Oxford Nanopore Sequencing for hard-to-treat TB: set up in resource-constrained settings

Dr Linzy Elton

Postdoctoral research associate

@LinzyElton @PandoraIDNet @AmrCovid

We will record these sessions and put them online so you can refer back to them later on

We will also put the slides up online so you can access the notes (links and image credits)

- Laboratory set up
- Introduction to the TB sequencing pipeline
- TB sequencing complications

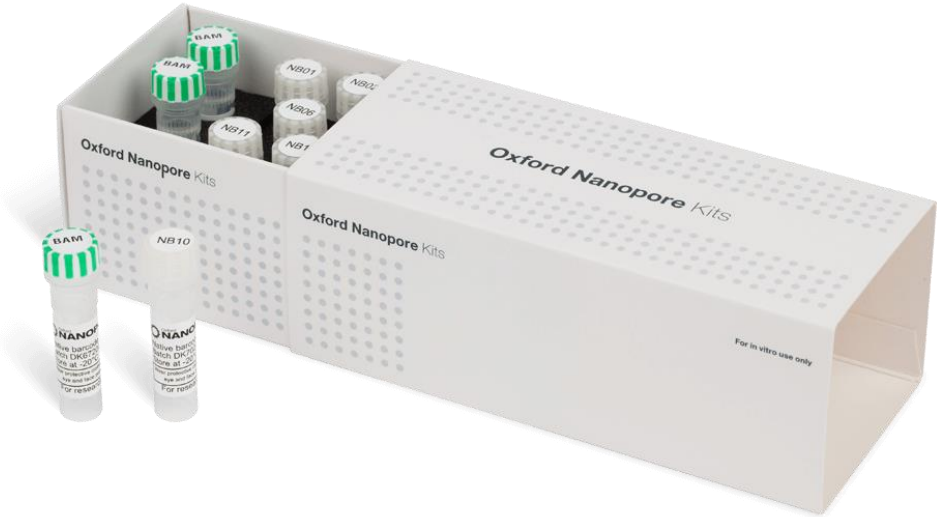
Laboratory set up

Lead and delivery times

- Some kits and equipment have a lead time
- Generally, the Nanopore website suggests allowing a working week for orders to arrive for ‘rest of world’
- Consider customs issues for your country

Reagents – Storage

- Most of the kit reagents need a cold chain and so must be kept in a fridge/ freezer
- Nanopore have created a field sequencing kit



Flow cells

- MinION flow cells can be stored (unopened) at room temperature for 4 weeks (from delivery date)
- They can be stored for 12 weeks (unopened) at 2-8°C
- Flow cells should be checked before library preparation (within 3 months of purchase) as part of the warranty

Sample Throughput

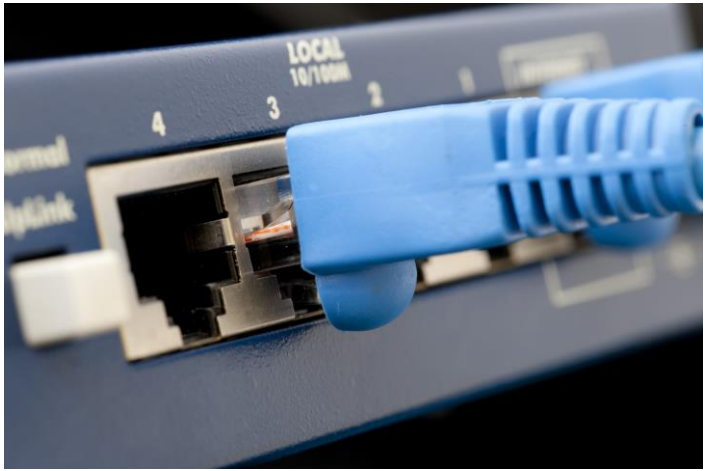
- If few samples are being processed at a time, or you are looking at amplicons, smaller genomes or targeted regions, the Flongle might be more useful
- Note that they are only guaranteed for 4 weeks after the shipment date (rather than 12 for a MinION)

Power issues

- Need power for the duration of your sequencing run (if it cuts out, it will save data up until the power cut)
- A laptop with a reliable battery may be more useful than a computer if power cuts are an issue
- UPS is something to consider
- Mk1C also needs an external power source

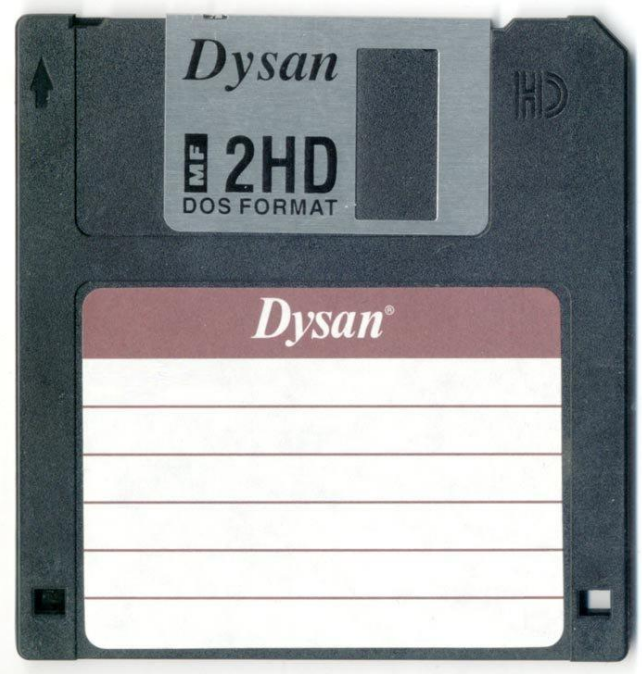
Internet Access

- MinION Mk1B needs a constant internet connection
- The Mk1C can sequence and basecall entirely without internet connectivity
- Many analysis programmes (e.g. TB-Profiler) need internet access
- Offline configurations are available



Data storage

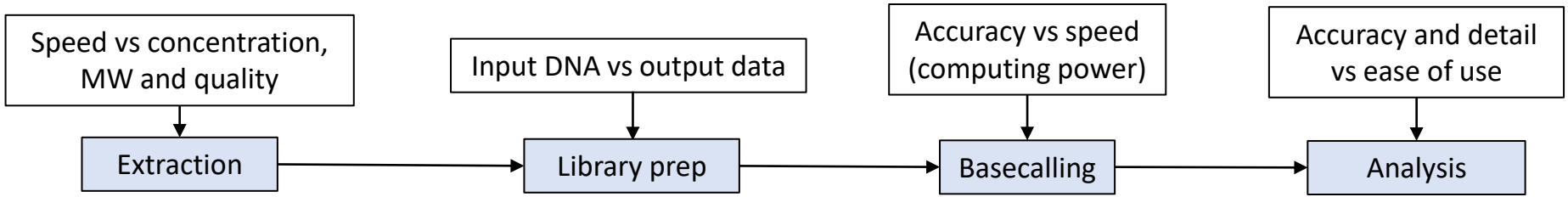
- A MinION flow cell can generate as much as 50 Gb of data, so storage may be an issue
- Running out of disk space will halt your sequencing experiment
- Unless you have GPUs, basecalling won't happen in real time
- The Mk1C has 1 TB SSD storage, 8 GB RAM + GPU
- Consider human sequence data (GDPR)



The TB sequencing pipeline

Aim of the TB sequencing pipeline

- Problem: can WGS be used to compliment routine diagnosis, especially in hard to treat cases e.g. relapse vs reinfection and drug resistant cases?
- Aim to build a pragmatic pipeline: balance between cost, accuracy and feasibility for research and routine TB laboratory scientists



Why ONT for sequencing TB?

- Cheap initial set up
- Smaller laboratory footprint
- Long reads can be useful for spanning traditionally hard to read sequences (TB is GC rich and has lots of repeats)
- Relatively straightforward library preparation
- ~$160 per sample (basic pipeline)

Problems

- ONT *generally* has lower accuracy than Illumina – so we need to validate it (new chemistry coming soon!)
- ONT kits *generally* need more input DNA (for native DNA sequencing)
- TB is difficult to work with:
  - Very slow growing!
  - Low yield from extraction (complex cell wall structure)
  - High GC content
  - Many repeat regions

TB extractions generally yield low concentrations (~5 ng/µl)…

Rapid PCR barcoding kit (SQK-RPB004)

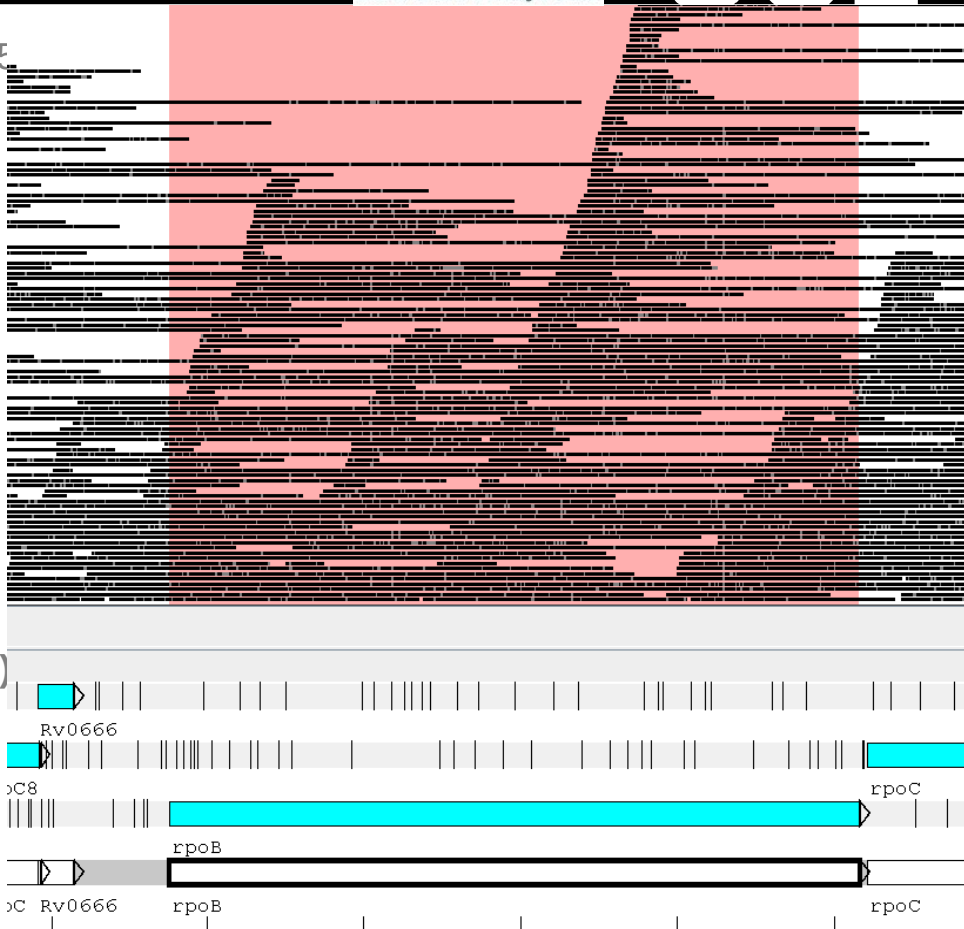
- <2 ng/µl input DNA
- Amplification was not even

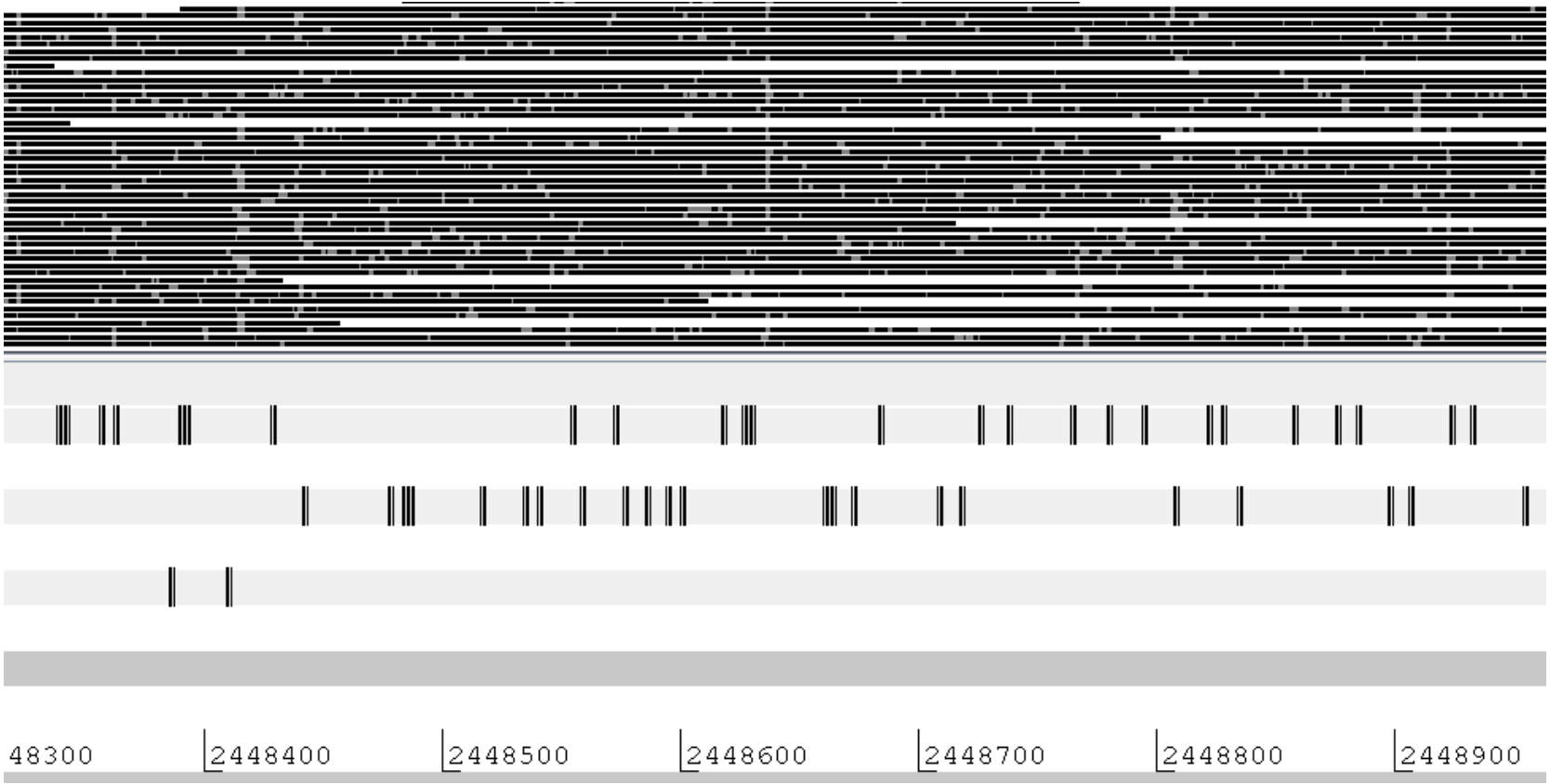
Rapid barcoding kit (native) (SQK-RBK004)

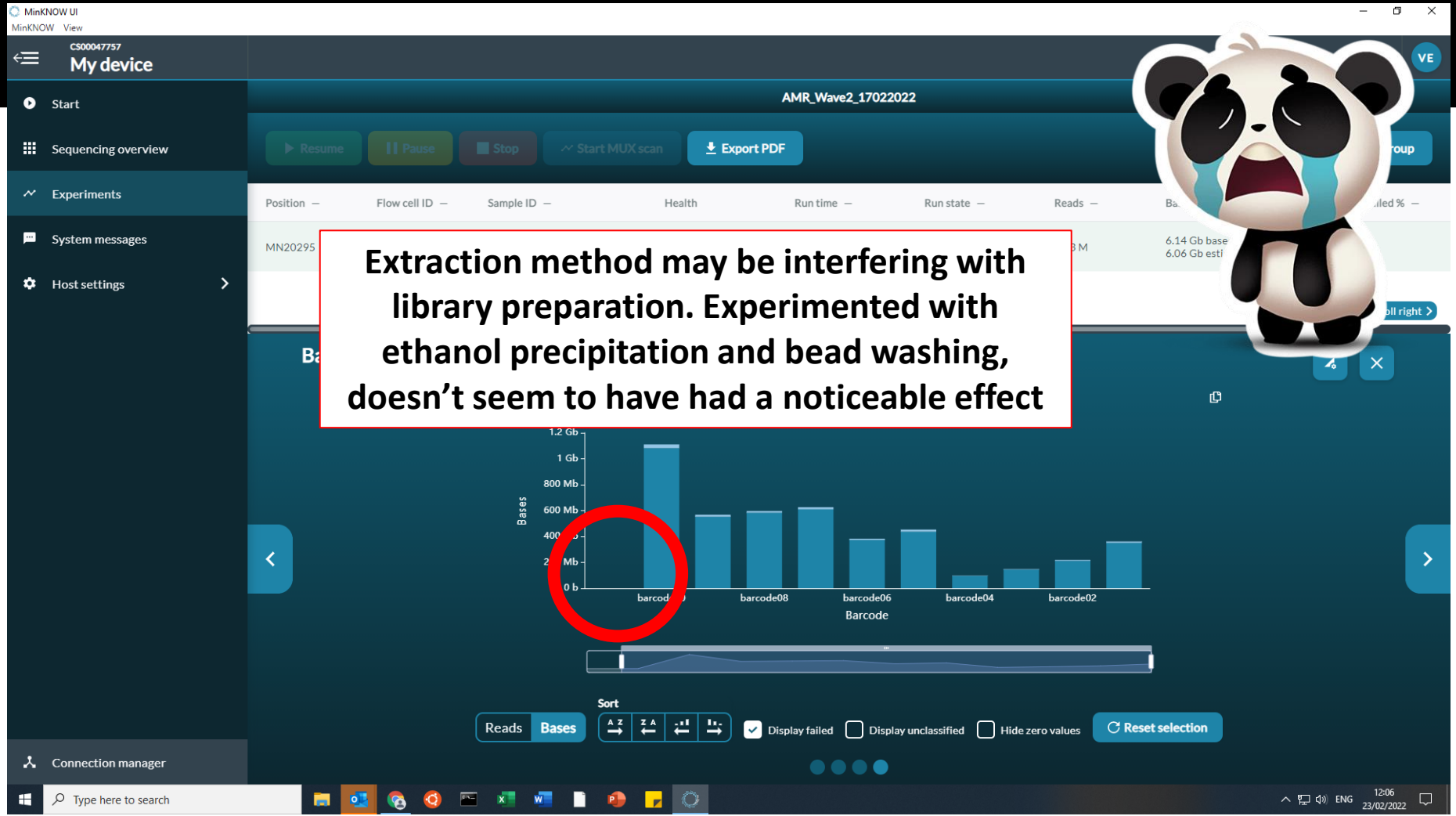
- 400 ng of >30kb DNA
- Normalise to fmol
- Better coverage, but needs improving

Rapid barcoding kit 96 (native) (SQK-RBK110-96)

- **50-400** ng of >30kb DNA
- Within 24h of sequencing, had >40x depth for 6
- After 48h all were between 54-128x depth




48300
2448400
2448500
2448600
2448700
2448800
2448900


MinKNOW UI
MinKNOW View
CS00047757
My device
Start
Sequencing overview
Experiments
System messages
Host settings
Connection manager
AMR_Wave2_17022022
Resume
Pause
Stop
Start MUX scan
Export PDF
Position
Flow cell ID
Sample ID
Health
Run time
Run state
Reads
MN20295
6.14 Gb base
6.06 Gb esti
Extraction method may be interfering with library preparation. Experimented with ethanol precipitation and bead washing, doesn't seem to have had a noticeable effect
1.2 Gb
1 Gb
800 Mb
600 Mb
0 b
Bases
barcode08
barcode06
barcode04
barcode02
Barcode
Sort
Reads
Bases
Display failed
Display unclassified
Hide zero values
Reset selection
Type here to search
ENG
12:06
23/02/2022

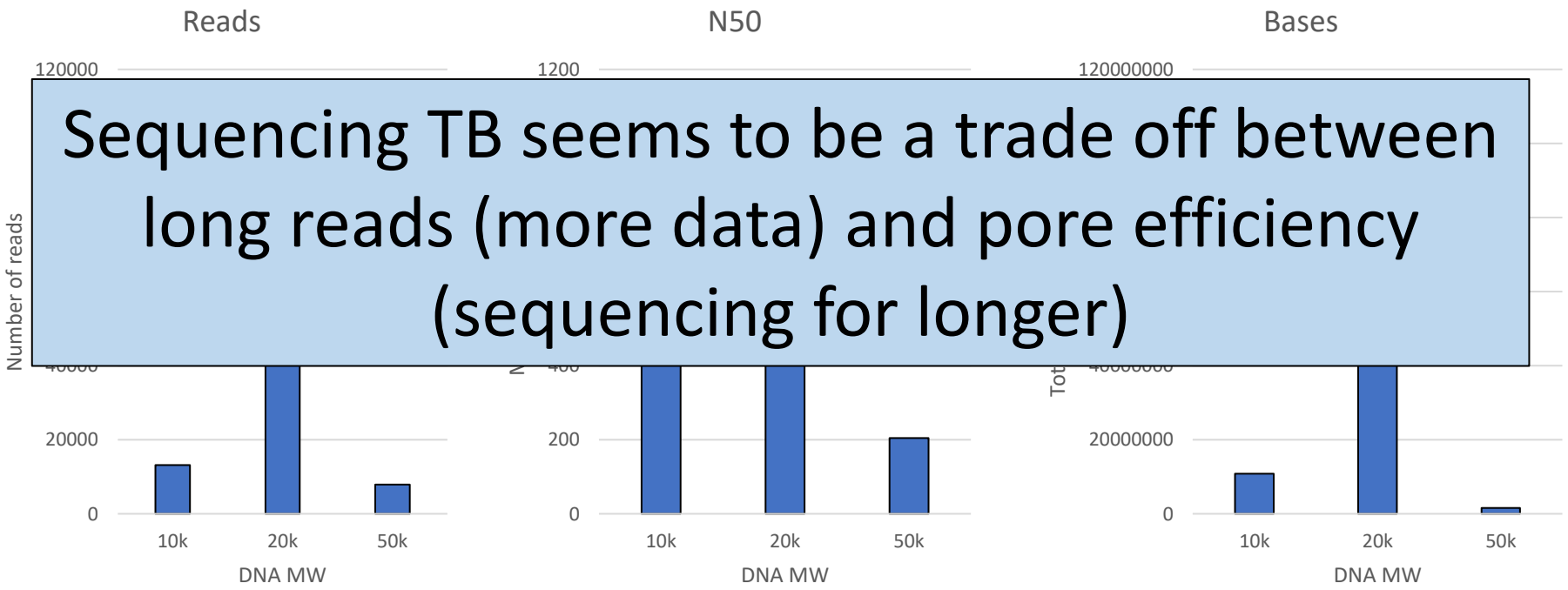
Is the DNA actually too long?

- Mycobacteria have a high CG content and lots of repeat regions
- The lower pore efficiency is possibly due to repeat regions binding to itself and blocking the pores
- Does reducing the length of the DNA help?

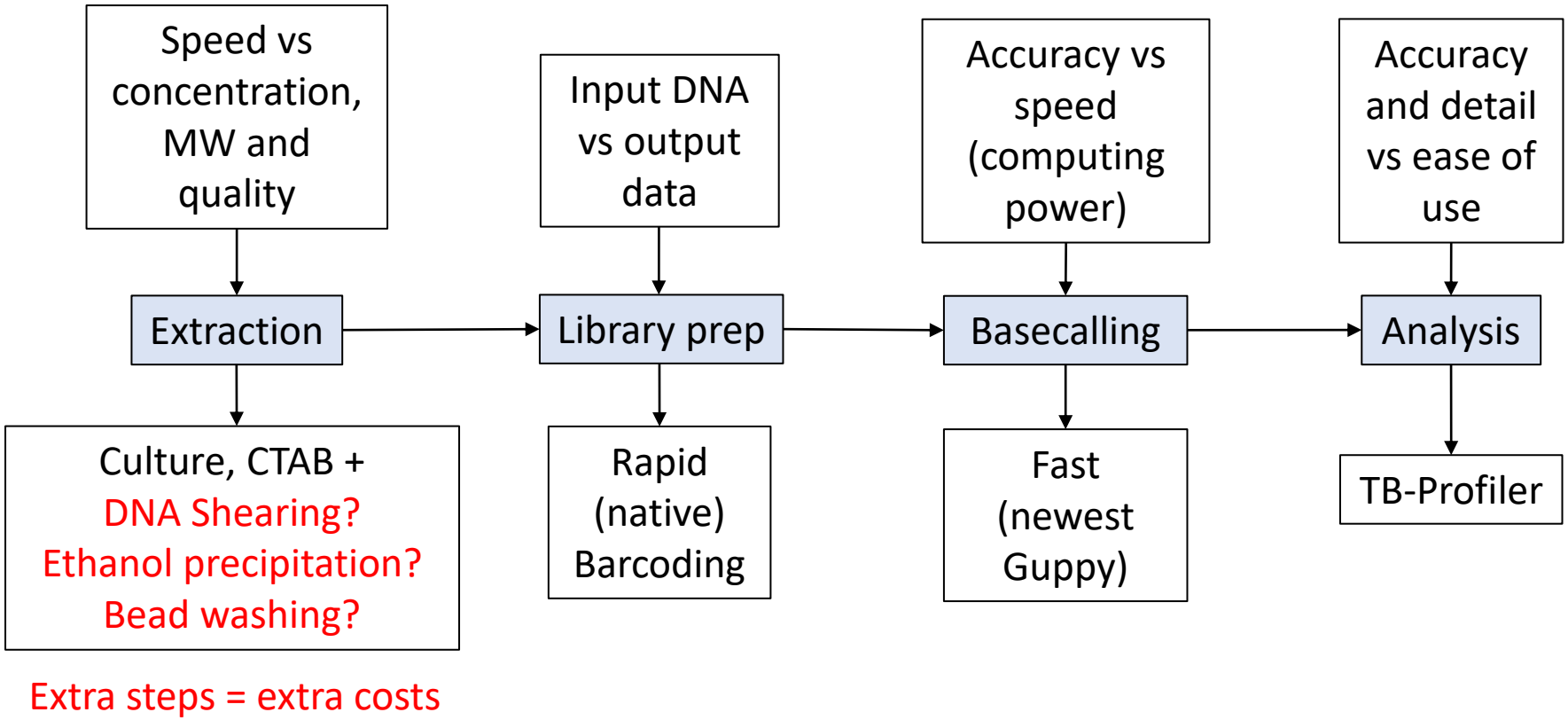
Is the DNA actually too long?

- Use Covaris tubes to shear the DNA to 10k and 20k
- Run each sample for 2 hours, wash the flow cell, compare to 50kb
- Covaris tubes are expensive

Is the DNA actually too long?



Sequencing TB seems to be a trade off between long reads (more data) and pore efficiency (sequencing for longer)



Extra steps = extra costs

Questions?