From Sample to Sequence: Wet-Lab Essentials for WGS

Webinar series on Pathogen Genomics for Antimicrobial Resistance (AMR) Surveillance

Niamh Lacy-Roberts, nlac@dtu.food.dk

Bioinformatician, Global Capacity Building Group, Technical University of Denmark

The Fleming Fund
Regional Grants

Introduction

Whole genome sequencing (WGS) plays a vital role in AMR surveillance, enabling identification, typing, and tracking of resistant pathogens.

But sequencing is only as good as the sample it starts from.

- This presentation focuses on the wet-lab workflow — the foundation of any successful WGS project.

What we'll cover today:

- Overview of sequencing workflows
- Culturing vs. direct metagenomic sequencing
- DNA extraction and quality control
- Short-read vs. long-read sequencing
- Common pitfalls and how to troubleshoot them

The Fleming Fund
Regional Grants

DNA

DNA, or **Deoxyribonucleic Acid**, stores the blueprints for all cellular components.

It is made up of **nucleotides**, the basic building blocks of DNA.

Each nucleotide contains:

- A **sugar-phosphate backbone**
- One of four **nitrogenous bases**:
  Adenine (A), Thymine (T), Guanine (G), or **Cytosine (C)**

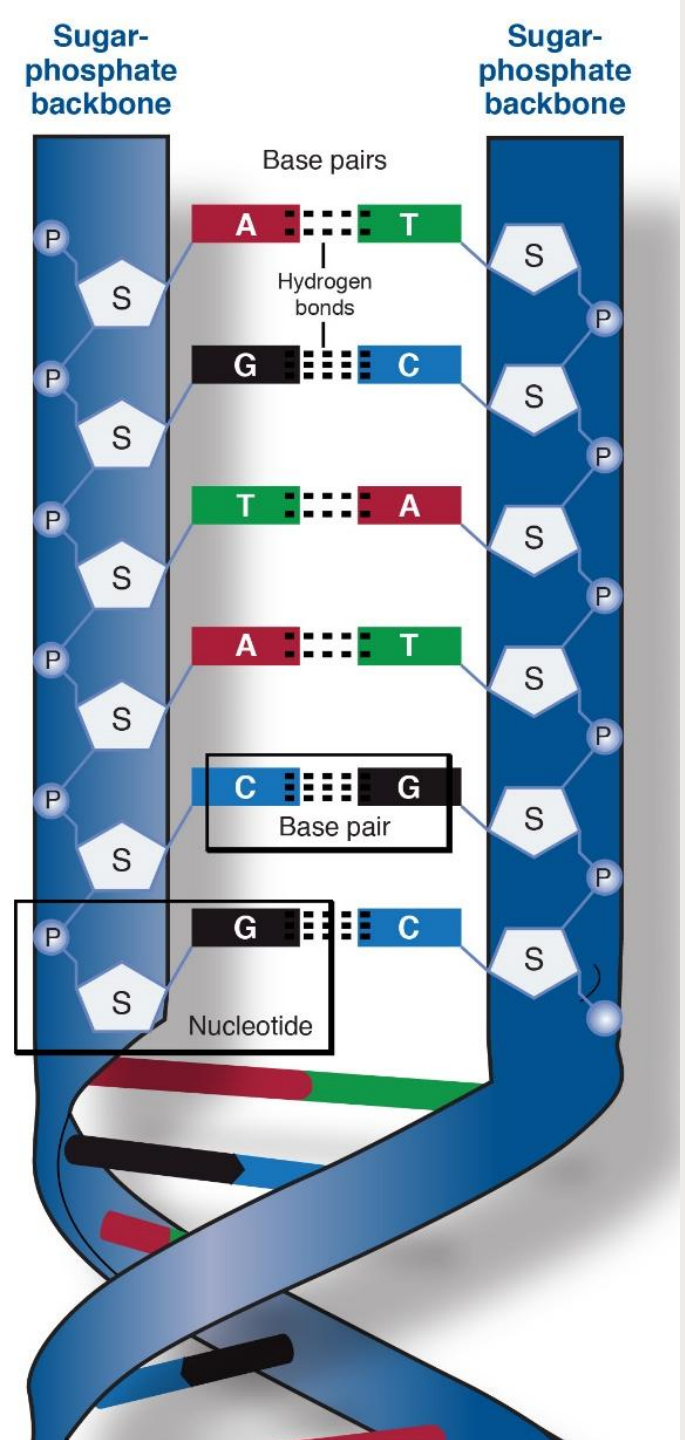
The sugar-phosphate units are **covalently bonded**, forming the stable backbone of the DNA strand.

DNA exists as a **double-stranded helix**, with each base pairing with a **complementary base**:

- A pairs with T
- G pairs with C

Because of this complementarity, knowing the sequence of one strand automatically tells us the sequence of the other.

- This **base pair** is the fundamental unit of DNA information.



Courtesy: National Human Genome Research Institute https://www.genome.gov/genetics-glossary/Phosphate-Backbone, *This image is a work of the National Institutes of Health and is in the public domain*

Bacterial genome

- In bacteria, DNA is stored in circular chromosomes.
- Most bacteria have 1 chromosome containing essential genes for survival, with notable exceptions e.g. Vibrio cholerae.
- Bacteria can also contain additional smaller circular DNA molecules, referred to as plasmids.
- Plasmids usually contain gene non-essential genes, which can confer advantages to the cell in specific cases, such as antimicrobial resistance genes.
- The genome refers to all the DNA in the cell
- Genome = Chromosome(s) + Plasmids

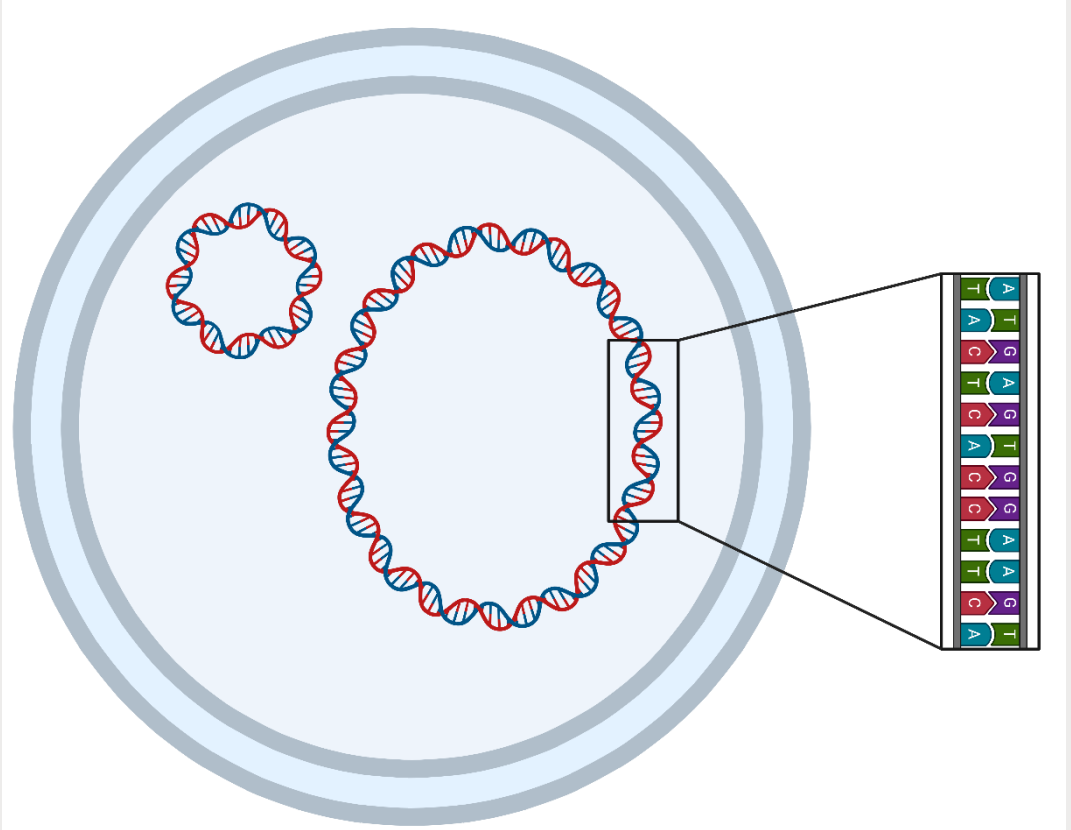


Figure created with BioRender.com

Central Dogma of Molecular Biology

- **Concept:** DNA → RNA → Protein.
- Genomics focuses on the first step: decoding DNA sequences.
- Bioinformatics connects these steps by analysing how genetic data translates into cellular function.
- **Examples**:
- DNA mutation → Faulty protein → Disease.
- DNA mutation → Protein conveying resistance→ AMR.

Figure created with BioRender.com

The Fleming Fund
Regional Grants

DNA Sequencing

Sequencing is the process of determining the exact order of base pairs (A, T, G, C) in a DNA molecule.

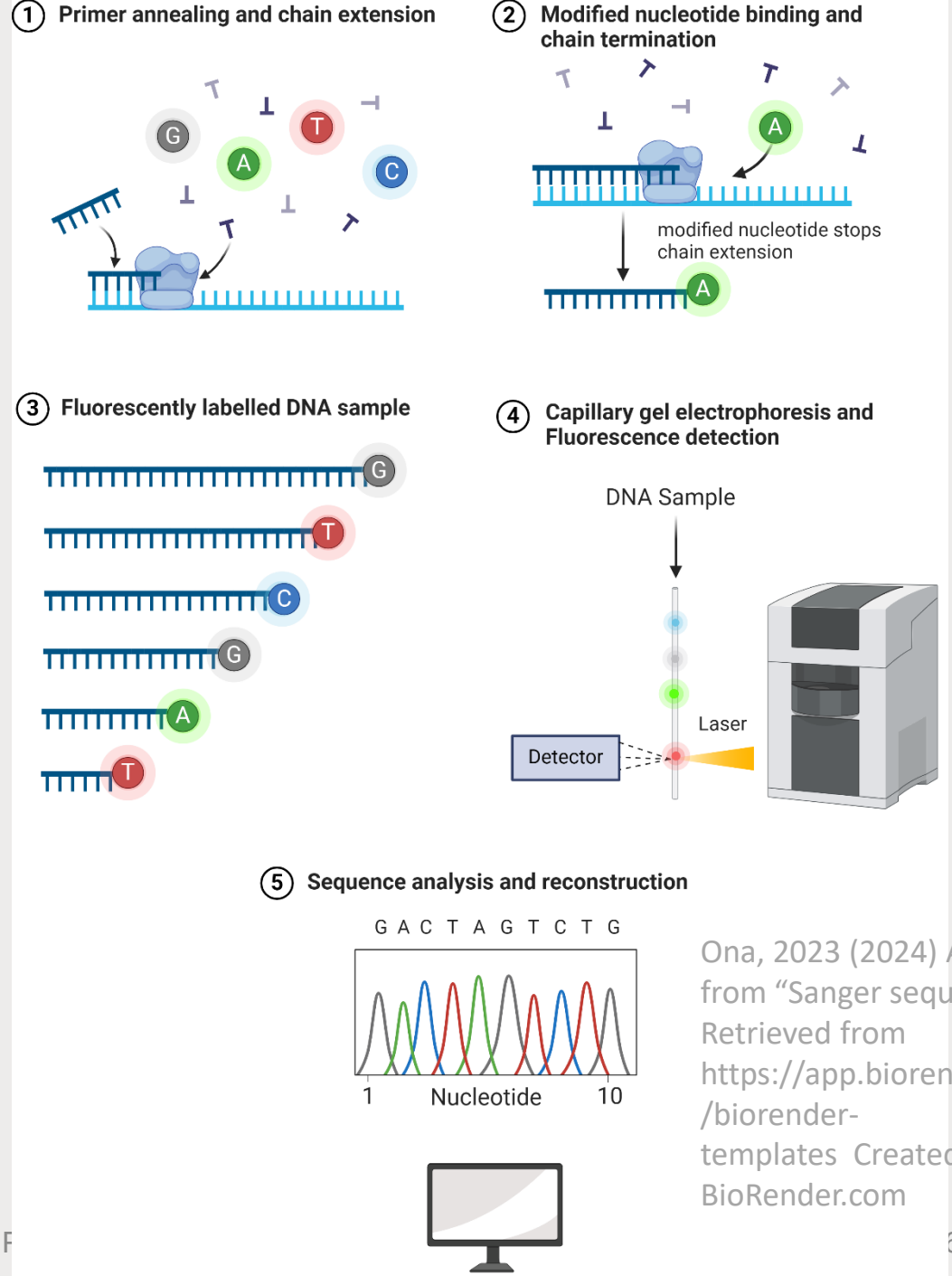
The Sanger Method (First published in **1977**):

- Based on the natural process of DNA replication.
- Uses DNA polymerase to synthesize new DNA strands.
- Incorporates modified nucleotides (dideoxynucleotides) that terminate synthesis at random points.

How does it work?

- Each terminating nucleotide is tagged with a fluorescent dye—one color for each base.
- DNA fragments of different lengths are generated and separated by size.
- A laser excites the dyes, and the emitted light reveals the sequence base-by-base.

“Sequencing-by-synthesis” — DNA is read as it is being built.



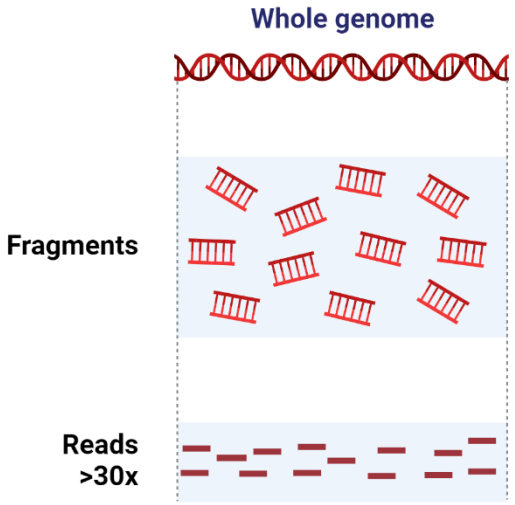
Ona, 2023 (2024) Adapted from “Sanger sequencing”. Retrieved from https://app.biorender.com/biorender-templates Created in BioRender.com

The Fleming Fund
Regional Grants

Sequencing approaches

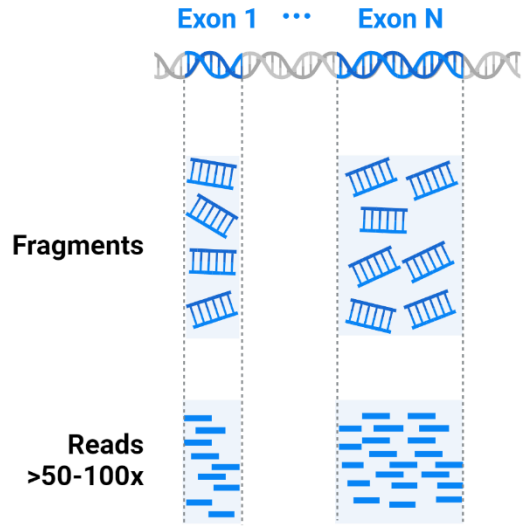
- Host of different approaches developed for sequencing at different levels and specificities.
- For AMR surveillance and pathogens, the most relevant methods are WGS and targeted gene panels, single isolate and metagenomic sequencing.

Whole Genome Sequencing



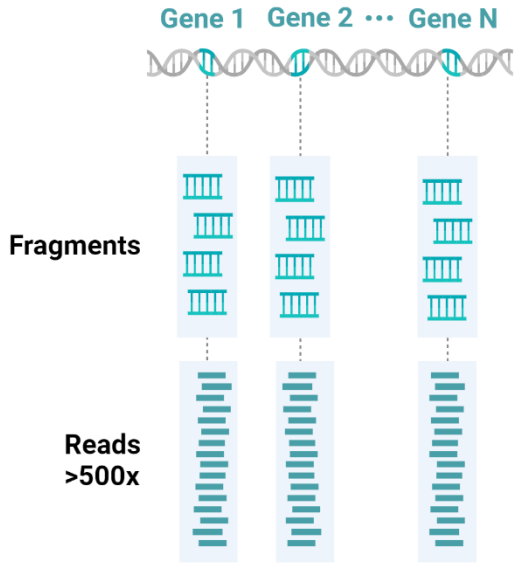
Coverage: All genes and non-coding DNA
Time: Longest turnaround time
Cost: Most expensive

Exome Sequencing



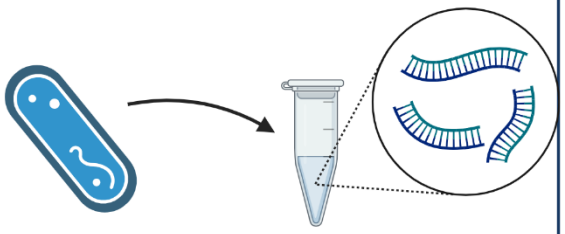
Coverage: Entire exome (20-25k genes)
Time: Long turnaround time
Cost: Cost-effective

Targeted Gene Panel



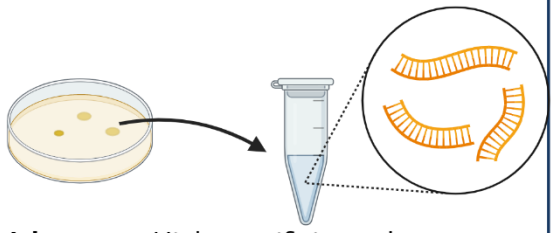
Coverage: 10-500 genes
Time: Relatively Rapid turnaround time
Cost: Most cost-effective

Single Cell



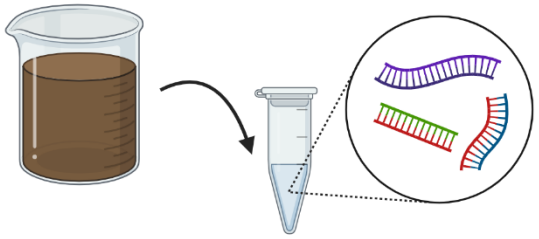
Advantage: Very high specificity
Disadvantage: Low representativeness and technically difficult

Single Isolate



Advantage: High specificity and easy analysis.
Disadvantage: strain must be culturable and pure

Metagenomic



Advantage: Represents entire bacterial population, even unculturable spp.
Disadvantage: Very difficult to analyse and ensure proper coverage

29 December 2017

Figure created with BioRender.com

Whole Genome Sequencing (WGS)

Whole Genome Sequencing (WGS) is the process of determining the complete DNA sequence of an organism's genome — whether bacterial, viral, human, or otherwise.

WGS advantages:

Comprehensive view:

- Captures the entire genetic content of the organism — including coding and non-coding regions.

Detects all variation types:

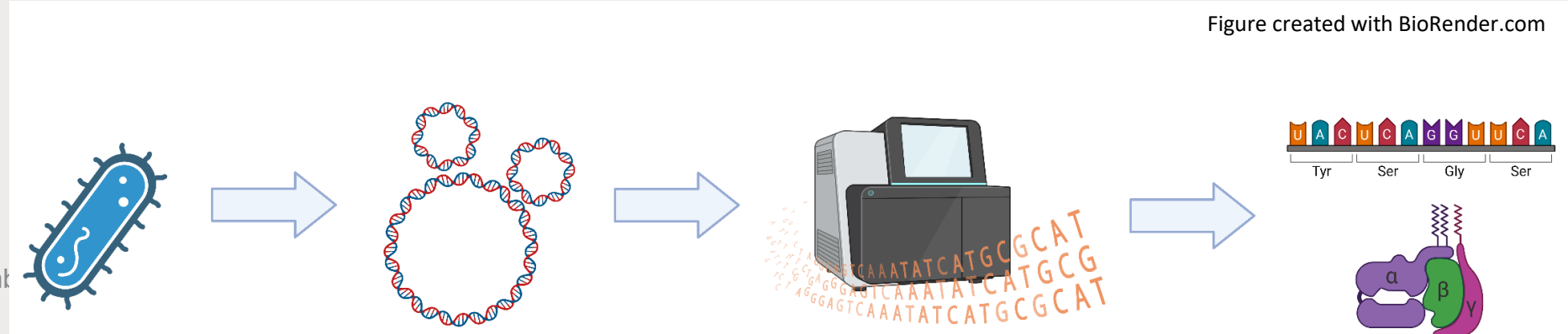
- Identifies both small (SNPs) and large (structural) variations that targeted methods might miss.

High-resolution typing:

- Enables pylogenetic comparison at single-base resolution — critical for outbreak detection and tracking.

Flexible data reuse:

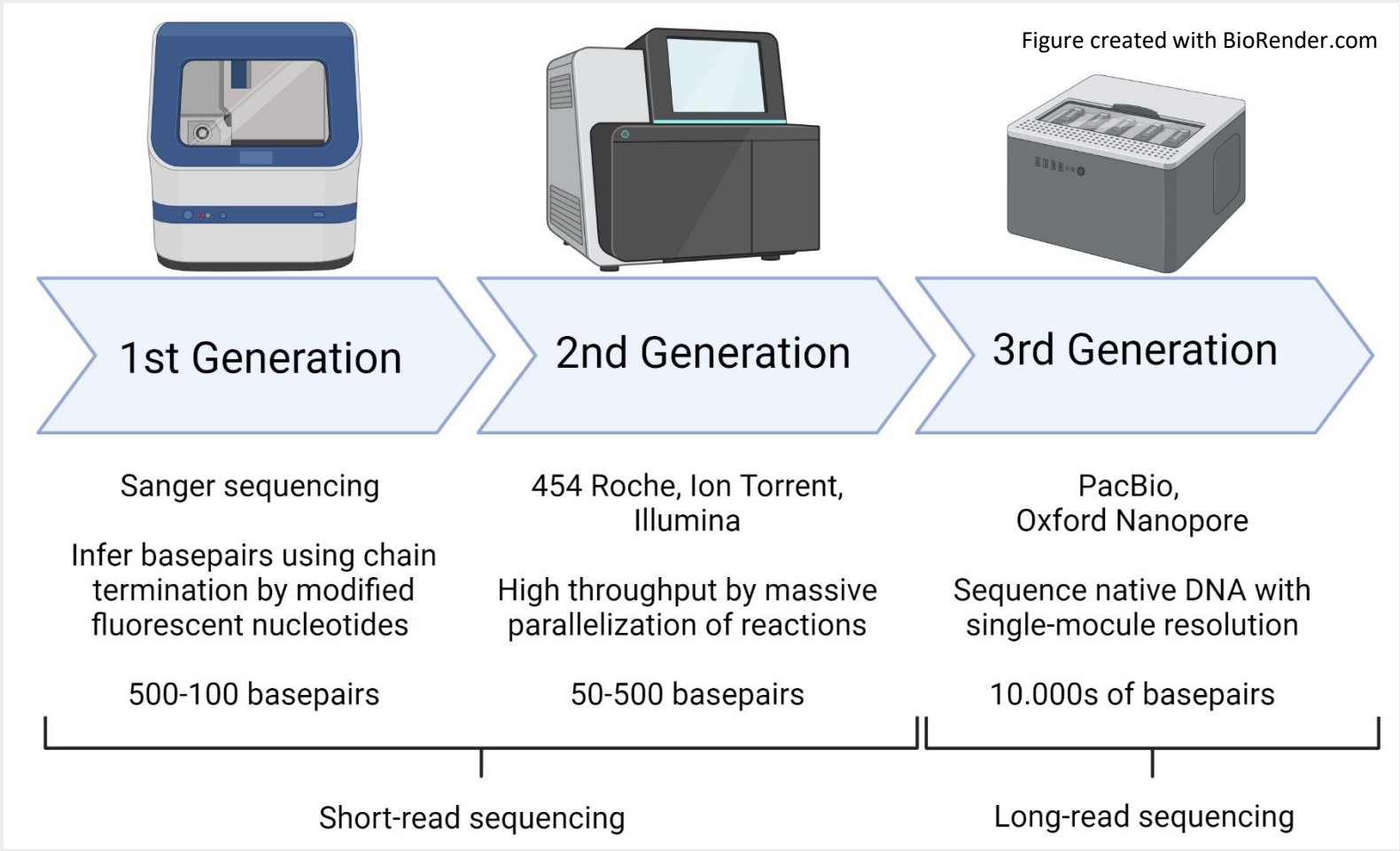
- Raw data can be re-analyzed when new resistance genes or epidemiological questions arise.



The Fleming Fund
Regional Grants

Sequencing platforms

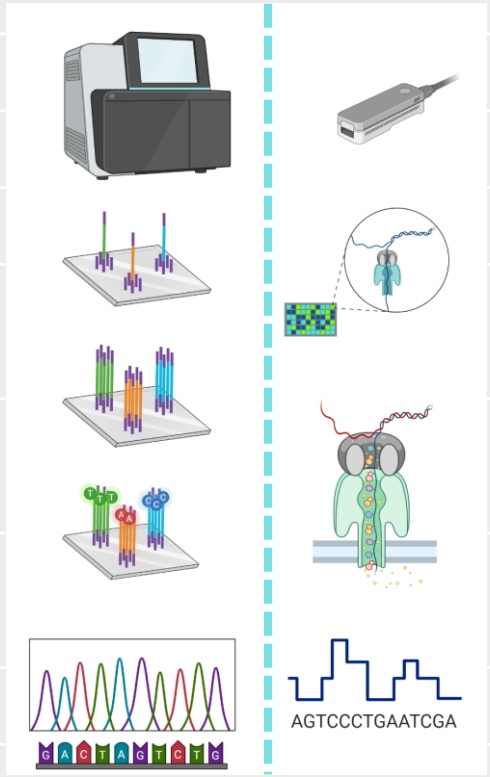
- There exists a variety of sequencing platforms divided into generations based on their methodology.
- 1st generations were based on the original Sanger sequencing method. Not suitable for WGS.
- 2nd generation improved sequencing by running millions of reactions in parallel, increasing the throughput of data
- 3rd generation utilized new methods to sequence ultra long segments of DNA



The Fleming Fund
Regional Grants

Short-Read vs. Long-Read Sequencing

Feature	Short-Read (e.g., Illumina)	Long-Read (e.g., ONT, PacBio)
Read length	100–300 bp	Thousands to >1 Mb
Accuracy	99.9% (very high)	~98–99.9% with new chemistries
DNA quality needed	Fragmented DNA OK	High Molecular Weight (HMW) needed
Ideal for	SNPs, MLST, AMR gene detection	Plasmids, mobile elements, hybrid assembly
Limitations	Fragmented assemblies	Sensitive to impurities; higher DNA needs
Cost & speed	Low cost, batch-friendly	Real-time, portable, cost decreasing

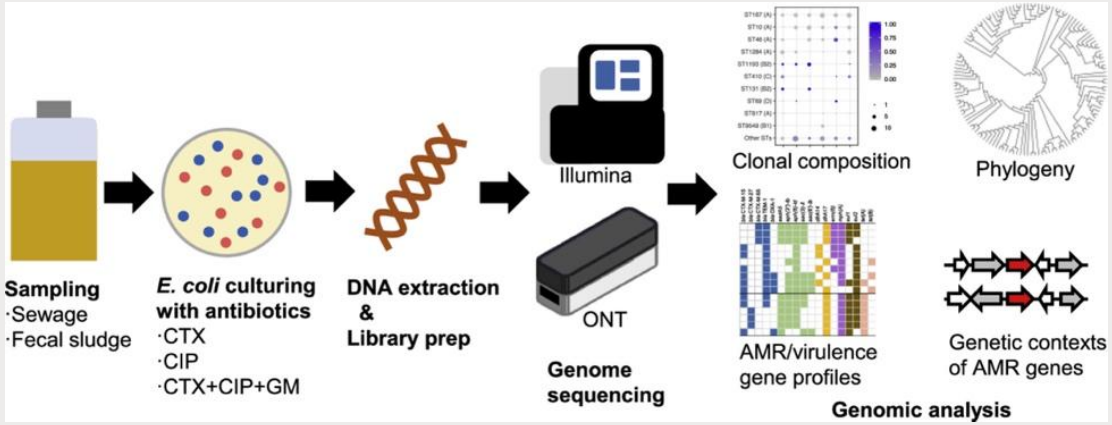


The Fleming Fund
Regional Grants

- Each step must be optimized to prevent data quality issues later.
- Important note: Workflow differs slightly for short-read vs. long-read platforms.

Sequencing Workflow Overview

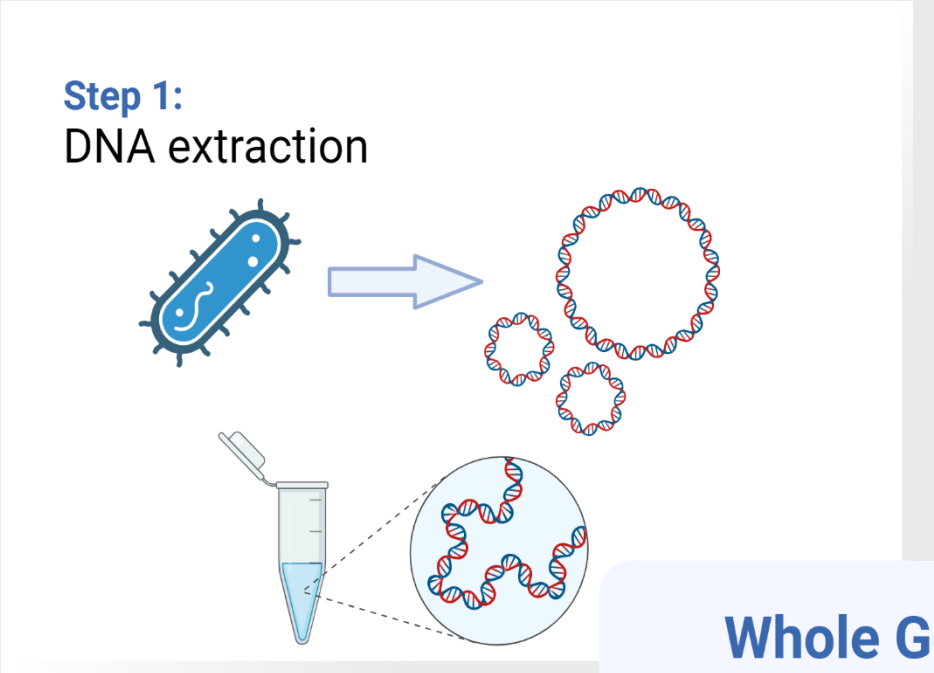
1. **Sample Collection & Transport**
   - Ensure cold chain & proper documentation
2. **Isolation or Direct Extraction**
   - Culture pathogens or use direct metagenomic DNA
3. **DNA Extraction**
   - Select method based on organism type
4. **Quality Control (QC)**
   - Measure purity, concentration, and integrity
5. **Library Preparation**
   - Tailored to sequencing platform (e.g., ONT, Illumina)
6. **Sequencing**
   - Generate reads for downstream analysis
7. **Bioinformatics**



Example of a WGS AMR surveillance workflow. Image from Gomi, R., Matsumura, Y., Yamamoto, M., Tanaka, M., Komakech, A. J., Matsuda, T., & Harada, H. (2024). Genomic surveillance of antimicrobial-resistant Escherichia coli in fecal sludge and sewage in Uganda. Water Research, 248, 120830. https://doi.org/10.1

The Fleming Fund
Regional Grants

1) Cells from a pure culture broken open. DNA is extracted, cleaned for proteins and cell debris.

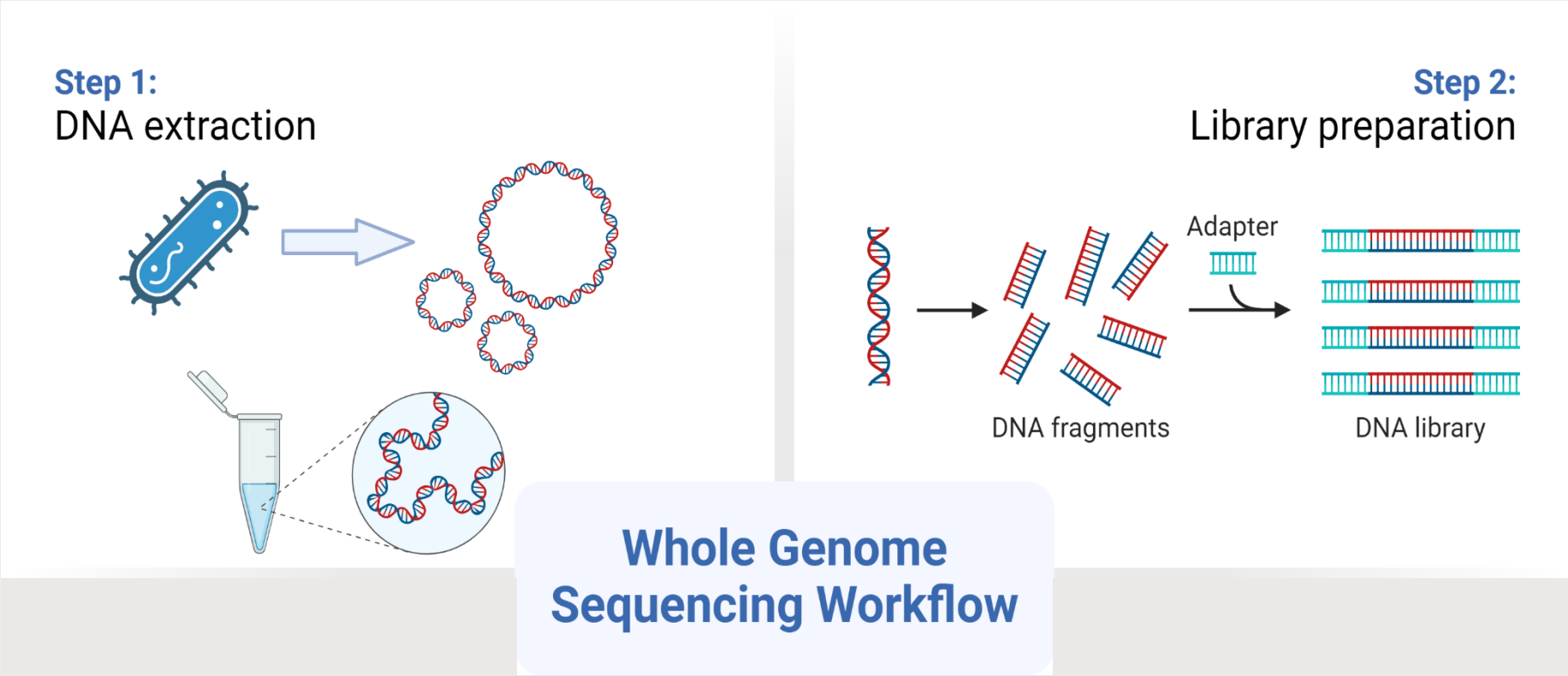


Whole Genome Sequencing Workflow

Ona, 2023 (2024) Adapted from "Next Generation Sequencing Workflow". Retrieved from https://app.biorender.com/biorender-templates Created in BioRender.com

The Fleming Fund
Regional Grants

1) Cells from a pure culture broken open. DNA is extracted, cleaned for proteins and cell debris.
2) DNA is fragmented to smaller pieces (optional) and adapters are attached.



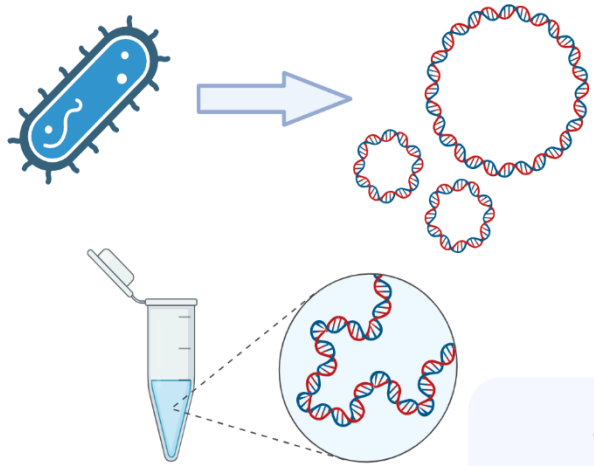
Ona, 2023 (2024) Adapted from "Next Generation Sequencing Workflow". Retrieved from https://app.biorender.com/biorender-templates Created in BioRender.com

The Fleming Fund
Regional Grants

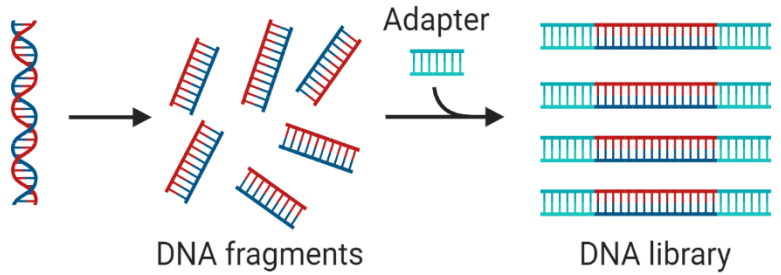
1) Cells from a pure culture broken open. DNA is extracted, cleaned for proteins and cell debris.
2) DNA is fragmented to smaller pieces (optional) and adapters are attached.
3) DNA library is loaded to sequencing platform and the sequence of nucleotides in each fragment determined.

Whole Genome Sequencing Workflow

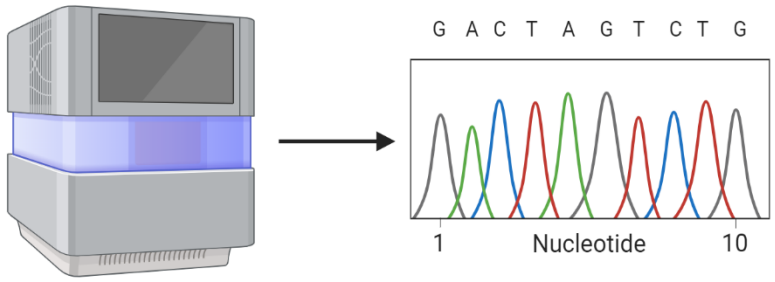
Step 1:
DNA extraction



Step 2:
Library preparation



Step 3:
Sequencing



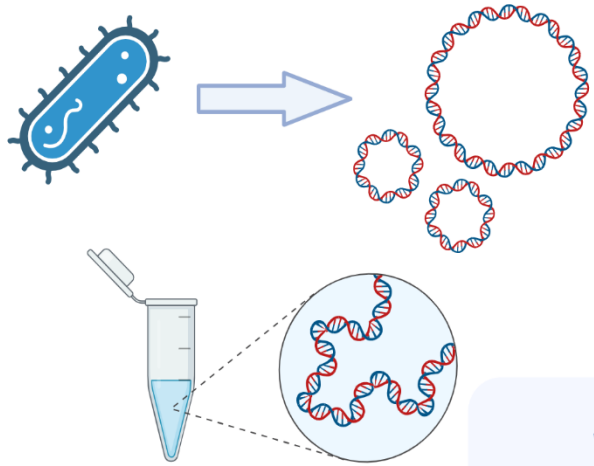
Ona, 2023 (2024) Adapted from "Next Generation Sequencing Workflow". Retrieved from https://app.biorender.com/biorender-templates Created in BioRender.com

The Fleming Fund
Regional Grants

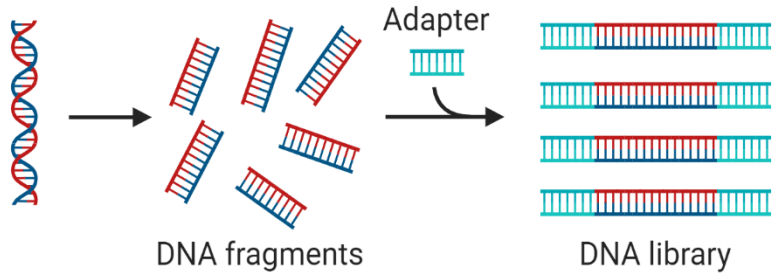
1) Cells from a pure culture broken open. DNA is extracted, cleaned for proteins and cell debris.
2) DNA is fragmented to smaller pieces (optional) and adapters are attached.
3) DNA library is loaded to sequencing platform and the sequence of nucleotides in each fragment determined.
4) The machine outputs each fragment as a “read”. Post-processing and quality control (QC) are conducted before analysis.

Whole Genome Sequencing Workflow

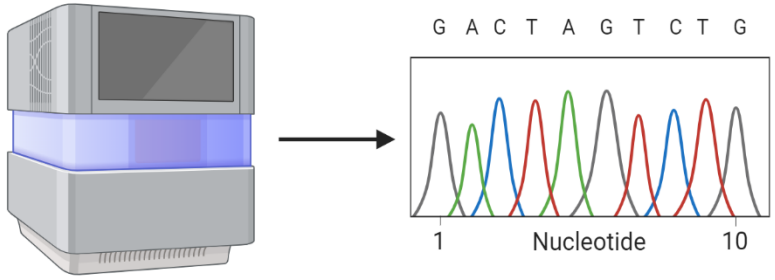
Step 1:
DNA extraction



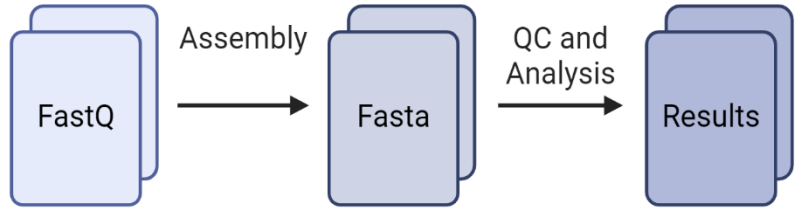
Step 2:
Library preparation



Step 3:
Sequencing



Step 4:
Analysis



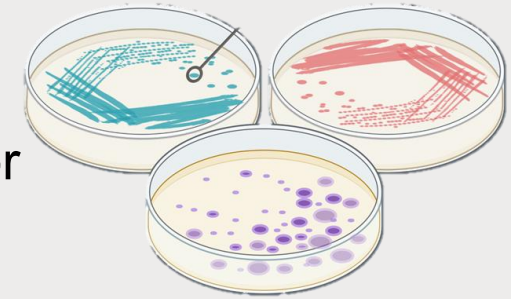
Ona, 2023 (2024) Adapted from “Next Generation Sequencing Workflow”. Retrieved from https://app.biorender.com/biorender-templates Created in BioRender.com

The Fleming Fund
Regional Grants

Isolation Techniques

Culture-Based Sequencing

- Involves isolating a single pathogen in pure culture
- Gold standard for clinical microbiology and surveillance
- **Advantages:**
  - High-quality, species-specific DNA
  - Enables strain-level comparison
- **Disadvantages:**
  - Time-consuming
  - May miss unculturable or fastidious organisms



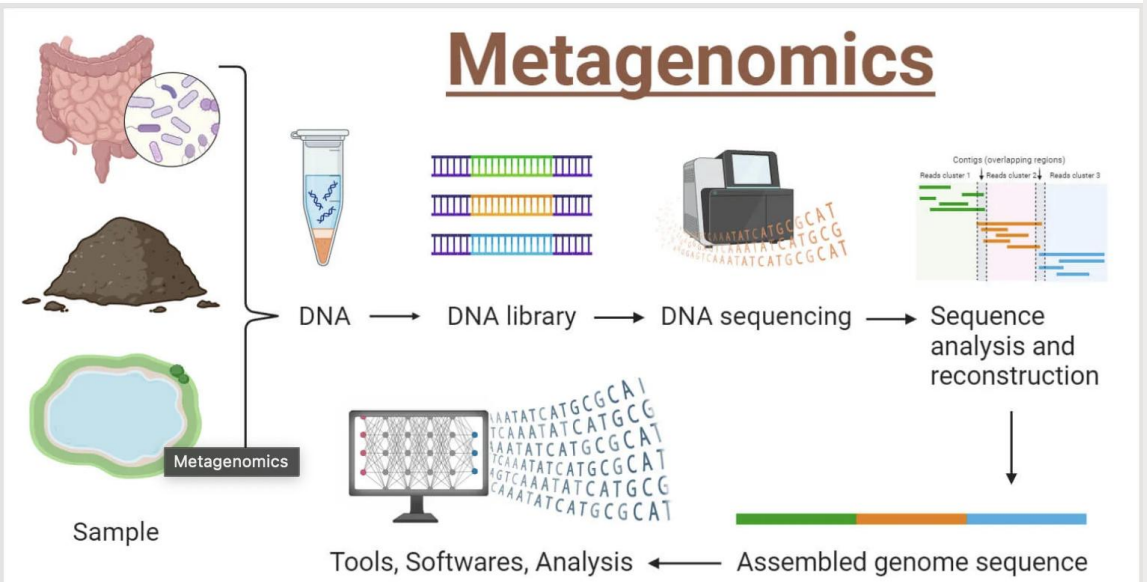


Image taken from https://microbenotes.com/metagenomics/

Direct Sequencing (Metagenomics)

- Extracts DNA directly from the sample (e.g., stool, sputum, swab)
- **Advantages:**
  - Detects mixed populations and unculturable pathogens
- **Disadvantages:**
  - Host DNA contamination
  - Complex data analysis
  - Lower DNA yield and purity

The Fleming Fund
Regional Grants

Manual vs Automated DNA Extraction

Manual DNA Extraction

Advantages:

- Lower startup cost (no equipment investment)
- Greater flexibility — protocols can be adapted for different organisms or sample types
- Ideal for troubleshooting or method development
- Often better for **long-read sequencing**, where gentle handling preserves high molecular weight DNA

Challenges:

- Labor-intensive and time-consuming, especially for large sample sets
- Greater risk of variability between runs or operators
- Higher chance of contamination if aseptic technique isn't consistent
- Requires trained, careful lab staff — pipetting accuracy matters

Automated DNA Extraction

Advantages:

- Higher reproducibility and consistency across samples
- Reduces human error and contamination risk
- Faster and scalable — some systems can process 96 samples at a time
- Integrates well into standardized WGS pipelines

Challenges:

- Higher upfront cost (equipment, service contracts)
- Less flexibility for tweaking protocols
- Some platforms may shear DNA more than manual methods — not ideal for long-read applications unless carefully validated

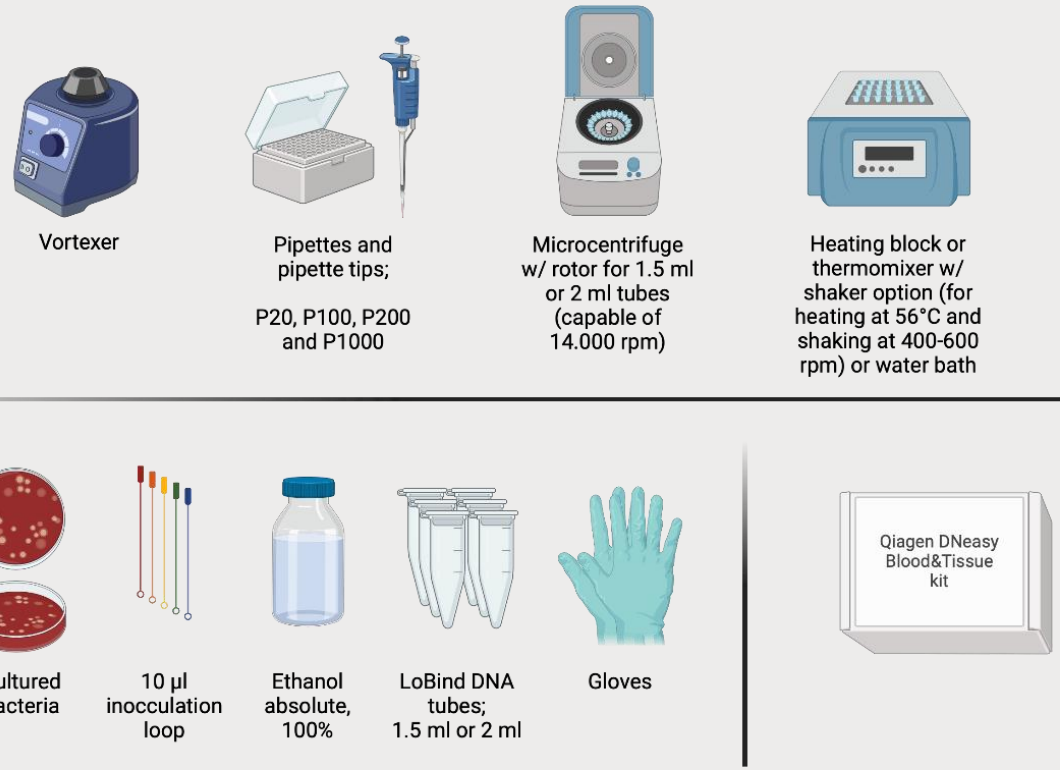
DNA Extraction - DNeasy Blood & Tissue Kit

Key Features:

- **Fast, column-based extraction** of high-quality DNA
- Suitable for PCR, Southern blotting, AFLP, RAPD, and RFLP
- No phenol/chloroform or alcohol precipitation required
- Minimal handling and rapid protocol: **~20 minutes after lysis**
- **Direct lysis + selective DNA binding** to silica membrane
- Simple centrifugation removes proteins, salts, and inhibitors
- Allows **parallel processing** of multiple samples
- DNA eluted in low-salt buffer or water – ready for downstream use

Typical DNA Output:

- Purity: **A260/A280 ~1.7–1.9**
- Fragment size: up to **50 kb** (majority ~30 kb)
- Efficient recovery of fragments as small as **100 bp**



Kit Storage

- DNeasy spin columns, DNeasy 96 plates and all buffers should be stored dry, at room temperature (15–25°C) and are stable for 1 year under these conditions, if not otherwise stated on the label.
- DNeasy Blood & Tissue Kits contain a ready-to-use Proteinase K solution, which is supplied in a specially formulated storage buffer.
- Proteinase K is stable for at least 1 year after delivery when stored at room temperature.
- For storage longer than one year or if ambient temperatures often exceed 25°C, we suggest storing Proteinase K at 2–8°C.

The Fleming Fund
Regional Grants

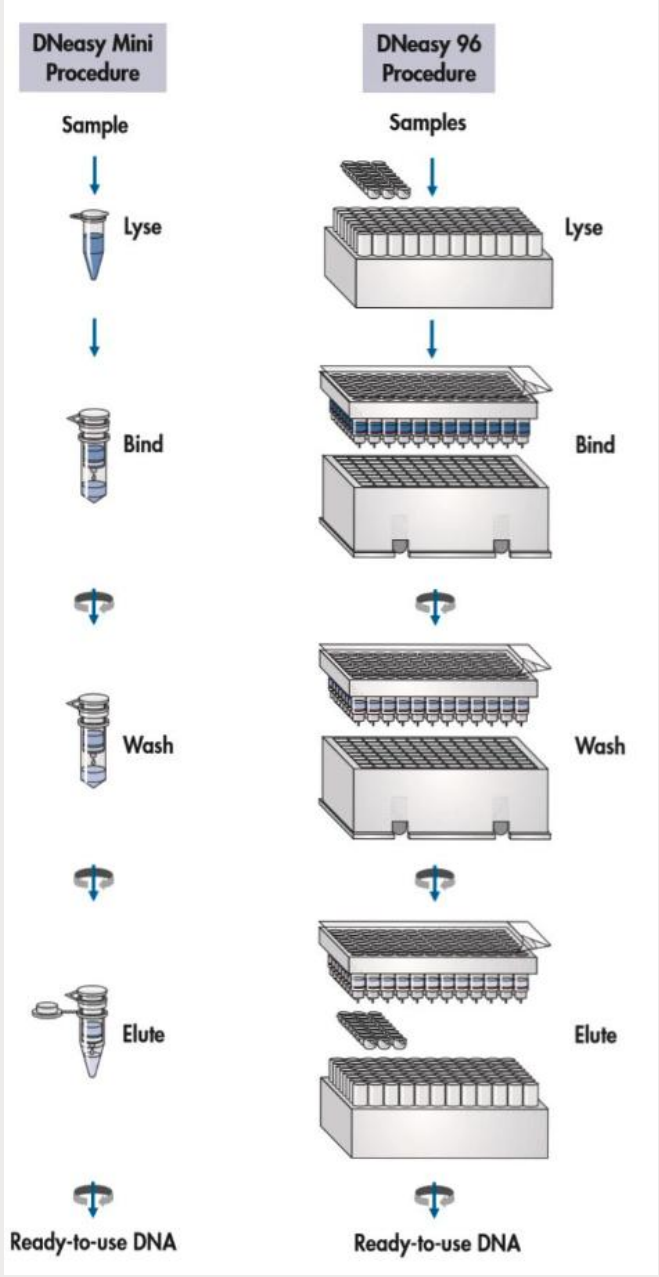
DNeasy DNA Extraction Workflow

Step-by-Step Process:

- **Sample lysis** with **Proteinase K** breaks down cells and proteins.
- **Buffering conditions** are optimized for DNA binding.
- Lysate is loaded onto a **DNeasy spin column or 96-well plate**.
- During **centrifugation**, DNA binds to the silica membrane, while contaminants pass through.
- **Two wash steps** remove remaining impurities and enzyme inhibitors.
- **DNA is eluted** in water or low-salt buffer, ready for downstream use.

DNA Quality:

- **A260/A280 ratio**: **~1.7–1.9**
- **UV absorbance scan**: sharp, symmetric peak at **260 nm**, indicating high purity



The Fleming Fund
Regional Grants

Why DNA Quality Control (QC) Is Essential

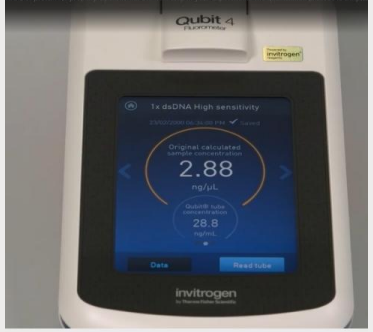
DNA QC ensures your sample is:

- Concentrated enough for library preparation
- Free from inhibitors (e.g., salts, proteins, phenol)
- Intact and appropriate for the sequencing platform (short vs long read)

Why it matters:

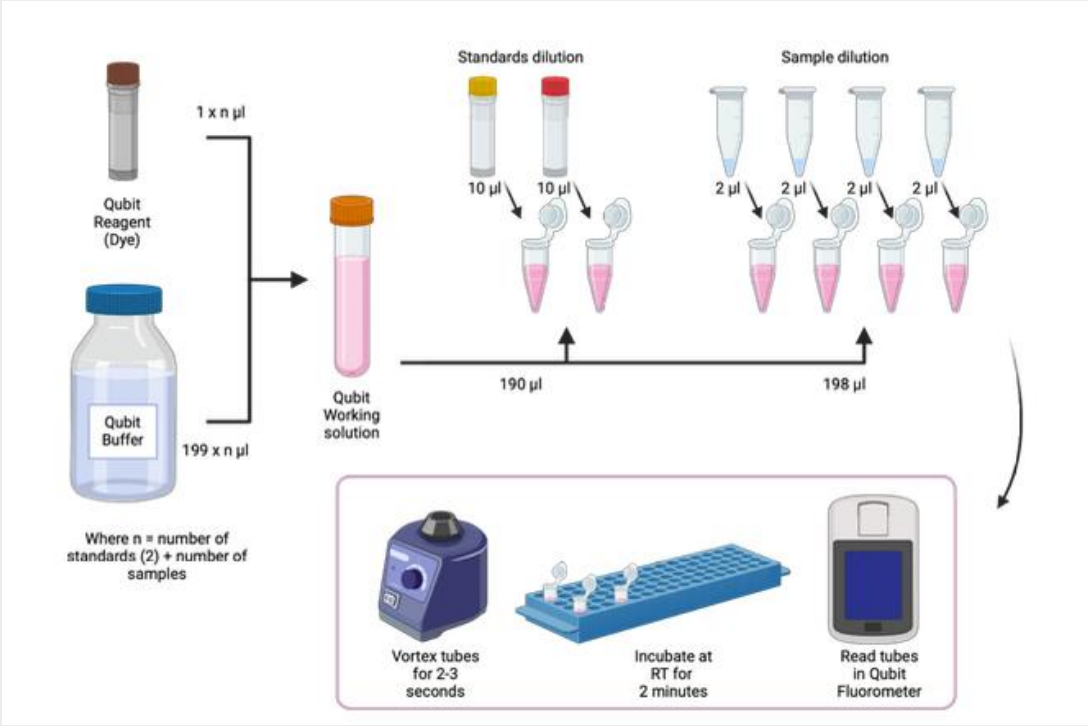
- Illumina tolerates some DNA fragmentation and impurities
- ONT & PacBio require high-purity, high-molecular-weight DNA — poor QC = failed runs
- Contaminants can inhibit enzymes during library prep
- Low or degraded DNA leads to poor yield, low coverage, or no sequencing output

The Fleming Fund
Regional Grants



DNA quantification with Qubit fluorometer

- Designed for accurate quantification of DNA, RNA, and proteins
- Supports RNA integrity and quality assessment
- Uses fluorescent dyes that bind specifically to target molecules
- Highly sensitive — ideal for low-concentration or precious samples
- Requires only 1–20 µL of sample
- Delivers results in under 3 seconds (for DNA)

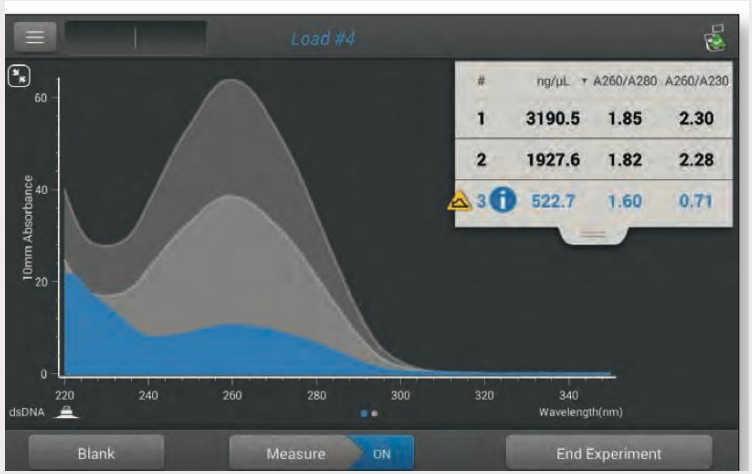


Common DNA Assay Kits:

Assay	Detection Range (total)	Concentration Range
Qubit dsDNA HS (High Sensitivity)	0.1–120 ng	0.005–120 ng/µL
Qubit dsDNA BR (Broad Range)	4–2,000 ng	0.2–2,000 ng/µL

The Fleming Fund
Regional Grants

DNA quality



High-quality DNA is essential for successful library preparation.

Poor-quality input (e.g., degraded, impure, or incorrectly quantified DNA) can lead to:

- Enzyme failure during library prep
- Low yield or poor complexity libraries
- Failed or low-quality sequencing runs

Common Contaminants to Avoid:

- Chemical impurities: detergents, phenol, ethanol, EDTA, salts
- Biological contaminants: RNA, single-stranded DNA, proteins, dyes
- These can inhibit enzymes used in fragmentation, ligation, and amplification

How to Assess DNA Quality:

Nanodrop spectrophotometry

- A260/280 ratio ~1.8 → indicates protein-free DNA
- A260/230 >2.0 → indicates low salt/solvent carryover
- UV absorbance scan should show a clean, symmetric peak at 260 nm

Note: Nanodrop is useful for purity, but not accurate for concentration or fragment size.

Tip: Always check both quantity and quality before starting library prep — especially for long-read sequencing platforms like ONT or PacBio, which are more sensitive to contaminants.

Troubleshooting – Low yield

Storage of starting material	DNA yield is dependent on the type, size, age and storage of starting material. Lower yields will be obtained from
Too much starting mater	
Insufficient mixing of sar AL and ethanol before bi	15 s each time
DNA inefficiently eluted	
Buffer AW1 or Buffer AW incorrectly	
Water used instead of Bu elution	
Bacteria insufficient lysis	
DNA not bound to DNea column.	

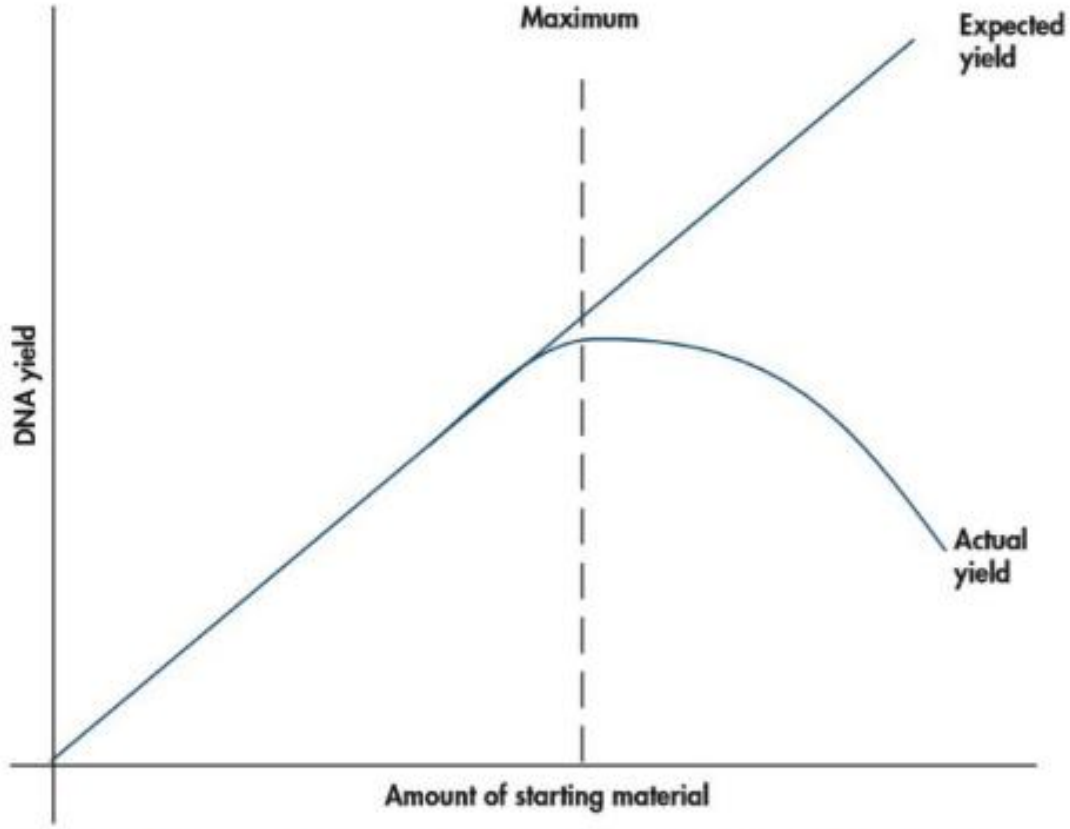


Figure 2. Schematic diagram of effect of sample size on DNA yield. If more than the maximum amount of starting material is used, DNA yield will be lower than expected.

Troubleshooting continued

DNeasy Mini spin column or DNeasy 96 plate clogged	Too much starting material and/or insufficient lysis	Increase g-force and/or duration of centrifugation step. In future preparations, reduce the amount of starting material used Insufficient lysis
Low concentration of DNA in the eluate	Second elution step diluted the DNA	Use a new collection tube for the second eluate to prevent dilution of the first eluate. Reduce elution volume to 50–100 µl
A260/A280 ratio of purified DNA is low	a) Water used instead of buffer to measure A260/A280 b) Inefficient cell lysis	Use 10 mM Tris·Cl, pH 7.5 instead of water to dilute the sample before measuring purity.
A260/A280 ratio of purified DNA is high	High level of residual RNA	Perform the optional RNase treatment in the protocol.

Troubleshooting continued

DNA does not perform well in downstream applications	a) Salt carryover	Ensure that Buffer AW2 has been used at room temperature. Ensure that Buffer AW1 and Buffer AW2 were added in the correct order.
	b) Ethanol carryover	Ensure that, when washing with Buffer AW2, the column is centrifuged for 3 min at 20,000 x g (14,000 rpm) to dry the DNeasy membrane. Following the centrifugation step, remove the Dneasy Mini spin column carefully so that the column does not come into contact with the flow-through. If ethanol is visible in the DNeasy Mini spin column (as either drops or a film), discard the flow-through, keep the collection tube, and centrifuge for a further 1 min at 20,000 x g.
	c) Too much DNA used	For PCR applications, a single-copy gene can typically be detected after 35 PCR cycles with 100 ng template DNA.
DNA sheared	a) Sample repeatedly frozen and thawed	Avoid repeated freezing and thawing of starting material.
	b) Sample too old	Old samples often yield only degraded DNA
White precipitate in Buffer ATL or Buffer AL	White precipitate may form at low temperature after prolonged storage	Any precipitate formed when Buffer ATL or Buffer AL are added must be dissolved by incubating the buffer at 56°C until it disappears.

The Fleming Fund
Regional Grants

Library Preparation – Illumina vs ONT

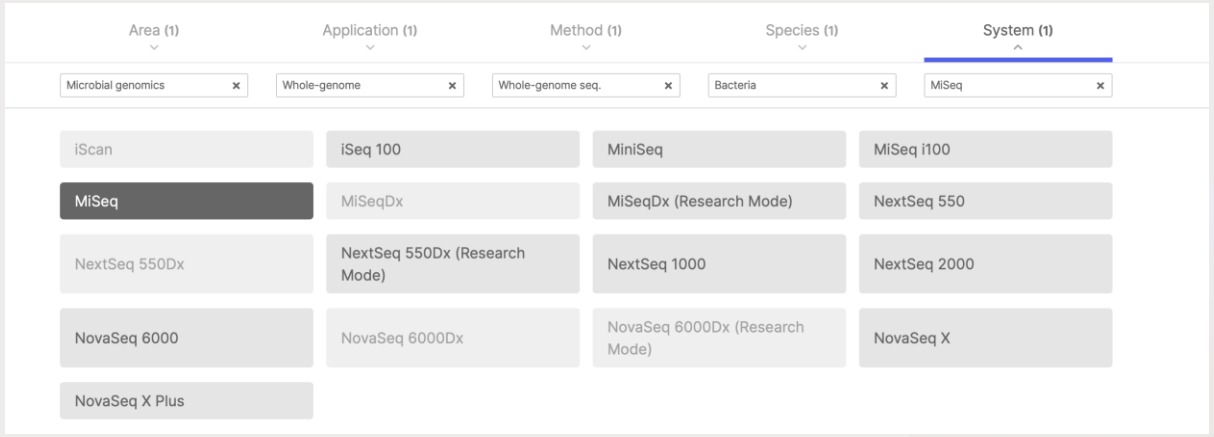
Illumina Library Prep

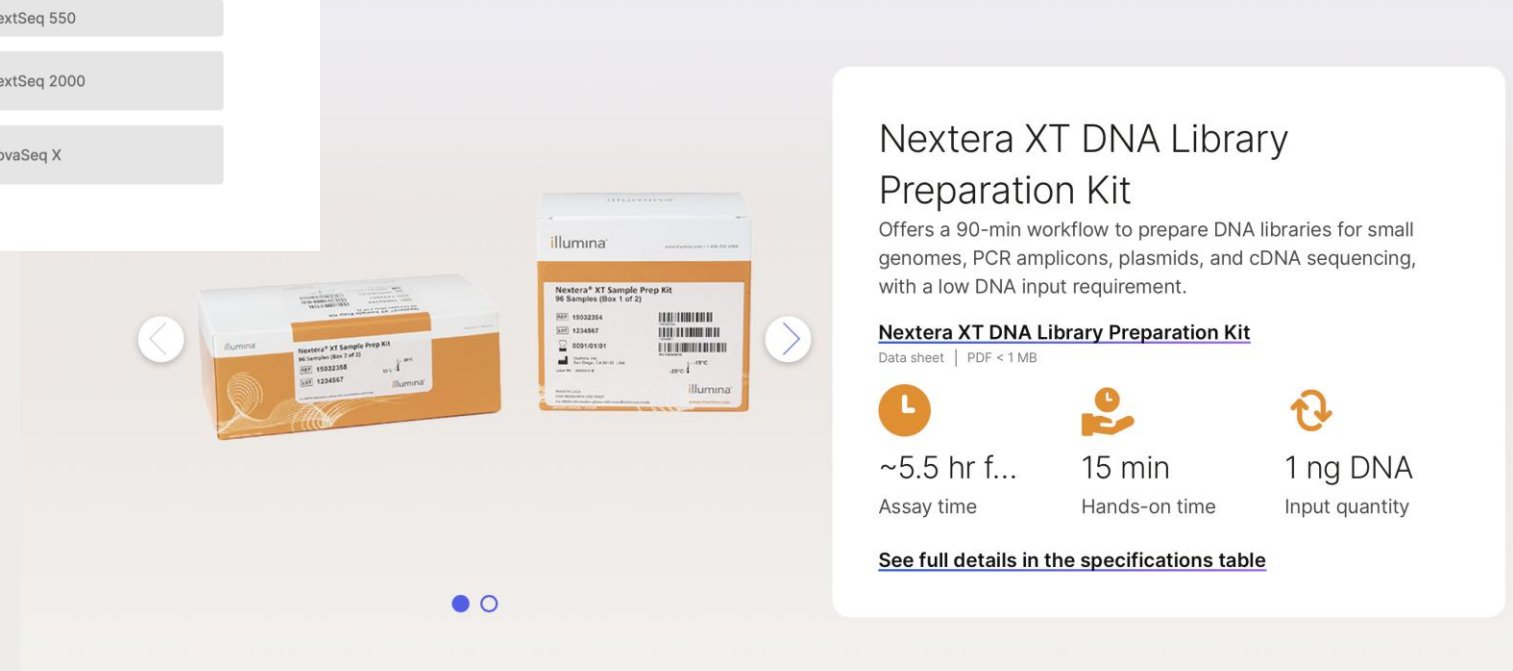
- Uses short DNA fragments (~200–600 bp)
- Typically involves enzymatic fragmentation and PCR amplification
- Adapters contain flow cell binding sites + barcodes
- Highly standardized and automation-friendly
- DNA input: ~1 ng – 1 µg
- Less sensitive to DNA shearing or impurities

ONT Library Prep

- Designed for long DNA fragments (often 10 kb–50+ kb)
- Uses ligation or transposase-based (rapid) kits
- Requires minimal DNA fragmentation — aim to preserve HMW DNA
- PCR-free options available (for unbiased sequencing)
- DNA input: ~200 ng – 1 µg (depending on kit)
- More sensitive to contaminants and shearing

Illumina - Nextera XT DNA Library Preparation Kit





https://emea.illumina.com/products/by-type/sequencing-kits/library-prep-kits/nextera-xt-dna.html

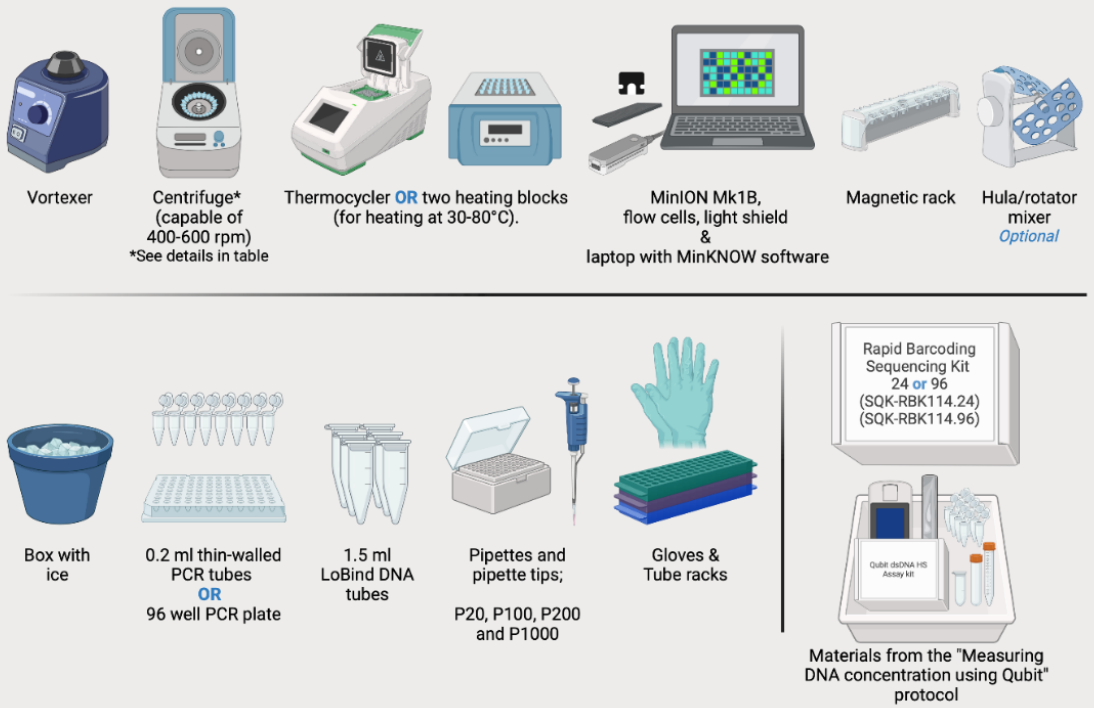
The Fleming Fund
Regional Grants

ONT library prep kits overview

	Output optimised	Speed optimised	Ultra-long reads optimised	Low input optimised	Targeted sequencing
	Ligation Sequencing Kit	Rapid Sequencing Kit	Ultra-long DNA Sequencing Kit	Rapid PCR Barcoding Kit	16S Barcoding Kit
Preparation time	60 minutes	10 minutes	200 mins + 1x O/N incubation	15 mins + PCR	25 mins + PCR
Input recommendation	~1000 ng gDNA or 100-200 fmol for amplicons	~200 ng gDNA or 50 ng for amplicons	6M cells	1 - 5 ng gDNA	10 ng gDNA
Fragmentation	Optional	Transposase-based	Transposase-based	Transposase-based	-
Amplification	No	No	No	Yes	Yes
Barcode options	Native Barcoding Kit 24 Native Barcoding Kit 96	Rapid Barcoding Kit 24 Rapid Barcoding Kit 96	-	24 plex	24 plex
Typical output	●●●	●●○	●●○	●●○	●●○
Adaptive sampling	Yes	Yes	Yes	Yes	-
Methylation included	Yes	Yes	Yes	-	-

ONT Rapid sequencing DNA V14 - barcoding (SQK-RBK114.24 or SQK-RBK114.96)

- This protocol uses genomic DNA
- For multiplexing 1-96 samples
- Library preparation time ~60 minutes
- High yield
- Fragmentation
- Compatible with R10.4.1 flow cells
- https://nanoporetech.com/document/rapid-sequencing-gdna-barcoding-sqk-rbk114



The Fleming Fund
Regional Grants

Common Wet-Lab Pitfalls

Issue	Potential Cause	Platform Most Affected
Low DNA yield	Incomplete lysis, low biomass	Both
DNA degradation	Repeated freeze-thaw, shearing	ONT/PacBio
Contamination	Poor aseptic technique	Both
Low purity	Salt or phenol carryover	ONT especially
Library prep failure	Incorrect ratios, pipetting error	Both
Adapter ligation issues	Dirty DNA ends, low molarity	ONT

The Fleming Fund
Regional Grants

Errors in Illumina Sequencing

- **Phasing and pre-phasing**: Some DNA strands lag or jump ahead during sequencing cycles, causing signal interference.
- **Signal bleed**: Fluorescence from neighboring clusters can overlap, especially with dense clustering.
- **GC bias**: GC-rich or AT-rich regions can be underrepresented due to amplification or sequencing inefficiencies.
- **Incomplete removal of fluorophores** can result in incorrect base calls.

The Fleming Fund
Regional Grants

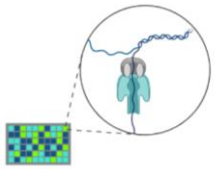
A note on ONT sequencing

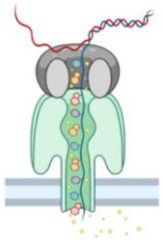
- Nanopore sequencing devices generate reads that reflect the lengths of the fragments loaded into the flow cell. To have control over the size of the fragments generated in the library prep, it is important to begin with high molecular weight (HMW) DNA.

- The shearing of HMW DNA can be minimised by:
  - Using wide-bore pipette tips to handle the gDNA
  - Mixing gently but thoroughly by flicking the tube, as opposed to vortexing or pipetting
  - Avoiding unnecessary freeze-thaw cycles
  - Avoiding pH <6 and >9
  - Avoiding high temperatures, which can lead to degradation

The Fleming Fund
Regional Grants

Why do errors occur for ONT?





- The pattern obtained from the nanopore needs to be interpreted.
  - Interpretation is based on machine learning models.
  - Signal varies depending on neighboring nucleotides, condition of pore (contamination, bubbles) and DNA speed variation (temperature issues).
- DNA string may slip in the pore during translocation.
- Short reads (<500 bp) tend to have worse quality and are often noise.
  - Longer reads provide better coverage but amplify homopolymer and repetitive regions issues.
- Stretches of homopolymers are difficult to call.
  - Repeated nucleotides (e.g., “TTTTT”) generate similar current disruptions.
  - Hard to distinguish between exact base counts (leading to insertions or deletions).
- Newer chemistry and updated basecalling models improves significantly on accuracy.

The Fleming Fund
Regional Grants

Examples of errors

Substitution Errors

- Incorrectly identifying one base as another (e.g., A → G).
    - Signal noise from current disruptions as DNA passes through the nanopore.
    - Inaccurate basecalling algorithms.
        - Affects gene sequences, especially for detecting SNPs and AMR mutations.

Insertion Errors

- Extra bases are called that do not exist in the actual sequence.
    - Signal misinterpretation due to homopolymer regions.
    - “AAAA” may be misread as “AAAAA.”

Deletion Errors

- Bases in the sequence are missed (e.g., a “C” is skipped).
    - Weak signal-to-noise ratio during strand passage.
    - Homopolymer stretches cause difficulty for nanopores.
        - Loss of genetic information; problematic for gene annotation and AMR detection.

The Fleming Fund
Regional Grants

Errors in downstream analysis

Downstream Analysis	Impact of Errors
Genome Assembly	Fragmented assemblies, reduced N50, inaccurate contigs.
Variant Calling	Misidentification of SNPs, insertions, or deletions.
AMR Detection	False positives or false negatives in AMR gene predictions.
Phylogenetics	Errors propagate into phylogenetic trees, misleading clusters.

The Fleming Fund
Regional Grants

Contaminants

- Chemical components used in DNA extraction methods can be carried over from the extraction procedure and contaminate the extracted DNA sample.
- These contaminants can have a significant effect on downstream library preparation efficiency, and therefore **sequencing throughput**.

The Fleming Fund
Regional Grants

Contaminants – Ethanol

- The presence of residual ethanol in extracted DNA can lead an overestimation of the concentration of DNA that is present in your sample.
- It can lead to a reduction in the measured A260/280 and A260/230 ratios.
- Sequencing runs can usually tolerate some ethanol contamination (10-20%) before adversely affected.

The Fleming Fund
Regional Grants

Contaminants - Isopropanol

- The presence of residual isopropanol in extracted DNA can lead an overestimation of the concentration of DNA that is present in your sample.
- Also, it can lead to a reduction in the measured A260/280 and A260/230 ratios.
- The presence of any % isopropanol might adversely affect sequencing performance

The Fleming Fund
Regional Grants

Contaminants - EDTA

- The presence of EDTA in extracted DNA can lead large perturbations in the nanodrop spectrum and A260/280 and A260/230 ratios.
- It can also lead to an overestimation of the concentration of DNA that is present in your sample.
- Up to 5 mM EDTA contamination can typically be tolerated before adverse affects

The Fleming Fund
Regional Grants

Contaminants - NaCl

- The presence of NaCl in extracted DNA does not appear to perturbation nanodrop spectra or A260/280 and A260/230 ratios.
- About 100 mM can typically be tolerated

The Fleming Fund
Regional Grants

Contaminants - Guanidinium chloride

- The presence of guanidinium chloride in extracted DNA can significantly perturb nanodrop spectra, particularly with respect to the A260/230 ratio.
- About 100 mM can typically be tolerated before adverse affects

The Fleming Fund
Regional Grants

Contaminants - Guanidinium isothiocyanate

- The presence of guanidinium isothiocyanate in extracted DNA can significantly perturb nanodrop spectra, giving atypical A260/280 and A260/230 ratios, and leading to mis-quantification of the DNA concentration.
- Up to 50 mM guanidinium isothiocyanate in the input sample can be tolerated

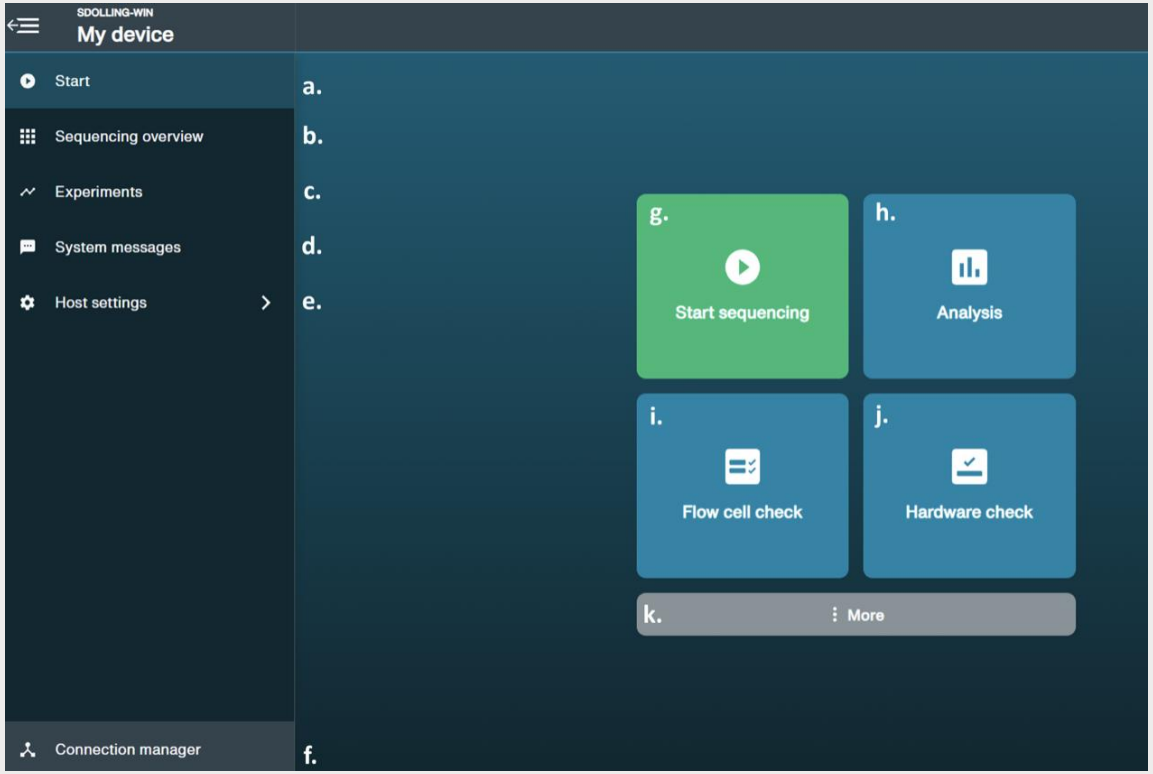
The Fleming Fund
Regional Grants

Contaminants - Phenol

- The presence of residual phenol in extracted DNA can significantly perturb nanodrop spectra, giving atypical A260/280 and A260/230 ratios, and lead to overestimation of the DNA concentration.
- Up to 1% phenol in the input sample can be tolerated

The Fleming Fund
Regional Grants

ONT sequencing – MinKnow Interface



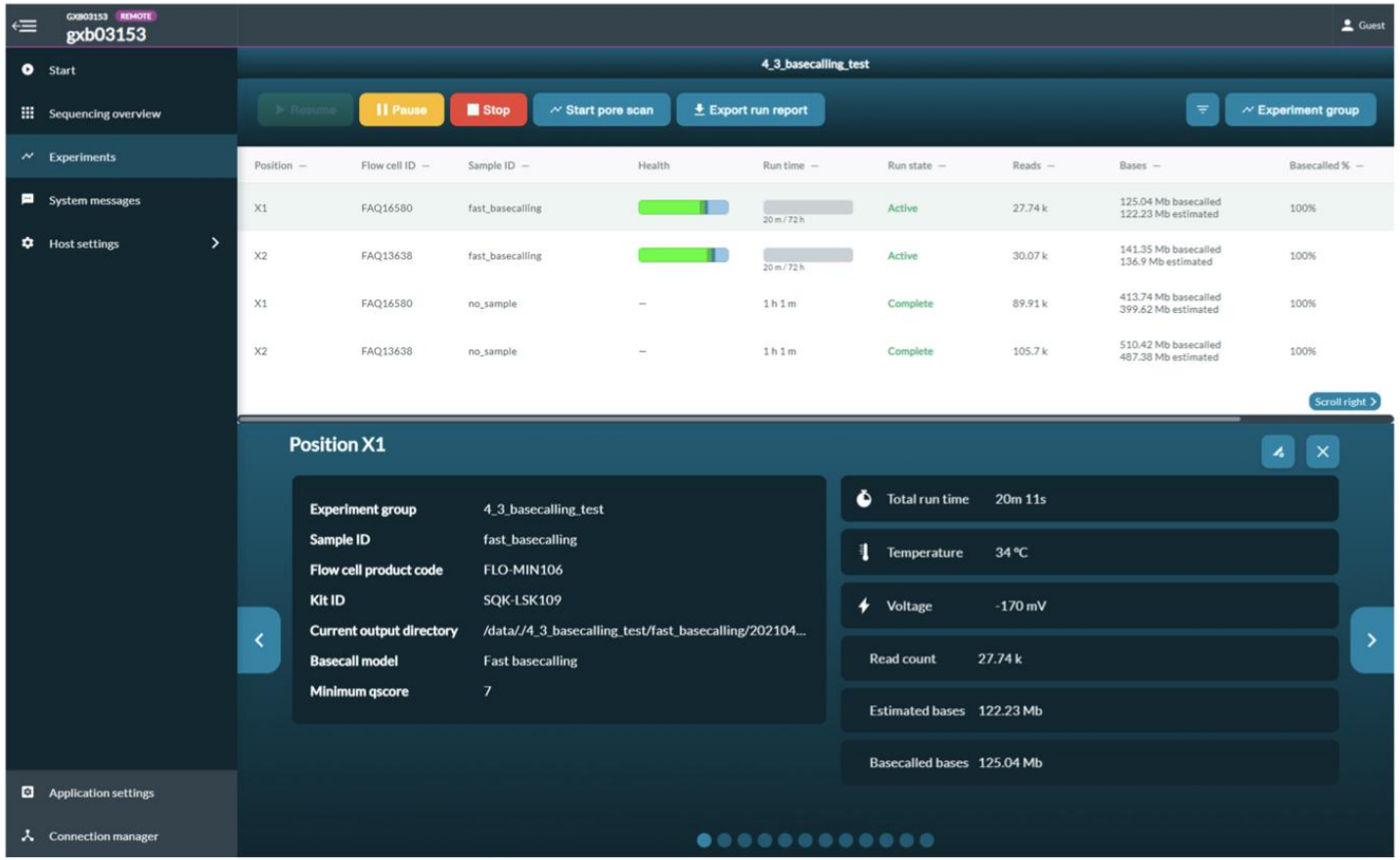
Homepage overview

The MinKNOW Homepage enables the user to navigate to:

a. **Start** homepage
b. **Sequencing Overview** of connected flow cells
c. Recent and current **Experiments**
d. **System Messages**
e. **Host Settings**
f. **Connection Manager** to connect with other available devices
g. **Start Sequencing** experiment
h. Post-run **Analysis**
i. **Flow Cell Check**
j. **Hardware Check**
k. **More** includes option to generate .mmi from .fasta file or to import a sample sheet
l. **Guest/initials** to logout

The Fleming Fund
Regional Grants

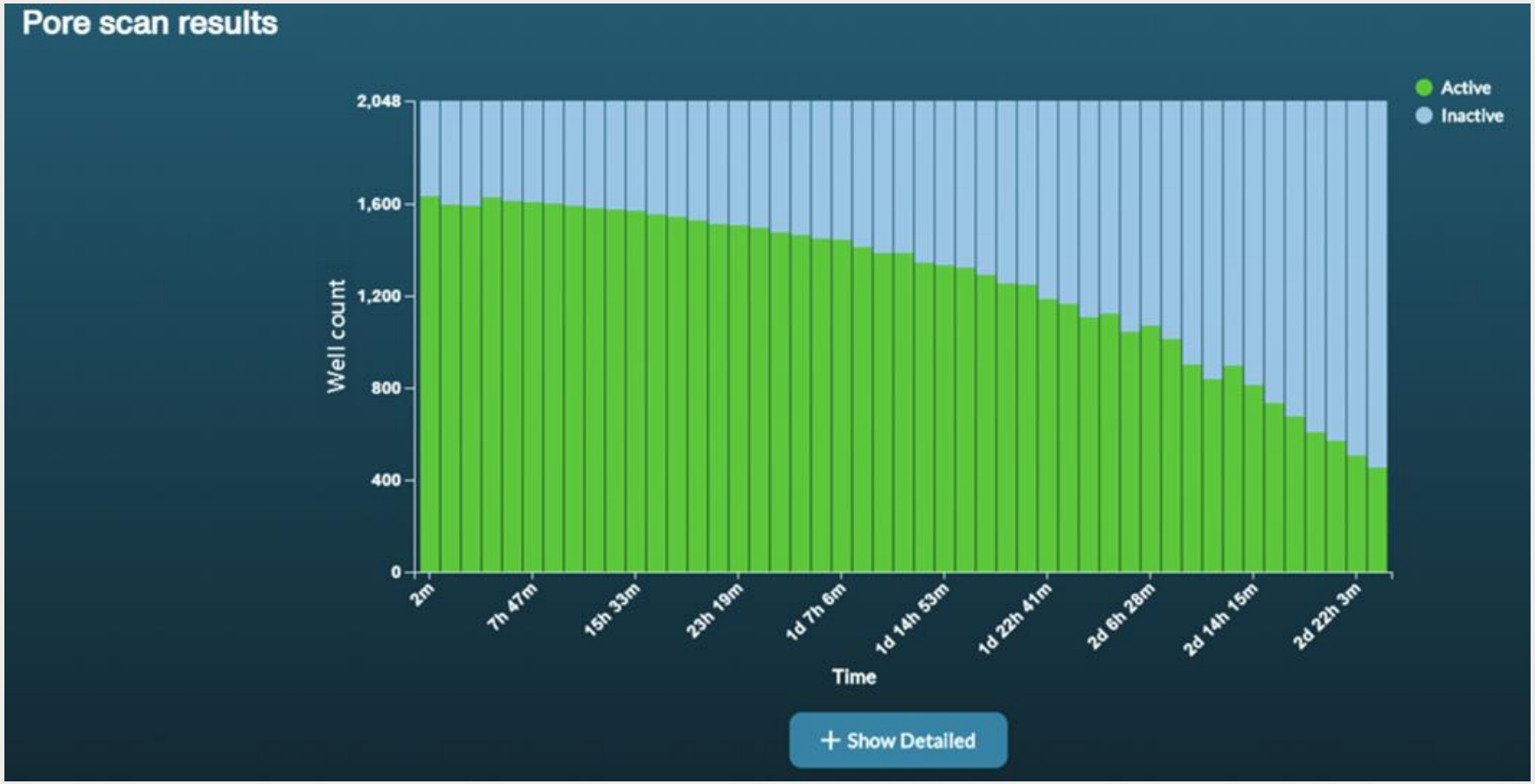
Experiment Summary Information



- Minknow will basecall and demultiplex live
- Real time information on flow cell health and sequencing

The Fleming Fund
Regional Grants

Pore scan

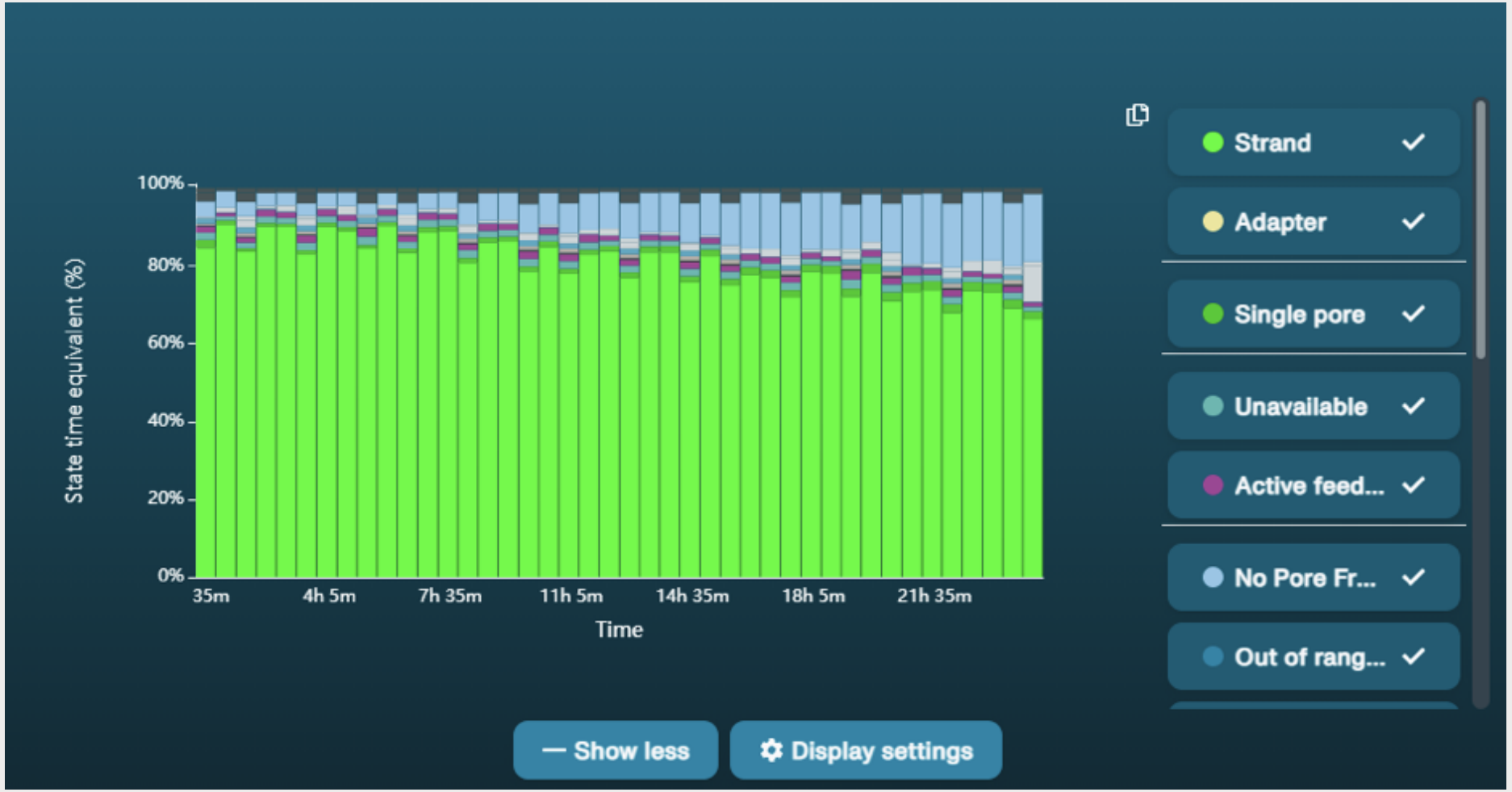


The Fleming Fund
Regional Grants

Pore Occupancy

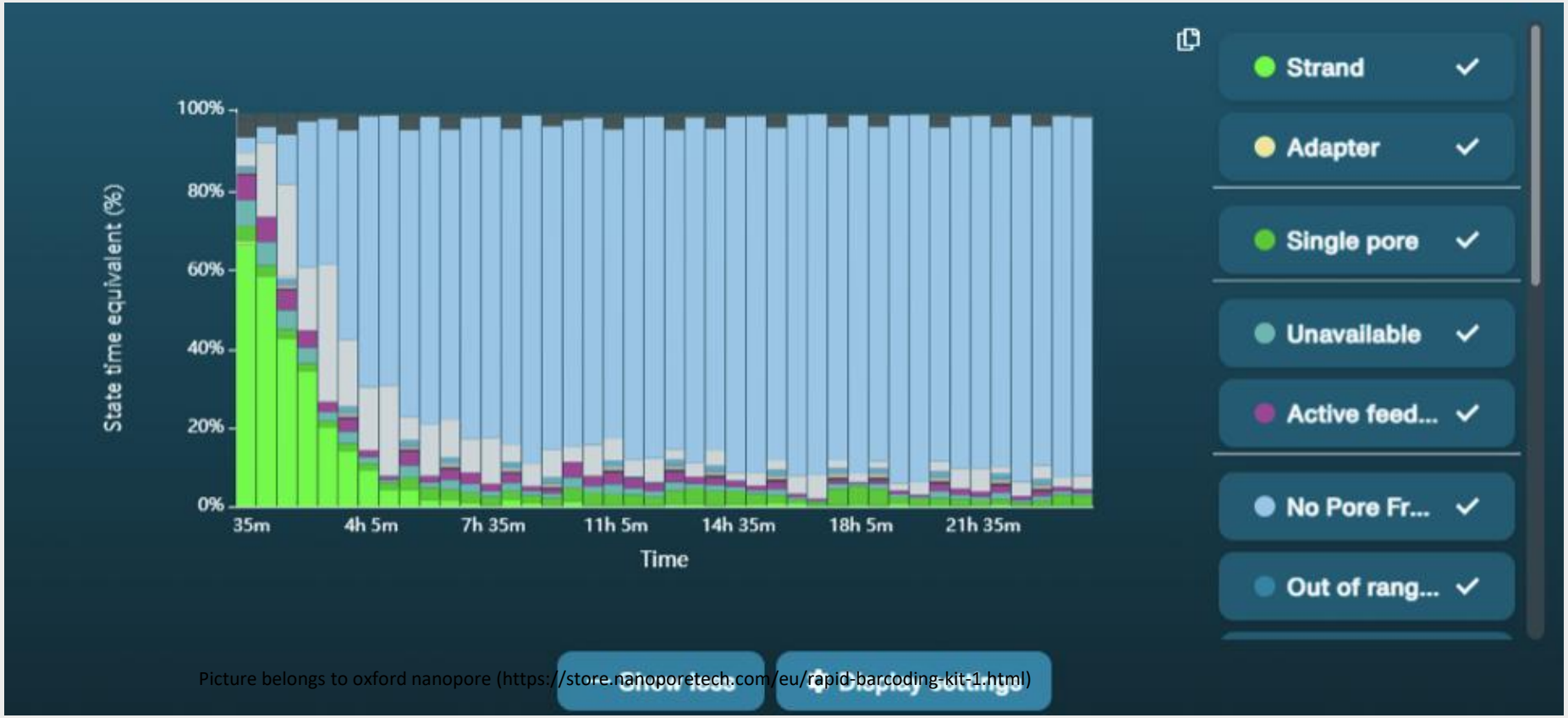
The Fleming Fund
Regional Grants

Good library



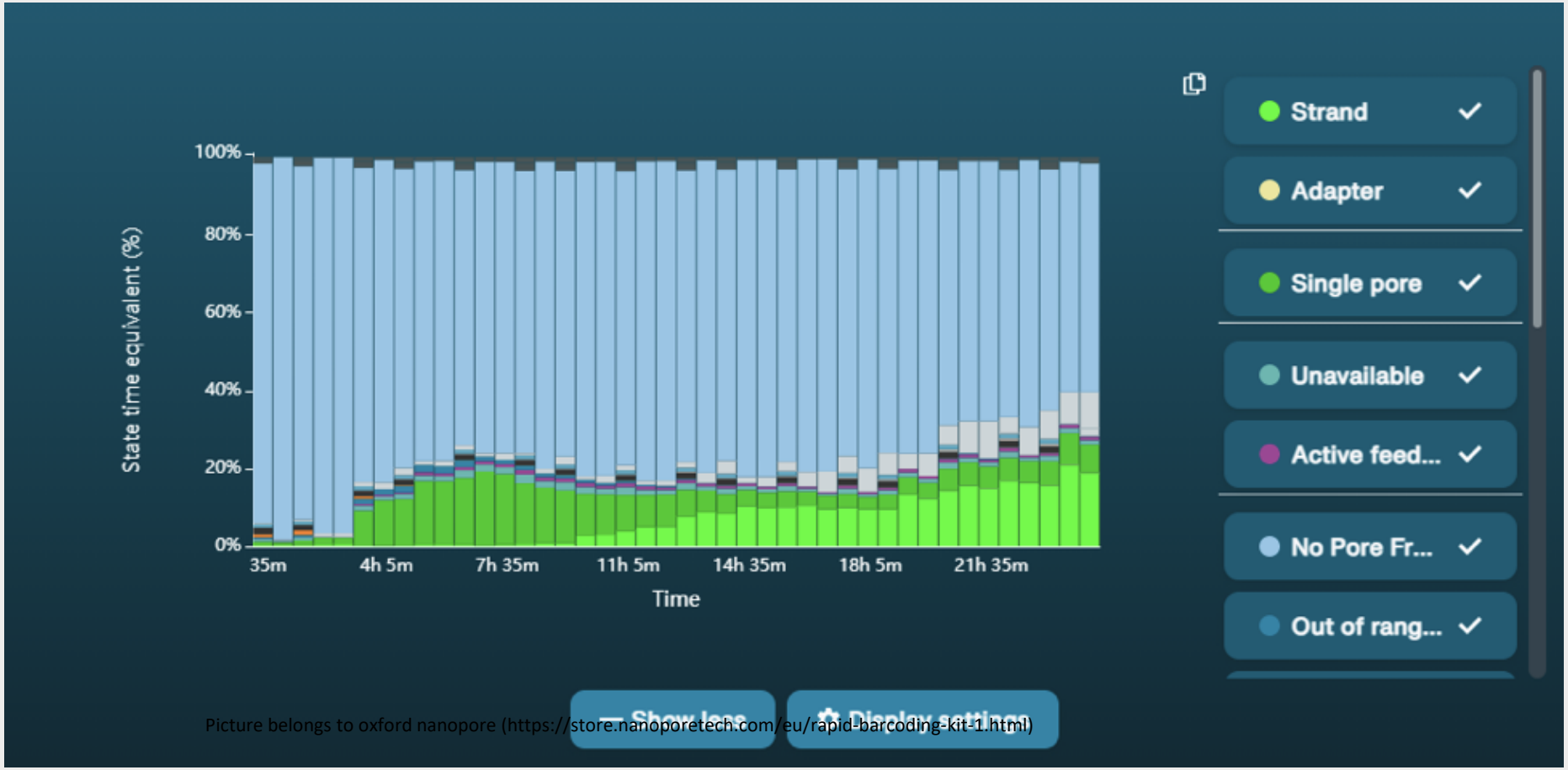
Picture belongs to oxford nanopore (https://store.nanoporetech.com/eu/rapid-barcoding-kit-1.html)

Channel Blocking



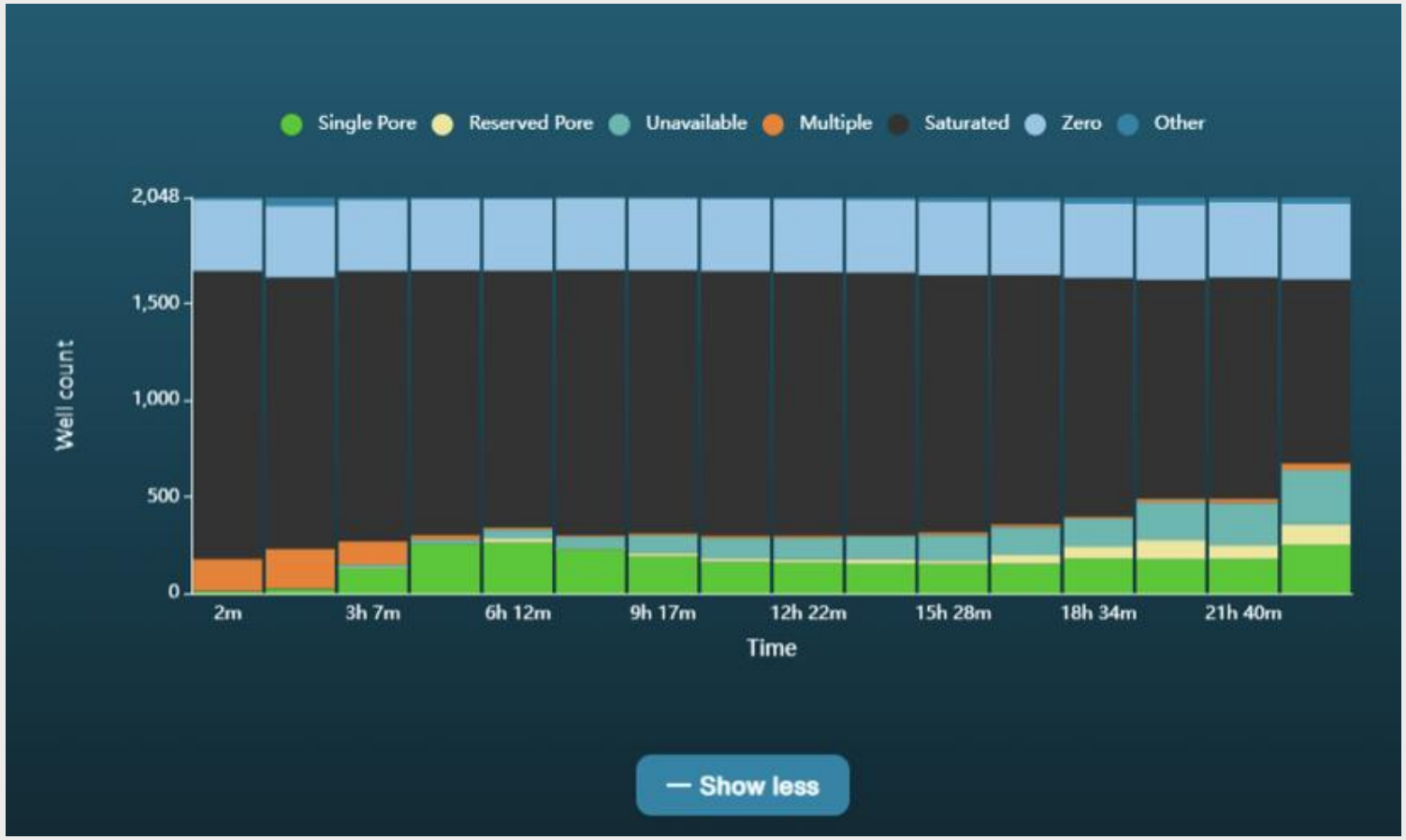
Picture belongs to oxford nanopore (https://store.nanoporetech.com/eu/rapid-barcoding-kit-1.html)

Osmotic Imbalance



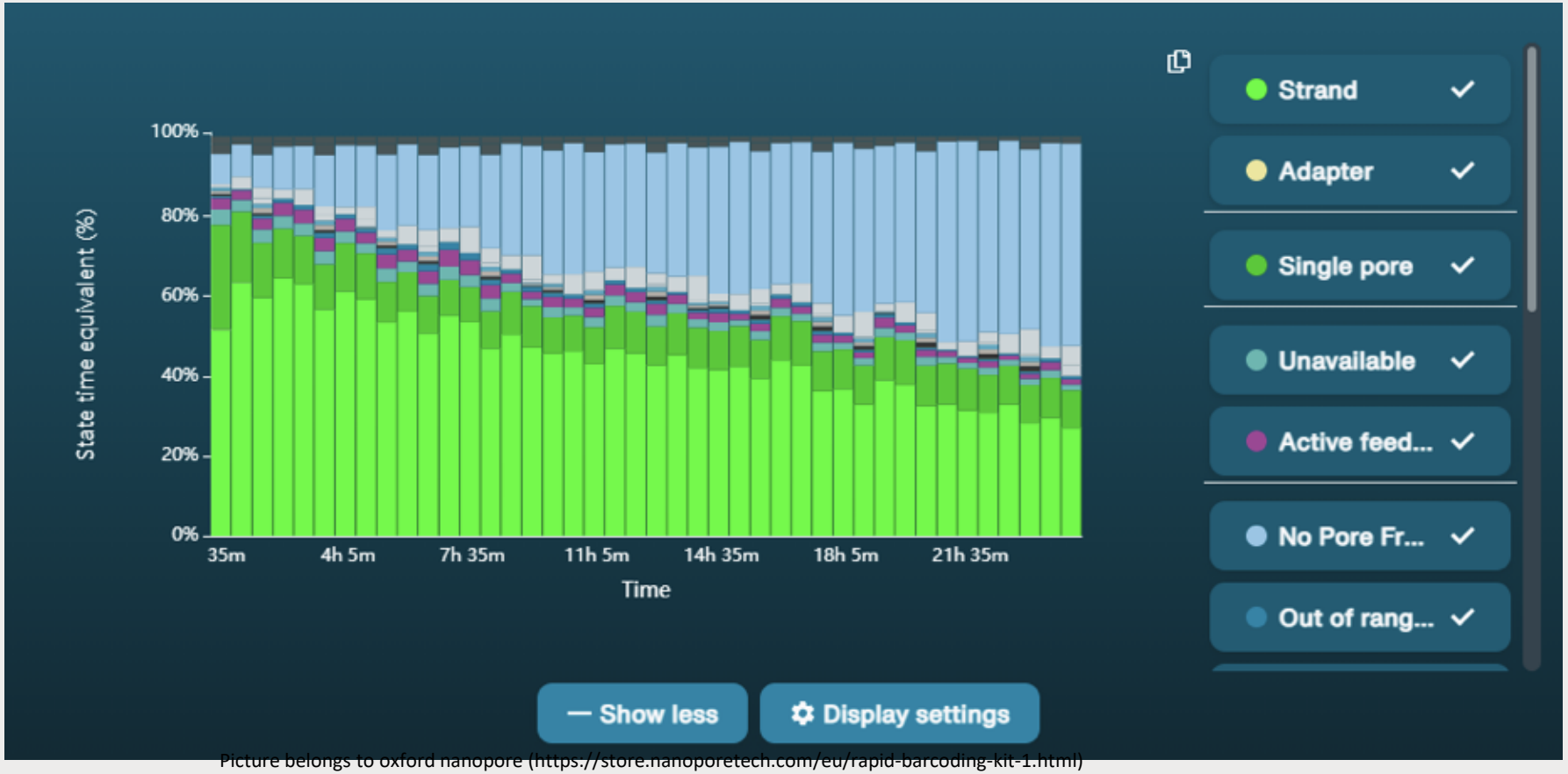
Picture belongs to oxford nanopore (https://store.nanoporetech.com/eu/rapid-barcoding-kit-1.html)

Osmotic Imbalance – channel scan



Picture belongs to oxford nanopore (https://store.nanoporetech.com/eu/rapid-barcoding-kit-1.html)

Low Pore Occupancy



Picture belongs to oxford nanopore (https://store.nanoporetech.com/eu/rapid-barcoding-kit-1.html)

The Fleming Fund
Regional Grants

Troubleshooting Tips

- **Low yield?**
  → Use mechanical lysis for tough organisms; increase biomass
- **Low purity?**
  → Do ethanol washes carefully, avoid carryover, repurify
- **Poor library prep?**
  → Check kit expiry, bead cleanup efficiency, and pipetting accuracy
- **Long-read issues?**
  → Avoid DNA shearing, do not vortex, use wide-bore tips
- **Batch failures?**
  → Check for contamination in shared reagents or barcodes

Tips:

- Use extraction and library controls in every batch
- Document and trace every deviation in SOPs

The Fleming Fund
Regional Grants

Key Takeaways

- Wet-lab steps are critical for producing high-quality, interpretable WGS data
- Know the strengths and requirements of short- and long-read sequencing
- DNA quality and purity determine success downstream — QC is not optional
- Library prep demands attention to detail, especially for long-read platforms
- Prevent and respond to wet-lab pitfalls with good documentation and controls

The Fleming Fund
Regional Grants

Thank you

Questions? ☺

This programme is being funded by the UK Department of Health and Social Care.
The views expressed do not necessarily reflect the UK Government's official policies.