minoTour - A real time LIMS platform for Oxford Nanopore data

Rory Munro • Roberto Santos • Alexander Payne • Matthew Loose
✉ rory.munro@nottingham.ac.uk ✉ matthew.loose@nottingham.ac.uk
School of Life Sciences, University of Nottingham

1. Abstract

minoTour is a web-based, real-time laboratory information management system (LIMS) for ONT sequencers, specifically MinION and GridION. Built using Python3 and Django; it can be hosted locally (offline) on a laptop via Docker, or remotely (online) on a server and accessed remotely. The client for monitoring and uploading data to minoTour, minFQ, can easily be installed via PyPI. minoTour can monitor the activity of a sequencer in real time, presenting a break down and analysis of live sequencing metrics and integrates with ONT's MinKNOW, allowing remote control and monitoring of sequencing.
Analysis, such as run summaries, alignment and metagenomics can be completed in real-time or after sequencing. minoTour is open source and freely available at **our github**, with community development strongly encouraged.

2. Architecture

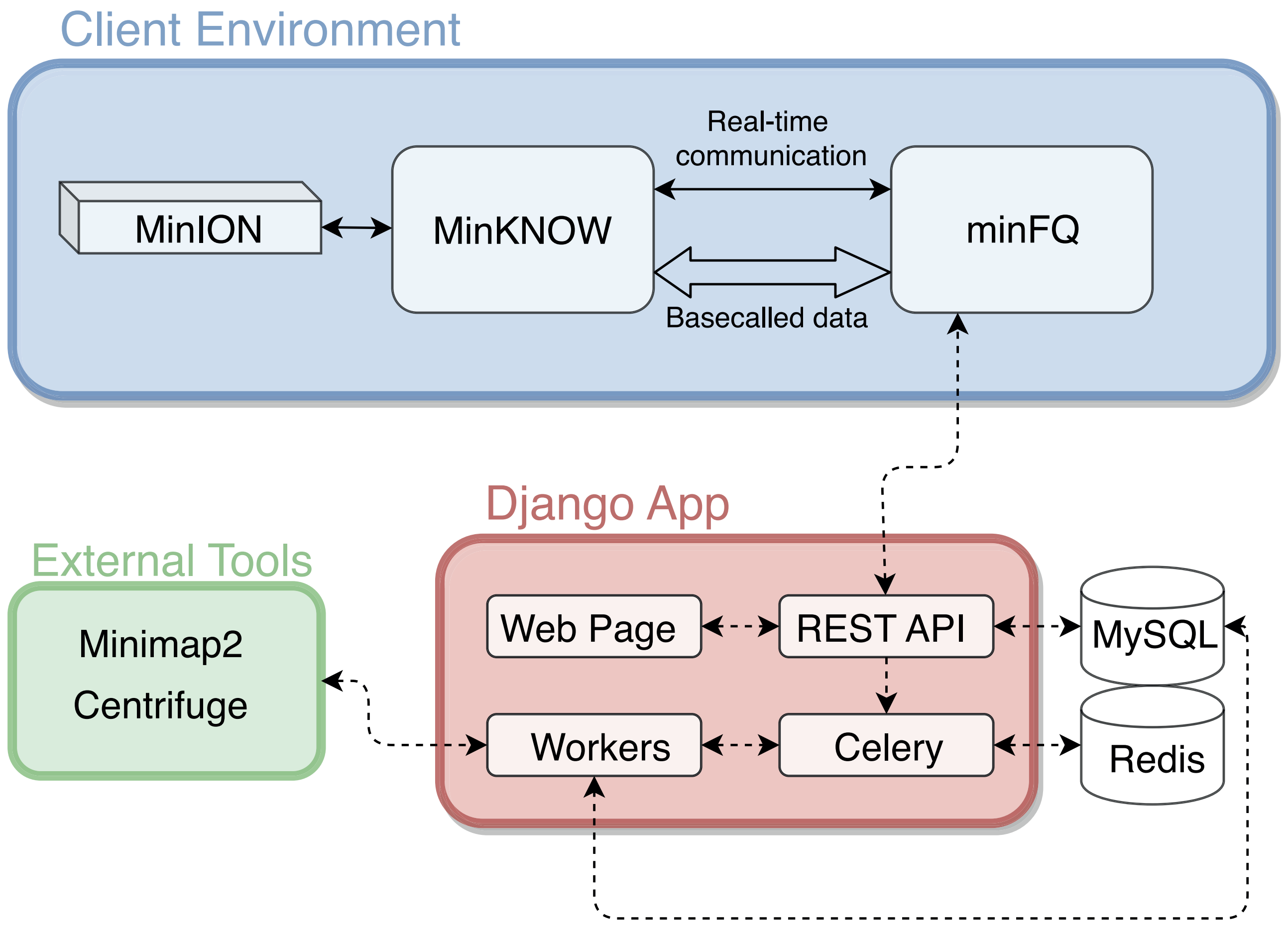

Client Environment
MinION
MinKNOW
Real-time communication
minFQ
Basecalled data
Django App
External Tools
Minimap2
Centrifuge
Web Page
REST API
MySQL
Workers
Celery
Redis

Figure 1: Application architecture - Different application environments are shown above, data flow is marked by arrows.

MinoTour data flow throughout the application is shown in Figure 1. MinKNOW and minFQ are in the client environment, where the sequencer is running. minFQ sends data from the client environment to the django app, which is minoTour itself. The data sent is saved in the mySQL database for historical browsing. minoTour can send commands and messages back to the client environment. The Django app retrieves data from mySQL and parses it, before displaying it to the site page.

3. MinKNOW and live metric monitoring

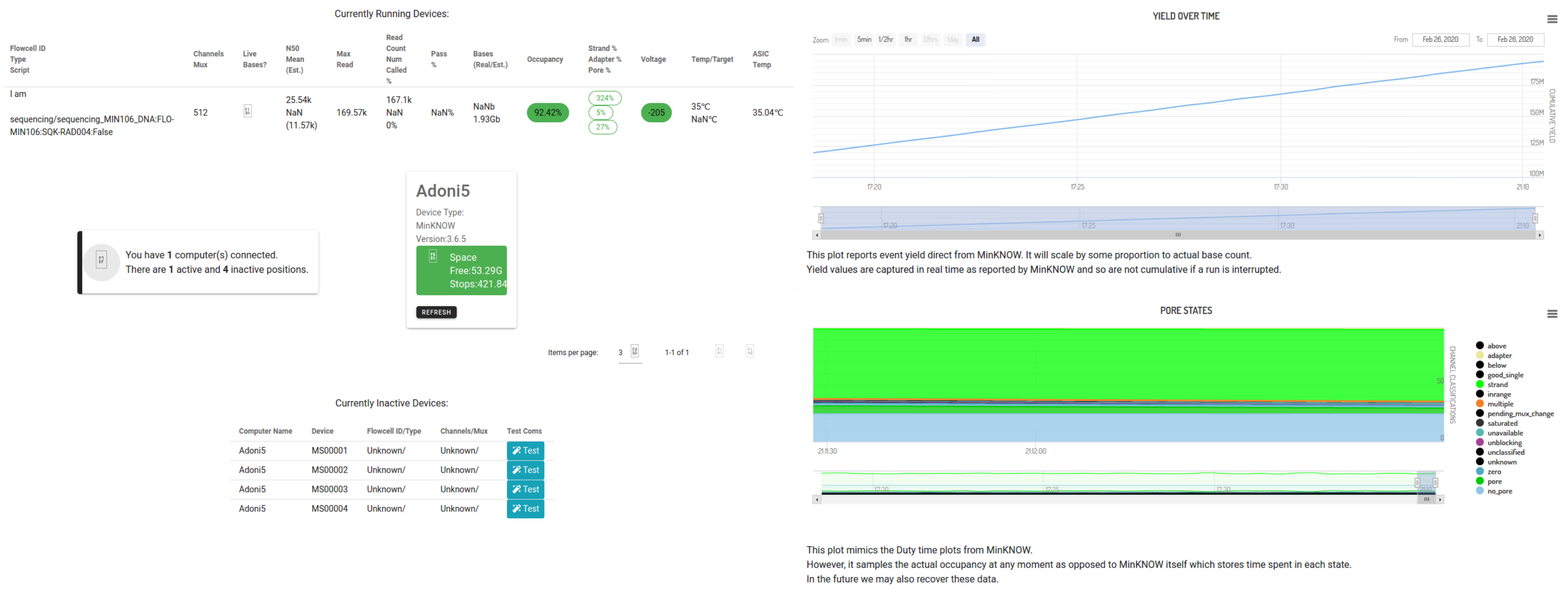


Figure 2: The MinKNOW control interface and example metrics visualisations.

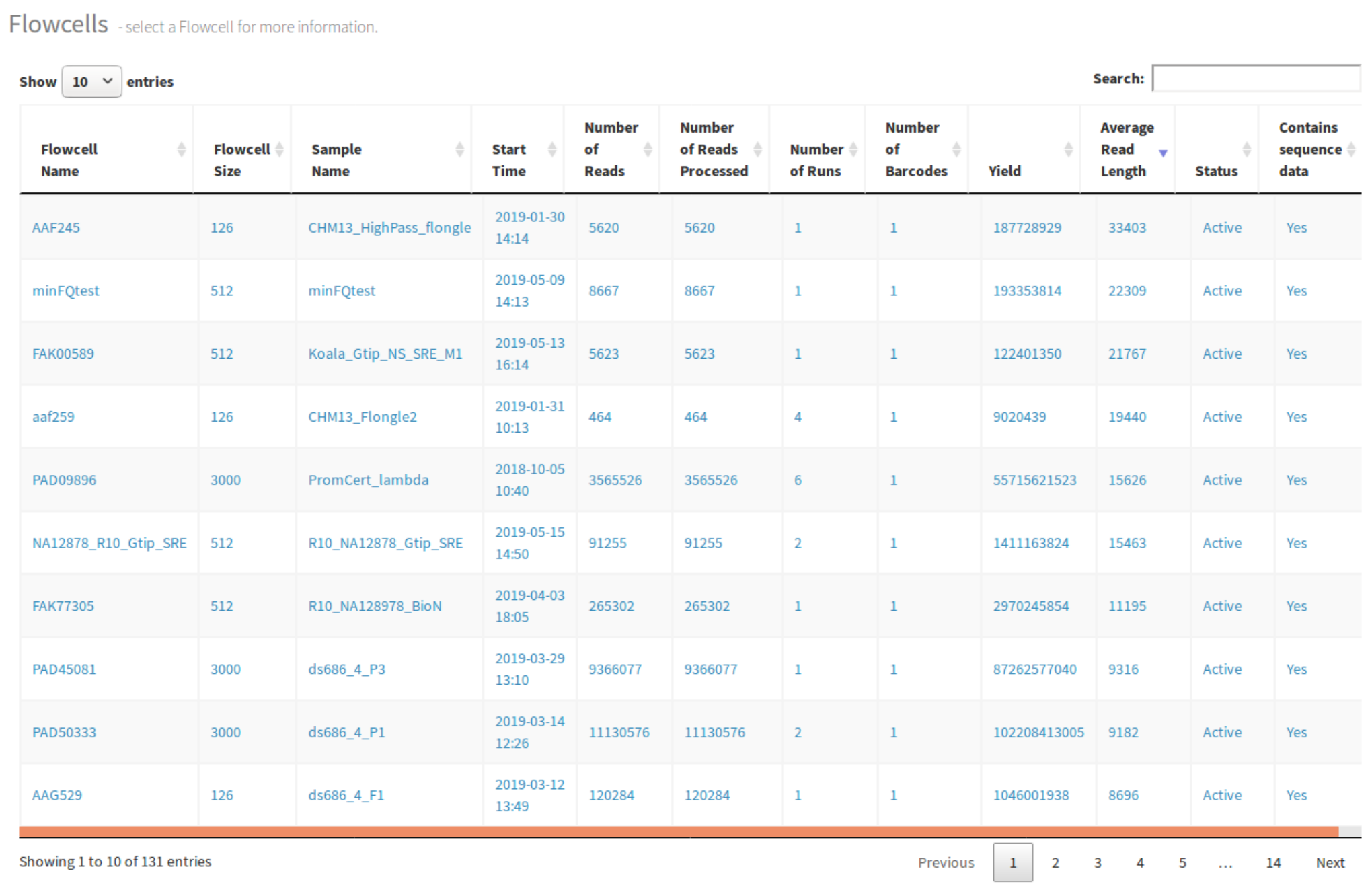
minoTour can capture live device metrics during sequencing from MinKNOW. This allows users to track how their device is performing, and keep a record of how the sequencing run went. minoTour can also remotely stop runs. Example metric charts are shown in Figure 2.

4. Real-time artic pipeline

Figure 3: Artic analysis - An example of Artic visualisations is shown above.

In light of the ongoing SARS-CoV-2 pandemic, we incorporated a real time Artic nCoV-2019 based analysis pipeline into minoTour. minoTour tracks SARS-CoV-2 coverage per barcode in real time (Figure 3), then automatically fires an asynchronous Artic nCoV-2019 pipeline once a high enough coverage has been reached. The user can then download the concensus FASTA file upon completion of the pipeline. Further steps are being integrated into this pipeline currently.

5. Flowcell management


Flowcells - select a Flowcell for more information.

Show 10 entries Search:

Flowcell Name	Flowcell Size	Sample Name	Start Time	Number of Reads	Number of Reads Processed	Number of Runs	Number of Barcodes	Yield	Average Read Length	Status	Contains sequence data
AAF245	126	CHM13_HighPass_flongle	2019-01-30 14:14	5620	5620	1	1	187728929	33403	Active	Yes
minFQtest	512	minFQtest	2019-05-09 14:13	8667	8667	1	1	193353814	22309	Active	Yes
FAK00589	512	Koala_Gtip_NS_SRE_M1	2019-05-13 16:14	5623	5623	1	1	122401350	21767	Active	Yes
aaf259	126	CHM13_Flongle2	2019-01-31 10:13	464	464	4	1	9020439	19440	Active	Yes
PAD09896	3000	PromCert_lambda	2018-10-05 10:40	3565526	3565526	6	1	55715621523	15626	Active	Yes
NA12878_R10_Gtip_SRE	512	R10_NA12878_Gtip_SRE	2019-05-15 14:50	91255	91255	2	1	1411163824	15463	Active	Yes
FAK77305	512	R10_NA128978_BioN	2019-04-03 18:05	265302	265302	1	1	2970245854	11195	Active	Yes
PAD45081	3000	ds686_4_P3	2019-03-29 13:10	9366077	9366077	1	1	87262577040	9316	Active	Yes
PAD50333	3000	ds686_4_P1	2019-03-14 12:26	11130576	11130576	2	1	102208413005	9182	Active	Yes
AAG529	126	ds686_4_F1	2019-03-12 13:49	120284	120284	1	1	1046001938	8696	Active	Yes

Showing 1 to 10 of 131 entries

Previous 1 2 3 4 5 … 14 Next

Figure 4: Flowcell management - the flowcell name, sample name and number of reads, alongside other metrics, are shown here in an interactive table.

minoTour can manage flowcells, allowing users to keep an archive of all the flowcells that have been used. It is not necessary to upload base-called sequence data for this functionality. Figure 4 shows minoTours flowcell tracking table.

6. Base-called data

minoTour can show users information about base-called reads produced during a run. This allows useful insights into sequencing runs all in one place. It can also be used to archive metrics about a run. Example metric visualisations for base-calling are shown in Figure 5.

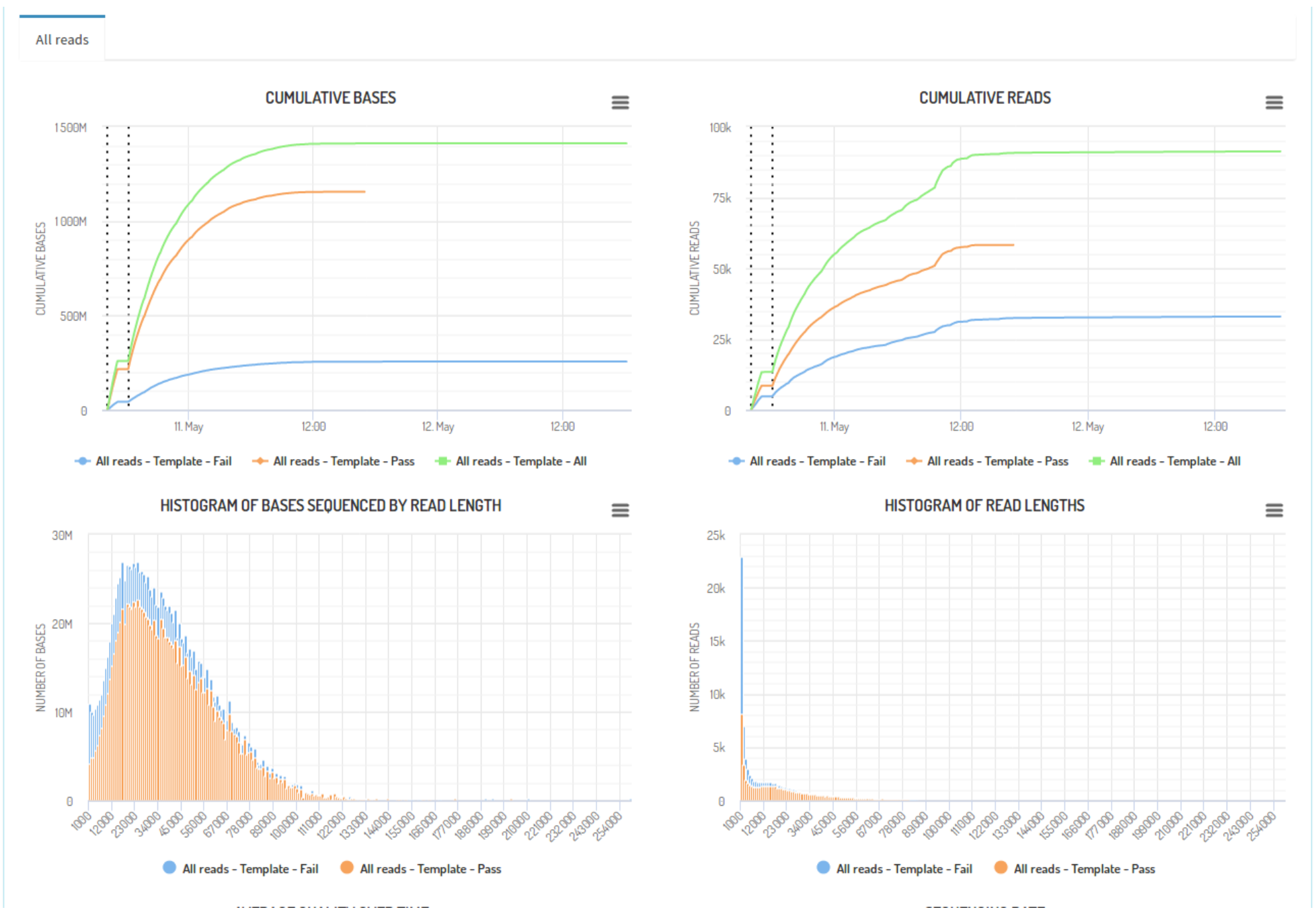


Figure 5: An example of basecalled sequence metrics for a run are shown above.

7. Alignment

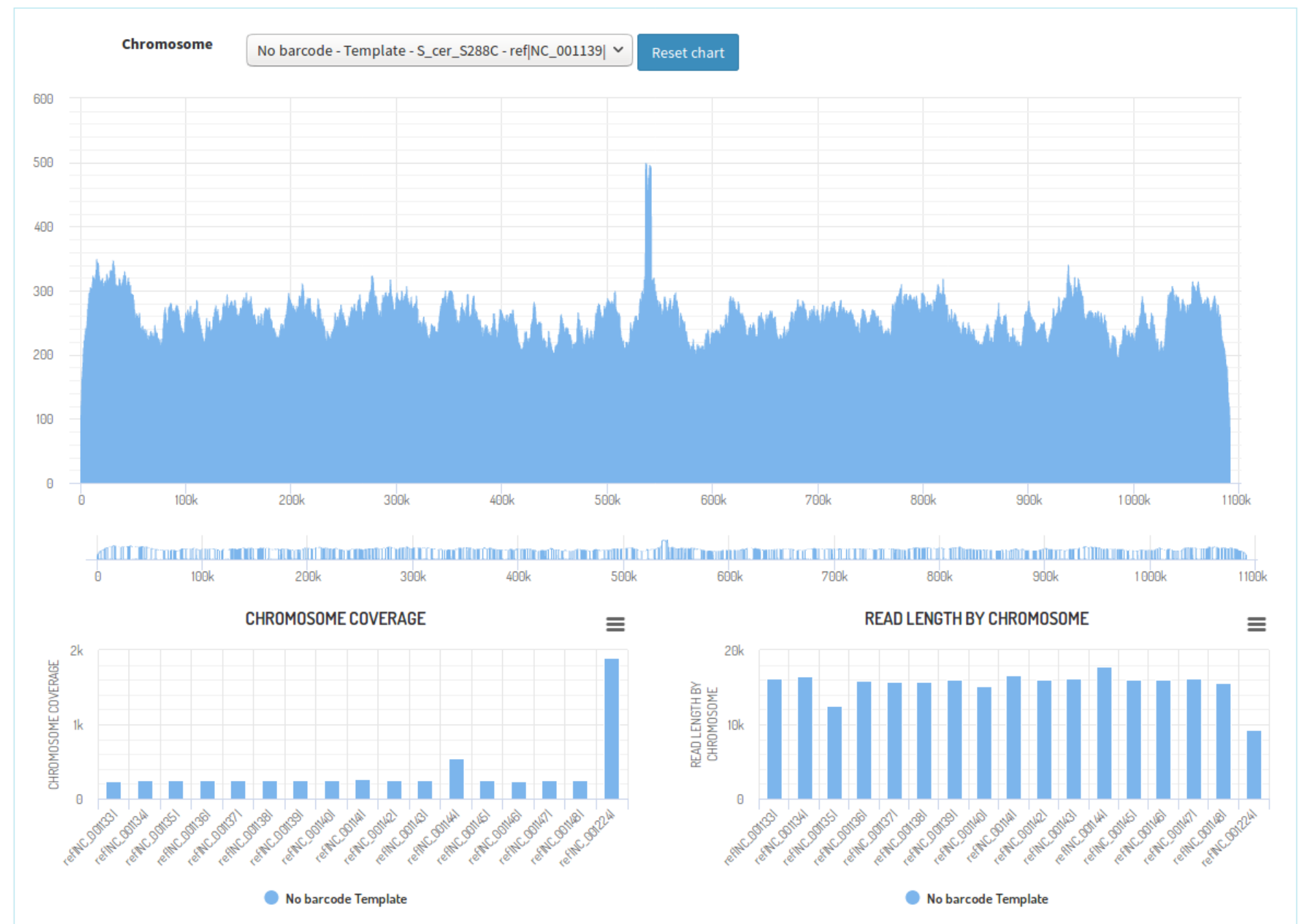


Figure 6: Mapping coverage - An example of alignment as performed by minoTour is shown above.

minoTour can run rapid alignments of base-called data against a reference genome, using minimap2 [1] against reads produced in real-time. This data is visualised to show to give rapid insights into the sequencing run. Example visualisations of minoTours' alignment are shown in Figure 6.

8. Metagenomics

minoTour can run metagenomics on a sample, using centrifuge [2] to classify its broad composition in real-time. If minoTour is running locally, a custom index can be chosen to classify runs against. Example metagenomics visualisations are shown in Figure 7.

Figure 7: Example metagenomics visualisations as shown by minoTour.

9. Conclusion

minoTour is accessible for both Docker and a server based setup; it offers a free, open source full LIMS for ONT sequencing. minoTour is currently under active development, and there are several improvements and features still in the queue for integration.

Acknowledgements

Thanks to Nadine Holmes and Victoria Wright for all their hard work sequencing!

Bibliography

[1] Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics, 34:3094-3100. doi:10.1093/bioinformatics/bty191

[2] Kim, D., Song, L., Breitwieser, F. and Salzberg, S. (2016). Centrifuge: rapid and sensitive classification of metagenomic sequences. Genome Res. doi:10.1101/gr.210641.116