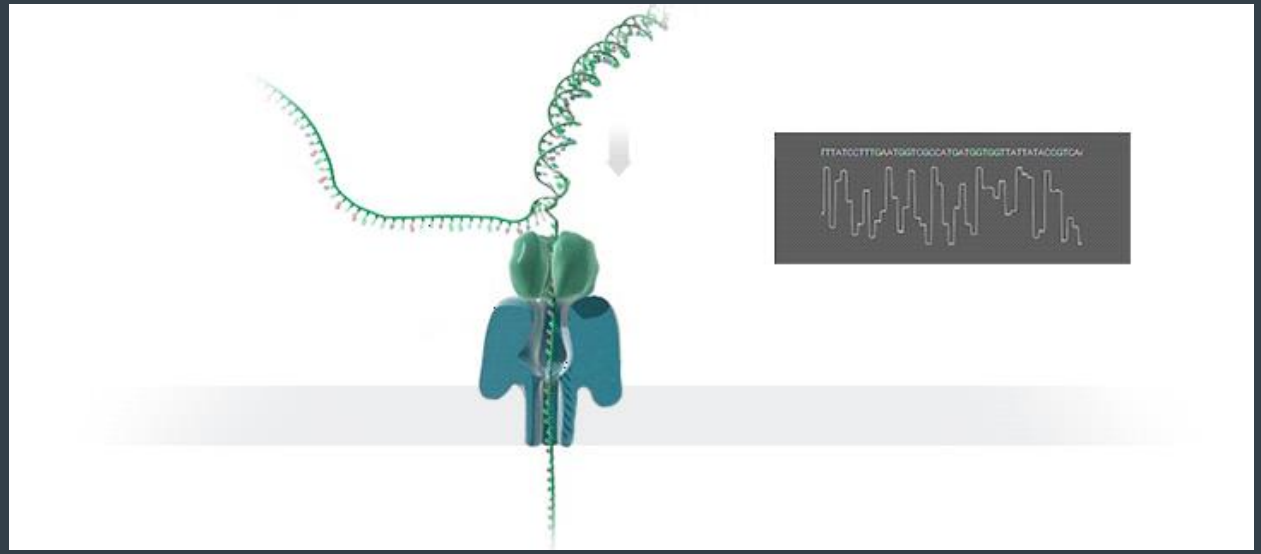
Quadram Institute Science Health Food Innovation

How to obtain good quality nucleic acid for MinION sequencing

Alp Aydin QIB

Quadram Institute Science Health Food Innovation

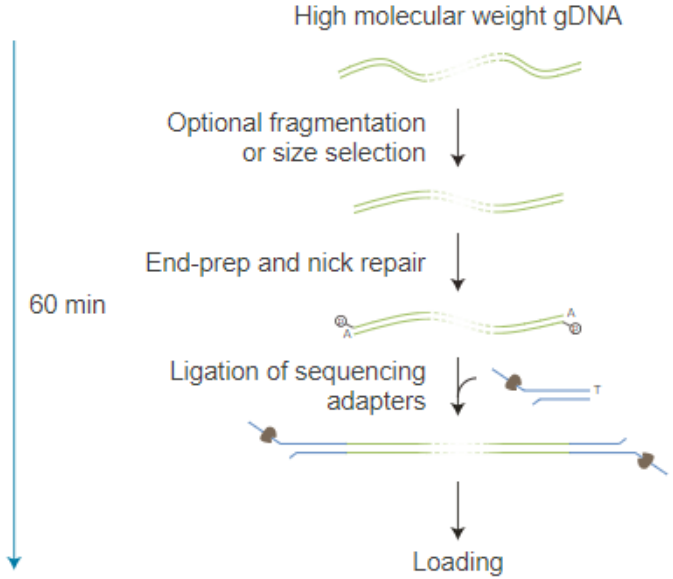
Sequencing chemistry



Types of kits

Broadly speaking, DNA:

1) Ligation Kits – adaptor is ligated to the end of DNA with ligase



PCR, multiplex, and 1D^2 options

Pros:

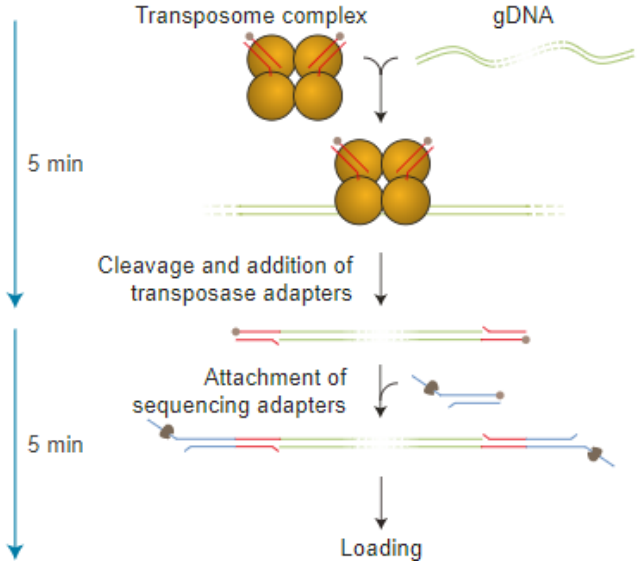
- Most control over size
- Highest yield

Cons:

- Longer
- More room for error
- More expensive (additional reagents)

Types of kits

2) Rapid Kits – transposase cuts DNA and allows adaptor to attach



PCR and multiplex options

Pros:

- Very fast
- Simple

Cons:

- Less control over size
- Lower yield

Types of kits

RNA:

1) Direct native RNA sequencing (Ligation based)

Pros:
- No reverse transcription or PCR bias

Cons:
- Much lower yield
- Early technology

2) cDNA sequencing (Ligation based)

Pros:
- Higher yield
- PCR option (lower input)

Cons:
- Takes much longer
- Requires reverse transcription

Extraction – up to you

- Extraction method will depend on what you are trying to extract and the purpose.
- **Generally**: Fresh cells and fresh extractions are best
- **Lysis**: Chemical, mechanical (bead-beating), enzymatic
- **Isolation**: phenol/chloroform, spin-column, SPRI beads
- **Consider**: Depletion or enrichment pre-extraction
- **Consider**: DNase or RNase step
- **Options**: Manual, automated

How much do you need? (DNA input)

- **Nanopore sequencing is not so much about mass (e.g. nanograms) but about molecules (e.g. femtomoles).**
- **If size and mass is known, calculate fmol. However, if size is difficult to (HMW) determine. Use input guidelines per kit.**

DNA Ligation kits

✓ **1000ng**: use as is *(per sample)*

➢ **100-500ng**: shear DNA

➢ **100ng**: use PCR (PSK004 kit)

➢ Short: calculate fmol, use **200 fmol**

DNA Rapid kits

✓ **400ng** (*per sample)*

➢ If you want longer reads: **1600+ ng**

➢ If low DNA, use PCR (RPB004 kit)

– input **1-10ng**

How much do you need? (RNA input)

Direct RNA sequencing

✓ **500ng** of poly-A RNA

cDNA sequencing

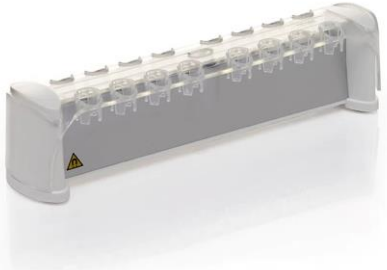
✓ **100ng** of input poly-A RNA

➢ If low, use PCR cDNA kit

- input **1ng** poly-A RNA or **50ng** total RNA

How much do you load?

- Follow all protocol-specified DNA washes. Purity is important for PCRs and sequencing.
  Biggest source of DNA loss



- If you end up with a surplus, do not be tempted to load all of it.
- Overloading will mean sequencing fuel is used up faster = lower yield
- If possible (e.g. for short DNA/Rapid Kits/PCR kits)
  measure size and quantity (mass), calculate to load ~50-100 fmol
- If still HMW after processing e.g. for Ligation kits, load ~400-600 ng

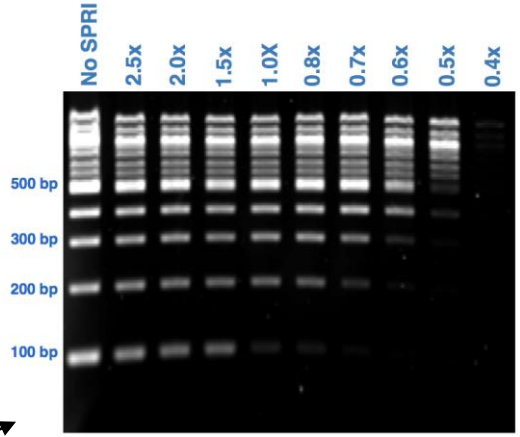
How long should it be? Up to you

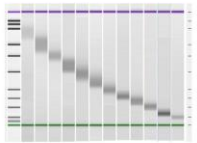
- No fragmentation or size selection required.
- Record read length: 2,200,000+ basepairs
- Longest reads achieved with Rapid Kit. More input = longer reads (but poor yield)
- Longer reads = better assemblies.
- More reliable sequencing with ligation kit if DNA is sheared (more ‘ligatable’ ends)
- Not degraded or too short - (Rapid Kit input recommendation >30kb)

Selecting the size

Size will depend on extraction but you can tweak average size after.

- **Shearing: Covaris g-Tubes**
- **Small fragment removal: SPRI beads (limited size range) 0.4x absolute minimum.**
- **Size fractionation device such as the SageELF**

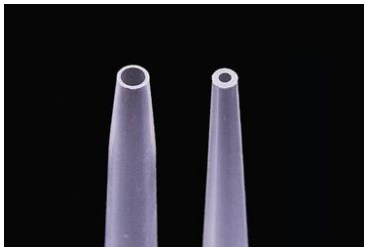




Handling DNA

To preserve long fragments (if a priority)

- **Avoid unnecessary pipetting or vortexing of DNA**
- **Avoid freeze-thawing DNA**
- **Use wide-bore pipette tips**



- **For 45kb+ DNA, drying step not recommended using SPRI beads**
- **Incubate DNA for longer and at higher temperatures during elution steps (e.g. at 37 C)**

Number of samples / yield

- Maximum MinION yield ~30 gigabases (normally lower)
- Don't think about how many samples you CAN multiplex, but how many samples you SHOULD multiplex
- Consider using other devices if yield is redundant or too low

PromethION max yield: ~148gb

Flongle max yield: ~1.8gb

Example:
- MinION run of stool sample yields **15gb**
- Detecting Campylobacter (1.6mb genome)
- Usually low proportion of reads (e.g. 1%)
- 1 Sample run = **93x coverage**
- 12 Sample run = **8x coverage**

Quadram Institute Science Health Food Innovation

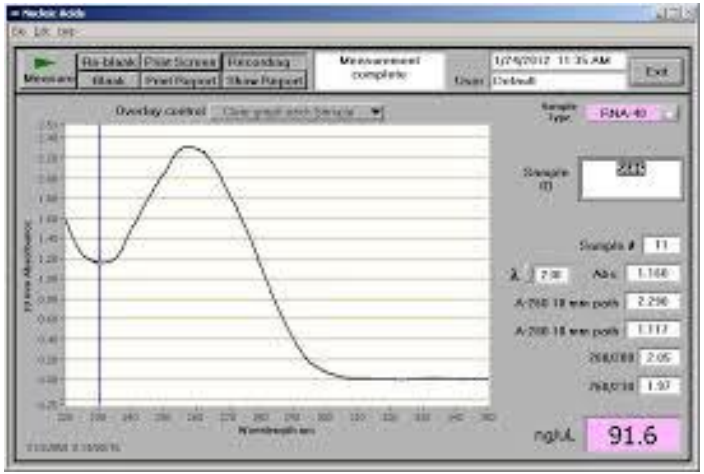
Checking your DNA/RNA

Quantity

- Check quantity with Qubit
- Use higher volumes (if possible to spare) for more accurate readings
- New standards each time you create the working buffer
- **Avoid Nanodrop** for quantification – overrepresents quantity

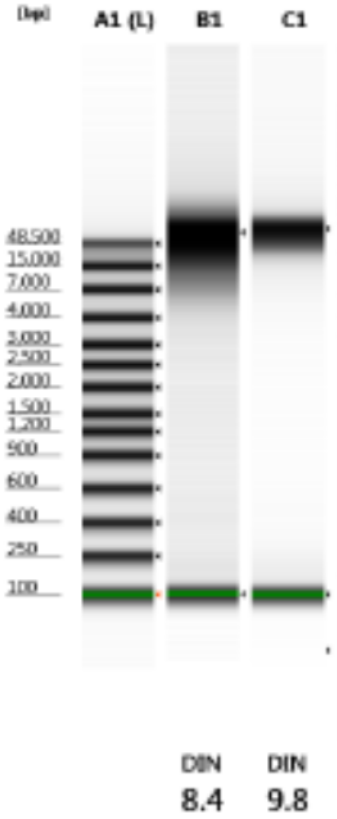
Quality

- Use Nanodrop to only check for potential contaminants
- 260/280 ratio of 1.8 = ‘pure DNA’
- 260/280 ratio of 2.0 = ‘pure RNA’
- 260/230 ratio of 2.0-2.2 = ‘clean’
  Lower = impure
- Carryover Ethanol not easily detected on Nanodrop
- Ratios unreliable at low concentrations

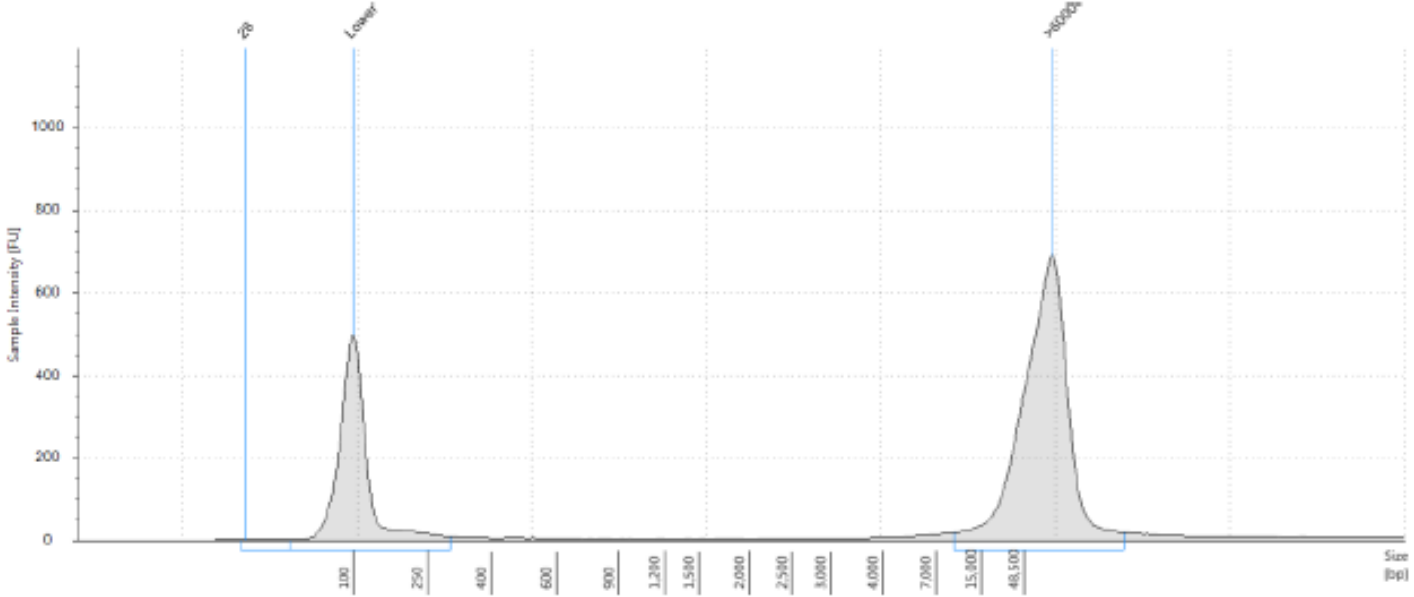


Size and integrity

- Use Tapestation for size and integrity (not degraded)
- DIN/RIN value = integrity (higher = better)
- Long fragments less reliable
  (don't assume 30kb = exactly 30kb)
- **Don't just focus on peak, focus on 100-1000bp area**



Tapestation – short fragments



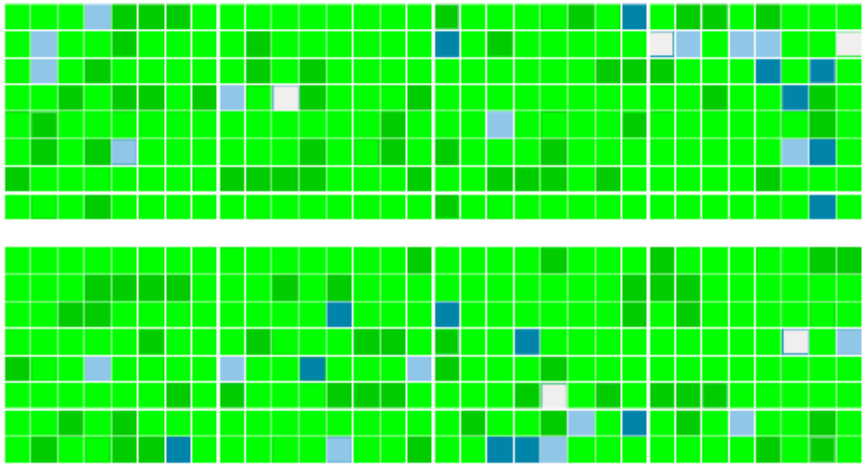
Checking quality by sequencing

- **Worth checking libraries on Flongle before expensive PromethION run**

Assessing your library in MinKNOW

Check for:

- **Available pores**
- **Pore occupancy**
- **Adaptors**



Quadram Institute Science Health Food Innovation

Remember for sequencing:

GARBAGE IN = GARBAGE OUT

Questions?