

# Genomics and Clinical Virology

## *Library preparation and Illumina sequencing*

### March 2025

## Metagenomic library preparation

### Monday 3rd March

CVR Genomics: **Ana Filipe, Lily Tong, Katherine Smollett**

UK HSA: Carmen Manso, **Dulcibella Boampong, Hodan Mohamed**

UCL Genomics: Charlotte Williams, Patricia Dyal Bynee, Cristina Venturini

Sanger: **Andrew Wang, Nashma Thesin**

# Day 1: Metagenomics

- Welcome & meet the team!
- The clinical scenario: You are running a hepatitis C clinic in a large hospital.
- You have received several samples that represent example cases.
- To begin, we will each take one sample and carry out NGS
- On Friday you will analyse the data to determine what treatment will be the best option for the patients.

Briefing



Medical  
Research  
Council

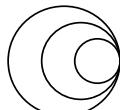


University  
of Glasgow  
VIA VERITAS VITA



CVR  
Centre for  
Virus Research

UK Health  
Security  
Agency



wellcome  
connecting  
science

# Day 1: Metagenomics

## A busy day: RNA > Sequencing

- You will receive nucleic acid samples, extracted from the serum of patients infected with HCV.
- Next, you will prepare libraries for **metagenomic** sequencing- **Kathy** will have her slides up and take you through each of the steps necessary to prepare the libraries, as described in the e-protocols. We will be around to help you at any stage.

# Lab protocol



## Genomics and Clinical Virology

2<sup>nd</sup> - 7<sup>th</sup> March 2025

## Laboratory Protocols – Illumina

### Table of Contents:

Introduction .....	4
Clinical scenario 1- Hepatitis C clinic .....	4
Sequencing non-enriched, metagenomic libraries in a MiSeq Illumina platform.....	5
Preparing capture-based target enrichment libraries .....	5
KITS AND REAGENTS .....	5
SAMPLES .....	5
Sequencing metagenomic libraries in a MiSeq Illumina platform .....	6
Extraction of samples and nucleic acid quality control .....	6
Library preparation using KAPA RNA HyperPrep Kit.....	6
INTRODUCTION .....	6
KITS AND REAGENTS .....	6
PROCESS WORKFLOW .....	7
PROTOCOL.....	8
Step 1. RNA Fragmentation and Priming .....	8
Step 2. 1 <sup>st</sup> Strand Synthesis .....	9
Step 3. 2 <sup>nd</sup> Strand Synthesis and A-tailing .....	9
Step 4. Adapter Ligation .....	10
Step 5. 1 <sup>st</sup> Post-ligation Cleanup .....	11
Step 6. 2 <sup>nd</sup> Post-ligation Cleanup .....	12
Step 7. Library Amplification .....	13
Step 8. Library Amplification Cleanup .....	14
Quality control, quantification and pooling of libraries.....	15
INTRODUCTION .....	15
PROTOCOLS .....	15
Step 1. Qubit quantification of NGS libraries.....	15
Step 2. Size distribution using TapeStation .....	17
Step 3 Calculation of molar concentration .....	21
Step 4. Prepare Equimolar pool.....	21
Quantification and dilution of pooled libraries .....	23
INTRODUCTION .....	23
PROTOCOL.....	23
Step 1. Qubit quantification of library pool .....	23
Step 2. Size distribution using TapeStation .....	23
Step 3. Calculate molar concentration of the pool.....	23
Step 4. Dilute library to 4 nM with Tris-HCl pH8, 10 mM .....	23
Step 5. Re-check the molar concentration of the pool.....	23

1



### Sequencing with Illumina MiSeq .....

MiSeq run setup .....	24
INTRODUCTION.....	24
DENATURING POOLED LIBRARIES AND PHIX CONTROL .....	24
Step 1. Denaturing pooled libraries.....	24
Step 2. Denature PhiX control V3 .....	25
Step 3. Dilute the 20 pM spiked library pool to 10 pM .....	25
Step 4. Dilute the 20 pM PhiX to 10pM .....	25
Step 5. Spike in the 10 pM library pool with 1% PhiX control .....	25
PREPARING THE REAGENT CARTRIDGE .....	26
PREPARING MISEQ INSTRUMENT .....	26
Step 1. Clean the Flow Cell .....	26
Step 2. Loading the Flow Cell .....	27
Step 3. Loading Reagents .....	27
Step 4. Load Sample Libraries onto Cartridge .....	27
Step 5. Sample Sheet Set-up .....	28
STARTING THE RUN.....	28
MiSeq Sequencing Metrics .....	29
Preparing capture-based target enrichment libraries .....	34
Quality control and pooling of NGS libraries for capture-based target enrichment .....	34
INTRODUCTION.....	34
PROTOCOL .....	34
Step 1. optional qPCR detection of target viral fragments .....	34
Step 2. Pooling of NGS libraries .....	34
Extra considerations when preparing the library pool .....	35
Capture-based target enrichment .....	36
INTRODUCTION.....	36
KITS AND REAGENTS .....	36
PROCESS WORKFLOW .....	37
PROTOCOL .....	38
HYBRIDISATION .....	38
Step 1. Prepare the Hybridisation Sample .....	38
Step 2. Prepare the Hybridization Master Mix .....	39
Step 3. Perform the hybridization incubation .....	39
WASHING AND RECOVERING CAPTURED DNA LIBRARY POOL .....	40
Step 1. Prepare Hybridization Wash Buffer and Bead Wash Buffers .....	40
Step 2. Prepare the Capture Beads .....	40
Step 3. Bind Hybridized DNA to the Capture Beads .....	41
Step 4. Wash the Capture Beads Plus Bound DNA .....	42
AMPLIFYING CAPTURED DNA LIBRARY POOL USING LM-PCR .....	43
Step 1. Set post-capture PCR .....	43
Step 2. Library amplification clean-up .....	44
Captured pool quantification and dilution .....	45
MiSeq run setup .....	45
MiSeq sequencing metrics .....	45
APPENDIX I Reagent list .....	46

2



### Metagenomic library preparation .....

Library Quality control .....	47
MiSeq Run .....	47
Sequence Capture .....	48

### APPENDIX II Workflows for alternative scenarios .....

#### Clinical scenario 2- Haemorrhagic fever .....

Metagenomic recovery of RNA virus genomes .....	49
INTRODUCTION .....	49
KITS AND REAGENTS .....	50
PROTOCOL .....	50
QUALITY CONTROL AND LIBRARY POOLING AND SEQUENCING .....	57
REFERENCES .....	57

#### Clinical scenario 3- HIV .....

Library preparation using Illumina NexteraXT Kit .....	59
KITS AND REAGENTS .....	59
INTRODUCTION .....	59
WORKFLOW .....	61
PROTOCOL .....	62
Step 1. Input DNA preparation .....	62
Step 2. DNA fragmentation .....	62
Step 3. Library amplification .....	63
Step 4. Library cleanup .....	64
QUALITY CONTROL AND LIBRARY POOLING AND SEQUENCING .....	65

#### Clinical scenario 4: Paediatric hepatitis outbreak .....

Metagenomic library preparation to sequence RNA and DNA viruses .....	66
INTRODUCTION .....	66
WORKFLOW .....	67
PROTOCOL .....	67
Targeted enrichment using pan-viral probe library preparation .....	67
INTRODUCTION .....	67
WORKFLOW .....	67
PROTOCOL .....	67

### APPENDIX III Modifications for target capture > 40 Mbp .....

### APPENDIX IV TapeStation analysis and troubleshooting:

Adapter Dimers .....	71
Daisy Chains .....	71
Large fragment contamination .....	72

### Appendix V Metagenomic library preparation results .....

3

# Day 1: Metagenomics

We will have several small lab talks during the day:

- **Kathy** will talk to you about **lab contaminations** and how to avoid them.
- After the morning break **Lily** will explain you how you can use **magnetic beads** during library preparations.
- In the afternoon **Kathy** will discuss **lab quality control** applied to NGS, followed by **molarity calculations** which are important to pool the NGS libraries together before loading in the sequencer.
- **Lily** will help you to do the quality control of your libraries using the **Qubit** and **TapeStation** instruments.
- We will **summarise the results** of your libraries.
- **Videos: diluting/denaturing pool of libraries, Miseq run, prepare cartridge/ flowcell/ samplesheet.**

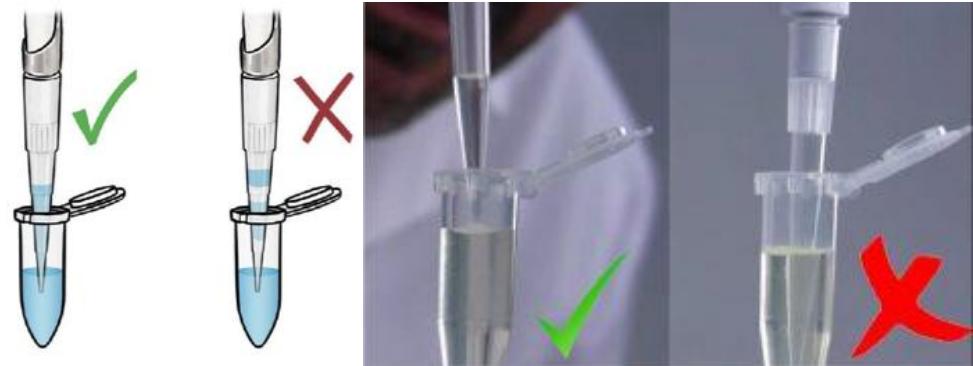
# Day 2: Targeted-enrichment

- **Kathy** will start by explaining how the **Illumina sequencer is working** to produce your dataset.
- **Lily** will let you have a peak at **your sequencing run** and talk you through **important metrics** associated to a sequencing run.
- **Ana** will **review** the main steps of the **library prep** of the previous day.
- That will lead into the **lab work of day 2**, where you will experience hands-on how to do enrichment of viral targets, using a **probe-capture approach, applied to HCV**.
- **Ana** will take you through each of the **steps necessary to perform the protocol**, using as starting material libraries, equivalent to the ones you prepared on day 1.
- **You will have a final chance to ask questions** before we say goodbye around 3 pm.

# Best lab practices to avoid variable results

(Reagent List, Appendix I, Manual page 46)

- ✓ Double-check volumes- look at the tip



- ✓ Vortex and spin down all reagents



- ✓ Record everything

- Write down the indices (barcode) associated to your samples- you will need them to identify your dataset
- Always label your sample tubes or mixes of reagents- use ethanol proof markers ☺

# Best lab practices to avoid variable results

**Magnetic beads** (Ampure/ HyperPure) or streptavidin magnetic beads are a special case

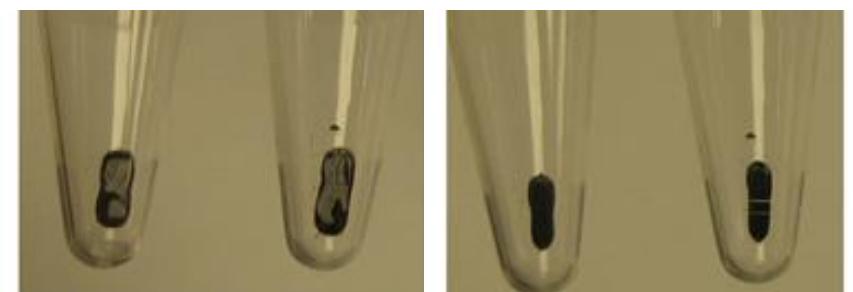
✓ Resuspend, resuspend, resuspend. They are heavy > sediment very quickly

- The first time you resuspend them it takes a bit longer
- Be careful with spin down (super fast and always check for clumps)



✓ When on the magnet they are clumped

- Do not disturb them during the ethanol washes
- Always make sure the supernatant is clear before removing it



✓ After removing ethanol, do not over-dry

# Best lab practices to avoid variable results

Ensure that all **temperature-sensitive** steps are conducted in the specified conditions

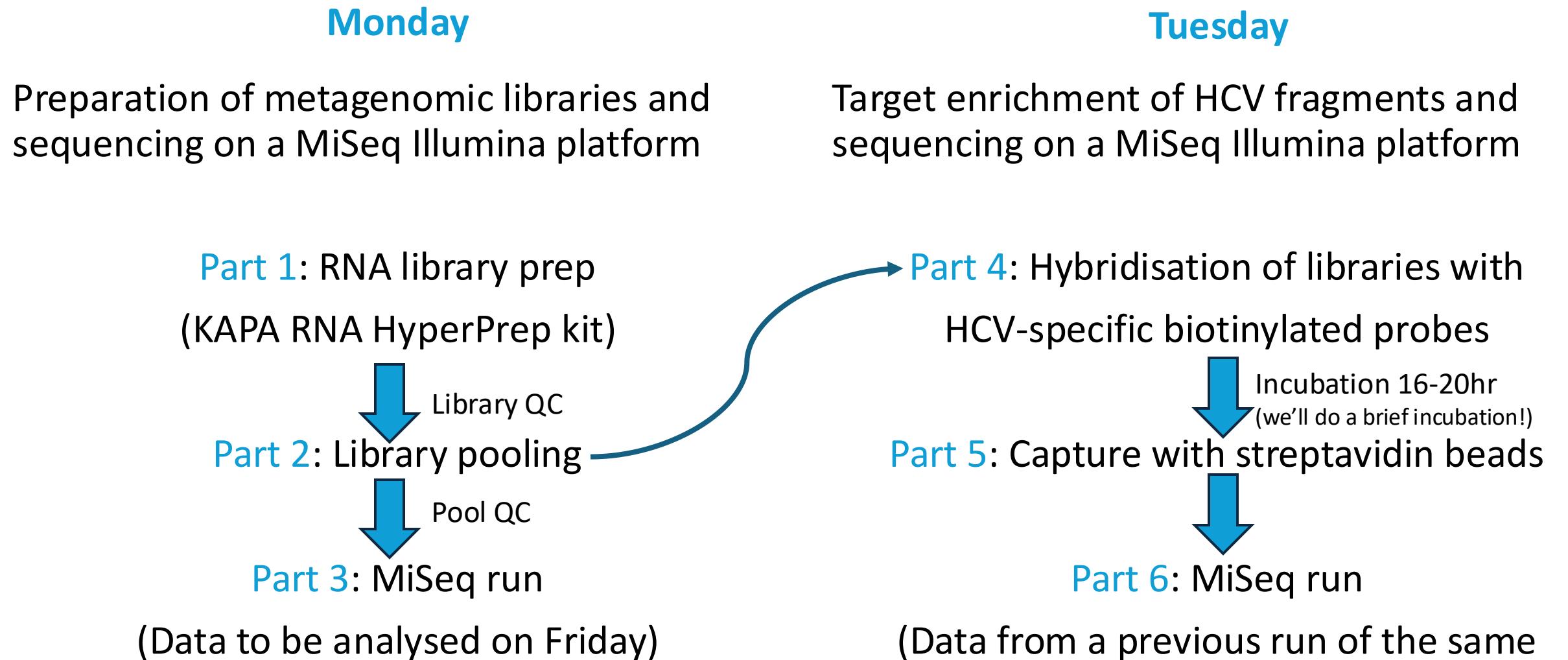
✓ Sample and reagents always on ice, except

- Magnetic beads and streptavidin magnetic beads
- Some reagents on day 2 will have to be incubated at 55°C before being used

✓ Replace ice as needed- ice machine location



# Workflow overview



# Workflow overview

9 working groups (2 people) :

Each group will be given 2 nucleic acid extracts (10µl)

## Nucleic acid extraction

- NUCLISENS easyMAG® extraction (other extraction methods are also suitable)
- Usually, extracts are treated with DNase to remove DNA present in the sample

Wide diversity among samples:

- HCV Ct ~21-33
- ~8-30 ng/µl DNA, undetectable RNA

We will discuss how the ratio of HCV to host material impacts upon the sequencing results

Sample	HCV Ct in the extracts	RNA Qubit quantification (ng/µl) of the extracts	DNA Qubit quantification (ng/µl) of the extracts
Sample 1a	21	Undetected	10
Sample 1b	23.1	Undetected	13.8
Sample 2a	22.5	Undetected	23
Sample 2b	27.2	Undetected	7.8
Sample 3a	25.4	Undetected	21.6
Sample 3b	20.96	Undetected	18.84
Sample 4a	25.82	Undetected	23.09
Sample 4b	26.04	Undetected	17.75
Sample 5a	31.54	Undetected	17.3
Sample 5b	22.11	Undetected	19.22
Sample 6a	32.79	Undetected	12.64
Sample 6b	25.62	Undetected	27.64
Sample 7a	24.13	Undetected	19.82
Sample 7b	26.48	Undetected	12.59
Sample 8a	27.49	Undetected	14.06
Sample 8b	22.24	Undetected	14.35
Sample 9a	25.47	Undetected	13.78
Sample 9b	29.9	Undetected	17.1

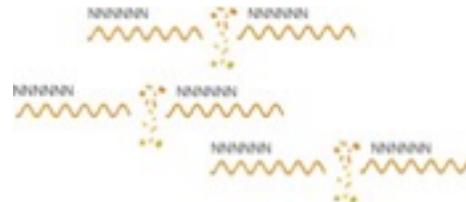
# Day 1: Metagenomics

## Part 1: RNA library prep using Kapa RNA HyperPrep Kit

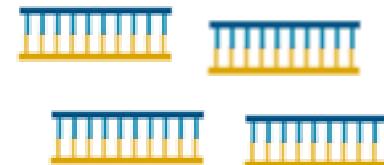
(Library prep kit compatible with Illumina sequencing platform)



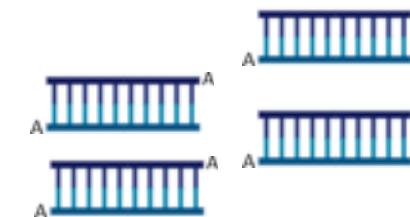
10µl of nucleic acid extract



Step 1: RNA fragmentation and priming.  
Heat in the presence of divalent cations ( $Mg^{2+}$ )



Step 2: cDNA first strand synthesis



Step 3: cDNA second strand synthesis and A tailing

Step 8: Library clean-up



Step 7: PCR amplification



Step 5&6: Double clean-up



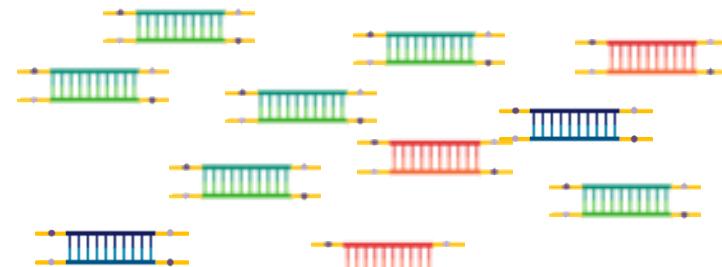
Step 4: Adapter ligation

Host origin

Viral origin

Others

e.g. bacteria, fungi,  
contaminants,  
artefacts



Index sequence: Usually 8-10 bp. Different between adapters

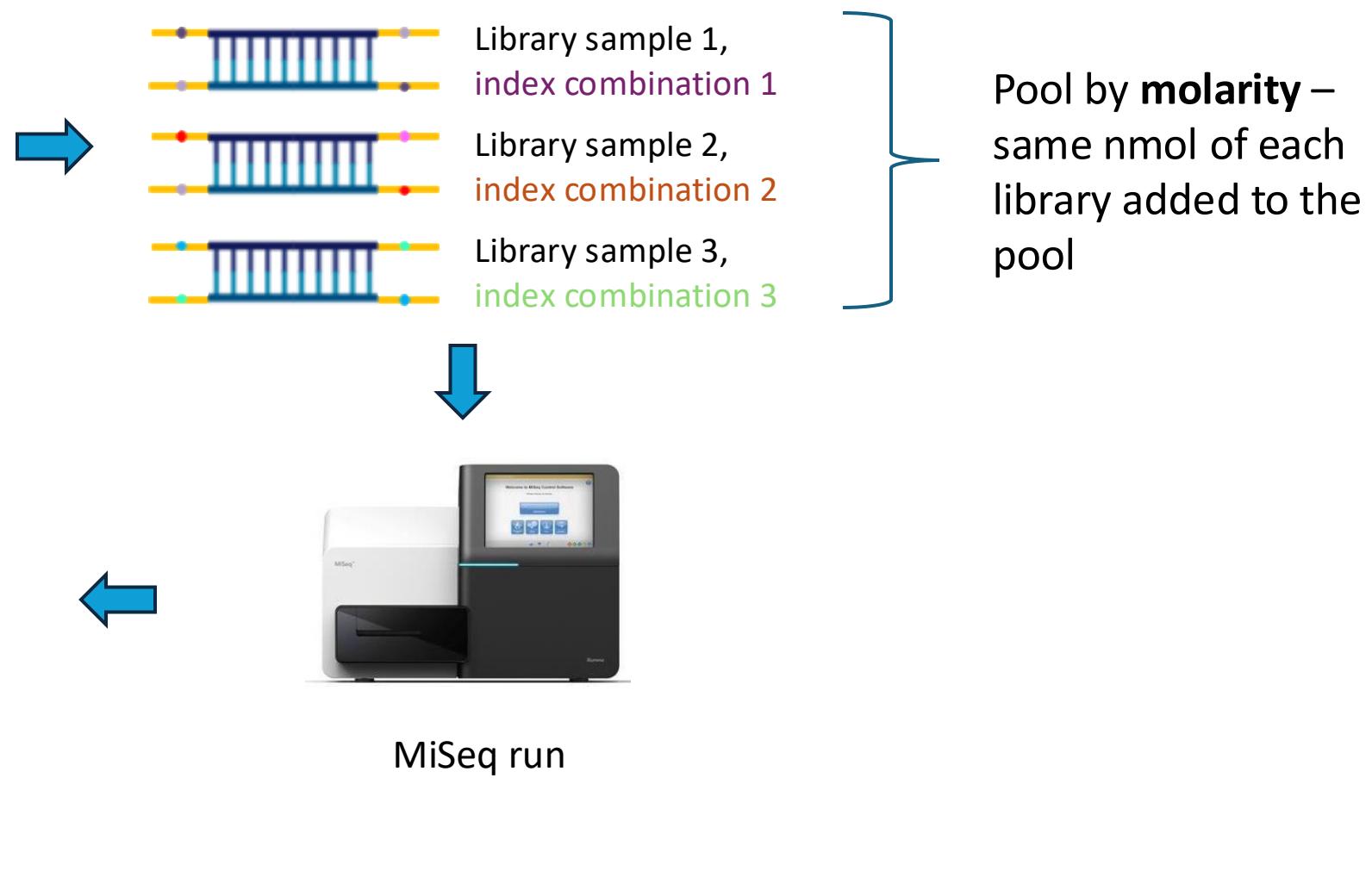
Each sample/library is labelled using a unique combination of i5-i7 adaptors

# Day 1: Metagenomics

## Part 2 & 3: library pooling and sequencing

### LIBRARY QC

- Qubit quantification of each library to determine total library yield (ng/ $\mu$ l)
- Mean fragment size (automated electrophoresis)





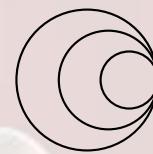
Medical  
Research  
Council



University  
of Glasgow



CVR  
Centre for  
Virus Research



wellcome  
connecting  
science

# Step-by-step

# Step 1: Fragmentation and priming

## What is the optimal fragment size and why?

- The required fragment size will depend on the chosen sequencing chemistry



*MiSeq Illumina platform:  
Reads one base per cycle*

- We will perform 150 bp paired end sequencing with 300 cycle MiSeq Reagent kit



*Shorter fragments would generate redundant data  
Longer fragments would not cover the middle of the sequence*

# Step 1: Fragmentation and priming

## How to get the optimal fragment size

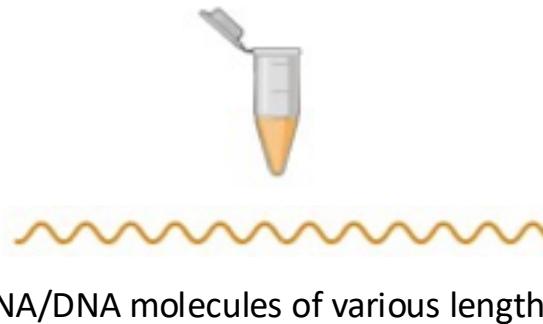
Input RNA type	Desired mean library insert size (bp)	Fragmentation
Intact	100-200	8 min at 94 °C
	200-300	6 min at 94 °C
	300-400	6 min at 85 °C
Partially degraded	100-300	1-6 min at 85 °C
Degraded (e.g. FFPE)	100-200	1 min at 65 °C

- Fragmentation step by incubation at high temperature in presence of Mg<sup>2+</sup>
- Incubation time dependent on input RNA quality
- Manufacturers' specification often refers to library preps for human DNA.
- RNA viral genomes are generally very small and fragmentation times will require optimization

# Step 1: Fragmentation and priming

- Nucleic acid fragmentation by heat incubation in the presence of divalent cations ( $Mg^{2+}$ )
- Hybridization of random primers

10 $\mu$ l of nucleic acid extracts  
(optional DNase treatment)

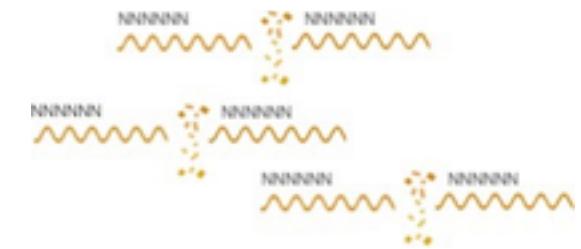


Mix with Fragment, Prime and Elute buffer, which contains:

- $Mg^{2+}$
- Random primers

Incubate 1 min at 85°C to achieve  
≈300-350bp fragments

Place in ice to stop fragmentation and encourage the hybridization of fragmented RNA and random primers



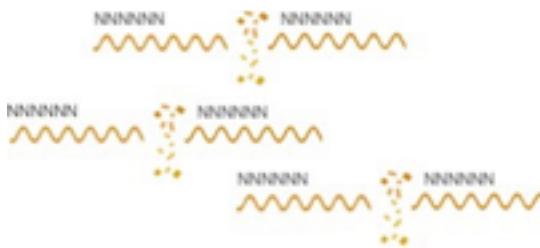
RNA/DNA molecules of various lengths

Lid Colour	Aliquot label	Component	Volume
		Purified RNA (1-100 ng)	10 $\mu$ l
Green	FPE	Fragment, Prime and Elute Buffer (2X)	10 $\mu$ l
		TOTAL:	20 $\mu$ l*

\*note typo in manual!

# Step 2: First strand cDNA synthesis

Fragmented RNA molecules and random primers

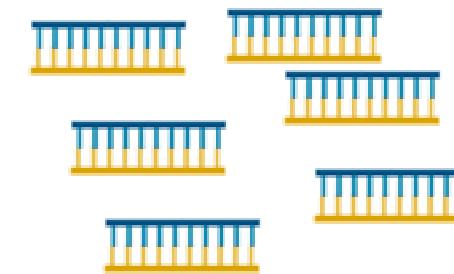


Add:

- 1<sup>st</sup> strand synthesis buffer
- KAPA script (reverse transcriptase)

Note: many RT have RNaseH-like domain to digest RNA template

dDNA – RNA duplex



Master mix:

Lid Colour	Aliquot label	Component	Master mix volume (inc. 20% excess)
Pink	1 <sup>st</sup> SS	1 <sup>st</sup> Strand Synthesis Buffer	11 µl
Blue	KP	KAPA Script	1 µl
		TOTAL:	12 µl

Reaction mix:

Component	Volume
Fragmented RNA (from step 1)	20 µl
1 <sup>st</sup> strand master mix	10 µl
TOTAL:	30 µl

Incubate:

Step	Temp.	Duration
Primer extension	25°C	10 min
1 <sup>st</sup> strand synthesis	42°C	15 min
Enzyme inactivation	70°C	15 min
HOLD	4°C	∞



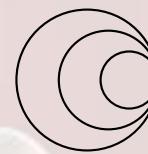
Medical  
Research  
Council



University  
of Glasgow



CVR  
Centre for  
Virus Research



wellcome  
connecting  
science

# Contamination control

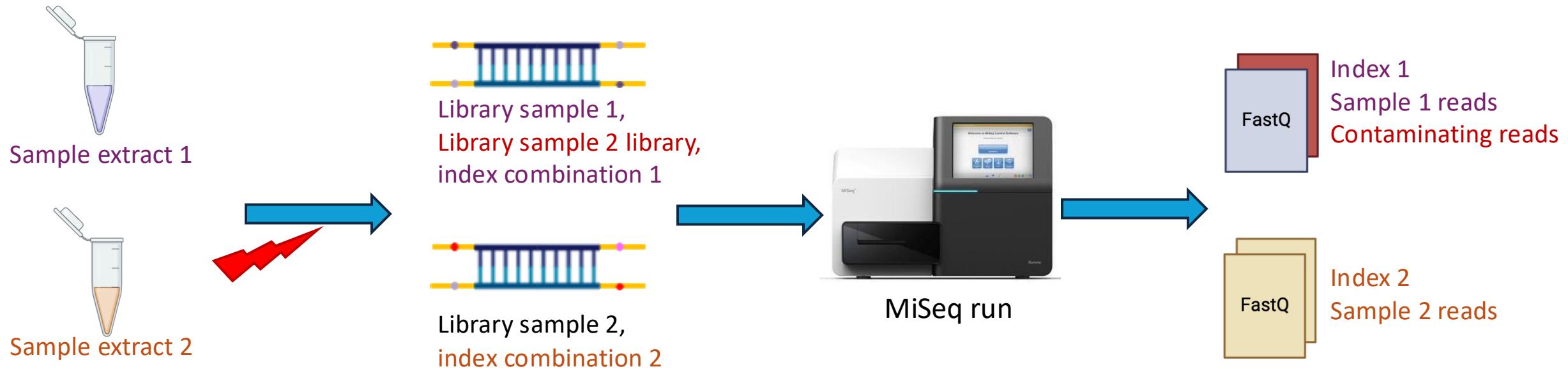
# Contamination control

NGS assays are highly susceptible to multiple sources of contamination

Contamination event:

→ sample 1 index with sample 2 fragment

→ sequencer assigns sample 2 fragment into sample 1 FASTQ file during demultiplexing



# Contamination control

**Effect of contamination dependant on project:**

- Sample/library type
- Throughput
- Viral load
- Known/unknown viral targets

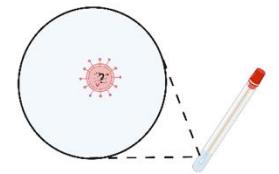
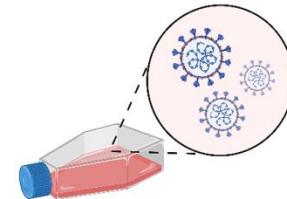
**Key problem for:**

- Minority variant studies
- Metagenomic studies

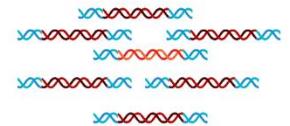
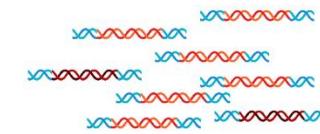
**Bioinformatic tools may help**

- Level of contaminating fragments often very low
- May not map to expected reference

Sample type:



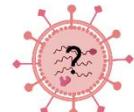
Library type:



Throughput:



Analysis:



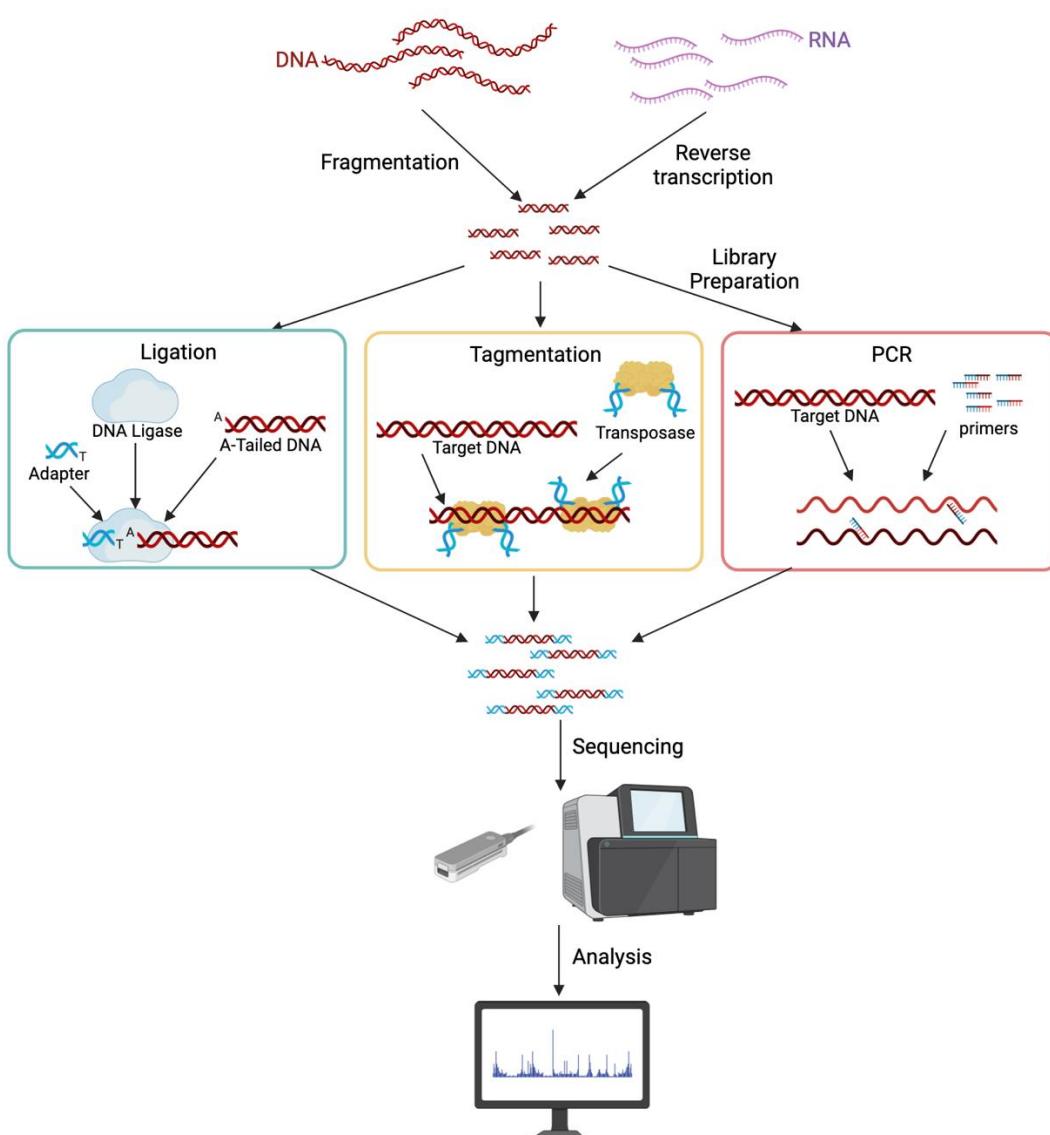
# Main sources of contamination

Contamination can occur at all stages

1. Contamination during library prep/seq capture

2. Contamination during sequencing

3. Index contamination



Increasing concentration, increasing viral load, more readily amplified

# Managing contamination

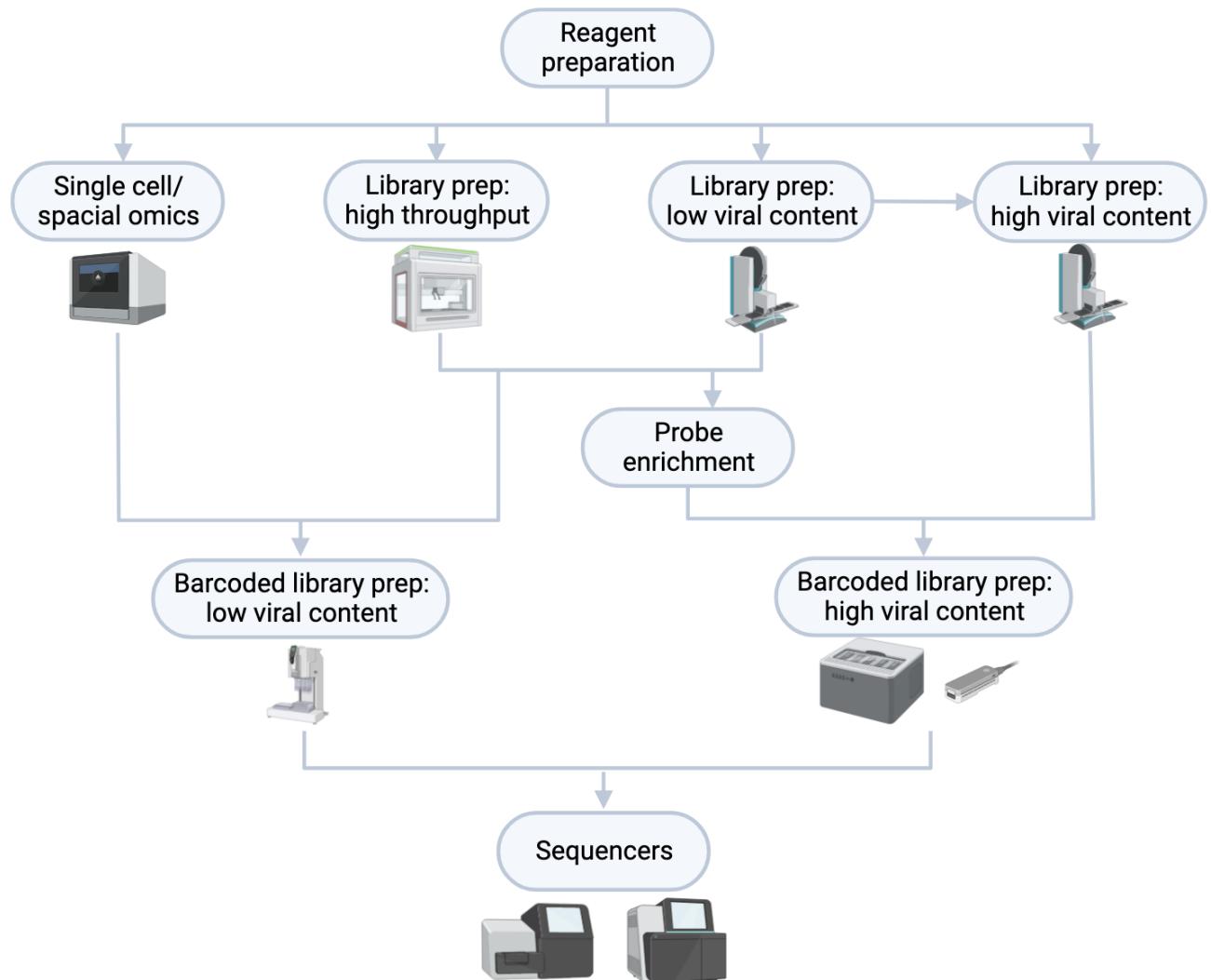
## Rigid procedures

- Unidirectional workflow, split by
  - Viral load
  - Stage of preparation
- Cleaning procedure
- Regular glove changing

## Guidelines

- Negative controls
- Spike-in controls
- Rotating indexes
- Use unique dual indexes

Unidirectional workflow at the CVR:



# Contamination during library prep

## Intra-run (within batch) contamination:

**Source:** Sample mishandling before indexed-adapter ligation reaction

### *How to monitor intra-run contamination?*

- Use of numerous, appropriate negative controls

### *How to minimise intra-run contamination?*

- Good laboratory practice
- Automation

## Inter-run (between batch) contamination:

**Source:** Current run contaminated with libraries from previous run

### *How to monitor inter-run contamination?*

- Rotate the positions of the negative controls in consecutive runs

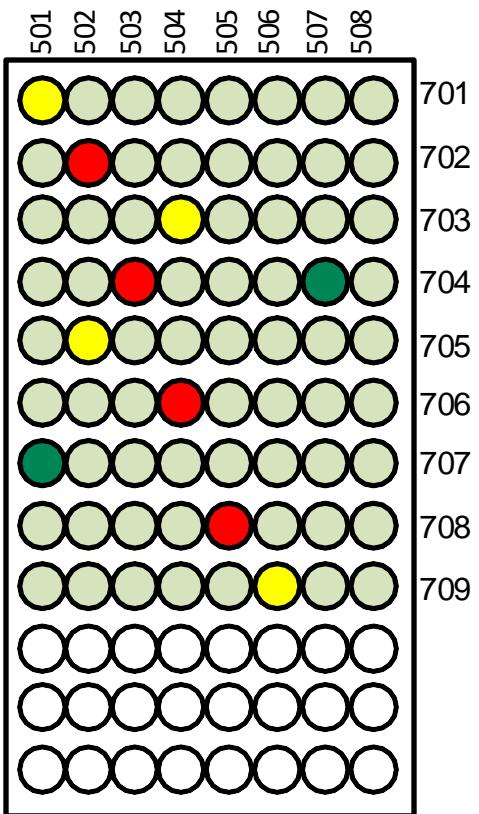
### *How to minimise inter-run contamination?*

- Cleaning workspaces, pipettes etc. between batches
- Changing adapter combinations

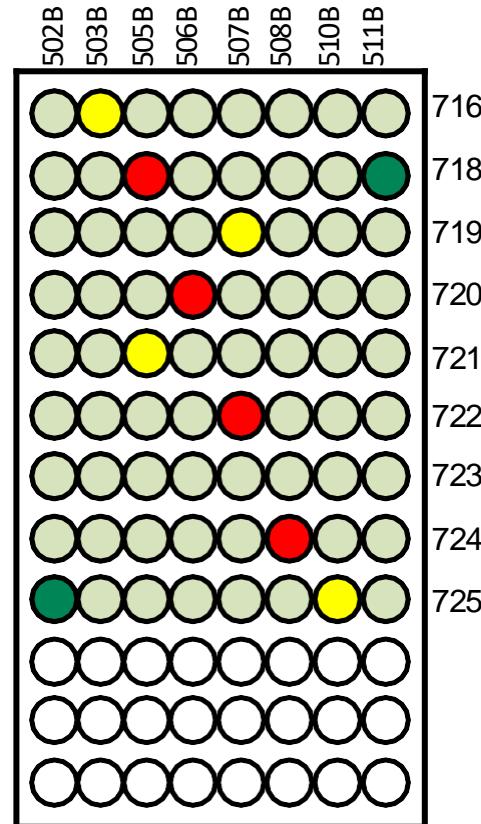
# Typical HCV WGS run in UKHSA

- 4 negative plasma control
- 4 water controls
- 2 Positive control
- 62 plasma samples

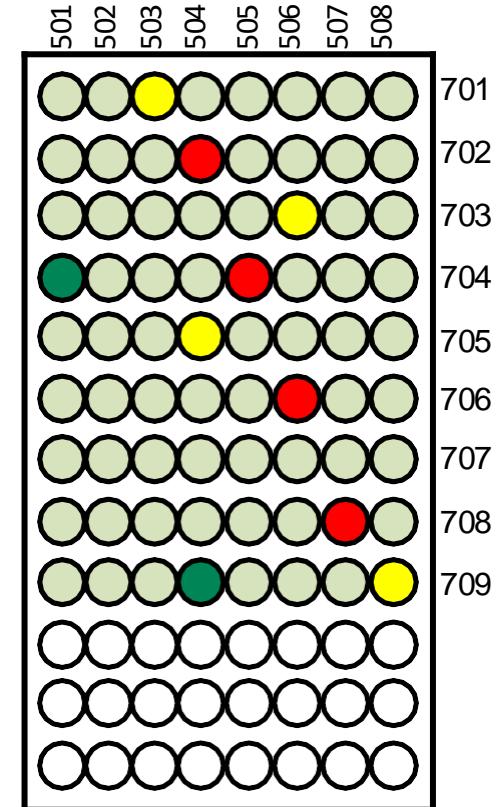
Run 1 (adapters set A)



Run 2 (adapters set B)



Run 3 (adapters set A)



Talk – contamination control



Medical  
Research  
Council

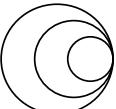


University  
of Glasgow  
VIA VERITAS VITA



CVR  
Centre for  
Virus Research

UK Health  
Security  
Agency



wellcome  
connecting  
science

# Contamination during sequencing

## Carry over contamination

**Source:** Libraries from a previous run present in the template line of a MiSeq instrument can contaminate the current run (typically <0.1% - issue has been acknowledged by Illumina)

## How monitor carry-over contamination?

- Use different adapter combinations in consecutive runs
- Rotate the positions of the negative controls

## How to minimise carry-over contamination?

- Rigorous cleaning of the MiSeq instrument

Table 2 Maintenance During Normal Operation

Activity	Frequency
Post-Run Wash	After every run
Maintenance Wash	Monthly
Standby Wash	To prepare for idle mode (if unused for $\geq 7$ days), and every 30 days the instrument remains idle
Instrument Shutdown	As needed

# Indexes contamination

## Intra-run contamination

**Source:** mishandling of the indexes leads to contamination of an index stock tube with another index

## *How to monitoring Indexes contamination?*

- Analysis of previous run. Reads from the sample labelled with the contaminated indexed will appear in the FastQ file from sample labelled with “contaminant” index

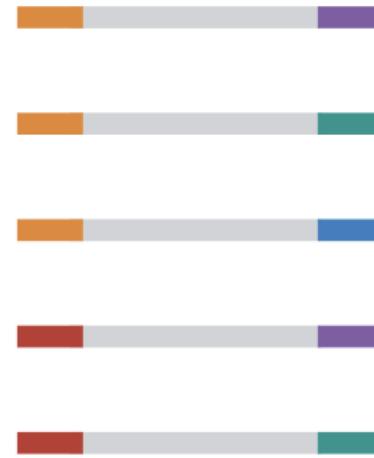
## *How to minimise carry-over contamination?*

- Good laboratory practice
  - spin adapter tubes/plates
  - clean bench surface
  - clean tips
  - etc
- Automation

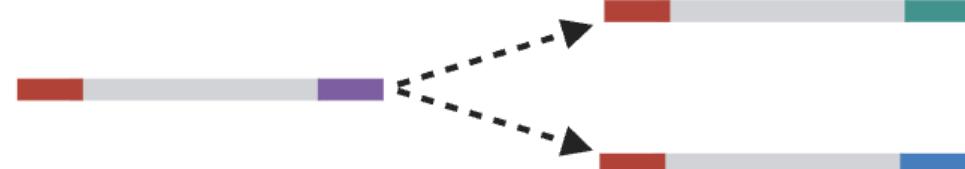
# Index hopping

- Index hopping also called:
  - index switching
  - index swapping
  - index cross-talk
  - index contamination
  - index mis-assignment
- Occurs when multiplexed samples are sequenced on NGS platforms
- Samples from an expected index incorrectly assigned to different index in the pool
- Caused by presence of free adapter dimers

Correct Indexes:



Index hopping:



# Index hopping

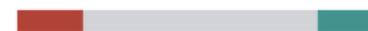
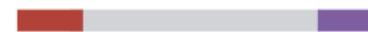
Affected by:

- Input concentration – higher for lower input samples
- Flow cell type - higher in patterned flow cells than random flow cells
- Library prep method – higher in ligation-based methods
- Level of PCR amplification – extra cycles can reduce free-adapter
- PCR of pooled samples – index swapping can occur during pooled PCR
- Index type – use unique dual indexes NOT combinational indexes
- Library storage – store at -20°C

Problem for:

- Very deep sequencing (needle in haystack e.g metagenomics where a couple of reads could result in false positives)
- Very low signal-to-noise ratio – e.g. some single cell or rare transcript detection

Correct Indexes:



Index hopping:



# Index hopping

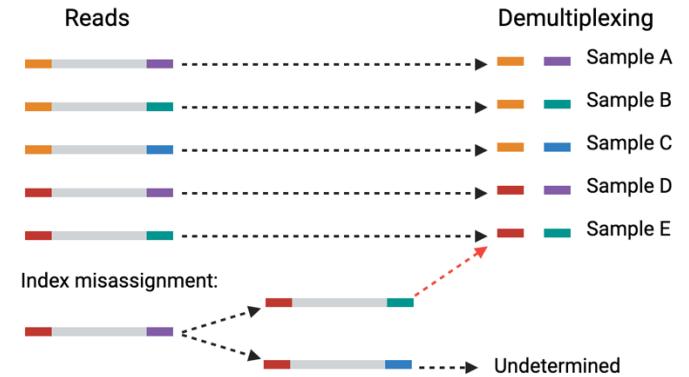
## How to monitor index hopping:

- Monitor by calculating the % single hopped indexes

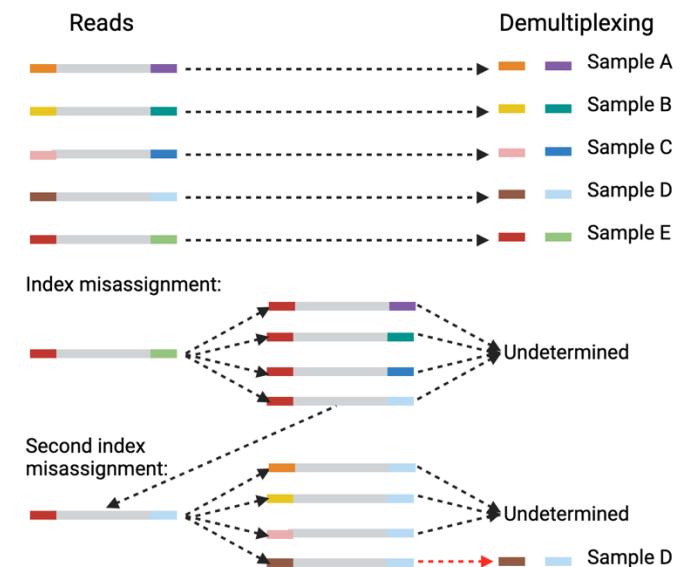
## How to minimize index hopping:

- Remove free adapters from library preps
- Store libraries individually at -20°C
- Pool libraries prior to sequencing
- Use unique dual indexing pooling combinations (unique i5 and i7 indexes)

### Combinational indexes

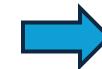
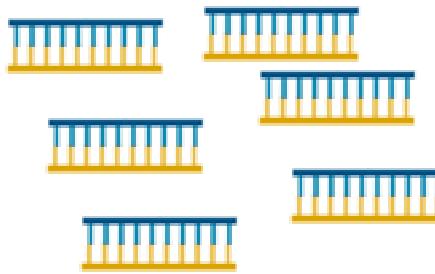


### Unique dual indexes



# Step 3: Second strand cDNA synthesis and A-Tailing

dDNA-RNA duplex

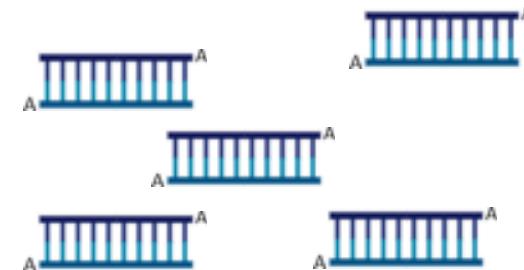


Add:

- 2<sup>nd</sup> strand synthesis buffer
- 2<sup>nd</sup> strand synthesis and A-Tailing enzyme mix

Note: during this step, the 2<sup>nd</sup> strand is marked with dUTP and dAMP is added to the 3'-ends

dsDNA with A-tail overhangs



Master mix:

Lid Colour	Aliquot label	Component	Master mix volume (inc. 10% excess)
Red	2 <sup>nd</sup> SM	2 <sup>nd</sup> Strand Marking Buffer	31 µl
Yellow	2 <sup>nd</sup> Enz	2 <sup>nd</sup> Strand Synthesis and A-Tailing Enzyme mix	2 µl
		TOTAL:	33 µl

Reaction mix:

Component	Volume
1 <sup>st</sup> strand product (step 2)	30 µl
2 <sup>nd</sup> strand master mix	30 µl
TOTAL:	60 µl

Incubate:

Step	Temp.	Duration
2 <sup>nd</sup> strand synthesis	16°C	30 min
A-tailing	62°C	10 min
HOLD	4°C	∞



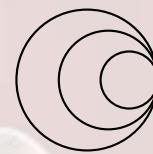
Medical  
Research  
Council



University  
of Glasgow



CVR  
Centre for  
Virus Research

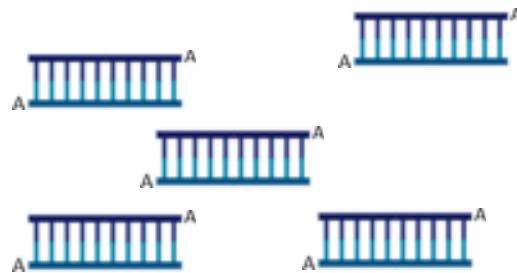


wellcome  
connecting  
science

# Coffee break!

# Step 4: Adapter ligation

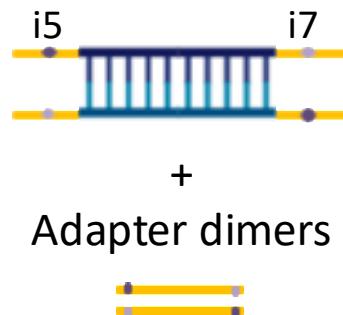
dsDNA with A-tail overhangs



Add:

- Ligation buffer
- T4 DNA ligase
- Adapters (with unique index sequences)

Adapter-ligated dsDNA fragments



Master mix:

Lid Colour	Aliquot label	Component	Master mix volume (inc. 10% excess)
Blue	LB	Ligation Buffer	40 µl
Pink	Lig	DNA Ligase	10 µl
		TOTAL:	50 µl

Reaction mix:

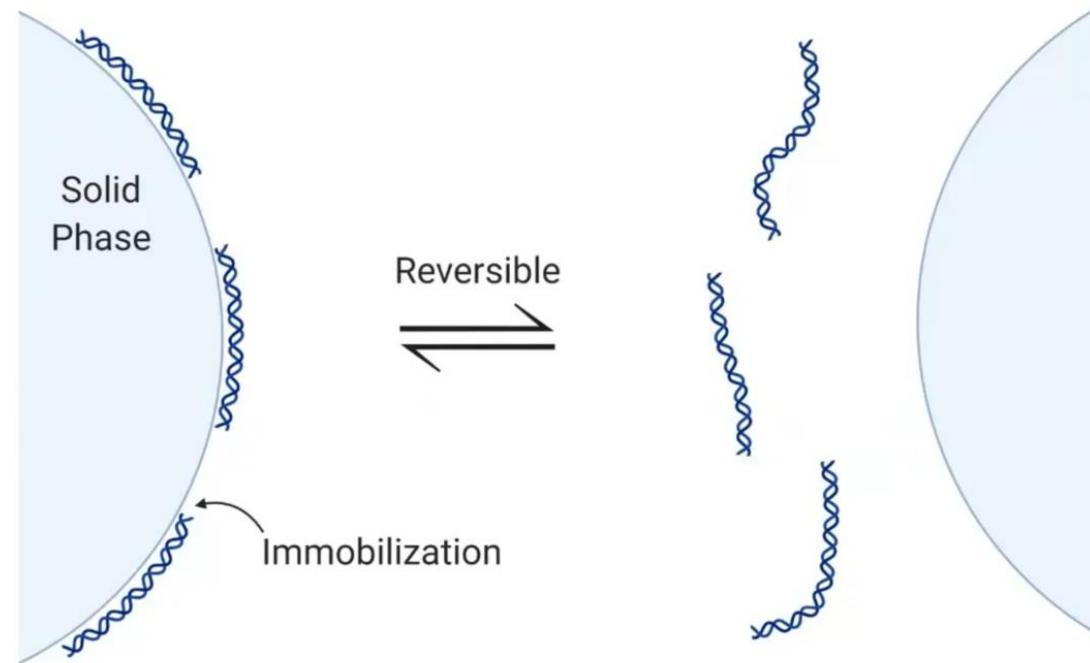
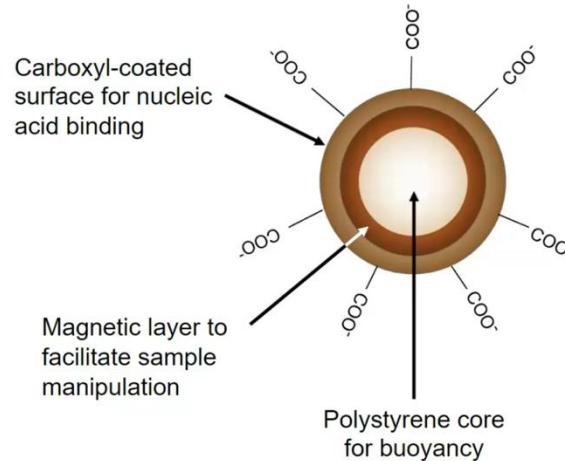
Component	Volume
2 <sup>nd</sup> strand synthesis product (from step 3)	60 µl
Adapter ligation master mix	45 µl
Adapter (1.5 mM)	5 µl
TOTAL:	110 µl

Incubate:

20°C for 15 min

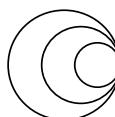
# Bead clean-ups for size selection

# SPRI Bead-Based Technology



Images from <https://www.mybeckman.uk/landing/genomics/spri-technology-webinar>

Talk – bead clean-up



# SPRI Reagent - Main Components

## 1. Magnetic bead

- Provides solid support for nucleic acid binding

## 2. Crowding agent

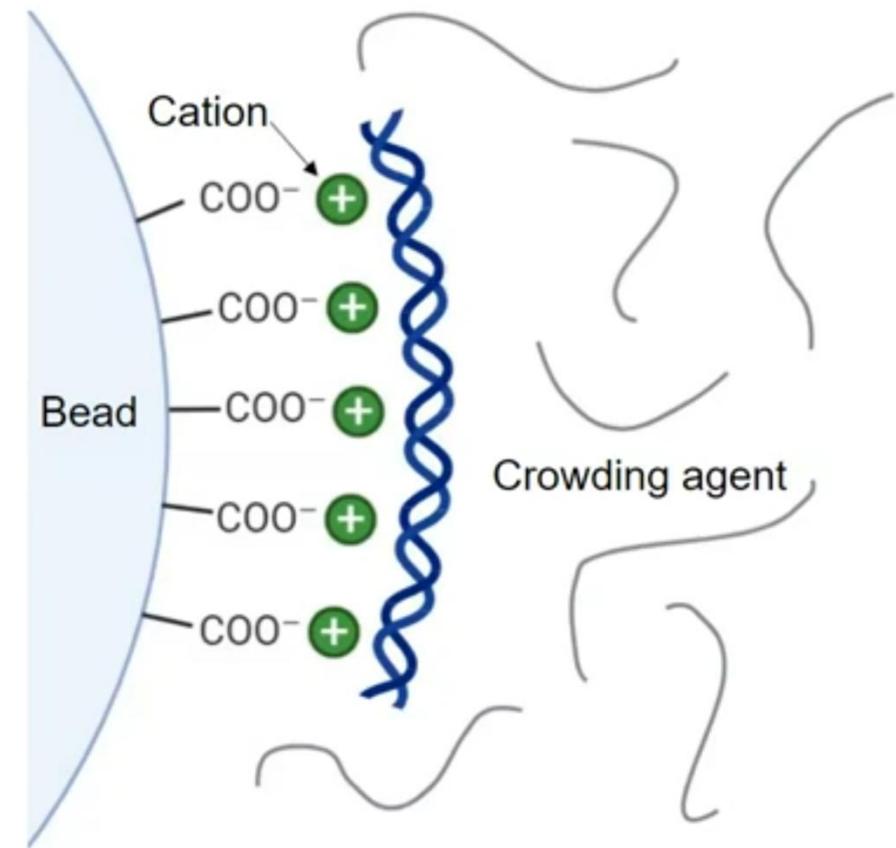
- Polyethylene glycol can promote nucleic acids of certain size out of solution.

## 3. Binding buffer

- Provides salt ions and adjusts pH to achieve the binding of nucleic acid and beads

Nucleic acid binding bead mixes:

- DNA mix: 10 mM Tris base, 1 mM EDTA, 2.5 M NaCl, 20% PEG 8000, 0.05% Tween 20, pH 8.0 @ 25 °C
- RNA mix: 1 mM trisodium citrate, 2.5 M NaCl, 20% PEG 8000, 0.05% Tween 20, pH 6.4 @ 25 °C



*DeAngelis MM, Wang DG, Hawkins TL. Solid-phase reversible immobilization for the isolation of PCR products. Nucleic Acids Res. 1995 Nov 25;23(22):4742-3. doi: 10.1093/nar/23.22.4742. PMID: 8524672; PMCID: PMC307455.*

Talk – bead clean-up



Medical  
Research  
Council

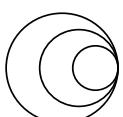


University  
of Glasgow  
*VIA VERITAS VITA*



CVR  
Centre for  
Virus Research

UK Health  
Security  
Agency



wellcome  
connecting  
science

# SPRI – main steps



## Step 1: BINDING



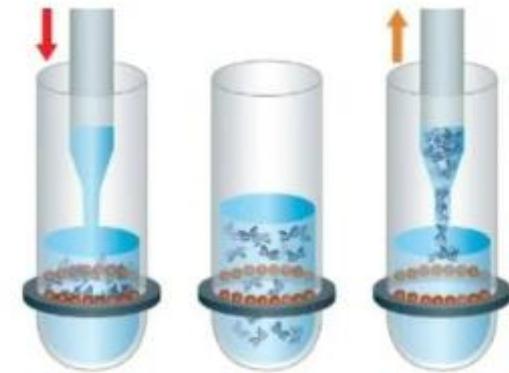
In the presence of a proprietary buffer, nucleic acids bind to the beads' surfaces

## Step 2: WASHING



Beads are immobilised with a magnet, facilitating contaminants removal

## Step 3: ELUTE



Nucleic acids are then released in an aqueous solution

Talk – bead clean-up



Medical  
Research  
Council

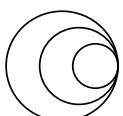


University  
of Glasgow  
VIA VERITAS VITA



CVR  
Centre for  
Virus Research

UK Health  
Security  
Agency

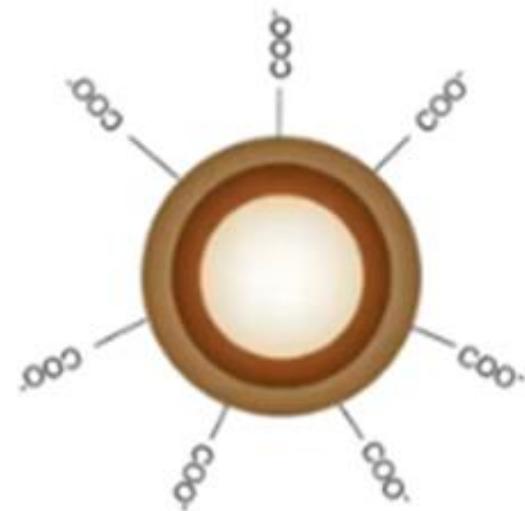


wellcome  
connecting  
science

# Quiz

Which of the following does not promote nucleic acid binding to the SPRI beads?

- Binding buffer
- Carboxyl-coated surface
- Crowding agent
- Magnetic layer



Talk – bead clean-up



Medical  
Research  
Council



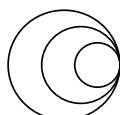
University  
of Glasgow  
VIA VERITAS VITA



CVR  
Centre for  
Virus Research

UK Health  
Security  
Agency

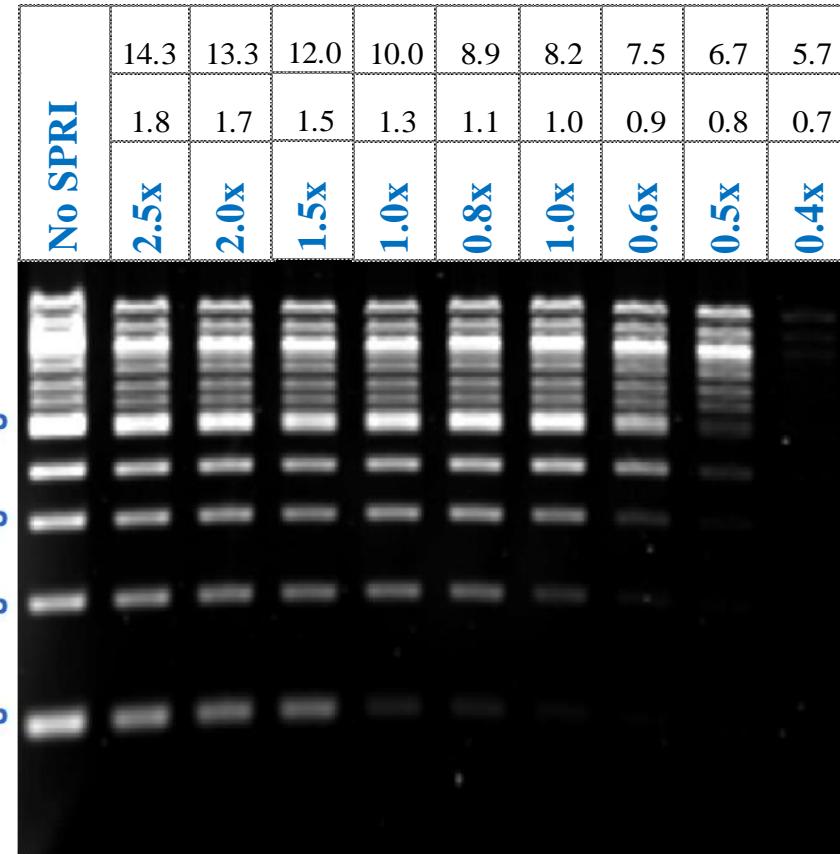
**UCL**



wellcome  
connecting  
science

# Size selection

- depends on the **CONCENTRATION** of the PEG/NaCl



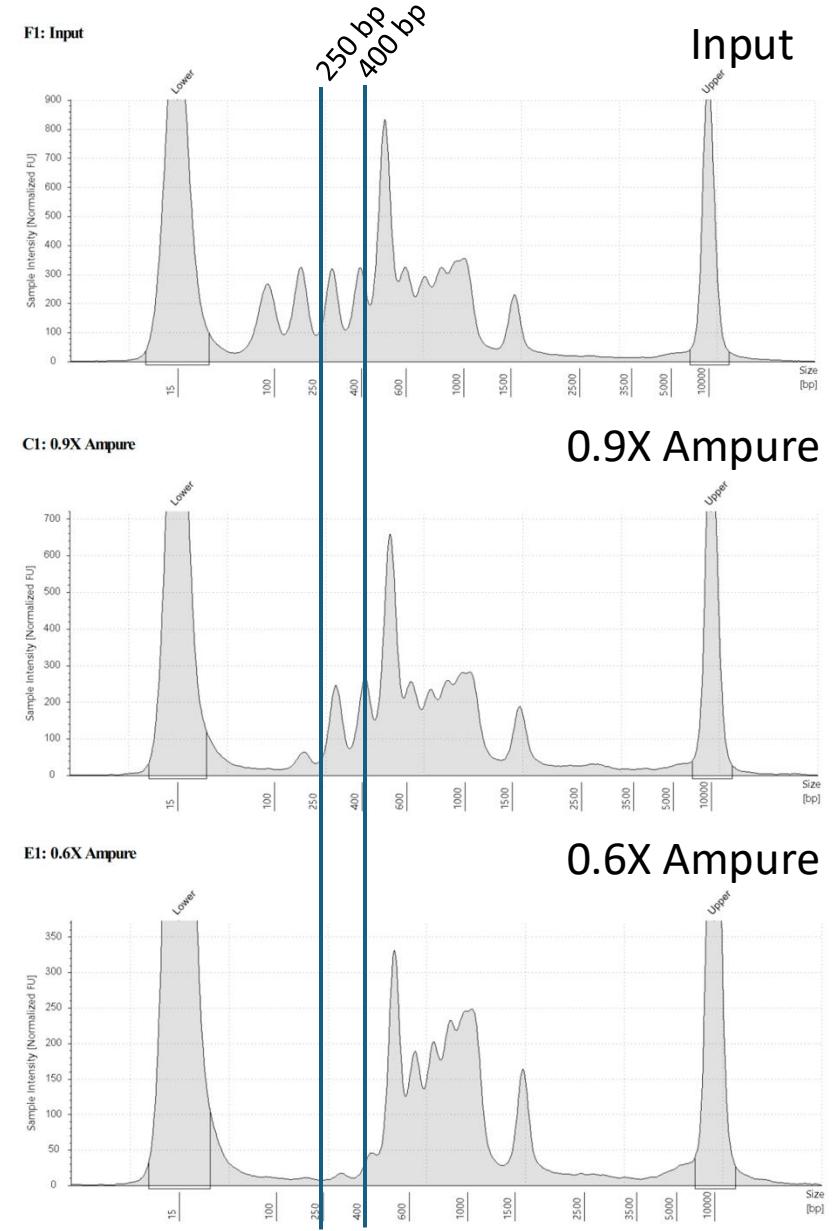
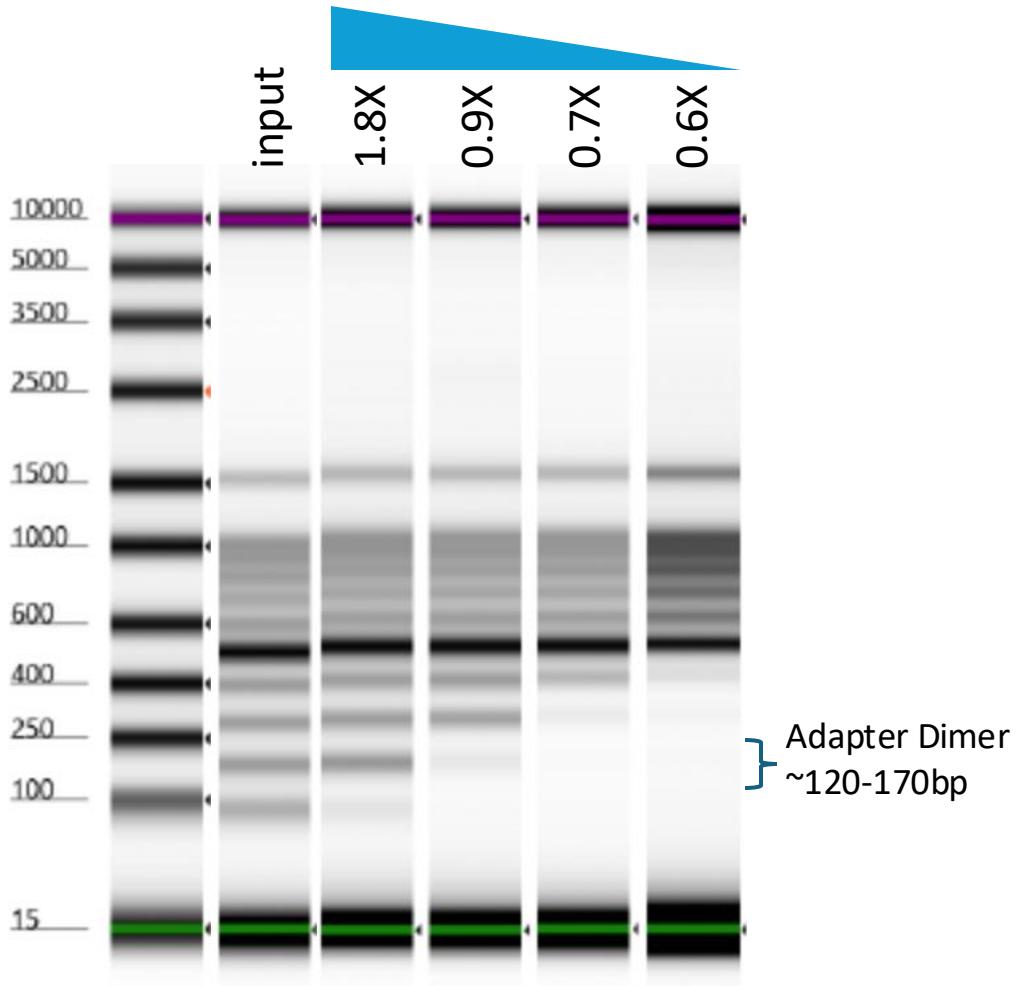
Conc. PEG%

Conc. [Na+] mol/L

SPRI Vol: Sample Vol



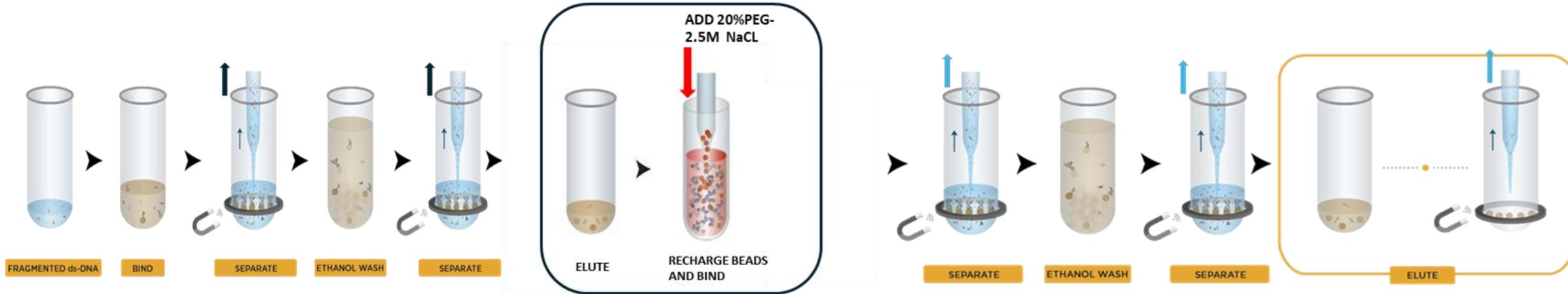
# Fine tuning by changing ratio



Talk – bead clean-up

# Double clean-up

- Recharge beads capacity binds to nucleic acid



- We will perform a double clean-up using 0.63X followed by 0.7X bead ratio

# Double-sided clean ups

## Step 1. Low conc. SPRI Binding

**0.5 x SPRI**  
*Fragments > 250 bp on beads*  
*Fragment < 250 bp in solution*

## Step 2. Keep Supernatant and discard beads

Keep supernatant  
0-300 bp

**Discard Beads**  
 (Reverse SPRI)

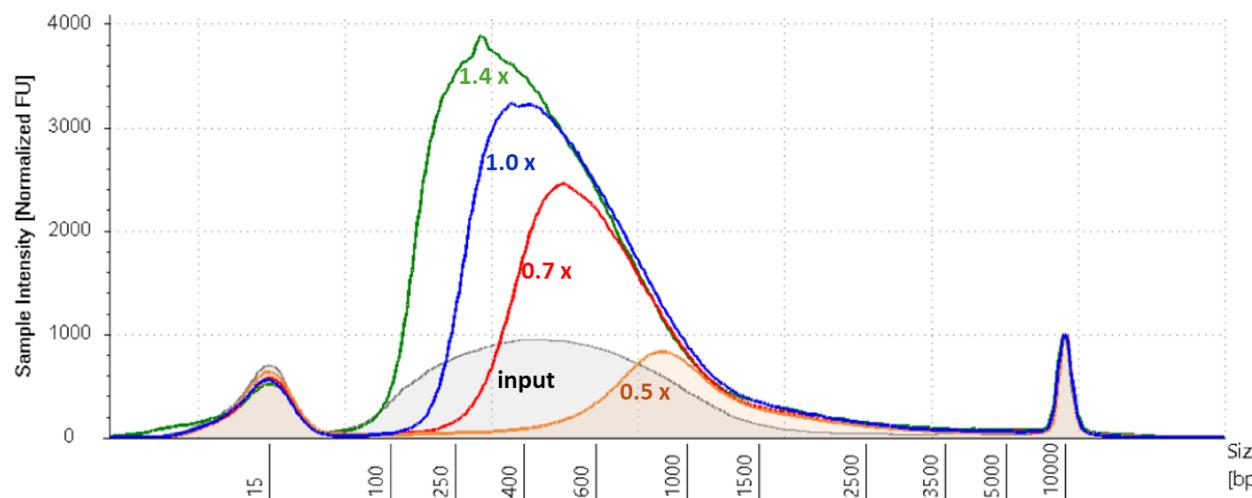
## Step 3. Top up fresh SPRI Higher conc. SPRI binding

**1.0 x SPRI**  
*Fragments > 100 bp on beads*  
*Fragment < 100 bp in solution*

## Step 4. Keep Beads, Wash and Elute

**Keep Beads**  
100 -250 bp

**Discard supernatant**



# Good practice when performing bead clean-ups

- ✓ Always equilibrate beads to RT before use)
- ✓ Beads should always be homogenous
  - resuspend with a pipette or vortex
- ✓ Ensure supernatant is clear before removing
- ✓ Do not disturb the beads on the magnet
- ✓ Don't over-dry beads
- ✓ Don't carry over beads in eluate.



BIND

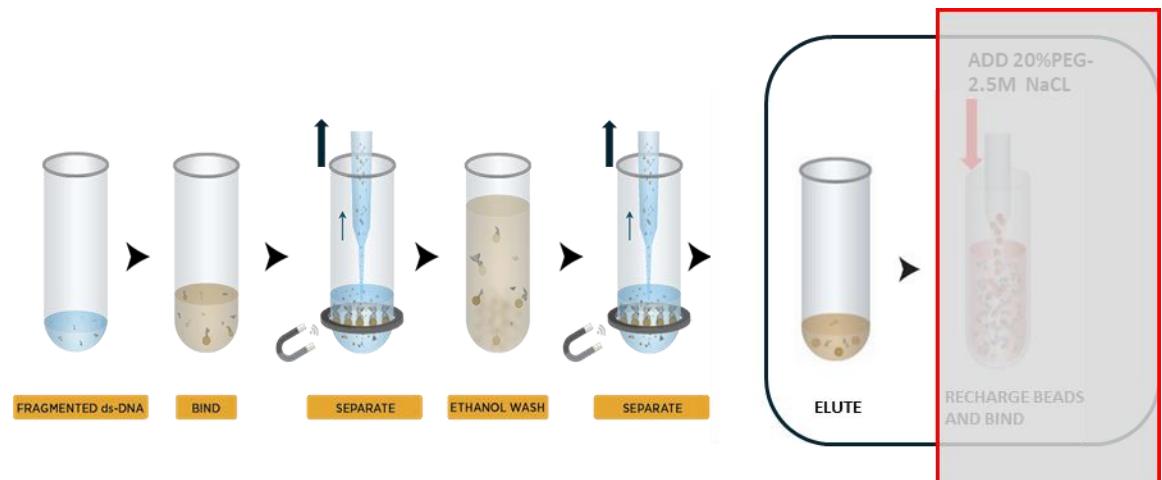


SEPARATE

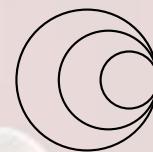
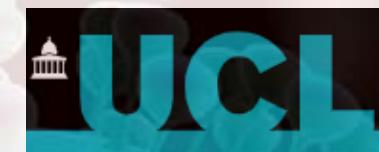
# Step 5: 1<sup>st</sup> post-ligation clean-up

- Add Ampure beads for 0.63X clean up
  - Adapter ligation reaction from step 6 = 110 µl
  - $0.63 \times 110 = 70 \mu\text{l}$  Ampure beads
- Incubate at Room temp. 5-10 min
- Collect beads on magnet and remove supernatant
- 2X washes in 200 µl 80% EtOH to remove salt and PEG
- Air-dry beads 3-5 min to remove EtOH
- **Resuspend** beads in 50 µl Tris-HCl (pH 8.0-8.5)

Lid Colour	Aliquot label	Component	Volume
		Adapter-ligated DNA (step 4)	110 µl
	AMP	Ampure XP beads	70 µl
		TOTAL:	180 µl



**Please Note!** We are performing a double clean-up, so do not transfer the first eluate. PEG/NaCl will be added to rebind the DNA to the beads for the second clean-up

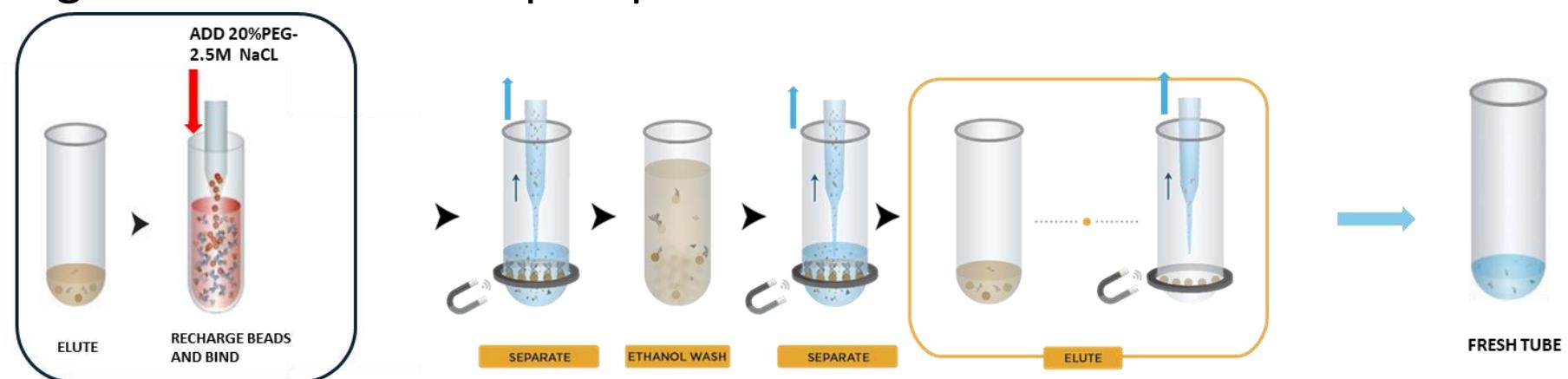


# Lunch break!

# Step 6: 2<sup>nd</sup> post-ligation clean-up

- Add PEG/NaCl solution for 0.7X clean up
  - Bead/Tris mix from Step 5 = 50 µl
  - $0.7 \times 50 = 35 \mu\text{l}$  PEG/NaCl
- Incubate at Room temp. 5-10 min
- Collect beads on magnet and remove supernatant
- 2X washes in 200 µl 80% EtOH to remove salt and PEG
- Air-dry beads 3-5 min to remove EtOH
- **Elute** beads in 22 µl Tris-HCl (pH 8.0-8.5)
- Collect on magnet and transfer 20 µl supernatant to a fresh tube

Lid Colour	Aliquot label	Component	Volume
		Beads with purified adapter-ligated DNA (step 5)	50 µl
Yellow	PEG	PEG/NaCl Solution	35 µl
		TOTAL:	85 µl



Step-by-step



Medical Research Council

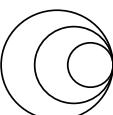


University  
of Glasgow  
*VIA VERITAS VITA*



CVR  
Centre for  
Virus Research

UK Health  
Security Agency



wellcome  
connecting  
science

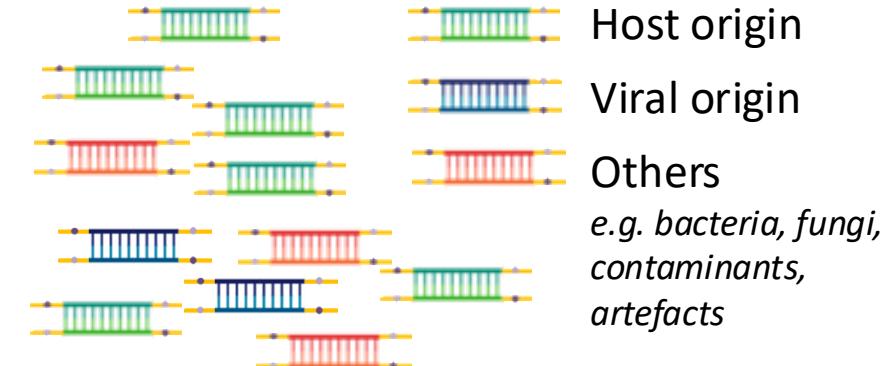
# Step 7: Library amplification

Adapter-ligated dsDNA fragments



Add:

- HiFi HotStart ReadyMix (enzyme included)
- Library amplification primer mix



Master mix:

Lid Colour	Aliquot label	Component	Master mix volume (10% excess)
Green	HiFi	KAPA HiFi HotStart Ready-mix (2X)	27.5 µl
Red	PM	Library Amplification Primer Mix (10X)	5.5 µl
		TOTAL:	33 µl

Reaction mix:

Component	Volume
Cleaned, adapter-ligated DNA (step 6)	20 µl
PCR master mix	10 µl
TOTAL:	20 µl

Incubate:

Step	Temp.	Duration	Cycles
Initial denaturation	98°C	45 sec	1
Denaturation	98°C	15 sec	16
Annealing*	60°C	30 sec	
Extension	72°C	30 sec	
Final extension	72°C	1 min	
HOLD	4°C	∞	

Step-by-step



Medical Research Council

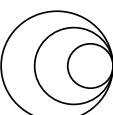


University  
of Glasgow  
*VIA VERITAS VITA*

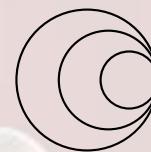


CVR  
Centre for  
Virus Research

UK Health  
Security Agency



wellcome  
connecting  
science

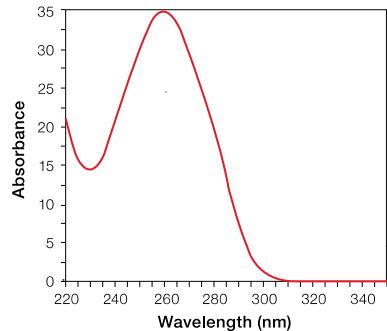


wellcome  
connecting  
science

# Quality control

# Quality control - quantification

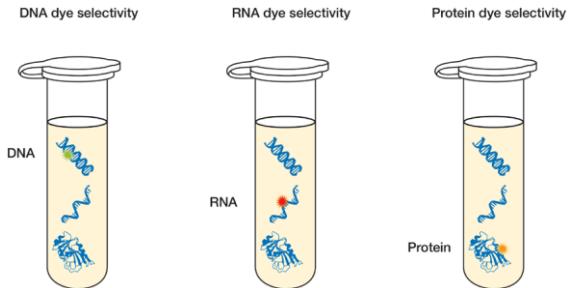
## Absorbance



## Nanodrop

- ✓ Simple - no preparation required
- ✓ Purity ratios - A260/A280, A260/A230
- ✓ Identify non-nucleic acid contaminants
- ✗ Lacks specificity
- ✗ Limited sensitivity

## Fluorescence



## Qubit

- ✓ Specific - DNA, dsDNA, ssDNA, RNA
- ✓ Sensitive - pg/ml
- ✓ Accurate - only quantifies target nucleic acid
- ✗ Requires preparation
- ✗ No purity information

# Quality control - visualisation

## Capillary electrophoresis

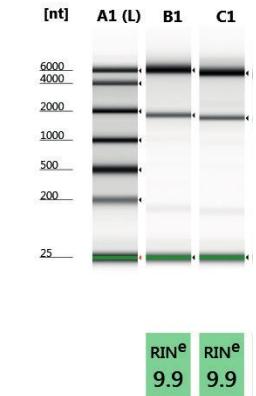
TapeStation, BioAnalyzer, Fragment Analyzer

- ✓ Specific - RNA, gDNA, hsDNA
- ✓ Sensitive - 100 pg/ $\mu$ l RNA, 5 pg/ $\mu$ l DNA

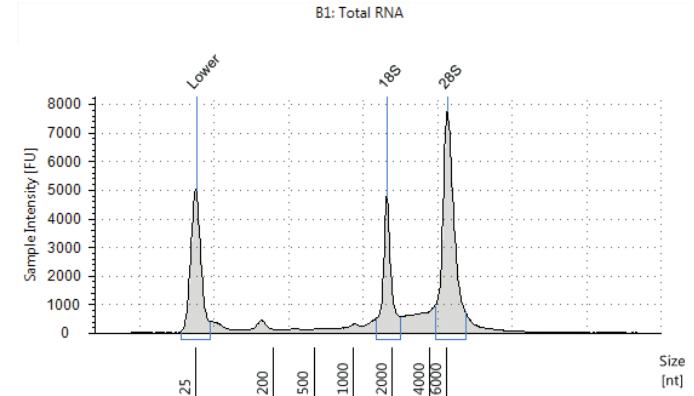
Evaluation of:

- Size - within expected range, accurate nM quantification
- Quality - primer dimers, PCR artefacts
- Integrity - RIN, DIN scores estimate degradation

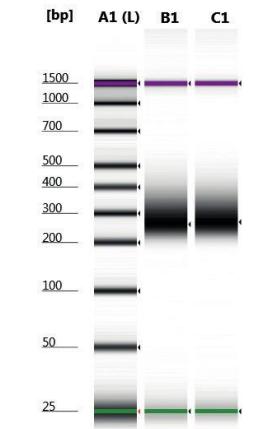
RNA ScreenTape assay



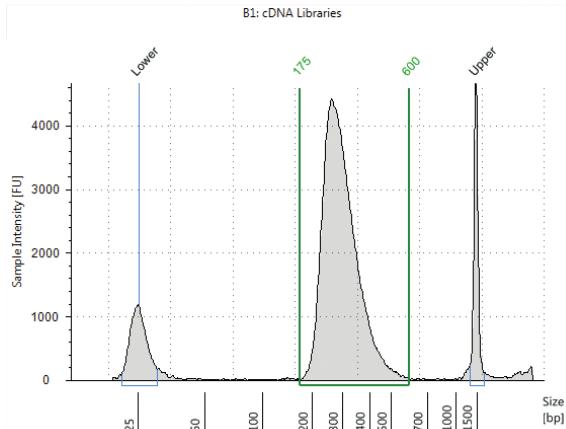
B1: Total RNA



D1000 ScreenTape assay



B1: cDNA Libraries



# Sample quality control

For optimal library preparation need to first check:

- Quantity - ensure optimal/equal amount of nucleic acid used for reactions
- Purity - chemical or protein contaminants, unwanted nucleic acids
- Integrity - fragmentation

Challenging samples can still be sequenced but may need different workflows or additional steps

- Low/undetectable input - more sensitive RT, increase conc adapter, increase library PCR cycles
- Unwanted nucleic acids - DNase treatment, remove host nucleic acid
- Degraded samples - minimise fragmentation steps during library prep

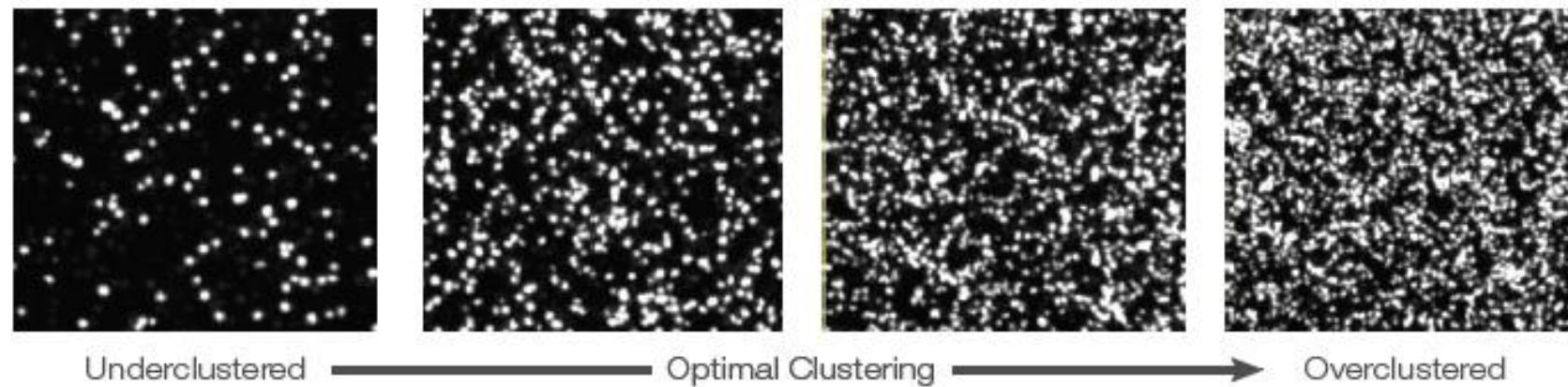
# Library quality control

Ensure good quality libraries for good quality sequencing

Accurate quantification - equal representation of each sample in the pool

- Qubit - gives concentration in ng/ $\mu$ l
- Tapestation shows the peak size in base pairs
- Combine to calculate Molar concentration
- qPCR to quantify only the amplifiable library molecules

# Pool quality control



Accurate quantification of pool crucial to run output

- Over-estimation of molarity => low cluster density, fewer reads, failure if too low
- Under-estimation of molarity => high cluster density, lower run quality, risk run failure

Dilute pool to 4 nM for illumina sequencing

- Final Qubit quantification - concentration low so small errors can make a large difference

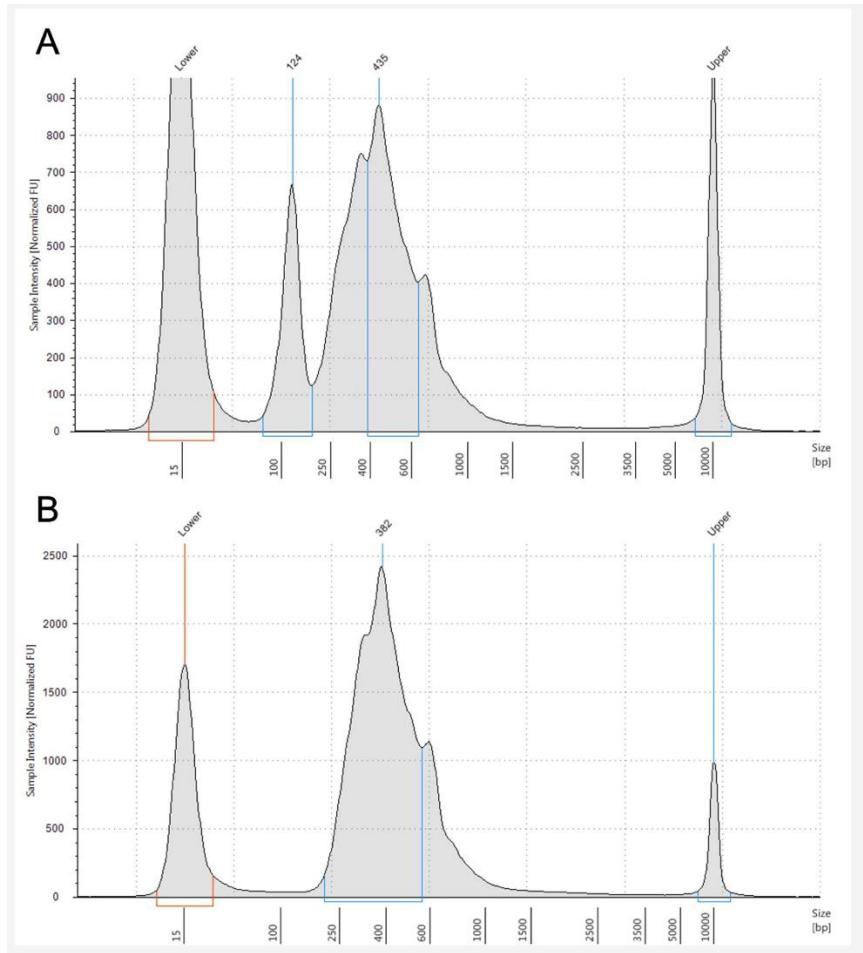
# Troubleshooting – adapter dimers

- Small fragments (typically ~120-170 bp)
- Contain full length adapter sequences - can be clustered
- Cluster more efficiently

Problems:

- Reduce the number of library specific reads
- Over clustering, reducing the data quality for all the samples
- Shorter length results in no signal in later run cycles reducing quality
- Free adapter can be incorporated into library clusters causing index hopping

**Removed by additional Ampure clean up steps**



Example of library with adapter dimer peak (A) and same library after additional round of Ampure XP clean up (B)

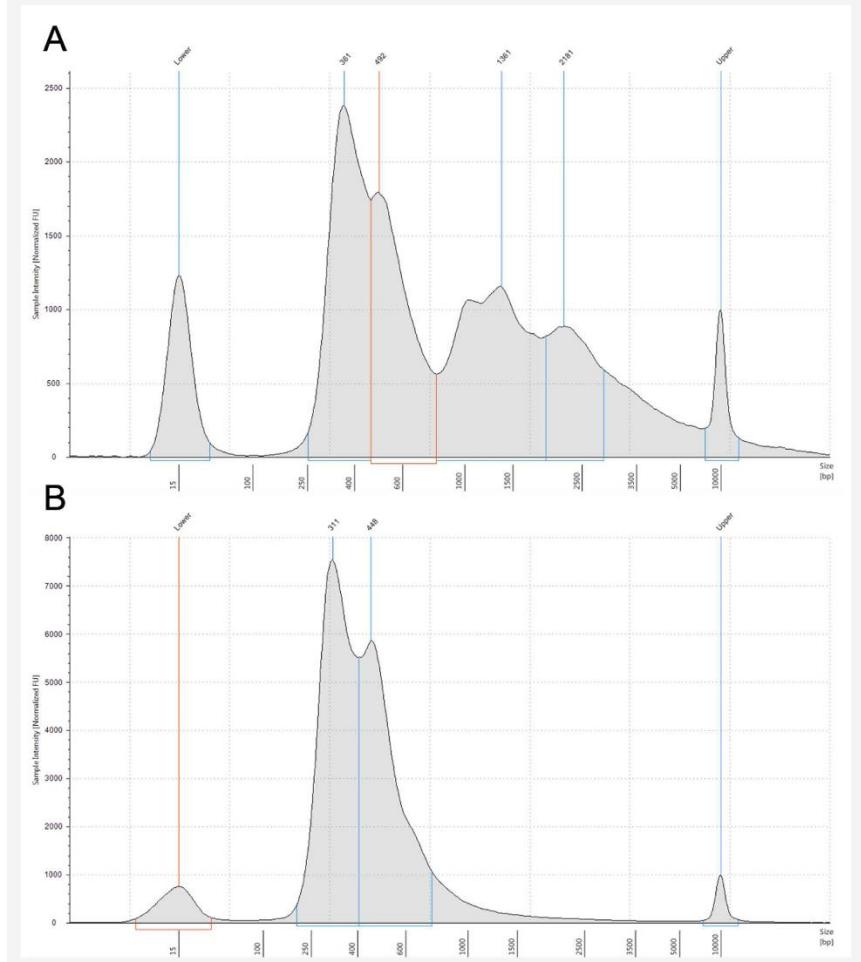
# Troubleshooting – daisy chains

Assemblies of improperly annealed, partially double-stranded, hetero-duplex DNA

Caused by:

- over amplification causing depletion of primers/dNTP's
- contaminants carry over into PCR reaction

Broken up by additional PCR cycles

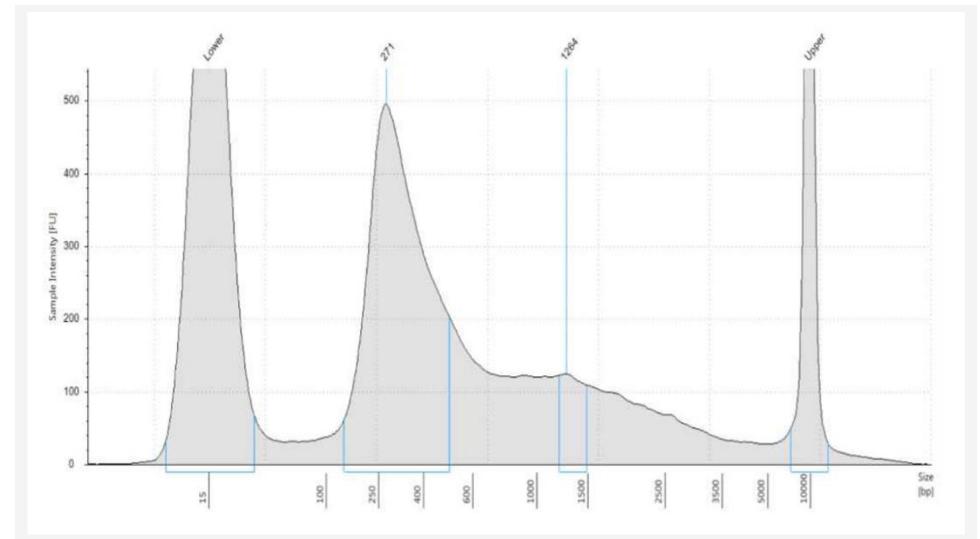


Example of library with daisy chain peaks (A) and same library after additional PCR cycles (B)

# Troubleshooting – large fragment contamination

- Larger fragments cluster less efficiently than the smaller, specific libraries.
- They do interfere with accurate quantification

Removed by reverse-ampure clean up.



Example of library with large fragment contamination

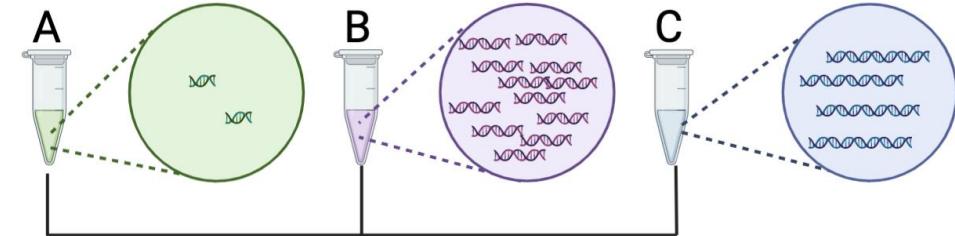
# Molarity calculations

# Library pooling

Accurate pooling important to ensure similar read depth

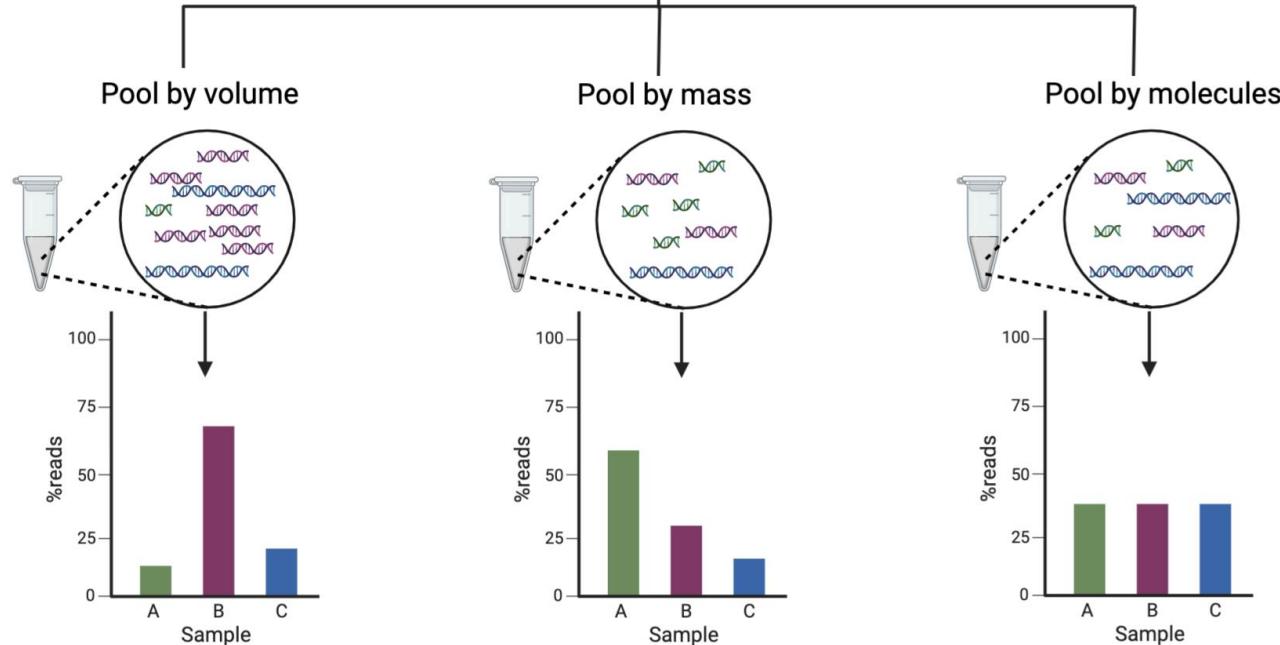
## Pool by volume - equal ul

- Libraries need similar concentration and profile
- Run on low output cartridge and rebalance before high output run



## Pool by mass - equal ng

- Libraries need the same profile



## Pool by molecules - equal nM

- Takes concentration and profile into account

# Calculating molarity of individual libraries

$$Molar\ conc.\ (nM) = \frac{sample\ (ng/\mu L)}{size\ (bp) * 660\ (g/mol)} * 1,000,000$$

To calculate the molarity need

- Concentration ng/μl (Qubit)
- Average fragment size bp (TapeStation)

*Example calculation:*

- Qubit concentration = 9.7 ng/μL
- TapeStation peak size = 382 bp
- **Molarity = ???? nM**

$$Molar\ conc.\ (nM) = \frac{9.7\ (ng/\mu L)}{382\ (bp) * 660\ (g/mol)} * 1,000,000$$

# Calculate library volume for equimolar pool

Determine pool requirements:

- Final Pool volume ( $\mu\text{L}$ )
- Final Pool conc (nM)
- Number of Libraries per pool

Determine the final concentration needed in the final pool:

$$\text{Equimolar library conc (nM)} = \frac{\text{Final pool conc (nM)}}{\text{number of libraries}}$$

Calculate the library volume needed:

$$\text{Volume } (\mu\text{L}) = \text{Equimolar library conc (nM)} * \frac{\text{Pool volume } (\mu\text{L})}{\text{Current conc (nM)}}$$

Make pool up to required volume with Tris:

$$\text{Tris Volume } (\mu\text{L}) = \text{Pool volume } (\mu\text{L}) - \text{sum(library Volumes } (\mu\text{L}))$$

# Calculate library volume for equimolar pool

*Example calculation:*

- Final Pool volume 50  $\mu\text{L}$
- Final Pool conc 25 nM
- 2 libraries per pool

Final concentration:

$$\text{Equimolar library conc (nM)} = \frac{\text{Final pool conc (nM)}}{\text{number of libraries}}$$

$$\text{Equimolar library conc (nM)} = \frac{25 \text{ (nM)}}{2} = ???????$$

Library volume:

$$\text{Volume (\mu L)} = \text{Equimolar library conc (nM)} * \frac{\text{Pool volume (\mu L)}}{\text{Current conc (nM)}}$$

$$\text{Volume (\mu L)} = 12.5 \text{ (nM)} * \frac{50 \text{ (\mu L)}}{38.5 \text{ (nM)}} = ???????$$

# Calculate library volume for equimolar pool

(A) Number of libraries	10	<i>adjust as needed</i>		
(B) Total Conc. of Pool (nM)	25			
(C) Total Volume of Pool ( $\mu$ L)	50			
(D) Final Conc. each library in Pool (nM)	2.5	= B/A		
SampleID	(E) Qubit (ng $\mu$ L)	(F) TapeStation (bp)	(G) Molarity (nM) = E/(F*660) *1,000,000	(H) Volume for pool ( $\mu$ L) = D*(C/G)
Library 1	21.8	300	110.10	1.14
Library 2	39	289	204.47	0.61
Library 3	28	270	157.13	0.80
Library 4	24	336	108.23	1.16
Library 5	25	335	113.07	1.11
Library 6	25.6	310	125.12	1.00
Library 7	25.6	289	134.21	0.93
Library 8	19.4	298	98.64	1.27
Library 9	17.3	277	94.63	1.32
Library 10	8.58	264	49.24	2.54
(I) SubTotal				11.86
Tris-HCl 10 mM = C-I				38.14
Total Volume				50.00

Talk – Molarity calculation



Medical  
Research  
Council



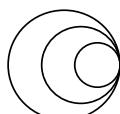
University  
of Glasgow  
*VIA VERITAS VITA*



CVR  
Centre for  
Virus Research

UK Health  
Security  
Agency

UCL



wellcome  
connecting  
science

# Confirm the pool molarity

$$Molar \ conc. (nM) = \frac{sample \ (ng/\mu L)}{size \ (bp) * 660 \ (g/mol)} * 1,000,000$$

Accurate quantification essential for accurate loading and good quality data!

Repeat the QC of the pool

- Concentration ng/ $\mu$ L (Qubit)
- Average fragment size bp (TapeStation)

*Example calculation:*

- Qubit concentration = 5.72 ng/ $\mu$ L
- TapeStation peak size = 320 bp
- Molarity = **?????** nM

$$Molar \ conc. (nM) = \frac{5.72 \ (ng/\mu L)}{320 \ (bp) * 660 \ (g/mol)} * 1,000,000$$

# Dilution to 4 nM

Immediately prior to sequencing need to dilute the pool to 4 nM

Accuracy essential – small errors > under/overloading and reduce data quantity and quality

$$Vol1 (\mu L) * conc1 (nM) = Vol2 (\mu L) * conc2 (nM)$$



$$Vol1 (\mu L) = \frac{Vol2 (\mu L) * conc2 (nM)}{conc1 (nM)}$$

## Example dilution to 4 nM:

- Molarity = 27.08 nM
- Final volume = 50 μL
- Final concentration = 4 nM
- **DNA volume = ????????**
- **Tris volume = ????????**

$$Vol1 (\mu L) = \frac{50 (\mu L) * 4 (nM)}{27.08 (nM)}$$

# Confirm the 4 nM pool molarity

$$Molar\ conc.\ (nM) = \frac{sample\ (ng/\mu L)}{size\ (bp) * 660\ (g/mol)} * 1,000,000$$

Accurate quantification essential for accurate loading and good quality data!

Repeat the QC of the pool

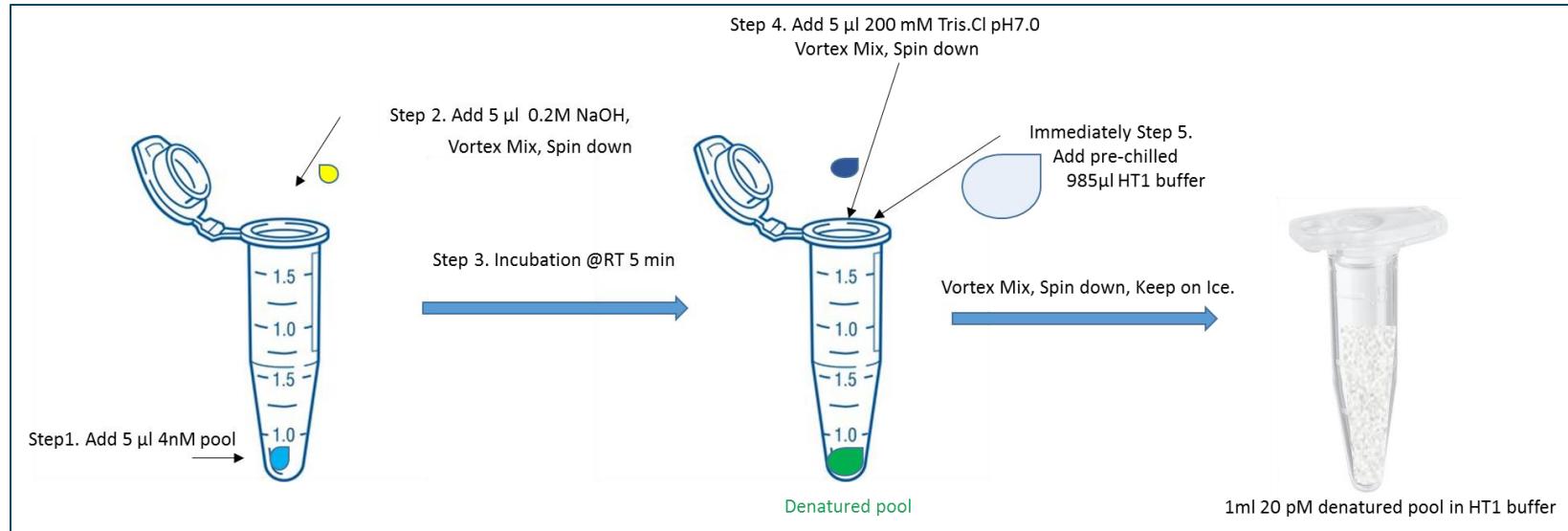
- Concentration ng/ $\mu$ L (Qubit)
- No need to repeat TapeStation – use fragment size of undiluted pool

*Example calculation:*

- Qubit concentration = 0.96 ng/ $\mu$ L
- TapeStation peak size = 320 bp
- **Molarity = ???? nM**

$$Molar\ conc.\ (nM) = \frac{0.96\ (ng/\mu L)}{320\ (bp) * 660\ (g/mol)} * 1,000,000$$

# Denaturation and loading



1. Add 5  $\mu$ L of 4 nM pool
2. Add 5  $\mu$ L of 0.2 M NaOH
3. Incubate 5 min
4. Neutralise with 5  $\mu$ L of 200 mM Tris pH 7 and make up to 1 mL with HT1 buffer – **20 pMol pool**
5. Dilute 20 pMol pool to desired loading concentration (we will do 10 pMol)

Talk – Molarity calculation



Medical  
Research  
Council



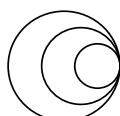
University  
of Glasgow  
*VIA VERITAS VITA*



CVR  
Centre for  
Virus Research

UK Health  
Security  
Agency

UCL



wellcome  
connecting  
science

# Altering dilution to ensure correct loading concentration

$$Vol1 (\mu L) * conc1 (nM) = Vol2 (\mu L) * conc2 (nM)$$

## Alter volume at step 1 pool denaturation

- Denaturation requires 5  $\mu L$  of 4 nM pool

$$Vol1 (\mu L) = \frac{5 (\mu L) * 4 (nM)}{4.54 (nM)} = 4.41 \mu L$$

- Need to also adjust volume of NaOH, Tris and HT1
- Results in 20 pMol pool
- Dilute to loading conc (10 pM in 1ml)

$$Vol1 (\mu L) = \frac{1000 (\mu L) * 10 (pM)}{20 (pM)} = 500 \mu L$$

$$HT1 (\mu L) = 1000 \mu L - 500 \mu L = 500 \mu L$$

## Alter volume at step 5 final loading conc

- Denature 5  $\mu L$  of 4.54 nM pool
- Add 5  $\mu L$  NaOH and Tris
- Make up to 1 ml with HT1 and calculate actual concentration

$$Conc1 (pM) = \frac{5 (\mu L) * 4.54 (nM)}{1000 (\mu L)} * 1000 = 22.7 \text{ pM}$$

- Dilute to loading conc (10 pM in 1 ml)

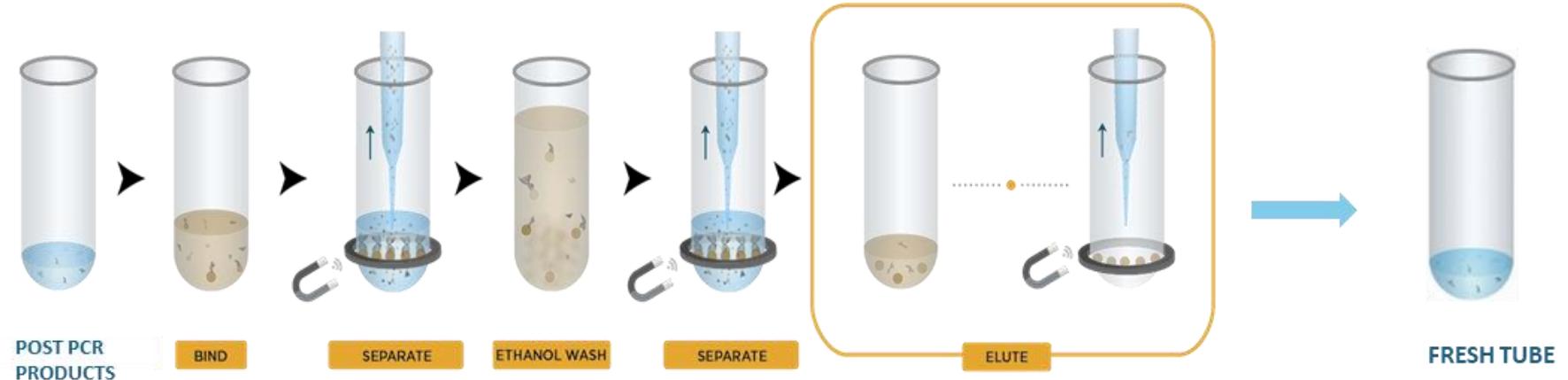
$$Vol1 (\mu L) = \frac{1000 (\mu L) * 10 (pM)}{22.7 (pM)} = 441 \mu L$$

$$HT1 (\mu L) = 1000 \mu L - 441 \mu L = 559 \mu L$$

# Step 8: Library amplification clean-up

- Add Ampure beads solution for 1X clean-up
  - Library amplification mix from Step 7 = 50 µl
  - 1 x 50 = 50 µl beads
- Incubate at Room temp. 5-10 min
- Separate beads on magnet and remove supernatant
- 2X washes in 200 µl 80% EtOH to remove salt and PEG
- Air-dry beads 3-5 min to remove EtOH
- Resuspend beads in 22 µl Tris-HCl (pH 8.0-8.5)
- Collect on magnet and transfer 20 µl supernatant to a fresh tube

Lid Colour	Aliquot label	Component	Volume
		Amplified library DNA (step 7)	50 µl
	AMP	Ampure XP beads	50 µl
		TOTAL:	100 µl



• *be careful not to carry over beads in eluate: bead carry-over will affect downstream QC!*

Step-by-step



Medical Research Council



University  
of Glasgow  
*VIA VERITAS VITA*

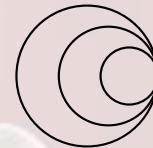


CVR  
Centre for  
Virus Research

UK Health  
Security Agency



wellcome  
connecting  
science

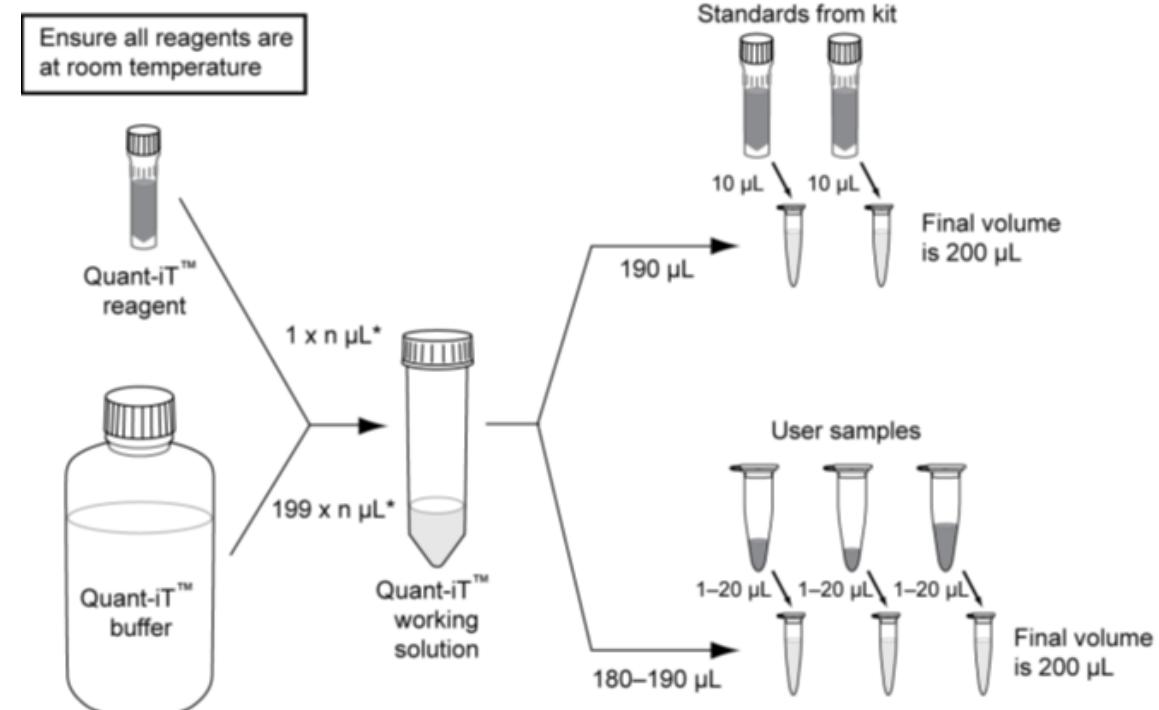


wellcome  
connecting  
science

# Coffee break

# Qubit quantification of prepared libraries

- We will prepare Quant-iT working solution and add 199 $\mu$ L to the sample tubes
- Label tubes (2 samples per group – we will prepare the standards)
  - NOTE – label lids only!
- Add 1  $\mu$ L sample to each tube
- Vortex and incubate ~2 min
- Read in Qubit
- Record concentration ng/ $\mu$ L

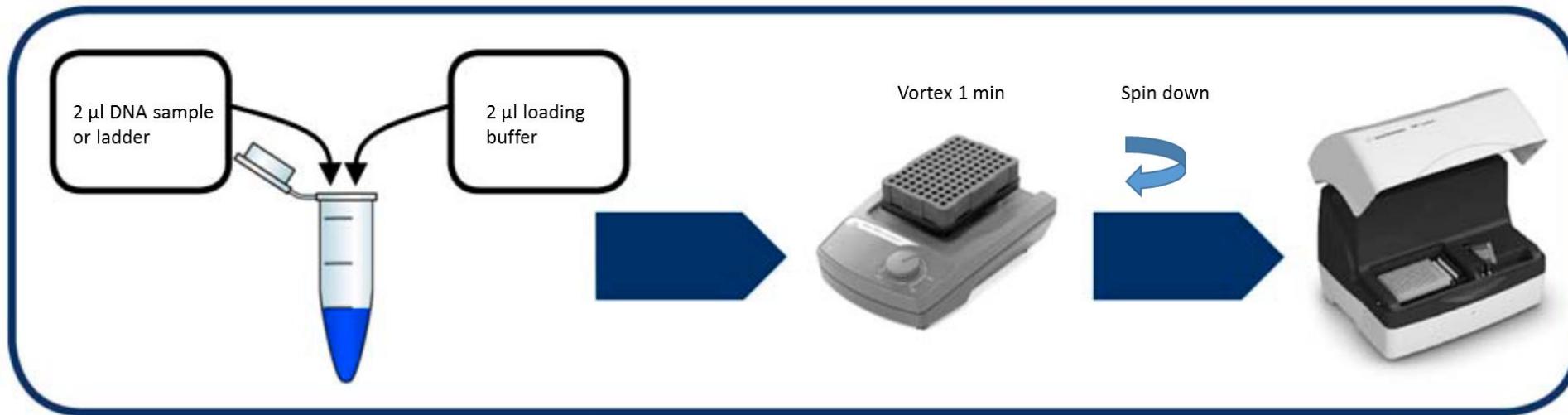


Vortex all assay tubes for 2–3 seconds

Incubate at room temperature for 2 minutes

Read tubes in the Qubit™ fluorometer

# TapeStation analysis of prepared libraries



- Make dilution <5 ng/µL (no more than 10 ng in 2 µL)
- Add 2 µL to preprepared tube containing loading dye
- We will run on TapeStation
- Analyse results
  - What is the library size (bp)?
  - What is the library quality?

# Calculations

1. Use Qubit and TapeStation results to calculate your library molarity

$$\text{Molar conc. (nM)} = \frac{\text{sample (ng}/\mu\text{L)}}{\text{size (bp)} * 660 (\text{g/mol})} * 1,000,000$$

2. Calculate the volume require for equimolar pooling

$$\text{Vol1 } (\mu\text{L}) = \frac{\text{Vol2 } (\mu\text{L}) * \text{conc2 } (\text{nM})}{\text{conc1 } (\text{nM})}$$

3. Give your results to the demonstrators

# Can we make libraries?

Sample	HCV Ct in the extracts	RNA Qubit quantification (ng/µl) of the extracts	DNA Qubit quantification (ng/µl) of the extracts	DNA Qubit quantification (ng/µl) of the libraries	TapeStation fragment size (bp) of the libraries	Library molarity (nM)
Sample 1a	21	Undetected	10	10	382	39.66
Sample 1b	23.1	Undetected	13.8	5	400	18.94
Sample 2a	22.5	Undetected	23			#DIV/0!
Sample 2b	27.2	Undetected	7.8			#DIV/0!
Sample 3a	25.4	Undetected	21.6			#DIV/0!
Sample 3b	20.96	Undetected	18.84			#DIV/0!
Sample 4a	25.82	Undetected	23.09			#DIV/0!
Sample 4b	26.04	Undetected	17.75			#DIV/0!
Sample 5a	31.54	Undetected	17.3			#DIV/0!
Sample 5b	22.11	Undetected	19.22			#DIV/0!
Sample 6a	32.79	Undetected	12.64			#DIV/0!
Sample 6b	25.62	Undetected	27.64			#DIV/0!
Sample 7a	24.13	Undetected	19.82			#DIV/0!
Sample 7b	26.48	Undetected	12.59			#DIV/0!
Sample 8a	27.49	Undetected	14.06			#DIV/0!
Sample 8b	22.24	Undetected	14.35			#DIV/0!
Sample 9a	25.47	Undetected	13.78			#DIV/0!
Sample 9b	29.9	Undetected	17.1			#DIV/0!

Fill in library Qubit and TapeStation readings to calculate molarity's

Step-by-step



Medical  
Research  
Council

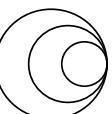


University  
of Glasgow  
VIA VERITAS VITA



CVR  
Centre for  
Virus Research

UK Health  
Security  
Agency



wellcome  
connecting  
science



Medical  
Research  
Council



University  
of Glasgow



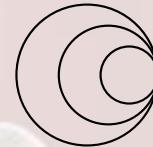
CVR  
Centre for  
Virus Research



UK Health  
Security  
Agency



UCL



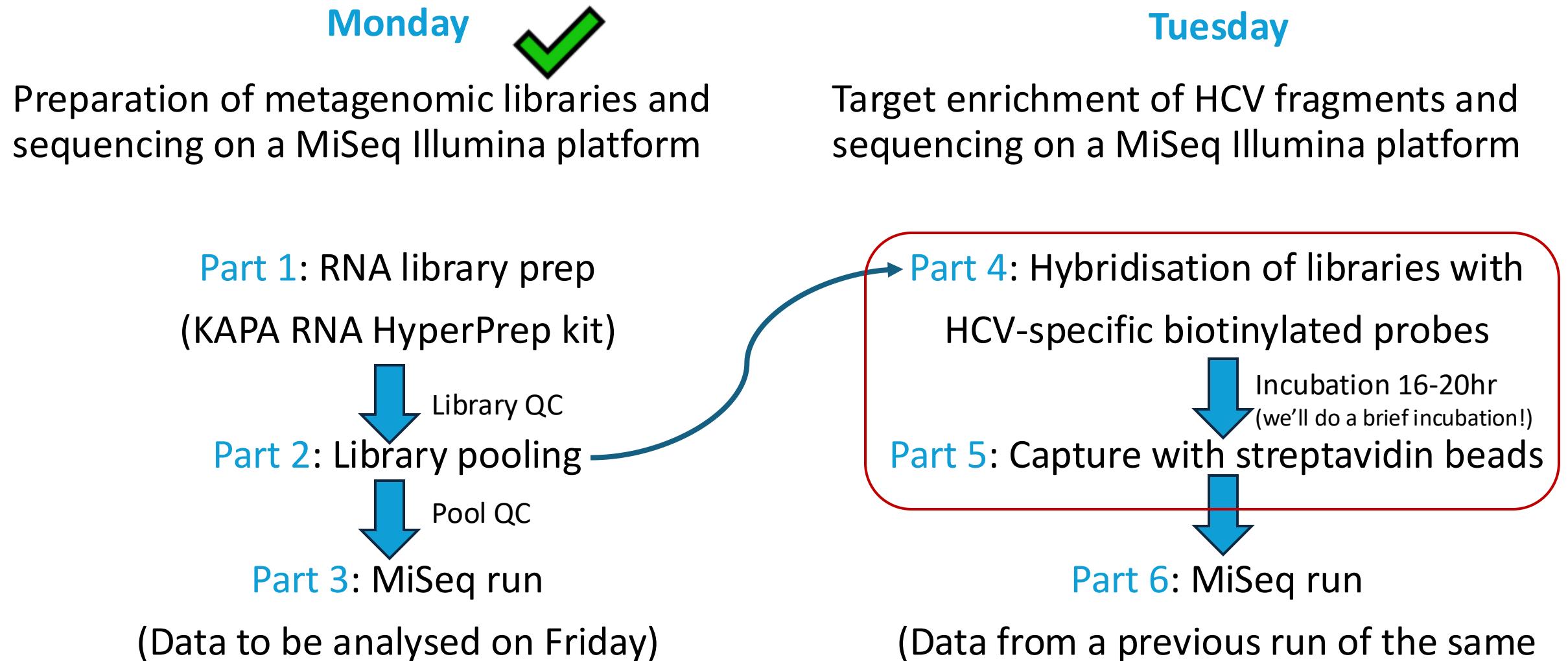
wellcome  
connecting  
science

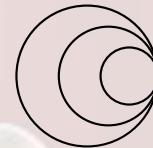
# Videos

# Illumina self-help videos

- Illumina MiSeq Sequencing reagents:
  - Cartridges <https://youtu.be/kHUsoczXv-xg?si=BbEmu6JEhE78N2jl>
  - Flow cells <https://youtu.be/XYU3kagGjI?si=zj7vHOJ3OZwzNdKe>
  - Denature and dilute libraries [https://youtu.be/1Ku9lodi\\_DQ?si=3mM3SNs8bXRxosRm](https://youtu.be/1Ku9lodi_DQ?si=3mM3SNs8bXRxosRm)
  - How to start a run [https://youtu.be/LX2U\\_bJRqfc?si=R9uX2Rra95XkJmaC](https://youtu.be/LX2U_bJRqfc?si=R9uX2Rra95XkJmaC)
- Using Illumina Experiment Manager (IEM)
  - <https://youtu.be/bfPpbJtj3LE?si=BRT4Lx0Yix81vbYa>

# Workflow overview





wellcome  
connecting  
science

# Dinner!