

Microbial typing Using Oxford Nnopore technology sequencing

Yaovi Mahuton Gildas Hounmanou, PhD

Assistant-Professor

APTI-Fellow, National Institutes of Health, MD, USA

<https://orcid.org/0000-0003-3991-0864>

University of Copenhagen

Faculty of Health and Medical Sciences

Section for Food Safety and Zoonoses

Department of Veterinary and Animal Sciences

Stigbøjlen 4, 1870 Frederiksberg, DK

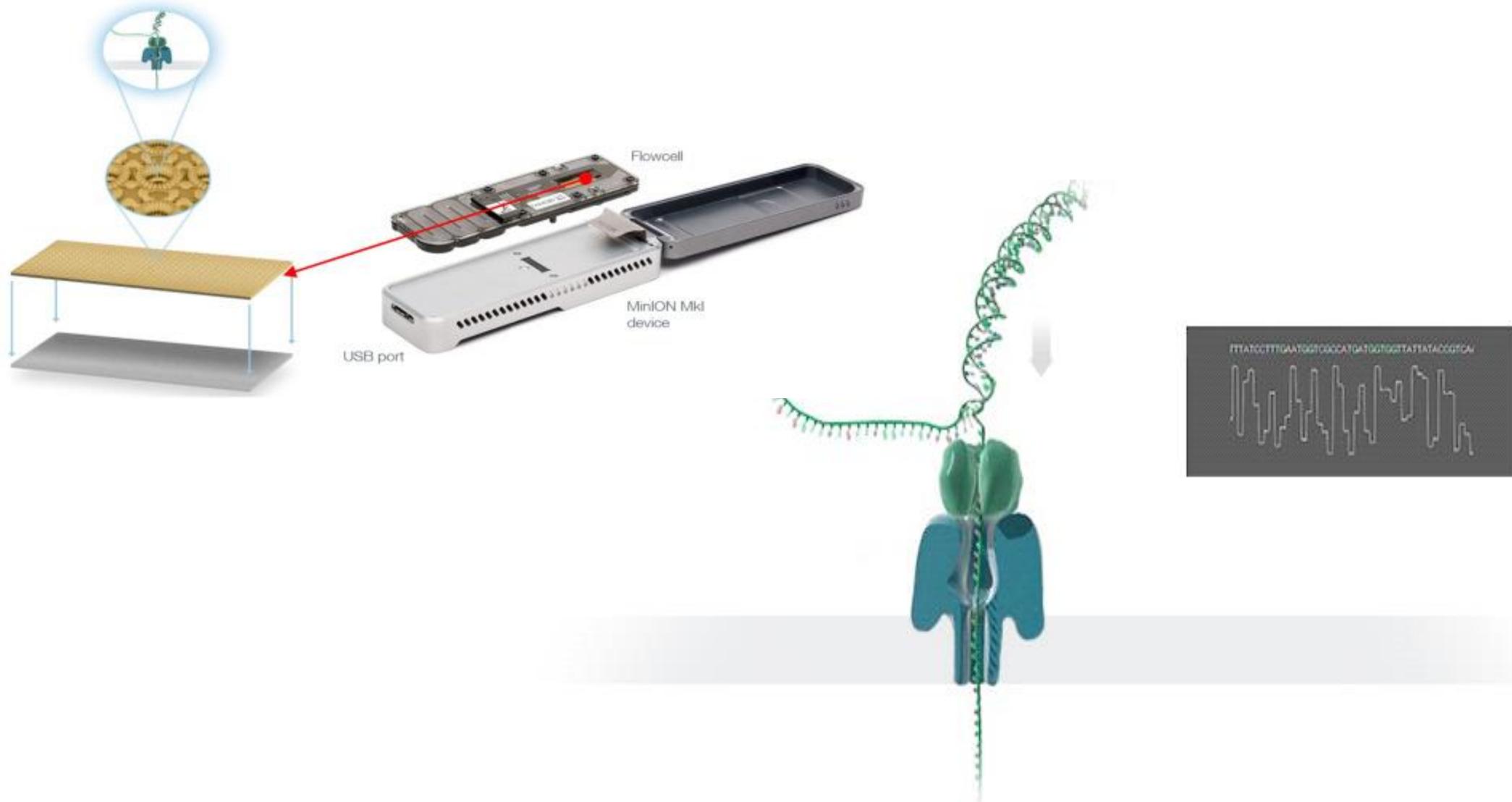
Cell Phone-WhatsApp: +4552641293 / +1 2402343714

Work phone: +4535326882

Email: gil@sund.ku.dk ; gilmahu@yahoo.fr

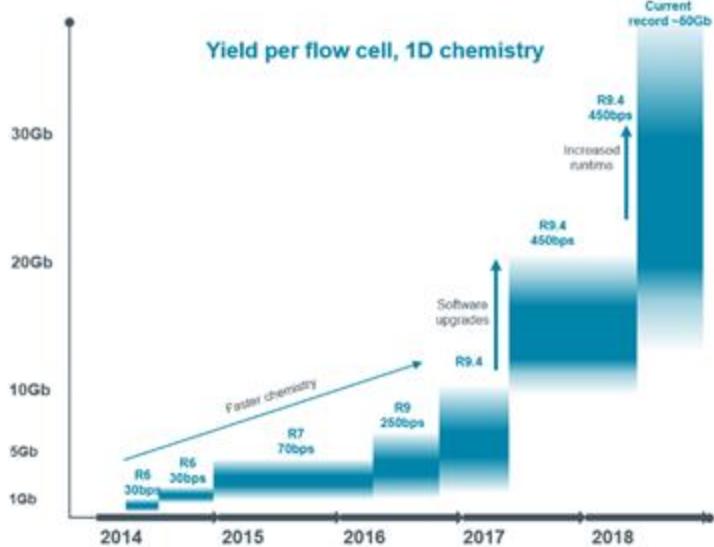


Sequencing mechanism

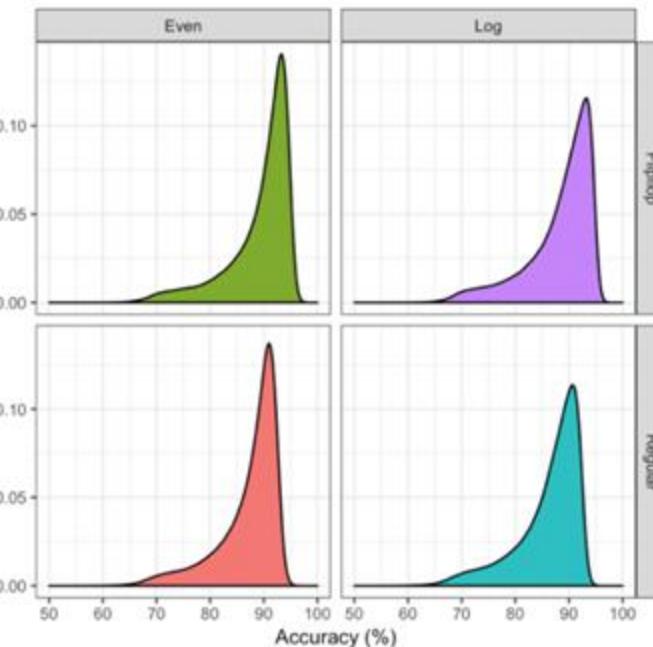


Nanopore sequencing - performance improvement

Throughput

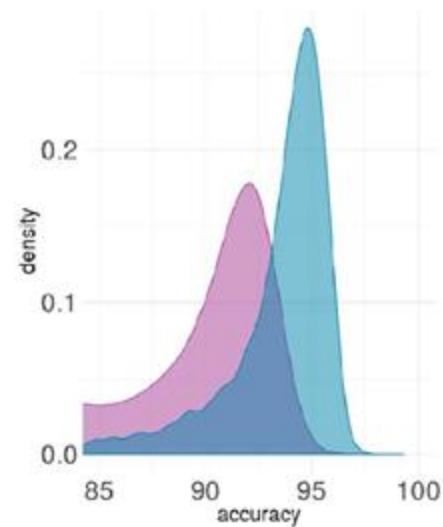


Software improvement



Quality

Flowcell improvement



Basecalling
with flip-flop

Basecalling
without flip-flop

Nanopore sequencing

MinION



Portable, USB powered
biological analysis

512* of 2048 channels



GridION



Five flow cells and
integrated computing

$$5 \times 512/2048 = 2,560^*$$

PromethION



High-throughput, versatile
benchtop system

$$48 \times 3,000^* = 144,000$$

Sequencing platforms

GridION and Flongle



Typical throughput: up to 1 Gbp

Best solution for:

- Viral genomes
- Single bacterial genome
- Amplicons (e.g. SARS-CoV 2 whole genome)

Sequencing platforms

GridION and R10 FlowCell

Typical throughput: 10 to 20 Gbp



- Best performance when:**
- Up to 12 bacterial metagenome samples for taxonomy identification
 - 1-2 metagenomic samples for pathway analysis
 - Up to 2-3 fungal genome per single flow cell (depending on genome size)

Sequencing platforms

GridION and R10 FlowCell



Best solution for:

- Small and medium size genomes
(bacteria, fungi)
- Multiplexed amplicons (e.g. 96 16S rRNA samples)

Sequencing platforms

PromethION

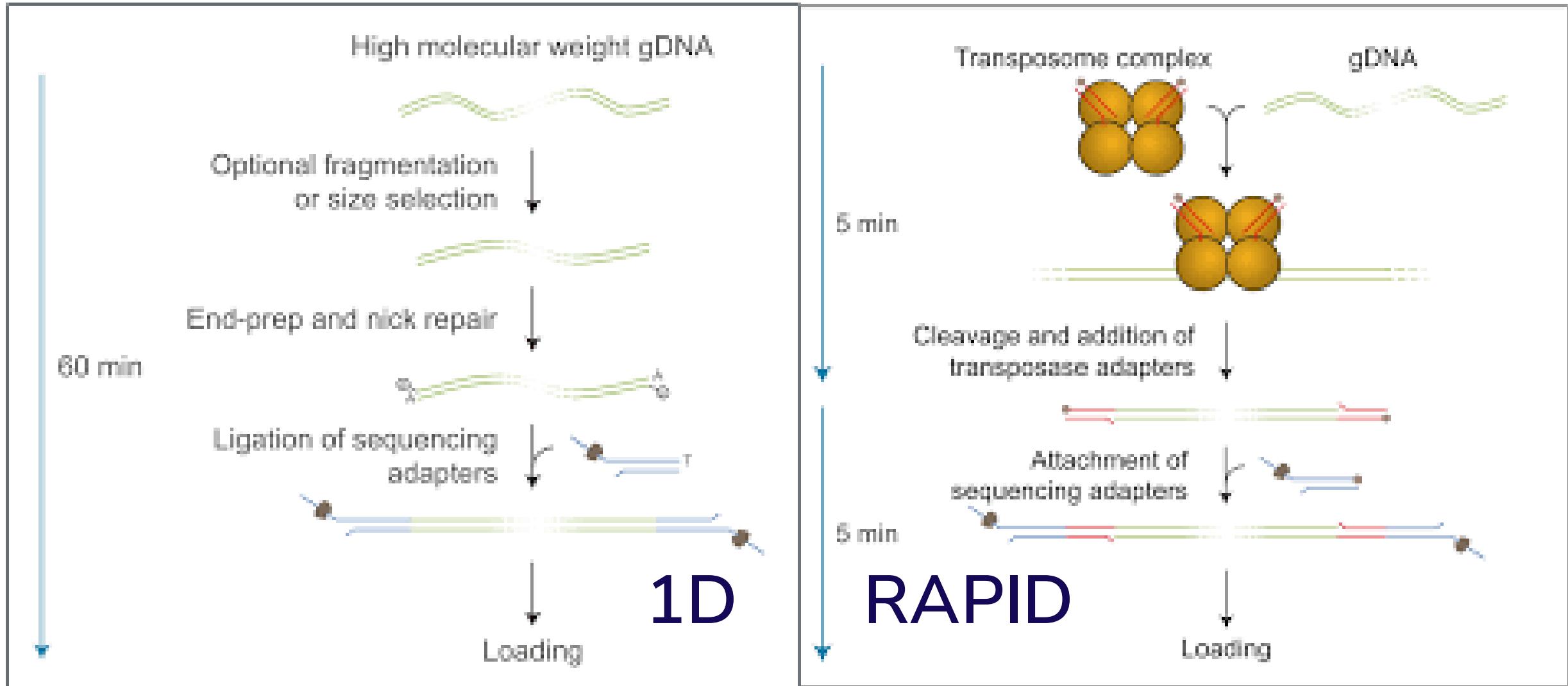
Typical throughput between 40 to 120 Gbp



Best solution for:

- Large genome sequencing projects (plant, animal, human)
- Deep sequencing of multiple metagenomic samples

Sequencing kits



Nanopore sequencing

SEQUENCING TRADE-OFFS

"You can't always get what you want (...)"

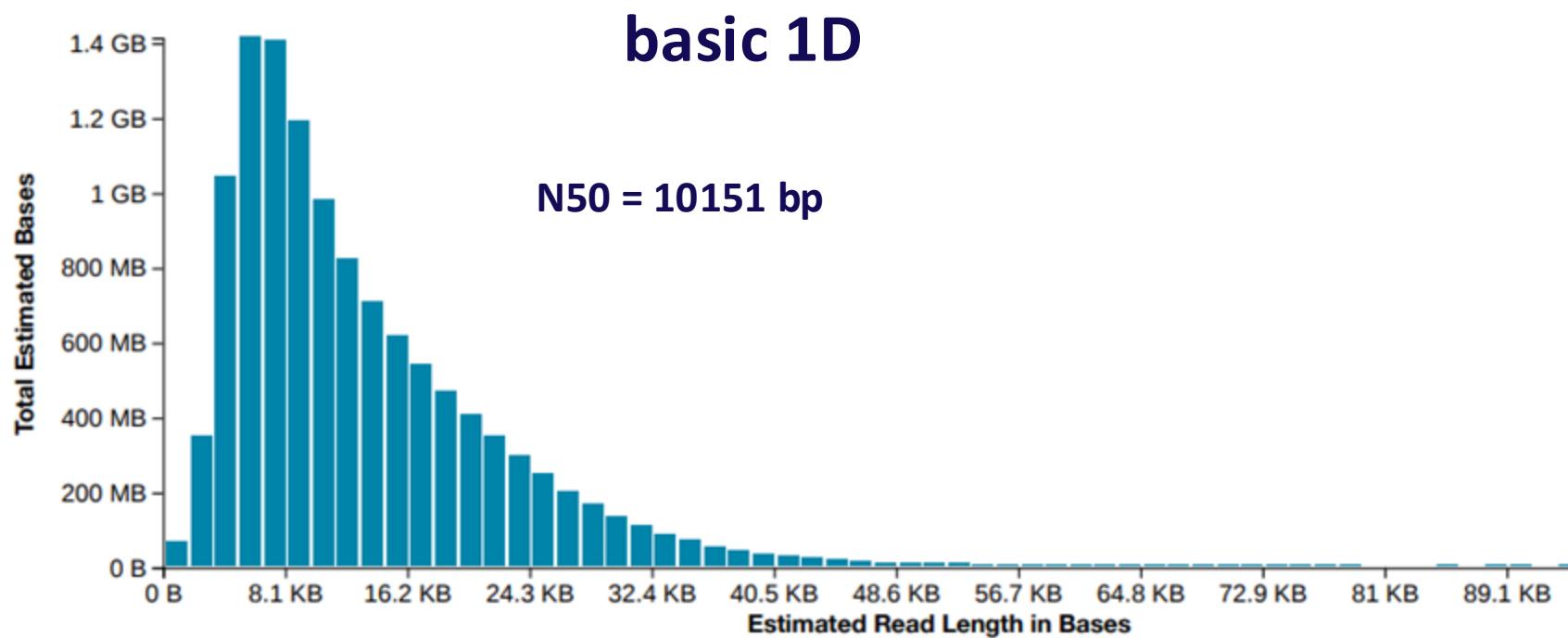
Three basic types of nanopore sequencing:

<u>Library type</u>	<u>reads length</u>	<u>advantage</u>
SQK-LSK109 – Genomic/Amplicons by Ligation	depending on DNA fragments length	versatility
SQK-LSK109 – Genomic by Ligation plus fragmentation	controlled by fragmentation parameters	throughput
SQK-RAD004/RBK004 – Rapid (Barcode) Sequencing	depending on DNA fragments length	simplicity

Nanopore sequencing

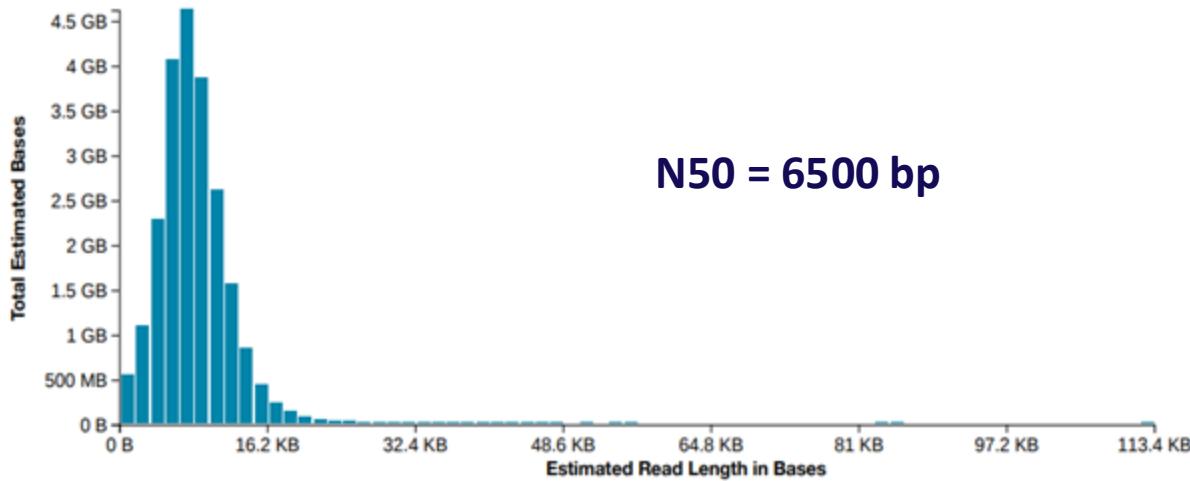
SEQUENCING TRADE-OFFS

"You can't always get what you want (...)"

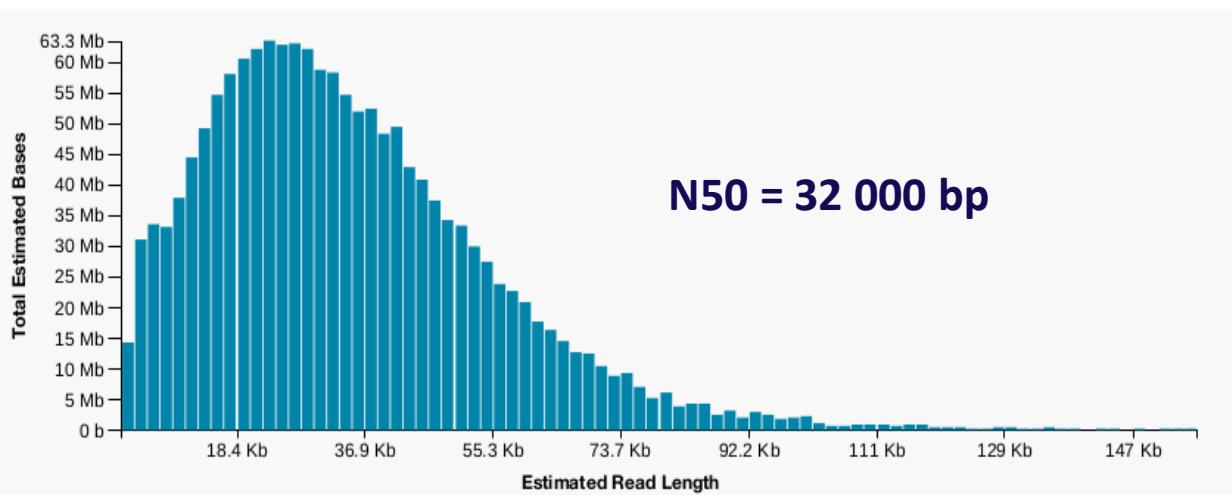


Nanopore sequencing

SEQUENCING TRADE-OFFS



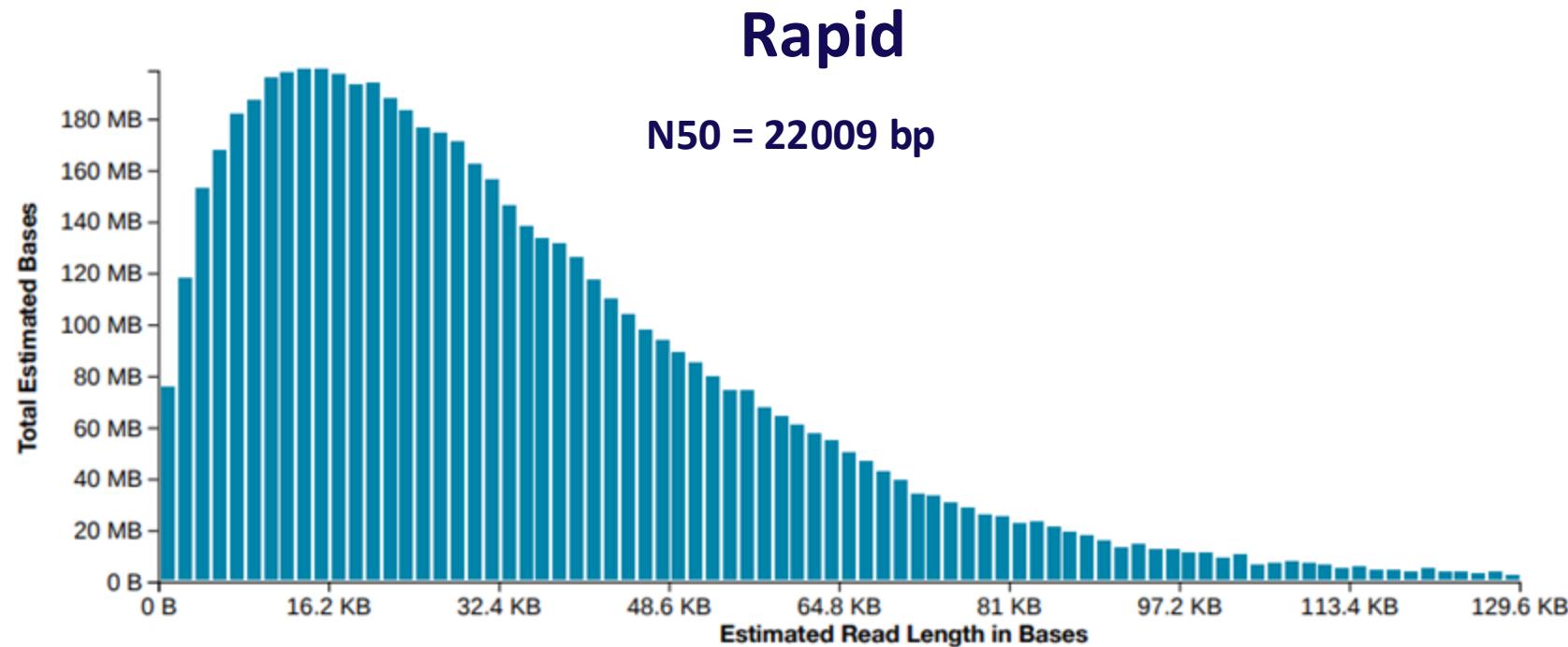
fragment + 1D



Nanopore sequencing

SEQUENCING TRADE-OFFS

"You can't always get what you want (...)"



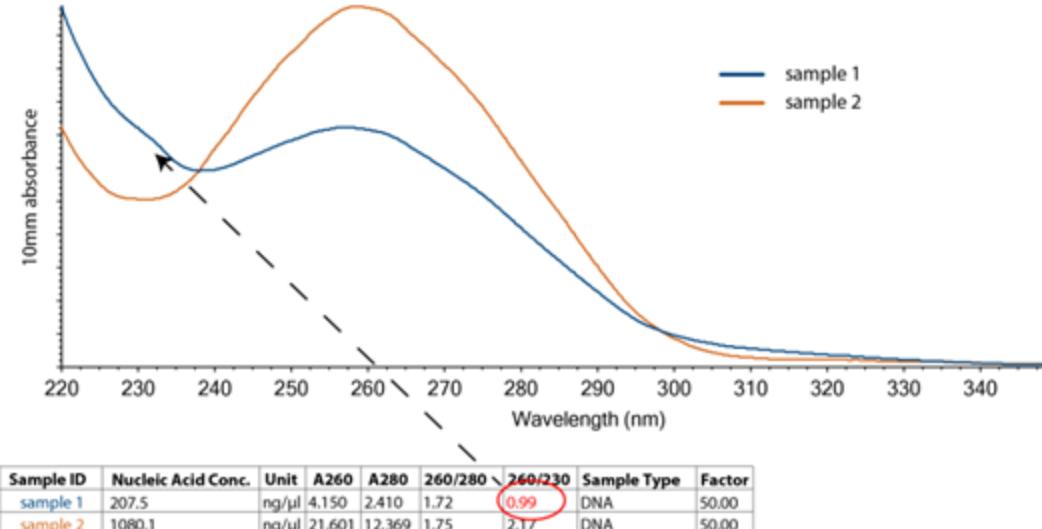
Nanopore sequencing – library preparation

Golden rules – library prep

1. use high quality DNA isolate

a) quality assessment by Nanodrop:

- 260/280 ~ 1.80
- 260/230 ~2.0-2.2



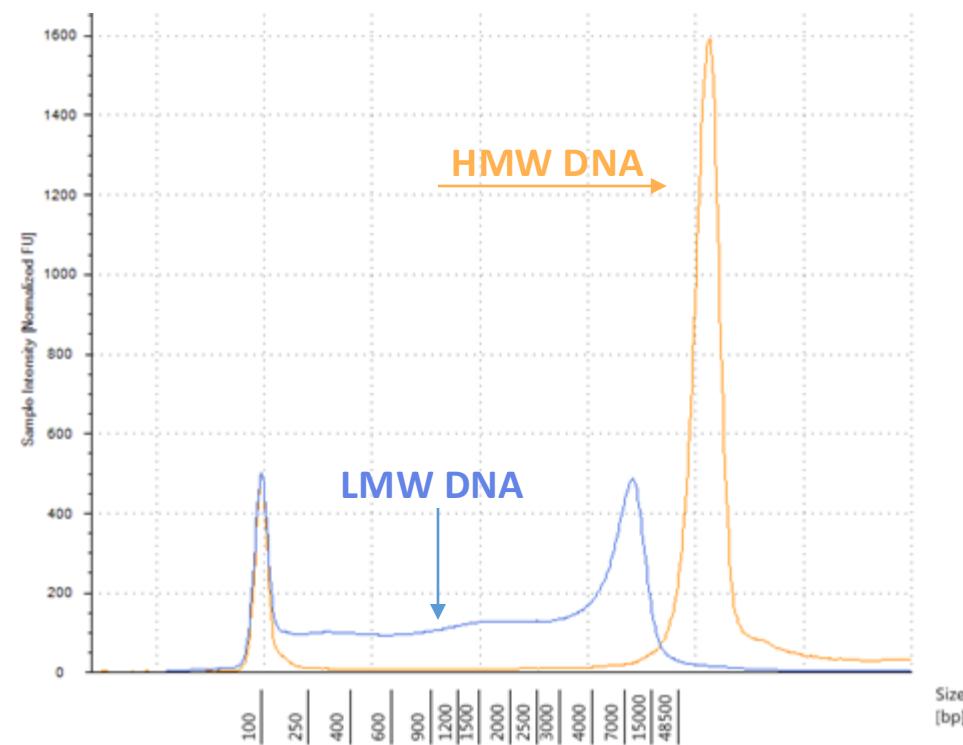
Nanopore sequencing – library preparation

Golden rules – library prep

1. use high quality DNA isolate

b) molecular weight distribution:

- TapeStation or Bioanalyzer
- FEMTO Pulse
- PFGE



Nanopore sequencing – library preparation

Golden rules – library prep

2. 0.2 pM ligation rule

adjust DNA input to 0.2 pM:

average fragment length = 2 kb – 260 ng

average fragment length = 10 kb – 1300 ng

average fragment length = 50 kb – 6600 ng

Use Biomath Calculators to determine molarity

Nanopore sequencing – library preparation and loading

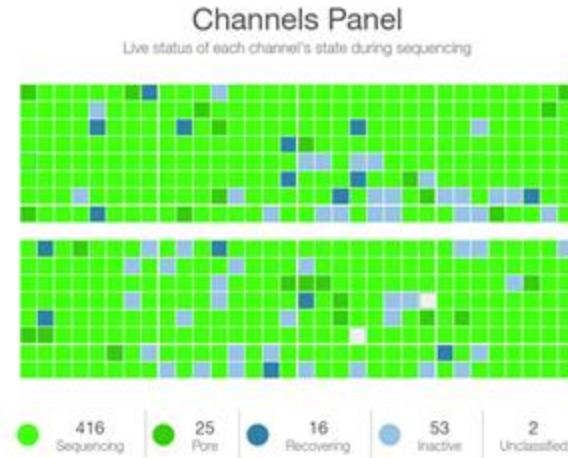
Golden rules – library prep and flowcell loading

3. avoid flowcell overload:

R9.4.1 pore - 5-50 fmol of good quality library load into the flow cell

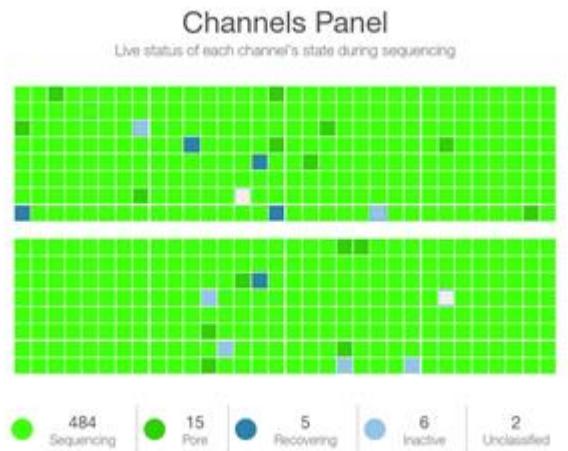
4. avoid introduction of air bubbles during flowcell priming

Nanopore sequencing – good library examples



$$\text{occupancy} = \frac{\text{Strand}}{\text{Strand} + \text{Single Pore}}$$

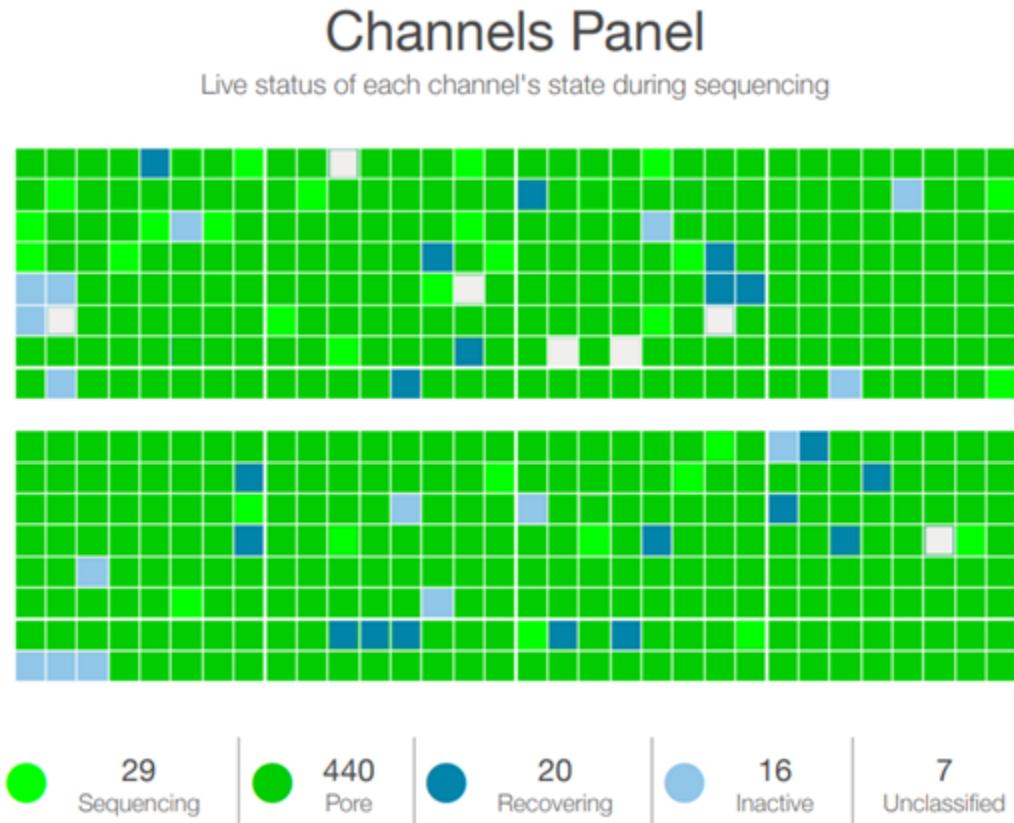
Good libraries – occupancy >80%



Excellent libraries – occupancy >95%

Nanopore sequencing – troubleshooting

Low pore occupancy

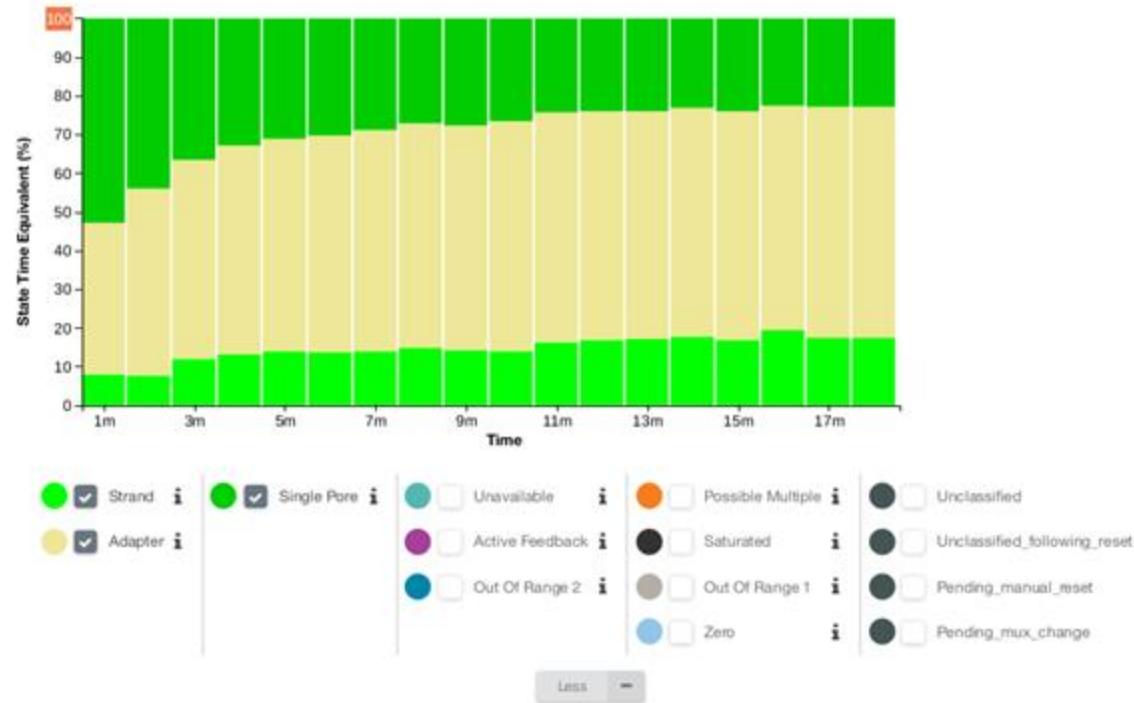


Cause:

- low number of threadable ends
- low DNA input
- poor ligation efficiency

Nanopore sequencing – troubleshooting

High level of adaptor state

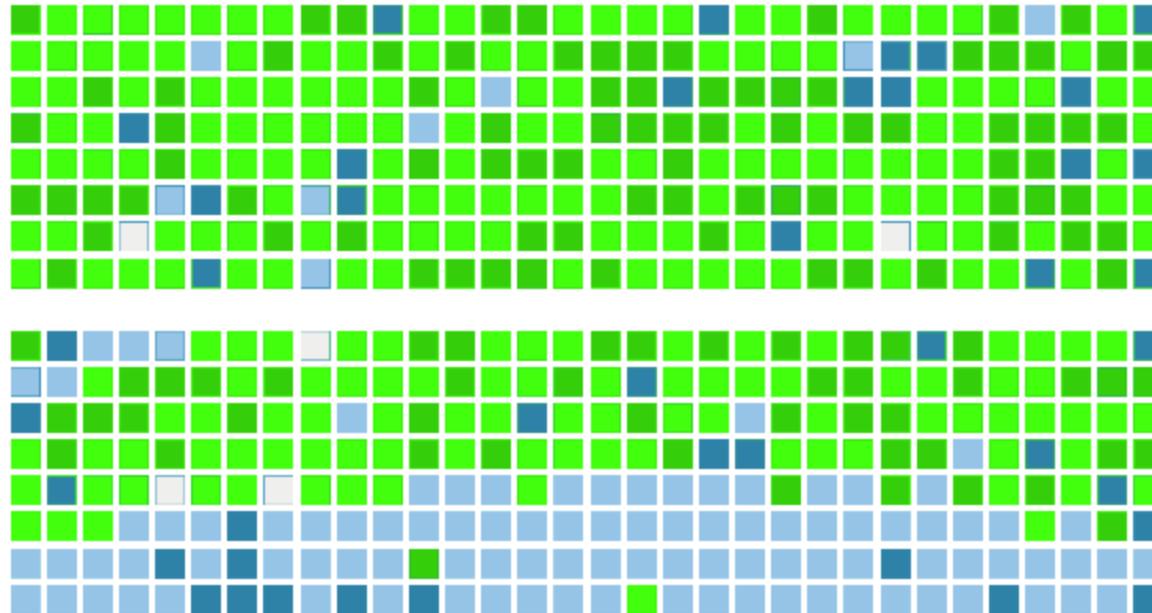


Cause:

- poor ligation efficiency

Nanopore sequencing – troubleshooting

Flow cell failure



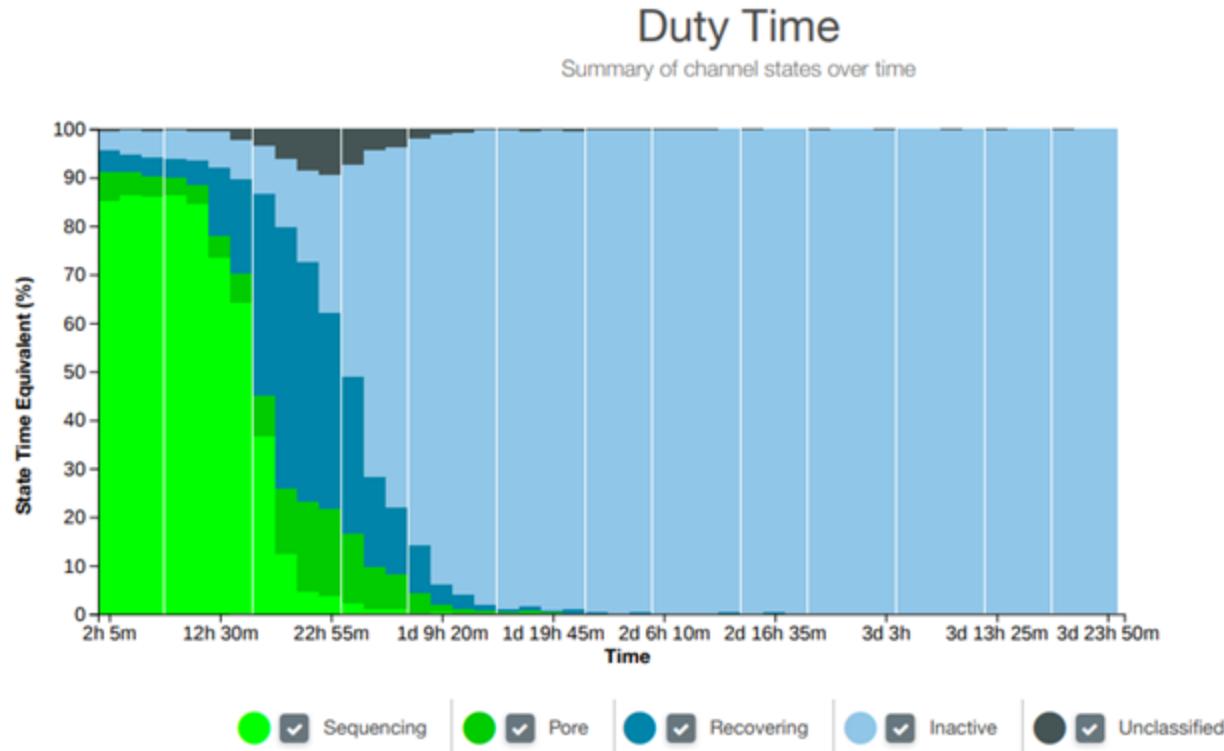
Cause:

- Introduction of bubbles
- Contamination in sample



Nanopore sequencing – troubleshooting

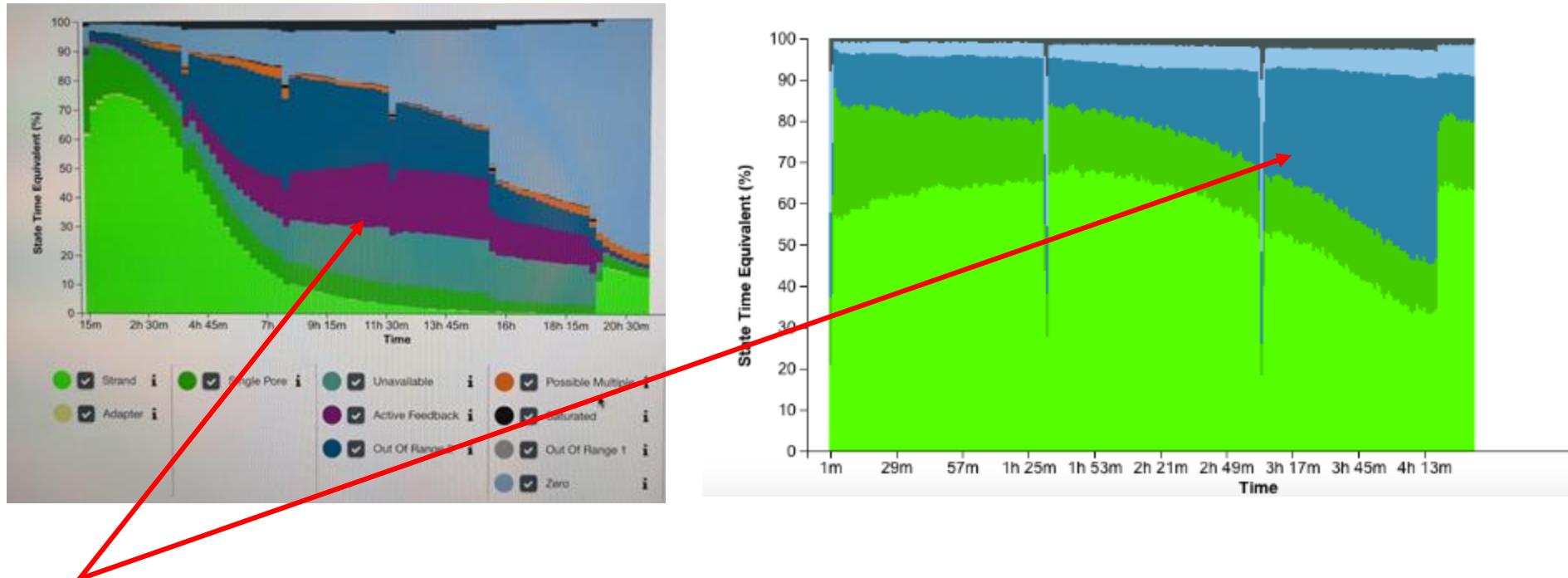
Flowcell overload



- Cause:
- depletion of ATP during sequencing run due to library overload

Nanopore sequencing – troubleshooting

Flowcell overload - refueling

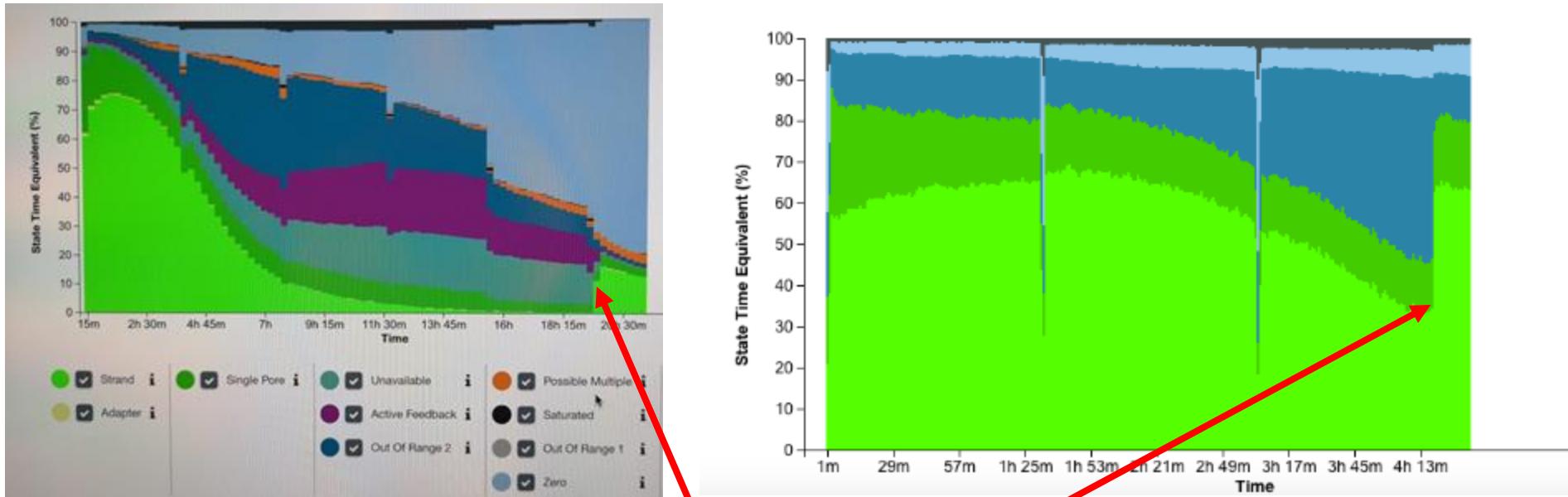


Problem: FLOWCELL OVERLOAD

(increasing RECOVERY STATE [unavailable, active feedback, and out of range 2] over-time)

Nanopore sequencing – troubleshooting

Flowcell overload - refueling



Solution: „refueling” – add 200µL FB buffer

DNA quality issues

- Optimum requirement 50 ng/ μ l in 10 μ l
 - A₂₆₀/280 – 1.8
- Hard to get with composite samples
- Several promising kits however most involve multiple cycles of centrifugation leading to DNA fragmentation
- Automated systems could be one solution ...



Gravity kit, potentially most adequate...



PRODUCTS ABOUT US CONTACT US

Search  POLSKI CURRENCY MY ACCOUNT 0.00 €

Home Page > Products > DNA purification > Genomic DNA purification

CORONAVIRUS SARS-CoV-2 2
NEW PRODUCTS 2
DNA purification 74
DNA cleanup after enzymatic reactions 12
Plasmid DNA purification 7
Genomic DNA purification 53
Bacteria 17
Blood 11
Body fluids 3
Feces 3
Food 2
Materials with a trace content of DNA 3
Plants 6
Soil 2
Swabs 4
Tissue and cell cultures 14
Viruses and phages 3
Yeast and fungi 3
DNA purification reagents 2
RNA purification 22
PCR and reverse transcription 45
Cloning and transformation 10
Methylation 2
Enzymes 24
Protein analysis 1
Molecular weight markers 11
Molecular biology reagents 10
Laboratory equipment 2



Bead-Beat Micro AX Gravity

Versatile, increased efficiency kit for genomic DNA purification from various sources. Mechanical lysis.

Sample size: up to 2 ml of bacterial culture, up to 2 ml of fluid or 100 mg of liquid yeast culture, up to 100 mg of plant tissue, up to 20 mg of animal tissue, up to 100 mg of faeces, up to 200 mg of environmental sample up to 50 mg of other biological material.

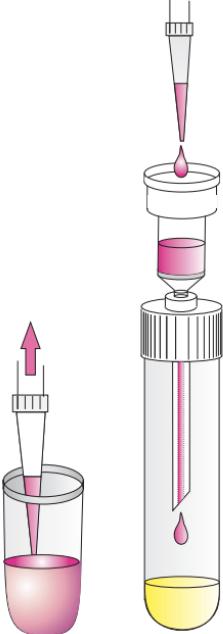
product code	size	price	quantity	ADD TO CART	MANUAL	MSDS
106-20	20 isolations	56.40 €	1			
106-100	100 isolations	225.60 €	1			



Blood Mini

Kit for genomic DNA purification from blood. Sample size: up to 1000 µl of fresh or frozen blood.

product code	size	price	quantity	ADD TO CART	MANUAL	MSDS
022-50	50 isolations	92.16 €	1			
022-250	250 isolations	368.64 €	1			





Fast DNA Plant Screen

Kit for rapid isolation of genomic DNA from plant material, to be used in PCR. Sample size: fragment of the plant tissue with a diameter of 2 - 3 mm.

product code	size	price	quantity	ADD TO CART	MANUAL	MSDS
050-192	192 isolations	186.00 €	1			

results.txt ^ Show all X

Kits, protocols, devices: <https://nanoporetech.com/> RBK, LSK,..

The screenshot shows a dual-monitor setup displaying the Oxford Nanopore Technologies website. The left monitor displays a dark-themed sidebar with navigation links for 'HOME / Sam', 'Devices', 'Consumables' (Flow cells, Sample prep, Legacy), 'Training' (MinION Rapid start, GridION Advanced, PromethION Advanced), and 'Services'. The right monitor displays the 'Flow cells' product page. The header includes the Oxford Nanopore logo, a search bar, and links for 'Products & Services', 'Applications', 'Resources', 'Investors', 'Careers', 'Store', and 'News'. A large banner image shows several flow cells labeled 'C', 'D', and 'E'. Below the banner, sections for 'MinION & GridION' and 'PromethION' are shown, each featuring an image of the respective flow cell and purchase options. The 'MinION & GridION' section indicates a price of €430.00 and a status of 'Fully released'. The 'PromethION' section indicates availability as packs containing 4 flow cells. A user profile icon is visible in the bottom right corner.

store.nanoporetech.com/eu/sample-prep.html

store.nanoporetech.com/eu/flow-cells.html

HOME / Sam

Devices

Consumables

- Flow cells
- Sample prep
- Legacy

Training

- MinION Rapid start
- GridION Advanced
- PromethION Advanced

Services

MinION & GridION

The MinION and GridION Flow Cell contains up to 512 nanopore channels for sequencing DNA or RNA in real time.

Purchase options:

- Buy single flow cells or save money by buying in bulk
- Available as R9.4.1, R10.4 and our latest chemistry, R10.4.1

Flow cells from: €430.00 [Buy >](#)

Fully released

PromethION

The PromethION Flow Cell contains up to 2675 nanopore channels for sequencing DNA or RNA in real time.

Purchase options:

- Available as packs, contains 4 PromethION flow cells
- Buy as annual or top-up packs to meet your requirements
- Available as R9.4.1, R10.4 and our latest chemistry, R10.4.1

Show all

Protocols all types ... display on site

 Oxford NANOPORE Technologies

Community ▾ Search... 

News About Store Community Contact   Yaovi Hounmanou ▾

 Yaovi Hounmanou Profile 72% complete 

University of Copenhagen Full Member

Following 0 Followers 0 Mentions 0 Interesting

[Create new post](#)



DOCUMENTATION

Nanopore Documentation

ALL / LIBRARY PREP PROT...  

Home Plan Prepare Sequence Analyse

All Sample Storage Extraction Protocols DNA/RNA Handling Library Prep Protocols Automation SARS-CoV-2

 [Chemistry Technical Document](#)

 [Ligation sequencing gDNA - Native Barcoding Kit 96 V14 \(SQK-NBD114.96\)](#)

We also have an FAQ section available on the Nanopore Community Support section. If you have tried our suggested solutions and the issue still persists, please contact Technical Support via

Last updated: 10/20/2022

 [Ligation sequencing gDNA - Native Barcoding Kit 24 V14 \(SQK-NBD114.24\)](#)

We also have an FAQ section available on the Nanopore Community Support section. If you have tried our suggested solutions and the issue still persists, please contact Technical Support via

Last updated: 10/20/2022

 [Ligation sequencing amplicons - Native Barcoding Kit 96 V14 \(SQK-NBD114.96\)](#)

We also have an FAQ section available on the Nanopore Community Support section. If you have tried our suggested solutions and the issue still persists, please contact Technical Support via

Last updated: 10/20/2022

 [Ligation sequencing amplicons - Native Barcoding Kit 24 V14 \(SQK-NBD114.24\)](#)

We also have an FAQ section available on the Nanopore Community Support section. If you have tried our suggested solutions and the issue still persists, please contact Technical Support via

Last updated: 10/20/2022

[Protocols](#) >

[Software Downloads](#) >

[Support](#) >

[NCM 2021 Blog](#) >

FILTERS

DEVICE 

- GridION
- MinION
- PromethION
- Flongle
- MiniT
- VoSTRAX

KIT 

Interactive exercise

- Get ONT- READ here:

<https://drive.google.com/drive/folders/19uCpc7ioC4UMk9xxPt8Hxn5jj9y1Yzuk?usp=sharing>

- Get Illumina reads here:

<https://figshare.com/s/4716724a13c2e7c28e0e>

Analysis from raw reads to assembly

- #basecalling (Guppy with GPU)

```
Path/to/guppy_basecaller -i path/to/fast5 -s /path/to/output/basecalled -c path/to/configurations/dna_r9.4.1_450bps_sup.cfg -x cuda:all
```

- # demultiplexing

```
Path/to/guppy_barcoder -i path/to/pass_basecalled_fastq/ -s path/to/output/demultiplexed -t 40 --num_extra_bases_trim 15 --barcode_kits SQK-RBK004 -r
```

Combine all nanopore fastq files in each barcode (sample) directory into one file per barcode

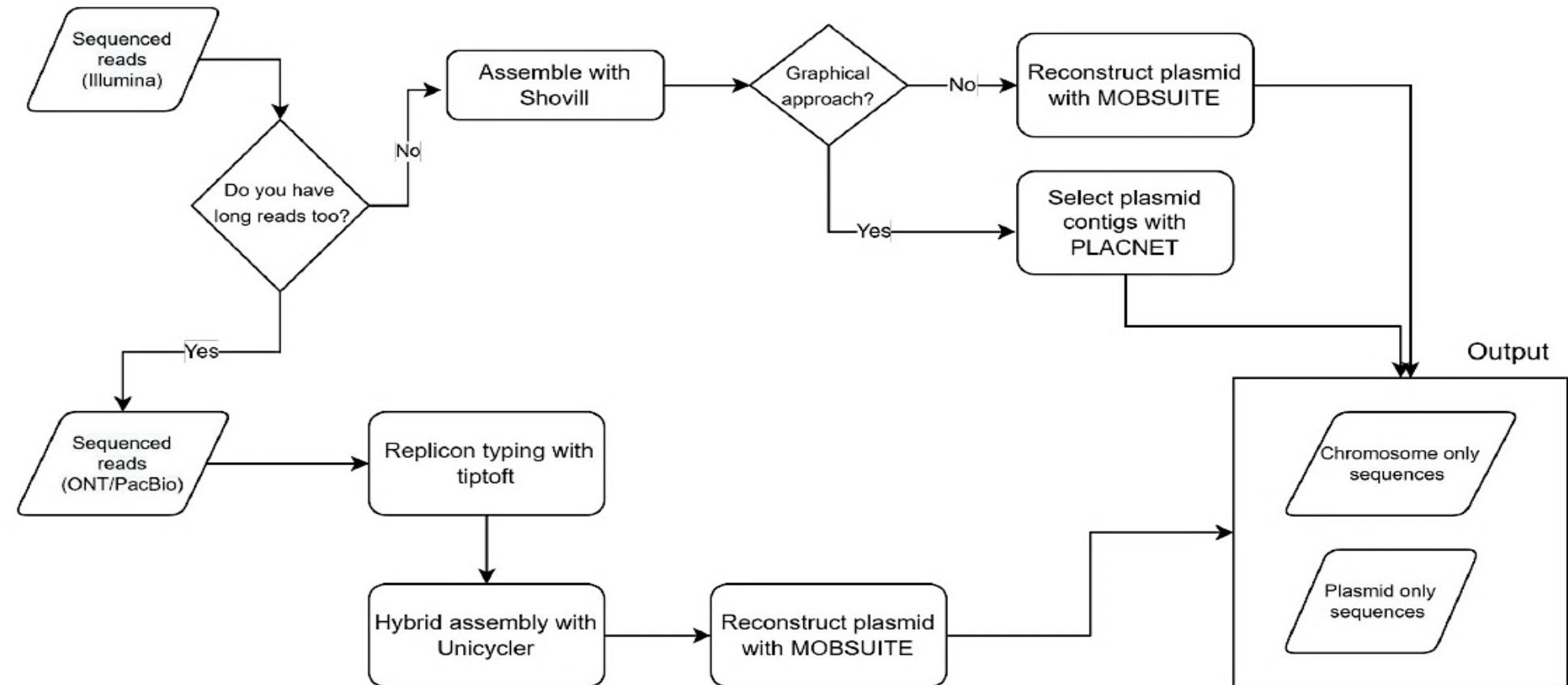
- for dir in \$(ls -d barcode*); do cat \${dir}/*.fastq | gzip > \${dir}/combined_\${dir}.fastq.gz ; done
- Or cat *.fastq > all.fastq
- # Add all combined files to one fast directory
- mkdir sample_fastqs
- cp barcode*/combined*.gz sample_fastqs
- # Run filtlong
- for each in \$(ls sample_fastqs/*barcode*.gz); do filename=\$(basename \$each); prefix=\${filename%%.*}; filtlong --min_length 1000 --keep_percent 95 \$each | gzip > assembly/filtlong/\${prefix}.filt.fastq.gz ; done
- Or filtlong --min_length 1000 --keep_percent 95 input_reads.fastq.gz > reads.fastq
- #Run flye on filtlong output
- for each in \$(ls filtlong/*filt.fastq.gz); do filename=\$(basename \$each); prefix=\${filename%%.*}; flye --nano-hq -o flye_assembly_\${prefix} --threads 16 \$each ; done
- Hibrid assembly
- **Illumina-only assembly:**
unicycler -1 short_reads_1.fastq.gz -2 short_reads_2.fastq.gz -o output_dir
- **Long-read-only assembly:**
unicycler -l long_reads.fastq.gz -o output_dir
- **Hybrid assembly:**
unicycler -1 short_reads_1.fastq.gz -2 short_reads_2.fastq.gz -l long_reads.fastq.gz -o output_dir

Long and short read *de novo* hybrid genome assembly

1. Why do hybrid assembly?
2. In AMR studies what is the importance of understanding which genomic regions i.e. MGEs vs chromosome a gene if interest if located on?
3. What is the genetic context of an ARG and why is that important?

Identifying plasmid sequence in genome assemblies

Through this process you should be able to separate your assembled contigs into plasmid and chromosome.



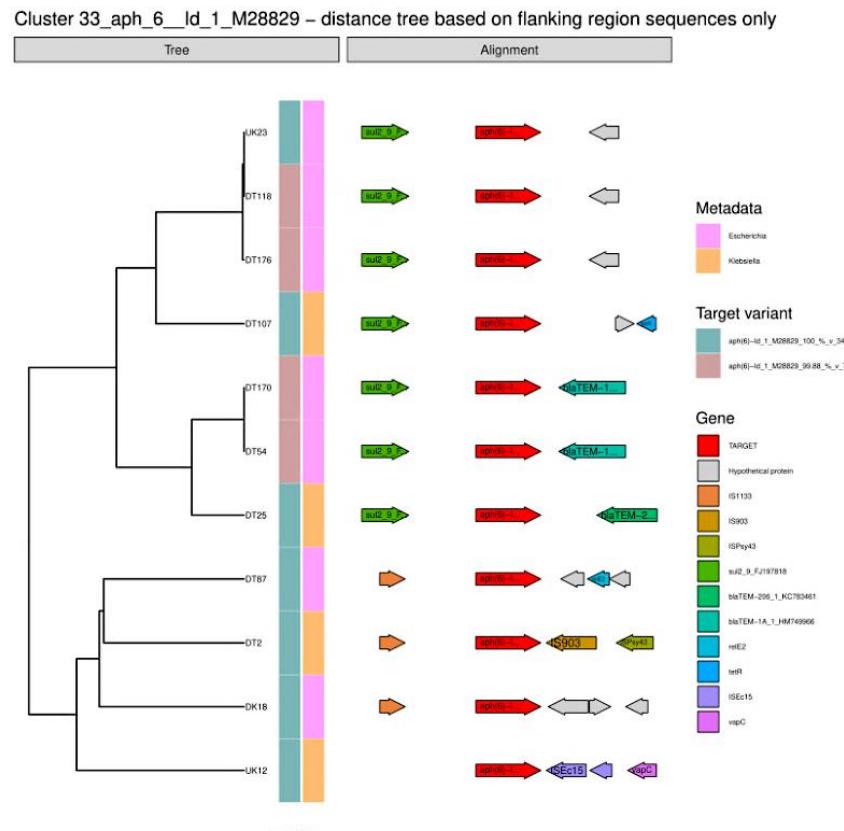
QUICK TOOLS TO SCREEN FOR ARGs location and context

<https://cge.food.dtu.dk/services/SourceFinder/>

<https://bitbucket.org/genomicepidemiology/flankophile/src/master/>

About Flankophile

Flankophile is a pipeline built for easy analysis and visualization of gene synteny - the genetic context of genes. Flankophile is especially useful for comparing the flanking regions of specific genes or other target sequences across different samples. Flankophile automatically gene synteny plots in pdf format. Flankophile also outputs the percentage identity of each hit relative to its reference sequence. This allows for incorporating gene variants into the analysis.



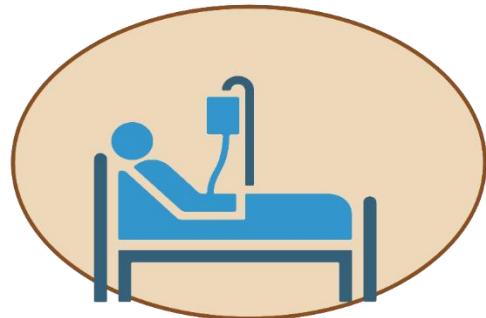
Case study from Uganda

- We tested the hypothesis that ESBL *E. coli* and/or their resistance determinants could spread between the healthcare and community settings through discharged patients that were still colonized.

Healthcare workers and healthcare contacts

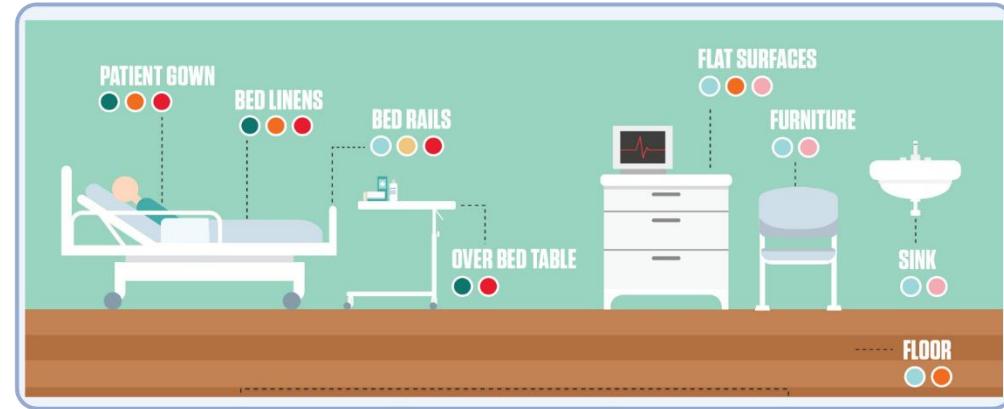


Index Patients

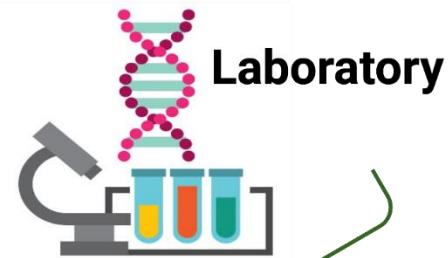


1600 samples
collected
268 MIC
67 Illumina
15 ONT

Healthcare Environments



Healthcare setting

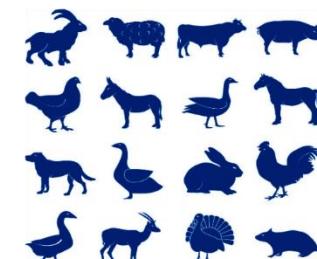


Laboratory

Community setting



Household Contacts

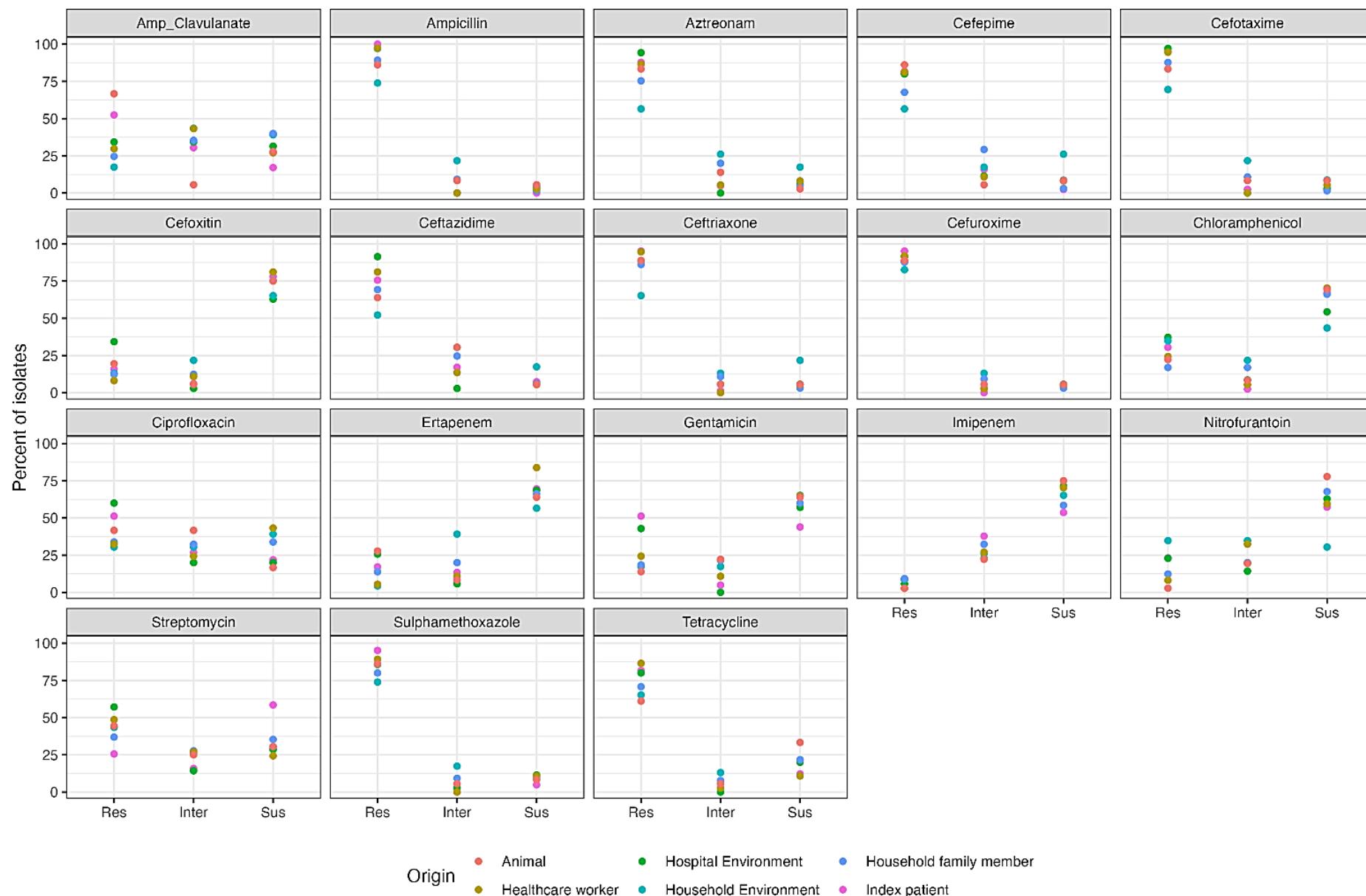


Household Animals

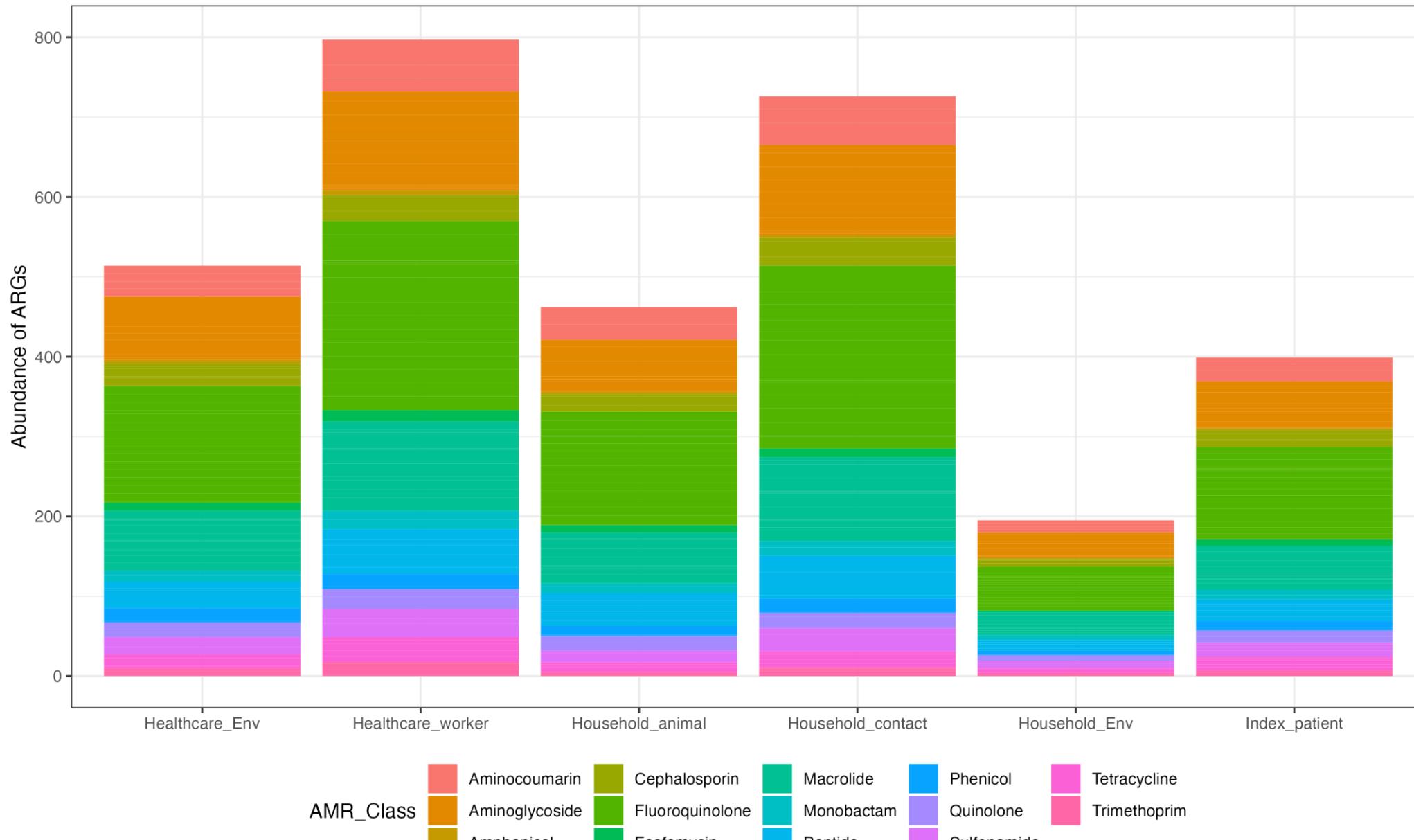


Household Environments

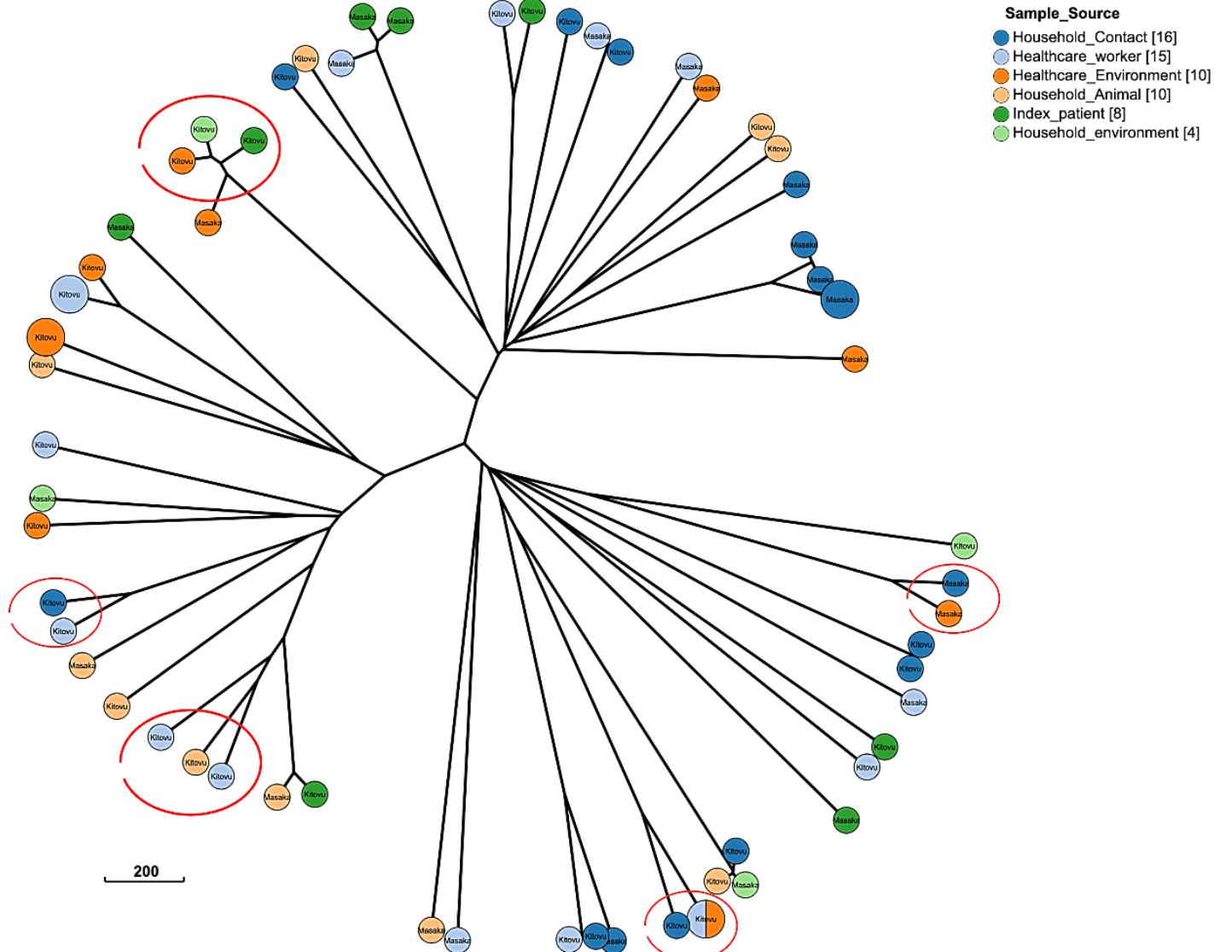
Phenotypic AMR profile by niche to each antimicrobial showing MDR with similar selection pressure for maintenance of strains



Relative abundance of antimicrobial resistance genes detected by class per niche showing the presence of 14 AMR classes in all the categories

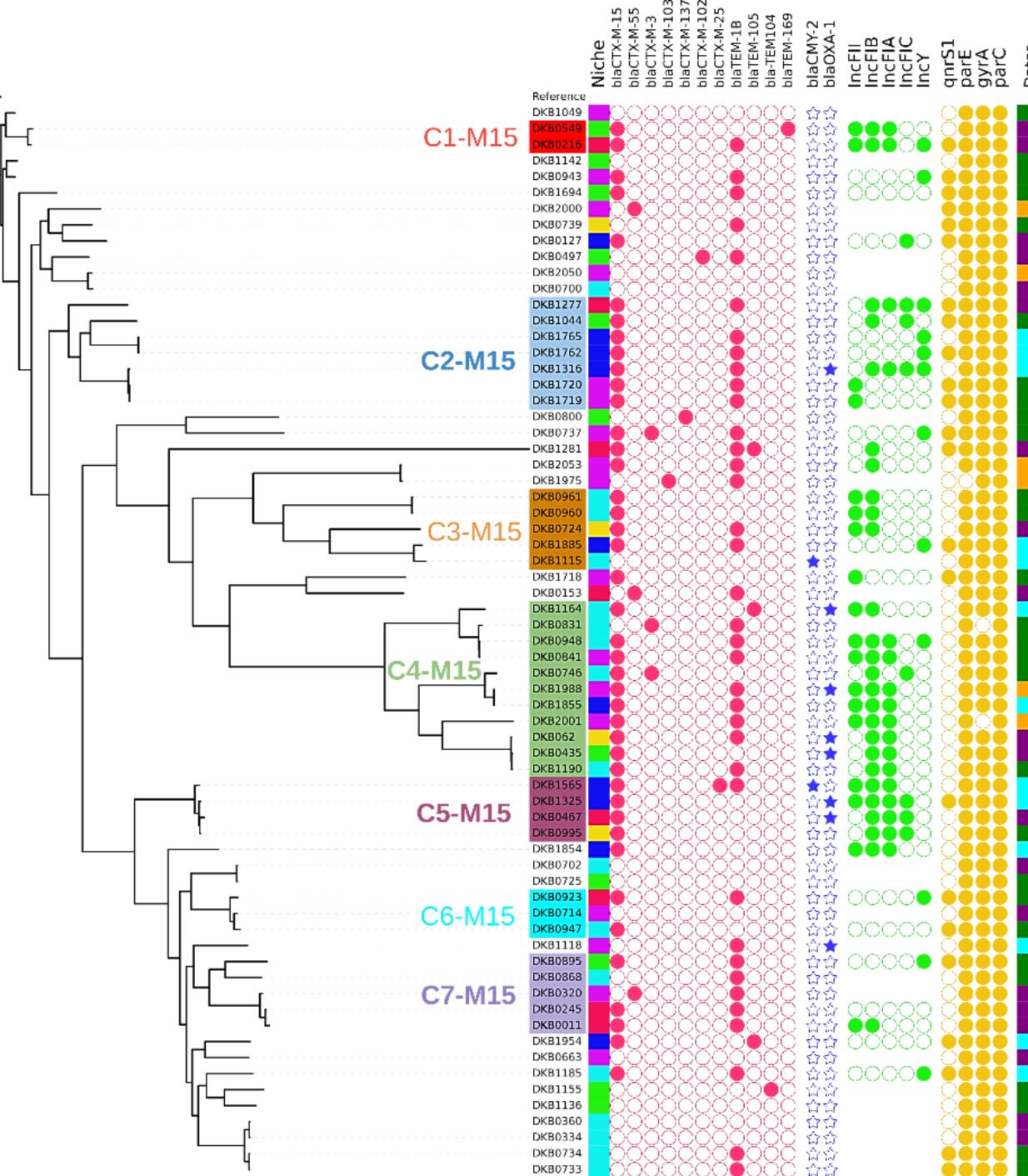


MST tree built with allelic distances generated by cgMLST from analysed strains – five sharing events

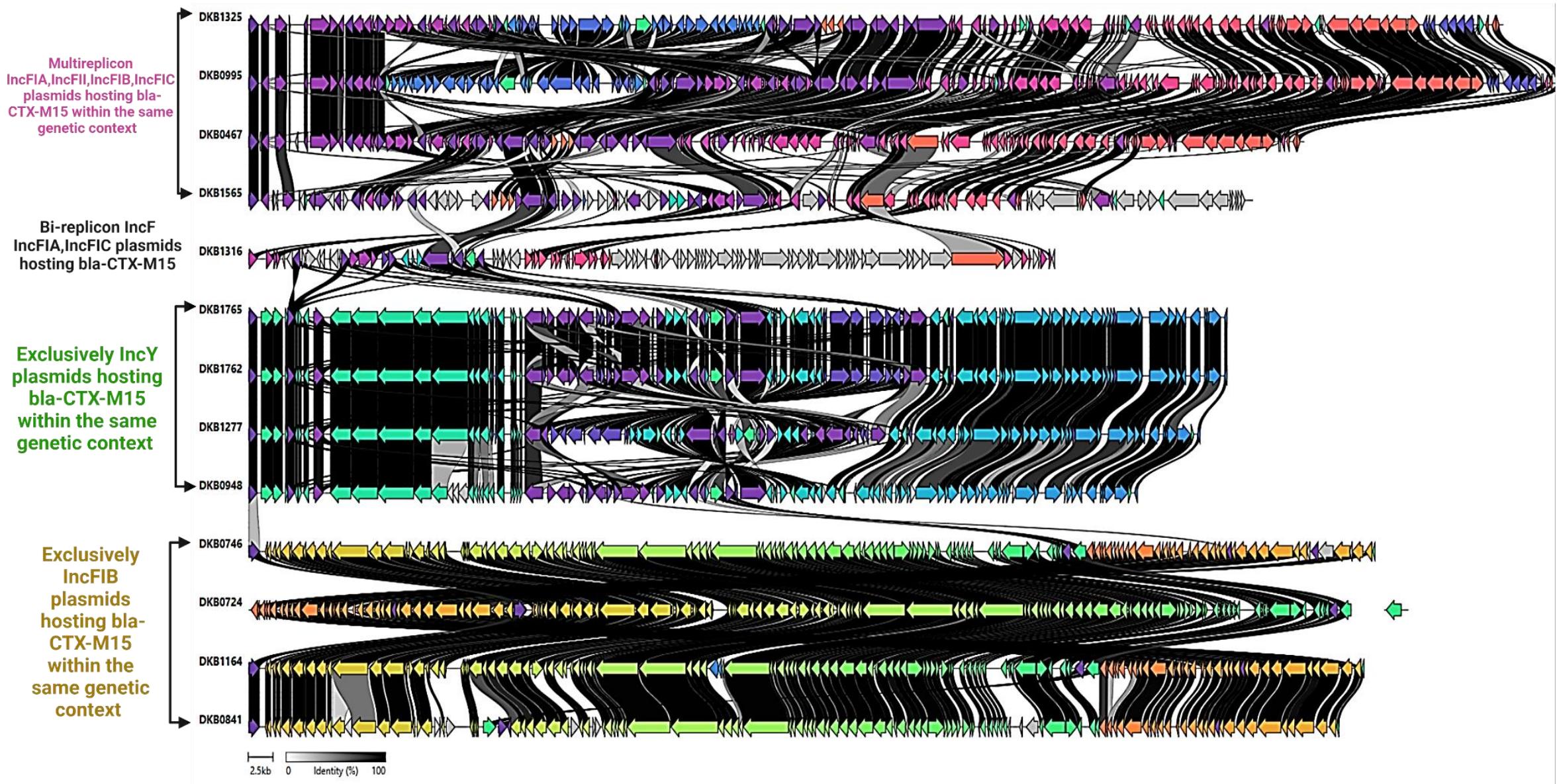


Core-genome SNP-based phylogeny of the ESBL *E. coli* strains and selected markers – one sharing event in C4

Tree scale: 0.1

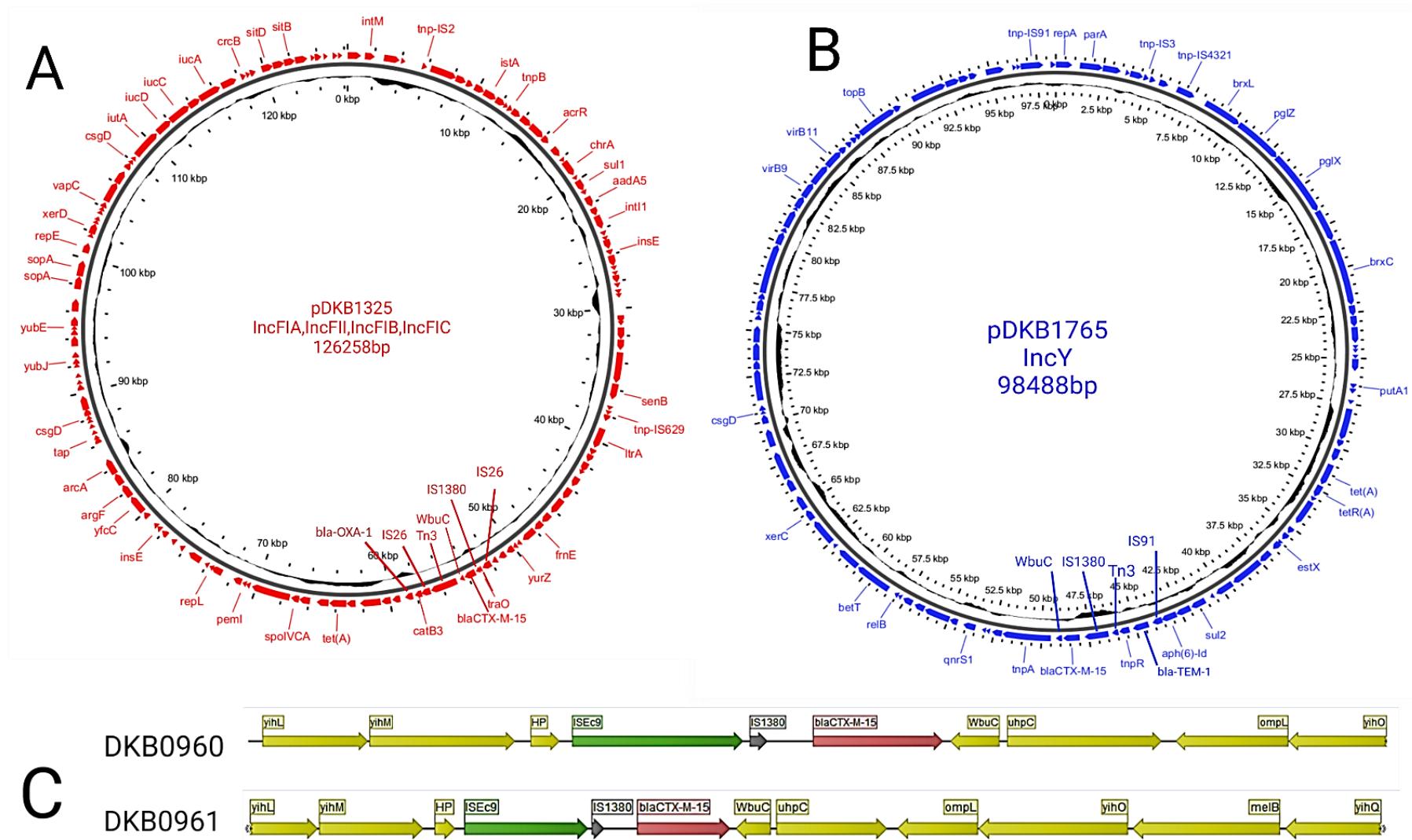


Overall comparison of plasmids categories hosting the *bla*-_{CTX-M-15} gene as reconstructed from hybrid assemblies in healthcare and community strains.



Context of representative complete ESBL-plasmids reconstructed by hybrid assembly from strains DKB1325 (A) and DKB1765 (B).

In panel C we show the genetic context around *bla*_{CTX-M-15} in the strains where we observed that the gene has settle on the chromosome through a transposon insertion sequence IS1380 (from IncFIB)

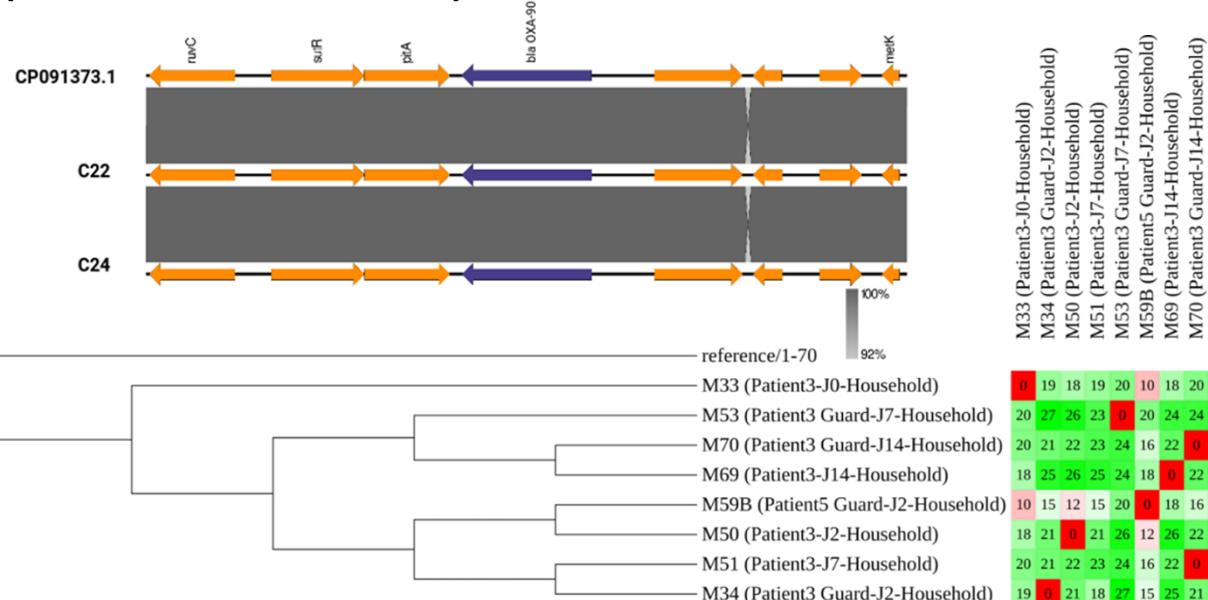


TRANSMISSION OF CARBAPENEMASE PRODUCING BACTERIA BETWEEN HOSPITAL AND HOUSEHOLD SETTING IN BENIN USING HURID ASSEMBLY



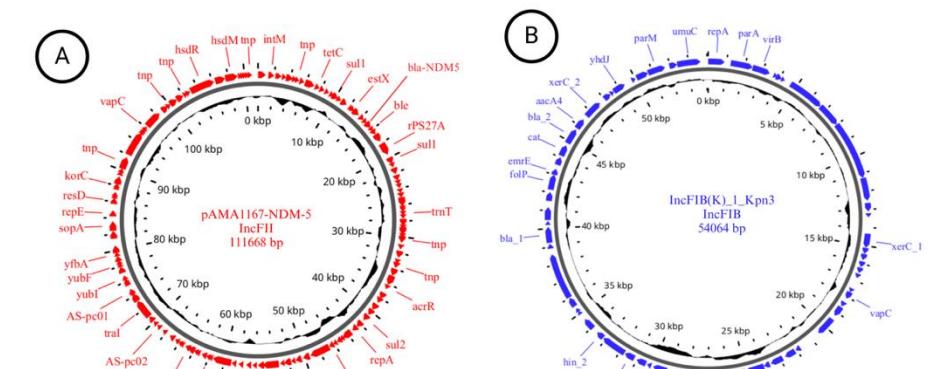
Kevin SINTONDJI

(PhD student, UAC, Benin)



- CRE colonization from admission (D0): 7.27% to 22.72% by D8 to 57.14% by D14 during follow-up after discharge.
 - Top species: *Escherichia coli* (34.14%), *Klebsiella pneumoniae* (29.27%), *Acinetobacter baumannii* (14.63%)
 - Key genes *blaNDM-5* and *blaOXA-181*
 - Transmission networks: Observed between patients, **caregivers**, and household contacts after discharge.
 - Resistance mechanism: Mainly IncF-mediated.

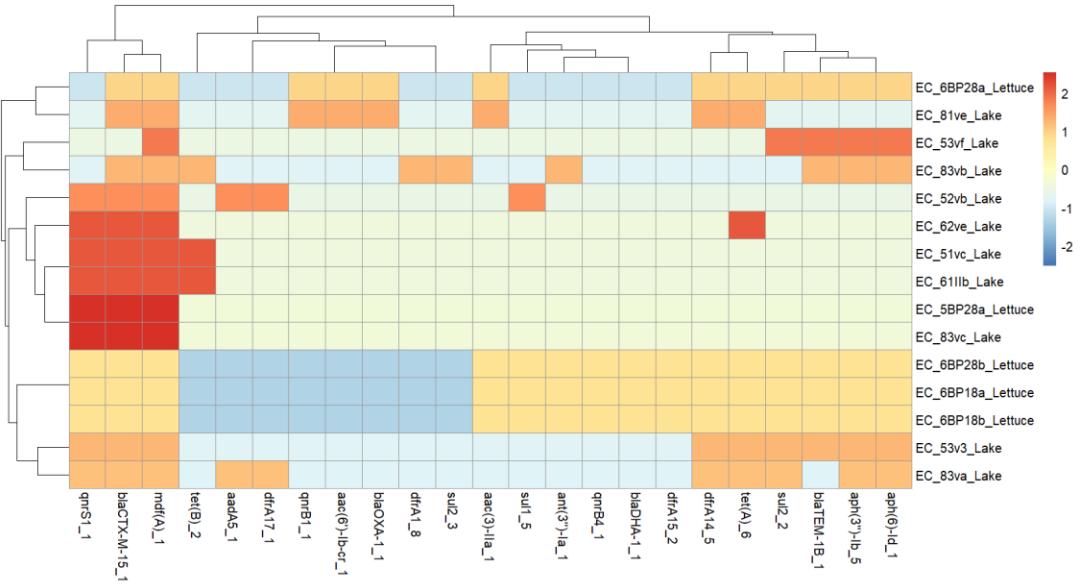
Transmission Network



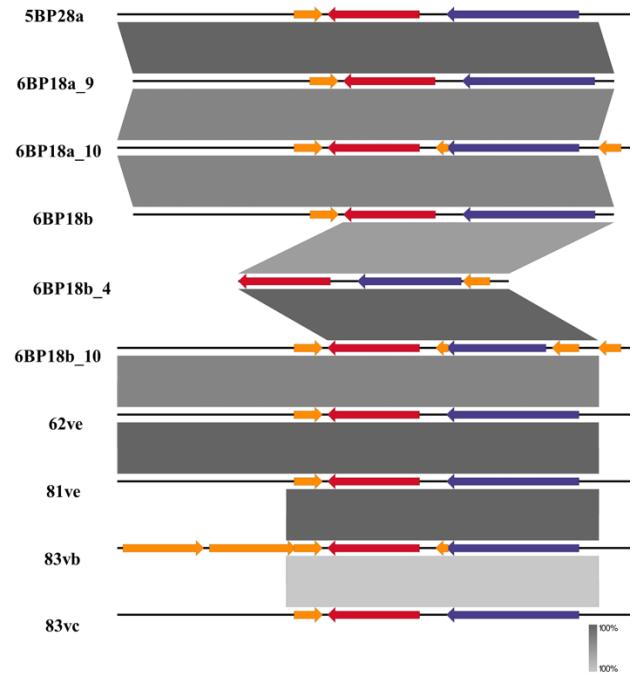
Most represented Plasmid

Extended-Spectrum β -Lactamase-Producing *Escherichia coli* recovered from Yamoussoukro manmade lakes and *Lactuca sativa* surrounding cultures

- The strains are unrelated and of diverse sequence types
- blaCTX-M15*, main ESBL gene despite strain diversity
- Highly mobile context: *blaCTX-M15* was flanked either by IS6 family transposase IS26 or structure IS1380 family transposase ISEcp1



Antibiotic Resistance genes in Lettuce and Lake Samples

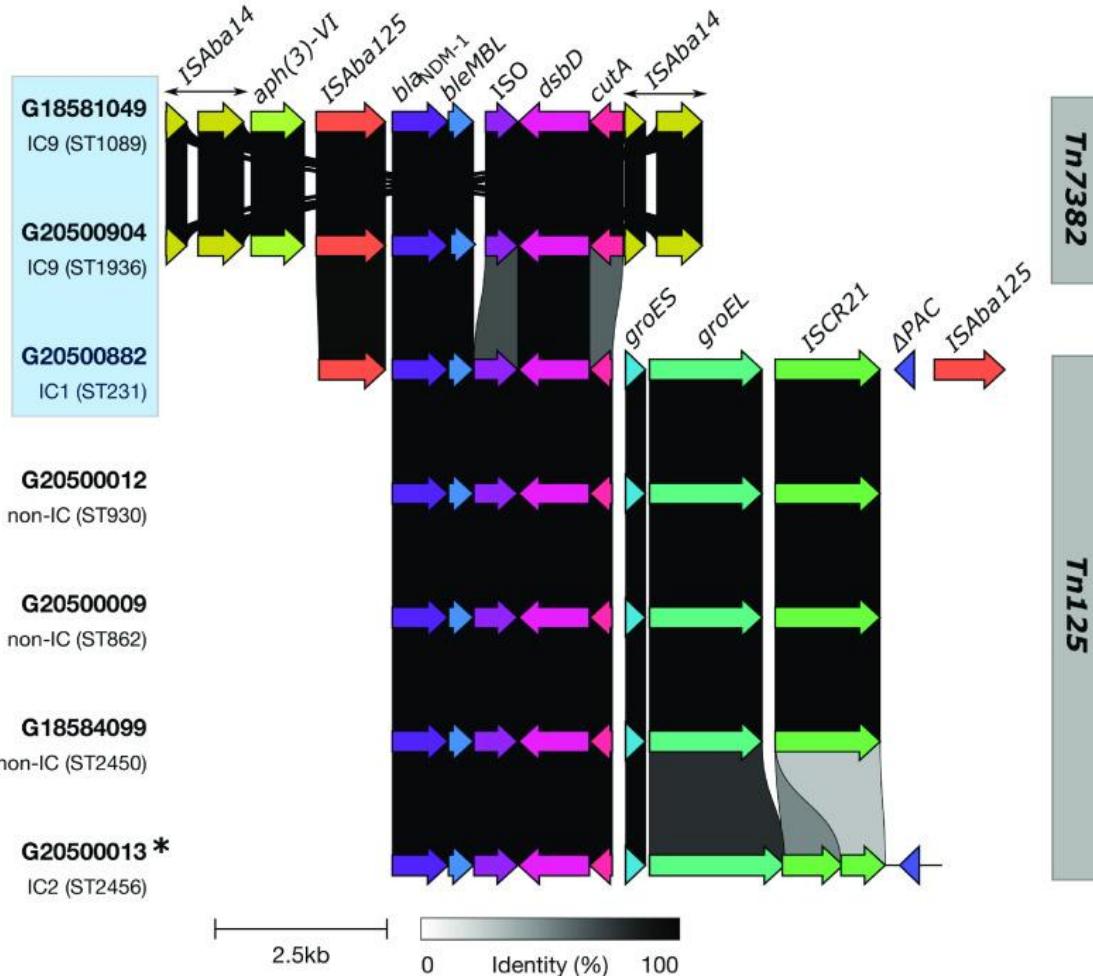


blaCTX-M-15 gene context



High Genetic Diversity of Carbapenem-Resistant *Acinetobacter baumannii* Isolates Recovered in Nigerian Hospitals in 2016 to 2020

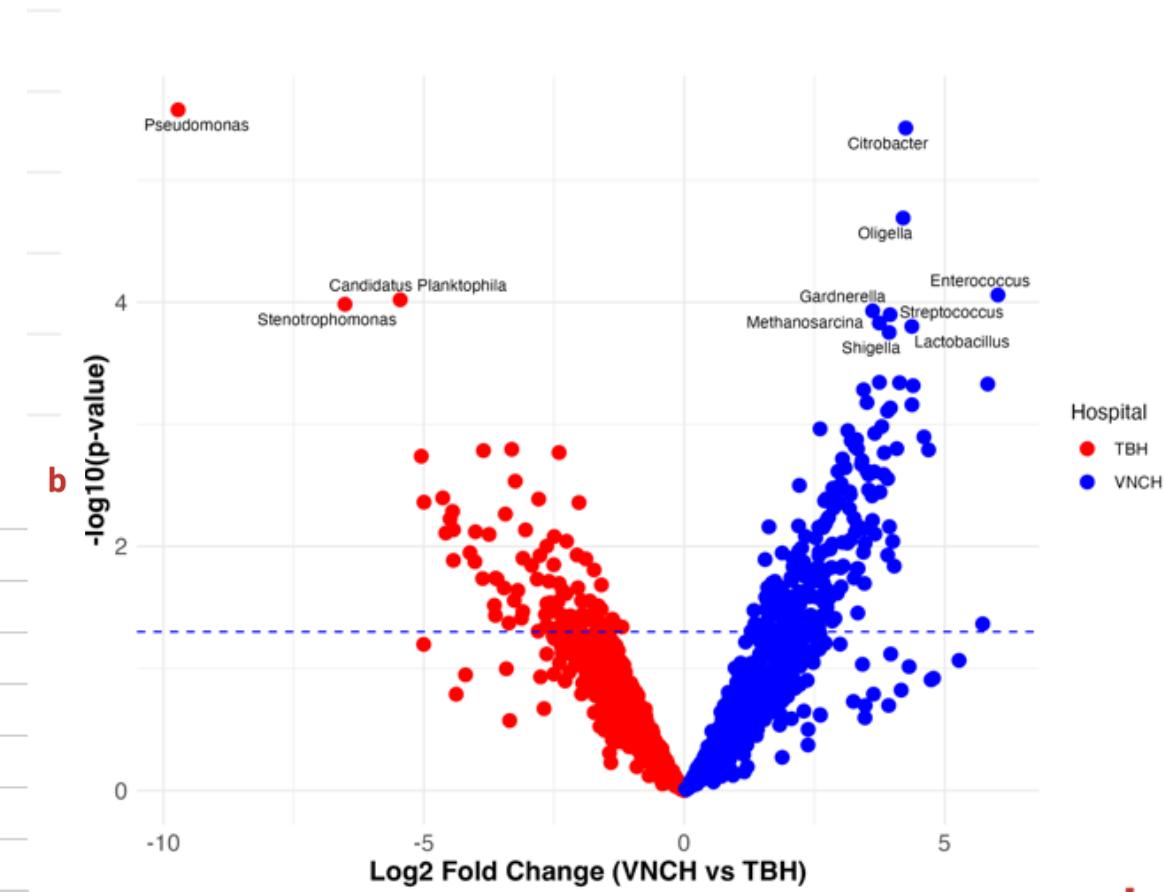
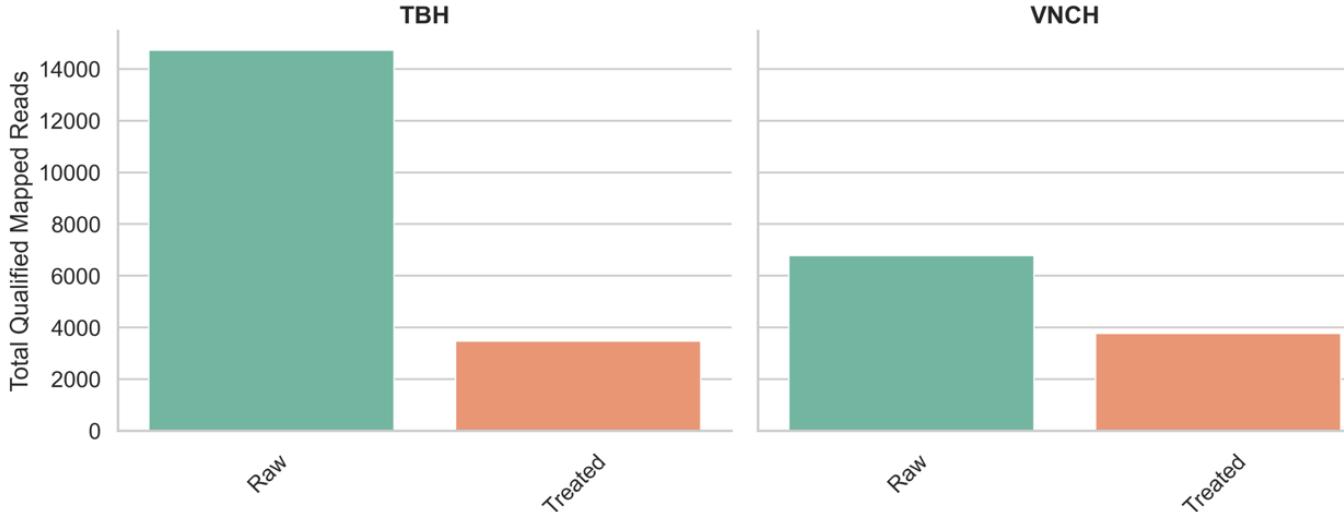
Erkison Ewomazino Odih,^{a,b} Anderson O. Oaikhena,^a Anthony Underwood,^c Yaovi Mahuton Gildas Hounmanou,^b



- Over 50% of isolates were resistant to 10 of 12 tested antimicrobials.
- 54/80 isolates were carbapenem-resistant:
- **Key resistance genes:** *blaOXA-23* (34.9%) and *blaNDM-1* (27.9%)
- *blaOXA-23* genes carried on Tn2006-like transposons,
- while a 10-kb Tn125 transposon facilitated *blaNDM-1* dissemination.

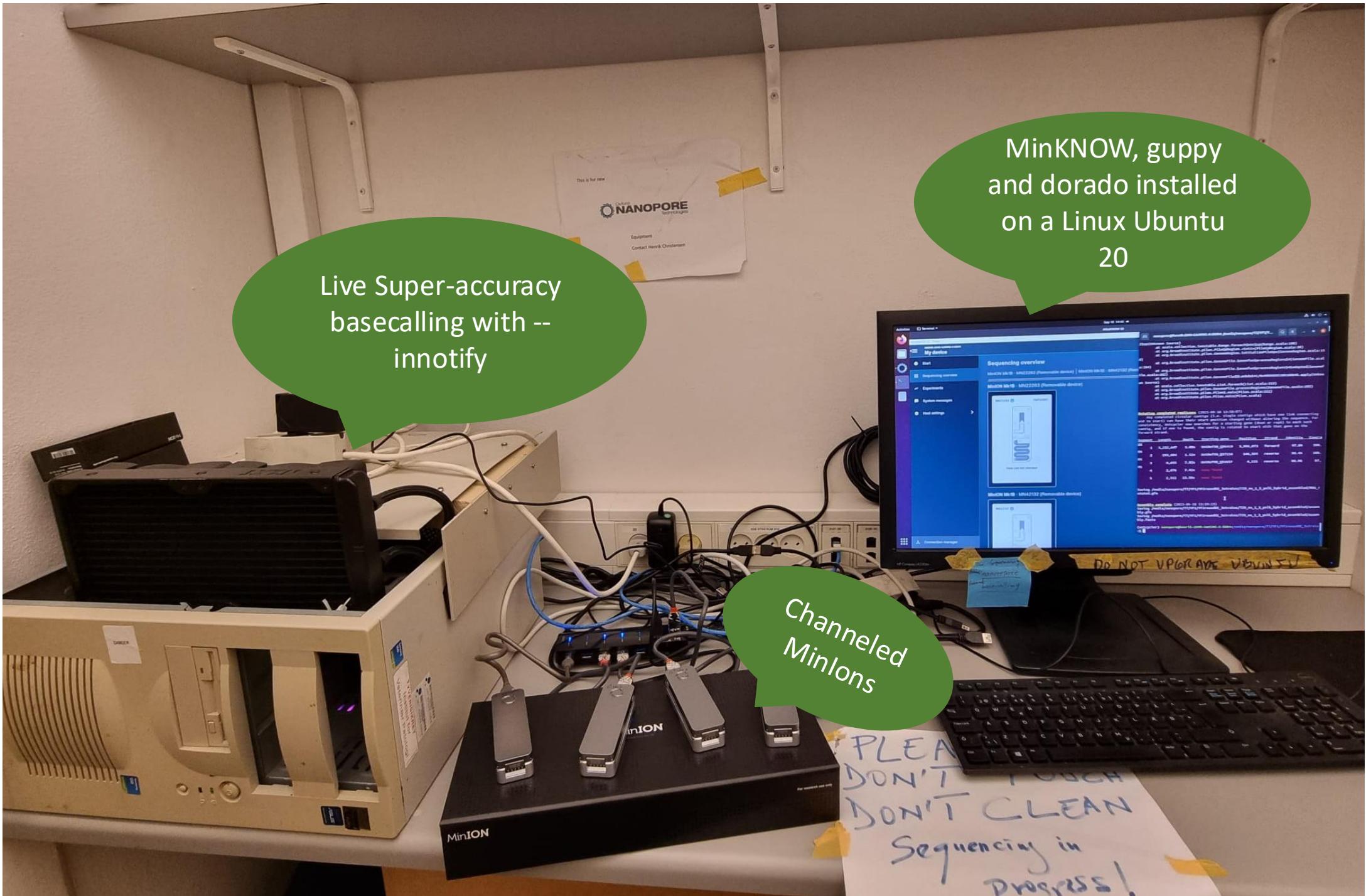
Interventions to decrease Carbapenem Resistant Enterobacteriaceae colonization and transmission between hospitals, households, communities and domesticated animals.

Onsite hospital wastewater treatment can decrease ARGs and the number of MDROs entering aquatic recipients and thereby the risk of human exposure and transmission downstream



d

What we propose and why?

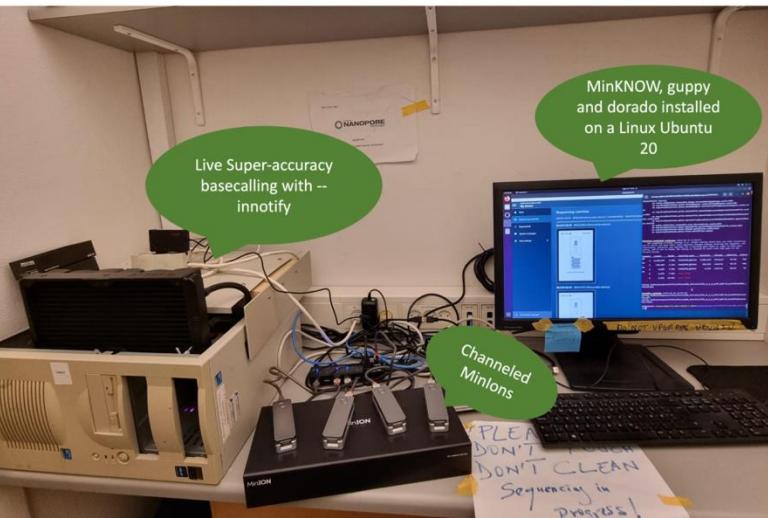




[gilmahu/Live_basecalling](#) Public

Projects Security Insights

master 1 Branch 0 Tags Q Go to file <> Code



gilmahu Update README.md 2075a07 · 3 weeks ago 3 Commits

README.md Update README.md 3 weeks ago

live_basecalling-gil.sh Initial commit: add basecalling script 3 weeks ago

README

About

This script is automating the real-time processing of .fast5 files for nanopore sequencing, using Guppy to basecall them as they appear in the observed directory, using inotifywait and organizing the outputs into batches. It can simply be adapted for .pod5 files basecalled with dorado the new ONT-basecalling tool.

Readme Activity 0 stars 1 watching 0 forks Report repository

Releases

No releases published

Packages

No packages published

Languages

Shell 100.0%

Live Basecalling Script for Guppy on fast5 files (Adaptable for pod5 files under dorado basecaller)

This repository contains a Bash script designed for real-time monitoring and basecalling of Nanopore .fast5 files using Guppy. The script leverages `inotifywait` to observe a specified directory for new .fast5 files, processes them using Guppy for super accuracy basecalling, and saves the resulting .fastq files to a designated output directory.



[gilmahu/Long_read_metagenomic_pipeline](#) Public

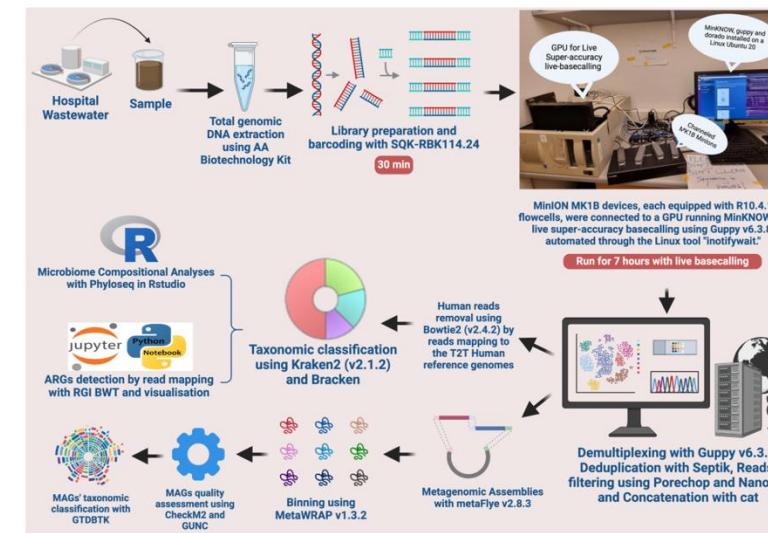
main 1 Branch 1 Tag Q Go to file <> Code About

gilmahu Update README.md 94ed50a · yesterday 28 Commits

- R_analysis** Add R Markdown analysis file yesterday
- notebooks** Add Jupyter Notebook analysis files (HTML and .ipynb) yesterday
- scripts** Add BIOM file export script yesterday
- Fig1_Overall_workflow.png** Add files via upload yesterday
- README.md** Update README.md yesterday
- run_pipeline.sh** Add files via upload yesterday

README

Pipeline Flow Diagram



Comprehensive pipeline for metagenomic analysis using Nanopore sequencing

Readme Activity 0 stars 1 watching 0 forks Report repository

Releases

1 tags

Packages

No packages published

Languages

- HTML 56.5%
- Jupyter Notebook 43.3%
- Shell 0.2%

Figure: Overview of the Long-Read Metagenomic Analysis Pipeline.

Long-Read Metagenomic Analysis Pipeline

Pros of Nanopore in Microbial Genomics

Feature	Benefit / Impact
Real-time sequencing	Immediate data streaming enables rapid pathogen detection and adaptive sampling.
Portability & field deployment	Devices like MinION can be used directly in remote or clinical settings.
Ultra-long reads	Enables high-quality genome assembly, plasmid tracking, and repeat resolution.
Direct detection of native nucleic acids	Identifies base modifications and RNA without amplification.
High microbial genome accuracy	ONT-only assemblies now exceed 99% identity for many microbes.
Emerging multi-omic capabilities	Supports direct protein sequencing and epigenetic analysis.

Cons of Nanopore in Microbial Genomics

Limitation	Challenge / Consideration
Higher per-base error rates	Despite improvements, still more indels and substitutions than short reads.
Workflow variability	Frequent software changes and protocol evolution demand constant adaptation.
Cost and throughput	Higher per-base cost and limited ultra-deep sequencing efficiency.
Flow cell contamination risk	Re-use can cause carryover, and wash kits don't fully eliminate residues.
Device durability	Sensitive nanopores and flow cells degrade under field conditions.
Computational demands	Requires strong infrastructure for basecalling, polishing, and storage.



Gildas Hounmanou, PhD

APTI-Fellow, National Institutes of Health, MD, USA

Assistant-Professor

<https://orcid.org/0000-0003-3991-0864>

University of Copenhagen

Faculty of Health and Medical Sciences

Section for Food Safety and Zoonoses

Department of Veterinary and Animal Sciences

***Thank you
for Listening!***