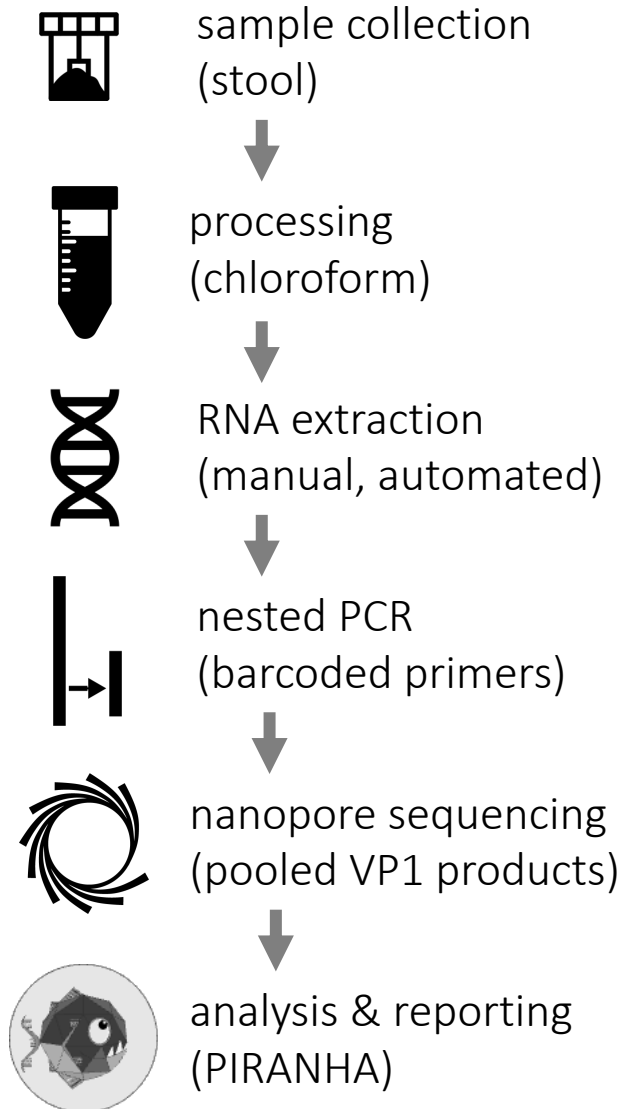


Overview of Direct Molecular Detection and Nanopore Sequencing (DDNS) of Poliovirus

Lecture and discussion

Poliovirus Direct Detection and Nanopore Sequencing (DDNS)



PoSeCo | Poliovirus Sequencing Consortium

Poliovirus Investigation Resource Automating Nanopore Haplotype Analysis

sample08 report 2022-09-12

Table 1 | Summary of sample content

Sample	Barcode	Reference Group
sample08	barcode33	Sabin1-related
sample08	barcode33	Sabin3-related

VP1 sequences

```
>sample08|barcode33|Sabin1-related|Poliovirus1-Sabin_AY184219|1|268:AT
GGGTTASGTGAGATGCTTGAAGAAGATGATTGACACACAGTCCGTGGAACGGTGGGGGGGACACGCTCTAGAGAGCGCTCCCAAAACAGTGAAGCAGTGGACAGCAGACTCCAAAGGAAATTCGGGCACTCACCAGTGGAACTGGGGCAC
AAATCCACTAGTCCCTTCTGATACAGTGAAGCCAGACATGTTGTACACATAGGTCAAGGTCAAGCTAGCATAGAGTCTTCTTCCGCGGGGGTGCATGCGTGGCATTTTAAACGTTGGATACTCACTCCACCAAGAAATAGGATAGAC
TATTTACAGTGTGGAAGATCACTATAAAGATACTGTCCAGTTACGGAGAAATGGAGTTCTTCCACTATTCTAGATTGATATGGAATTTACCTTTGTGGTTACTGCAAAATTTCACTGGACTAAACAATGGGATGCTTAAATCAAGTGTAC
CAAATATGTCAGTACACCAAGGGGCTCCAGTGCCTCCGAGAAATGGGAGACTACACATGGCAAACTCATCAAAATCCATCAATCTTTACACTACGGAAACAGCTCCAGCCGGATCTGGTACCGATGTTGGTATTTCGAAACGCTATTACAC
CTTTTACAGCGGTTTTTCCAAAGTACACTGAAGGACAGTCCGGAGCAGCTAGGTGACTCCCTCTATGGTGCAGCATCTCTAAATGACTTGGTATTTGGCTGTAGAGTAGTCAATGATCACAAACCGAACCAAGTCACTCCAAATCAGAG
```

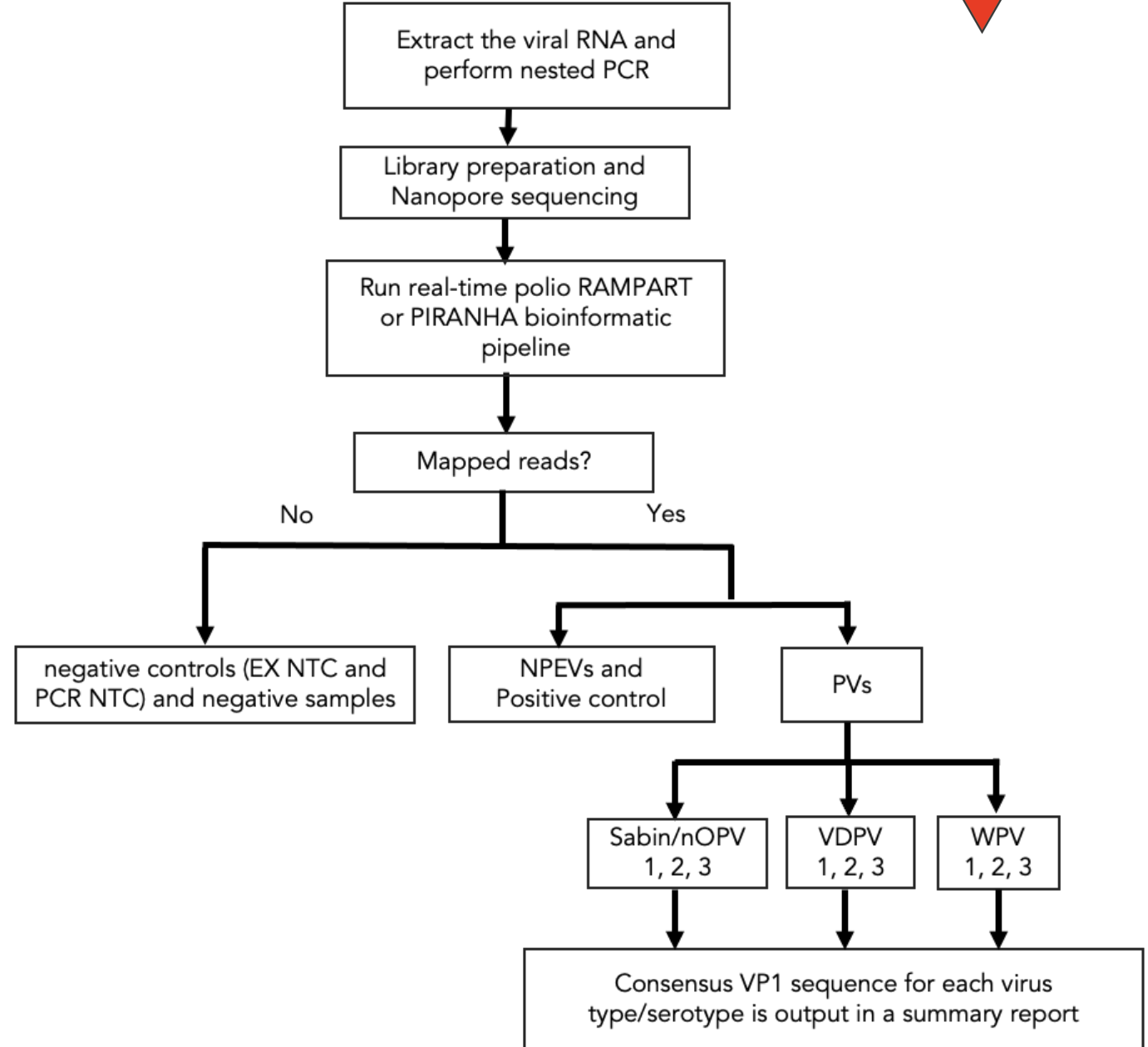
Download

Advantages of DDNS over cell-culture-ITD-Sanger sequencing

- Protocol detects and generates a VP1 sequence for poliovirus in stool samples 2-3 days after arrival in the laboratory
- Non-inferior sensitivity and specificity compared with cell-culture
- Accurate consensus sequences for every poliovirus detected
- Faster detection of outbreaks
- Cost-effective at ~\$15 per sample assuming >90 samples/week
- Ongoing training in GPLN labs to implement method and generate further data on performance and costs -> accepted/recommended by GPLN SWG



Diagnostic algorithm



Example Piranha report

Piranha report 2022-12-08

a.

Sample	Barcode	Sample Classification	Reference group	Number of mutations
ENV001	barcode01	Sabin-like	Sabin3-related	1
ENV001	barcode01	Sabin-like	Sabin2-related	2
ENV002	barcode02	VDPV	Sabin2-related	12
ENV003	barcode03	WPV1	WPV1	NA
ENV004	barcode04	VDPV	Sabin2-related	10

b.

Sample	Barcode	Sabin1-related	Sabin2-related	Sabin3-related	WPV1	WPV2	WPV3	NonPolioEV	unmapped
ENV001	barcode01	0	488	252	0	0	0	0	25
ENV002	barcode02	0	1100	0	0	1	0	35	12
ENV003	barcode03	0	0	0	339	0	0	0	0
ENV004	barcode04	0	0	0	138	0	0	0	0
negative	barcode05	0	0	0	0	0	0	1	10

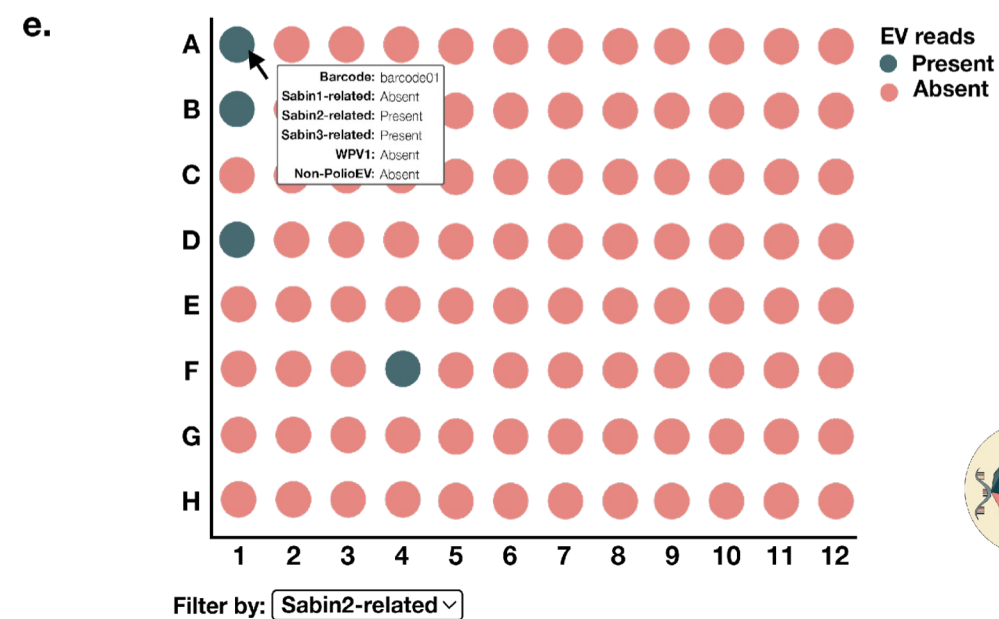
Showing 5 of 96 entries

c.

Identical Sequences	Sequence IDs
1	ENV001 barcode01 Sabin2-related 2 161:CT;427:GA ENV030 barcode30 Sabin2-related 2 161:CT;427:GA
2	ENV022 barcode22 Sabin3-related 1 17:CT ENV024 barcode24 Sabin3-related 1 17:CT

d.

Pass	Sample	Barcode	Sabin1-related	Sabin2-related	Sabin3-related	WPV1	WPV2	WPV3	NonPolioEV	unmapped
✓	negative	barcode05	0	0	0	0	0	0	1	10
	positive	barcode05	0	0	0	0	0	0	45	0



<http://polionanopore.org>

- Protocols maintained on protocols.io – can be reached through www.polionanopore.org

- Forums on the protocols.io site to raise any queries you have about the protocol

- New version of the protocols will be published here.



Free, open source protocols and software for poliovirus detection and sequencing using nanopore.

You can find all you need here for direct detection and nanopore sequencing (DDNS) of poliovirus from stool and environmental samples. Helping you detect and respond to poliovirus outbreaks faster. Funded by the Bill and Melinda Gates Foundation.



QA/QC

- Careful QA/QC essential (clean PCR, data management)
- QC guidelines available and QA programme under development
- SOPs include recommended best practices for preventing and detecting cross-contamination, minimising sample handling
- Positive control (lyophilized CV-A20 standard prepared by NIBSC) and negative (water) control included in each sample batch from RNA extraction step
- Sequencing accuracy can be confirmed by testing reference panel and repeats/comparison with alternative sequencing platforms in a subset
- Long term monitoring of lab performance (including NPEV detection) and proficiency testing
- Accreditation of GPLN laboratories for DDNS will follow a similar process as for virus isolation, ITD, sequencing and environmental surveillance based on minimum requirements checklist

How this training is organised

- Scheduled presentations and guided lab work
- An overview of the method
- Guided use of the protocol by trainee lab teams
- Independent performance of the protocol by the lab teams with support
- Practice the use of Oxford nanopore MinKNOW sequencing software
- Practice the use of the graphical user interface for Piranha ('piranhaGUI')

Expected outcomes

- Be able to perform the DDNS protocol including RNA extraction, nested PCR, barcode file preparation, nanopore sequencing, report generation by PIRANHA
- Understand the QC/QA process
- Be ready to implement in your own laboratories with GPLN and PSC support
- Appreciate the other potential applications of poliovirus nanopore sequencing including generation of whole genomes

Post training expectations

- Trainees should identify any additional equipment their lab would require to perform DDNS
- Trainees should identify the areas in their lab where the steps of the DDNS protocol could be performed
- Trainers should explain the DDNS method to their other lab members
- Trainees should liaise with the WHO for procurement of required equipment and reagents for DDNS implementation in their lab
- Trainees should provide a point of contact for the PSC to provide additional reagents (e.g. the positive control virus)
- Further training should be arranged to support implementation of DDNS in their own laboratory

Acknowledgments

Study team

Imperial College London: Alex Shaw, Catherine Troman, Isobel Blake, Joyce Akello, Laura Cooper, Shannon Fitz, Nicholas Grassly

NIBSC: Manasi Majumdar, Erika Bujaki, Javier Martin

NIH Islamabad: Adnan Khurshid, Yasir Arshad, Masroor Alam

University of Edinburgh: Áine O'Toole, Rachel Colquhoun, Corey Ansley, Zoe Vallance, Andrew Rambaut

INRB DRC: Tresor Kabeya, Emmanuel Lokilo, Yogolelo Riziki, Amuri Aziza, Eddy Lusamaki, Jean-Claude Makangara, Marceline Alonga, Yvonne Lay, Bibiche Nsunda, Elisabeth Pukuta, Steve Ahuka, Jean-Jacques Muyembe, Placide Mbala

Biosurv International: Catherine Pratt

Collaboration and funding

BMGF: Ananda Bandyopadhyay, Kathleen Rankin and colleagues

WHO HQ: Ousmane Diop, GPLN SWG, Surveillance Group

WHO regional offices: Salmaan Sharif (EMRO), Anfumbom Kitu Womeyi Kfutwah (Jude) (AFRO), Julius Chia (AFRO), Eugene Saxentoff (EURO)

□

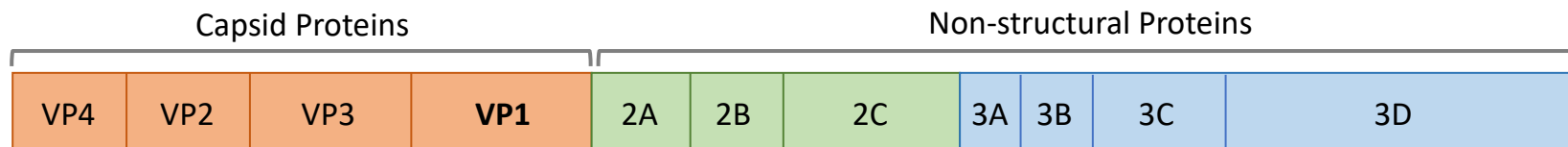
2. Sequencing Run preparation, RNA extraction and RT-PCR, nested PCR, Library preparation and flow cell loading

Lecture and discussion

Planning a run

- The protocol requires 2 days once RNA is extracted
- Can process up to 96 samples (including negatives) by barcoding samples
- Negative and positive run controls should be included in every run
- Samples are pooled to allow simultaneous sequencing
- Samples currently only tested once, but contamination likely to be identified through identical sequences for non-vaccine strains.

Our PCR strategy



**Pan-Enterovirus RT-PCR (primers for A+C enteroviruses, inc. nOPV2 : 5'NTR
+ Cre/nOPV-MM-R**
- SuperScript III One-Step RT-PCR System



VP1 PCR Nest

- DreamTaq PCR Master Mix
- Barcoded primers

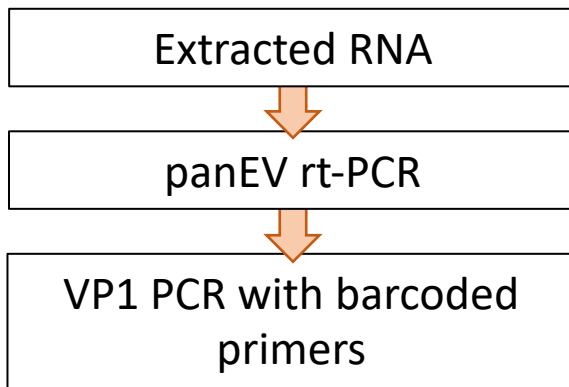


Barcode with flanking
sequences

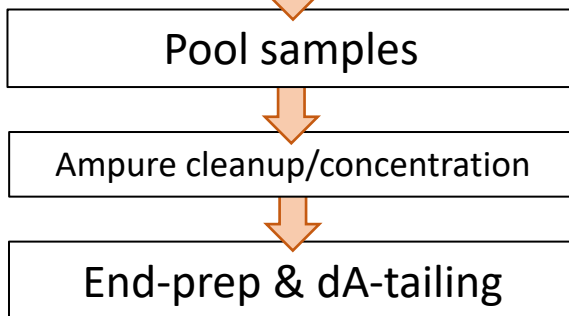


Library preparation overview

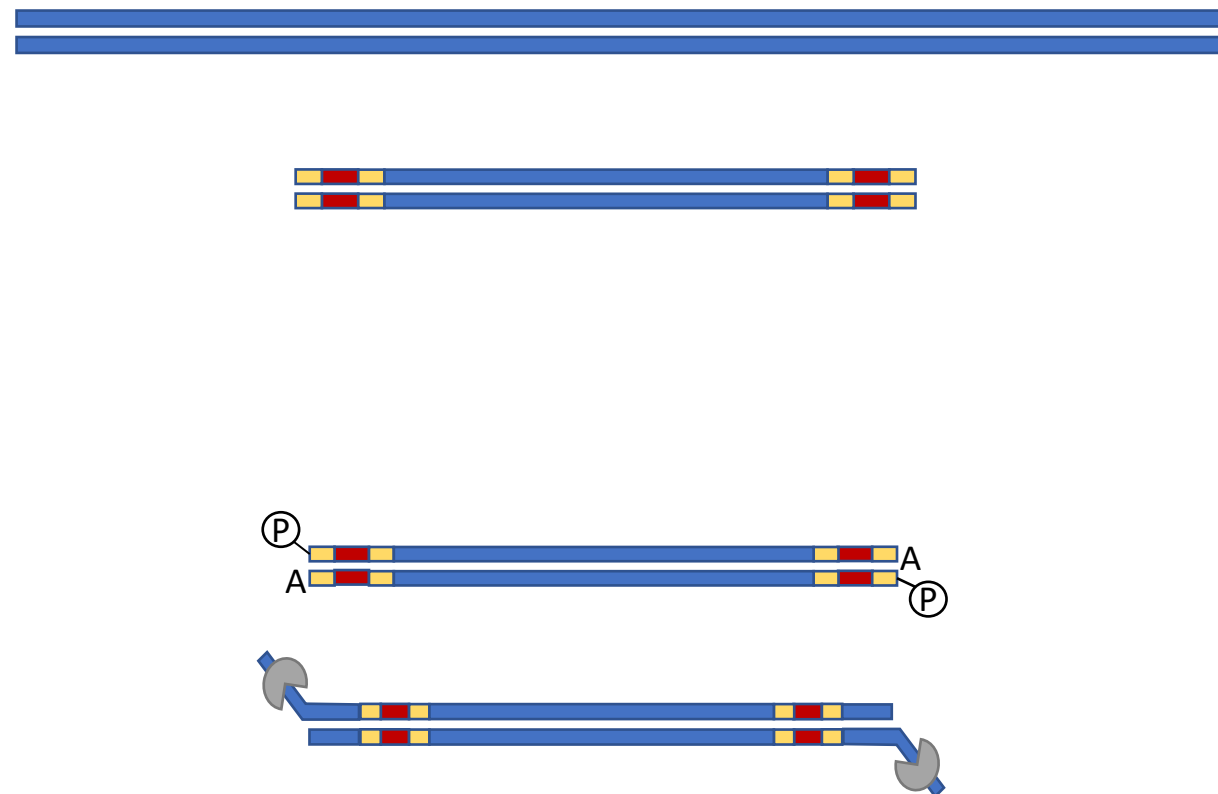
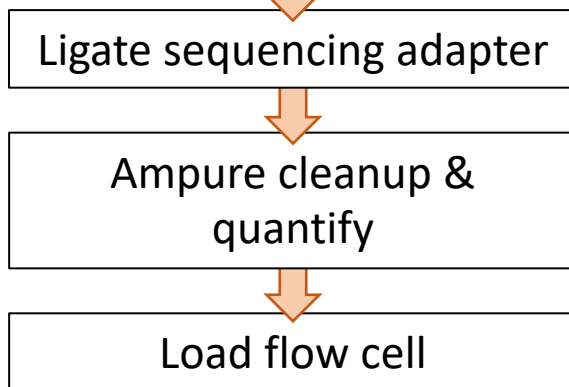
Part 1:
RNA
extraction,
RT-PCR and
PCR



Part 2:
Pooling,
cleaning and
end-prep



Part 3:
Preparation
for
sequencing



Run preparation

- All sample metadata can be entered into a csv file and each sample assigned a barcode.
- The analysis software (PIRANHA) will append the sequencing results to each sample.

	A	B	C	D	E
1	barcode	sample	EPID	institute	...
2	barcode01	sample01	ARA-HIG-TOR-22-01		
3	barcode02	sample02	ARA-HIG-TOR-22-01		
4	barcode03	sample03	ARA-KAN-ERF-22-01		
5	barcode04	sample04	ARA-KAN-ERF-22-01		
6	barcode05	sample05	ARA-HIG-TOR-22-02		

The barcoded primers are stored in the format shown below, each number being a unique barcoded sequence attached to both Q8 and Y7

	1	2	3	4	5	6	7	8	9	10	11	12
A	1	9	17	25	33	41	49	57	65	73	81	89
B	2	10	18	26	34	42	50	58	66	74	82	90
C	3	11	19	27	35	43	51	59	67	75	83	91
D	4	12	20	28	36	44	52	60	68	76	84	92
E	5	13	21	29	37	45	53	61	69	77	85	93
F	6	14	22	30	38	46	54	62	70	78	86	94
G	7	15	23	31	39	47	55	63	71	79	87	95
H	8	16	24	32	40	48	56	64	72	80	88	96

Run controls

- Both negative and positive controls go through the whole process from RNA extraction to sequencing
- Negative run control: Nuclease free water used for RNA extraction
- Positive run control: CVA20 formulated by NIBSC, distributed freeze-dried and known to amplify in both PCR steps
- Instructions for use of the positive control are provided with the vials
- The controls demonstrate if the required sensitivity and specificity of the method is achieved
- They confirm negative or positive signals and can indicate if any issues arise during processing steps

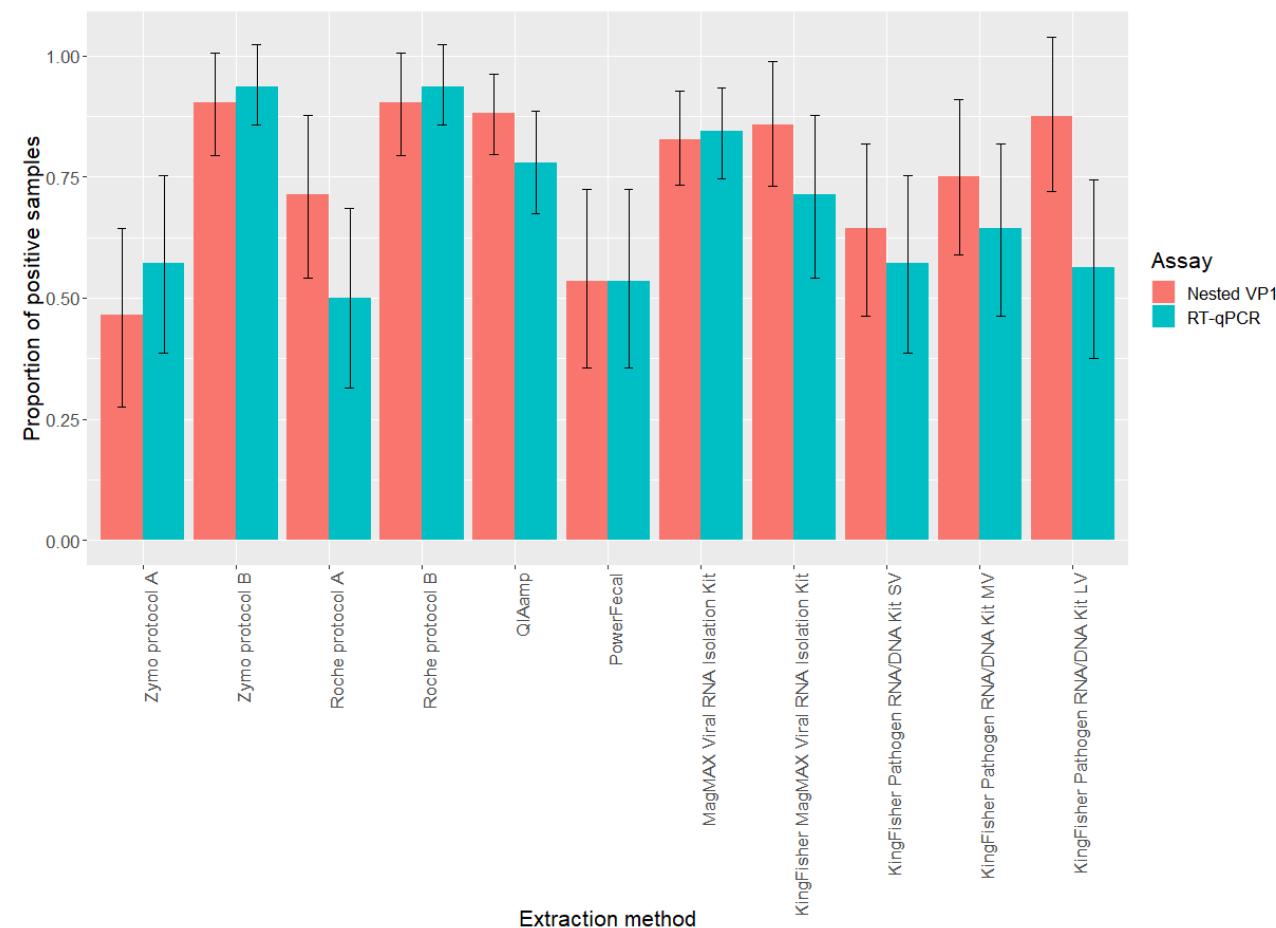
Performance of different RNA extraction kits

Approach

- 11 different RNA extraction methods have been compared
- Performance for nested VP1 PCR, pan-poliovirus qPCR (part of ITD) and near full-genome PCR assays examined
- We also considered processing time, flexibility, ease of use and cost

Conclusions

- There is significant variability between RNA extraction methods in their assay performance
- Optimal kits were:
 - manual: Roche HighPure, MagMAX Viral RNA
 - automated: Kingfisher MagMAX Viral RNA



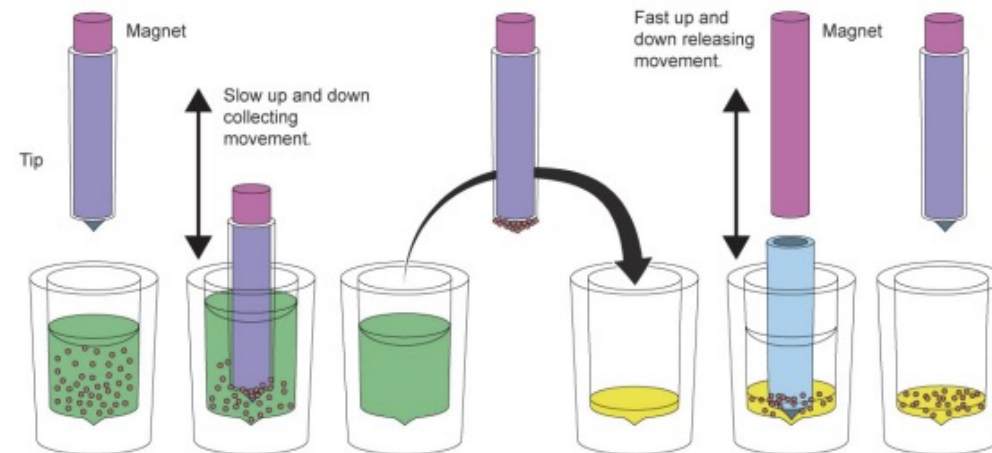
full details: Akello, Bujaki et al. 2023 *Microbiol.Spectr.* e0425222

RNA extraction

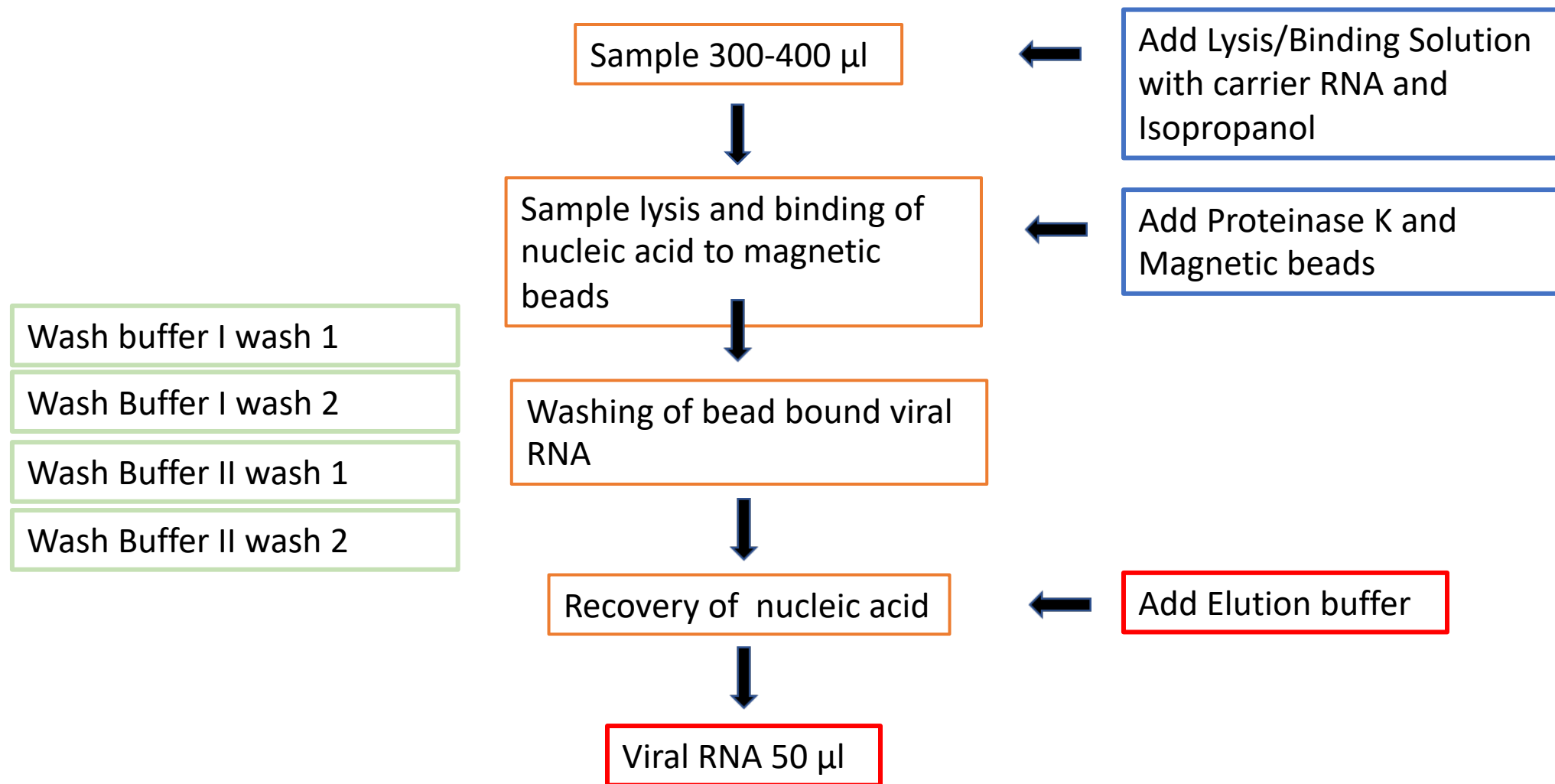
- Recommended kit: MagMAX Viral RNA Isolation kit Cat no: AM1939
Individual kit components are also available in large format.
- Magnetic bead-based method
- Can be performed manually in microcentrifuge tubes or in deep-well plates using KingFisher equipment



or



RNA extraction - MagMAX Viral RNA Isolation Kit



Setting up the panEV RT-PCR

Mastermix plate
(19 μ l reaction mix per well)

	1	2	3	4	5	6	7	8	9	10	11	12
A	1	9	17	25	33	41	49	57	65	73	81	89
B	2	10	18	26	34	42	50	58	66	74	82	90
C	3	11	19	27	35	43	51	59	67	75	83	91
D	4	12	20	28	36	44	52	60	68	76	84	92
E	5	13	21	29	37	45	53	61	69	77	85	93
F	6	14	22	30	38	46	54	62	70	78	86	94
G	7	15	23	31	39	47	55	63	71	79	87	95
H	8	16	24	32	40	48	56	64	72	80	88	96

← Add 5 μ l

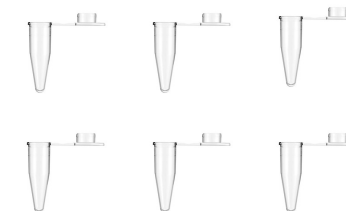
Extracted RNA

Automated

	1	2	3	4	5	6	7	8	9	10	11	12
A	1	9	17	25	33	41	49	57	65	73	81	89
B	2	10	18	26	34	42	50	58	66	74	82	90
C	3	11	19	27	35	43	51	59	67	75	83	91
D	4	12	20	28	36	44	52	60	68	76	84	92
E	5	13	21	29	37	45	53	61	69	77	85	93
F	6	14	22	30	38	46	54	62	70	78	86	94
G	7	15	23	31	39	47	55	63	71	79	87	95
H	8	16	24	32	40	48	56	64	72	80	88	96

Manual

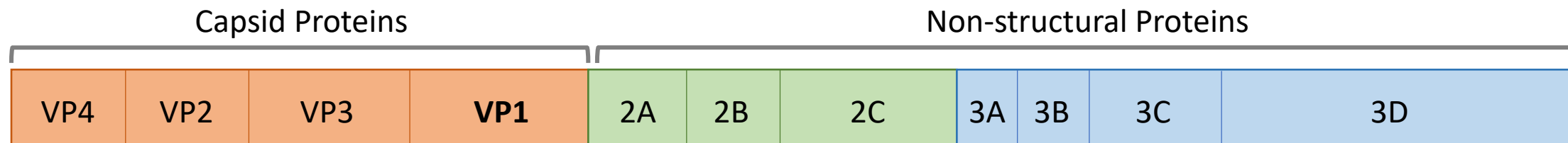
or



	1 Reaction (μ l)
2x Master Mix	12.5
SSIII Platinum Taq mix	1
Reverse Primer (10 mM, Cre/nOPV2-MM-R mix)	1
Nuclease free Water	4.5
Total	19

- 30 minutes of incubation at 50 °C (reverse transcription)
- Add 1 μ l of 10 μ M forward primer (5'NTR) for the PCR

panEV PCR



Pan-Enterovirus RT-PCR – Cre/nOPV-MM-R (rev) and 5’NTR (for)

- SuperScript III One-Step RT-PCR System

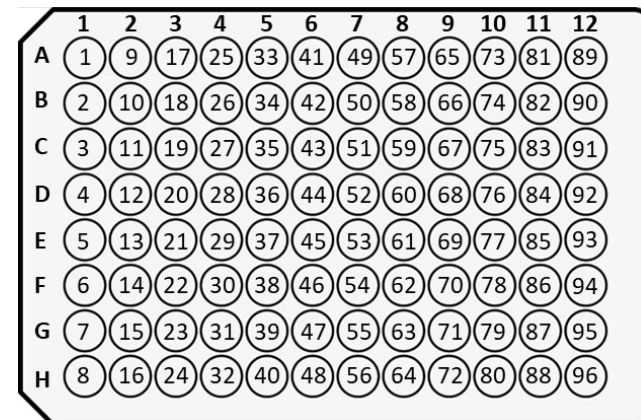


- Can run a gel to confirm, expect a band around 4kb.
- Only indicates the presence of an enterovirus.
- We do not run this gel routinely
- We do not exclude samples at this stage

	1	2	3	4	5	6	7	8	9	10	11	12
A	1	9	+	25	+	41	49	57	65	73	81	89
B	2	10	18	26	34	42	50	+	66	74	82	90
C	+	11	19	27	+	43	51	59	67	75	83	91
D	4	12	20	28	36	44	52	60	68	76	84	+
E	5	13	21	29	37	45	53	+	69	77	85	93
F	6	+	22	30	38	46	54	62	70	78	86	94
G	7	15	23	31	39	47	55	63	71	79	87	95
H	8	16	24	+	40	48	56	64	72	80	88	96

Nested VP1 PCR

- Performed with Dreamtaq
- Uses 2 μ l of each of the panEV reaction (no cleaning required)
- Primers Q8 and Y7 for poliovirus specific amplification of VP1
- Primers have a barcode to allow multiplexing of samples; each sample is assigned a barcode, which is the same in each of the primers.
- Barcoded Q8 and Y7 primers can be organised in a 96 well plate for ease of addition.
- Duration: 2 hrs for PCR



	1	2	3	4	5	6	7	8	9	10	11	12
A	1	9	17	25	33	41	49	57	65	73	81	89
B	2	10	18	26	34	42	50	58	66	74	82	90
C	3	11	19	27	35	43	51	59	67	75	83	91
D	4	12	20	28	36	44	52	60	68	76	84	92
E	5	13	21	29	37	45	53	61	69	77	85	93
F	6	14	22	30	38	46	54	62	70	78	86	94
G	7	15	23	31	39	47	55	63	71	79	87	95
H	8	16	24	32	40	48	56	64	72	80	88	96

VP1 PCR primers

Flanking sequence

Y7 **GGTGCTGACCGAGATCCTACGAATGGAGTGTTAACCTGGGTTTGTGTCAGCCTGTAATGA**

Barcode sequence

Primer sequence

TACACCTTRTCTCTGGAGAATCCAATTACCGAGATCCTACGAATGGAGTGTCGTGG Q8

Primer sequence

Barcode sequence

	1	2	3	4	5	6	7	8	9	10	11	12
A	1	9	17	25	33	41	49	57	65	73	81	89
B	2	10	18	26	34	42	50	58	66	74	82	90
C	3	11	19	27	35	43	51	59	67	75	83	91
D	4	12	20	28	36	44	52	60	68	76	84	92
E	5	13	21	29	37	45	53	61	69	77	85	93
F	6	14	22	30	38	46	54	62	70	78	86	94
G	7	15	23	31	39	47	55	63	71	79	87	95
H	8	16	24	32	40	48	56	64	72	80	88	96

Setting up the nested VP1 PCR

Mastermix plate
(21 μ l reaction mix per well)

	1	2	3	4	5	6	7	8	9	10	11	12
A	1	9	17	25	33	41	49	57	65	73	81	89
B	2	10	18	26	34	42	50	58	66	74	82	90
C	3	11	19	27	35	43	51	59	67	75	83	91
D	4	12	20	28	36	44	52	60	68	76	84	92
E	5	13	21	29	37	45	53	61	69	77	85	93
F	6	14	22	30	38	46	54	62	70	78	86	94
G	7	15	23	31	39	47	55	63	71	79	87	95
H	8	16	24	32	40	48	56	64	72	80	88	96

	1 Reaction (μ L)
Water	8.5
DreamTaq 2x master mix	12.5
Total volume	21

Add 2 μ l



Q8/Y7 barcoded primer plate (10 μ M)

	1	2	3	4	5	6	7	8	9	10	11	12
A	1	9	17	25	33	41	49	57	65	73	81	89
B	2	10	18	26	34	42	50	58	66	74	82	90
C	3	11	19	27	35	43	51	59	67	75	83	91
D	4	12	20	28	36	44	52	60	68	76	84	92
E	5	13	21	29	37	45	53	61	69	77	85	93
F	6	14	22	30	38	46	54	62	70	78	86	94
G	7	15	23	31	39	47	55	63	71	79	87	95
H	8	16	24	32	40	48	56	64	72	80	88	96

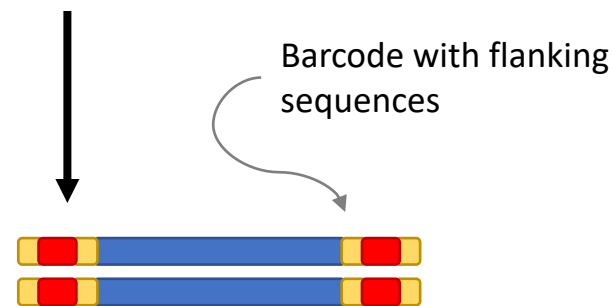
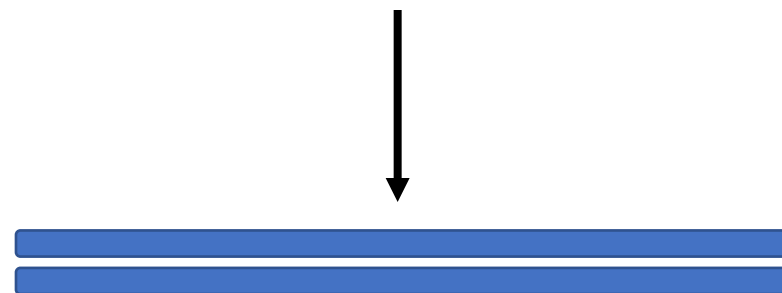
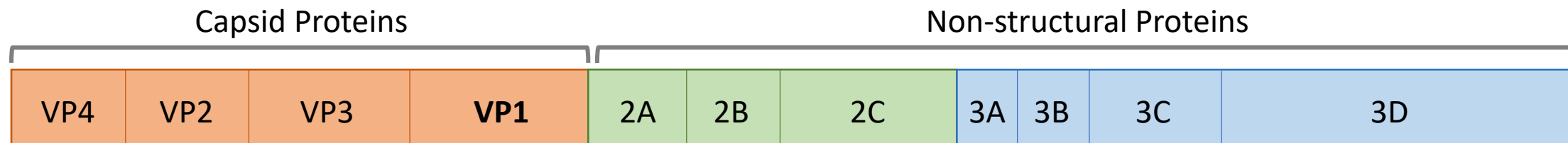
Add 2 μ l



panEV reaction plate

	1	2	3	4	5	6	7	8	9	10	11	12
A	1	9	17	25	33	41	49	57	65	73	81	89
B	2	10	18	26	34	42	50	58	66	74	82	90
C	3	11	19	27	35	43	51	59	67	75	83	91
D	4	12	20	28	36	44	52	60	68	76	84	92
E	5	13	21	29	37	45	53	61	69	77	85	93
F	6	14	22	30	38	46	54	62	70	78	86	94
G	7	15	23	31	39	47	55	63	71	79	87	95
H	8	16	24	32	40	48	56	64	72	80	88	96

Post PCR



	1	2	3	4	5	6	7	8	9	10	11	12
A	1	9	+	25	+	41	49	57	65	73	81	89
B	2	10	18	26	34	42	50	+	66	74	82	90
C	+	11	19	27	+	43	51	59	67	75	83	91
D	4	12	20	28	36	44	52	60	68	76	84	+
E	5	13	21	29	37	45	53	+	69	77	85	93
F	6	+	22	30	38	46	54	62	70	78	86	94
G	7	15	23	31	39	47	55	63	71	79	87	95
H	8	16	24	+	40	48	56	64	72	80	88	96

- Can run a gel to confirm, expect a band around 1.2kb.
- Only indicates the presence of poliovirus or enterovirus with similar primer binding sites
- We typically run a selection of samples to confirm PCR success

Library Pooling

- Barcoded samples equally pooled by volume (2 μ l each)
- Previously pooled equal ng of product, we have found for routine stool testing that it is easier to pool by volume without losing sensitivity

	1	2	3	4	5	6	7	8	9	10	11	12
A	1	9	+	25	33	41	49	57	65	73	81	89
B	2	10	18	26	34	42	50	58	66	74	82	90
C	3	11	19	27	+	43	51	59	67	75	83	91
D	4	12	20	28	36	44	52	60	68	76	84	+
E	5	13	21	29	37	45	53	+	69	77	85	93
F	6	+	22	30	38	46	54	62	70	78	86	94
G	7	15	23	31	39	47	55	63	71	79	87	95
H	8	16	24	32	40	48	56	64	72	80	88	96



Why pool everything?

- Equal volume pooling removes quantification and dilution steps
 - Reduces cost
 - Reduces scope for error
 - Quicker
- Samples tend to have a fairly narrow range of concentrations after the nested PCR- generally 10-100 ng/ μ l
- Do not need great sequencing depth with stool samples- likely low complexity
- Negative samples do not detract from the success of sequencing

Library cleaning and concentration

- Clean and concentrate the pool using 1:1 ratio of AmpureXP beads.
- In detail:
 - Add ampure beads to the pool which will bind the DNA and allow it to be separated by a magnet
 - Wash the pool with 80 % ethanol twice
 - Allow to briefly dry (bead pellet will become matt brown)
 - Resuspend pellet in 51 μ l of water
 - Use magnet to pull the beads out of solution
 - Retain 50 μ l of the eluate (the concentrated pool).

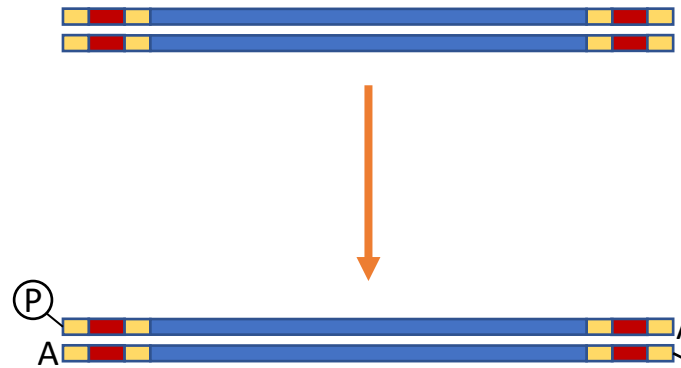


Ampure beads,
Beckman Coulter

Pause point – store library at 4°C

End preparation

- Treatment of DNA with Ultra II End-prep
- Ensures DNA has 3' dA tailed and 5' phosphorylated ends making them sticky for ligation of the sequencing adapter



- Clean sample with ampure beads to remove enzyme

Library preparation

- Performed using the Ligation Sequencing Kit from Oxford Nanopore
- Current version is LSK-114
- Contents:



There vials in the kit that we will not use in this protocol (these are crossed out to the right)

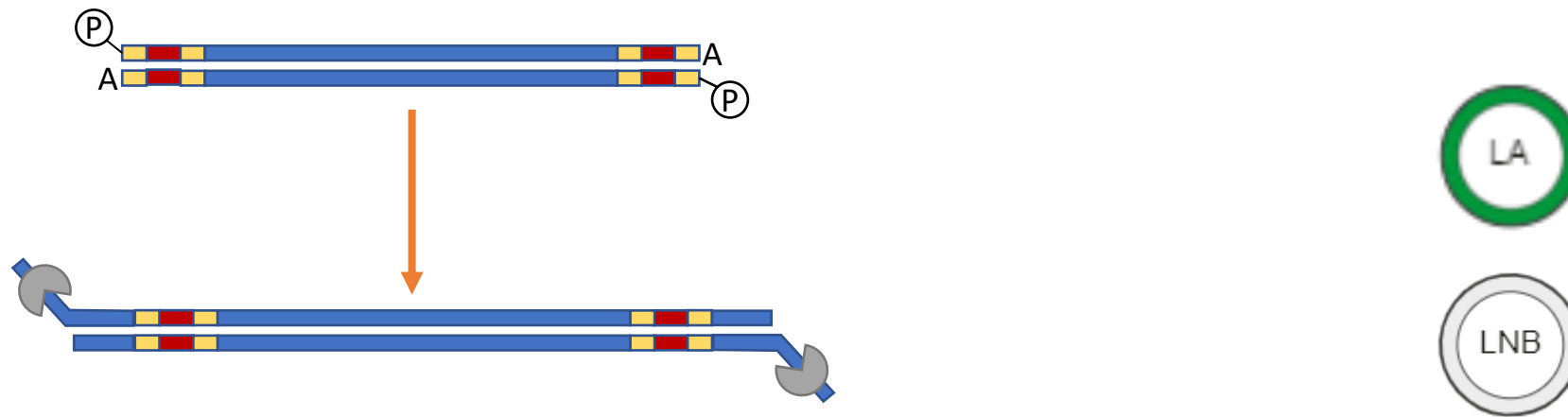


DCS : DNA Control Strand
 LA : Ligation Adapter
 LNB : Ligation Buffer
 LFB : Long Fragment Buffer
 SFB : Short Fragment Buffer
 AXP : AMPure XP Beads

SB : Sequencing Buffer
 EB : Elution Buffer
 LIB : Library Beads
 LIS : Library Solution
 FCF : Flow Cell Flush
 FCT : Flow Cell Tether

Adaptation for Sequencing

- Ligate the Ligation Adapter (LA) which includes the motor protein



- The motor protein controls the speed at which the DNA passes through the pore
- Ligation performed with Quick T4 Ligase (supplied by NEB) and the LNB ligation buffer

Cleaning of Library

- Clean using Ampure beads

BUT

washing is performed with Short Fragment Buffer (SFB) and elution in elution buffer (EB)

- SFB used for washing where no size selection is required.
- Long fragment buffer (LFB) can be used to select for DNA >3kb



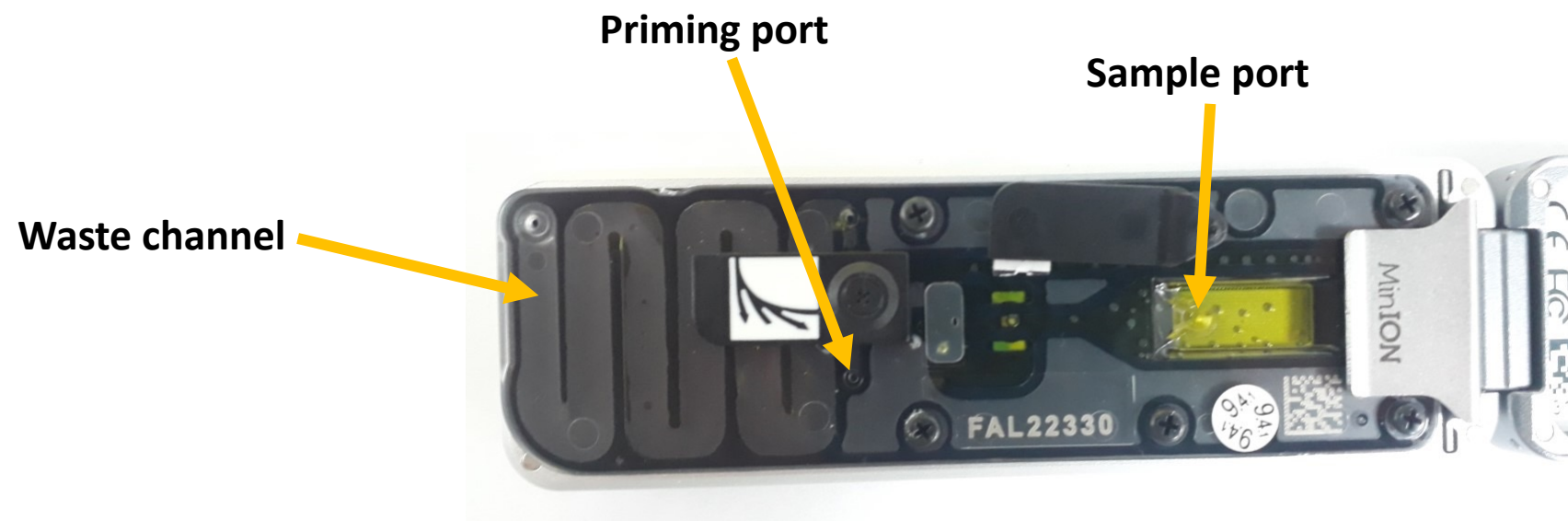
Priming of the flow cell

- Flowcell priming – Flow Cell Flush (FCF), Flow Cell Tether (FCT), and BSA



- The function of the tethers is to help pull DNA down towards the pores for sequencing
- The BSA (not included in the kit) is said to improve stability and throughput
- Added via the priming port in two steps
 - First time with sample port closed
 - Second time with sample port open

Loading the MinION



In this picture both the priming and sample loading ports are open.

Final Library

- The mix loaded onto the MinION contains:
 - Your adapted DNA library
 - Sequencing buffer (SB)
 - Library beads (LB)



- Loaded via the spot-on port, one drop at a time.

Loading the MinION

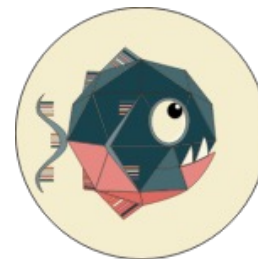
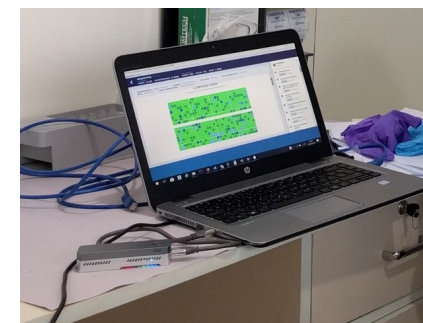
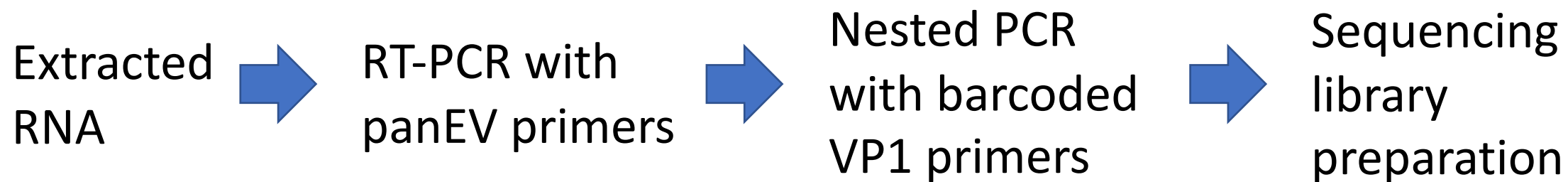


- Before flowcell priming/loading any small air bubbles are removed from beneath the priming port ()
- Waste reagents are removed from the waste port () with both priming port and sample port covers **closed**
- *Video guide: See [“Loading the MinION”](#)*

3. Nanopore sequencing

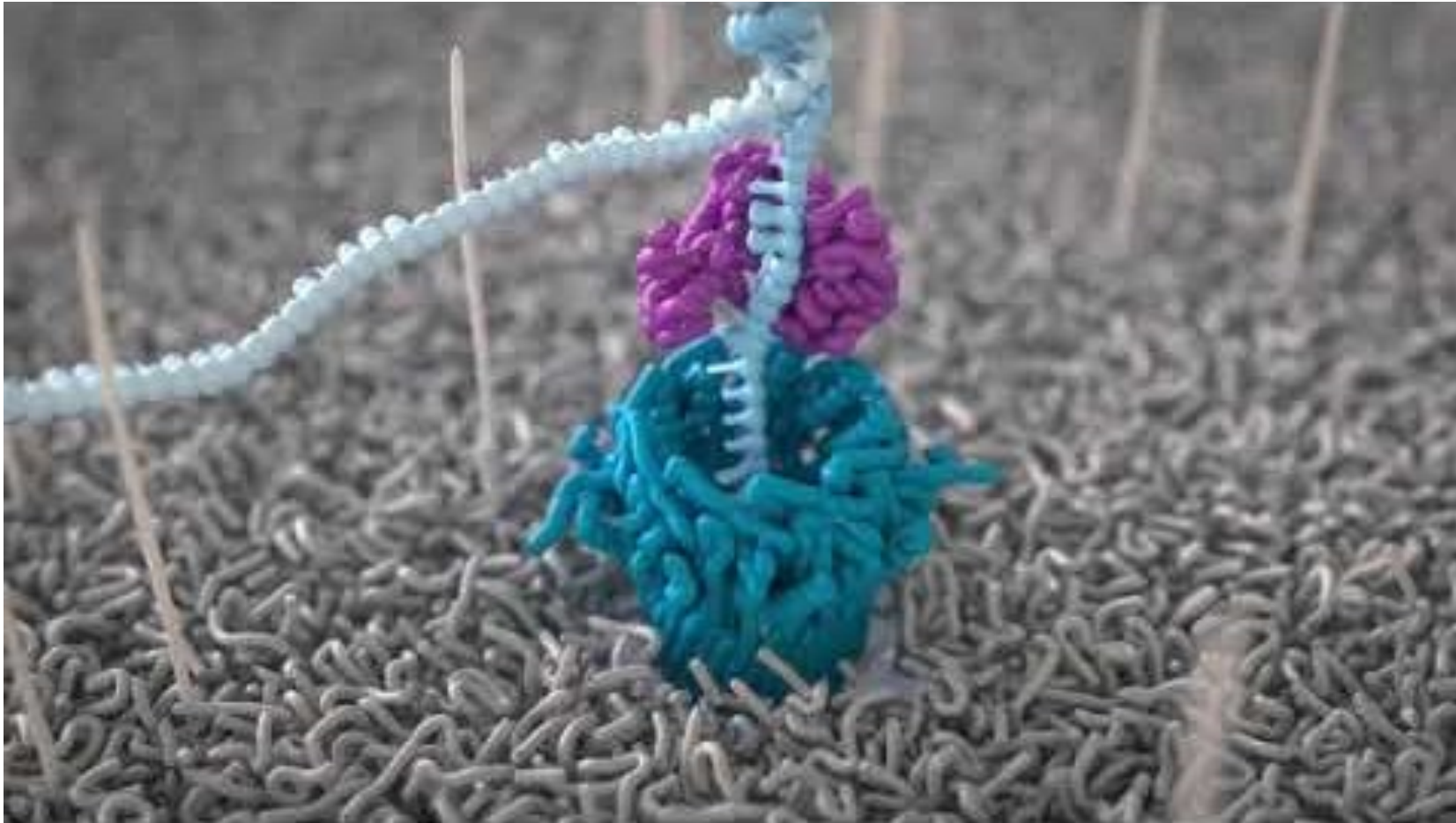
Lecture and discussion

Library Preparation and Nanopore Sequencing



PIRANHA (command line) or
piranhaGUI (point and click)

Nanopore Sequencing



<https://youtu.be/RcP85JHLmnl>

MinION Sequencing

- Long Reads- potentially 100,000s of bp
- Fast- 1,000,000 reads per hour and can process in real time
- “Cheap” (depending on multiplexing)

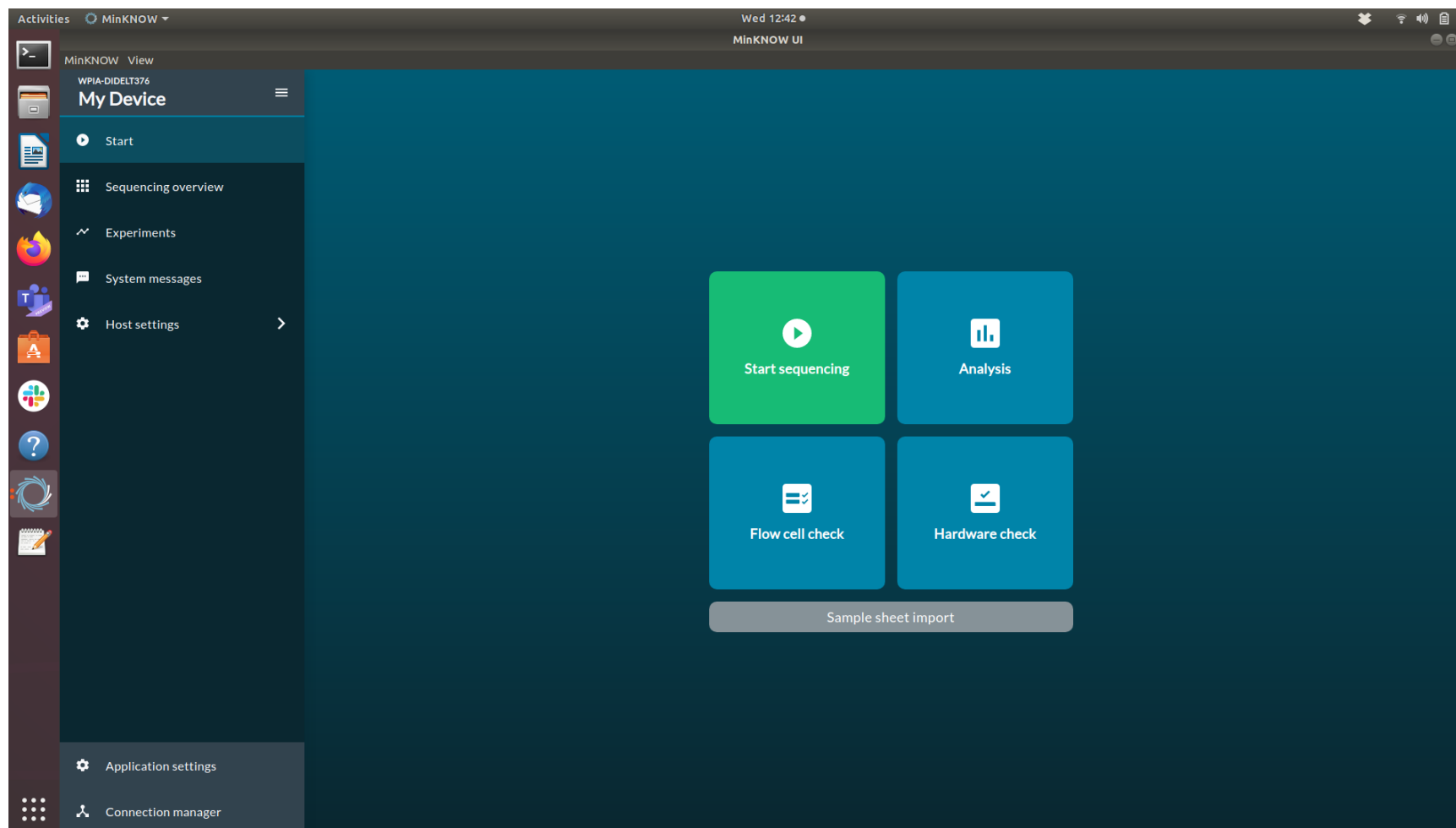
but

- Error rate per base (~0.5% with current highest accuracy settings) - can still give accurate consensus sequences.

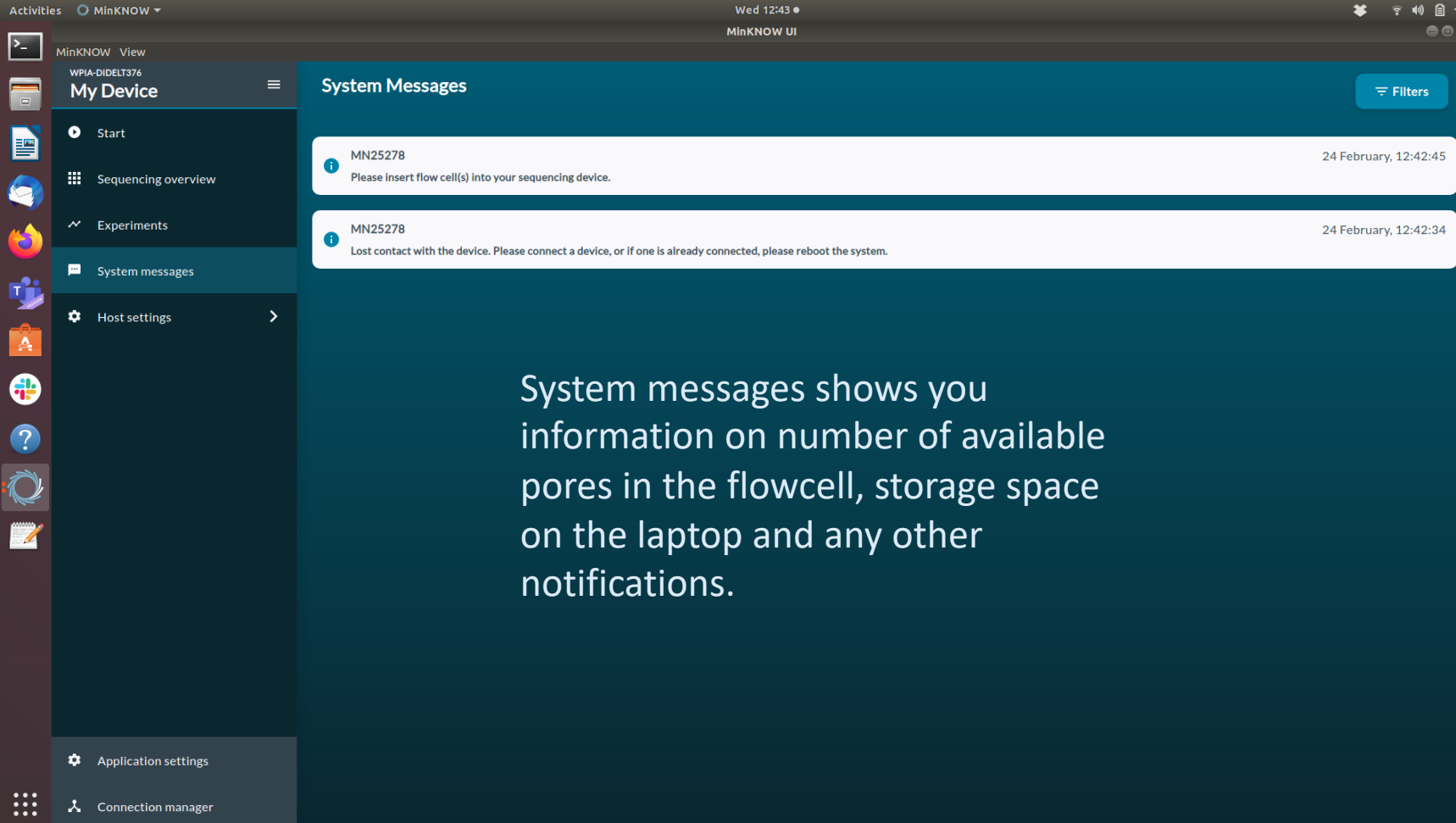
MinKNOW

- Easy to use for setting up and running your sequencing run
- Install following the instructions on the Nanopore website
- Basecalling and demultiplexing performed within MinKNOW

MinKNOW GUI



System messages



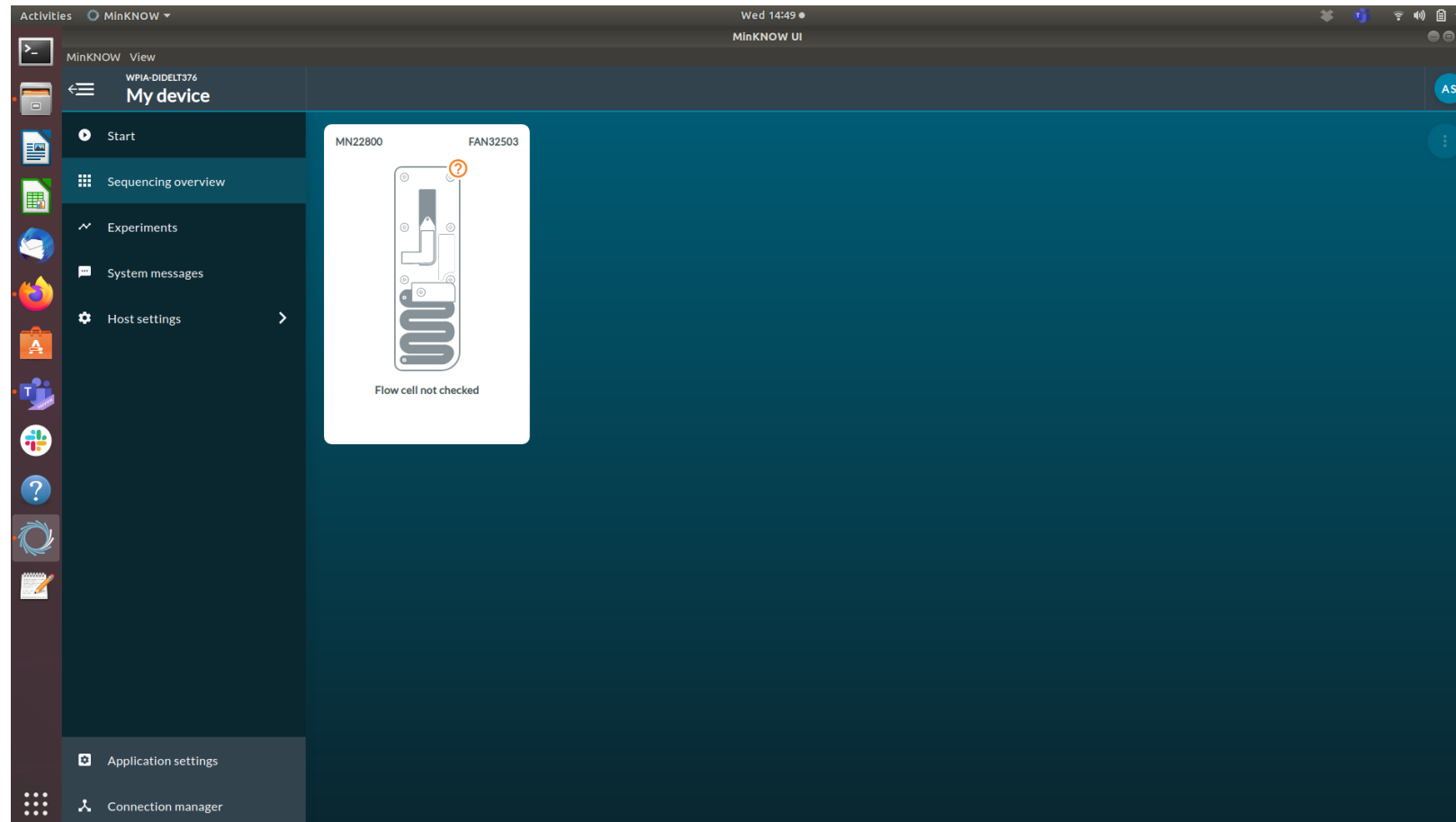
The screenshot shows the MinKNOW UI interface. The top bar displays 'Activities', 'MinKNOW', and the time 'Wed 12:43'. The main content area is titled 'System Messages' and contains two messages:

- MN25278** (24 February, 12:42:45): Please insert flow cell(s) into your sequencing device.
- MN25278** (24 February, 12:42:34): Lost contact with the device. Please connect a device, or if one is already connected, please reboot the system.

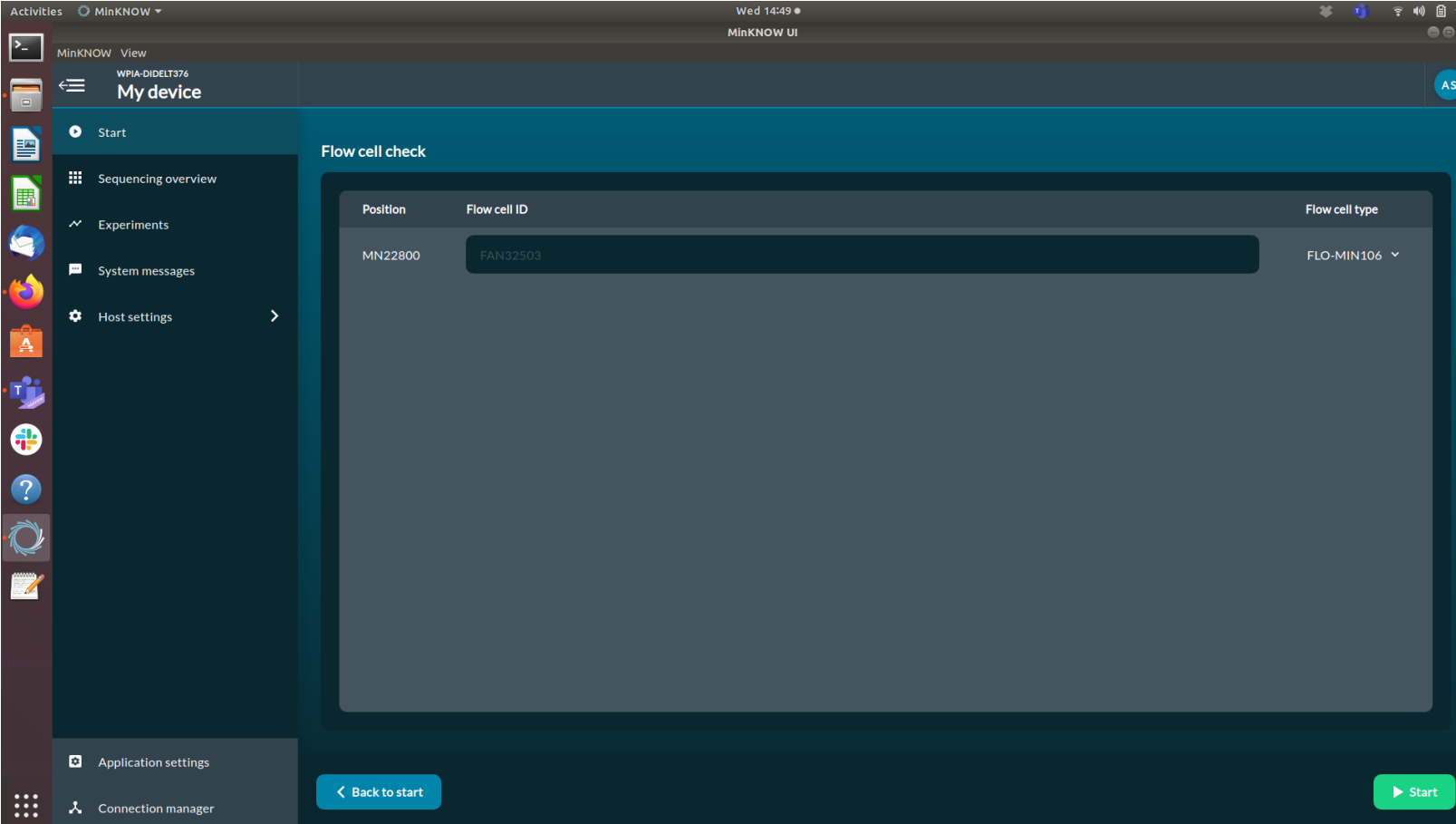
A 'Filters' button is located in the top right corner of the messages list.

System messages shows you information on number of available pores in the flowcell, storage space on the laptop and any other notifications.

Sequencing overview



Running a flow cell check



Activities MinKNOW Wed 14:49 MinKNOW UI

MinKNOW View
WPIA-DIDELT376
My device AS

Start
Sequencing overview
Experiments
System messages
Host settings

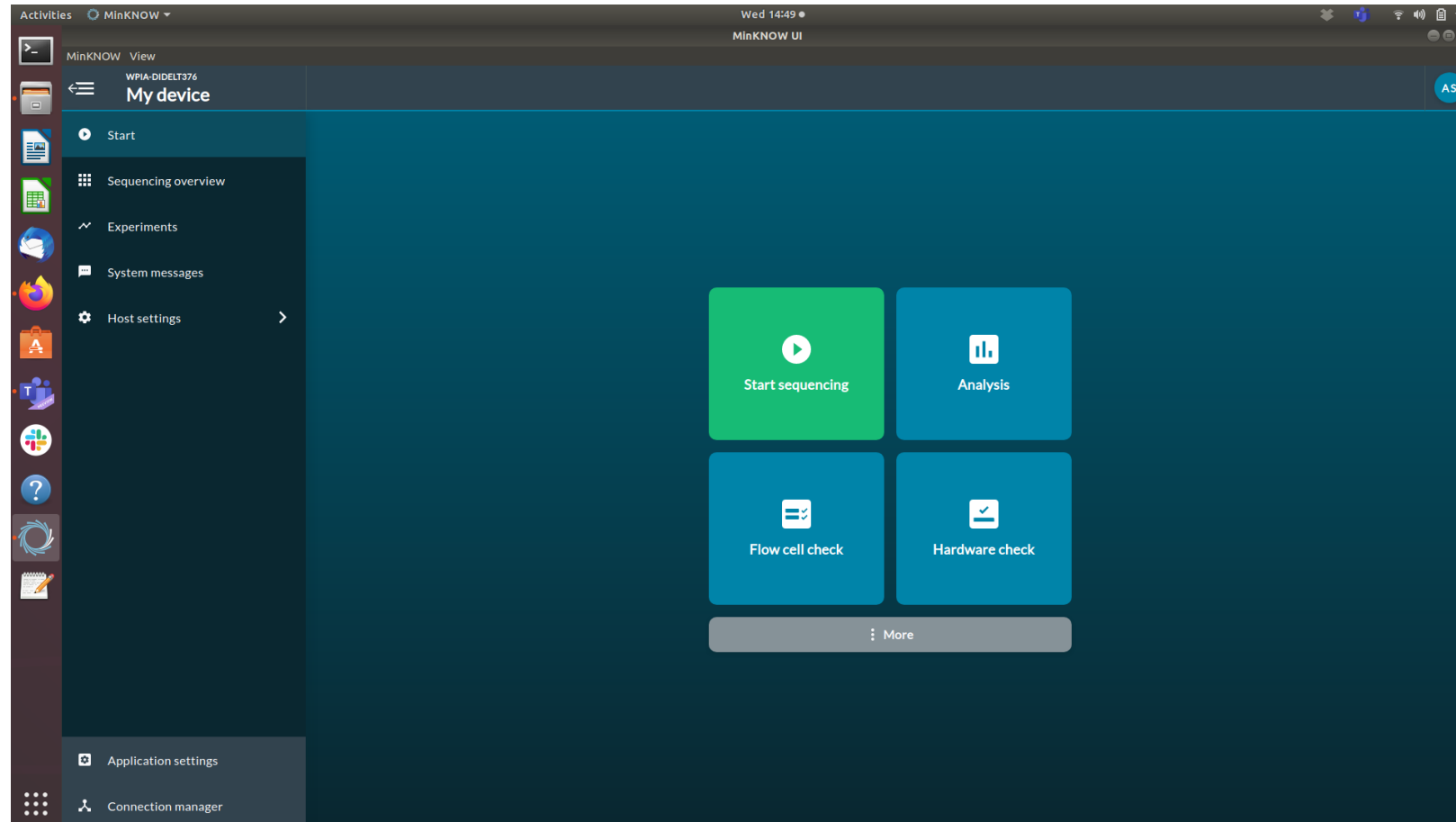
Application settings
Connection manager

Flow cell check

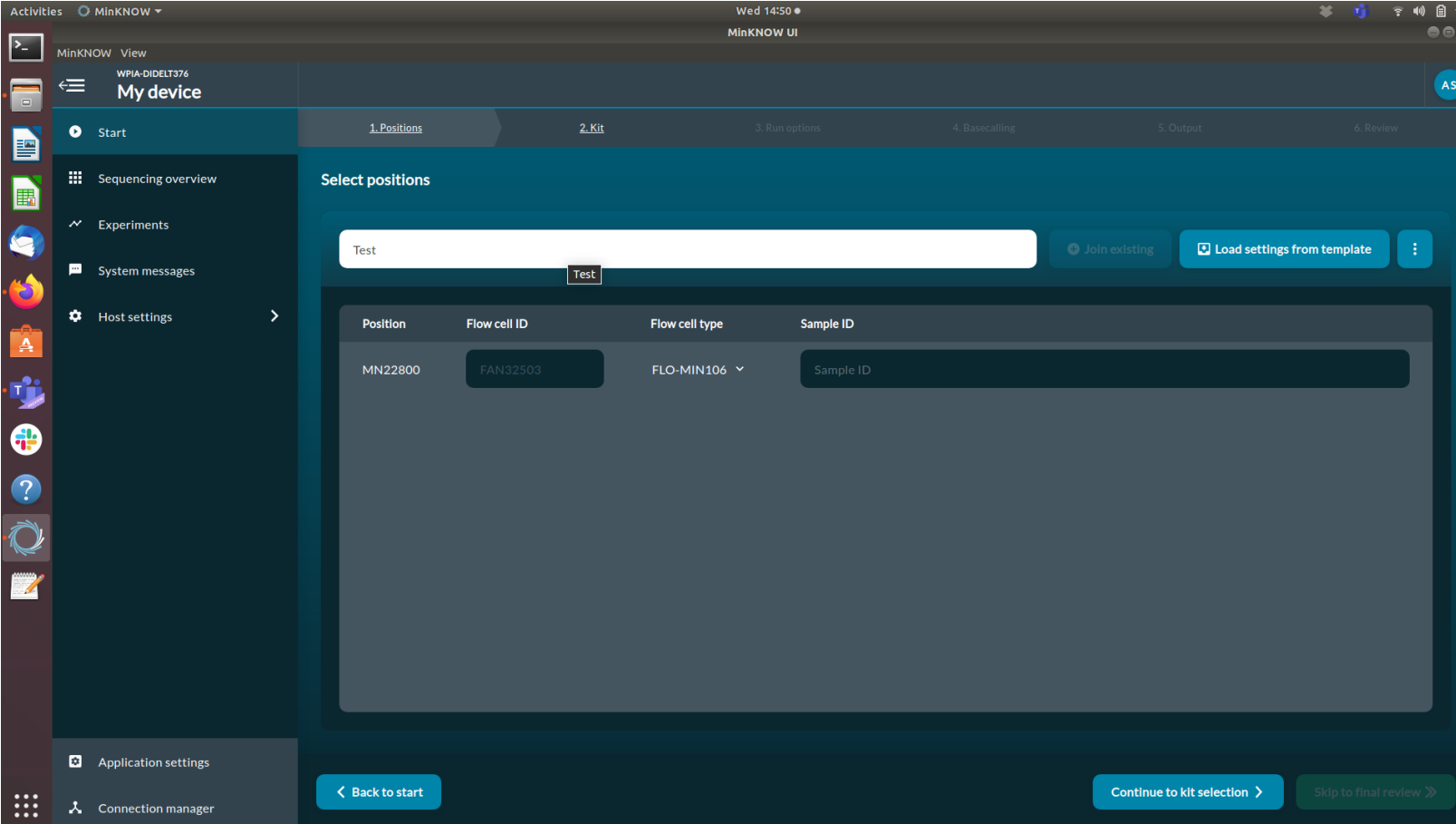
Position	Flow cell ID	Flow cell type
MN22800	FAN32503	FLO-MIN106

< Back to start Start

Starting a sequencing run



Starting a sequencing run



Activities MinKNOW Wed 14:50 MinKNOW UI

WPIA-DIDELT376 My device AS

1. Positions 2. Kit 3. Run options 4. Basecalling 5. Output 6. Review

Select positions

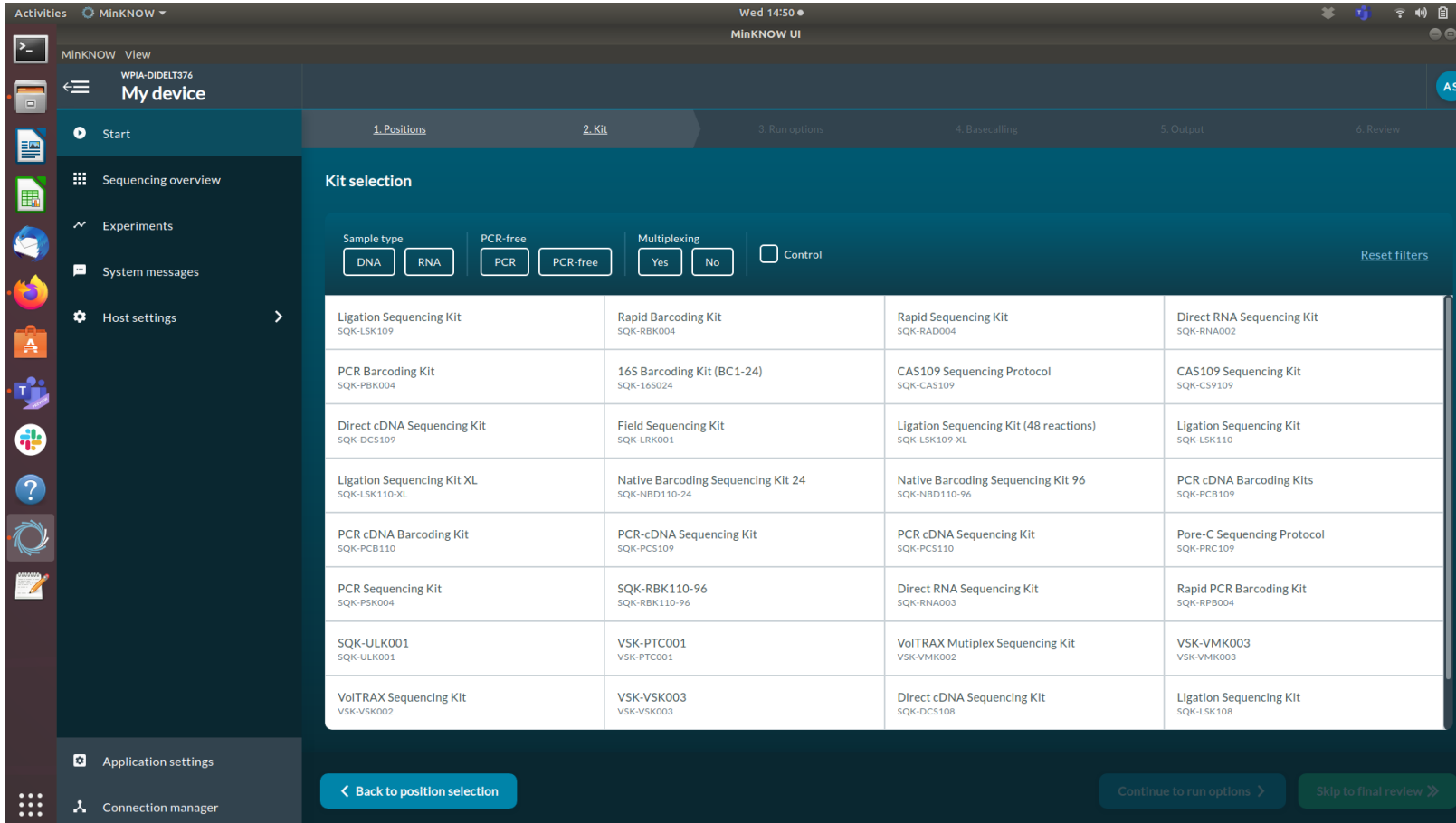
Test

Join existing Load settings from template

Position	Flow cell ID	Flow cell type	Sample ID
MN22800	FAN32503	FLO-MIN106	Sample ID

Back to start Continue to kit selection Skip to final review

Selecting the sequencing kit



Kit selection

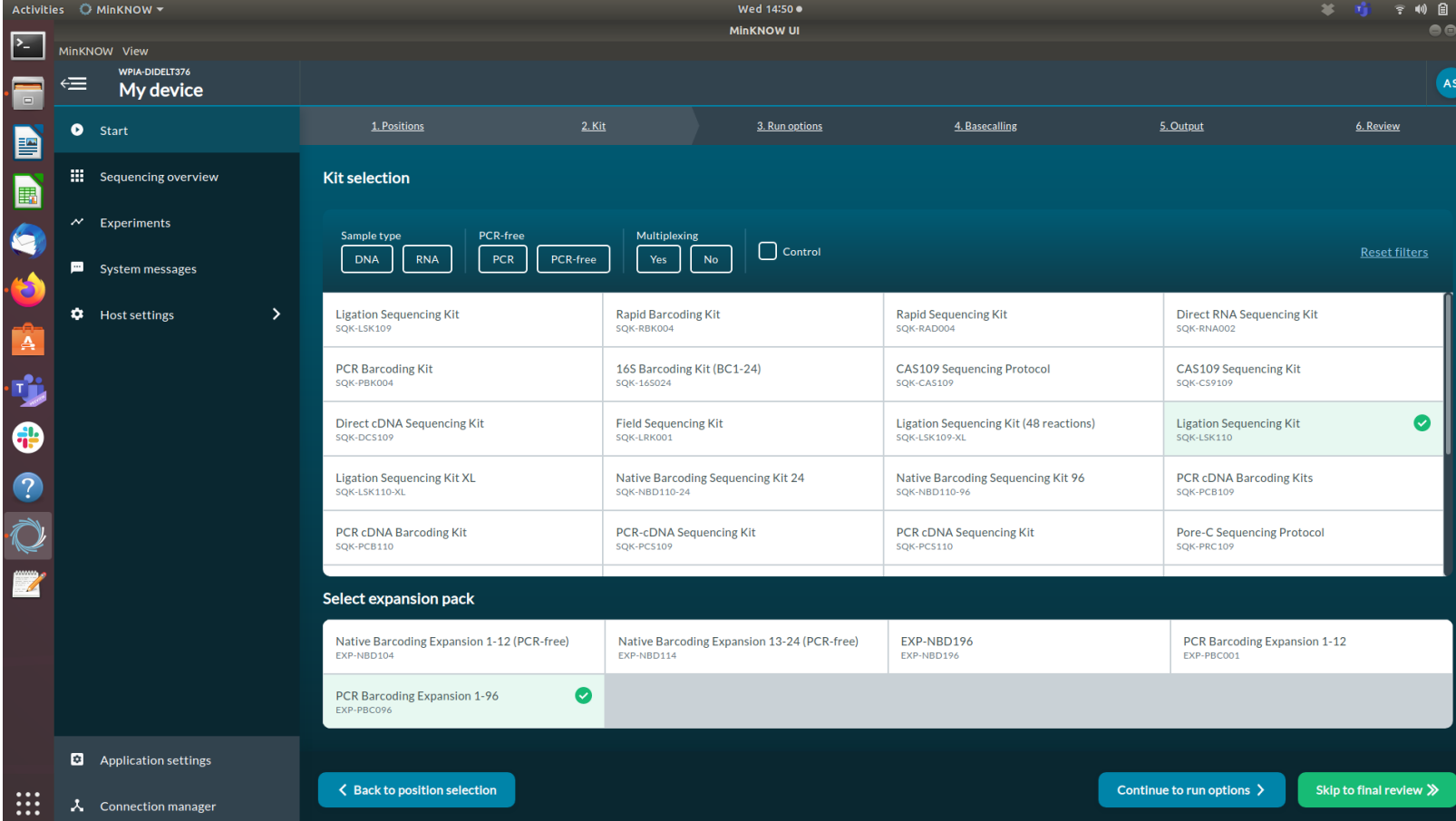
Sample type: PCR-free: Multiplexing: Control [Reset filters](#)

Ligation Sequencing Kit SQK-LSK109	Rapid Barcoding Kit SQK-RBK004	Rapid Sequencing Kit SQK-RAD004	Direct RNA Sequencing Kit SQK-RNA002
PCR Barcoding Kit SQK-PBK004	16S Barcoding Kit (BC1-24) SQK-16S024	CAS109 Sequencing Protocol SQK-CAS109	CAS109 Sequencing Kit SQK-CS9109
Direct cDNA Sequencing Kit SQK-DCS109	Field Sequencing Kit SQK-LRK001	Ligation Sequencing Kit (48 reactions) SQK-LSK109-XL	Ligation Sequencing Kit SQK-LSK110
Ligation Sequencing Kit XL SQK-LSK110-XL	Native Barcoding Sequencing Kit 24 SQK-NBD110-24	Native Barcoding Sequencing Kit 96 SQK-NBD110-96	PCR cDNA Barcoding Kits SQK-PCB109
PCR cDNA Barcoding Kit SQK-PCB110	PCR-cDNA Sequencing Kit SQK-PCS109	PCR cDNA Sequencing Kit SQK-PCS110	Pore-C Sequencing Protocol SQK-PRC109
PCR Sequencing Kit SQK-PSK004	SQK-RBK110-96 SQK-RBK110-96	Direct RNA Sequencing Kit SQK-RNA003	Rapid PCR Barcoding Kit SQK-RPB004
SQK-ULK001 SQK-ULK001	VSK-PTC001 VSK-PTC001	VolTRAX Multiplex Sequencing Kit VSK-VMK002	VSK-VMK003 VSK-VMK003
VolTRAX Sequencing Kit VSK-VSK002	VSK-VSK003 VSK-VSK003	Direct cDNA Sequencing Kit SQK-DCS108	Ligation Sequencing Kit SQK-LSK108

[Back to position selection](#)
[Continue to run options](#)
[Skip to final review](#)

For v14 chemistry and sequencing with the barcoded VP1 primers select SQK-LSK114

Selecting barcodes



Activities MinKNOW View Wed 14:50 MinKNOW UI

WPIA-DIDELT376 My device AS

1. Positions 2. Kit 3. Run options 4. Basecalling 5. Output 6. Review

Kit selection

Sample type: PCR-free: Multiplexing: Control [Reset filters](#)

Ligation Sequencing Kit SQK-LSK109	Rapid Barcoding Kit SQK-RBK004	Rapid Sequencing Kit SQK-RAD004	Direct RNA Sequencing Kit SQK-RNA002
PCR Barcoding Kit SQK-PBK004	16S Barcoding Kit (BC1-24) SQK-16S024	CAS109 Sequencing Protocol SQK-CAS109	CAS109 Sequencing Kit SQK-CS9109
Direct cDNA Sequencing Kit SQK-DCS109	Field Sequencing Kit SQK-LRK001	Ligation Sequencing Kit (48 reactions) SQK-LSK109-XL	Ligation Sequencing Kit SQK-LSK110 <input checked="" type="checkbox"/>
Ligation Sequencing Kit XL SQK-LSK110-XL	Native Barcoding Sequencing Kit 24 SQK-NBD110-24	Native Barcoding Sequencing Kit 96 SQK-NBD110-96	PCR cDNA Barcoding Kits SQK-PCB109
PCR cDNA Barcoding Kit SQK-PCB110	PCR-cDNA Sequencing Kit SQK-PCS109	PCR cDNA Sequencing Kit SQK-PCS110	Pore-C Sequencing Protocol SQK-PRC109

Select expansion pack

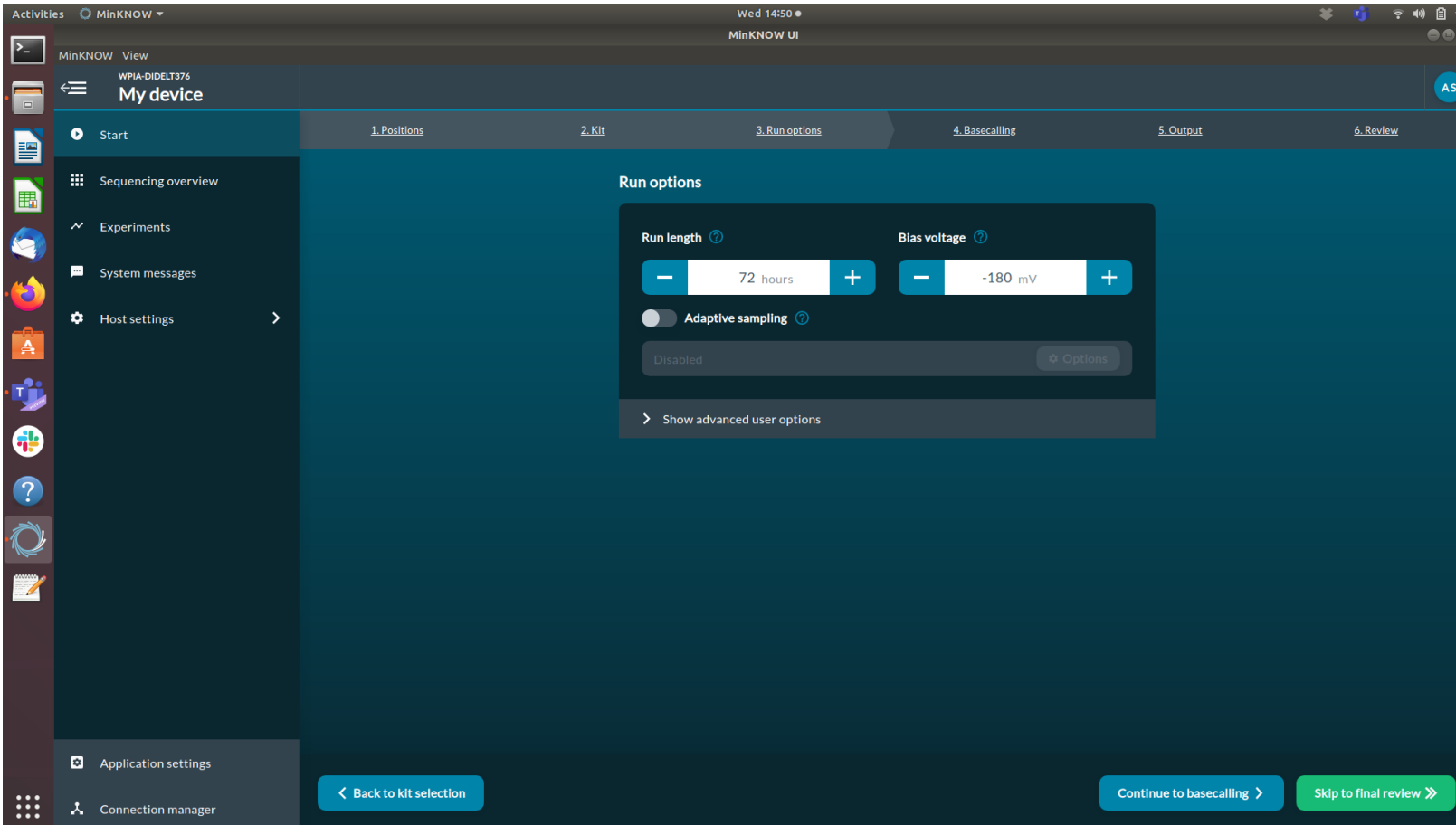
Native Barcoding Expansion 1-12 (PCR-free) EXP-NBD104	Native Barcoding Expansion 13-24 (PCR-free) EXP-NBD114	EXP-NBD196 EXP-NBD196	PCR Barcoding Expansion 1-12 EXP-PBC001
PCR Barcoding Expansion 1-96 EXP-PBC096 <input checked="" type="checkbox"/>			

[Back to position selection](#) [Continue to run options](#) [Skip to final review](#)

For the barcoded VP1 primers select EXP-PBC096

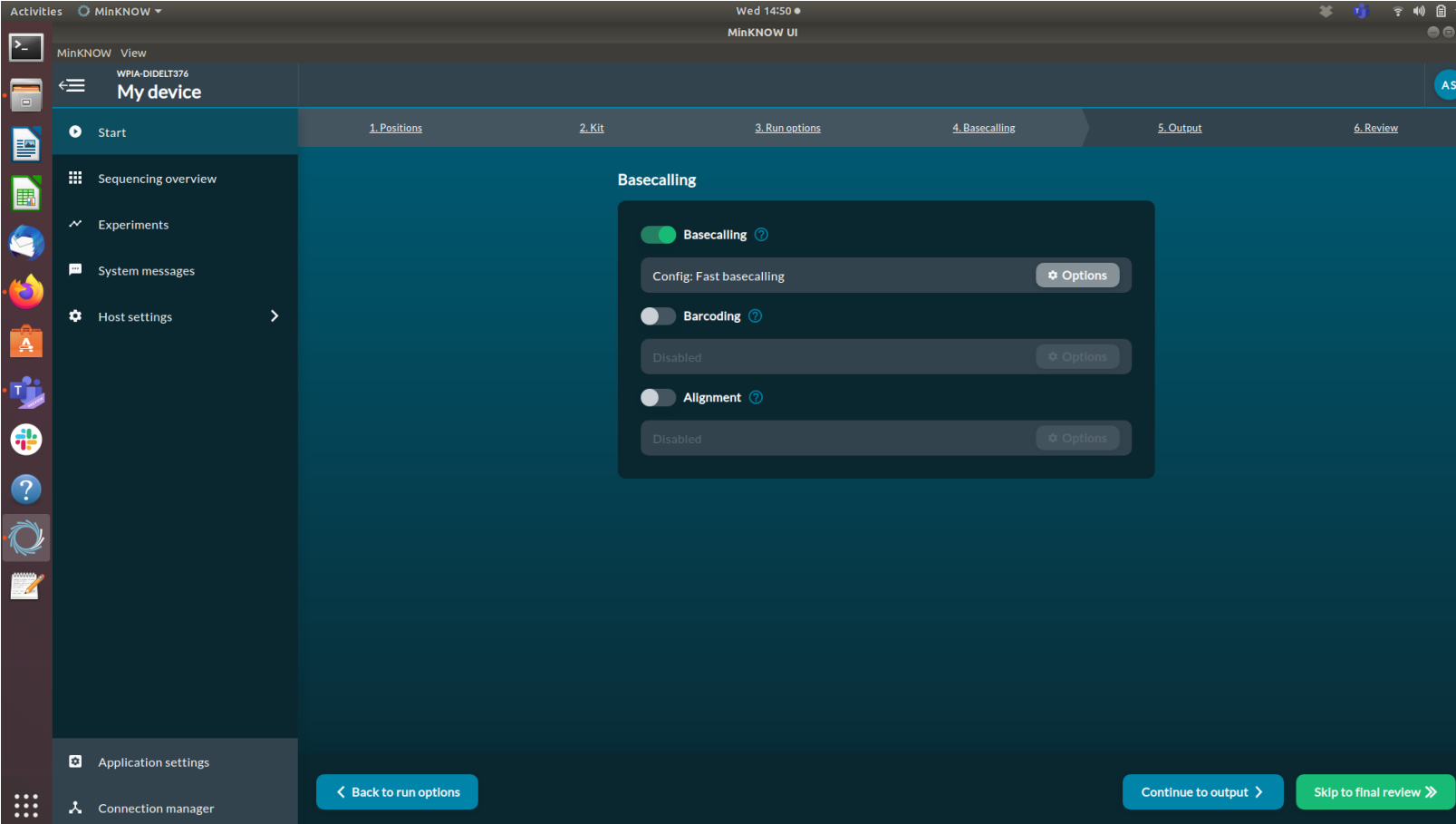
(note- this step may not be available in the most recent software version, but can be performed in the analysis tab after the run)

Setting run length and voltage



The screenshot shows the MinKNOW UI interface. The top navigation bar includes 'Activities', 'MinKNOW', and 'View'. The main header displays 'WPIA-DIDELT376' and 'My device'. A progress bar at the top indicates the current step: '3. Run options', with other steps being '1. Positions', '2. Kit', '4. Basecalling', '5. Output', and '6. Review'. The left sidebar contains navigation options: Start, Sequencing overview, Experiments, System messages, Host settings, Application settings, and Connection manager. The central 'Run options' panel features two input fields: 'Run length' set to '72 hours' and 'Bias voltage' set to '-180 mV'. Below these is a toggle for 'Adaptive sampling' which is currently disabled. At the bottom of the panel is a 'Show advanced user options' link. The bottom navigation bar contains three buttons: 'Back to kit selection', 'Continue to basecalling', and 'Skip to final review'.

Basecalling options

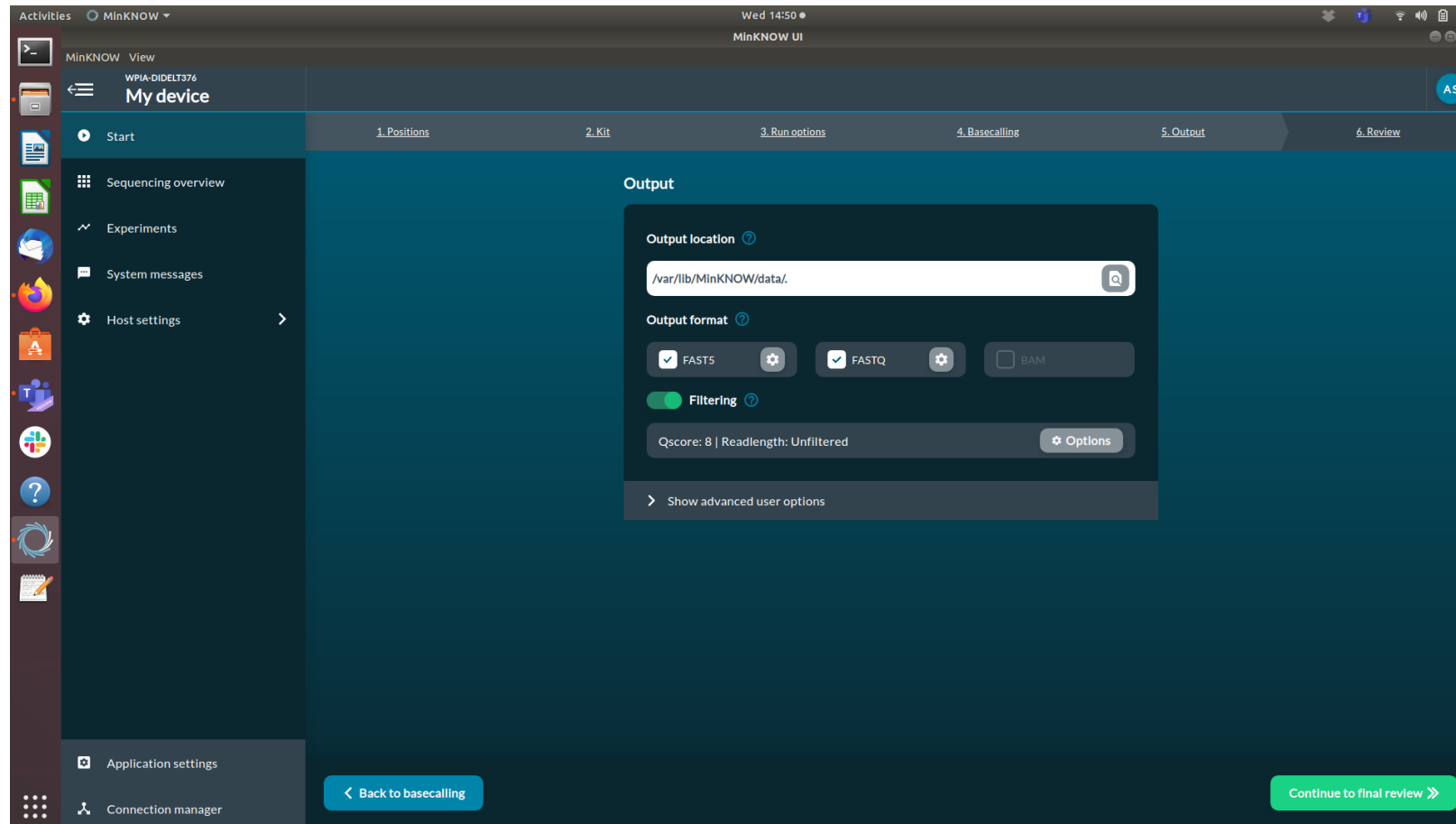


The screenshot shows the MinKNOW UI interface. The top bar displays 'Activities', 'MinKNOW', and 'Wed 14:50'. The main window title is 'MinKNOW UI'. The left sidebar shows 'My device' with a sub-header 'WPIA-DIDELT376'. The sidebar menu includes 'Start', 'Sequencing overview', 'Experiments', 'System messages', 'Host settings', 'Application settings', and 'Connection manager'. The main content area is titled 'Basecalling' and features a progress bar with steps: '1. Positions', '2. Kit', '3. Run options', '4. Basecalling', '5. Output', and '6. Review'. The 'Basecalling' settings are displayed in a dark panel with the following options:

- Basecalling** ⓘ
Config: Fast basecalling Options
- Barcoding** ⓘ
Disabled Options
- Alignment** ⓘ
Disabled Options

At the bottom of the screen, there are three buttons: 'Back to run options', 'Continue to output', and 'Skip to final review'.

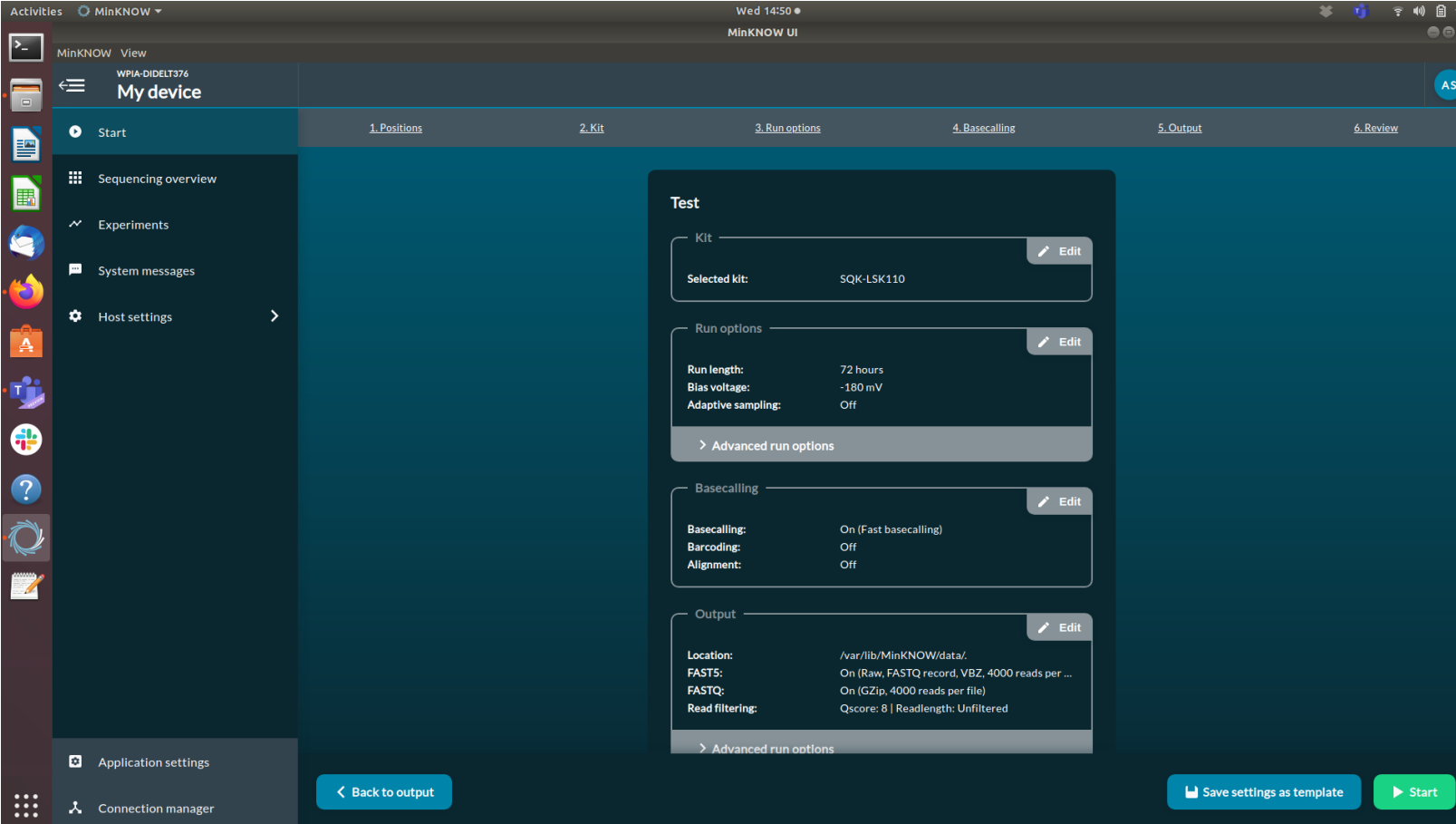
Output options



Select an analysis folder on your desktop for easy access

On a Linux computer the default is `var/lib/minknow/data`

Review the settings



The screenshot shows the MinKNOW UI interface. The top bar displays 'Wed 14:50' and 'MinKNOW UI'. The main header includes 'MinKNOW View', 'WPIA-DIDELT376', and 'My device'. A navigation bar at the top contains six steps: 1. Positions, 2. Kit, 3. Run options, 4. Basecalling, 5. Output, and 6. Review. The left sidebar lists various system and application settings. The main content area is titled 'Test' and contains four sections: 'Kit' (Selected kit: SQK-LSK110), 'Run options' (Run length: 72 hours, Bias voltage: ~180 mV, Adaptive sampling: Off), 'Basecalling' (Basecalling: On (Fast basecalling), Barcoding: Off, Alignment: Off), and 'Output' (Location: /var/lib/MinKNOW/data/, FAST5: On (Raw, FASTQ record, VBZ, 4000 reads per ...), FASTQ: On (GZip, 4000 reads per file), Read filtering: Qscore: 8 | Readlength: Unfiltered). At the bottom, there are buttons for 'Back to output', 'Save settings as template', and 'Start'.

Activities MinKNOW Wed 14:50 MinKNOW UI

MinKNOW View WPIA-DIDELT376 My device AS

Start 1. Positions 2. Kit 3. Run options 4. Basecalling 5. Output 6. Review

Sequencing overview Experiments System messages Host settings

Application settings Connection manager

Test

Kit [Edit](#)

Selected kit: SQK-LSK110

Run options [Edit](#)

Run length: 72 hours
Bias voltage: ~180 mV
Adaptive sampling: Off

> Advanced run options

Basecalling [Edit](#)

Basecalling: On (Fast basecalling)
Barcoding: Off
Alignment: Off

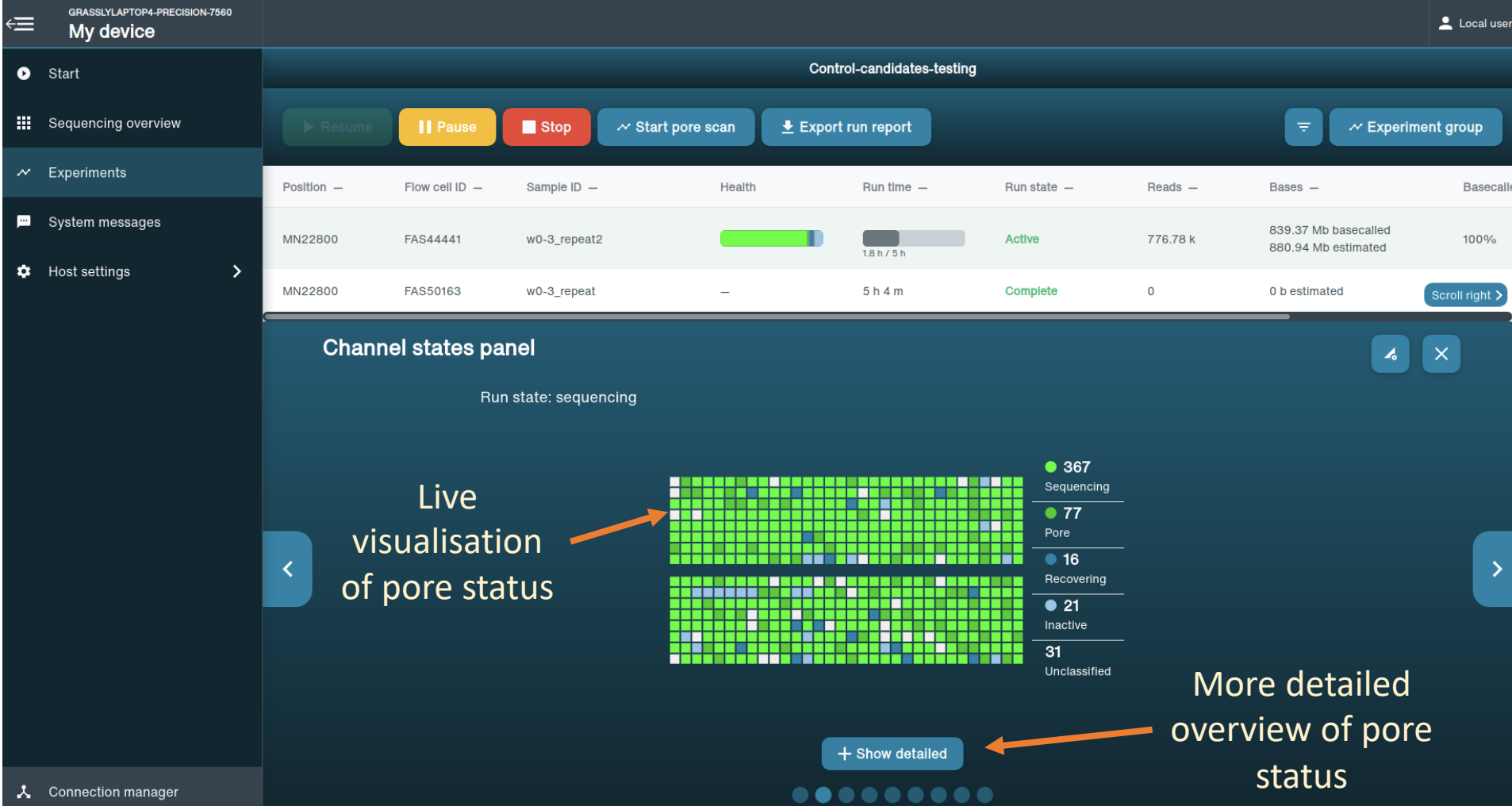
Output [Edit](#)

Location: /var/lib/MinKNOW/data/
FAST5: On (Raw, FASTQ record, VBZ, 4000 reads per ...
FASTQ: On (GZip, 4000 reads per file)
Read filtering: Qscore: 8 | Readlength: Unfiltered

> Advanced run options

< Back to output Save settings as template Start

Sequencing Run



GRASSLYLAPTOP4-PRECISION-7560
My device Local user

Control-candidates-testing

▶ Resume **Pause** **Stop** ⌵ Start pore scan **Export run report** ☰ Experiment group

Position	Flow cell ID	Sample ID	Health	Run time	Run state	Reads	Bases	Basecall
MN22800	FAS44441	w0-3_repeat2	<div style="width: 80%; background-color: #28a745;"></div>	1.8 h / 5 h	Active	776.78 k	839.37 Mb basecalled 880.94 Mb estimated	100%
MN22800	FAS50163	w0-3_repeat	—	5 h 4 m	Complete	0	0 b estimated	Scroll right >

Channel states panel 🔍 ✕

Run state: sequencing

Live visualisation of pore status

- 367 Sequencing
- 77 Pore
- 16 Recovering
- 21 Inactive
- 31 Unclassified

More detailed overview of pore status

Show detailed

Connection manager

Sequencing Run

GRASSLYLAPTOP4-PRECISION-7560 Local user

My device

Control-candidates-testing

▶ Resume
⏸ Pause
■ Stop
⌵ Start pore scan
📄 Export run report
☰
⌵ Experiment group

Position	Flow cell ID	Sample ID	Health	Run time	Run state	Reads	Bases	Basecall
MN22800	FAS44441	w0-3_repeat2	<div style="width: 80%; background-color: #00ff00;"></div>	1.8 h / 5 h	Active	777.98 k	840.66 Mb basecalled 882.41 Mb estimated	100%
MN22800	FAS50163	w0-3_repeat	—	5 h 4 m	Complete	0	0 b estimated	Scroll right >

Channel states panel

Run state: sequencing

- 363 Strand
- 6 Adapter
- 77 Single Pore
- 8 Unavailable
- 8 Active Feedback
- 19 No Pore From Scan
- 1 Out Of Range 2

— Show less

Connection manager

Sequencing Run

GRASSLYLAPTOP4-PRECISION-7560 Local user

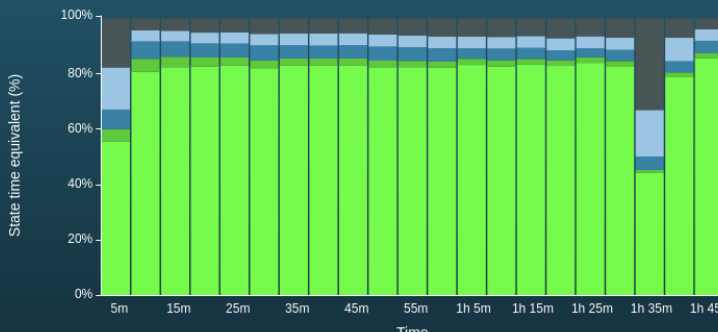
My device

Control-candidates-testing

▶ Resume
⏸ Pause
■ Stop
⌵ Start pore scan
⬇ Export run report
⌵ Experiment group

Position	Flow cell ID	Sample ID	Health	Run time	Run state	Reads	Bases	Basecall
MN22800	FAS44441	w0-3_repeat2	<div style="width: 80%; height: 10px; background-color: #00ff00;"></div>	1.8 h / 5 h	Active	780.32 k	843.26 Mb basecalled 885.35 Mb estimated	100%
MN22800	FAS50163	w0-3_repeat	—	5 h 4 m	Complete	0	0 b estimated	Scroll right >

Pore activity



State time equivalent (%)

Time

- sequencing ✓
- pore ✓
- recovering ✓
- inactive ✓
- unclassified ✓

+ Show detailed
⚙ Display settings

Connection manager

Sequencing Run

GRASSLYLAPTOP4-PRECISION-7560 Local user

My device

Control-candidates-testing

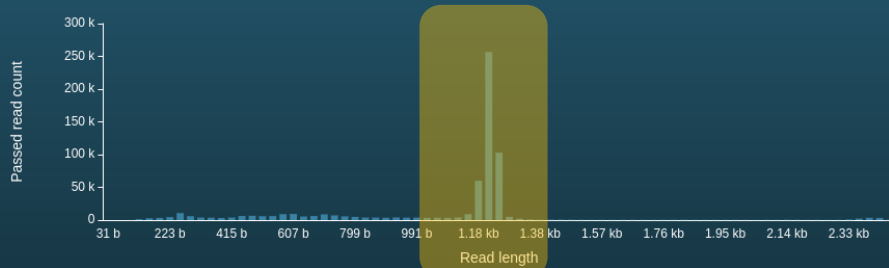
▶ Resume
⏸ Pause
■ Stop
↶ Start pore scan
📄 Export run report
☰
↶ Experiment group

Position	Flow cell ID	Sample ID	Health	Run time	Run state	Reads	Bases	Basecall
MN22800	FAS44441	w0-3_repeat2	<div style="width: 80%; height: 10px; background-color: #00ff00;"></div>	1.8 h / 5 h	Active	781.45 k	844.48 Mb basecalled 886.77 Mb estimated	100%
MN22800	FAS50163	w0-3_repeat	—	5 h 4 m	Complete	0	0 b estimated	Scroll right >

Read length histogram 🔍 ✕

Assess size of reads being sequenced

Estimated N50: 1.2 kb



Adjusted read length range: 1.18 kb to 1.38 kb

Adjust here to focus on specific read lengths

Read
Bases
Lengths
Counts
Estimated
Basecalled
🔄 Reset
 Hide outliers
 Split by read end reason

Connection manager

Sequencing Run

GRASSLYLAPTOP4-PRECISION-7560
Local user

My device

- Start
- Sequencing overview
- Experiments
- System messages
- Host settings

Control-candidates-testing

Resume
Pause
Stop
Start pore scan
Export run report

Position	Flow cell ID	Sample ID	Health	Run time	Run state	Reads	Bases	Basecall
MN22800	FAS44441	w0-3_repeat2	<div style="width: 80%; height: 10px; background-color: #4caf50;"></div>	1.8 h / 5 h	Active	782.6 k	845.78 Mb basecalled 888.33 Mb estimated	100%
MN22800	FAS50163	w0-3_repeat	—	5 h 4 m	Complete	0	0 b estimated	Scroll right >

Barcode hits

● Passed reads



Barcode

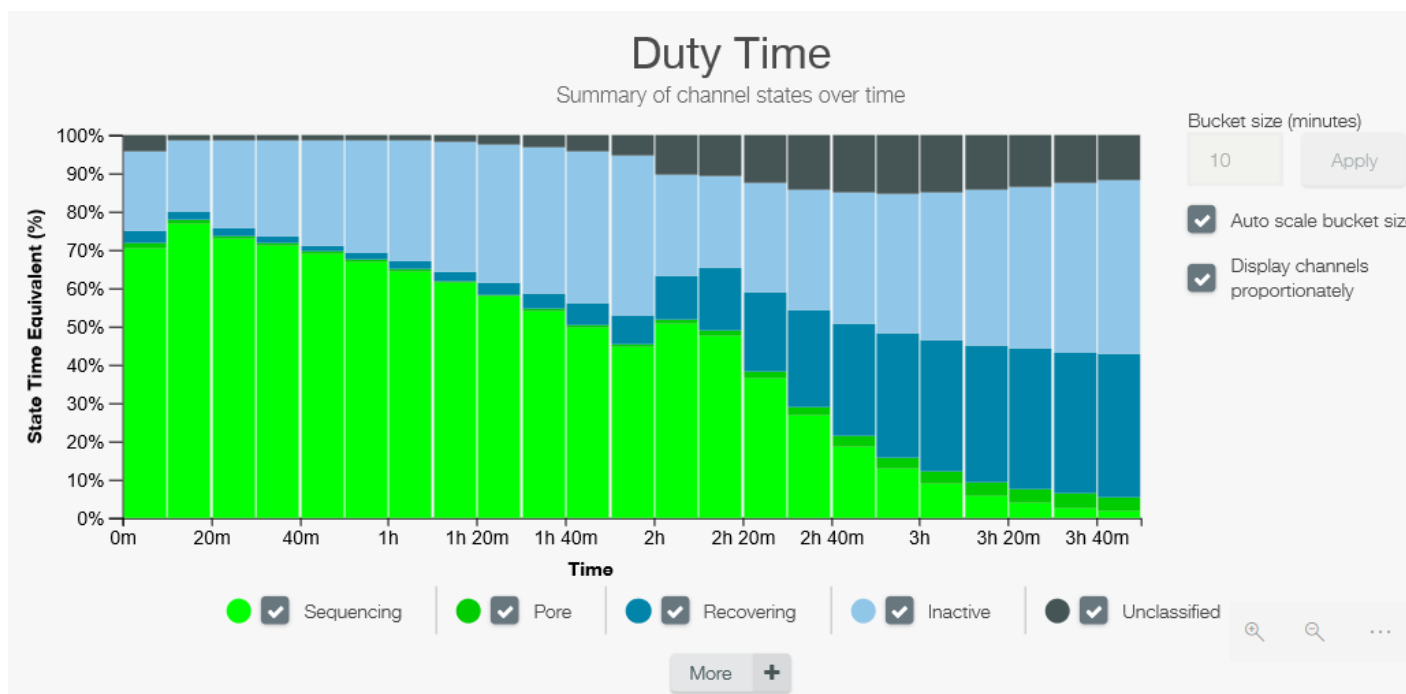
Reads
Bases

Sort
A-Z
Z-A
↕
↕

Display failed
 Display unclassified
 Hide zero values
 Reset selection

See which barcodes have been detected and how many reads each

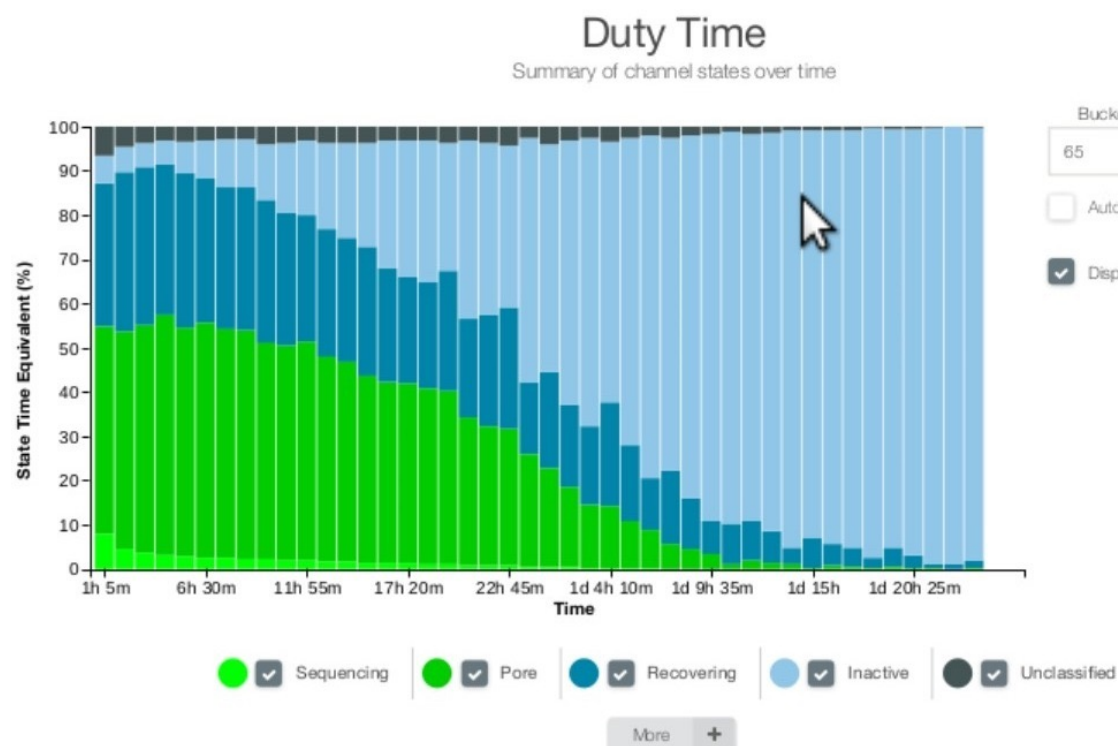
Examples of a bad start



Rapid pore death may be caused by:

- Contamination
- Using an old flowcell
- Software fault

Examples of a bad start



Low pore occupancy may be caused by:

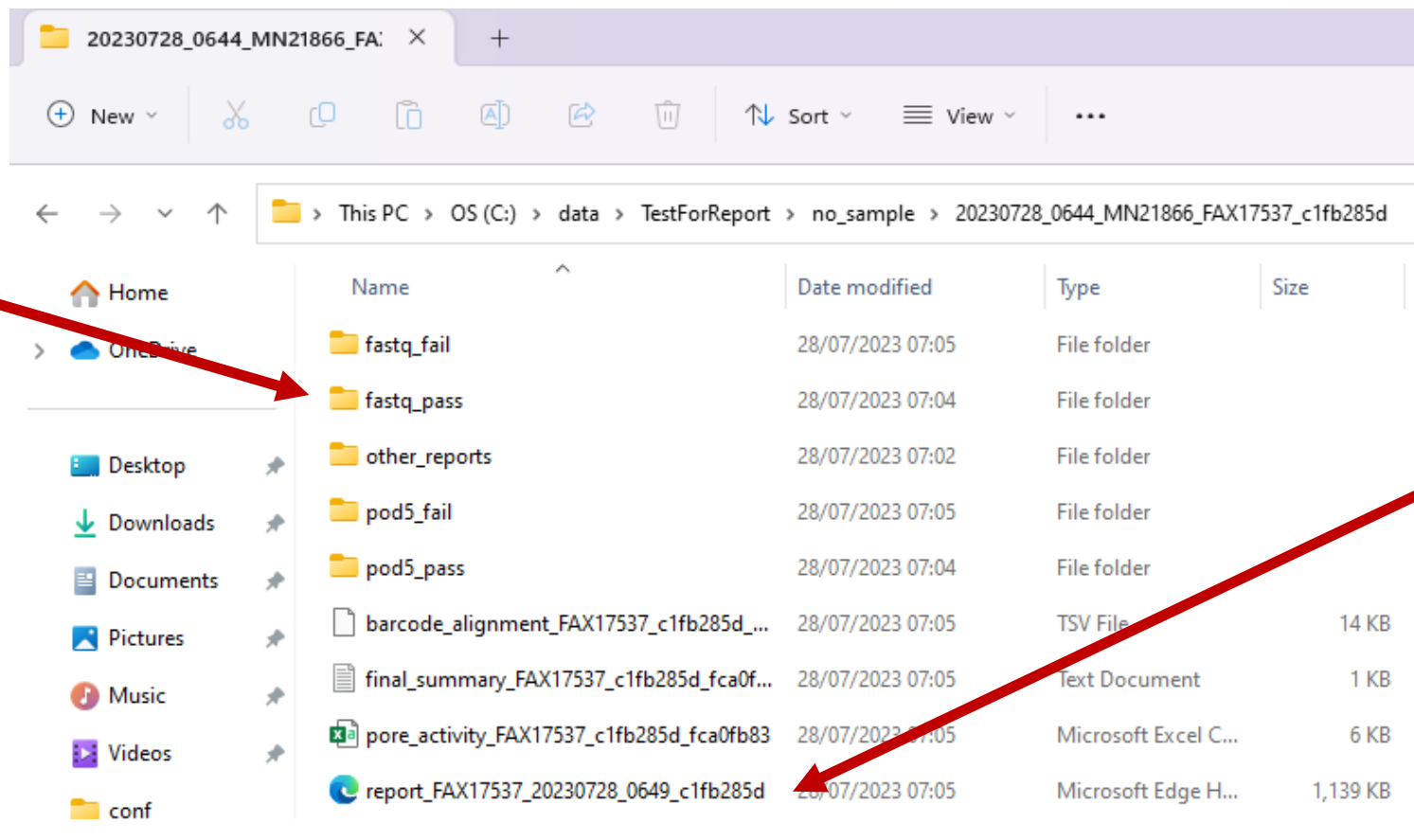
- Not loading enough sample onto the flowcell
- Failed adaptation for sequencing

4. Post-sequencing run checks

Lecture and discussion

1. Confirm that your run created data

Confirm that there is data in the fastq_pass folder

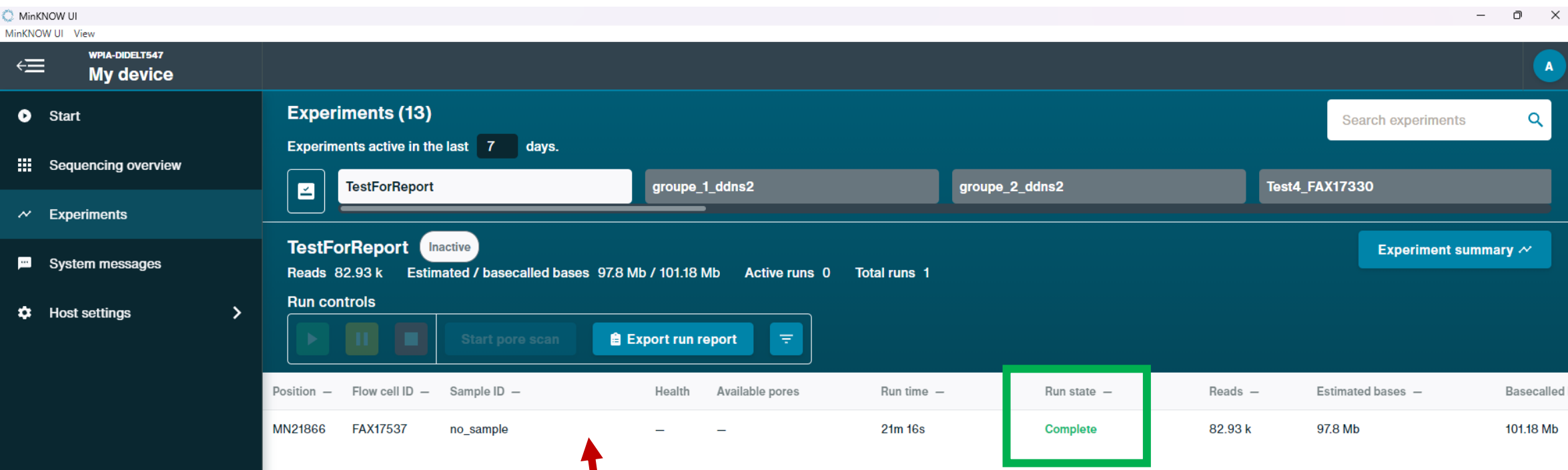


Confirm that your run report has been created (this contains a summary of the run performance)

If there is no data, you can restart your run in MinKNOW.

If there is not enough data, you can click “join existing” when you start a run to ADD the data to your prior run

2. Check your run is complete in MinKNOW



MinKNOW UI

WPIA-DIDELT547
My device

Start

Sequencing overview

Experiments

System messages

Host settings

Experiments (13)

Search experiments

Experiments active in the last 7 days.

TestForReport groupe_1_ddns2 groupe_2_ddns2 Test4_FAX17330

TestForReport **Inactive** Experiment summary

Reads 82.93 k Estimated / basecalled bases 97.8 Mb / 101.18 Mb Active runs 0 Total runs 1

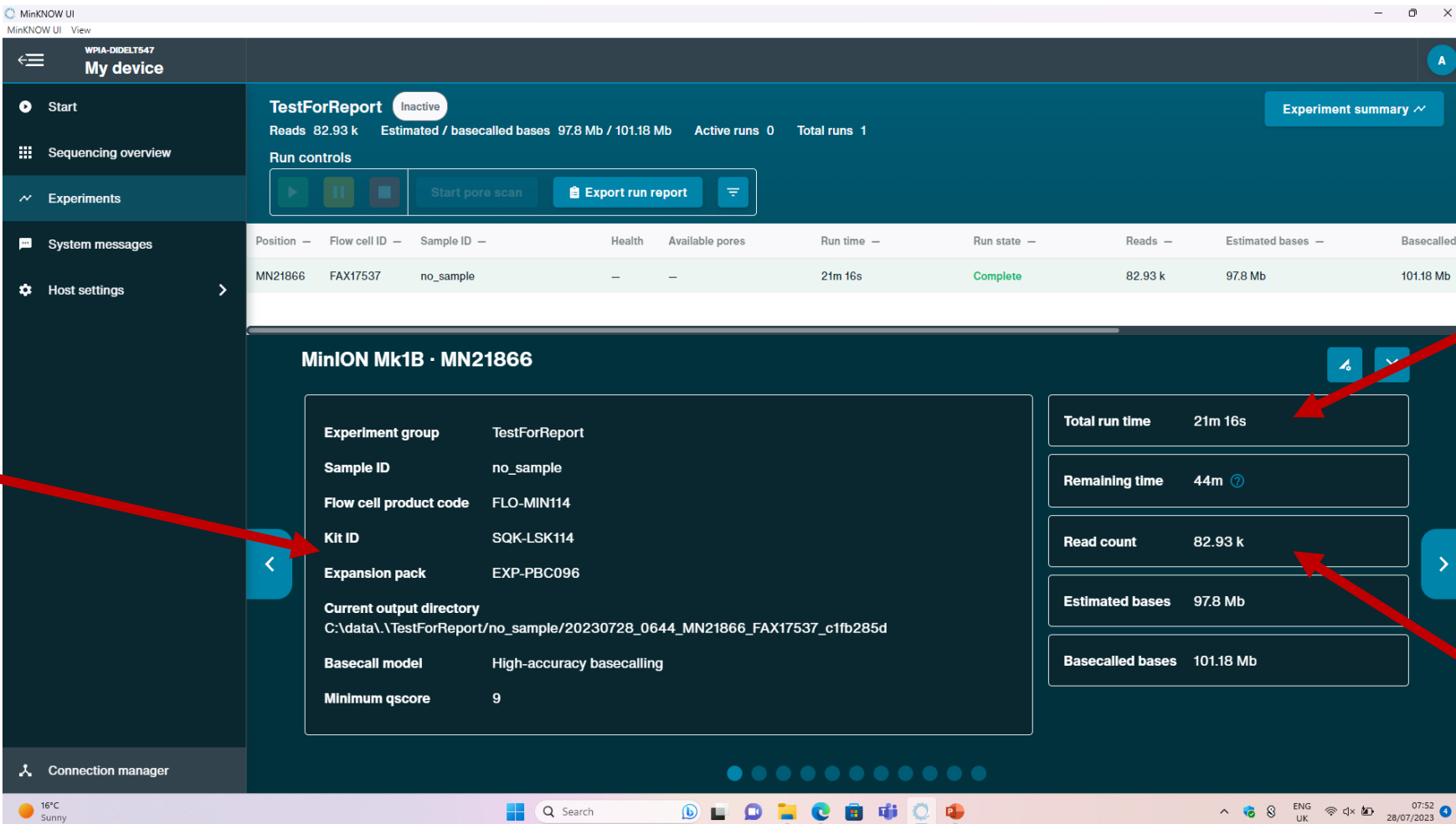
Run controls

Start pore scan Export run report

Position	Flow cell ID	Sample ID	Health	Available pores	Run time	Run state	Reads	Estimated bases	Basecalled
MN21866	FAX17537	no_sample	—	—	21m 16s	Complete	82.93 k	97.8 Mb	101.18 Mb

Click on your run to see more details (will take you to the screen on the next slide)

3. Check your run performance



The screenshot shows the MinKNOW UI interface for a MinION Mk1B device. The main display area is titled 'TestForReport' and shows the following metrics:

- Reads: 82.93 k
- Estimated / basecalled bases: 97.8 Mb / 101.18 Mb
- Active runs: 0
- Total runs: 1

Below these metrics is a 'Run controls' section with buttons for 'Start pore scan' and 'Export run report'. A table below that shows the run details:

Position	Flow cell ID	Sample ID	Health	Available pores	Run time	Run state	Reads	Estimated bases	Basecalled bases
MN21866	FAX17537	no_sample	—	—	21m 16s	Complete	82.93 k	97.8 Mb	101.18 Mb

The bottom section of the interface is titled 'MinION Mk1B · MN21866' and contains two panels. The left panel shows 'Your run settings' with the following details:

- Experiment group: TestForReport
- Sample ID: no_sample
- Flow cell product code: FLO-MIN114
- Kit ID: SQK-LSK114
- Expansion pack: EXP-PBC096
- Current output directory: C:\data\...\TestForReport\...
- Basecall model: High-accuracy basecalling
- Minimum qscore: 9

The right panel shows performance metrics:

- Total run time: 21m 16s
- Remaining time: 44m
- Read count: 82.93 k
- Estimated bases: 97.8 Mb
- Basecalled bases: 101.18 Mb

Your run settings

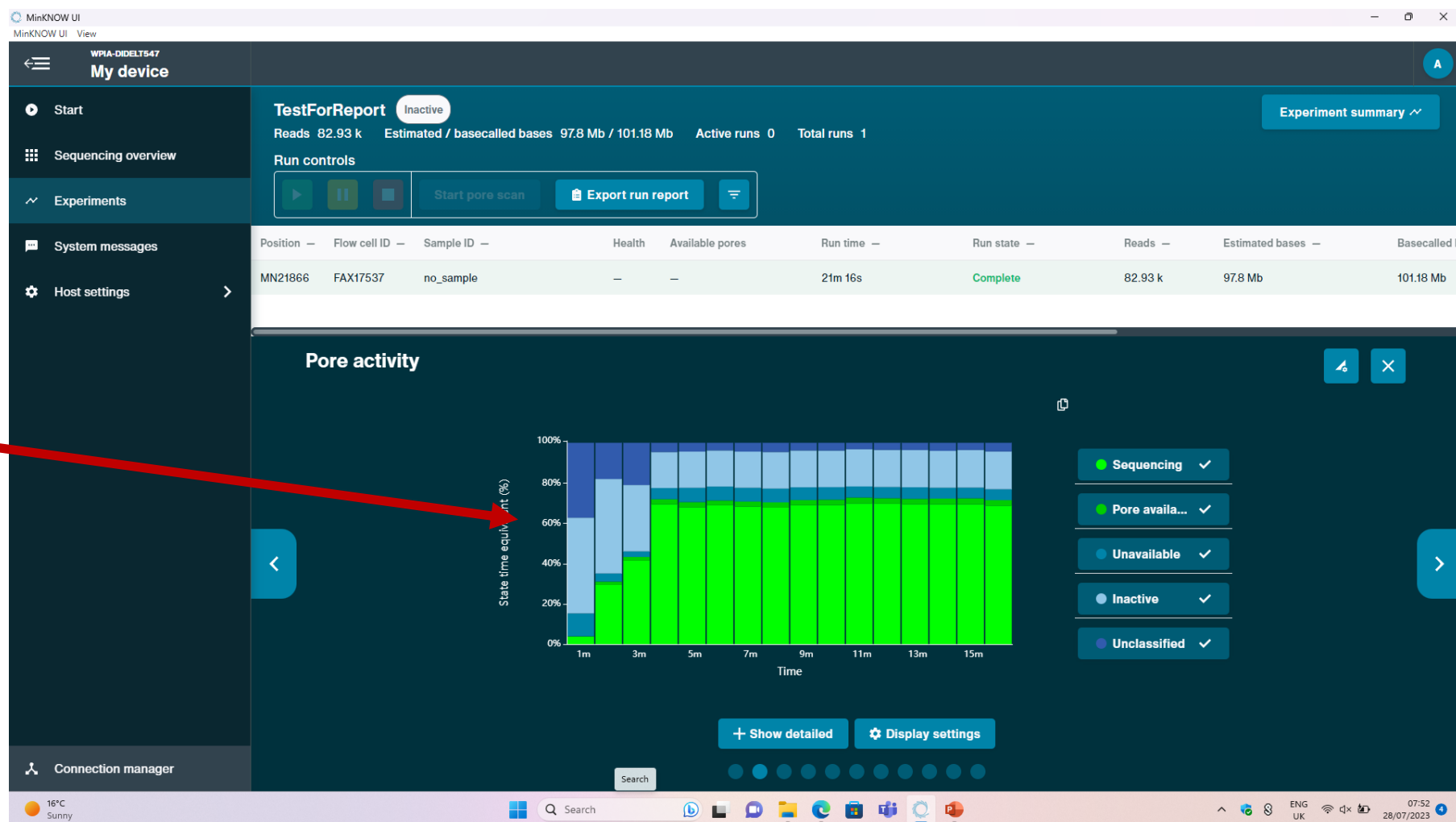
Confirm your run length (4 hours for routine stool testing by DDNS)

Click for more figures

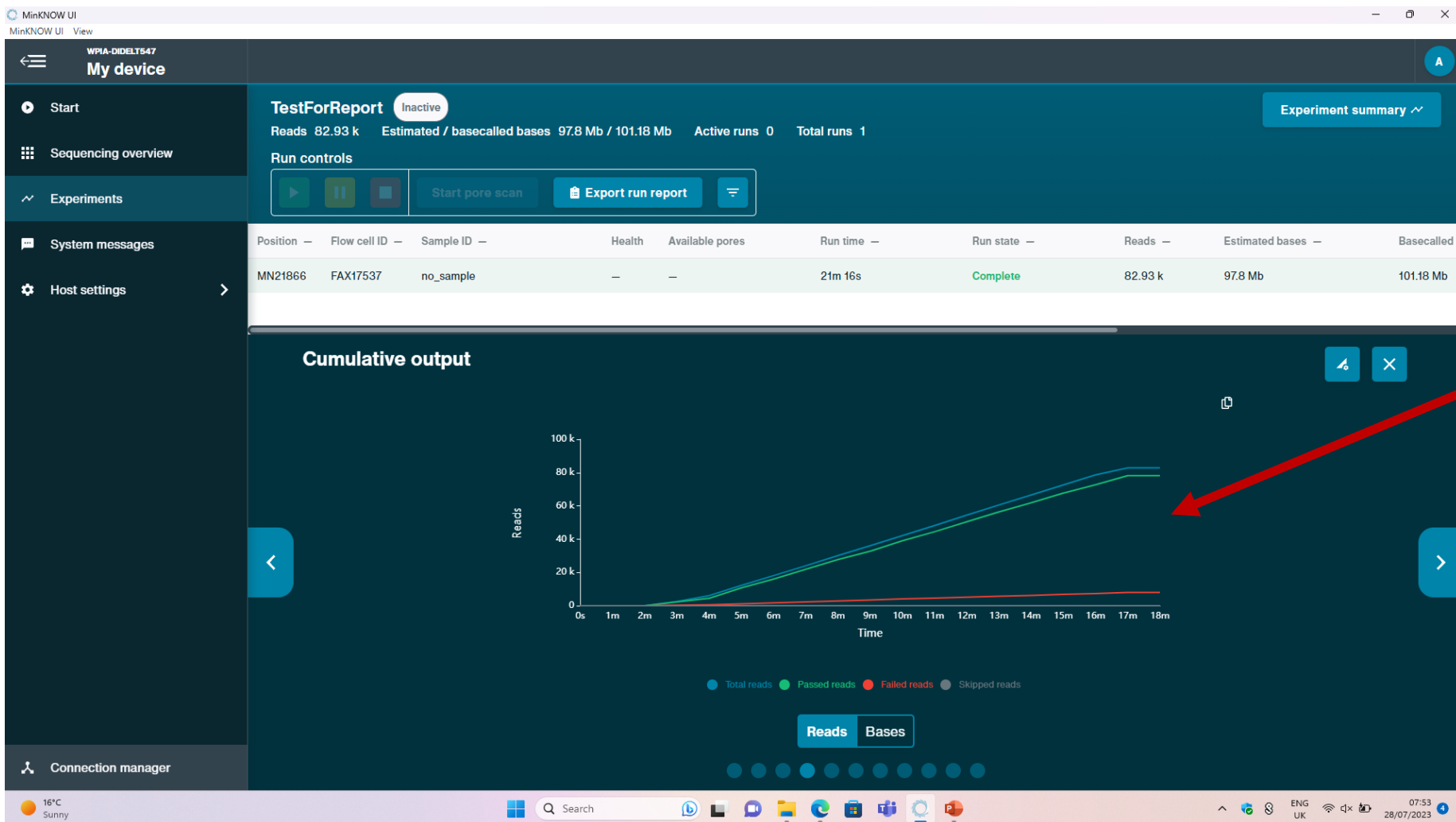
The total number of sequencing reads generated

4. Check the number of pores remaining

Check how many pores are remaining at the end of your run (and again after washing). Green pores are available to sequence.

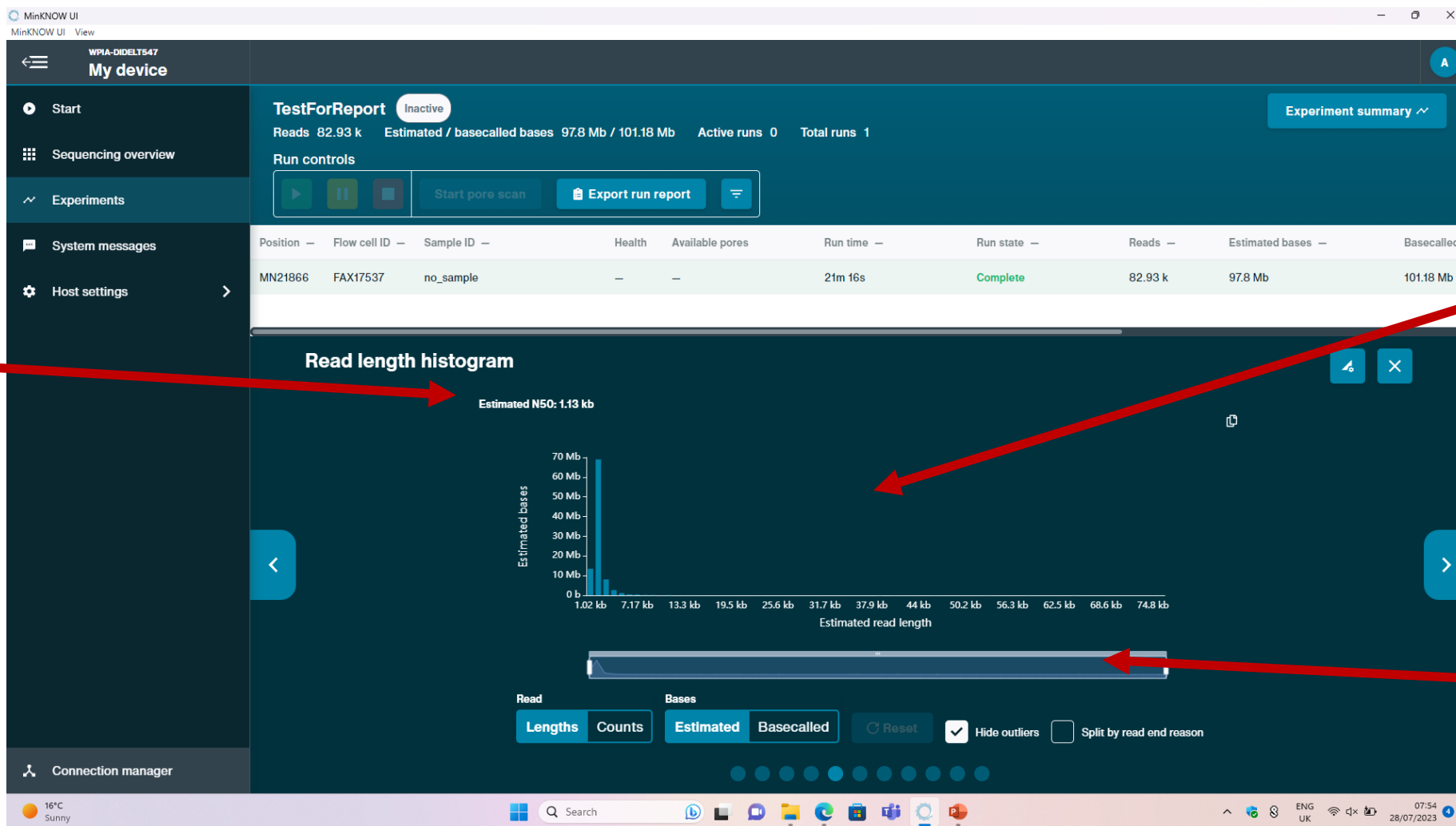


5. Check the quality of your data



Check the proportion of reads that are passing quality checks (green line). The vast majority of reads should pass.

6. Check the length of your sequences

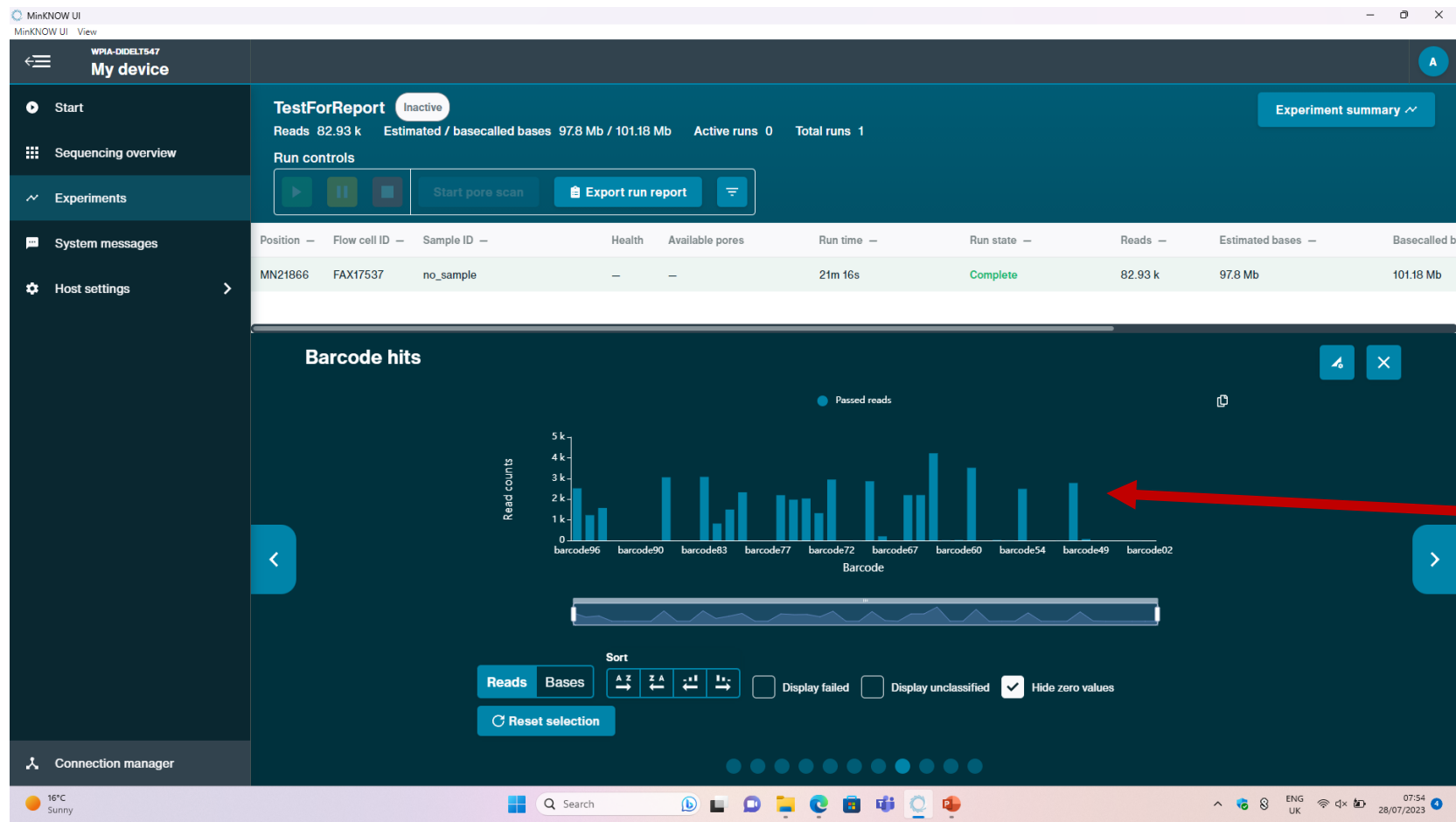


Check the average read length (N50)- for VP1 by DDNS this should be ~ 1.1 kb

This chart shows the distribution of read lengths

Use this slider to change the width of the chart

7. Check your data has been demultiplexed



If barcoding was turned on, this chart will show the data assigned to each barcode

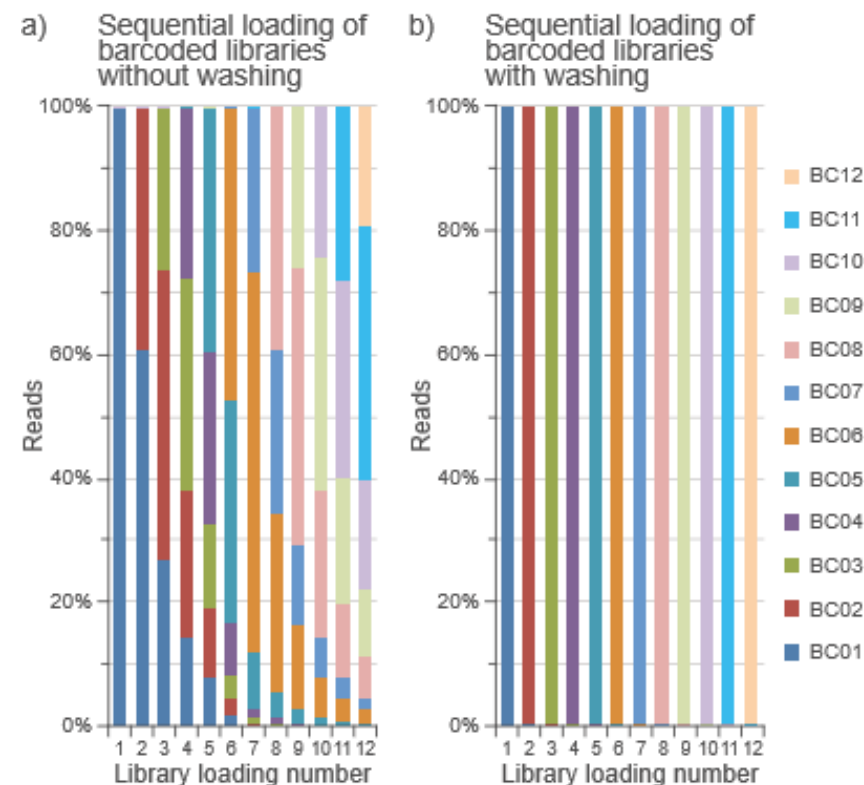
If your data has not been demultiplexed you can go to “Start”, “Analysis” and “Barcoding” in MinKNOW to perform the demultiplexing.

5. Flow cell washing

Lecture and discussion

Washing a flow cell after a sequencing run

- Wash kit (WSH004) contains
 - Wash mix (DNase I, WMX)
 - Wash diluent (enzyme buffer, DIL)
 - Storage buffer (S)
- DNase I removes DNA that remains in the pores after the run ends
 - Avoids contaminating the next run

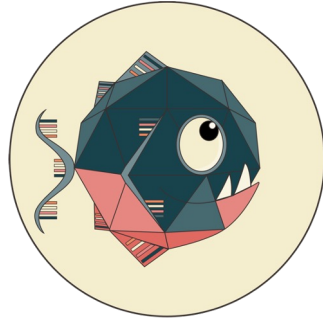


Protocol

- Combine 2ul of the DNase enzyme (WMX) with 398ul of the wash buffer (DIL)
- Remove any air bubbles under the priming port
- Put the 400ul wash solution onto the flow cell through the priming port with the SpotOn port closed
- Incubate at room temperature for 60 minutes

Protocol

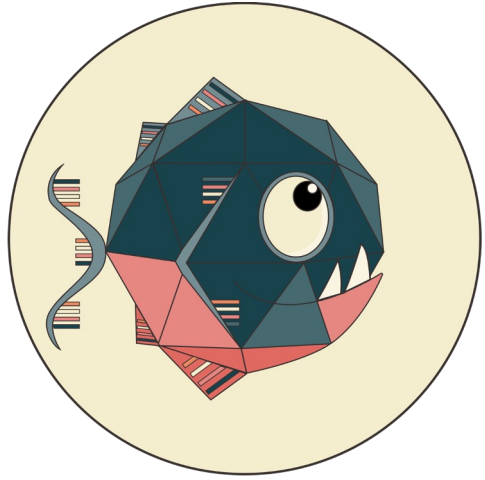
- For immediate use:
 - Flush the flow cell with Flush Buffer as described for flow cell priming
- For storage:
 - Add 500ul of the storage buffer (S) through the priming port, put the flow cell back into its packaging
 - Store in the fridge until the next use



6. Installation and use of piranha

Poliovirus investigation resource
automating nanopore haplotype analysis

Lecture and discussion



piranha

Poliovirus investigation resource automating nanopore haplotype analysis

Áine O'Toole^{1,*}, Rachel Colquhoun¹, Corey Ansley¹, Catherine Troman², Daniel Maloney¹, Zoe Vance¹, Joyce Akello², Erika Bujaki³, Manasi Majumdar³, Adnan Khurshid⁴, Yasir Arshad⁴, Muhammad Masroor Alam⁴, Javier Martin³, Alexander G Shaw², Nicholas C Grassly^{2,\$}, Andrew Rambaut^{1,\$}

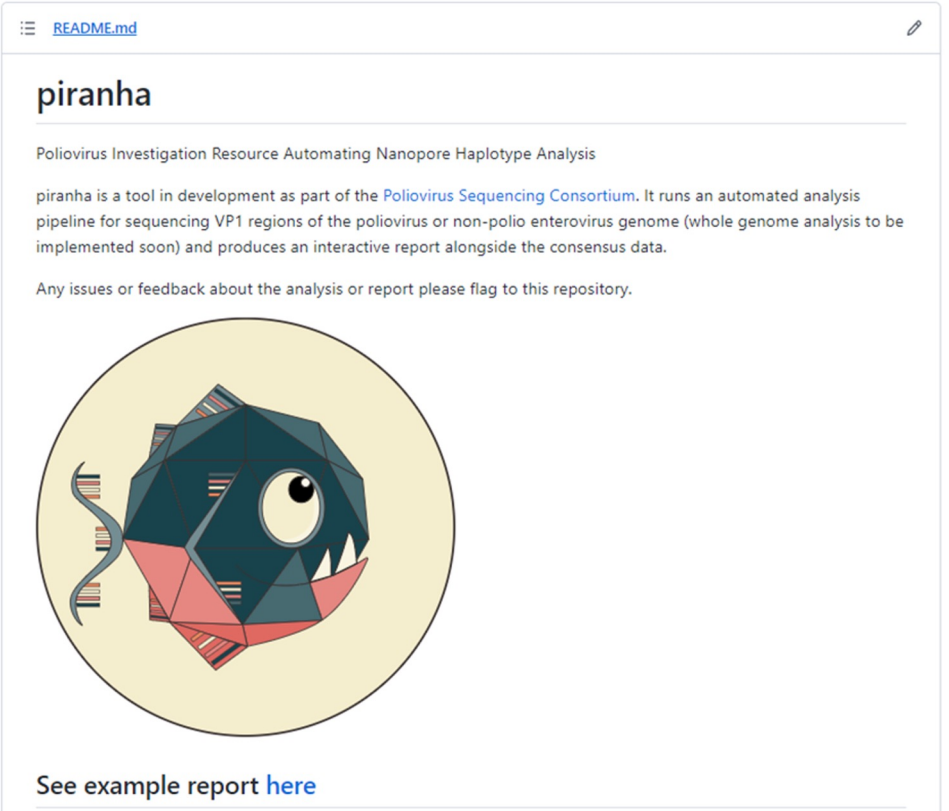
Affiliations

1. Institute of Ecology & Evolution, University of Edinburgh, United Kingdom
 2. Department of Infectious Disease Epidemiology, Imperial College London, United Kingdom
 3. Division of Virology, National Institute for Biological Standards and Control (NIBSC), Hertz, United Kingdom
 4. Department of Virology, National Institute for Health, Islamabad, Pakistan
-

PIRANHA - overview

- Takes demultiplexed fastq files, maps to reference sequences, filters by length, and produces consensus sequences and reports as output
- Requires command line
- Installation instructions and basic usage available on github:

<https://github.com/polio-nanopore/piranha>



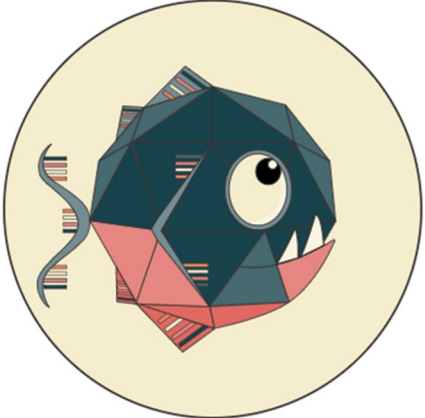
☰ [README.md](#) ✎

piranha

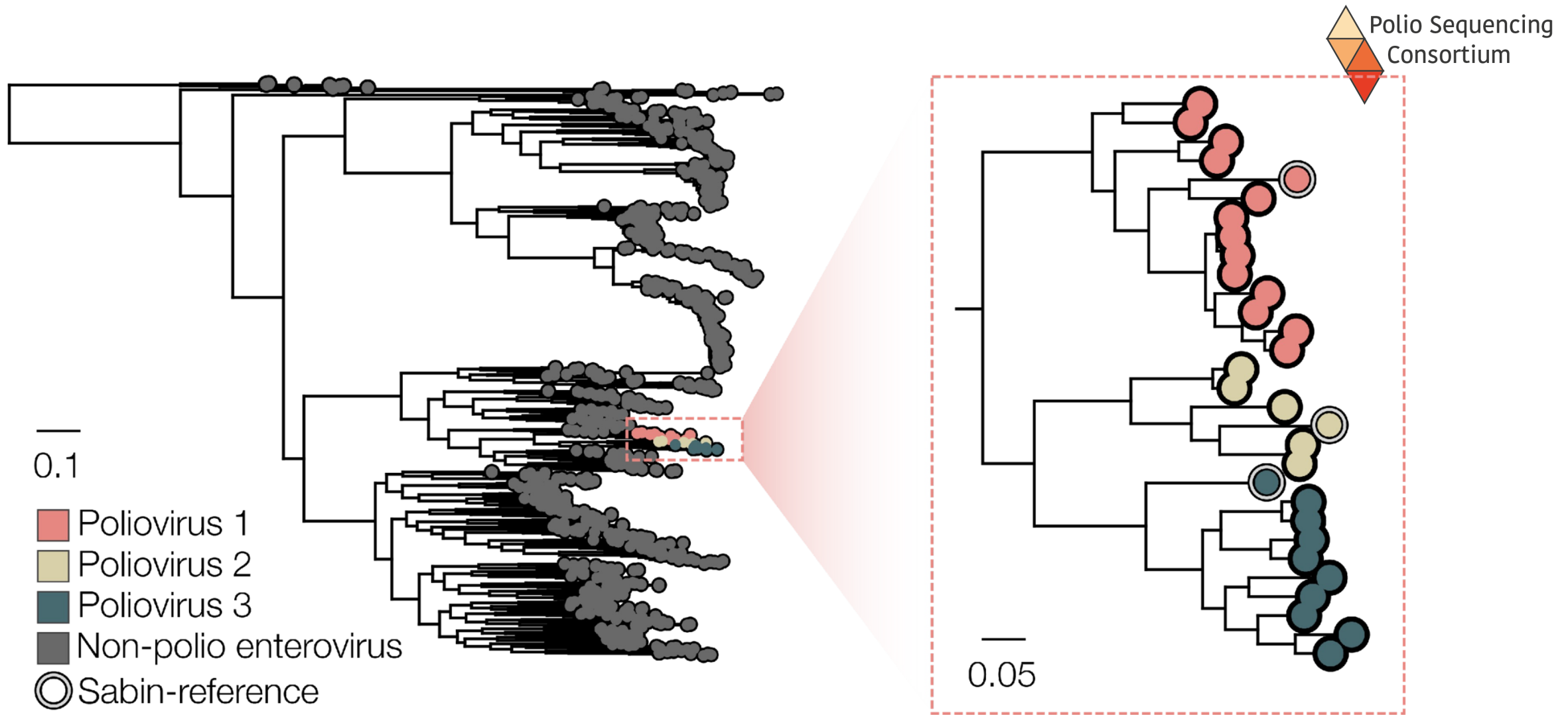
Poliovirus Investigation Resource Automating Nanopore Haplotype Analysis

piranha is a tool in development as part of the [Poliovirus Sequencing Consortium](#). It runs an automated analysis pipeline for sequencing VP1 regions of the poliovirus or non-polio enterovirus genome (whole genome analysis to be implemented soon) and produces an interactive report alongside the consensus data.

Any issues or feedback about the analysis or report please flag to this repository.

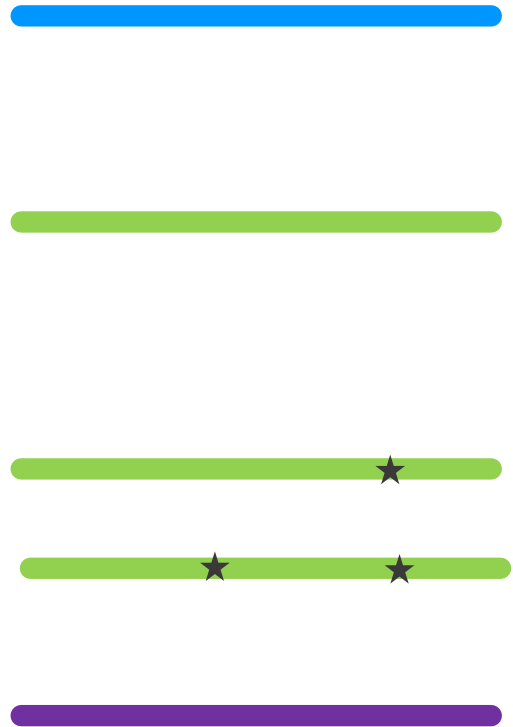


See example report [here](#)

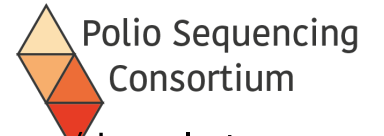
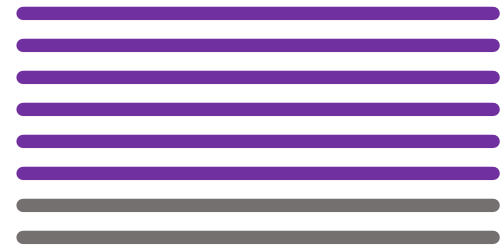
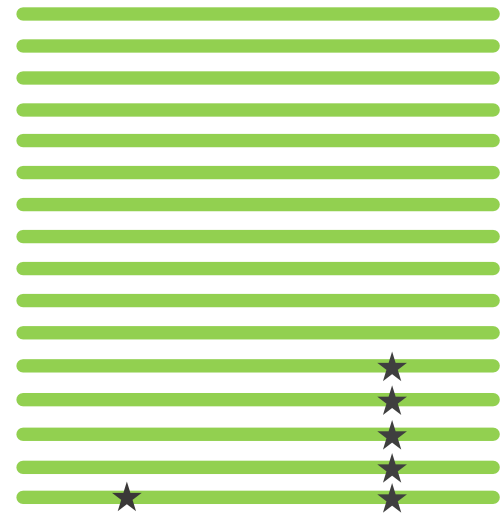
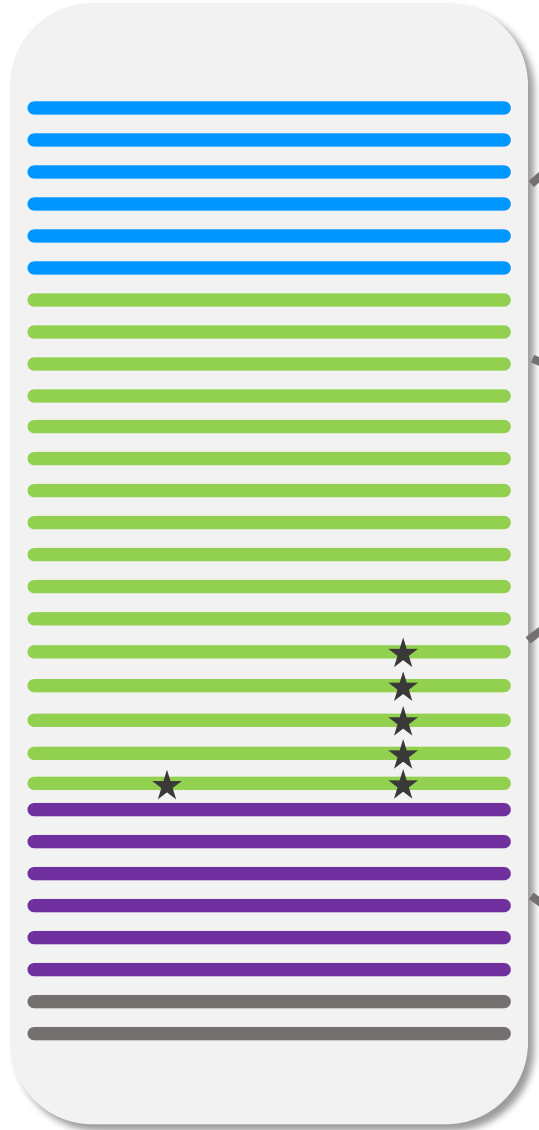


Background database: 959 VP1 sequences

Mixed Sample



Reference groups



Consensus/ haplotype



PiranhaGUI

- Uses *Docker* to allow running of PIRANHA on different operating system (e.g. Windows, Linux) and to avoid using the command line!
- Available on github here:
<https://github.com/polio-nanopore/piranha>
- Requires same input as piranha and produces the same output

PiranhaGUI Installation

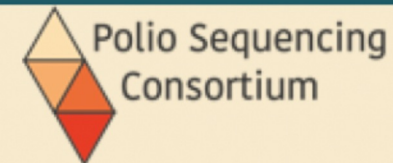
- Download from <https://github.com/polio-nanopore/piranha/releases>
- To install for windows, download: [PiranhaGUIv1.5.0_installer_windows.exe](#)
- Double click the downloaded file to install
- Open the PiranhaGUI when installed
- Click on the “Install Docker” button to go to the Docker website to download Docker
- Click on the “Install piranha” button to install piranha inside piranhaGUI.



Piranha

Polio Direct Detection by Nanopore Sequencing (DDNS)
analysis pipeline and reporting tool

About



Bill & Melinda Gates Foundation OPP1171890 and OPP1207299

Options

Docker software installed

Docker is free software used to install and run the analysis pipelines.

PIRANHA software installed

Piranha is the primary analysis pipeline for the DDNS polio detection platform.

An internet connection and a Docker install is required to install or update software

Continue

Running PiranhaGUI

- Select your barcode.csv file
- and your demultiplexed sequencing data
- Select an output folder for your results
- Click “Continue”



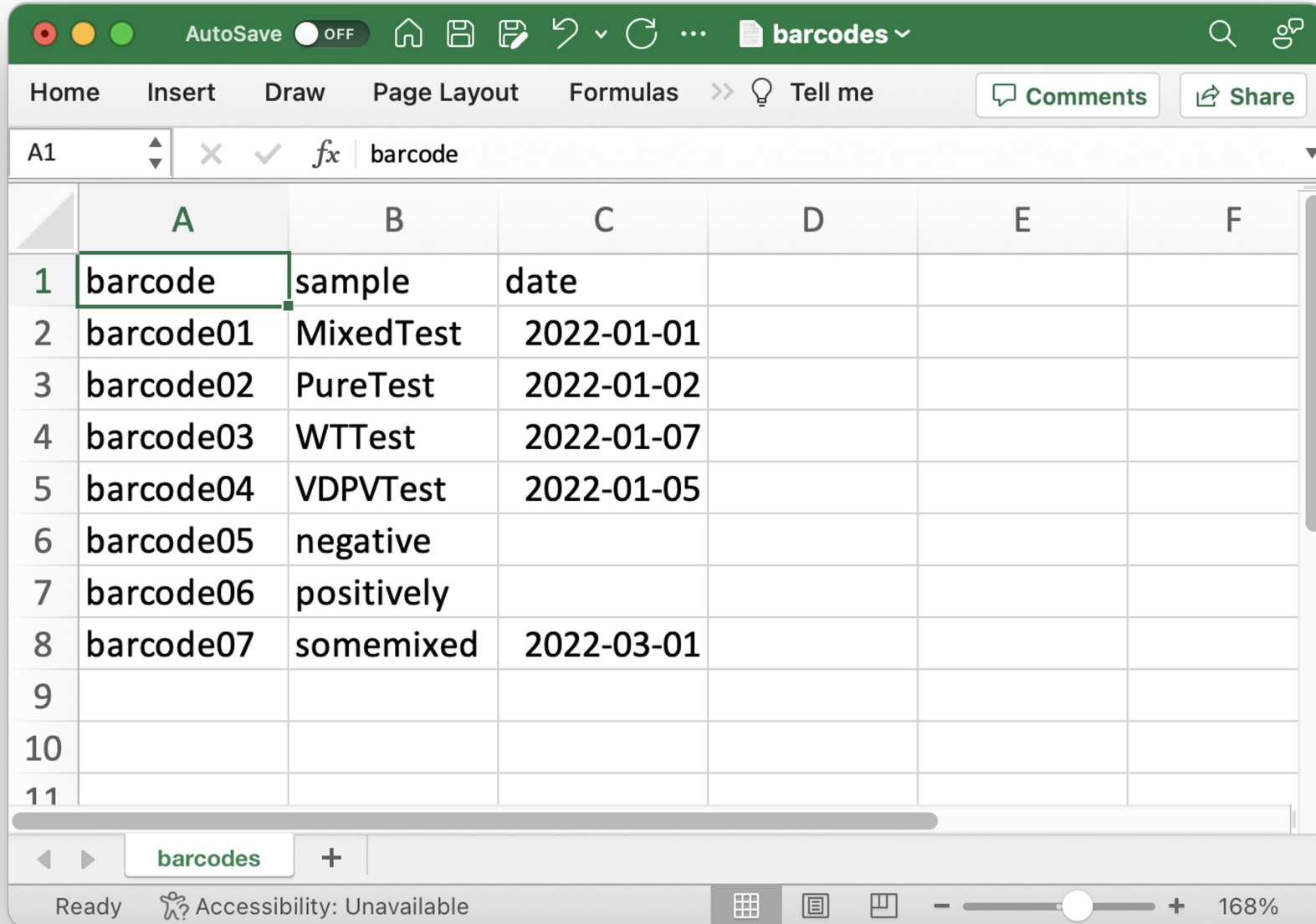
The screenshot shows the PiranhaGUI web interface. At the top, it says "Powered by ARTIFICE | ARTICnetwork: http://artic.network" and "Piranha v1.0.0". The main section is titled "Sequencing Run:" and contains three input fields, each with a "Select" button:

- Samples:** `work/piranha/test_data/barcodes.csv` (with a "View" button)
- MinKnow run:** `/piranha/test/pak_run/demultiplexed`
- Output Folder:** `ries/ARTICNetwork/piranha/test_data`

At the bottom, there are two buttons: "Run Options" and "Continue". A footer at the very bottom reads "Welcome Trust Award 206298/Z/17/Z".

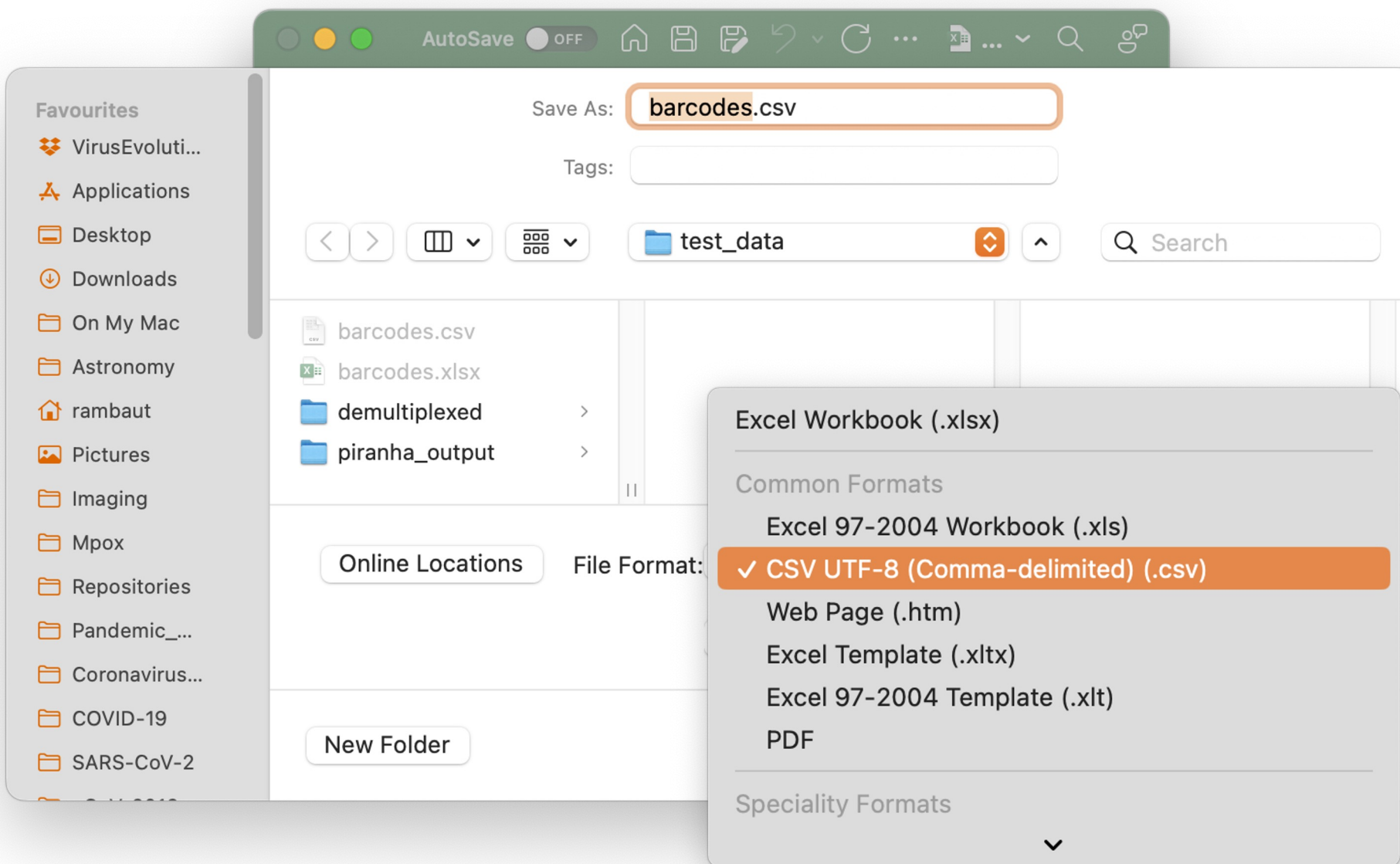
Red arrows from the instructions point to the "Select" buttons for the Samples, MinKnow run, and Output Folder fields, and to the "Continue" button.

Barcodes and Samples



The screenshot shows a Microsoft Excel spreadsheet with the following data:

	A	B	C	D	E	F
1	barcode	sample	date			
2	barcode01	MixedTest	2022-01-01			
3	barcode02	PureTest	2022-01-02			
4	barcode03	WTTest	2022-01-07			
5	barcode04	VDPVTest	2022-01-05			
6	barcode05	negative				
7	barcode06	positively				
8	barcode07	somemixed	2022-03-01			
9						
10						
11						



The image shows a macOS file save dialog window. The title bar includes window controls and an 'AutoSave' toggle set to 'OFF'. The 'Save As:' field contains the filename 'barcodes.csv'. Below it is a 'Tags:' field. The breadcrumb path shows the current location is 'test_data'. A search bar is on the right. The file list shows 'barcodes.csv', 'barcodes.xlsx', 'demultiplexed', and 'piranha_output'. At the bottom, there are buttons for 'Online Locations', 'New Folder', and 'File Format:'. The 'File Format:' dropdown menu is open, showing options under 'Common Formats' and 'Speciality Formats'. The 'CSV UTF-8 (Comma-delimited) (.csv)' option is selected and highlighted in orange.

AutoSave OFF

Save As: **barcodes.csv**

Tags:

< > [Grid View] [Compare View] test_data [Dropdown] Search

barcodes.csv
barcodes.xlsx
demultiplexed
piranha_output

Online Locations File Format:

New Folder

Excel Workbook (.xlsx)

Common Formats

- Excel 97-2004 Workbook (.xls)
- ✓ CSV UTF-8 (Comma-delimited) (.csv)**
- Web Page (.htm)
- Excel Template (.xltx)
- Excel 97-2004 Template (.xlt)
- PDF

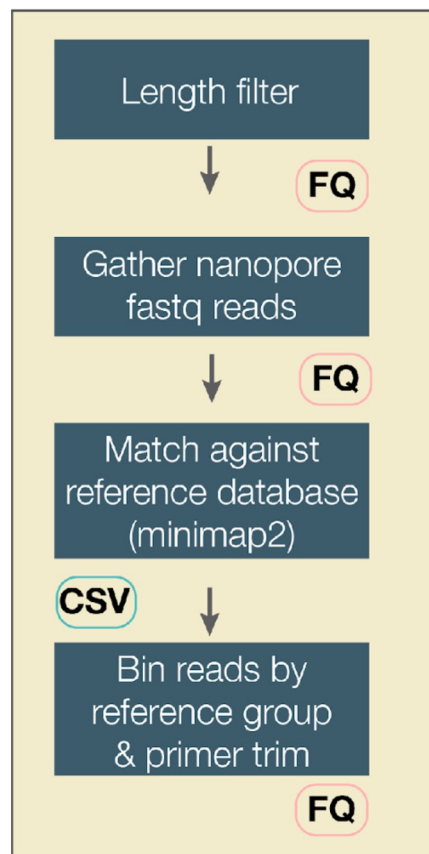
Speciality Formats

barcodes.csv — Edited

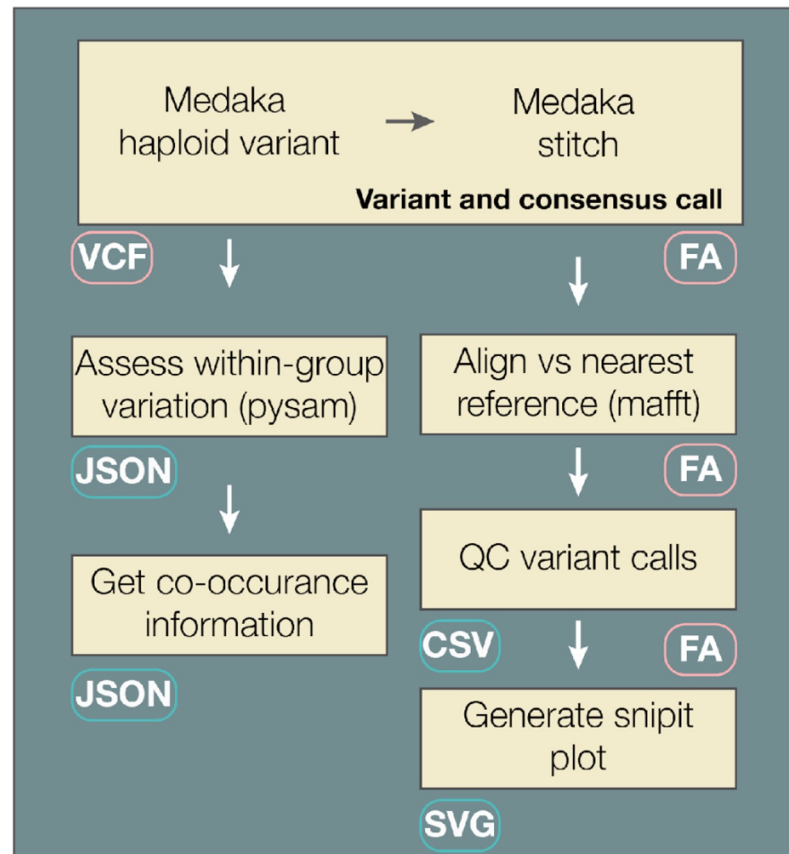
```
barcode, sample, date, well  
barcode01, MixedTest, 2022-01-01, A01  
barcode02, PureTest, 2022-01-02, B01  
barcode03, WTTest, 2022-01-07, C01  
barcode04, VDPVTest, 2022-01-05, D05  
barcode05, negative, , D06  
barcode06, positively, , D07  
barcode07, somemixed, 2022-03-01, H12
```


Workflow schema

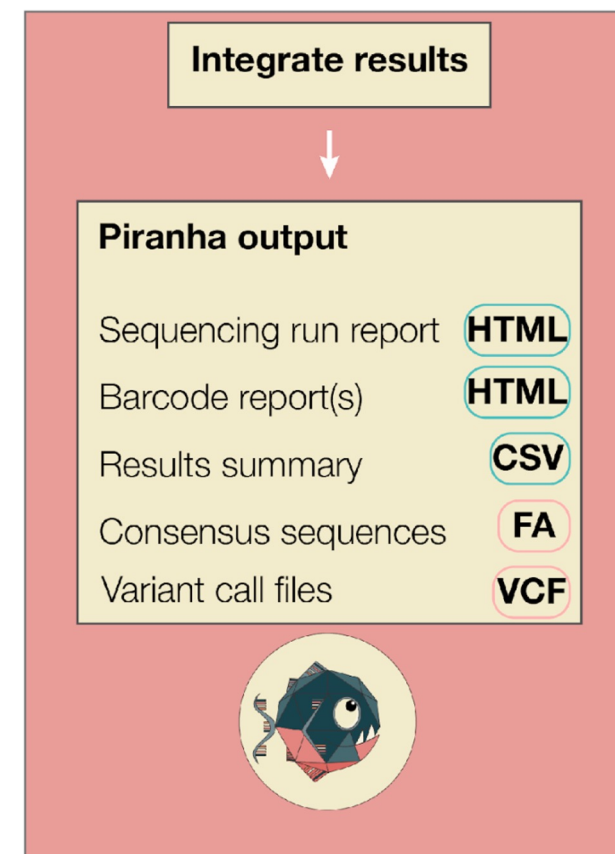
A. Barcode analysis



B. Reference group analysis



C. Results reporting



Example report

a.

Sample	Barcode	Sample Classification	Reference group	Number of mutations
ENV001	barcode01	Sabin-like	Sabin3-related	1
ENV001	barcode01	Sabin-like	Sabin2-related	2
ENV002	barcode02	VDPV	Sabin2-related	12
ENV003	barcode03	WPV1	WPV1	NA
ENV004	barcode04	VDPV	Sabin2-related	10

b.

Sample	Barcode	Sabin1-related	Sabin2-related	Sabin3-related	WPV1	WPV2	WPV3	NonPolioEV	unmapped
ENV001	barcode01	0	488	252	0	0	0	0	25
ENV002	barcode02	0	1100	0	0	1	0	35	12
ENV003	barcode03	0	0	0	339	0	0	0	0
ENV004	barcode04	0	0	0	138	0	0	0	0
negative	barcode05	0	0	0	0	0	0	1	10

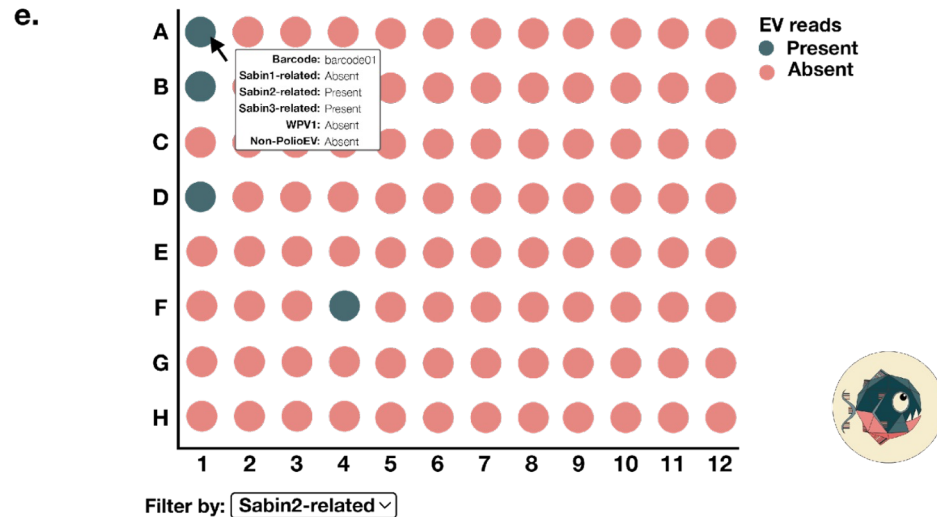
Showing 5 of 96 entries

c.

Identical Sequences	Sequence IDs
1	ENV001 barcode01 Sabin2-related 2 161:CT;427:GA ENV030 barcode30 Sabin2-related 2 161:CT;427:GA
2	ENV022 barcode22 Sabin3-related 1 17:CT ENV024 barcode24 Sabin3-related 1 17:CT

d.

Pass	Sample	Barcode	Sabin1-related	Sabin2-related	Sabin3-related	WPV1	WPV2	WPV3	NonPolioEV	unmapped
✓	negative	barcode05	0	0	0	0	0	0	1	10
	positive	barcode05	0	0	0	0	0	0	45	0



ENV001 report 2022-12-08

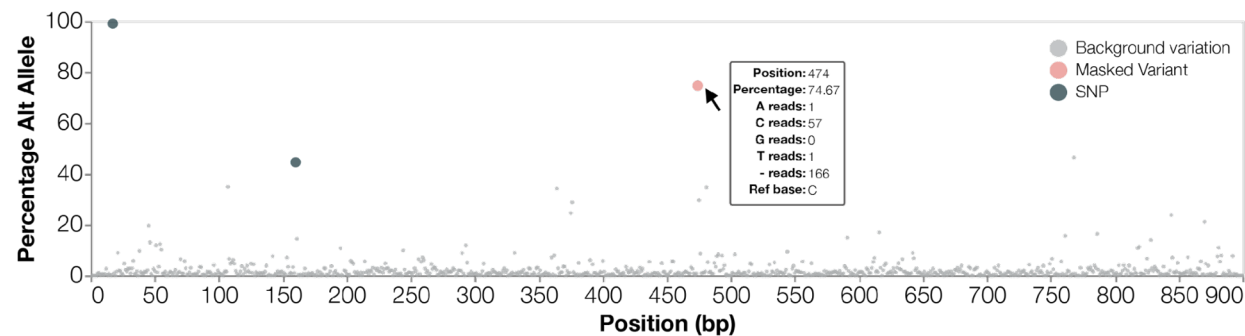
a.

Sample	Barcode	Reference group
ENV001	barcode01	Sabin3-related
ENV001	barcode01	Sabin2-related

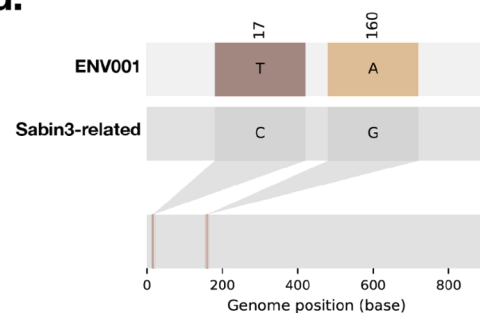
b. VP1 sequences

```
>ENV001|barcode01|Sabin3-related|Poliovirus3-Sabin_AY184221|2|17:CT;160:GA|date=2022-03-01
GGTATTGAAGATTTGATTTCTGAAGTTGCACAGGGCGCCCTAACTTTGTCACCTCCCGAAGCAACAGGATAGCTTACCTGATACTAAGGCCAG
TGGCCCGGCGCATTCCAAGGAGGTACCTGCACTCACTGCAGTCGAGACTGGAGCCACCAATCCTCTGACACCATCCGACACAGTTCAAA
CGCGCCACGTAGTCCAACGACGCAGCAGGTGAGTCCACAATAGAATCATTCTTCCGACGCGGGGCGTGCCTGCTATTATTGAGGTGG
ACAATGAACAACCAACCACCCGGGCACAGAACTATTTGCCATGTGGCGCATTACATACAAGATACAGTGCAGTTGCGCCGTAAGTTGGA
```

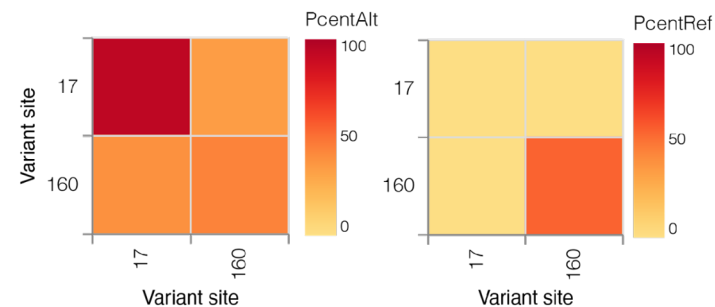
c.



d.

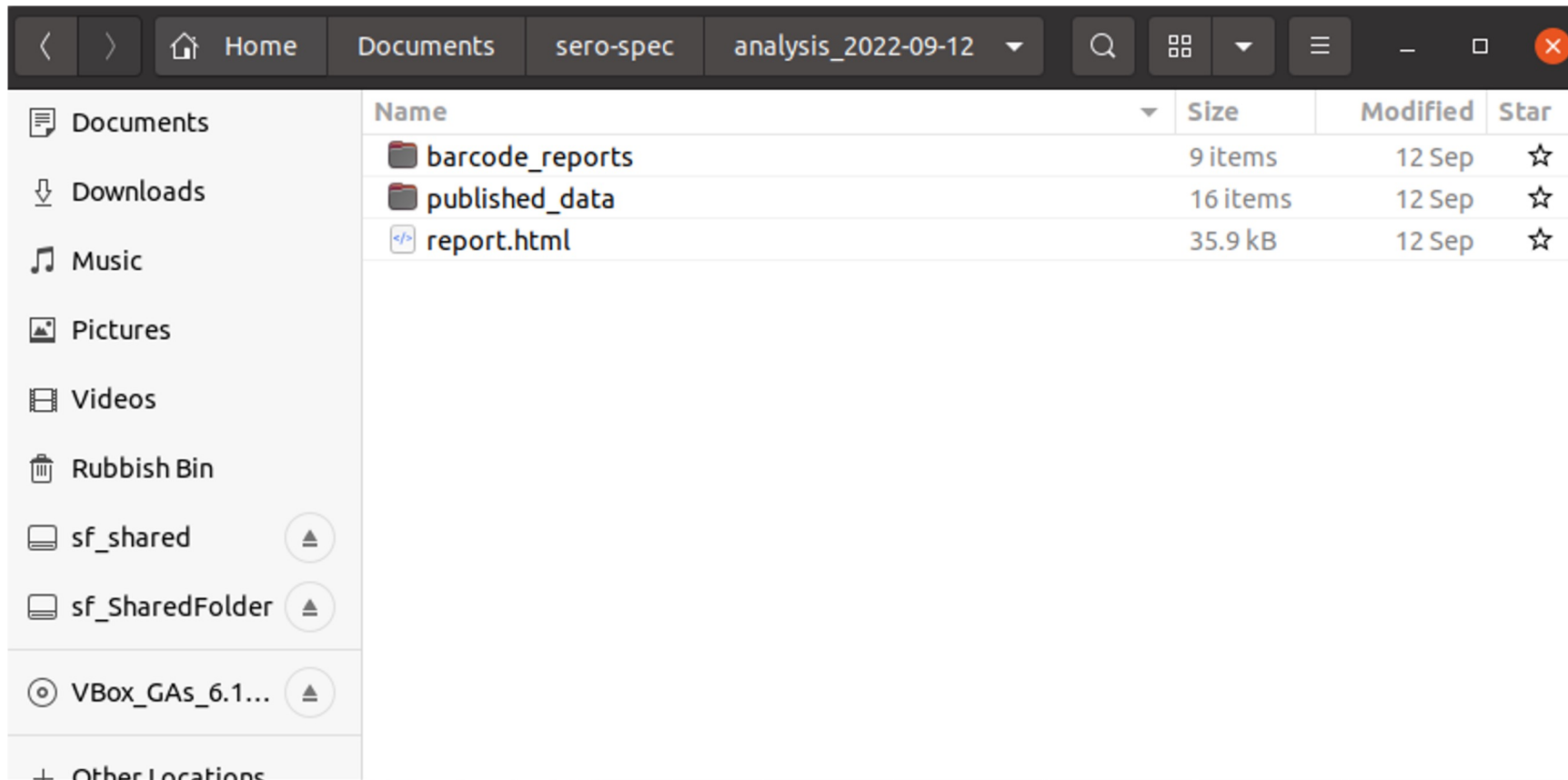


e.

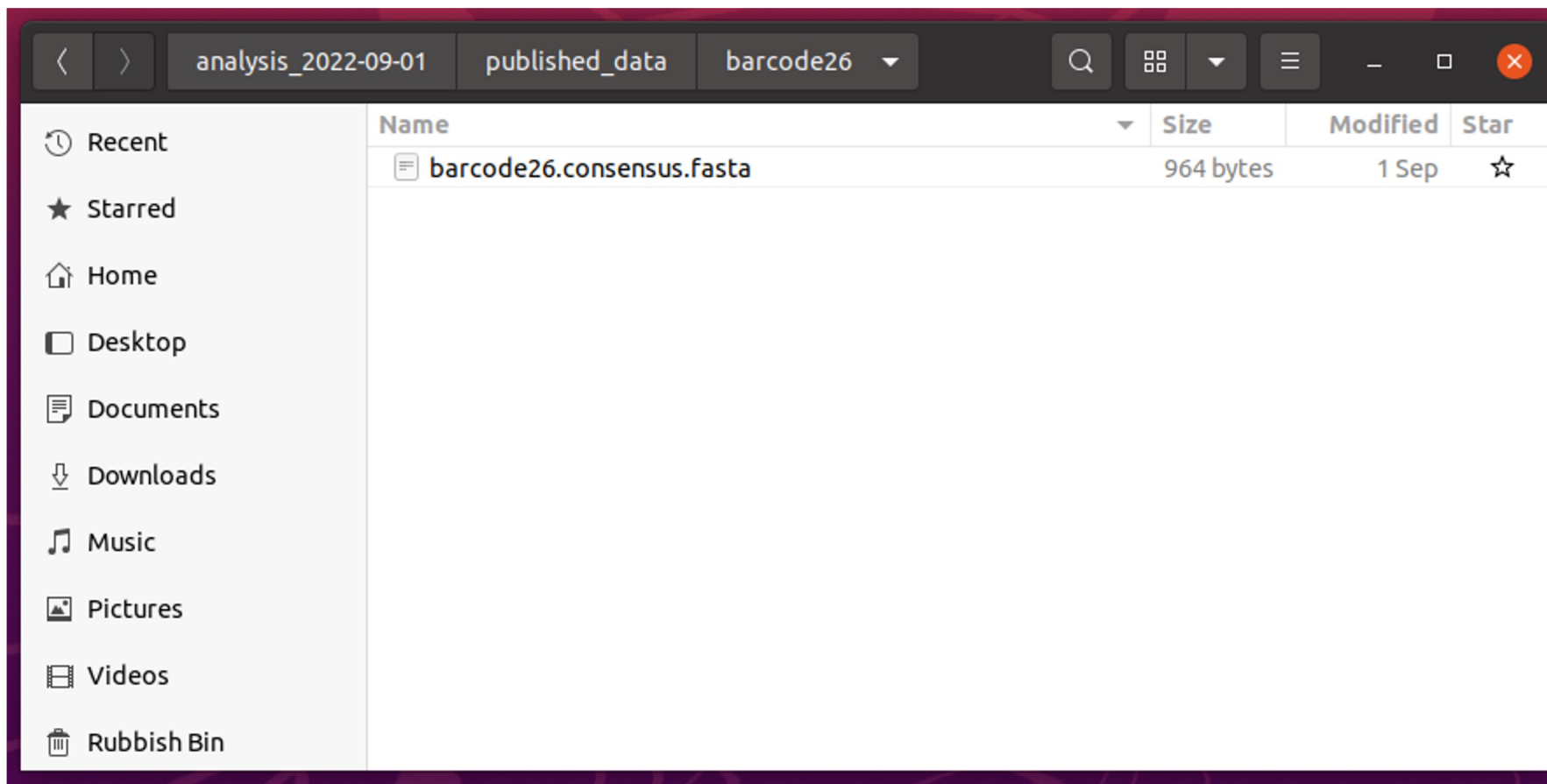


PIRANHA - output

- Outputs a summary report html, individual barcode reports, and consensus sequences as a fasta (both individual and in one file)



PIRANHA fasta output



Things to look out for

- Spaces
 - Don't put spaces in folder names
- Special characters
 - Try to stick to using only '-' and '_' when naming samples, files, and directories: e.g., `run_2023-07-26_barcodes.csv`
- Capital letters
 - Some bioinformatics software is case sensitive so be careful when writing commands and file paths
- Dates: always use the ISO 8601 date format: `2023-07-26`

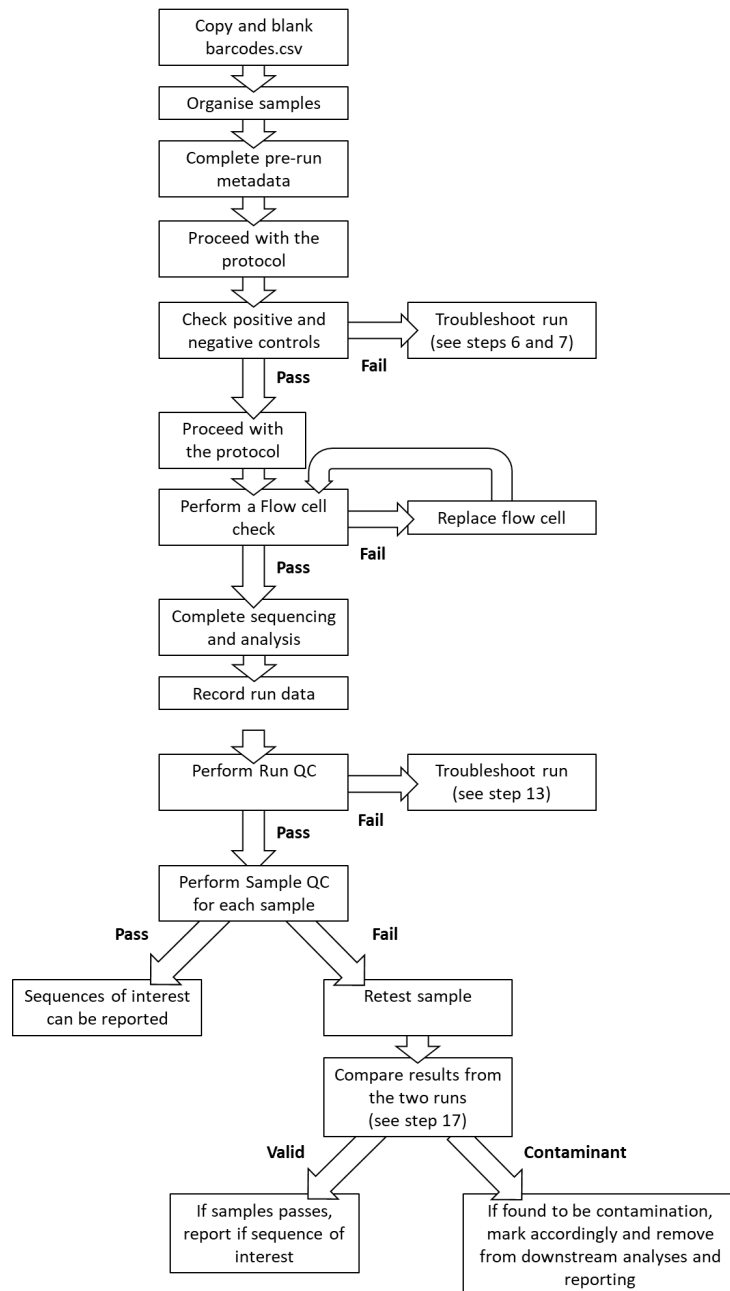
7. Quality Control and Reporting

Lecture and discussion

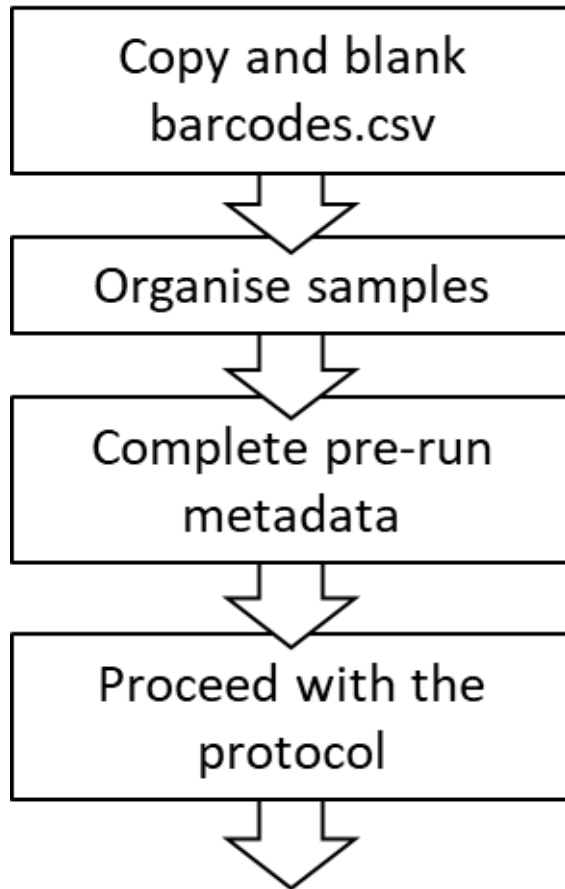
QC for NS of PV

- QC documents in progress:
SOP_Metadata and QC v4.0.docx

- Feedback very welcome!
- Spreadsheet contains an example output file with QC data. This is built by PIRANHA using your barcodes.csv file.
- Word document explains the procedure in detail
- This QC routine is designed for routine direct testing of poliovirus from stool samples where the majority of samples will be negative. Other applications may require other QC thresholds (e.g. minimum pore numbers, run duration)



Run setup



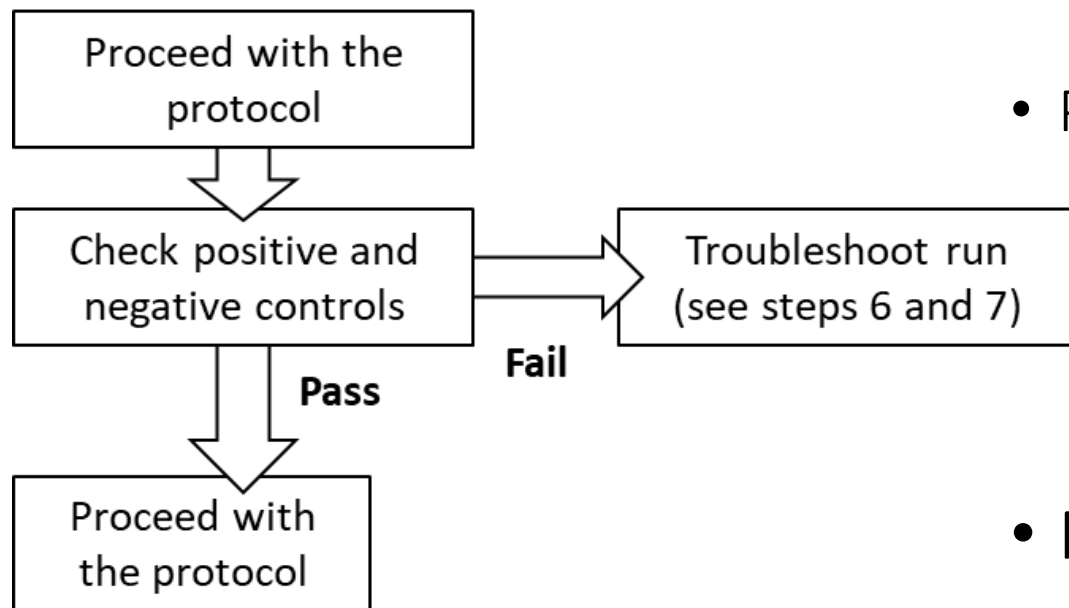
	A	B	C	D	E	F	G	H	I
1	sample	barcode	EPID	institute	...	Sabin1-related closest_refere	Sabin1-related num_reads	Sabin1-re	Sabin1-rel!
2	sample01	barcode01	ARA-HIG-TOR-22-01				0		
3	sample02	barcode02	ARA-HIG-TOR-22-01			Poliovirus1-Sabin_AY184219	3556	1	99.89
4	sample03	barcode03	ARA-KAN-ERF-22-01			Poliovirus1-Sabin_AY184219	2243	1	99.89
5	sample04	barcode04	ARA-KAN-ERF-22-01			Poliovirus1-Sabin_AY184219	1453	0	100
6	sample05	barcode05	ARA-HIG-TOR-22-02			Poliovirus1-Sabin_AY184219	1278	0	100

(optional pre-run metadata in columns D,E, and can inset more)

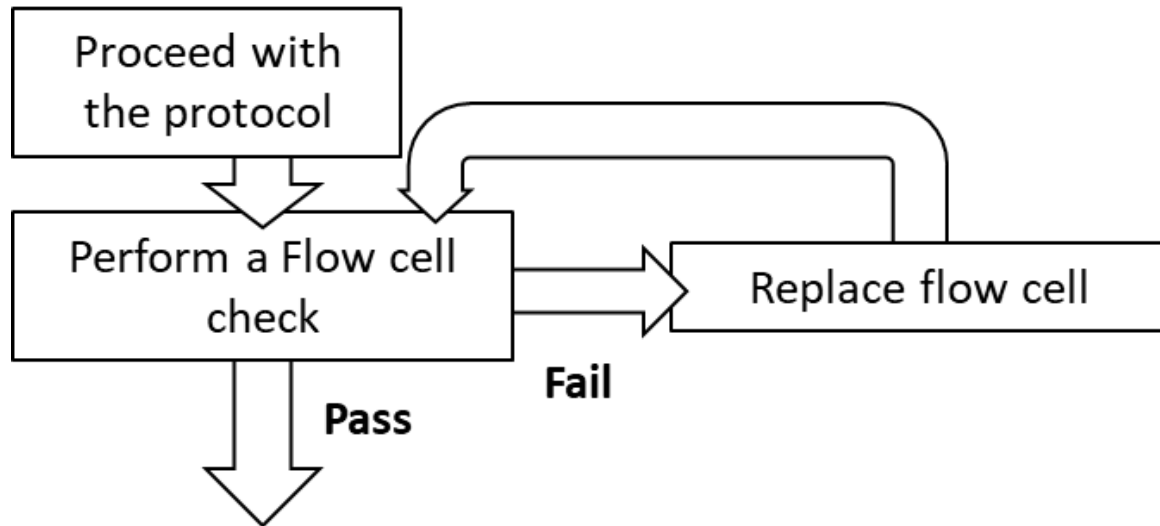
- First columns are the barcode.csv file- more metadata can now be added to this file
- Aim to separate cases from the same region if possible- these viruses may be similar, making QC difficult.

RNA extraction and nested RT-PCR QC

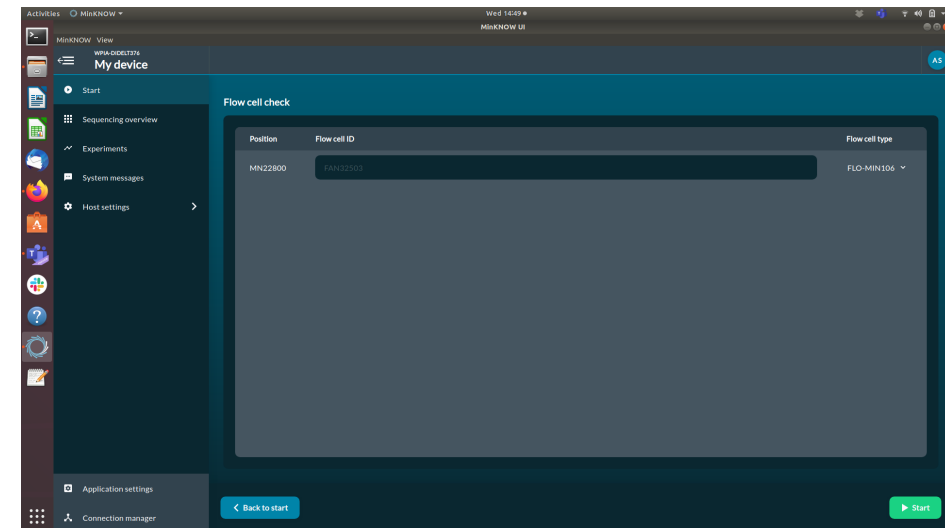
- Failed RT-PCR or failed PCR reactions
 - Note these in the barcodes.csv (they need to be repeated on later runs)
- Positive control
 - Two per day of RNA extractions performed (first and last extractions)
 - Should yield a ~1,400 bp amplicon on a gel or tapestation
- Negative control
 - Two per day of RNA extractions performed (first and last extractions)
 - Should have no amplicon.



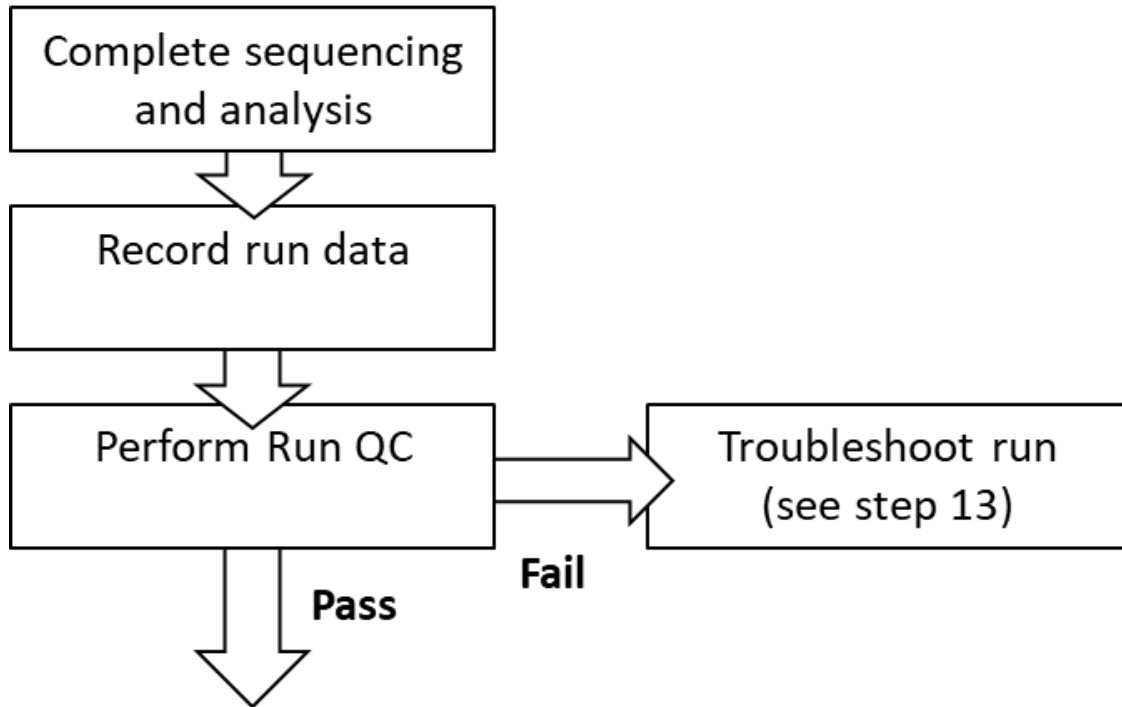
Flow cell check



- Should be > 400 pores available for routine direct testing from stool.
- Record flow cell ID- can be useful to identify contamination between runs.

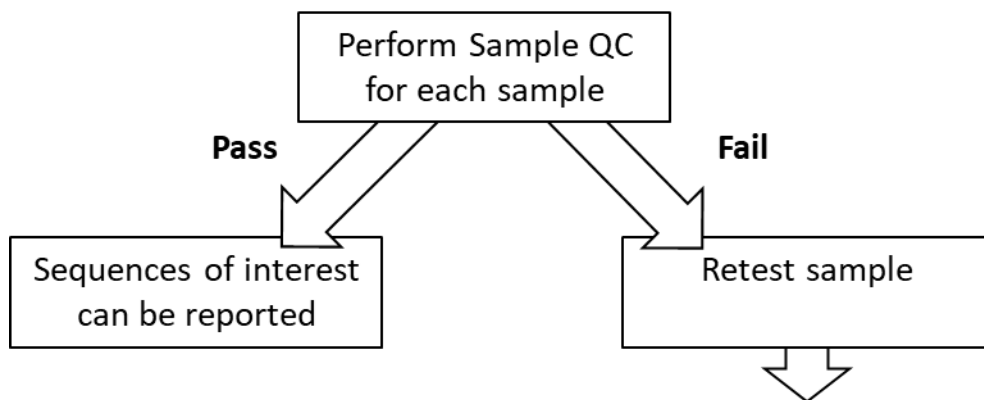


Sequencing Run QC



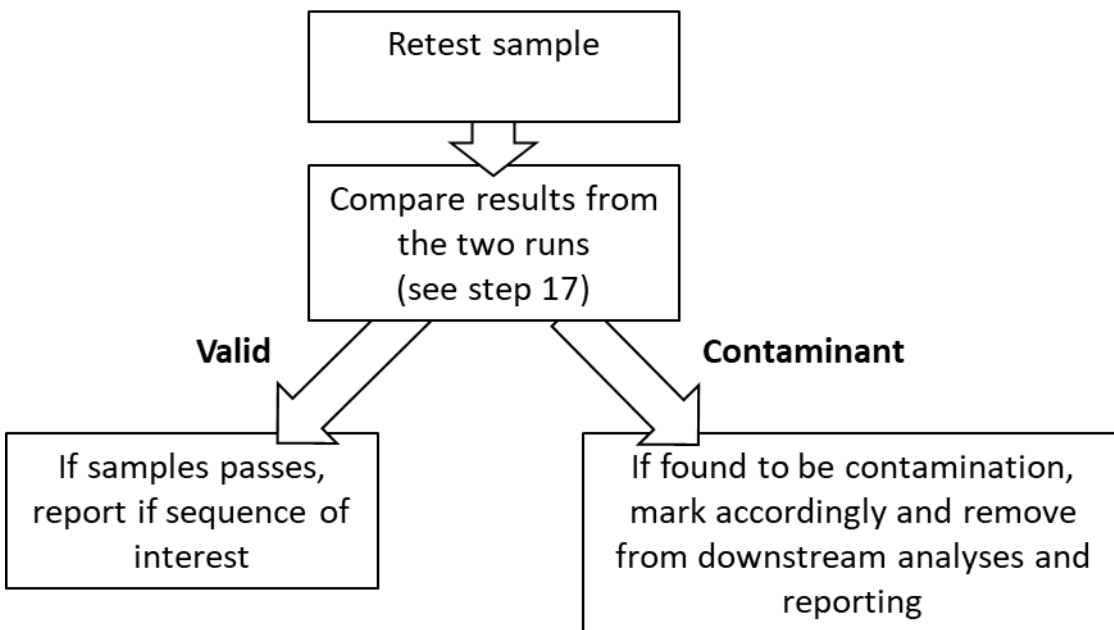
- PIRANHA will add sequencing and QC data to the barcodes.csv and save the file as a report.
- Check run QC
 - Did Positive control have > 500 reads mapping to A20?
 - Does the Negative control have <50 reads mapping to PV or NPEVs?
 - Did the run continue to completion and with at least 400 pores available (excluding shifts between pore groups)

Sample QC



- If run passes, check samples where PV was detected. For VDPVs:
 - Samples with the same EPID (i.e. from the same case) should be no more than 2 nucleotides different from each other over VP1.
 - Samples adjacent to each other on the plate (on either axis) with different EPIDs should differ by >1 nucleotides over VP1.
 - Samples with <1000 reads should be >1 nucleotides different over VP1 from sequencing reads from previous runs that were generated using the same barcode and the same flow cell ID.

Sample retesting



- Try to avoid using the same layout if checking a group of samples.
- Flag as a QC-check on the new sequencing run.
- Annotate the original run depending on the result- can report the sample if it passes the repeat.
- Talk to the appropriate person to arrange the retesting

End