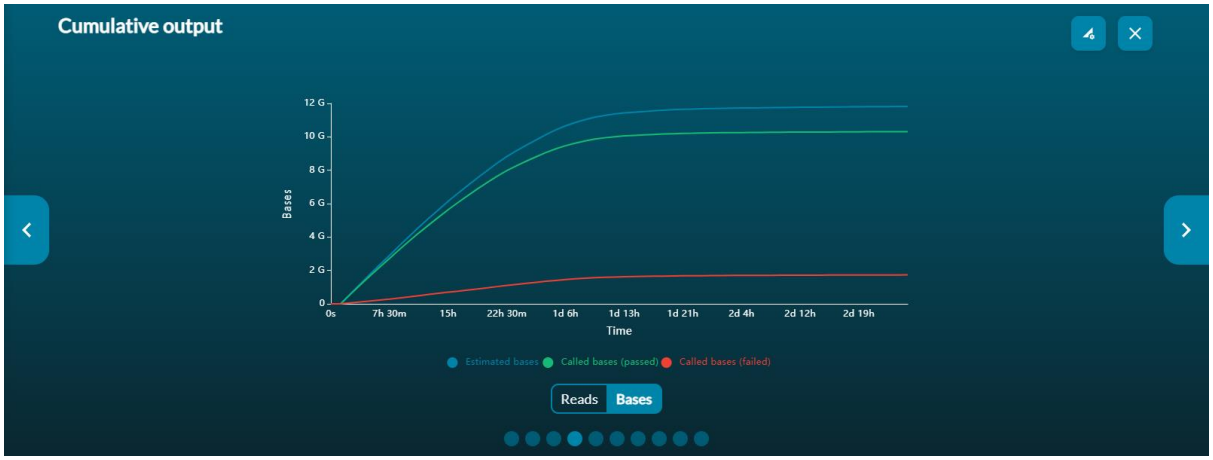
The N50 value is presented (only for the whole set of passed reads) in the top left corner of the histogram.



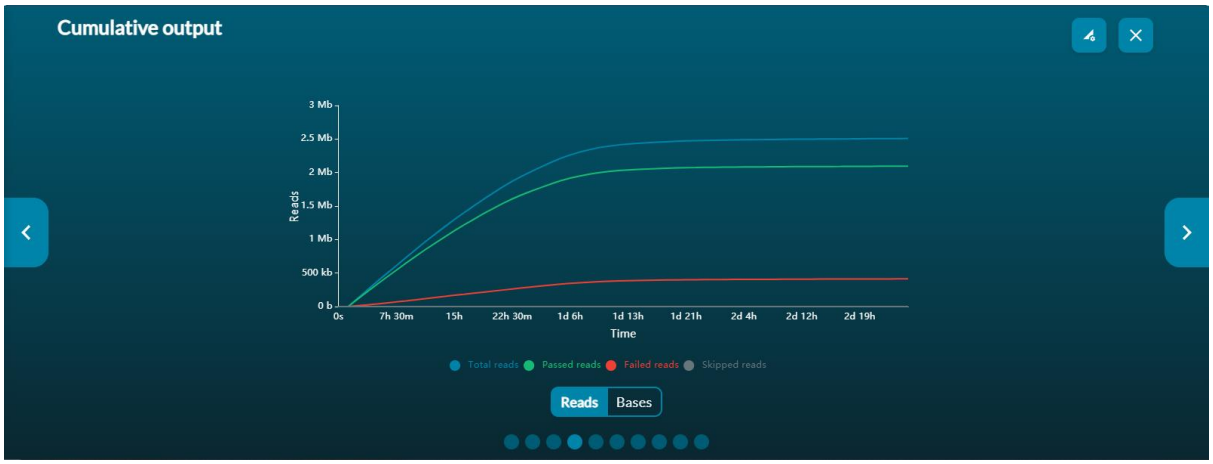

## Cumulative output

The cumulative output graph shows:

- the number of bases that have been sequenced and basecalled.

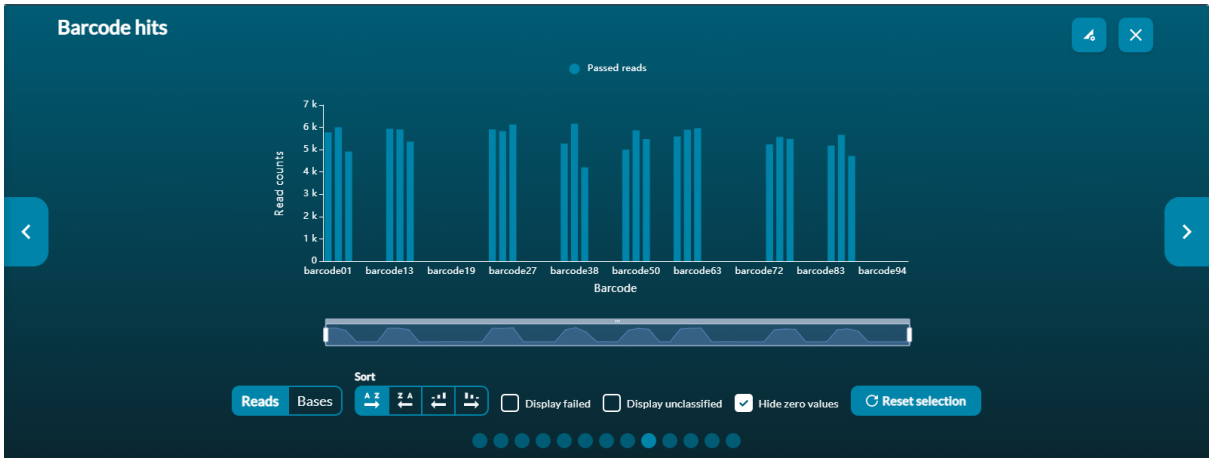



- the number of reads that have been sequenced and basecalled; and whether the reads have passed or failed the quality filters.




## Barcode read counts

The Barcode Read Counts graph shows the breakdown of barcoded reads, if barcoding was used for the experiment. The default view only shows reads that have passed the quality score filters. However, selecting the Display failed box will show all reads.



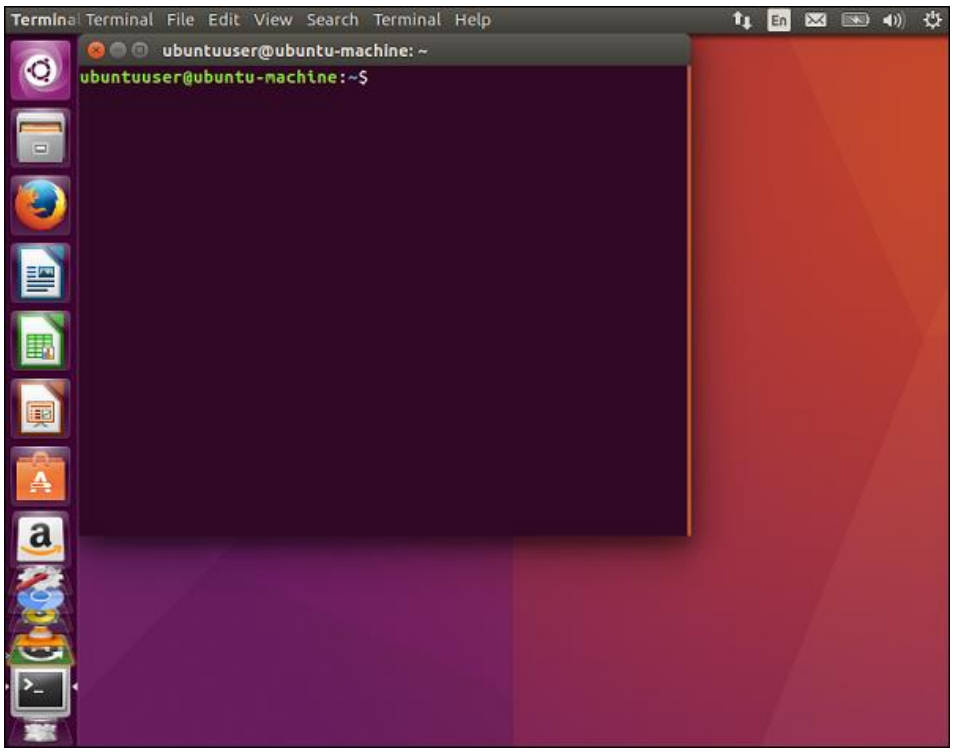
## Bioinformatics Analysis

### Pre-requisites

a) Ubuntu, version 20.04 LTS
b) MinKNOW UI
c) InterARTIC
d) The following files:
    a. .fast5,
    b. .fastq,
    c. sequencing_summary.txt,
    d. sample-barcodes.csv

### Introduction to command line interface

Open the **terminal window** in ubuntu (Image below).




#### Basic Unix commands to navigate the system

a) To know which directory you are in, you can use the ***"pwd"***
b) Use the ***"ls"*** command to know what files are in the directory you are in.
c) Use the ***"cd"*** command to go to a directory.
d) Use the ***mkdir*** command when you need to create a folder or a directory
e) Use the ***rm*** command to delete files and directories.
f) Use the ***cp*** command to copy files through the command line.

g) Use the ***mv*** command to move files through the command line.

h) The **locate** command is used to locate a file in a Linux system

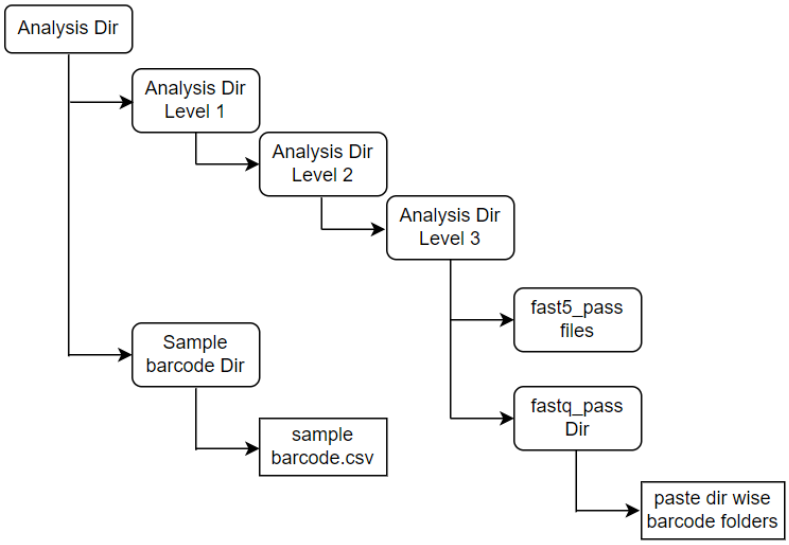
i) Use ***zip*** to compress files into a zip archive, and **unzip** to extract files from a zip archive.

## MinKNOW UI:

1. Open MinKNOW software using login credentials
2. On the left, click start -> analysis ->basecalling
3. Verify the input directory in the input page.
4. Ensure sufficient space is available (min = 100Gb) for analysis
5. Click "*continue to output*"
6. Unselect "*Compress .fastq*" files and select "*Output.fast5"* files
7. Click "*continue to basecalling*"
8. In the drop-down menu of "*configuration*", select FLO-MIN106 / FLO-FLG001 DNA – Fast (verify the flow cell name and select accordingly)
9. Click "*continue to barcoding*"
10. From the drop-down menu select the appropriate barcoding kit used for library prep (example SQK-RBK110-96).
11. Continue to Alignment
12. Continue to review
13. Note down the input and output directory
14. Click start

## Structure of Input directory:

After the MinKNOW demultiplexing is finished go the output directory. Select the fast5, fastq and summary files of the sequencing run for which the analysis is required. Select the files which has the name of the barcodes used for the run & create a directory as per hierarchy shown below.



## interARTIC usage:

There are broadly two sections into which the interARTIC analysis can be divided.

1. Genome assembly
2. Variant calling

URL to the interARTIC software and tutorial for reference (https://github.com/Psy-Fer/interARTIC )

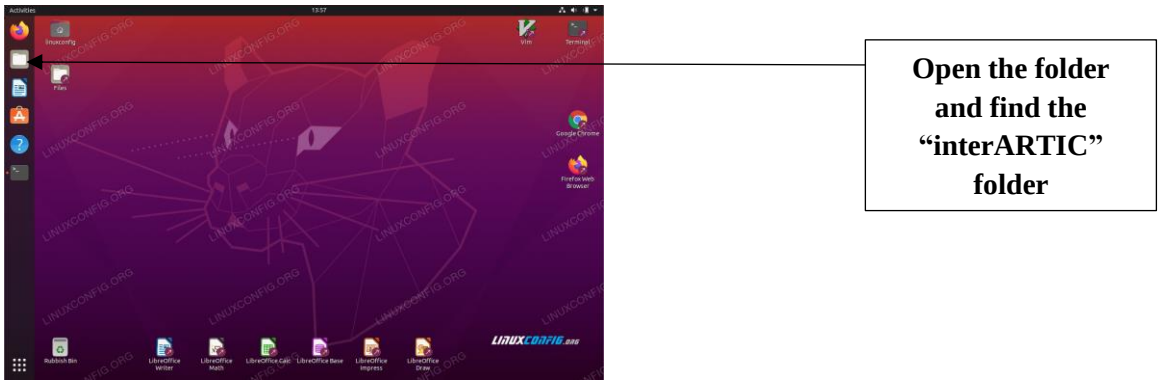
Two inbuilt tools namely "Medaka" and "Nanopolish" are employed to create consensus sequences and variant calls from nanopore sequencing data. This task is performed using neural networks applied to a pileup of individual sequencing reads against a draft assembly. It provides state-of-the-art results outperforming sequence-graph based methods and signal-based methods, whilst also being faster.

The interARTIC pipeline is a package that combines the above two tools to provide a consensus genome sequence and the variants in the sample with respect to the Wuhan CoV-19 genome.

**Steps:**

### A. Initializing the software

1. Click on the **home** button and go the folder interARTIC




2. Inside "***interartic_bin***" folder identify a file named "***run.sh***"

3. Go to you Linux command line and type ***./run.sh*** (Image below)

```
root@DESKTOP-3MOGA03:/mnt/d/FIND/software/interartic_bin# ./run.sh
```

4. The following output will be visible (Image below)

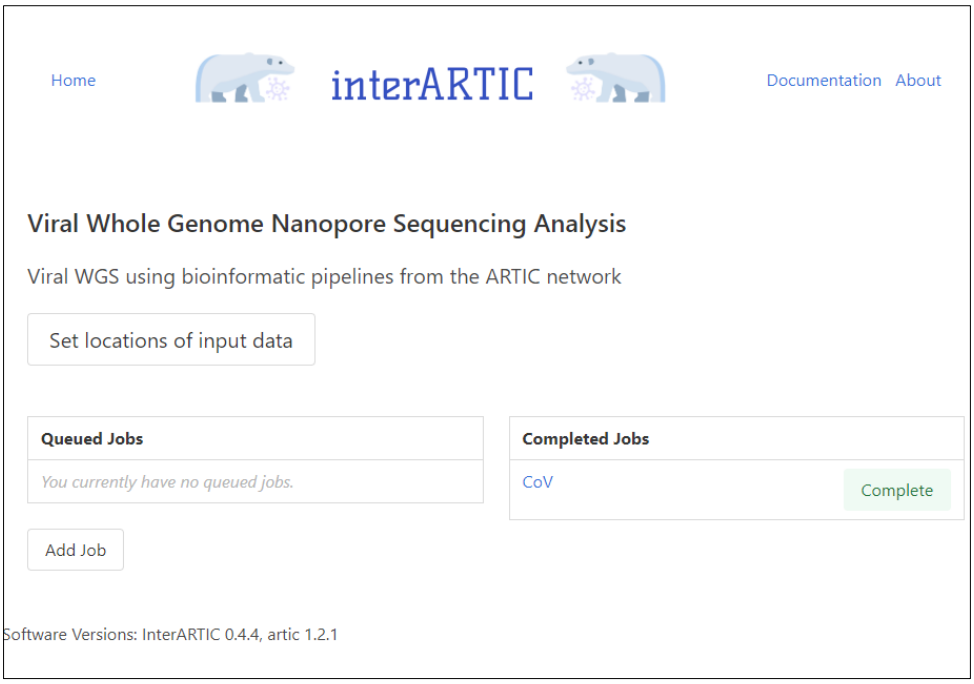
```
Starting redis server on port 7777. Log location: /mnt/d/FIND/software/interartic_bin/redis.log
Starting interartic on 127.0.0.1:5000. Log location: /mnt/d/FIND/software/interartic_bin/interartic.log
Starting celery. Log location: /mnt/d/FIND/software/interartic_bin/celery.log

InterARTIC is now running on your machine :)
To launch InterARTIC web interface visit http://127.0.0.1:5000 on your browser
To keep your InterARTIC active this terminal must remain open.
To terminate InterARTIC type CTRL-C or close the terminal.
```

5. Copy the IP address as shown here as **127.0.0.1:5000**

6. After copying, paste it in the address tab of your google chrome browser

7. Wait for few seconds and the following page will open




## B.Setting up the directory:

1. On your web page, click on the option "***Set locations of input data***"
2. In the first box type the ***"/mnt/d/FINDD/software/interartic_bin/data"*** (type your directory based on your input files)
3. In the second box enter: ***"/mnt/d/FINDD/software/interartic_bin/data/sample-barcodes"*** (type your directory based on your input files)
4. Click **confirm**

## C. Adding Job

1. You should on the interartic home page to initiate analysis. If not, click "**Home**" on the top left corner
2. Click on "**Add Job**"
3. Fill out the page using the following details:
    a. Job name: **sample_1**
    b. Select the input data directory for your experiment: **Folder_1**
    c. This input contains: **Multiple samples**
    d. Output folder: "leave blank"
    e. Select the virus you want to analyse: SARS-CoV-2
    f. Select your primer scheme: Eden V1/Midnight/etc.
    Note: The options "*Primer scheme top directory*" & "*Name of primer scheme*" will be filled automatically once you give Eden V1/Midnight/etc.
    g. Leave the Demultiplexing option unchecked
    h. Which library preparation method was used: Ligation library prep or **Rapid library prep**
    i. Select a pipeline to run: Both

## D. Hit Submit Job(s)

The run will take a while depending in the size of your input files. The following page will appear.


## D. Analysis Results

You will receive the following page after the run is over


1. Hit the button "**Go to output**"
2. On the new page hit the "**View**" button
3. You will reach the data visualisation page
4. You can go through the following outputs in detail for your sample
    a. Variants and coverage

b. Relative position of reference Vs alternate allele position

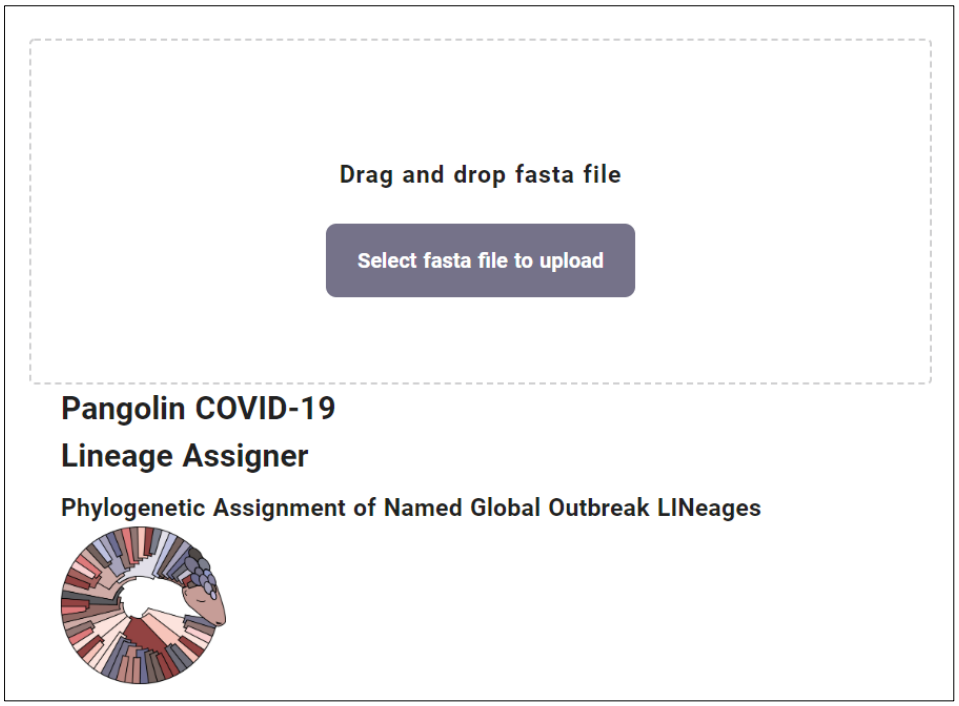
c. Sequencing related metrics

## Interpretation:

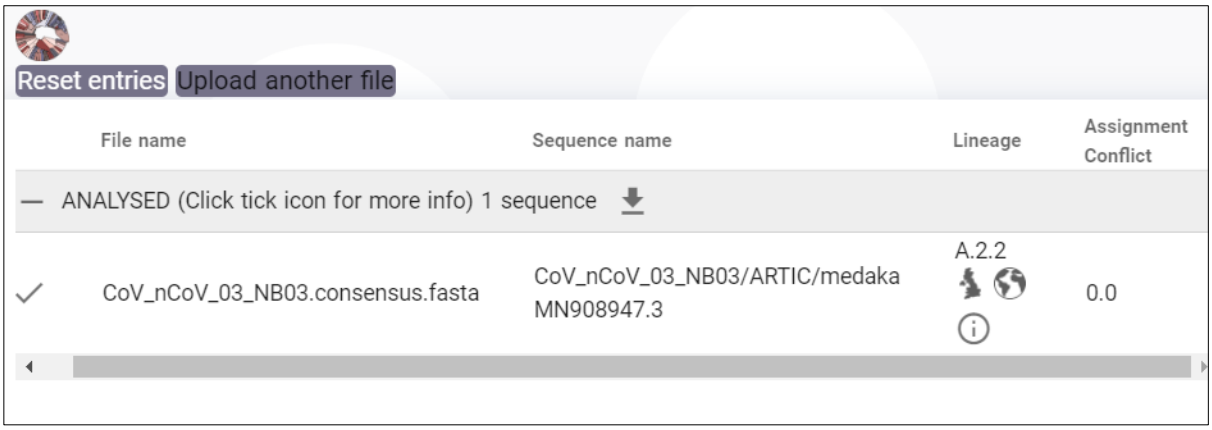
**Further interpretation (or variant annotation) can be done using two online tools namely:**

### Pango (https://pangolin.cog-uk.io/):

1. Go to the above link
2. From your output folder of the analysed run that is produced from the interARTIC pipeline, select the folder on interest (Medaka or Nanopolish) then to the sample specific folder
3. Open the folder and search for the file that ends with the name "**.consensus.fasta**"
4. This consensus file should used as input file for Pango to annotate the variants.
5. Come back to the web page of Pango
6. Drag and drop or select the above file into the page which looks like the image shown below



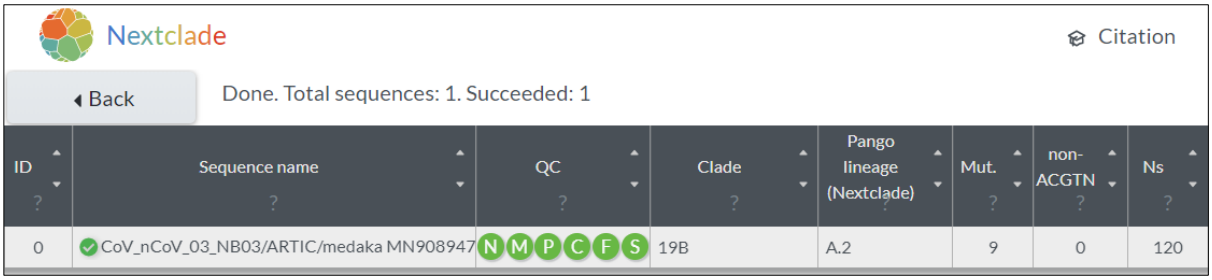

7. After importing, a new page will open with the option "**Start analysis**" on left top corner.
8. Click on it and wait for the run to finish.
9. The following page will appear




10. Click on the two symbols and see the output separately
11. The file will take some time to load
12. Analyse the output on the web page further.

Nextclade (https://clades.nextstrain.org/ )

1. Open the above link
2. Similar to the above tool, drag and drop the same file (".consensus.fasta") to the window.
3. Click the button "**Run**"
4. A new webpage will open. Wait for the below results to appear once the run is over




5. You can analyse further based on your requirement and study.

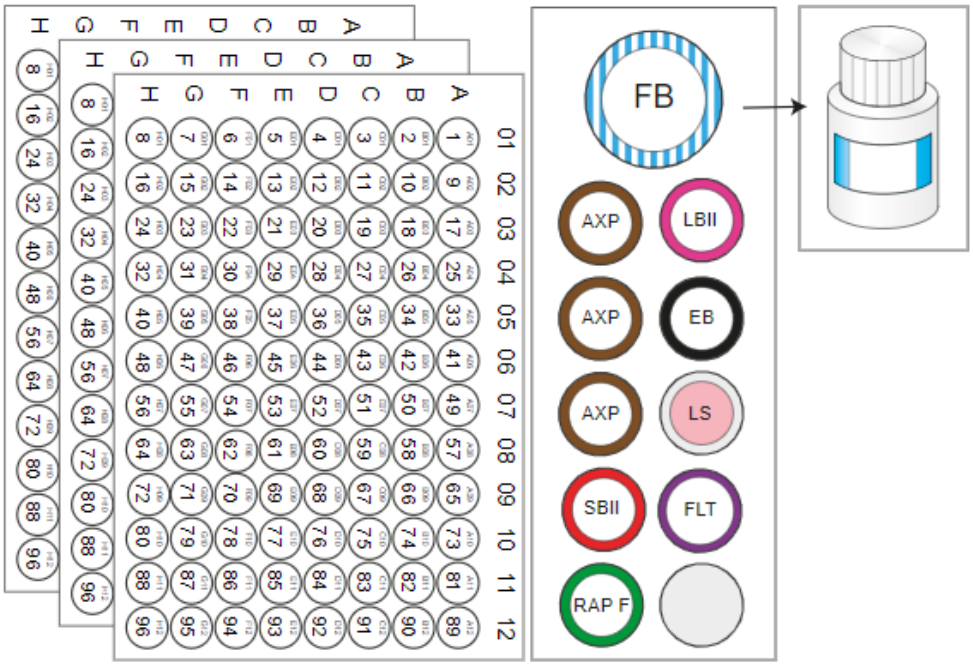
## References


1. PCR tiling of SARS-CoV-2 virus - rapid barcoding (SQK-RBK110.96). https://community.nanoporetech.com/docs/prepare/library_prep_protocols/pcr-tiling-of-sars-cov-2-virus-with-rapid-barcoding-sqk-rbk110/v/pctr_9125_v110_revh_24mar2021, requires ONT community access.
2. Post-run basecalling. https://community.nanoporetech.com/docs/prepare/library_prep_protocols/experiment-companion-minknow/v/mke_1013_v1_revcq_11apr2016/post-run-basecalling, requires ONT community access.

3. Ferguson JM, Gamaarachchi H, Nguyen T, Gollon A, Tong S, Aquilina-Reid C, Bowen-James R, Deveson IW. InterARTIC: an interactive web application for whole-genome nanopore sequencing analysis of SARS-CoV-2 and other viruses. Bioinformatics. 2022 Feb 7;38(5):1443-1446. doi: 10.1093/bioinformatics/btab846. PMID: 34908106; PMCID: PMC8826086.
4. Aksamentov, I., Roemer, C., Hodcroft, E. B., & Neher, R. A., (2021). Nextclade: clade assignment, mutation calling and quality control for viral genomes. Journal of Open Source Software, 6(67), 3773, https://doi.org/10.21105/joss.03773 (https://clades.nextstrain.org).

## Appendix

### Rapid Barcoding Kit 96 (SQK-RBK110.96)




FB : Flush Buffer
AXP : AMPure XP Beads
SBII : Sequencing Buffer II
RAP F : Rapid Adapter F
LBII : Loading Beads II
EB : Elution Buffer
LS : Loading Solution
FLT : Flush Tether

| Name | Acronym | Cap color | No. of vials | Fill volume per vial (µl) |
|---|---|---|---|---|
| Rapid Barcode plate | RB96 | - | 3 plates | 8 µl per well |
| AMPure XP Beads | AXP | Brown | 3 | 1,200 |
| Sequencing Buffer II | SBII | Red | 1 | 500 |
| Rapid Adapter F | RAP-F | Green | 1 | 25 |
| Elution Buffer | EB | Black | 1 | 500 |
| Loading Beads II | LBII | Pink | 1 | 360 |
| Loading Solution | LS | White cap, pink label | 1 | 400 |
| Flush Tether | FLT | Purple | 1 | 400 |

| Name | Acronym | Cap color | No. of vials | Fill volume per vial (µl) |
|---|---|---|---|---|
| Flush Buffer | FB | White | 1 bottle | 15,500 |

### Rapid barcode sequences

| Component | Sequence |
|---|---|
| RB01 | AAGAAAGTTGTCGGTGTCTTTGTG |
| RB02 | TCGATTCCGTTTGTAGTCGTCTGT |
| RB03 | GAGTCTTGTGTCCCAGTTACCAGG |
| RB04 | TTCGGATTCTATCGTGTTTCCCTA |
| RB05 | CTTGTCCAGGGTTTGTGTAACCTT |
| RB06 | TTCTCGCAAAGGCAGAAAGTAGTC |
| RB07 | GTGTTACCGTGGGAATGAATCCTT |
| RB08 | TTCAGGGAACAAACCAAGTTACGT |
| RB09 | AACTAGGCACAGCGAGTCTTGGTT |
| RB10 | AAGCGTTGAAACCTTTGTCCTCTC |
| RB11 | GTTTCATCTATCGGAGGGAATGGA |
| RB12 | CAGGTAGAAAGAAGCAGAATCGGA |
| RB13 | AGAACGACTTCCATACTCGTGTGA |
| RB14 | AACGAGTCTCTTGGGACCCATAGA |
| RB15 | AGGTCTACCTCGCTAACACCACTG |
| RB16 | CGTCAACTGACAGTGGTTCGTACT |

| Component | Sequence |
|---|---|
| RB17 | ACCCTCCAGGAAAGTACCTCTGAT |
| RB18 | CCAAACCCAACAACCTAGATAGGC |
| RB19 | GTTCCTCGTGCAGTGTCAAGAGAT |
| RB20 | TTGCGTCCTGTTACGAGAACTCAT |
| RB21 | GAGCCTCTCATTGTCCGTTCTCTA |
| RB22 | ACCACTGCCATGTATCAAAGTACG |
| RB23 | CTTACTACCCAGTGAACCTCCTCG |
| RB24 | GCATAGTTCTGCATGATGGGTTAG |
| RB25 | GTAAGTTGGGTATGCAACGCAATG |
| RB26 | CATACAGCGACTACGCATTCTCAT |
| RB27 | CGACGGTTAGATTCACCTCTTACA |
| RB28 | TGAAACCTAAGAAGGCACCGTATC |
| RB29 | CTAGACACCTTGGGTTGACAGACC |
| RB30 | TCAGTGAGGATCTACTTCGACCCA |
| RB31 | TGCGTACAGCAATCAGTTACATTG |
| RB32 | CCAGTAGAAGTCCGACAACGTCAT |
| RB33 | CAGACTTGGTACGGTTGGGTAACT |
| RB34 | GGACGAAGAACTCAAGTCAAAGGC |

| Component | Sequence |
|---|---|
| RB35 | CTACTTACGAAGCTGAGGGACTGC |
| RB36 | ATGTCCCAGTTAGAGGAGGAAACA |
| RB37 | GCTTGCGATTGATGCTTAGTATCA |
| RB38 | ACCACAGGAGGACGATACAGAGAA |
| RB39 | CCACAGTGTCAACTAGAGCCTCTC |
| RB40 | TAGTTTGGATGACCAAGGATAGCC |
| RB41 | GGAGTTCGTCCAGAGAAGTACACG |
| RB42 | CTACGTGTAAGGCATACCTGCCAG |
| RB43 | CTTTCGTTGTTGACTCGACGGTAG |
| RB44 | AGTAGAAAGGGTTCCTTCCCACTC |
| RB45 | GATCCAACAGAGATGCCTTCAGTG |
| RB46 | GCTGTGTTCCACTTCATTCTCCTG |
| RB47 | GTGCAACTTTCCCACAGGTAGTTC |
| RB48 | CATCTGGAACGTGGTACACCTGTA |
| RB49 | ACTGGTGCAGCTTTGAACATCTAG |
| RB50 | ATGGACTTTGGTAACTTCCTGCGT |
| RB51 | GTTGAATGAGCCTACTGGGTCCTC |
| RB52 | TGAGAGACAAGATTGTTCGTGGAC |

| Component | Sequence |
|---|---|
| RB53 | AGATTCAGACCGTCTCATGCAAAG |
| RB54 | CAAGAGCTTTGACTAAGGAGCATG |
| RB55 | TGGAAGATGAGACCCTGATCTACG |
| RB56 | TCACTACTCAACAGGTGGCATGAA |
| RB57 | GCTAGGTCAATCTCCTTCGGAAGT |
| RB58 | CAGGTTACTCCTCCGTGAGTCTGA |
| RB59 | TCAATCAAGAAGGGAAAGCAAGGT |
| RB60 | CATGTTCAACCAAGGCTTCTATGG |
| RB61 | AGAGGGTACTATGTGCCTCAGCAC |
| RB62 | CACCCACACTTACTTCAGGACGTA |
| RB63 | TTCTGAAGTTCCTGGGTCTTGAAC |
| RB64 | GACAGACACCGTTCATCGACTTTC |
| RB65 | TTCTCAGTCTTCCTCCAGACAAGG |
| RB66 | CCGATCCTTGTGGCTTCTAACTTC |
| RB67 | GTTTGTCATACTCGTGTGCTCACC |
| RB68 | GAATCTAAGCAAACACGAAGGTGG |
| RB69 | TACAGTCCGAGCCTCATGTGATCT |
| RB70 | ACCGAGATCCTACGAATGGAGTGT |

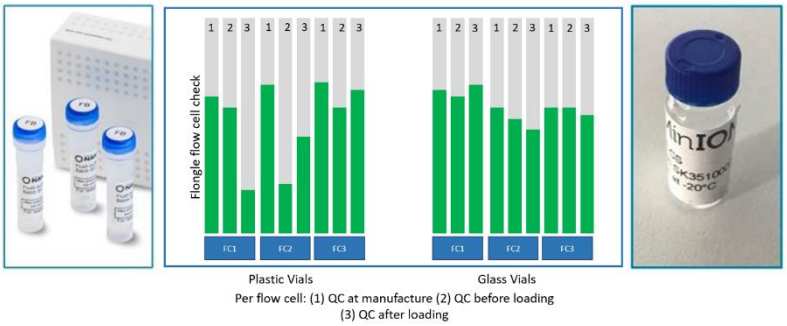
| Component | Sequence |
|---|---|
| RB71 | CCTGGGAGCATCAGGTAGTAACAG |
| RB72 | TAGCTGACTGTCTTCCATACCGAC |
| RB73 | AAGAAACAGGATGACAGAACCCTC |
| RB74 | TACAAGCATCCCAACACTTCCACT |
| RB75 | GACCATTGTGATGAACCCTGTTGT |
| RB76 | ATGCTTGTTACATCAACCCTGGAC |
| RB77 | CGACCTGTTTCTCAGGGATACAAC |
| RB78 | AACAACCGAACCTTTGAATCAGAA |
| RB79 | TCTCGGAGATAGTTCTCACTGCTG |
| RB80 | CGGATGAACATAGGATAGCGATTC |
| RB81 | CCTCATCTTGTGAAGTTGTTTCGG |
| RB82 | ACGGTATGTCGAGTTCCAGGACTA |
| RB83 | TGGCTTGATCTAGGTAAGGTCGAA |
| RB84 | GTAGTGGACCTAGAACCTGTGCCA |
| RB85 | AACGGAGGAGTTAGTTGGATGATC |
| RB86 | AGGTGATCCCAACAAGCGTAAGTA |
| RB87 | TACATGCTCCTGTTGTTAGGGAGG |
| RB88 | TCTTCTACTACCGATCCGAAGCAG |

| Component | Sequence |
|---|---|
| RB89 | ACAGCATCAATGTTTGGCTAGTTG |
| RB90 | GATGTAGAGGGTACGGTTTGAGGC |
| RB91 | GGCTCCATAGGAACTCACGCTACT |
| RB92 | TTGTGAGTGGAAAGATACAGGACC |
| RB93 | AGTTTCCATCACTTCAGACTTGGG |
| RB94 | GATTGTCCTCAAACTGCCACCTAC |
| RB95 | CCTGTCTGGAAGAAGAATGGACTT |
| RB96 | CTGAACGGTCATAGAGTCCACCAT |

### Flongle Flow Cell Priming Kit (EXP-FSE001)

There are three buffers that come into direct contact with a flow cell at point of loading (SBII: Sequencing Buffer II, FB: Flush Buffer and LB II: Loading Beads II or LS: Loading Solution). When looking at these buffers, we found that there are a very low level of contaminants seeping out of the plastic vials that impacts the robustness of the Flongle flow cell system (MinION and PromethION are not impacted by this).

ONT have found that when storing these buffers in glass vials instead of plastic, incidence of deterioration is reduced.




To load a library onto your Flongle flow cell, you will need to use the following components:

**Flongle Flow Cell Priming Kit (EXP-FSE001) components**

- Sequencing Buffer II (SBII)

- Flush Buffer (FB)

- Loading Beads II (LBII) or Loading Solution (LS)

**Sequencing or Flow Cell Priming Kit components**

- Flush Tether (FLT)

Oxford Nanopore Technologies deem the useful life of the Flow Cell Expansion to be 6 months from receipt by the customer.

# SOP-002

# ARTIC amplicon sequencing using Native barcoding protocol for SARS-CoV-2 Genome

| | LABORATORY STANDARD OPERATING PROCEDURES | | | |
|---|---|---|---|---|
| Country and Lab Demo Site | India | **DOCUMENT NUMBER** | **VERSION** | **EFFECTIVE DATE** |
| | | SOP-002 | 3.6 | |
| **PROCEDURE** | **ARTIC amplicon sequencing using Native barcoding protocol for SARS-CoV-2 Genome** | | | |

| | | | | |
|---|---|---|---|---|
| Approved By | Name, Title | | Signature | Date |
| | | | | |
| | | | | |
| SOP Annual Review | Name, Title | | Signature | Date |
| | | | | |
| | | | | |
| | | | | |
| Revision History | Version # [0.0] | Revision Date [dd/mm/yy] | Description (notes) | |
| | | | | |
| | | | | |
| | | | | |
| Distributed Copies to | Name (or location) | # of copies | Name (or location) | # of copies |
| | | | | |
| | | | | |
| | | | | |

***I acknowledge that I have read, understand and agree to follow this SOP.***

| Name (print) | Signature | Designation | Date |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

## Table of Contents

## 1. PURPOSE

- This Standard Operating Procedure provides guidance to perform library preparation and sequencing of SARS-CoV-2 genome on the MinION nanopore sequencer using ARTIC protocol.
- This SOP contains instructions to prepare cDNA libraries using extracted RNA from human nasopharyngeal swab specimens positive for SARS-CoV-2. The tiling amplicon method was used to enrich the viral target from cDNA.
- This SOP also contains instructions for priming and loading of libraries onto the flongle and initiate sequencing.

## 2. INTRODUCTION

The SARS-CoV-2 has evolved continuously since its first identification in January 2020. Several variants have since emerged including Alpha (B.1.1.7), Beta (B.1.351), Gamma (P1), Delta (B.1.617.2) and Omicron (B.1.1.529). Since the beginning of the COVID-19 pandemic, multiple variants have shown higher transmissibility, virulence and immune escape resulting in higher rates of infection and subsequent morbidity and mortality globally.

Therefore, it is critical to identify variants early to inform the public health response through genomic surveillance for pandemic control and apply a systematic approach to provide a representative indication of the extent of transmission of SARS-CoV-2 variants based on the local context.

## 3. PRINCIPLE

Next Generation Sequencing (NGS) is a universal tool for rapidly sequencing genomes of microorganisms. It can help identify and track variants of concern and variants of interest (VoI). (3) Additionally, NGS provides valuable critical evidence for:

- Tracking of transmission routes of the virus globally
- Detection of mutations in a timely manner, to prevent the spread of new strains
- Identification of viral mutations that can affect vaccine potency or avoid detection by established molecular diagnostic assays
- Screening of targets for possible COVID-19 therapeutics
- Identification/characterization of respiratory co-infections and antimicrobial resistance alleles (4).

### 3.1 Sample acceptance criteria

- Viral RNA should be extracted freshly and not more than 24 hours prior to library preparation are preferred or stored appropriately at -80°C and not freeze-thawed.
- CT value for the extracted RNA samples should be preferably below 30.

### 3.2 Specimen rejection criteria

- Samples not maintained at -80°C during storage and underwent multiple freeze-thaw.
- Ct Value of the RNA sample above 30.

## 4. PERSONNEL QUALIFICATIONS

### 4.1 Medical fitness

- All personnel involved in specimen receipt and handling should be tested for COVID-19 beforehand. Only those who test negative should be involved in performing the test procedure.
- Options for reassignment of personnel with COVID-19 comorbid conditions, such as diabetes, chronic respiratory diseases, high blood pressure, or immuno-suppressed individuals away from the high-risk areas of the COVID laboratory should be considered.
- If the resources are limited, they should be made aware of the risk of experiencing severe symptoms of the disease.

### 4.2 Education and training

Education and training must be given on the following topics:

- Potential risks to health (symptoms of COVID-19 disease and transmission)
- Precautions to be taken to minimize droplets & aerosol formation and prevent exposure
- Hygiene requirements
- Donning and Doffing of PPE
- Laboratory biosafety, specifically handling of potentially infectious materials
- Laboratory design, including airflow conditions
- Use of biological safety cabinets (operation, identification of malfunctions and maintenance)
- Use of autoclaves, microcentrifuge, micropipettes & refrigerators, (operation, identification of malfunctions and maintenance)
- Prevention of incidents and steps to be taken by workers in the case of incidents (biohazard incidents, chemical, electrical and fire hazards)
- Good laboratory practice and good microbiological techniques
- Workflow in laboratory
- Procedure to be performed
- Waste handling
- Importance of laboratory results for individual patient and COVID pandemic management

## 5. RESPONSIBILITIES

It is the responsibility of the lab personnel to correctly understand and perform this procedure. All users of this procedure who do not understand the procedure or unable to carry it out as described are responsible for seeking advice from their supervisor.

## 6. LIBRARY PREPARATION

### 6.1 Materials**

1) Template – Extracted RNA (not more than 24hrs and with a CT value below 30)
2) LunaScript RT SuperMix Kit (NEB, E3010L)
3) Q5 Hot Start High-Fidelity 2X Master Mix (NEB, M0494L)
4) COVID ARTIC WGS Primer Pools A & B (300 rxns)
5) Ligation Sequencing Kit (SQK-LSK110)
6) Native Barcoding Expansion 1-12 (EXP-NBD104) or the Native Barcoding Expansion 13-24 (EXP-NBD114)
7) NEBNext Companion Module for Oxford Nanopore Technologies Ligation Sequencing (E7180S)
8) NEBNext Quick Ligation Module (E6056L)
9) SPRI Beads (Ampure XP 60ml) (Beckman Coulter, A63881)
10) Nuclease-free water
11) Freshly prepared 80% ethanol in nuclease-free water
12) Qubit dsDNA HS Assay Kit (ThermoFisher, Q32851)

#### 6.1.1 Recommended Storage

Please refer to annexure IV for recommended storage condition for all the reagents used in the workflow.

### 6.2 Consumables**

1) 1.5 ml micro centrifuge tubes
2) 8-well PCR strips or 0.1 ml PCR tubes

### 6.3 Equipments**

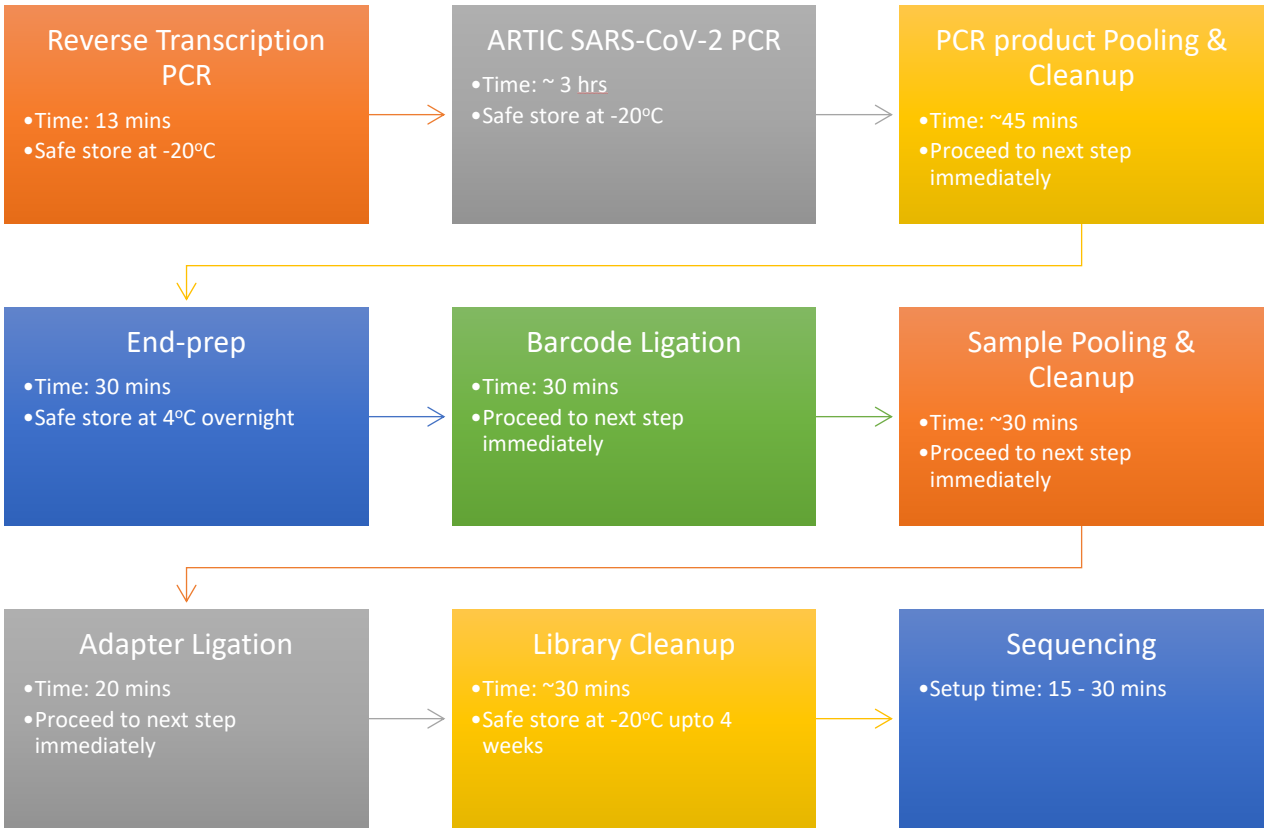
1) Magnetic stand, suitable for 1.5 ml Eppendorf tubes
2) 1.5 ml Centrifuge
3) Microfuge (compatible for 0.1 ml PCR tube and PCR strip)
4) Vortex mixer
5) Thermal cycler (conventional/ real time instrument)
6) Multichannel pipettes suitable for dispensing 0.5–10 µl, 2–20 µl and 20–200 µl,
7) P1000 pipette
8) P200 pipette
9) P100 pipette
10) P20 pipette
11) P10 pipette
12) 100 – 1000 µl filter tips
13) 20 – 200 µl filter tips

14) 10 – 100 µl filter tips
15) 2 – 20 µl filter tips
16) 0.5 –10 µl filter tips
17) Timer
18) Qubit fluorometer (or equivalent for QC check)
19) PCR hood with UV sterilizer (optional but recommended to reduce cross-contamination)
20) PCR-Coolant or ice bucket separate for RNA samples and Libraries

Please refer to Annexure III for room wise consumables & reagents requirement for the SARS-CoV-2 genome sequencing workflow.

### 6.4 Procedure

SARS-CoV-2 Genome Sequencing involves several steps as follows. Overall, it takes 1- 1.5 days to complete the preparation and loading of library onto the flongle/flow cell based on the number of samples processed.



### 6.4.1 Reverse Transcription PCR

Note: Keep the RNA sample on ice as much as possible to prevent nucleolytic degradation, which may affect sensitivity.

This protocol guides for the preparation of library for a single sample. However, reaction volumes & reagents may be calculated according to the number of samples to be processed. In a PCR strip, maximum of 8 samples can be processed. For more than 8 samples, 96 well PCR plate is suggested.

6.4.1.1 In a clean pre-PCR hood, using a single or a multichannel pipette, add 2 µl of LunaScript RT SuperMix to a fresh 8-well PCR strip or 0.1 ml PCR tube.

6.4.1.2 To each well containing LunaScript reagent of the RT plate, add 8 µl of RNA sample and gently mix by pipetting 10 times.

6.4.1.3 Close the PCR tubes and spin down. Return the tube to ice.

6.4.1.4 Preheat the thermal cycler to 25°C.

6.4.1.5 Incubate the samples in the thermal cycler using the following program:

| Step | Temperature | Time | Cycles |
|---|---|---|---|
| Primer annealing | 25°C | 2 min | 1 |
| cDNA synthesis | 55°C | 10 min | 1 |
| Heat inactivation | 95°C | 1 min | 1 |
| Hold | 4°C | ∞ | |

While the reverse transcription reaction is running, prepare the Pool A & Pool B PCR master mixes as described in the next section.

Safe Stopping Point: RT-PCR product can be stored at -20°C up to 7 days.

### 6.4.2 ARTIC SARS-CoV-2 PCR

These primers are designed to generate 400 bp amplicons.

6.4.2.1 In the pre-PCR hood, prepare the following master mixes (Pool A & Pool B) in two separate 1.5 ml micro centrifuge tubes and mix thoroughly as follows:

Reaction volume per sample:

| Reagent | Pool A | Pool B |
|---|---|---|
| Nuclease free water | 3 µl | 3 µl |
| Primer pool A | 3 µl | - |
| Primer pool B | - | 3 µl |
| Q5 Hot Start HF 2x Master Mix | 10 µl | 10 µl |
| **Total** | **16 µl** | **16 µl** |

6.4.2.2 Using a single or multichannel pipette, aliquot 16 µl of Pool A and Pool B into a clean 8-well PCR strip or tubes based on number of samples as follows:

| X3 sample | X5 samples |
|---|---|
| Strip 1 - Pool A: 1-3<br>Strip 1 - Pool B: 4-6 | Strip 1 - Pool A: 1-5<br>Strip 2 - Pool B: 1-5 |

6.4.2.3 Using a single or multichannel pipette, transfer 4 µl of each cDNA from the PCR tube to the corresponding well for both Pool A and Pool B of the PCR plate(s) or strip(s). Mix by pipetting the contents of each well up and down for 10 times.

6.4.2.4 Seal the PCR strip(s) or tubes and spin down briefly.

6.4.2.5 Incubate using the following program, with the heated lid set to 105°C:

| Step | Temperature | Time | Cycles |
|---|---|---|---|
| Initial denaturation | 98°C | 30 sec | 1 |
| Denaturation,<br>Annealing &<br>Extension | 98°C<br>65°C | 15 sec<br>5 min | 32** |
| Hold | 4°C | ∞ | |

** No of cycles recommendation based on RNA Ct value

| Ct Value | No of Cycles |
|---|---|
| <20 | 25 |
| 20 - 30 | 30 |

Optional: Check the quality of the amplification by running an 1.2% agarose gel using 3 µl of PCR products & quantity by Qubit fluorometer.

Safe Stopping Point: PCR product be stored at -20°C up to 7 days.

### 6.4.3 PCR Product Pooling & Clean-up

Note: Equilibrate SPRI beads at room temperature for at least 30 minutes prior to use.

6.4.3.1 Using a single or a multichannel pipette, transfer 15 µl of each well of PCR Pool A product & the corresponding well of PCR Pool B product (total volume of pooled PCR product should be 30 µl) to the fresh 1.5 ml micro centrifuge tube individually.

6.4.3.2 Resuspend the SPRI beads by vortexing vigorously.

6.4.3.3 To the pooled PCR product of each sample, add an equal volume (1X) of resuspended SPRI beads and mix by flicking the tube or pipette mixing for at least 10 times.

Example: For 30 µl of a sample (PCR A & B pooled), 30 µl (1X) of SPRI beads should be added.

6.4.3.4 Incubate for 5 minutes at room temperature (pulse flick the tube if the beads are settled at the bottom of the tube).

6.4.3.5 Prepare at least 2 – 10 ml of fresh 80% ethanol in nuclease-free water.

6.4.3.6 Spin down the sample and pellet on a magnetic separation rack. Keep the tube on the magnet, and pipette off the supernatant without touching the beads once the solution is transparent.

6.4.3.7 Keep the tube on the magnet and wash the beads with 200 µl of freshly prepared 80% ethanol without disturbing the pellet for 30 seconds. Remove the ethanol using a pipette and discard.

Note: 80% Ethanol volume varies based on the size of the pellet formed. Ensure to immerse the beads with sufficient ethanol.

6.4.3.8 Repeat the previous step (total of two washes).

6.4.3.9 Briefly spin down and place the tube back on the magnetic stand. Pipette off any residual ethanol. Allow to dry the beads by keeping the tube cap open for 30 seconds, but do not dry the pellet to the point of cracking.

6.4.3.10 Remove the tube from the magnetic rack and resuspend the pellet by pipetting in 15 µl of Nuclease free water and Incubate for 5 minutes at room temperature.

6.4.3.11 Pellet the beads on a magnetic stand until the eluate is clear and colorless.

6.4.3.12 Remove and retain 15 µl of eluate containing the DNA library into a clean 1.5 ml micro centrifuge tube.

6.4.3.13 Quantify pooled PCR product using the Qubit dsDNA HS Assay Kit (Refer to Annexure 1 for the SOP).

### 6.4.4 End-prep

6.4.4.1 Determine the volume of the cleaned-up PCR product that yields 300 - 500 ng of DNA per sample and aliquot in a clean PCR strip or 0.1 ml PCR tubes.

6.4.4.2 Make up each sample per well to 12.5 µl using nuclease-free water.

6.4.4.3 Prepare the following end-prep master mix in a 1.5 ml micro centrifuge tube and mix thoroughly.

| Reagent | Volume per Sample |
|---|---|
| NEBNext Ultra II End Prep Reaction Buffer | 1.75 µl |
| NEBNext Ultra II End Prep Enzyme Mix | 0.75 µl |
| **Total** | **2.5 µl** |

6.4.4.4 Using a single channel pipette, aliquot 2.5 µl of the end-prep master mix to the PCR strip or tubes containing PCR product.

6.4.4.5 Pipette mix well and spin down briefly.

6.4.4.6 Using a thermal cycler, incubate at 20°C for 5 minutes and 65°C for 5 minutes.

Safe Stopping Point: End-Prepped DNA product can be stored at 4ºC overnight.

### 6.4.5 Native Barcode Ligation

6.4.5.1 Thaw the native barcodes at room temperature. Use one barcode per sample. Individually mix the barcodes by pipetting, spin down, and place them on ice.

6.4.5.2 Take 1.5 µl of end-prepped DNA into a fresh PCR strip or tube and add 2.5 µl of unique barcode for every sample.

6.4.5.3 Prepare the following native barcode ligation master mix in a 1.5 ml micro centrifuge tube and mix thoroughly.

| Reagent | Volume per Sample |
|---|---|
| Ligation Buffer (LNB) | 5 µl |
| Quick T4 DNA Ligase | 5 µl |
| Nuclease free water | 6 µl |
| **Total** | **16 µl** |

6.4.5.4 Using a single channel pipette, aliquot 16 µl of the native barcode ligation master mix to the PCR strip or tube containing end-prepped DNA & a unique barcode.

6.4.5.5 Pipette mix well and spin down briefly.

6.4.5.6 Using a thermal cycler, incubate at 20°C for 20 minutes and 65°C for 10 minutes.

### 6.4.6 Sample pooling and clean-up

Note: Equilibrate SPRI beads at room temperature for at least 30 minutes prior to use.

6.4.6.1 Briefly spin down the barcoded samples to collect the liquid at the bottom of the wells prior to opening.

6.4.6.2 Pool the barcoded samples into a 1.5 ml micro centrifuge tube (expected to have about ~20 µl per sample).

Example: if 5 samples were used, the total volume should make up to 100 µl in the 1.5ml micro centrifuge tube.

6.4.6.3 Resuspend the SPRI beads by vortexing vigorously.

6.4.6.4 To the entirely pooled barcoded sample, add 0.8X volume of resuspended SPRI beads and mix by flicking the tube or pipette mixing for at least 10 times.

Example: To the 100 µl of pooled samples, 80 µl (0.8X) of SPRI beads should be added.

6.4.6.5 Incubate for 5 minutes at room temperature (pulse flick the tube if the beads are settled at the bottom of the tube).

6.4.6.6 Prepare at least 0.5 - 2 ml of fresh 80% ethanol in nuclease-free water.

6.4.6.7 Spin down the sample and pellet on a magnet. Keep the tube on the magnet, and pipette off the supernatant without touching the beads once the solution is transparent.

6.4.6.8 Keep the tube on the magnet and wash the beads with 200 µl of freshly prepared 80% ethanol without disturbing the pellet for 30 seconds. Remove the ethanol using a pipette and discard.

Note: 80% Ethanol volume varies based on the size of the pellet formed. Ensure to immerse the beads with sufficient ethanol.

6.4.6.9 Repeat the previous step (total of two washes).

6.4.6.10 Briefly spin down and place the tube back on the magnetic stand. Pipette off any residual ethanol. Allow to dry the beads by keeping the tube cap open for 30 seconds, but do not dry the pellet to the point of cracking.

6.4.6.11 Remove the tube from the magnetic rack and resuspend the pellet by pipetting in 12 µl Nuclease free water. Incubate for 5 minutes at room temperature.

6.4.6.12 Pellet the beads on a magnetic stand until the eluate is clear and colorless.

6.4.6.13 Remove and retain 12 µl of eluate containing the pooled barcoded sample into a clean 1.5 ml micro centrifuge tube.

6.4.6.14 Quantify pooled barcoded sample concentration using the Qubit dsDNA HS Assay Kit (Refer to Annexure 1 for the SOP).

### 6.4.7 Adapter Ligation

6.4.7.1 Take forward 10 - 15 ng of barcoded sample and make up the volume to 14.5 µl with Nuclease free water.

6.4.7.2 Prepare the following adapter ligation master mix in a 1.5 ml micro centrifuge tube and mix thoroughly.

| Reagent | Volume per Sample |
|---|---|
| NEBNext Quick Ligation Buffer (5X) | 5 µl |
| Quick T4 DNA Ligase | 3 µl |
| Adapter Mix II (AMII) | 2.5 µl |
| **Total** | **10.5 µl** |

6.4.7.3 Using a single channel pipette, aliquot 10.5 µl of the adapter ligation master mix to the fresh PCR strip or tubes containing pooled barcoded sample.

6.4.7.4 Pipette mix well and spin down briefly.

6.4.7.5 Using a thermal cycler, incubate at 20°C for 20 minutes.

### 6.4.8 Adapter Ligation Clean-up

Note: Clean-up step uses SFB (Short Fragment Buffer) and not 80% ethanol to wash the beads. The use of ethanol will be detrimental to the sequencing reaction.

6.4.8.1 Briefly spin down the Adapter Ligated library to collect the liquid at the bottom of the wells prior to opening.

6.4.8.2 Transfer the adapter ligated library into a 1.5 ml micro centrifuge tube (expected to have about ~25 µl).

6.4.8.3 Resuspend the SPRI beads by vortexing vigorously.

6.4.8.4 To the adapter ligated library, add 0.8X volume of resuspended SPRI beads and mix by flicking the tube or pipette mixing for at least 10 times.

Example: For 25 µl of adapter ligated library, 20 µl (0.8X) of SPRI beads should be added.

6.4.8.5 Incubate for 5 minutes at room temperature (pulse flick the tube if the beads are settled at the bottom of the tube).

6.4.8.6 Thaw the tube of Short Fragment Buffer (SFB) at room temperature, mix by vortexing, spin down and place on ice.

6.4.8.7 Spin down the sample and pellet on a magnet. Keep the tube on the magnet, and pipette off the supernatant without touching the beads once the solution is transparent.

6.4.8.8 Wash the beads by adding 100 µl Short Fragment Buffer (SFB). Flick the beads to resuspend, then return the tube to the magnetic rack and allow the beads to pellet. Remove the supernatant using a pipette and discard.

6.4.8.9 Repeat the previous step (total of two washes).

6.4.8.10 Briefly spin down and place the tube back on the magnetic stand. Pipette off any residual SFB buffer. Allow to dry the beads by keeping the tube cap open for 30 seconds, but do not dry the pellet to the point of cracking.

6.4.8.11 Remove the tube from the magnetic rack and resuspend the pellet by pipetting in 7 µl Elution buffer (EB). Incubate for 5 minutes at room temperature.

6.4.8.12 Pellet the beads on a magnetic stand until the eluate is clear and colorless.

6.4.8.13 Remove and retain 7 µl of eluate containing the purified library into a clean 1.5 ml micro centrifuge tube.

6.4.8.14 Quantify the library concentration using the Qubit dsDNA HS Assay Kit (Refer to Annexure 1 for the SOP).

The final library is used for loading into the Flongle flow cell. Store the library on ice until ready to load.

Safe Stopping Point: Final library product should be stored at -20°C upto four weeks.

## 7. SEQUENCING

### 7.1 Priming & Loading the Flongle flow cell

#### 7.1.1 Materials**:

1) Flush Buffer (FB)
2) Flush Tether (FLT)
3) Loading Beads II (LBII)
4) Sequencing Buffer II (SBII)

** The above reagents are either available in Ligation Sequencing Kit (SQK-LSK110) or Flongle Sequencing Expansion kit (EXP-FSE001).

#### 7.1.2 Consumables

1) 1.5 ml micro centrifuge tubes
2) Flongle flow cell (FLO-FLG001)

#### 7.1.3 Equipment

1) MinION Mk1B
2) P200 pipette
3) P20 pipette

4) P10 pipette
5) 20 – 200 µl filter tips
6) 2 – 20 µl filter tips
7) 0.5 –10 µl filter tips

### 7.1.4 Procedure

7.1.4.1 Thaw the Sequencing Buffer II (SBII), Loading Beads II (LBII), Flush Buffer (FB) and Flush Tether (FLT) from the Ligation Sequencing kit (SQK-LSK110) at room temperature.

7.1.4.2 Mix the Sequencing Buffer II (SBII), Flush Buffer (FB) by vortexing and Flush Tether (FLT) tubes by pipette mixing and spin down at room temperature.

7.1.4.3 In the post PCR hood, prepare the following solution mixes in a fresh 1.5 ml micro centrifuge tubes separately & mix thoroughly as follows.

**Priming Solution:**

| Reagents | Volume |
|---|---|
| Flush Buffer (FB) | 117 µl |
| Flush Tether (FLT) | 3 µl |

**Sequencing Mix:**

| Reagents | Volume |
|---|---|
| Sequencing Buffer II (SBII) | 15 µl |
| Loading Beads II (LBII) mixed immediately before use | 10 µl |
| DNA library | 4 – 5 ng |
| Elution Buffer | Up to 5 µl |
| **Total** | **30 µl** |

Note: Load the library onto the flongle flow cell immediately after adding the **Sequencing Buffer II (SBII)** and **Loading Beads II (LBII)** because the fuel in the buffer will start to be consumed by the adapter.

7.1.4.4 Open MinION Mk1B lid & slide the Flongle adapter under the clip.

7.1.4.5 Press down firmly on the adapter to ensure correct thermal and electrical contact.

7.1.4.6 Place the flow cell into the Flongle adapter and press the flow cell down until you hear a click.

7.1.4.7 Perform flow cell QC check prior to the sequencing (Refer to Annexure II for SOP).

7.1.4.8 Peel back the seal tab from the Flongle flow cell, up to a point where the sample port is exposed and pull the seal tab to open access to the sample port.

7.1.4.9 Hold the seal tab open by using adhesive on the tab to stick to the MinION Mk1B lid.

Note: Ensure that there is no air gap in the pipette tip before loading and retain few µl volume in the pipette tip to avoid introduction of bubbles into the sample port. To avoid flushing the flow cell too vigorously, load the priming mix by twisting the pipette plunger down.

7.1.4.10 Place the P200 pipette tip inside the sample port and slowly dispense the priming solution that was prepared into the Flongle flow cell.

Note: Ensure that there is no air gap in the pipette tip before loading and retain few µl volume in the pipette tip to avoid introduction of bubbles into the sample port.

7.1.4.11 Place the P200 tip inside the sample port and slowly dispense the 30 µl sequencing mix into the flow cell by twisting the pipette plunger down.

7.1.4.12 Seal the Flongle flow cell using the adhesive on the seal tab by sticking the transparent adhesive tape to the sample port.

7.1.4.13 Step 13: Replace the top (Wheel icon section) of the seal tab to its original position & close the sequencing platform lid.

7.1.4.14 Step 14: Open the MinKNOW application and click **Start** from the left side panel then click **Start Sequencing** option from the menu to setup and start the sequencing run.

## 8. QUALITY CONTROL

8.1. Only the sample or library that pass the acceptation criteria should be considered for processing.

8.2. Sample should always be stored at -80°C and on ice while processing.

8.3. Final library should be stored at -20°C.

8.4. Optional QC check can be performed after PCR amplification to confirm the expected PCR product (~400 bp).

## 9. BIOHAZARD WASTE DISPOSAL

9.1. All solid waste (tips, gloves, packaging, etc) collected in the specimen processing room should be discarded only in labelled biohazard bags (Labelled as COVID-19 WASTE) inside the biosafety cabinet. Filled Biohazard Bags should be tied inside the biosafety cabinet with tag.

9.2. Removed PPE should be discarded in marked designated bins. Bags should be tied and labelled.

9.3. Tied and labelled biohazard bags should be autoclaved at 121°C and 15 psi for 60 minutes (gravity flow) and 45 minutes in vacuum autoclave.

9.4. Note: Waste containing sodium hypochlorite should never be autoclaved.

9.5. Autoclaved waste should be weighed and clearly labeled as "COVID-19 waste" and handed over to Housekeeping Staff.

9.6. Housekeeping Staff should take the autoclaved waste to designated area for pickup and incineration.

9.7. Any incidents including spills, mechanical breakdowns, failure in bio-containment or any other maintenance problem should be reported immediately to the biosafety officer.

9.8. Any incidence of exposure to personnel should be reported to the officer in charge.

Note: For detailed Biomedical Waste Management guidelines, please refer to the tool kit – "A comprehensive manual on SARS-CoV-2 diagnostics" from the link: https://www.finddx.org/wp-ontent/uploads/2022/08/20220819_usaid_covid_toolkit_FV_EN.pdf.

## 10. REFERENCES


10.1. Marco Cascella., et al. Features, Evaluation, and Treatment of Coronavirus (COVID-19 - StatPearls - NCBI Bookshelf (nih.gov)

10.2. Euro Surveill. 2021;26(24):pii=2100509. https://doi.org/10.2807/1560-7917.ES.2021.26.24.2100509

10.3. Coronavirus Sequencing | NGS for COVID-19 (SARS-CoV-2) identification (Illumina.com)

10.4. PCR tiling of SARS-CoV-2 virus - classic protocol (SQK-LSK109 with EXP-NBD104, EXP-NBD114 or EXP-NBD196. https://community.nanoporetech.com/docs/prepare/library_prep_protocols/classic-pcr-tiling-SARS-CoV-2/v/ptc_9096_v109_revx_06feb2020, requires ONT community access).

10.5. Indian COVID-19 Genome Surveillance. https://clingen.igib.res.in/covid19genomes/

10.6. Pangolin COVID-19 Lineage Assigner. https://pangolin.cog-uk.io/

# SOP-003

# Whole Genome Analysis of SARS-CoV-2 using InterARTIC pipeline

| | LABORATORY STANDARD OPERATING PROCEDURES | | | |
|---|---|---|---|---|
| Country and Lab Demo Site | India | **DOCUMENT NUMBER** | **VERSION** | **EFFECTIVE DATE** |
| | | SOP-003 | 2.3 | |
| **PROCEDURE** | **Whole Genome Analysis of SARS-CoV-2 using InterARTIC pipeline** | | | |

| | | | | |
|---|---|---|---|---|
| Approved By | Name, Title | | Signature | Date |
| | | | | |
| | | | | |
| SOP Annual Review | Name, Title | | Signature | Date |
| | | | | |
| | | | | |
| | | | | |
| Revision History | Version # [0.0] | Revision Date [dd/mm/yy] | Description (notes) | |
| | | | | |
| | | | | |
| | | | | |
| Distributed Copies to | Name (or location) | # of copies | Name (or location) | # of copies |
| | | | | |
| | | | | |
| | | | | |

***I acknowledge that I have read, understand and agree to follow this SOP.***

| Name (print) | Signature | Designation | Date |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

## Table of Contents

# 1. PURPOSE

- This Standard Operating Procedure provides guidance to perform bioinformatics analysis of SARS-CoV-2 genome data generated through MinION nanopore sequencer using interARTIC analysis pipeline.
- This SOP contains instructions to perform basecalling, demultiplexing, adapter trimming and removal, alignment with the reference genome, variant calling and interpretation.

# 2. INTRODUCTION

The SARS-CoV-2 has evolved continuously since its first identification in January 2020. Several variants have since emerged including Alpha (B.1.1.7), Beta (B.1.351), Gamma (P1), Delta (B.1.617.2) and Omicron (B.1.1.529). Since the beginning of the COVID-19 pandemic, multiple variants have shown higher transmissibility, virulence and immune escape resulting in higher rates of infection and subsequent morbidity and mortality globally.

Therefore, It is critical to identify variants early to inform the public health response through genomic surveillance for pandemic control and apply a systematic approach to provide a representative indication of the extent of transmission of SARS-CoV-2 variants based on the local context.

# 3. PRINCIPLE

There are broadly two sections into which the interARTIC analysis can be divided.

1. Genome assembly
2. Variant calling

URL to the interARTIC software and tutorial for reference (https://github.com/Psy-Fer/interARTIC).

Two inbuilt tools namely "Medaka" and "Nanopolish" are employed to create consensus sequences and variant calls from nanopore sequencing data. This task is performed using neural networks applied on pileup of individual sequencing reads against a draft assembly. It provides state-of-the-art results outperforming sequence-graph based methods and signal-based methods, whilst also being faster.

The interARTIC pipeline is a package that combines the above two tools to provide a consensus genome sequence and the variants in the sample with respect to the Wuhan CoV-19 genome.

# 4. PERSONNEL QUALIFICATIONS

## Education and training

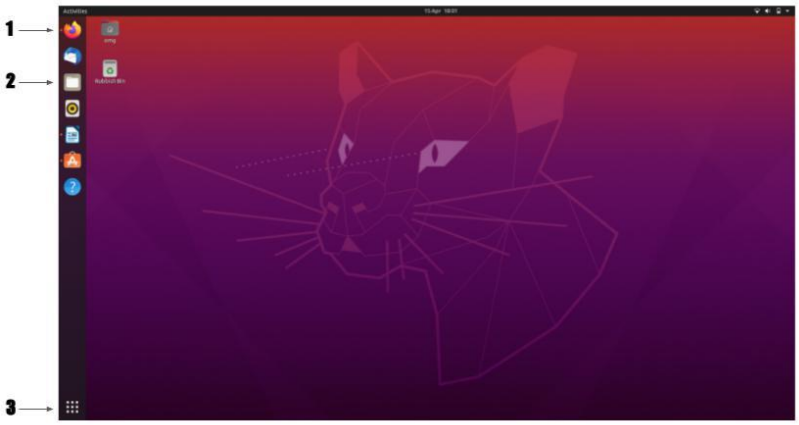
- Education and training must be given on the following topics:
- Potential risks to health (symptoms of COVID-19 disease and transmission).
- Precautions to be taken to minimize droplets & aerosol formation and prevent exposure.
- Hygiene requirements.
- Donning and Doffing of PPE.
- Laboratory biosafety, specifically handling of potentially infectious materials.
- Laboratory design, including airflow conditions.
- Use of biological safety cabinets (operation, identification of malfunctions and maintenance).
- Use of autoclaves, microcentrifuge, micropipettes & refrigerators, (operation, identification of malfunctions and maintenance).
- Prevention of incidents and steps to be taken by workers in the case of incidents (biohazard incidents, chemical, electrical and fire hazards).
- Good laboratory practice and good microbiological techniques.
- Workflow in laboratory .
- Procedure to be performed.
- Waste handling.
- Importance of laboratory results for individual patient and COVID pandemic management.

# 5. RESPONSIBILITIES

It is the responsibility of the lab personnel to correctly understand and perform this procedure. All users of this procedure who do not understand the procedure or unable to carry it out as described are responsible for seeking advice from their supervisor.

# 6. PROCEDURE

## 6.1 Introduction to Ubuntu:



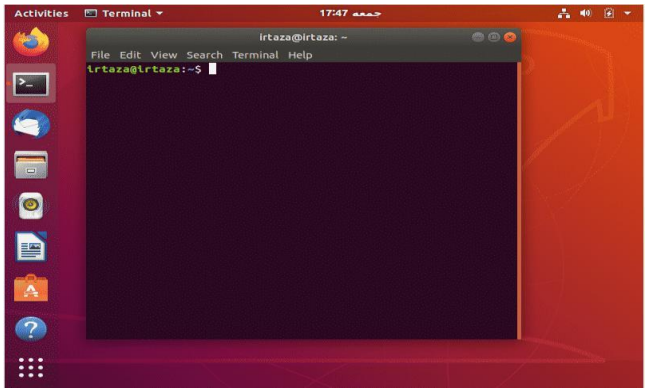

The above image shows the desktop screen of Ubuntu 20.04 LTS.

1. **Browser icon -** To open the browser, place your mouse cursor on the browser icon and **click the left mouse button once.**
2. **File/Folder icon -** To view the files in your system, place your mouse cursor on the file/folder icon and **click the left mouse button once.**

3. **Search icon -** To search for programs/tools/applications in the system, place your mouse cursor on the search icon and **click the left mouse button once.**

## 6.2 Basic file operations in Ubuntu Terminal:

1. We perform all our file operations and analysis using the terminal (Command Line Interface) which looks like the below image:




2. To open the terminal screen, using your keyboard, press hold the keys **Ctrl + Alt + T** one by one. The terminal screen similar to the above image pops open.

3. We can type commands into the terminal screen only when a blinking white rectangle appears immediately after the **$** symbol.

4. **Naming a file/folder can include underscores (_) and hyphens (-) if the name is larger than one word. Names should not have spaces.**

   **Eg: File_1_15_06_2022 or File1_15-06-2022**

5. Some of the basic commands to work with the terminal are listed below:

   a) To know the current directory/folder you're working in, type '**pwd**' and press ENTER. This step lists the path of the directory you're currently in.

   b) To know the list of files and folders in your current working directory, type '**ls**' and press ENTER.

   c) To create a new directory/folder, type '**mkdir name_of_directory**' and press ENTER.

      name_of_directory - any directory name that you wish to navigate.

   d) To move/navigate into a folder/directory, type '**cd name_of_directory**' and press ENTER.

   e) To make a copy of a file, type '**cp old_file_name new_file_name**' and press ENTER.

   f) To move a file from the current directory to another directory, type '**mv file path_of_preferred_directory**' and press ENTER.

6. To get the path of the preferred directory, type 'cd' to return to the home directory or open a new terminal screen. From the home directory, go to your preferred directory using step d) and once inside the directory, do step a). Copy the displayed path and proceed with step f).

   g) To REMOVE FILES, type '**rm file_name**' and press ENTER.

   h) To REMOVE FOLDERS, type '**rm -r folder_name**' and press ENTER.

## 6.3 Working as a superuser in Ubuntu:

Certain files and folders in ubuntu are password protected and maintained securely by default. Those files/folders can be accessed by having admin/superuser privileges. To access and work with those files, we use the term '**sudo**' as a prefix in every command.

**Eg;** To copy a file inside admin access directories,
type '**sudo cp old_file_name new_file_name**'.

**Note:**

Once you type the command with **sudo** and press ENTER, the terminal screen will request you for password. Type your system admin password.

Your password will not be visible when you type, but after typing the system password and pressing ENTER, the command will work.

In the terminal, to navigate to the folders/directories which have superuser access, we type, '**cd /name_of_directory**'.

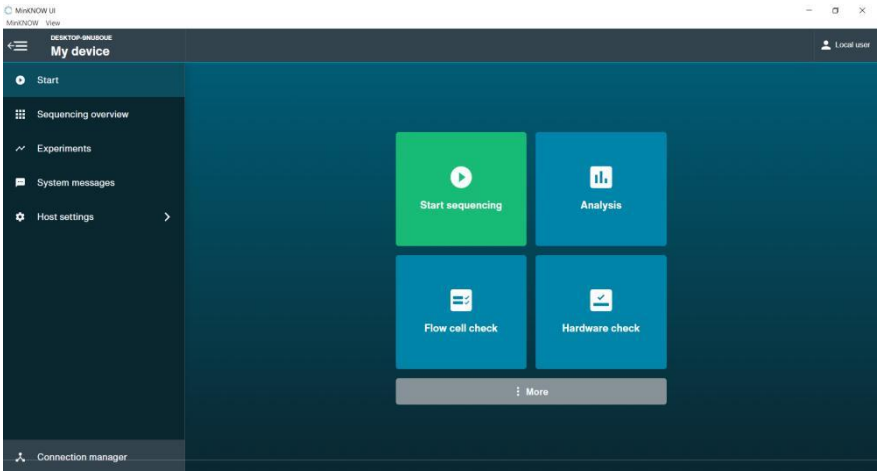
**Note:** Here, we use a **Forward slash (/)** before typing the name of the directory.

## 6.4 Basecalling & Demultiplexing - The MinKNOW UI:

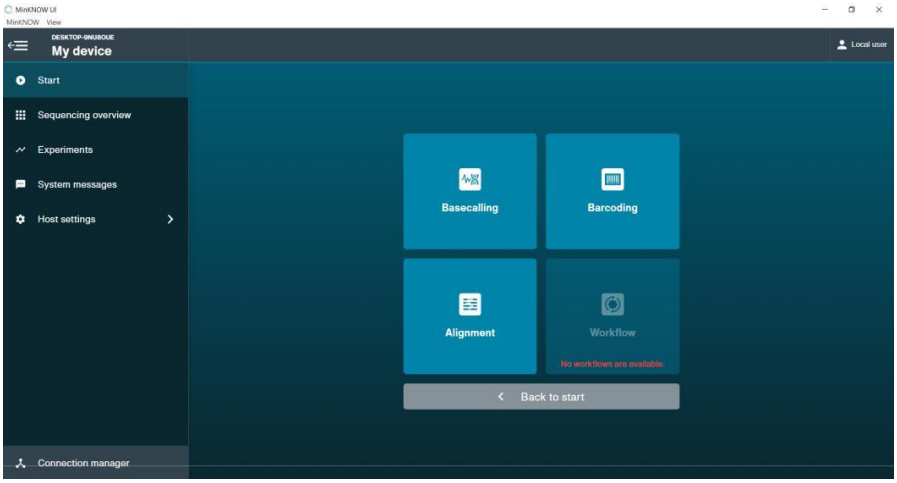
All the sequenced FAST5 files will be stored in the **/var/lib/minknow/data** directory which we had set up while running the sequencing for our samples. This is the default output directory for all the sequencing runs. To perform basecalling, we use the FAST5 files as input in MinKNOW.

6.4.1 In Ubuntu Desktop, find the following **MinKNOW icon**:

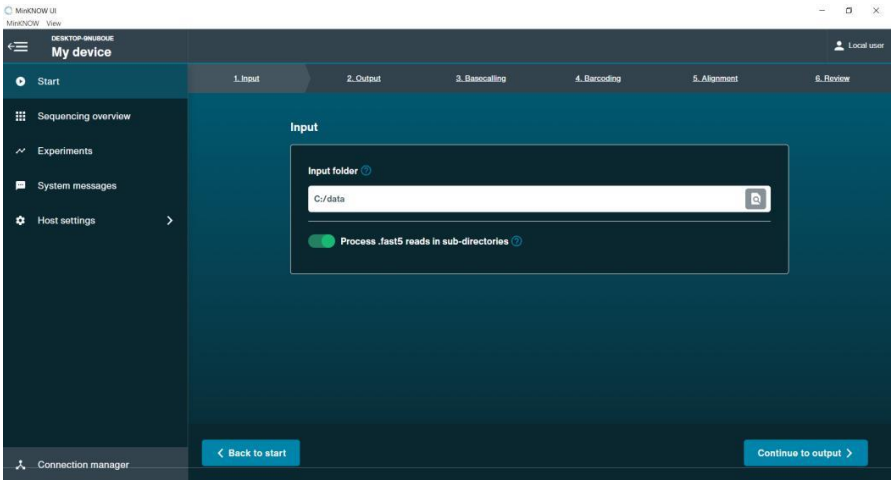
6.4.2 Left click the icon once to open **MinKNOW**. Click **Start** from the left side panel and the following screen will appear:




6.4.3 Step 3: Click **Analysis**, then click **Basecalling**.




6.4.4 The following screen appears.



The above screen has 6 simple steps to be set to do basecalling and further steps.

Before proceeding to the steps, do the following:

6.4.5 Open a terminal screen (Ctrl + Alt + T), type **cd /var/lib/minknow/data** and press **ENTER**.

6.4.6 You are now inside the directory, where the input FAST5 data files are present inside the sample data directory. Now you have to provide accessing permission to use the sample directory.

So, type **"sudo chmod 777 -R sample_directory_name".**

And press ENTER.

6.4.7 Step 6: Type your **admin/system password**.

6.4.8 Step 7: Next, type '**ls**' and press ENTER. You should see a color change for your sample_directory, which means the folder can now be accessed for analysis.

## Step 1: Setting the input directory:

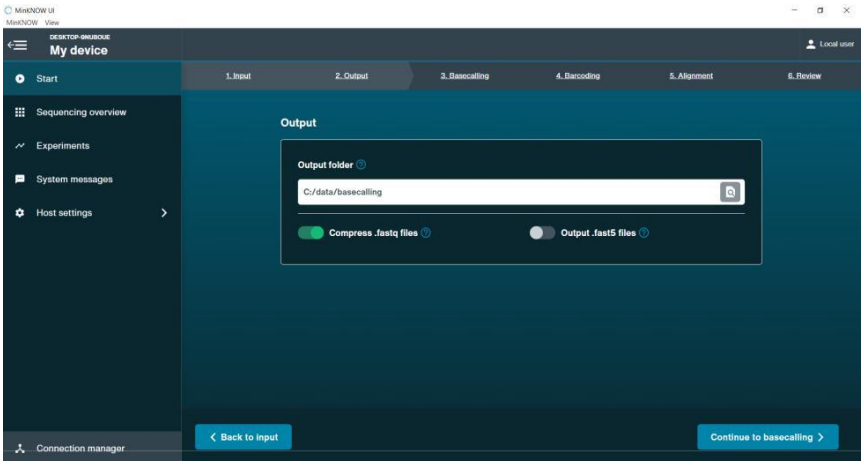
6.4.9 In the Input screen in MinKNOW, click the white space and select the "**/var/lib/minknow/data/sample_directory".**

**Note:** Here, sample_directory is the directory where your sequenced FAST5 sample files are present.

6.4.10 After selecting the input directory, click **Continue** to output button seen in the lower right corner of MinKNOW screen.

Now, the input directory is set successfully.

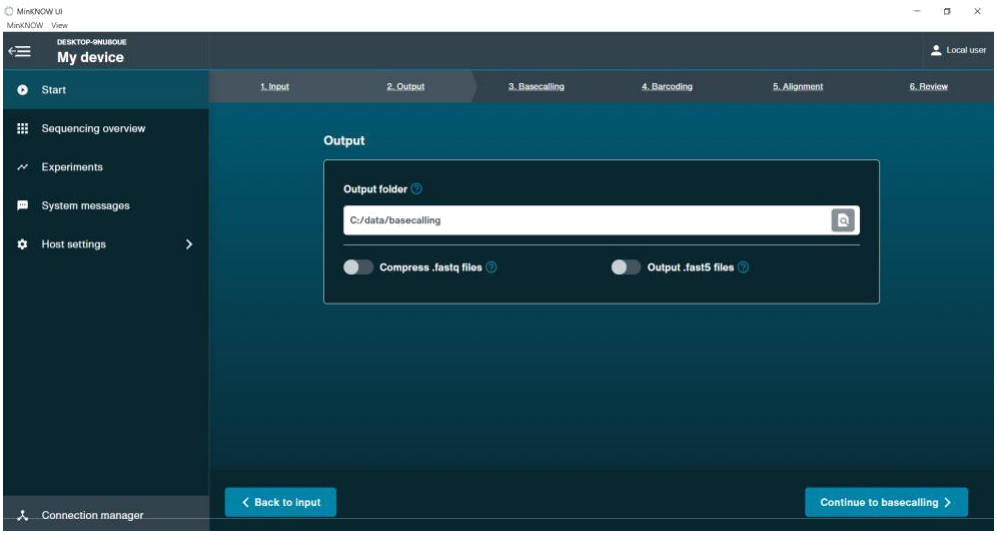
## Step 2: The output directory:




6.4.11 In this step, the output directory is automatically set under the **/var/lib/minknow/data/basecalling** directory.

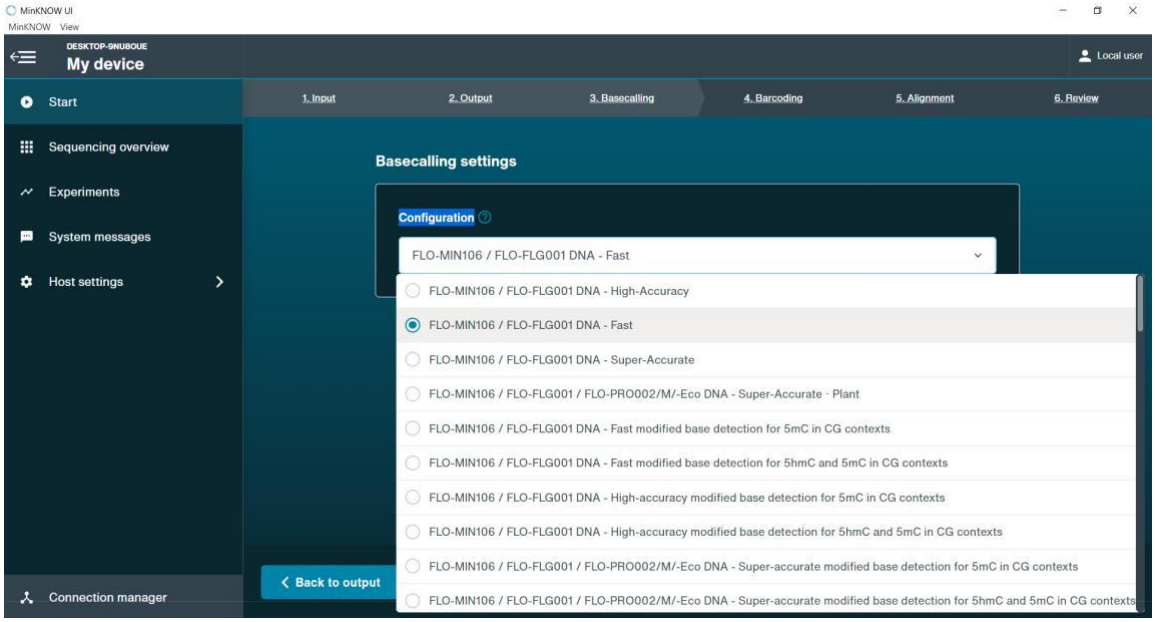
6.4.12 After the basecalling step, all output files and folders will be stored under a folder named '**basecalling'**.

6.4.13 In this step, turn off the green color toggle button which says, '**Compress .fastq** files'.

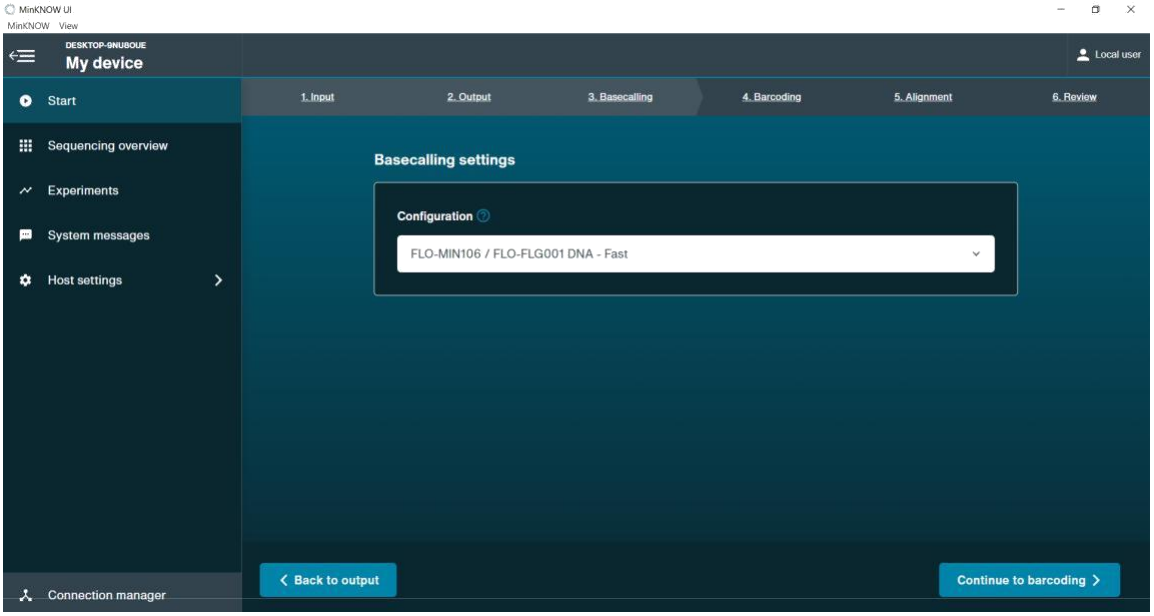



6.4.14 Click the **Continue to basecalling** button in the lower right corner of MinKNOW screen.

## Step 3: Basecalling settings:

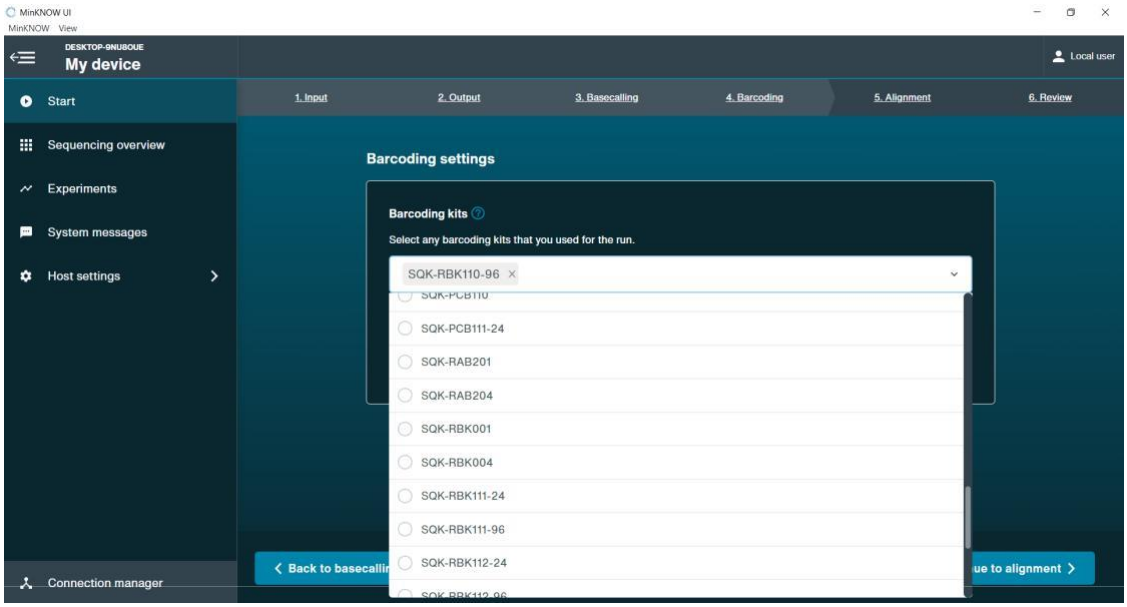



6.4.15 Here, click to display a dropdown. From the dropdown list, select the option, "**FLO-MIN 106/FLO-FLG001 DNA – FAST".**

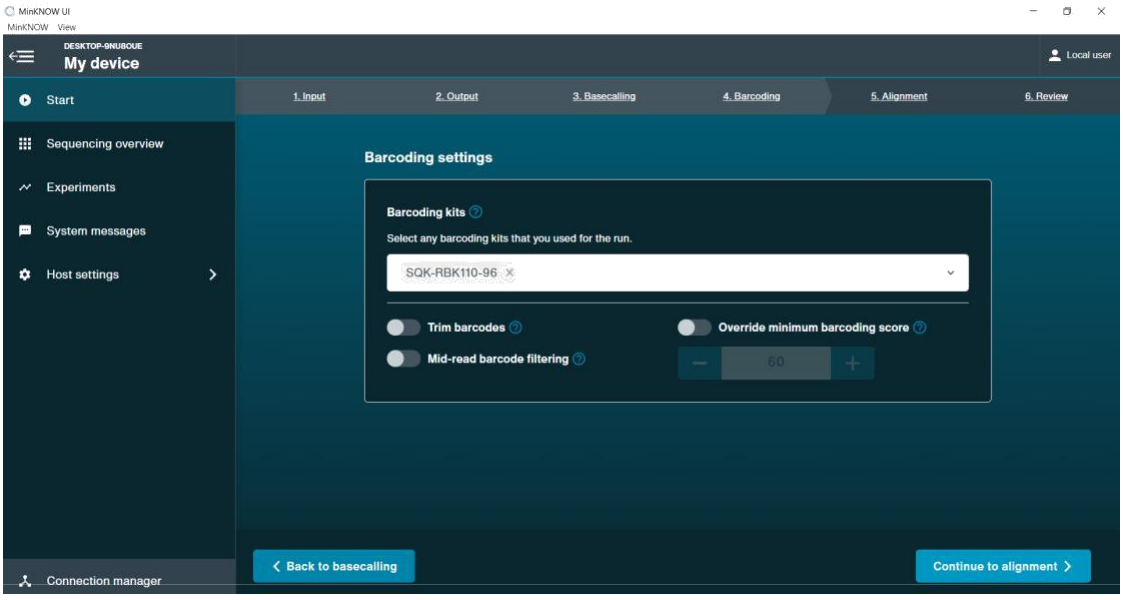



6.4.16 Click **Continue to barcoding** button in the lower right corner to proceed to barcoding/demultiplexing step.

## Step 4: Selecting Barcoding kit:



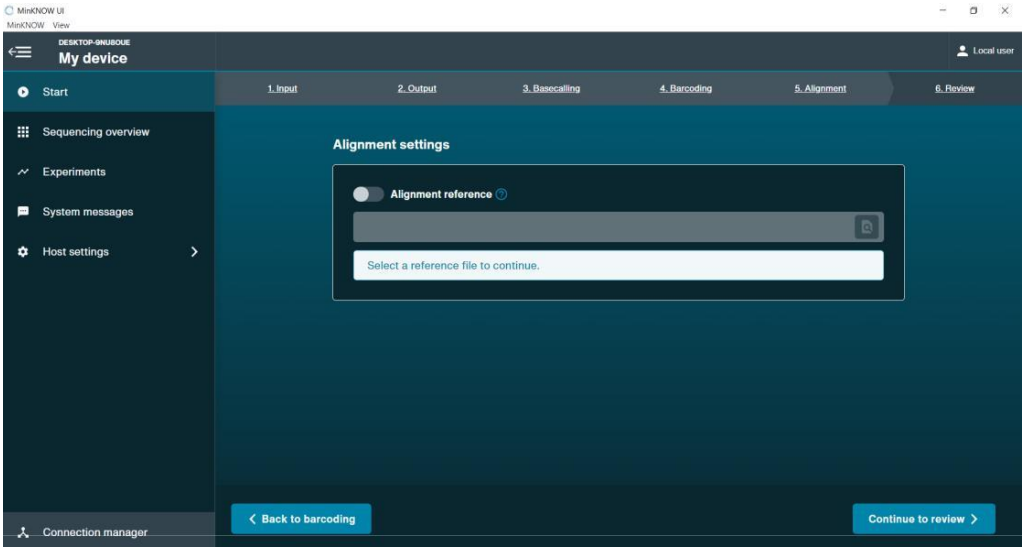

6.4.17 Click on the option to display dropdown menu. From the menu, select the appropriate barcode kit used for the library preparation (eg: SQK-RBK 110-96 or SQK-LSK-110 with an expansion kit of EXP-NBD-104/114).




6.4.18 Leave all other options as default and click on **Continue to alignment** button at the lower right corner.

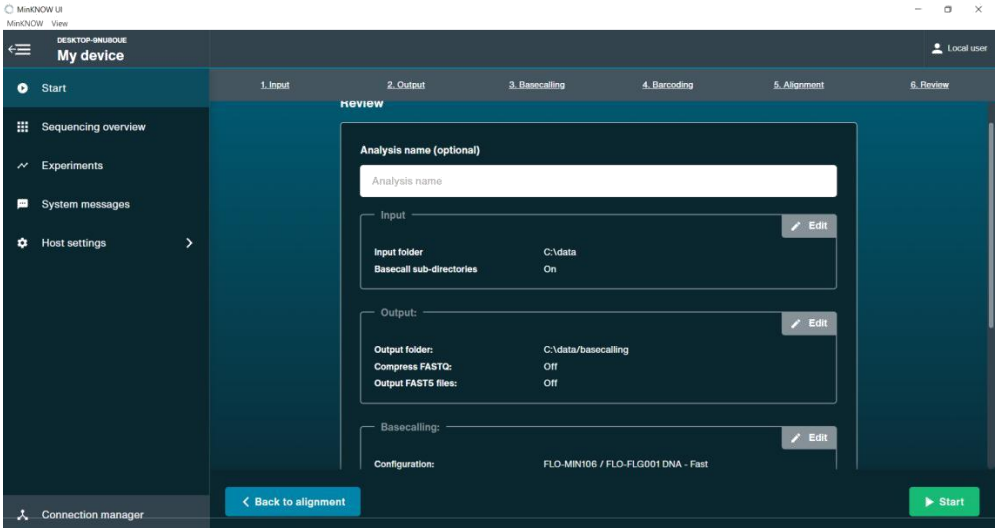
## Step 5: Alignment:

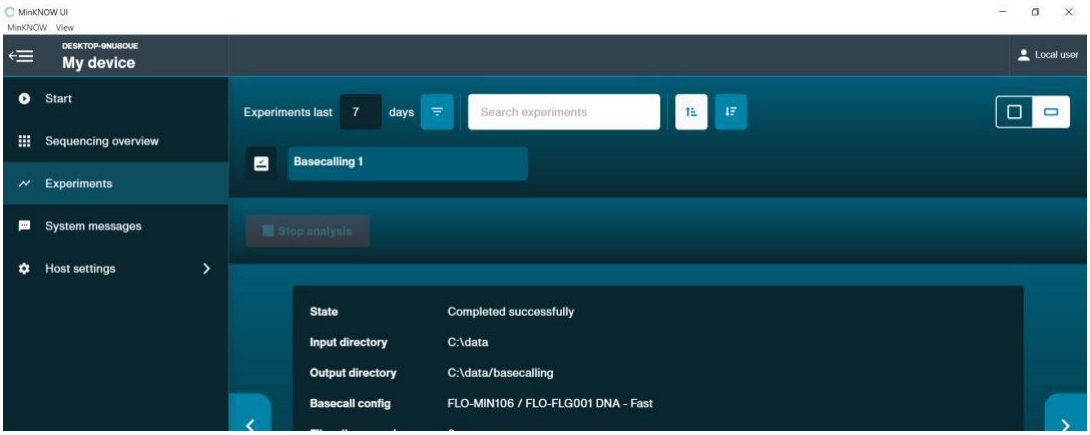
6.4.19 In this step, we don't have to set any option and click "**continue to review**" button in the lower right corner.




## Step 6: Reviewing our settings:

6.4.20 In this step, we can check the settings we had set just before proceeding to basecalling. and click **Start** in the lower right corner to begin analysis.






6.4.21 After clicking start, our basecalling analysis job will begin. Once the job is complete, we get the status of the job displayed as seen in the above image.

## 6.5 Preparing input files for further analysis using output from MinKNOW:

From the above step, the basecalling and demultiplexing steps are successfully completed and all our output files are saved in the **/var/lib/minknow/data/basecalling** directory.

The next step, we have to create several folders/directories for analysis.

The folders and files have to be created by following the steps below:

6.5.1 Go to Ubuntu Desktop.

6.5.2 Place your mouse cursor on the screen and click on the mouse right button once. Select **New folder** option to create a new folder in Desktop.

6.5.3 Give a name to the folder.

6.5.4 This folder will be your parent directory.

6.5.5 Now, click that parent directory 2 times using your mouse cursor. The folder will open which should be empty.

6.5.6 Now inside the folder, right click your mouse button again. Again, select **New folder** option to create another new folder. Name this folder as "**runid".**

6.5.7 Create another new second folder inside parent directory. Name this folder as "**sample_barcodes".**

After this step, the parent directory should have 2 folders namely, **'runid'** and **'sample_barcodes'.**

6.5.8 Now, enter into **runid** folder by clicking it twice. Here, create another new folder and name it as **runid_subdir** again.

6.5.9 Go into this 2nd **runid** folder again and create another new folder and name it as "**runid_subdir".**

6.5.10 Go into this **runid_subdir2** folder and create two folders and name them as "**fast_q"** and **"fast_5".**

6.5.11 Using Terminal, Go to the **/var/lib/minknow/data/sample_directory** and copy all your **.fast5** sample files and paste it into the **fast_5** folder created just now in step 10). (Use command **sudo cp *.fast5 /path_of_fast_5_folder_in_Desktop/fast_5/** ).

6.5.12 In the same way, using terminal, Goto the **/var/lib/minknow/data/basecalling/pass** directory and copy all the barcoding folders corresponding to your samples.

6.5.13 Paste the folders into the **fast_q** folder created in step 10).

6.5.14 If the command **sudo cp** did not work for copying folders, use **"sudo cp -r".**

6.5.15 After this step, in lower left corner of your desktop, click the following icon:

6.5.16 After this, type **libreoffice calc**

6.5.17 This icon will appear. Click on it to open an excel sheet.

6.5.18 6.5.18 Fill the first column with sample data (name of sample) and second column with barcode data.

6.5.19 To save the excel sheet, Move your mouse cursor to the top left corner of screen.

6.5.20 Click **File** → **Save as** → A popup will appear. Select the file destination by clicking **Desktop** → **parent directory** → **sample_barcodes** on the left side panel. After selecting file destination, type file name as **sample_barcodes**.

6.5.21 Move your mouse cursor to the lower end of the **Save as** popup. Place your cursor by dragging without clicking in a semi-freeze dropdown. Continue dragging the cursor until you find an option, **.csv**

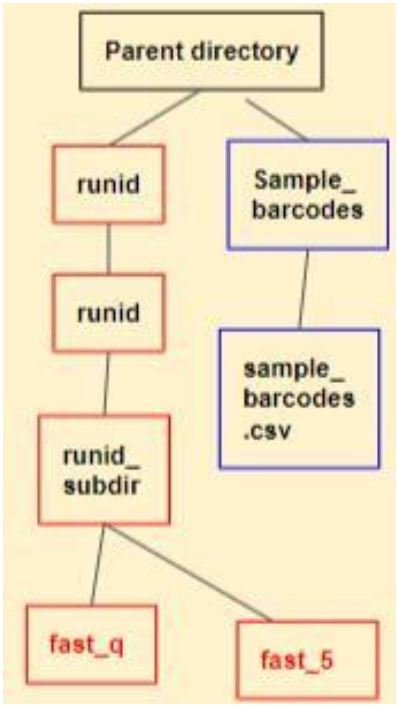
6.5.22 Select that **.csv** option.

6.5.23 Click **Save.**

6.5.24 Now, the **sample_barcodes.csv** is saved.

The file preparation step is completed.

The overall file/folder creation layout should look something like this:




We just completed preparation of input files for analysis to use with interARTIC pipeline.

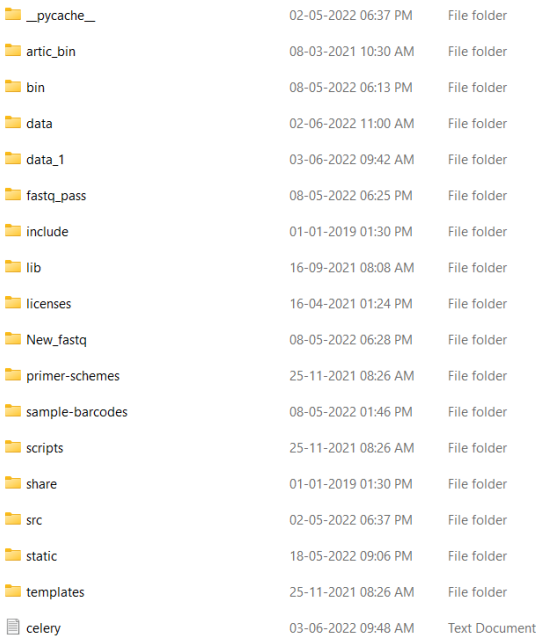
## 6.6 interARTIC Usage

### 6.6.1 Initialising the software

6.6.1.1 Click on the folder icon marked 2 in the below image and when the folder opens, enter the "home" option.

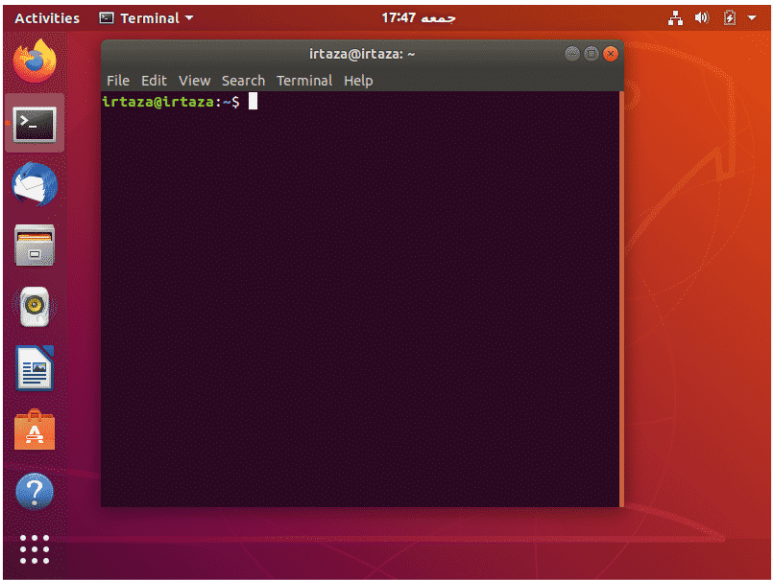

6.6.1.2 Go to the folder **interARTIC** and you will see a folder named "interartic_bin".

6.6.1.3 Inside **interartic_bin** you will see the following files as represented in the image below.




6.6.1.4 Amongst all the files and folders, identify a file named "**run.sh**".

6.6.1.5 Go to you Linux command line by clicking the icon shown below.




and type **./run.sh** (Image below).

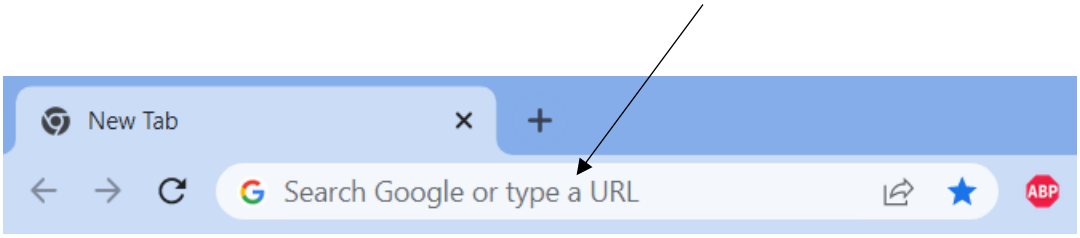
```
root@DESKTOP-3MOGA03:/mnt/d/FINDD/software/interartic_bin# ./run.sh
```

6.6.1.6 The following output will be visible (Image below).

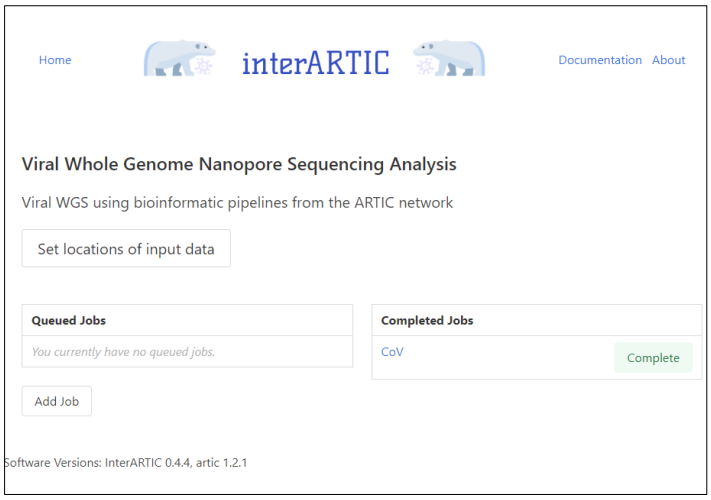
```
InterARTIC is now running on your machine :)
To launch InterARTIC web interface visit http://127.0.0.1:5000 on your browser
To keep your InterARTIC active this terminal must remain open.
To terminate InterARTIC type CTRL-C or close the terminal.
```

6.6.1.7 Copy the IP address show here as "**127.0.0.1:5000**".

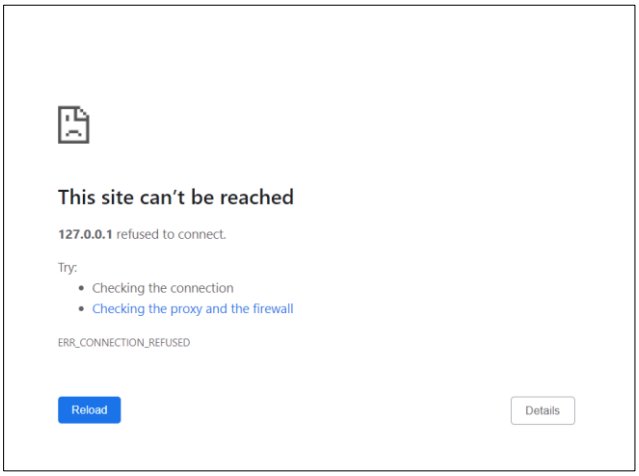
6.6.1.8 After copying, paste it in the address tab of your google chrome browser.




6.6.1.9 Wait for few seconds and the following page will open.

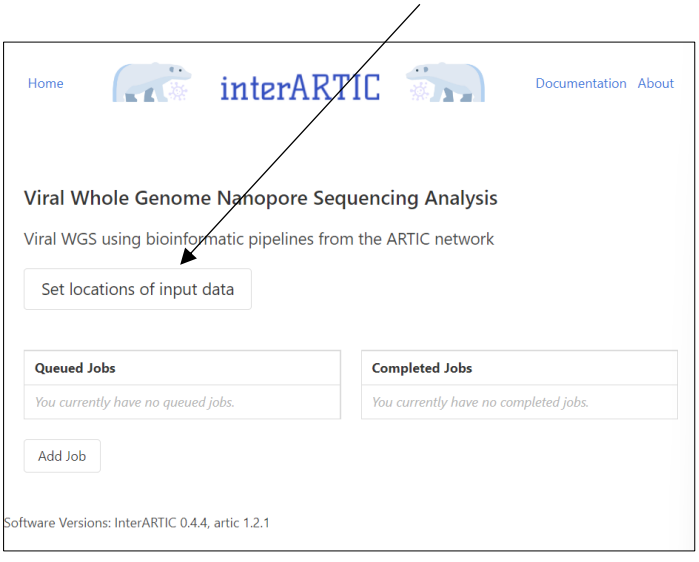



Note: Sometimes, it may show like the below given image. In that case, please check the interARTIC terminal is running and reload the page.




## 6.6.2 Setting up the directory

6.6.2.1 On your web page, click on the option "Set locations of input data".

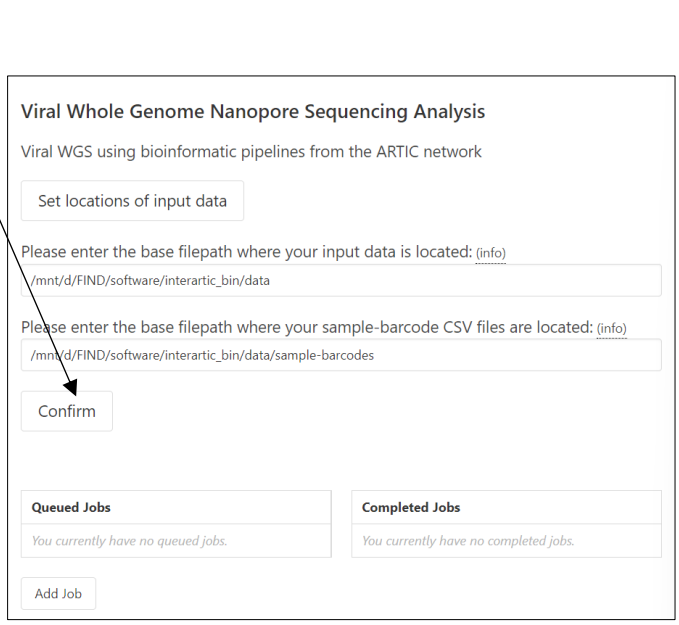


6.6.2.2 In the first field, type the location of the directory you have created for analysis. **"/mnt/d/FINDD/software/interartic_bin/data".**

**The above folder name "data" is the name given to the "parent_directory" (refer page 10).**

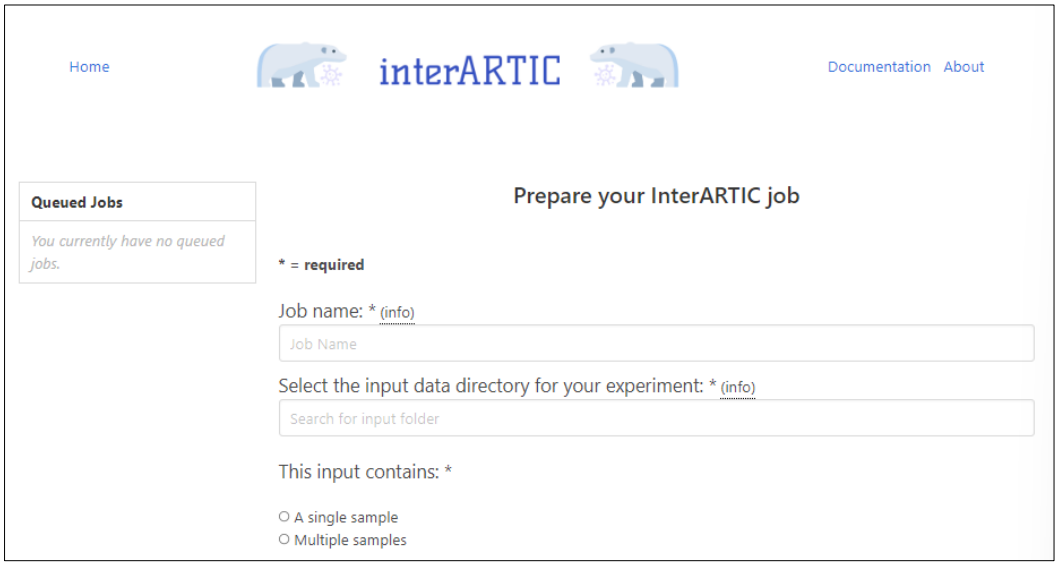
6.6.2.3 In the second field, enter: "**/mnt/d/FINDD/software/interartic_bin/data/sample-barcodes**".

6.6.2.4 Click **Confirm**.




## 6.6.3 Adding Job

6.6.3.1 You should be on the interARTIC home page. If not, click "**Home**" on the top left corner.



6.6.3.2 Click on "**Add Job**".

6.6.3.3 Fill out the page using the following details:

a. Job name: **Analysis_1**

b. Select the input data directory for your experiment: **Folder_1**

c. This input contains: **Multiple samples**

d. Output folder: "leave blank"

e. Select the virus you want to analyze: **SARS-CoV-2**

f. Select your primer scheme: **ARTICV3**

Note: The options "Primer scheme top directory" & "Name of primer scheme" will be filled automatically once you give ARTICV3.

g. Leave the Demultiplexing option unchecked

h. Which library preparation method was used: **Ligation library prep**

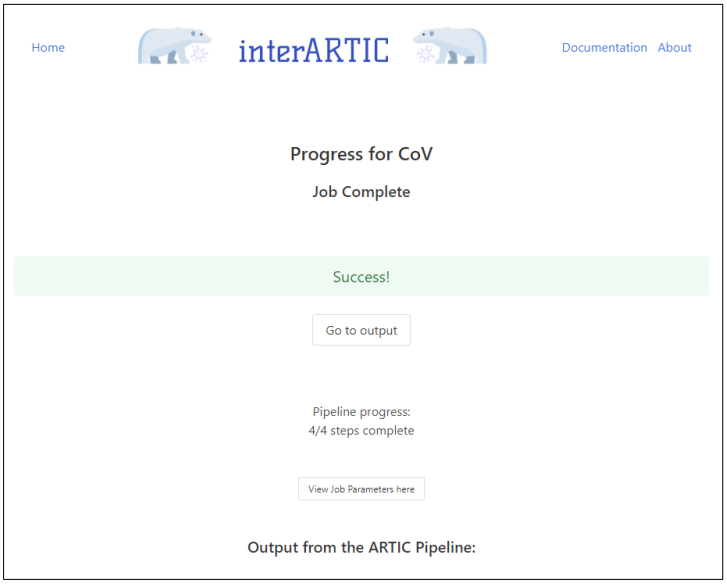
i. Select a pipeline to run: Both

6.6.3.4 Hit Submit Job(s)

The run will take a while depending on the number of samples and the size of your input files. The following page will appear.

### 6.6.4 Analysis

You will receive the following page after the analysis is over.




6.6.4.1 Hit the button "**Go to output**".

6.6.4.2 On the new page hit the "**View**" button.

6.6.4.3 You will reach the data visualization page.

6.6.4.4 You can go through the following outputs in detail for the selected sample.

a. FastQC Metrics

b. Coverage plot & Variant information of the sample

c. SARS-CoV-2 genome based metrics

## 6.7. Interpretation

**Further interpretation (or variant annotation) can be done using two online tools namely:**

6.7.1 **Pango (https://pangolin.cog-uk.io/).**

6.7.1.1 Go to the above link.

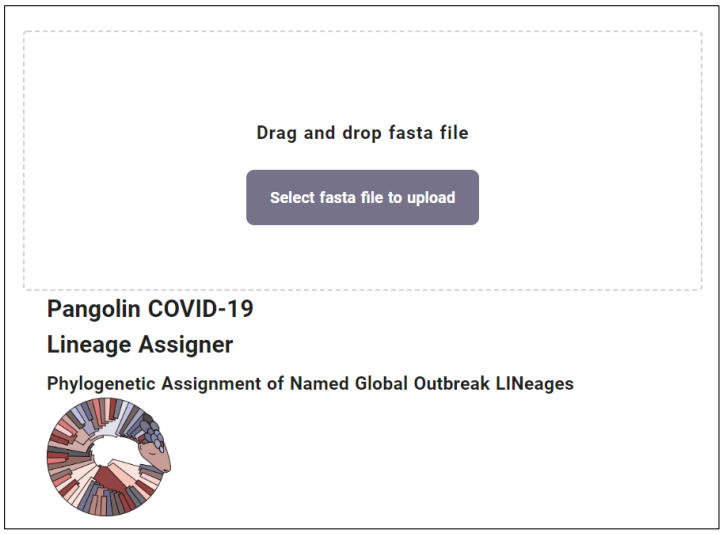
6.7.1.2 From your output folder that is produced from the interARTIC analysis, Go to the folder of interest (Medaka or Nanopolish) then to the sample specific folder.

6.7.1.3 Open the folder and search for the file that ends with the name "**.consensus.fasta**".

6.7.1.4 This consensus file should used as input file for Pango to annotate the variants.

6.7.1.5 Come back to the web page of Pango.

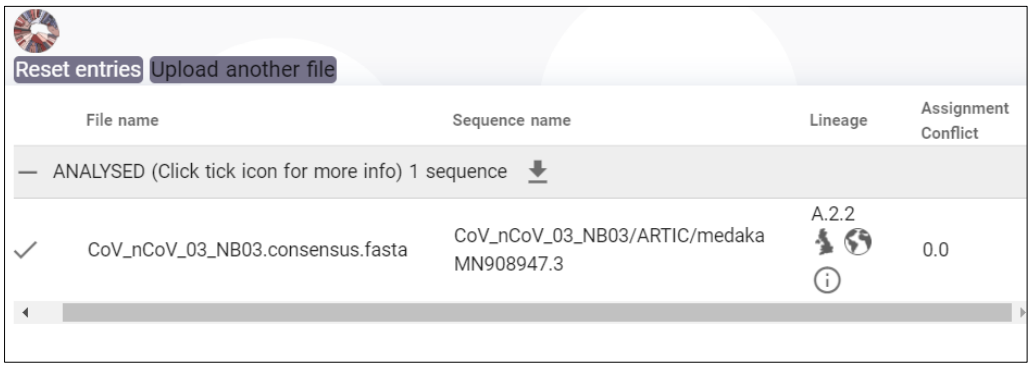
6.7.1.6 Drag and drop or select the above file into the page which looks like the image shown below.




6.7.1.7 After importing, a new page will open with the option "**Start analysis**" on left top corner.

6.7.1.8 Click on it and wait for the run to finish.

6.7.1.9 The following page will appear.



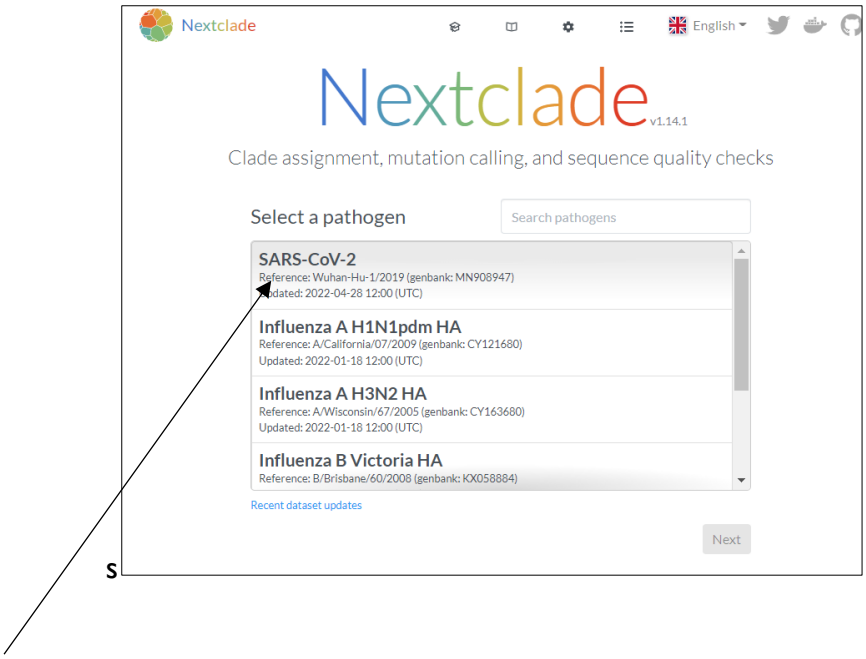

6.7.1.10 Click on the two symbols and see the output separately.

6.7.1.11 The file will take some time to load.

6.7.1.12 Analyze the output on the web page further.

### 6.7.2 **Nextclade (https://clades.nextstrain.org/ )**

6.7.2.1 Open the above link and you will see the following web page.

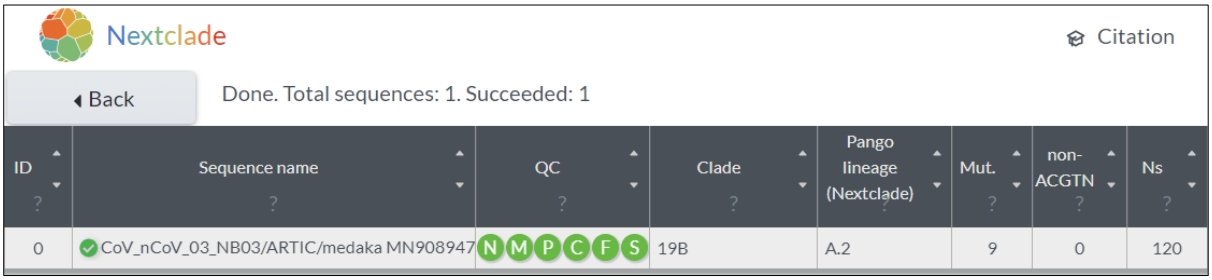



6.7.2.2 Click on the option SARS-CoV-2.

6.7.2.3 Similar to the above tool drag and drop the same file ("**.consensus.fasta**") to the window.

6.7.2.4 Click the button "**Run**".

6.7.2.5 A new webpage will open. Wait for the below results to appear once the run is over.




6.7.2.6 You can analyze further based on your requirement and study.

# 7 References


7.7.1 Post-run basecalling. https://community.nanoporetech.com/docs/prepare/library_prep_protocols/experiment-companion-minknow/v/mke_1013_v1_revcq_11apr2016/post-run-basecalling, requires ONT community access.

7.7.2 Ferguson JM, Gamaarachchi H, Nguyen T, Gollon A, Tong S, Aquilina-Reid C, Bowen-James R, Deveson IW. InterARTIC: an interactive web application for whole-genome nanopore sequencing analysis of SARS-CoV-2 and other viruses. Bioinformatics. 2022 Feb 7;38(5):1443-1446. doi: 10.1093/bioinformatics/btab846. PMID: 34908106; PMCID: PMC8826086.

7.7.3 6.8.3 Aksamentov, I., Roemer, C., Hodcroft, E. B., & Neher, R. A., (2021). Nextclade: clade assignment, mutation calling and quality control for viral genomes. Journal of Open Source Software, 6(67), 3773, https://doi.org/10.21105/joss.03773 (https://clades.nextstrain.org).

7.7.4 6.8.4 Indian COVID-19 Genome Surveillance. https://clingen.igib.res.in/covid19genomes/

7.7.5 6.8.5 Pangolin COVID-19 Lineage Assigner. https://pangolin.cog-uk.io/

# SOP-004

# Flushing, Reloading and Storing of Mk I B Flow cell

| | LABORATORY STANDARD OPERATING PROCEDURES | | | |
|---|---|---|---|---|
| Country and Lab Demo Site | India | **DOCUMENT NUMBER** | **VERSION** | **EFFECTIVE DATE** |
| | | SOP-004 | 1.0 | |
| **PROCEDURE** | **Flushing, Reloading and Storing of Mk1B flow cell** | | | |

| | | | | |
|---|---|---|---|---|
| Approved By | Name, Title | | Signature | Date |
| | | | | |
| | | | | |
| SOP Annual Review | Name, Title | | Signature | Date |
| | | | | |
| | | | | |
| | | | | |
| Revision History | Version # [0.0] | Revision Date [dd/mm/yy] | Description (notes) | |
| | | | | |
| | | | | |
| | | | | |
| Distributed Copies to | Name (or location) | # of copies | Name (or location) | # of copies |
| | | | | |
| | | | | |
| | | | | |

***I acknowledge that I have read, understand and agree to follow this SOP.***

| Name (print) | Signature | Designation | Date |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

# Table of Contents

## Objective

To perform washing and reuse of Mk1b flow cell on Nanopore (Mk1b) platform.

## Principle

The Wash Kit allows sequential runs of multiple sequencing libraries on the same flow cell. It works by washing out the first library, and refreshing the system ready for a subsequent library to be loaded. This procedure provides the opportunity to utilise the same flow cell a number of times, maximising the available run time, particularly for cases where less data per library is required. Following the wash step, Storage Buffer can be introduced into the flow cell, allowing storage of the flow cell before subsequent library additions.

Although the wash procedure should remove 99.9% of the library, some residual DNA may remain on the flow cell. For this reason, users may prefer to barcode their libraries when used in conjunction with the Wash Kit, such that reads from different libraries can be separated from each other. Successful deconvolution of DNA reads has been demonstrated in Oxford Nanopore's internal development.

## Procedure

**Materials Required:**

- Flow Cell Wash Kit (EXP-WSH004)

Flow Cell Wash Kit contents (EXP-WSH004):


WMX : Wash Mix
DIL : Wash Diluent
S : Storage Buffer

| Contents | Volume (µl) | No. of tubes | No. of uses |
|---|---|---|---|
| Wash Mix (WMX) | 15 | 1 | 6 |
| Wash Diluent (DIL) | 1,300 | 2 | 6 |
| Storage Buffer (S) | 1,600 | 2 | 6 |

- Wash Mix (WMX) contains DNase I.
- Wash Diluent (DIL) contains the exonuclease buffer that maximises activity of the DNase I.
- The Storage Buffer allows flow cells to be stored for extended periods of time.

## Flushing a MinION Flow Cell

Preparation to run the washing procedure.

- This protocol assumes that the flow cell has already had a DNA/RNA library run on it
- The aim is to remove most of this initial library and prepare the flow cell for the loading of a subsequent library
- The Wash Kit contains all solutions required for removal of the initial library
- Data acquisition in MinKNOW should be stopped (if loading a new library or storing the flow cell), or paused (if loading more of the same library after the wash) during the wash procedure and also during subsequent library addition
- After the flow cell has been washed, a new library can be loaded or the flow cell can be stored at 4°C

**Step 1:** Place the tube of Wash Mix (WMX) on ice. Do not vortex the tube.

**Step 2:** Thaw one tube of Wash Diluent (DIL) at room temperature.

**Step 3:** Mix the contents of Wash Diluent (DIL) thoroughly by vortexing, spin down briefly and place on ice.

**Step 4:** In a clean 1.5 ml Eppendorf tube, prepare the following Flow Cell Wash Mix:

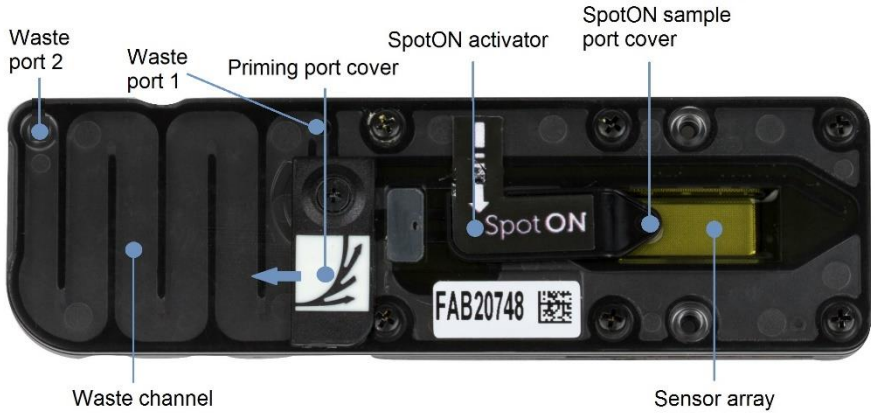
| Component | Volume |
|---|---|
| Wash Mix (WMX) | 2 µl |
| Wash Diluent (DIL) | 398 µl |
| **Total** | **400 µl** |

**Step 5:** Mix well by pipetting, and place on ice. Do not vortex the tube.

**Step 6:** Stop or pause the sequencing experiment in MinKNOW, and leave the flow cell in the device.

**Step 7:** Ensure that the priming port cover and SpotON sample port cover are in the positions indicated in the figure below.

**Step 8:** Using a P1000, remove all fluid from the waste channel through Waste port 1. As both the priming port and SpotON sample port are closed, no fluid should leave the sensor array area.




Note: It is vital that the flow cell priming port and SpotON sample port are closed to prevent air from being drawn across the sensor array area, which would lead to a significant loss of sequencing channels.

**Step 9:** Rotate the flow cell priming port cover clockwise so that the priming port is visible.

**Step 10:** Check for air between the priming port and the sensor array. If necessary, using a P1000 draw back a maximum of 30 µl volume to remove any air:

1. Set a P1000 pipette to 200 µl
2. Insert the tip into the priming port
3. Turn the wheel until the dial shows 220-230 µl, or until you can see a small volume of buffer/liquid entering the pipette tip.
4. Visually check that there is continuous buffer from the priming port across the sensor array.

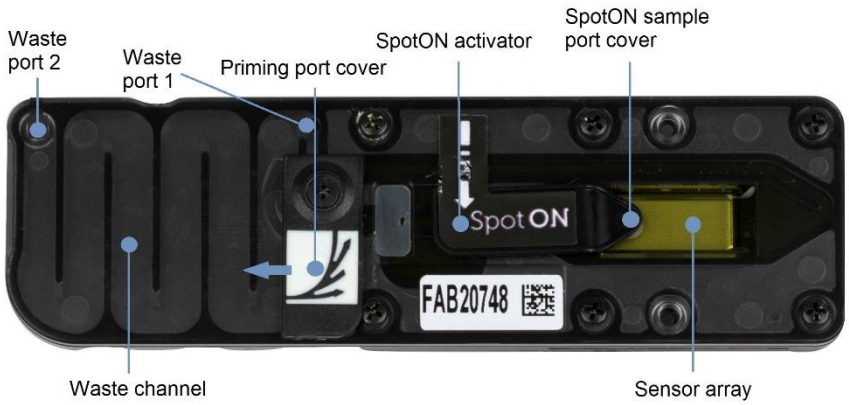
Take care when drawing back buffer from the flow cell. Do not remove more than 20-30 µl, and make sure that the array of pores are covered by buffer at all times. Introducing air bubbles into the array can irreversibly damage pores.

**Step 11:** Load 400 µl of the prepared Flow Cell Wash Mix into the flow cell via the priming port, avoiding the introduction of air.

**Step 12:** Close the priming port and wait for 60 minutes.

**Step 13:** Ensure that the priming port cover and SpotON sample port cover are in the positions indicated in the figure below.

**Step 14:** Using a P1000, remove all fluid from the waste channel through Waste port 1. As both the priming port and SpotON sample port are closed, no fluid should leave the sensor array area.




Note: It is vital that the flow cell priming port and SpotON sample port are closed to prevent air from being drawn across the sensor array area, which would lead to a significant loss of sequencing channels.

**Follow one of the two options described in the next steps of the protocol:**

## To run a second library on a MinION flow cell straight away

**Note:** The buffers used in this process are incompatible with conducting a Flow Cell Check step prior to loading the subsequent library. In order to check your flow cell, follow the instructions in the next section "To store the MinION flow cell for later use" before priming and loading the flow cell.

**Step 1:** To run a second library straight away, follow the instructions in the "**Priming and loading the flow cell**" section of the relevant protocol.

Note: As part of this process the flow cell will need priming using the Flow Cell Priming Kit.

Once the flow cell has been primed and loaded, either resume the run in MinKNOW or start a new sequencing experiment.

### Reloading a library

Additional buffers for reloading a diluted library, following the washing of a flow cell, can be found in the one of the following expansion kits:

- Sequencing Auxiliary Vials expansion (EXP-AUX001). This expansion contains vials of Elution Buffer (EB), Sequencing Buffer (SQB) and Loading Beads (LB), additional to those found in DNA sequencing kits prior to '110' chemistry.
- Sequencing Auxiliary Vials expansion (EXP-AUX002). This expansion contains vials of Sequencing Buffer II (SBII), Elution Buffer (EB), Loading Solution (LS) and Loading Beads II (LBII).

## To store the MinION flow cell for later use

Materials: Flow Cell Wash Kit (EXP-WSH004)

**Step 1:** Thaw one tube of Storage Buffer (S) at room temperature.

**Step 2:** Mix contents thoroughly by pipetting and spin down briefly.

**Step 3:** Rotate the flow cell priming port cover clockwise so that the priming port is visible.

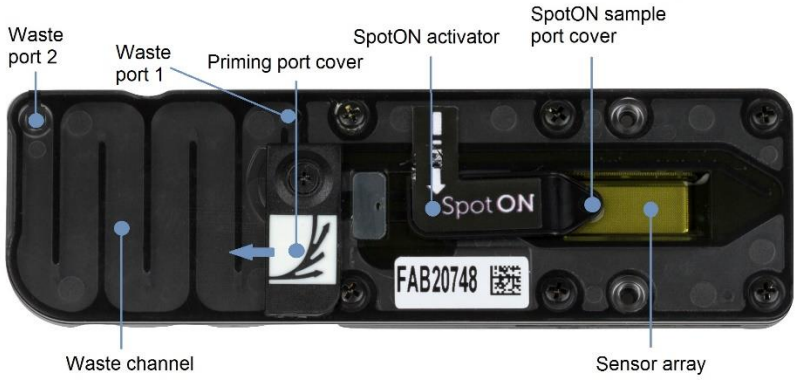
**Step 4:** Check for air between the priming port and the sensor array. If necessary, using a P1000 draw back a small volume to remove any air (a few µls):

1. Set a P1000 pipette to 200 µl
2. Insert the tip into the priming port
3. Turn the wheel until the dial shows 220-230 µl, or until you can see a small volume of buffer/liquid entering the pipette tip.
4. Visually check that there is continuous buffer from the priming port across the sensor array.

**Step 5:** Slowly add 500 µl of Storage Buffer (S) through the priming port of the flow cell.

**Step 6:** Close the priming port.

**Step 7:** Using a P1000, remove all fluid from the waste channel through Waste port 1. As both the priming port and SpotON sample port are closed, no fluid should leave the sensor array area.




Note: It is vital that the flow cell priming port and SpotON sample port are closed to prevent air from being drawn across the sensor array area, which would lead to a significant loss of sequencing channels.

**Step 8:** The flow cell can now be stored at 4-8°C.

**Step 9:** When you wish to reuse the flow cell, remove the flow cell from storage, and allow it to warm to room temperature for ~5 minutes. You will need to perform a Flow Cell Check before loading the next library.

## Refuelling your flow cell

### General advice

Refuelling is the replenishment of motor fuel in the sequencing experiment through the addition of Flush Buffer (FB) from the Flow Cell Priming Kit (EXP-FLP002). The translocation speed graph in MinKNOW can indicate whether it is necessary to top up fuel. Volumes and processes of refuelling are specific to the flow cell type. Please follow the instructions relevant to your flow cell type e.g. Flongle & MinION.

### When to refuel

If the DNA translocation speed drops below 300 bases per second, you may start to see a reduction in quality of data reflected in the Qscore. We therefore recommend topping up the flow cell with fuel, using the Flush Buffer (FB) from the Flow Cell Priming Kit. Please follow the instructions below if you wish to top up the fuel during an experiment.

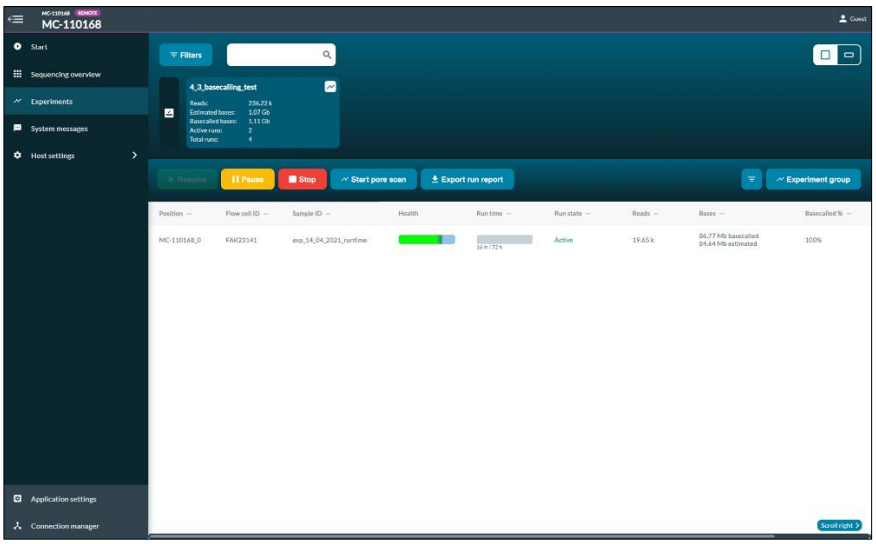
### Refuelling a Flongle flow cell

- Remove one tube of FB from the freezer and thaw by bringing to room temperature
- Pause the experiment on the MinION Mk 1B Flow Cell that is being refuelled:

  a) Navigate to Experiments and open the experiment running

  b) Click Pause

  c) A dialogue box will open. Choose the flow cell to pause and click Pause to confirm

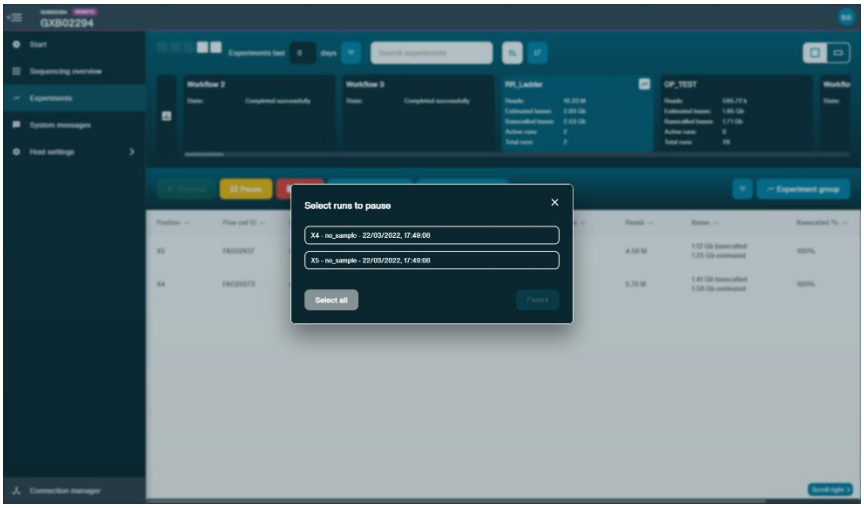



- Pull back the seal tab to reveal the sample port
- Pass 30 µl of FB through sample port

- Reseal the tab, ensuring the drain ports are covered
- Unpause the experiment on the relevant GridION position/on the MinION Mk 1B by clicking Resume
- (optional) Click Start pore scan to pick a new set of the best channels for the remainder of the sequencing experiment

## Refuelling a MinION flow cell

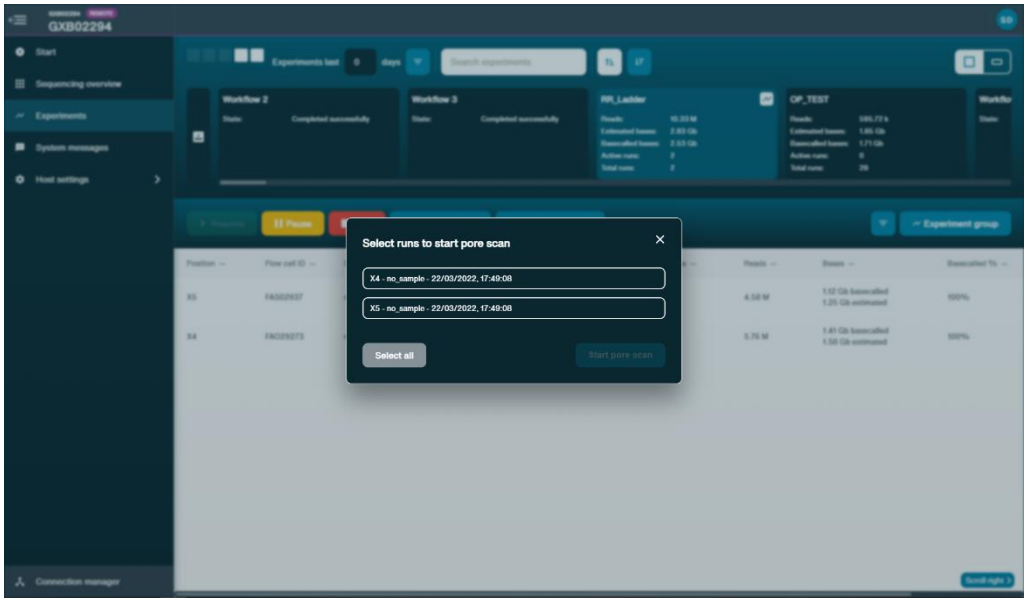
- Remove one tube of Flush Buffer (FB) from the freezer and thaw by bringing to room temperature
- Pause the experiment on the MinION Mk1B Flow Cell that is being refuelled:
    a) Navigate to Experiments and select the experiment running
    b) Click Pause
    c) A dialogue box will open. Choose the flow cell to pause and click Pause to confirm




- After opening the priming port, check for a small bubble under the cover. Draw back a small volume to remove the bubble (a few µl):
    a) Set a P1000 pipette to 200 µl
    b) Insert the tip into the priming port
    c) Turn the wheel until the dial shows 220-230 µl, or until you can see a small volume of buffer entering the pipette tip
- Visually check that there is continuous buffer from the priming port across the sensor array.

## Complete the flow cell refuelling

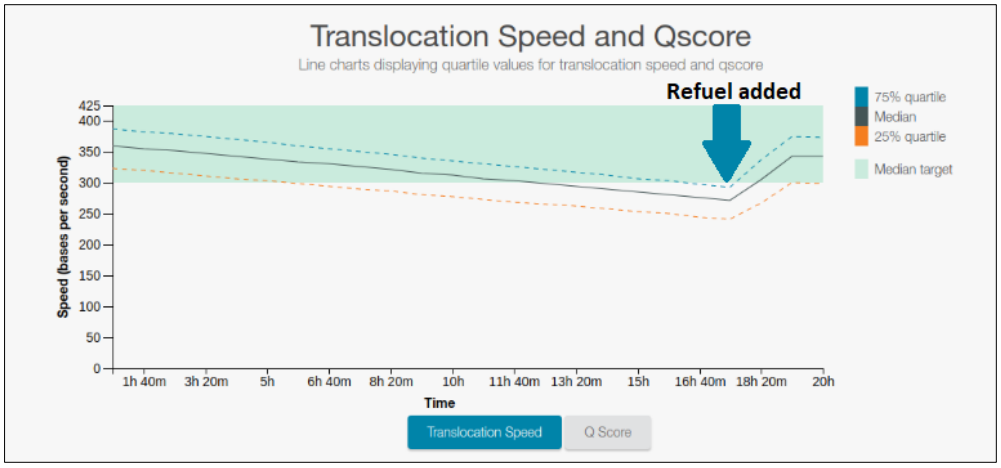
- Load 250 µl of the FB into the flow cell via the priming port (i.e. not the SpotON port), avoiding the introduction of air bubbles
- Close the priming port and replace the MinION Mk1B/GridION lid
- Unpause the experiment on the relevant GridION position/on the MinION Mk1B by clicking Resume
- (optional) Click Start pore scan and choose your flow cell, to pick a new set of the best channels for the remainder of the sequencing experiment




## The outcome and benefits of refuelling

### Translocation speed and Qscore over time

Below is a graph that shows what is expected for translocation speed after the addition of FB to the flow cell in a previous version of MinKNOW.



### Speed of bases through nanopores before and after refuelling

As the speed drops below 300 bases per second, the Qscore will begin to decline for the reads processed through the nanopores at this speed. After refuelling at the 17.5 hour mark, the speed begins to increase and returns to an improved rate (~400 bases per second), which is similar to the speed at the start of the experiment. After the addition of fuel using FB, the quality of the data may increase and return to Qscores equivalent to those seen at the start of your run.

### Refuelling multiple times in a run

You can refuel a sequencing run multiple times over an experiment. When you should refuel will depend on when the translocation speed drops below 300 bases per second on the speed graph in the MinKNOW GUI for Kit 9 (e.g. SQK-LSK109). Kit 12 (e.g. SQK-LSK112) adapter contains fuel fix technology, enabling users to run long experiments without the need for fuel addition.

### Warning - overloading the flow cell when refuelling

If you refuel your flow cell multiple times, the waste reservoir of the flow cell will gradually fill up with buffer.

The array and waste reservoir of the below platforms (Flongle & MinION Mk 1B) will take the following fluid loads:

- ✓ a prime
- ✓ a library
- ✓ one refuel - for Flongle
- ✓ three refuels - for MinION/GridION

## References


1. Flushing a MinION/GridION Flow Cell. https://community.nanoporetech.com/docs/prepare/library_prep_protocols/flow-cell-wash-kit-xl-exp-wsh004-xl/v/wfcx_9146_v1_revh_01dec2021/flushing-a-minion-gridion-flow-cell?devices=minion, requires ONT community access.

# Annexures: I – V

# ANNEXURES

## 1. Annexure 1: Fluorometric Quantification of dsDNA

### 11.1 Introduction:

The Qubit dsDNA HS (High Sensitivity) Assay Kits make DNA quantitation easy and accurate. The kit includes concentrated assay reagent, dilution buffer, and prediluted DNA standards. The assay is highly selective for double-stranded DNA (dsDNA) and is accurate for initial sample concentrations from 0.5ng/µL to 100 ng/µL.

### 11.2 Materials:

1) Qubit assay tubes (500 tubes, Life Technologies, Cat. no. Q32856) or Axygen PCR-05-C tubes (VWR, part no. 10011-830).

2) 15 or 50 ml Centrifuge tubes for aliquoting dilution buffer

### 11.3 Storage Recommendation:

The Qubit dsDNA HS Reagent and Buffer are designed for room temperature storage and the HS reagent (orange color solution) is light sensitive & should be stored in dark environment.

### 11.4 Procedure:

**General note: while carrying out the experiment do not expose the HS reagent or the qubit working solution to direct light.**

11.4.1 Set up the required number of 0.5-mL tubes for standards and samples. The Qubit dsDNA HS Assay requires 2 standards.

11.4.2 Label the tube lids as Std 1, Std 2 & samples

11.4.3 Each standard tube requires 190 µL (189 µL of buffer & 1 µL of HS reagent) of Qubit working solution, and each sample tube requires 199 µL (198 µL of buffer & 1 µL of HS reagent). The final volume in each tube must be 200 µL.

For example, for 8 samples, prepare enough working solution for the samples and 2 standards: ~200 µL per tube in 10 tubes yields 2 mL of working solution (10 µL of Qubit reagent plus 1990 µL of Qubit buffer).

11.4.4 Add 10 µL of each qubit standard to the appropriate tube containing 190 µL of Qubit working solution, then mix by vortexing 2–3 seconds. Be careful not to create bubbles.

11.4.5 Add 1 µL of each sample to the assay tubes containing 199 µL of Qubit working solution, then mix by vortexing 2–3 seconds.

11.4.6 Allow all tubes to incubate at room temperature for 2 minutes.

11.4.7 Proceed to "Reading standards and samples"; follow the procedure appropriate for your instrument.

11.4.8 On the Home screen of the Qubit 4.0 Fluorometer, press DNA, then select dsDNA High Sensitivity as the assay type. The "Read standards" screen is displayed. Press Read Standards to proceed.

11.4.9 Note: If you have already performed a calibration for the selected assay, the instrument prompts you to choose between reading new standards and running samples using the previous calibration. If you want to use the previous calibration, skip to step 12. Otherwise, continue with step 9.

11.4.10 Insert the tube containing Standard #1 into the sample chamber, close the lid, then press Read standard. When the reading is complete (~3 seconds), remove Standard #1.

11.4.11 Insert the tube containing Standard #2 into the sample chamber, close the lid, then press Read standard. When the reading is complete, remove Standard #2.

11.4.12 The instrument displays the results on the Read standard screen. To verify the calibration, quantify standard 2 as described below. Standard 2 concentration should be between 98 – 100 ng for 10 µL. If not, repeat the calibration.

11.4.13 Press Run samples.

11.4.14 On the assay screen, select the sample volume and units:

a. Press the + or – buttons on the wheel to select the sample volume added to the assay tube (from 1–20 µL).

b. From the dropdown menu, select the units for the output sample concentration.

11.4.15 Insert a sample tube into the sample chamber, close the lid, then press Read tube. When the reading is complete (~3 seconds), remove the sample tube. The instrument displays the results on the assay screen. The top value (in large font) is the concentration of the original sample. The bottom value is the dilution concentration.

11.4.16 Repeat step 2.6 until all samples have been read.

## 11.5 Reference:

1. https://tools.thermofisher.com/content/sfs/manuals/Qubit_dsDNA_HS_Assay_UG.pdf

## 2. Annexure II: Flongle/Flow cell QC Check:

This protocol is used to check the number of pores in your flow cell.

Consumables: Flongle device - flow cell and adapter or Mk1B flow cell

Equipment: MinION with a host computer connected to the Internet with MinKNOW installed.

### 12.1 Procedure:

Note: if Mk1B flow cell is used, skip step 1 to 3, and proceed directly with step 4.

12.1.1 Place the Flongle adapter into the MinION and press down firmly on the adapter to ensure correct thermal and electrical contact.

12.1.2 The adapter needs to be plugged into your device, and the device should be plugged in and powered on before inserting the Flongle flow cell.

12.1.3 Place the flow cell into the Flongle adapter and press the flow cell down until you hear a click.

12.1.4 Double-click the MinKNOW icon located on the desktop to open the MinKNOW GUI.

12.1.5 To log in, you must be connected to the internet.

12.1.6 The sequencing overview should display on the screen indicating that the flow cell did not had any checks carried out earlier.

12.1.7 Navigate to the start homepage and select 'Flow Cell Check'.

12.1.8 Fill in the flow cell ID (only for flongle) and select the flow cell type from the drop-down menu.

12.1.9 Select 'Start' to begin the flow cell check and this will automatically get navigated to the sequencing overview page.

12.1.10 The flow cell check should take a few minutes and after completion, green tick should appear on the flongle image in the sequencing overview page.

### 12.2 Reference:

1. https://community.nanoporetech.com/docs/prepare/library_prep_protocols/experiment-companion-minknow/v/mke_1013_v1_revcl_11apr2016/flow-cell-check

## 3. Annexure III: Consumables & Reagent requirements for SARS-CoV-2 Nanopore sequencing workflow

14.1 Equipment/consumables Required for processing <24 samples:

Room No 1: Master Mix Preparation

1. 8 well PCR Strips or individual 0.1 ml PCR tubes
2. 1.5 or 2.0 ml microcentrifuge tubes
3. 0.5 - 10 µl or 2 - 20 ul Pipette (single channel) & filter tips
4. 20 – 200 µl Pipette (single channel) & filter tips
5. -20°C PCR cooler & racks (96 well & 1.5 or 2.0 ml boxes) (transferrable between rooms)
6. Ice Bucket (transferrable between rooms)

Room No 2: Library Preparation (Template Addition Room if available can be utilized for this protocol)

1. 8 well PCR Strips or individual 0.1 ml PCR tubes
2. 1.5 or 2.0 ml micro centrifuge tubes
3. 0.5 ml thin walled, clear PCR tubes
4. 15 or 50 ml centrifuge tubes (for aliquoting the qubit buffers)
5. 1.5- or 2.0 ml magnetic Stand
6. 1.5 or 2.0 ml tube spinner/minifuge/centrifuge
7. 8 well PCR strip spinner
8. Qubit Instrument
9. Thermal cycler (PCR machine) or RT-PCR instrument
10. 0.5 - 10 µl or 2 - 20 µl Pipette (single channel) & filter tips
11. 20 – 200 µl Pipette (single channel) & filter tips
12. 100 – 1000 µl Pipette (single channel) & filter tips
13. 96 well Plate adhesive seal or aluminum foil
14. scalpel blade or scissor
15. Vortex mixer
16. Digital Timer
17. -20°C PCR cooler & racks (96 well & 1.5 or 2.0 ml boxes) (transferrable between rooms)

18. Ice Bucket (transferrable between rooms)

Room No 3: Sequencing Room

1. 1.5 or 2.0 ml microcentrifuge tubes
2. 20 – 200 µl Pipette (single channel) & filter tips
3. 100 – 1000 ul Pipette (single channel) & filter tips
4. 1.5 or 2.0 ml tube spinner/minifuge/centrifugation
5. Vortex mixer
6. Ice Bucket (transferrable between rooms)

14.2 Equipment/consumables Required for processing >24 samples:

Room No 1: Master Mix Preparation

1. 8 well PCR strips
2. 96 well PCR Plate
3. 1.5 or 2.0 ml microcentrifuge tubes
4. 20 – 200 µl Pipette (single channel) & filter tips
5. 0.5 – 10 µl Pipette (multi-channel) & filter tips
6. 8 well PCR strip spinner
7. Ice Bucket (transferrable between rooms)
8. -20°C PCR cooler & racks (96 well & 1.5 or 2.0 ml boxes) (transferrable between rooms)

Room No 2: Library Preparation (Template Addition Room if available can be utilized for this protocol)

1. 1.5 or 2.0 ml micro centrifuge tubes
2. 15 or 50 ml Centrifuge tubes
3. 0.5 ml thin walled, clear PCR tubes
4. 0.5 - 10 µl or 2 - 20 µl Pipette (single channel) & filter tips
5. 0.5 – 10 µl pipette (multi-channel) & filter tips
6. 10 – 100 µl pipette (multi-channel) & filter tips
7. 20 – 200 µl Pipette (single channel) & Filter tips
8. 100 – 1000 µl Pipette (single channel) & filter tips

9. 1.5 or 2.0 ml Magnetic Stand

10. 1.5 or 2.0 ml tube spinner/minifuge/centrifugation

11. Qubit Instrument

12. Thermal cycler (PCR machine) or RT-PCR instrument

13. 96-well plate spinner or plate centrifuge

14. Vortex mixer

15. -20°C PCR cooler & racks (96 well & 1.5 or 2.0 ml boxes) (transferrable between rooms)

16. 96 well Plate adhesive seal or aluminum foil

17. scalpel blade or scissor

18. Ice Bucket (transferrable between rooms)

19. Digital Timer or stopwatch

Room No 3: Sequencing Room

1. 1.5 or 2.0 ml microcentrifuge tubes

2. 20 – 200 µl Pipette (single channel) & filter tips

3. 100 – 1000 µl Pipette (single channel) & filter tips

4. 1.5 or 2.0 ml tube spinner/minifuge/centrifuge

5. Vortex mixer

6. Ice Bucket (transferrable between rooms)

Reagents Required:

Room No 1: Master Mix Preparation

1. LunaScript RT SuperMix Reagent

2. Nuclease-free water

3. SARS CoV2 ARTIC/Midnight Primer Vials A & B

4. Q5 Hot Start HF 2X Master Mix

Room No 2: Library Preparation (Template Addition Room if available can be utilized for this protocol)

1. Ligation Sequencing Kit

2. Nuclease free water

3. SPRI Beads
4. 80% Ethanol (freshly prepared)
5. Elution Buffer
6. NEBNext Companion Module for Oxford Nanopore Technologies Ligation Sequencing (E7180S)
7. Qubit dsDNA HS reagents
8. NEBNext Quick Ligation Module
9. Native Barcoding Expansion 1 – 12 or 13-24 or Rapid Barcoding 96 Kit

Room No 3: Sequencing Room

1. Sequencing buffer II
2. Loading Beads II or loading Solution
3. Flush Buffer
4. Flush Tether
5. Flongle/Minion flow cell*
6. *Flow cell wash kit (required only for flow cell)

## 4. Annexure IV: Recommended storage conditions of the reagents/consumables used in the process

### 4.1 For Native barcoding based ARTIC Protocol:

| Reagents | Recommended Storage |
|---|---|
| LunaScript RT SuperMix Kit | -20°C |
| Q5 Hot Start High-Fidelity 2X Master Mix | -20°C |
| COVID ARTIC WGS Primer Pools A & B | -20°C |
| Ligation Sequencing Kit | -20°C |
| Native Barcoding Expansion 1-12 | -20°C |
| Native Barcoding Expansion 13-24 | -20°C |
| NEBNext Companion Module for Oxford Nanopore Technologies Ligation Sequencing | -20°C |
| NEBNext Quick Ligation Module | -20°C |
| Flongle/Minion flow cell* | 2 - 8°C |
| Flow cell wash Kit (required only for flow cell) | -20°C |
| SPRI Beads | 2 - 8°C |
| Nuclease-free water | 2 - 8°C |
| 80% Ethanol | Room Temperature |
| Qubit dsDNA HS Assay Kit | Room Temperature |

### 4.2 For Rapid barcoding based Midnight Protocol:

| Reagents | Recommended Storage |
|---|---|
| LunaScript RT SuperMix Kit | -20°C |
| Q5 Hot Start High-Fidelity 2X Master Mix | -20°C |
| COVID Midnight Primer Pools A & B | -20°C |
| Ligation Sequencing Kit | -20°C |
| Rapid Barcoding Kit 96 | -20°C |
| Flongle/Minion flow cell* | 2 - 8°C |
| *Flow cell wash Kit (required only for flow cell) | -20°C |
| SPRI Beads | 2 - 8°C |
| Nuclease-free water | 2 - 8°C |
| 80% Ethanol | Room Temperature |
| Qubit dsDNA HS Assay Kit | Room Temperature |

## Annexure V: Sample processing sheet template

| | | | SARS-CoV-2 WGS library prepration & Sequencing<br>Sample Processing worksheet | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Country and Lab Demo Site | | | Document No. | | Version | | Effective Date | | Page # |
| India | | | | | 1.0 | | | | 1 of 1 |
| S.No | Sample ID/Name | Ct Value | RNA QC P/F | RT PCR Well Id | PCR - Pool A Well ID | PCR - Pool B Well ID | Barcode Name | Library Name | Remarks |
| 1 | | | P | | | | | | |
| 2 | | | P | | | | | | |
| 3 | | | P | | | | | | |
| 4 | | | P | | | | | | |
| 5 | | | P | | | | | | |
| 6 | | | P | | | | | | |
| 7 | | | P | | | | | | |
| 8 | | | P | | | | | | |
| 9 | | | P | | | | | | |
| 10 | | | P | | | | | | |
| 11 | | | P | | | | | | |
| 12 | | | P | | | | | | |
| 13 | | | P | | | | | | |
| 14 | | | P | | | | | | |
| 15 | | | P | | | | | | |
| 16 | | | P | | | | | | |
| 17 | | | P | | | | | | |
| 18 | | | P | | | | | | |
| 19 | | | P | | | | | | |
| 20 | | | P | | | | | | |
| 21 | | | P | | | | | | |
| 22 | | | P | | | | | | |
| 23 | | | P | | | | | | |
| 24 | | | P | | | | | | |

| Processed | Checked | Verified | Date | Comments |
|---|---|---|---|---|
| Samples are added to the correct well ID/ strip | | | | |
| PCR- Pool A & Pool B Master mixes prepared correctly | | | | |
| Corresponding Pool A & B products are pooled & barcode was added correctly | | | | |
| RT-PCR Machine Used | | | | |
| PCR Machine Used | | | | |
| Barcoded libraires pooled | | | | |
| | | | | |
| Sample Processing Worksheet | Version#1.0 | | | Page 1 of 1 |

# Chapter 4: Training Evaluation

**This chapter provides details on training evaluation for the master trainers.**

- Pre/Post Training Quiz
- Trainees Feedback form

# Pre/Post Training Quiz

**Instructions for the trainer:**

The trainer/instructor will administer the test before the training and collect the response

The same test will be given to the trainee after the completion of the training

The difference in scores will be recorded for training evaluation purpose

**Trainee Name:** **Date:**

**Designation:**

**Test taken before the training (yes/no):**

**Test taken after the training (yes/no):**

**Instructions for the trainee:**

The test quiz carries …25… questions

Attempt all the questions.

Put a tick mark on the correct option or fill the blanks or True/ False

Please check whether all questions are answered before submitting

## Questionnaire

Q1. Which of these is important for preparing templates for Next Generation Sequencing?

a) Isolating quality DNA/RNA from sample
b) Breaking up DNA into smaller fragments
c) Checking the quality and quantity of the fragment library
d) All of the above

Q2. What are SPRI beads?

a) Is used for purifying PCR products
b) Is used for size selection
c) Is used for purifying library fragments
d) All of the above
e) None of the above

Q3. What is barcode in NGS?

a) It is generated by the Nanopore software to label the sample
b) It is added during library preparation to identify the sample

c) It is tagged with the sequencing data for data submission
d) All of the above
e) None of the above

Q4. State if True or False: Nanopore can be used for quantification of SARS CoV-2-

a) True
b) False

Q5. What is insert?

a) A fragmented DNA or region of interest in the DNA library
b) A short oligo nucleotide sequences attached to both the ends of library
c) A primer sequence used to amplify the libraries

Q6. What is ONT?

a) Organization for Nanopore Technologies
b) Oxford Network Technologies
c) Oxford Nanopore Technologies
d) None of the above

Q7. Which generation Nanopore belongs to?

a) First Generation
b) Second Generation
c) Third Generation
d) Fourth Generation

Q8. What are the advantages of Nanopore Sequencing?

a) Can be sequenced anywhere (portability)
b) Can sequence only short fragments
c) Can sequence long reads
d) Both a & b
e) Both a & c

Q9. State if True or False: Nanopore can allow multiple samples to be sequenced

a) True
b) False

Q10. Applications of Nanopore Sequencing

a) Sars-CoV-2 Genome Sequencing
b) Targeted Exome Sequencing
c) Amplicon sequencing
d) Transcriptome sequencing
e) All of the above

Q11. How many times can you reuse a flongle?

a) Cannot be reused
b) One time
c) It depends

Q12. After Sequencing, washed flow cells should be stored at?

a) Room Temperature
b) 2 - 8 deg Celcius
c) –20 deg Celcius & Above
d) Ambient temperature

Q13. After Sequencing, instrument should be stored at?

a) Room Temperature
b) 2 - 8 deg Celcius
c) –20 deg Celcius & Above
d) Ambient temperature

Q14. What are loading beads?

a) Used for purification
b) Used for hybridization capture
c) Used for sequencing
d) None of the above

Q15. Once the sequences are obtained from your Next Generation Sequencing experiment what is the first thing you should do?

a) Perform a bioinformatics analysis of your data
b) Check your data using a different method
c) Publish your results
d) Further investigate the sequences of interest.

Q16. What is the format of data that Oxford Nanopore instrument produces?

a) FASTQ
b) BAM
c) FAST5
d) All of the above

Q 17. What is the ultimate aim of oxford Nanopore sequencing?

a) Calculating variants
b) Generating consensus genome of the virus
c) Producing FAST5 & FASTQ files
d) Both a & b

Q18. A variant for which there is evidence of an increase in transmissibility, more severe disease, increased hospitalizations or deaths is known as

a) Variant of Interest
b) Variant of Concern
c) Variant of epidemic
d) Variant of pandemic

Q19. Which is the database that is jointly initiated by the Union Health Ministry of Health, and Department of Biotechnology (DBT) with Council for Scientific & Industrial Research (CSIR) and Indian Council of Medical Research (ICMR) for SARS-CoV-19?

a) GISAID
b) NCBI
c) Genbank
d) INSACOG

Q20. Name a tool to compute the quality of sequencing reads

a) Sequence analysis viewer
b) FastQC
c) Bowtie
d) Porechop

Q21. For the accession number NC_000023.10, what is the reference chromosome number

a) 10
b) 01
c) 02
d) X

Q22. How many non-structural proteins have been annotated form hCoV-19

a) 16
b) 10
c) 10.5
d) 15

Q23. Which is a tool to perform base calling from the sequenced data of Nanopore?

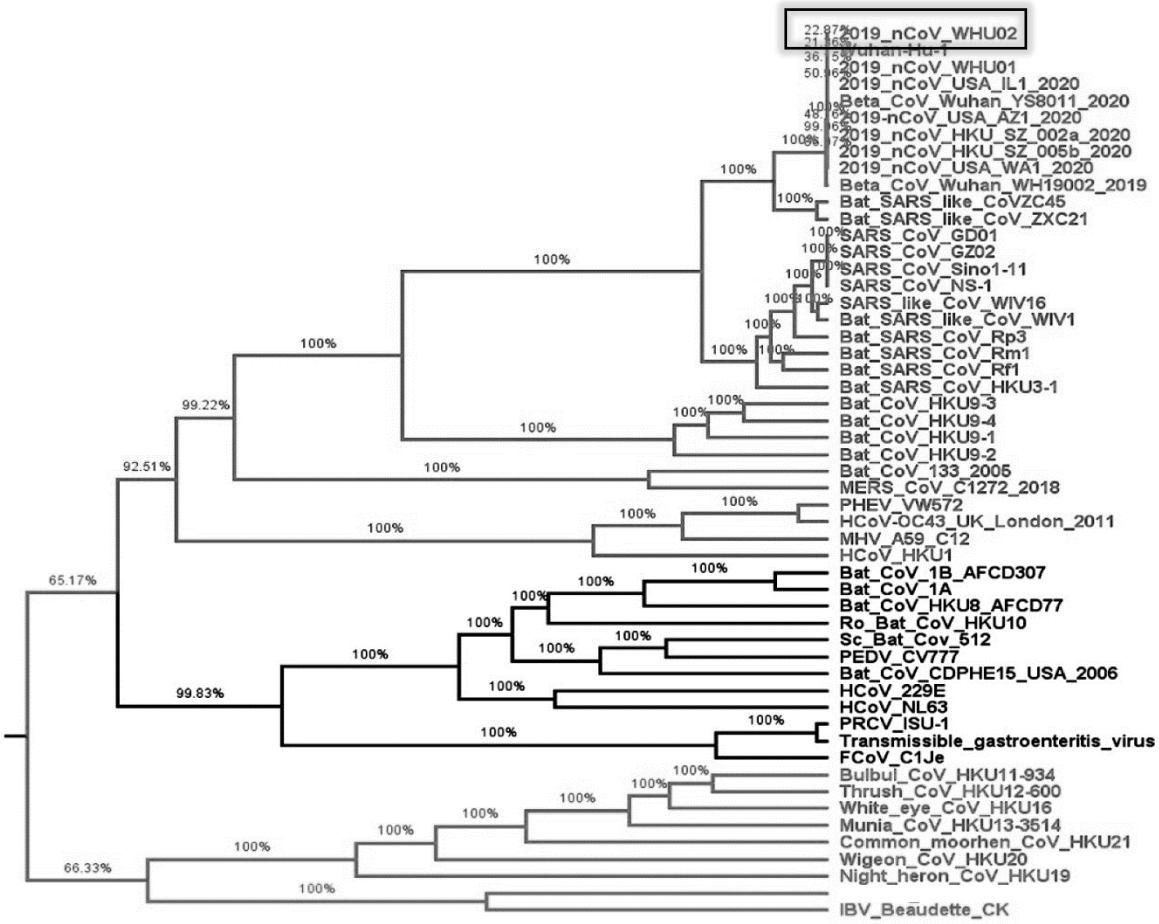
a) Nanopore seq
b) Guppy
c) Bowtie
d) NCBI

Q24. Which tool is used to visualize SARS-CoV variant?

a) Pango
b) Nextclade
c) NCBI

d) All the above

Q25. Using the image below, find out the most distant (taxonomically different) pathogen strain to the 2019_nCoV_WHU02




a) Night_heron_CoV_HKU19
b) Bat_CoV_1A
c) IBV_Beaudette_CK
d) MERS_CoV_C1272_2018

# Pre/Post Training Quiz - Answers

| Question No | Answer |
|---|---|
| 1. | d |
| 2. | d |
| 3. | b |
| 4. | b |
| 5. | a |
| 6. | c |
| 7. | c |
| 8. | e |
| 9. | a |
| 10. | e |
| 11. | a |
| 12. | b |
| 13. | a or d |
| 14. | c |
| 15. | a |
| 16. | c |
| 17. | d |
| 18. | b |
| 19. | d |
| 20. | b |
| 21. | d |
| 22. | a |
| 23. | b |
| 24. | d |
| 25. | c |

## USAID RISE "Workshop on Genome Sequencing of SARS-CoV-2 using Oxford Nanopore Sequencer"

## FEEDBACK

| | |
|---|---|
| Name: | Designation: |
| Training site: | Dates: |

## Training Overall

| | **Excellent: 5** | **Good: 4** | **Satisfactory: 3** | **Fair: 2** | **Poor: 1** |
|---|---|---|---|---|---|
| Content was relevant to me | | | | | |
| Material covered was sufficient | | | | | |
| Media were used appropriately which made learning easy | | | | | |
| I am confident of using the concepts covered | | | | | |
| Duration of the training was appropriate | | | | | |
| The training met my expectations | | | | | |

## Faculty

| | Excellent: 5 | Good: 4 | Satisfactory: 3 | Fair: 2 | Poor: 1 |
|---|---|---|---|---|---|
| Faculty had a good grasp of the subject | | | | | |
| The concepts were clearly explained | | | | | |
| Faculty involved all participants | | | | | |
| My questions were answered adequately | | | | | |

| | Excellent: 5 | Good: 4 | Satisfactory: 3 | Fair: 2 | Poor: 1 |
|---|---|---|---|---|---|
| Faculty was supportive and encouraging | | | | | |

## Presentations

| Sessions | Excellent: 5 | Good: 4 | Satisfactory: 3 | Fair: 2 | Poor: 1 |
|---|---|---|---|---|---|
| **Please write appropriate number as your answer to each question** | | | | | |
| Basic introduction on NGS | | | Content: | Presenter: | Overall: |
| Library preparation methodologies | | | Content: | Presenter: | Overall: |
| Nanopore Sequencing & its application | | | Content: | Presenter: | Overall: |
| Nanopore Kit contents | | | Content: | Presenter: | Overall: |
| Flow cell concepts | | | Content: | Presenter: | Overall: |
| Sars-CoV2 Sequencing Workflow | | | Content: | Presenter: | Overall: |
| Library Preparation process | | | Content: | Presenter: | Overall: |
| Setting up of Nanopore Instrument | | | Content: | Presenter: | Overall: |
| Flow cell loading & Priming | | | Content: | Presenter: | Overall: |
| Sequencing run monitoring & completion | | | Content: | Presenter: | Overall: |
| Flow cell washing & Re-fueling | | | Content: | Presenter: | Overall: |
| Bioinformatics Data Analysis | | | Content: | Presenter: | Overall: |
| Variant Interpretation & Results | | | Content: | Presenter: | Overall: |
| Data handling & management | | | Content: | Presenter: | Overall: |

## Hands-on sessions

| Sessions | Excellent: 5 | Good: 4 | Satisfactory: 3 | Fair: 2 | Poor: 1 |
|---|---|---|---|---|---|
| **Please write appropriate number as your answer to each question** | | | | | |
| Master mix preparation & RT-PCR Setup | | | Content: | Presenter: | Overall: |
| Library Preparation Process | | | Content: | Presenter: | Overall: |
| Sequencing Instrument Preparation & Setup | | | Content: | Presenter: | Overall: |
| Flow cell loading, priming & washing | | | Content: | Presenter: | Overall: |
| Sequencing monitoring & completion | | | Content: | Presenter: | Overall: |
| Bioinformatics Data Analysis | | | Content: | Presenter: | Overall: |
| Variant Interpretation & Results | | | Content: | Presenter: | Overall: |
| Data handling & management | | | Content: | Presenter: | Overall: |

1. **Three most useful aspect of the training program for me were:**
   - 
   - 
   - 
2. **Please write how you are going to use this learning in connection to below activity:**

Day to day job:

Improvement of my performance:

For my future growth:

**3. Three least useful aspect of the training program for me were:**

- 
- 
- 

**4. Please give your suggestions for improving the training program.**

**5. Would you recommend this training program to your colleagues?** YES / NO

## Overall Rating of the Training Program

| Excellent: 5 | Good: 4 | Satisfactory: 3 | Fair: 2 | Poor: 1 |
|---|---|---|---|---|

Signature: Date:

**Disclaimer:** Images used herein have been either sourced from the public domain, creative commons or are in accordance with fair use terms under the copyright law, and **not for distribution**.

Any trade/product names used in this document is for identification purposes only and does not imply endorsement by the USAID RISE.