

Introduction to MinNOW

Introduction to MinNOW

The MinNOW™ software carries out several core tasks: data acquisition, real-time analysis and feedback, basecalling, data streaming, controlling the device, and ensuring that the platform chemistry is performing correctly to run the samples. MinNOW takes the raw data and converts it into reads by recognition of the distinctive change in current that occurs when a DNA strand enters and leaves the pore. MinNOW then basecalls the reads, and writes out the data into .fast5 or FASTQ files.

MinNOW for MinION, GridION and PromethION

All Oxford Nanopore devices use MinNOW as the primary software. The user interface differs depending on the device, however the underlying instrument control and data acquisition functions are the same.

Computer specification required to run nanopore experiments

Computer specification

pdf

Uploaded file

[MinION Mk1B Computer Requirements.pdf](#)

pdf

Uploaded file

[GridION Mk1 Computer Requirements.pdf](#)

pdf

Uploaded file

[PromethION Computer Requirements.pdf](#)

MinNOW GUI

MinNOW GUI introduction

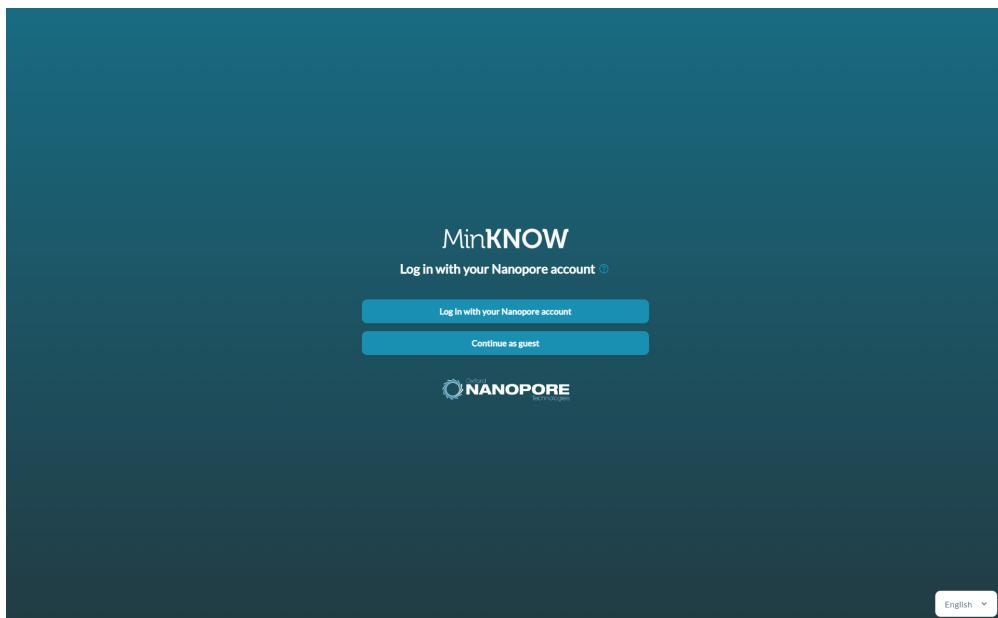
The MinNOW™ software carries out several core tasks: data acquisition, real-time analysis and feedback, basecalling, data streaming, providing device control, and ensuring that the platform chemistry is performing correctly to run the samples. MinNOW takes the raw data and converts it into reads by recognition of the distinctive change in current that occurs when a DNA strand enters and leaves the pore. The reads are then basecalled, and written into .fast5 or FASTQ files. Post-run analysis can now also be carried out through the GUI without needing to use the command line.

Login screen

Before a device is connected, the following screen appears:

We recommend users to log into the MinKNOW software using their Community credentials.

If you experience login issues, please contact Technical Support via email (support@nanoporetech.com) or via LiveChat in the Nanopore Community and use Continue as guest for temporary use.



Sequencing Overview

This page displays the inserted flow cell state and progress of a current sequencing experiment, including pausing, pore scan and basecalling.

Flow cell with no checks:

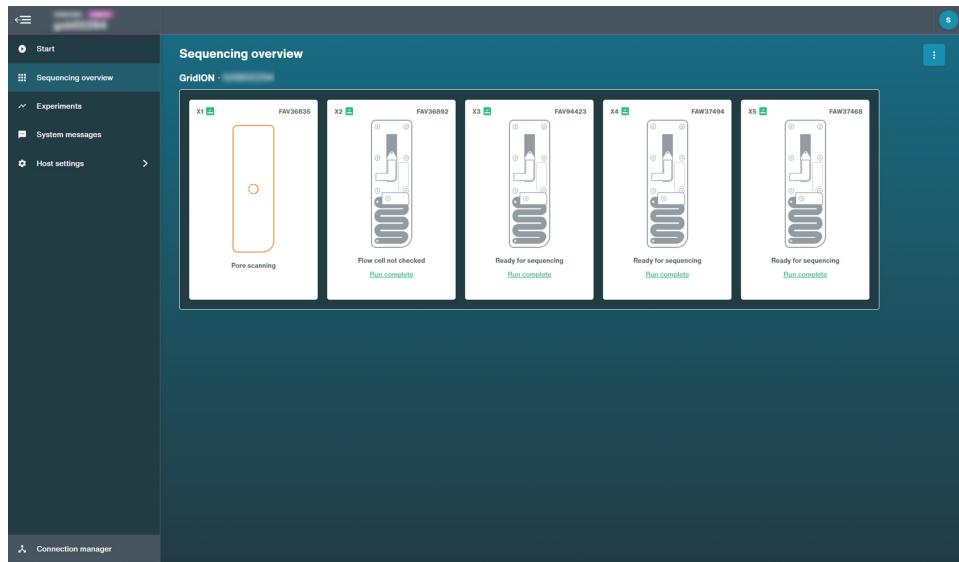
This example is on a GridION device.

The image shows the "Sequencing overview" page for a GridION device. On the left, a sidebar menu includes "Start", "Sequencing overview" (which is selected and highlighted in blue), "Experiments", "System messages", and "Host settings". The main content area is titled "Sequencing overview" and "GridION - [REDACTED]". It displays five flow cell slots, each represented by a small diagram of a flow cell cartridge. The flow cells are labeled X1, X2, X3, X4, and X5. Each slot has a green checkmark icon and a unique identifier: FA-08164, FAV36892, FAV94423, FAW37494, and FAW37463 respectively. Below each slot, the status message "Flow cell not checked" is displayed. At the bottom of the page, there is a "Connection manager" section.

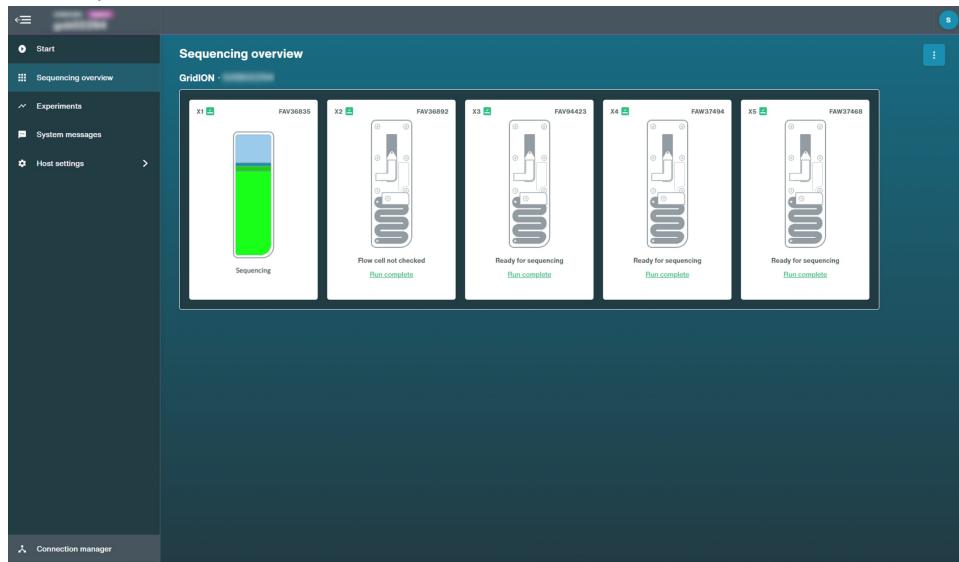
Sequencing states:

These examples are on a GridION device.

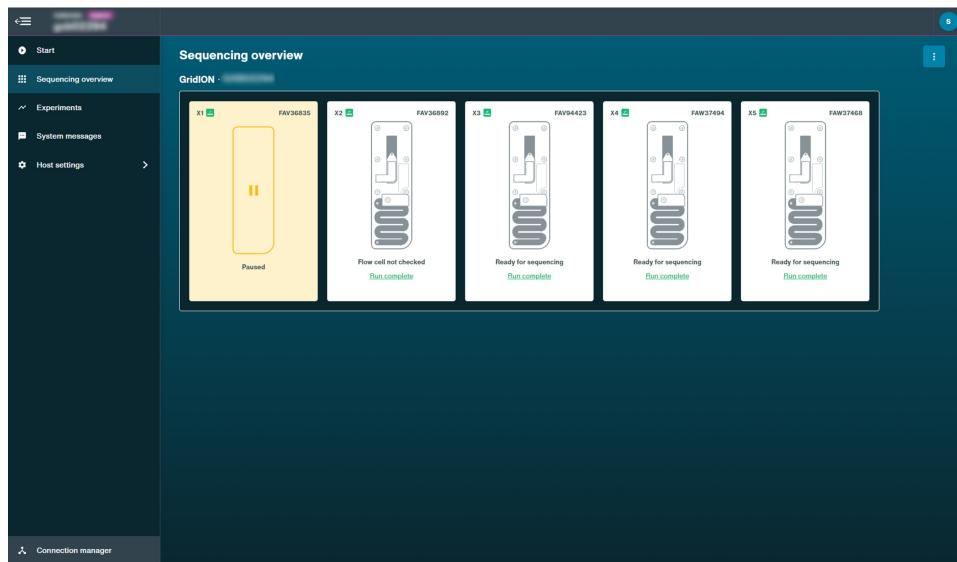
- Pore scan:



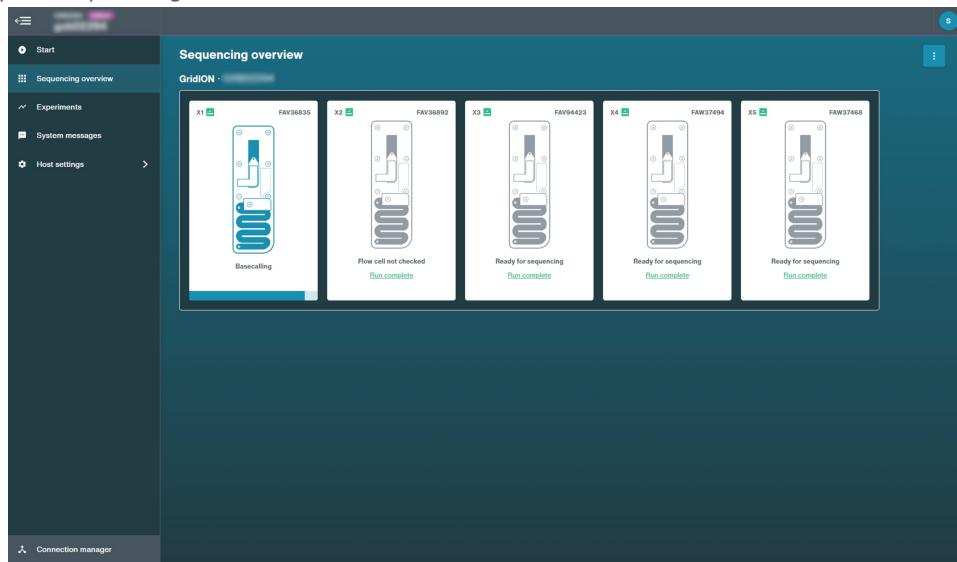
- Flow cell health after the first pore scan:



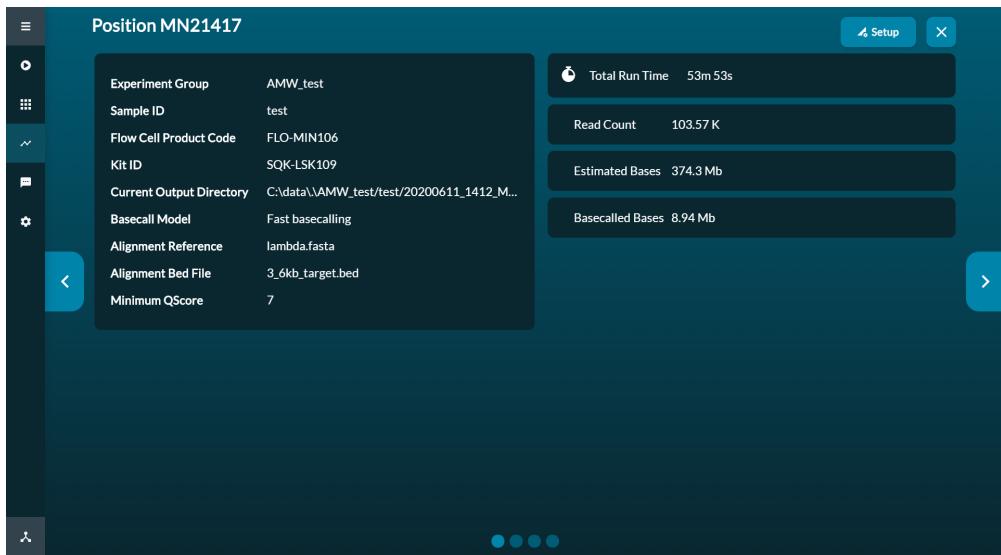
- Run paused:



- Basecalling catch up post-sequencing:

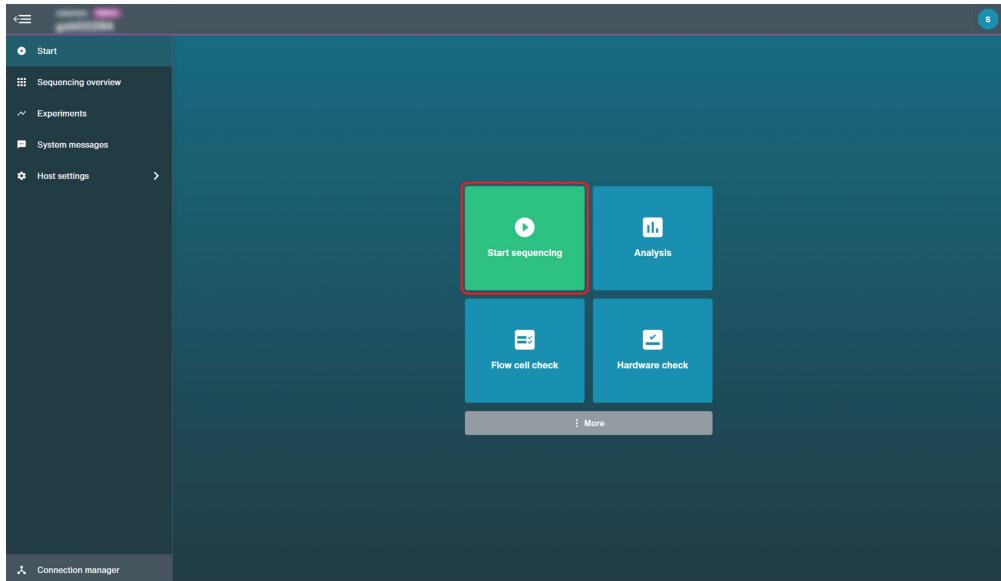


Click the flow cell to open the quick view of the current experiment.



Experiment configuration

To start an experiment, the user clicks **Start Sequencing** to open the experiment configuration options.



There are a number of parameters to be configured:

- **Flow cell position and experiment name**

The "Positions" tab will show the chosen flow cell. An experiment name can then be assigned.

The other tabs will not become available until an experiment name has been provided.

- **Kit selection**

The "Kit" tab will provide a table of available kits. If a kit selected is compatible with a barcoding expansion pack, the barcoding options will become available for selection.

- **Run Options**

The "Run options" tab provides variables for run time, minimum read length and the option to use adaptive sampling. Adaptive sampling can be enabled from here on MinION Mk1C, GridION and PromethION.

Analysis

Select whether data will be basecalled live on the instrument.

Note: MinKNOW provides the option to re-basecall .fast5 files during post-run analysis.

Output

Specify where to save the output data for the run, and the file type.

Barcode demultiplexing

During kit selection, if you have used a barcoding kit or a barcoding expansion pack for your library preparation, MinKNOW will split your reads by barcode without having to use commandline tools. Demultiplexing places reads into barcode-specific folders.

The screenshot shows the 'Kit selection' step of the sequencing workflow. On the left, there's a sidebar with 'Start', 'Sequencing overview', 'Experiments', 'System messages', 'Host settings', 'Application settings', and 'Connection manager'. The main area has tabs for '1 Positions', '2 Kit', '3 Run options', '4 Basecalling', '5 Output', and '6 Review'. Under 'Kit selection', there are filters for 'Sample type' (DNA, RNA, PCR-free, PCR), 'Multiplexing' (Yes, No, Control), and a 'Reset filters' button. A table lists various sequencing kits: Ligation Sequencing Kit (selected), PCR Barcoding Kit, Direct cDNA Sequencing Kit, Ligation Sequencing Kit XL, PCR cDNA Barcoding Kit, PCR Sequencing Kit, SQK-ULK001, VSK-VSK003, and Direct cDNA Sequencing Kit. Below this is a section for 'Select barcode expansion pack' with two rows: Native Barcoding Expansion 1-12 (PCR-free) and PCR Barcoding Expansion 1-96. At the bottom are buttons for 'Continue to run options' and 'Skip to final review'.

In the "Analysis" tab, the barcoding option can be toggled on and off when setting up an experiment. Further options are also available by selecting **Edit options**, including trimming barcodes, which can be used to enable barcode trimming in read files. Please note that some primer sequences may also be trimmed together with the barcodes.

The screenshot shows the 'Analysis' tab of the workflow. It includes sections for 'Basecalling', 'Barcode' (with 'Barcode ON' selected), 'Alignment', and 'Options' (with 'Model high accuracy basecalling' and 'Edit options'). A red box highlights the 'Barcode' section. To the right, a modal dialog titled 'Barcode options' is open, showing settings for 'Trim barcodes', 'Mid-read barcode filtering', 'Barcode both ends', and 'Override minimum barcoding score' (set to 60). Buttons for 'Cancel' and 'Save' are at the bottom.

Alignment

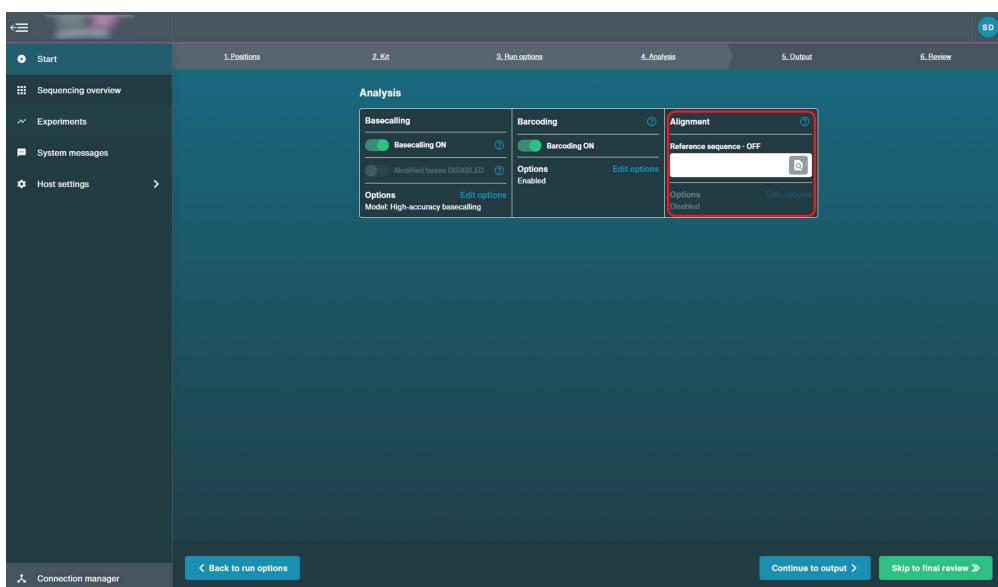
Alignment can be performed live during sequencing alongside basecalling on MinKNOW. An alignment reference must be uploaded

locally on bacterial-sized genomes as a .fasta or minimap index file.

The .fasta or minimap index file can contain multiple entries within the same file (e.g. multiple chromosome). The .fasta and minimap index file alignment hits will populate the alignment graphs.

A .bed file may be uploaded alongside the reference .fasta or minimap index file. The .bed file can be used when the user is interested in a particular region of the reference (e.g. specific gene in the chromosome). The .bed file alignment hits will be highlighted in the sequencing .txt file generated in the data folder

Navigate to the "Analysis" tab and add a reference sequence in the Alignment box to turn on alignment. Basecalling must be enabled for live alignment.



Alignment can be carried out in post-run analysis with basecalling analysis. Alignment can also be carried out independently from basecalling to generate .bam files.

For further information on post-run alignment, refer to the post-run analysis section of this protocol.

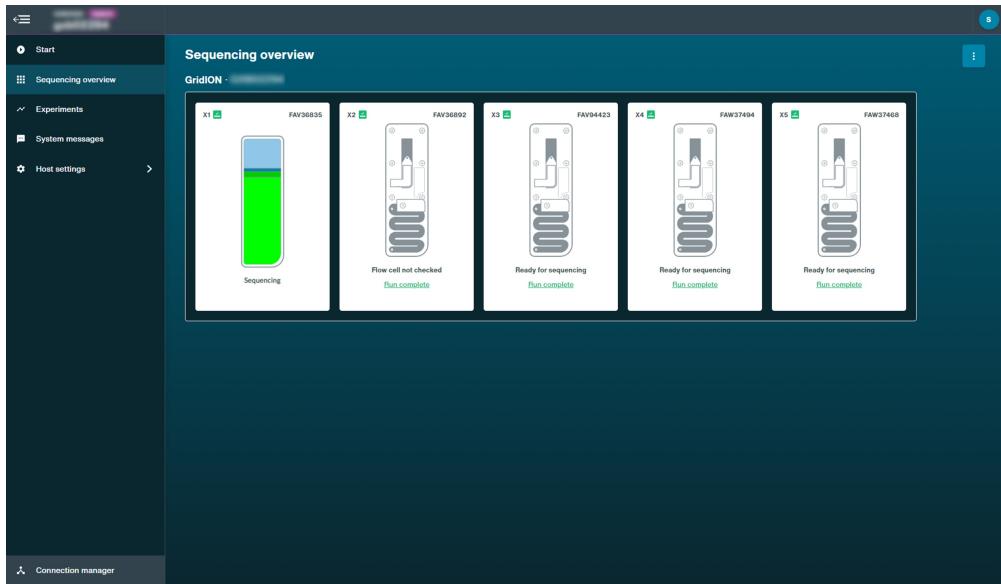
Flow cell health

During a sequencing experiment, the MinKNOW Sequencing Overview page shows a flow cell icon with coloured bars. The bars represent the combined health of all pores in a flow cell, and indicate how well the flow cell is performing. The colours are:

- Light green: sequencing
- Dark green: open pore
- Dark blue: pore recovering
- Light blue: pore inactive

This information is identical to the last bar of the duty time plot.

GridION flow cell health diagram



Experiments page

The experiments page displays summary information for all sequencing flow cells and device checks carried out on the device.

Previous runs can be viewed on the UI at the top of the screen and the number of days to view the last active experiments can be altered by typing in a different integer.

From this page, you can control specific runs and identify real-time information, including flow cell health and reads.

- **Run statistics:** The total number of reads, estimated and basecalled bases across an experiment, and number of active and total runs
- **Run time:** The duration of the experiment
- **Run state:** The current state of the sequencing run; 'Active', 'Basecalling', 'Complete', 'Stopped with error'
- **Health:** The current flow cell health

The white panel displays a summary of sequencing experiments and the blue panel displays status information of a specific run.

Example of experiments page on GridION:

Experiments (6)

Device activity: Experiments active in the last 7 days.

Position	Flow cell ID	Sample ID	Health	Run time	Run state	Reads	Estimated bases	Basecalled bases	Basecalled %
X3	FAU00546	67	Green	26 m / 12 h	Active	18.1 k	98.73 Mb	—	—
X1	FAT97523	54	 	26 m / 12 h	Active	17.96 k	65.7 Mb	—	—
X2	FAT97517	55	Green	26 m / 12 h	Active	19.15 k	96.93 Mb	—	—
X4	FAT97530	68	 	26 m / 12 h	Active	360	1.23 Mb	—	—

For more status information of a specific run, click the run to open the quick view, including current temperature and voltage. In the example below, the run in position X3 was clicked to open the quick view.

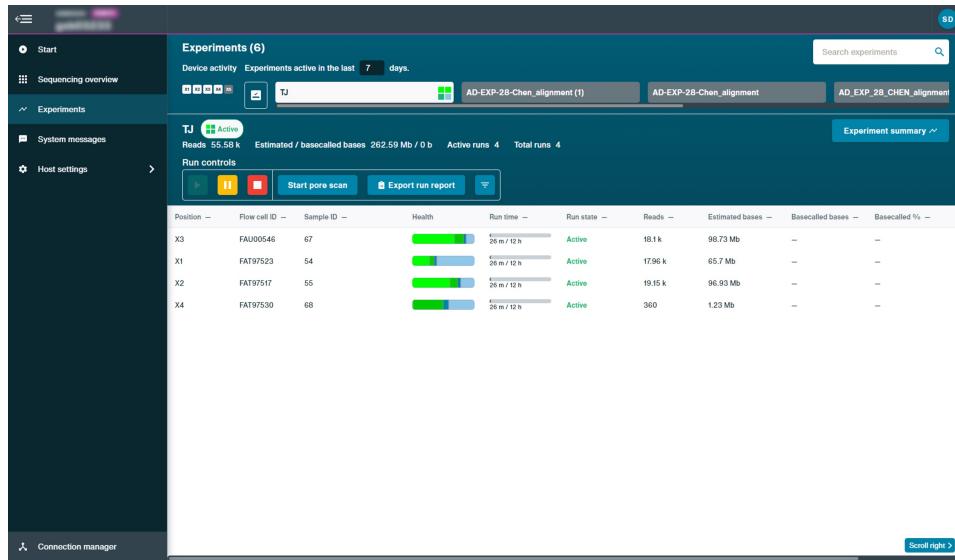
Position X3

Experiment group	TJ	Total run time	25m 38s
Sample ID	67	Temperature	35 °C
Flow cell product code	FLO-MINI14	Read count	17.66 k
Kit ID	SQK-LSK114-XL	Estimated bases	96.15 Mb
Current output directory	/data/TJ/67/20221014_1522_X3_FAU00546_23ede9e	Basecalled bases	0 b

Experiment page: stats and physical layout

The experiment page shows how the run is progressing in real-time. Select a run to view the real-time data in the quick view and use the arrows to move through the graphs.

- The quick view below displays information such as the number of generated and basecalled reads, run time, temperature, and voltage.
- The run time bar shows how far the experiment has progressed.



- Channel panel: a real-time representation of each of the channels (e.g. 512 for MinION Mk1B)



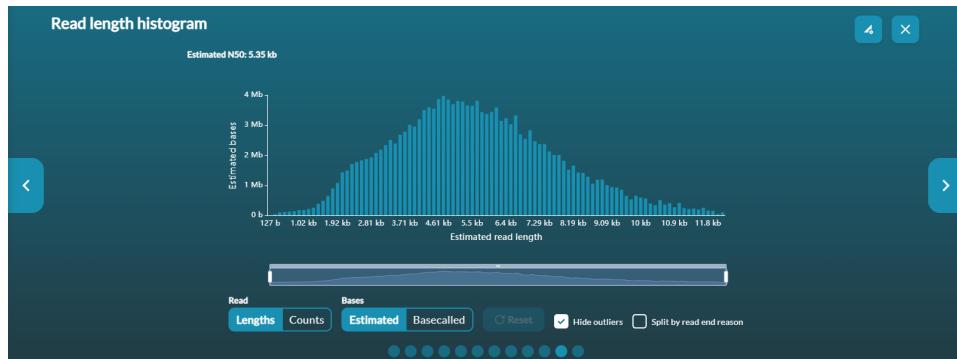
- Pore activity plot: a summary of the channel states over time. This can be used to assess the quality of the sequencing experiment



- Pore scan results: a summary of the channel states during the most recent pore scan



- Read length histogram: a real-time representation of the reads generated so far



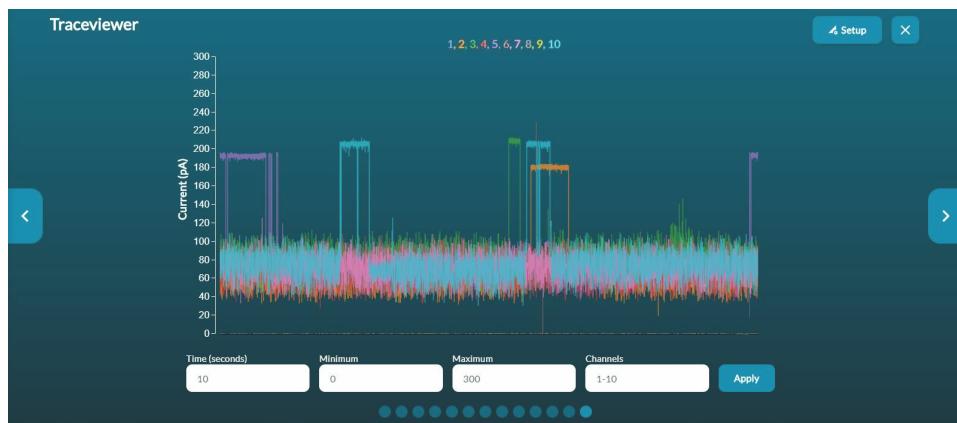
- Basecalling: shows the translocation speed and Qscore against time.



- Cumulative output: shows the number of bases that have been sequenced and basecalled

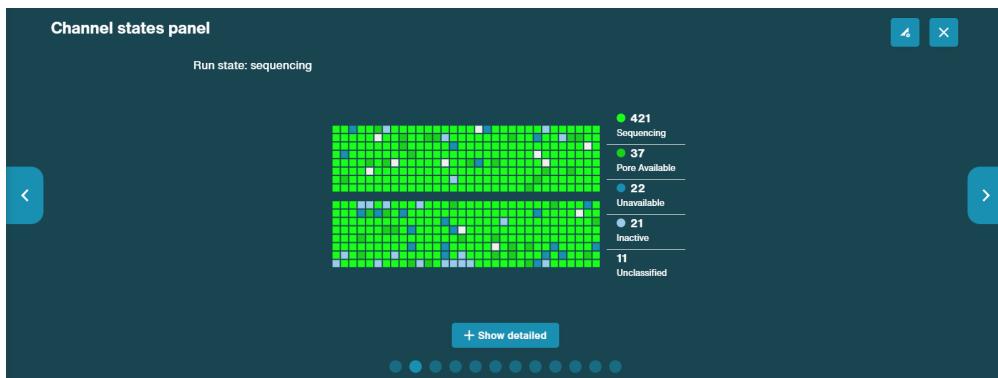


- Trace viewer: shows the current level fluctuations in the selected channels



Pore occupancy

The channel states panel gives an overview of the states the flow cell pores are in to give the user an idea of how well the sequencing run is performing in real time. A good library will be indicated by a higher proportion of light green channels in "Sequencing" than are in "Pore available". The combination of "Sequencing" and "Pore available" indicates the number of active pores at any point in time. A low proportion of "Sequencing" channels will reduce the output of the run.



Clicking on the **Show Detailed** button reveals a more detailed array of channel states:

- **Sequencing:** Pore currently sequencing.
- **Pore Available:** Pore available for sequencing.
- **Adapter:** Pore currently sequencing adapter.
- **Active feedback:** Channel ejecting analyte.

- **No pore:** No pore detected in channel.
- **Multiple:** Multiple pores detected. Unavailable for sequencing.
- **Unavailable:** Pore unavailable for sequencing.
- **Unclassified:** Pore status unknown.
- **Saturated:** The channel has switched off due to current levels exceeding hardware limitations.
- **Out of range-high:** Current is positive but unavailable for sequencing.
- **Out of range-low:** Current is negative but unavailable for sequencing.
- **Zero:** Pore currently unavailable for sequencings.

System Messages

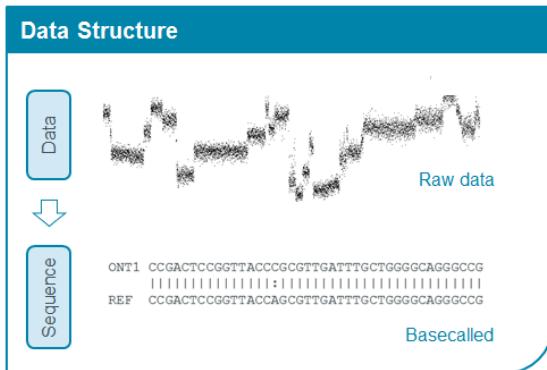
All device reports and messages are displayed here.

ID	Message	Date
X2	Channel scan for flow cell AFX570 has found a total of 46 pores. 46 pores available for immediate sequencing	16 April, 11:20:30
X2	Performing Channel Scan	16 April, 11:19:10
X1	Channel scan for flow cell AGE388 has found a total of 53 pores. 53 pores available for immediate sequencing	16 April, 11:18:35
X1	Performing Channel Scan	16 April, 11:17:23
X4	Mux scan for flow cell FAQ13638 has found a total of 732 pores. 396 pores available for immediate sequencing	16 April, 11:14:49
X3	Mux scan for flow cell FAQ16580 has found a total of 577 pores. 321 pores available for immediate sequencing	16 April, 11:14:35
X4	Performing Mux Scan	16 April, 11:12:25
X3	Performing Mux Scan	16 April, 11:12:11
X2	Channel scan for flow cell AFX570 has found a total of 53 pores. 53 pores available for immediate sequencing	16 April, 09:48:03
X2	Performing Channel Scan	16 April, 09:46:50
X1	Channel scan for flow cell AGE388 has found a total of 55 pores. 55 pores available for immediate sequencing	16 April, 09:46:18

Analysis workflow

Analysis workflow

MinNOW acquires raw signal from the device, and sends it to the analysis pipeline (basecaller) in chunks of a defined size. Before being sent to the pipeline, MinNOW assesses the data and determines where the individual reads start and end by detecting the abrupt signal change when DNA enters and leaves the nanopore. Only data that is considered to be within a read is sent for analysis.



The format and location of your data will depend on the options chosen in the new experiment settings screen:

All files output directly from your experiment are located in the same directory. The directory has the structure:

`{output_dir}/{experiment_id}/{sample_id}/{start_time}_{device_ID}_{flow_cell_id}_{short_protocol_run_id}/`

- `output_dir` is the configured output directory

- `experiment_id` is the user-entered identifier for a group of runs

- `sample_id` is the user-entered value for a specific sample or run. It is intended that multiple `sample_ids` may exist beneath an individual `experiment_id`

- `start_time` is the time at which the protocol started in YYYYMMDD_HHMM format

- `device_id` is the serial ID of the MinION Mk1B or device position for GridION/PromethION

- `flow_cell_id` is the flow cell id (eg: FAH12345), either programmed on the ASIC or entered by the user

- `short_protocol_run_id` is a unique identifier of 7 characters from the protocol ID

Examples of the above naming convention:

/data/MyExperiment/Sample1/20181011_1759_X1_FAH12345_0ffe109

Individual read files will be split into .fast5 pass and fail folders, as well as .fastq pass and fail folder within this directory:

`{flowcell_id}_{basecall_state}_{short_run_id}_{batch_number}.fastq`

`{flowcell_id}_{basecall_state}_{short_run_id}_{batch_number}.fast5`

If the reads are barcoded, the barcode will be included in the file name before the `short_run_id`.

Examples of the above file naming:

fast5_pass/FAK12345_pass_9aad0448_0.fast5

fasq_pass/FAK12345_pass_9aad0448_0.fastq

Run reports can be found in the data folder (e.g. C:\data\reports).

Local basecalling

Introduction to live basecalling in MinKNOW

For MinION Mk1B and Flongle (on MinION Mk1B), the MinKNOW software presents an option to basecall reads on the local computer. The basecalling is carried out live, as the read files are generated during a sequencing experiment.

Basecalling results are displayed in real-time in the MinKNOW user interface, and data is written out in the BAM or FASTQ file format.



The MinION Mk1C, GridION and PromethION devices have the capacity to perform live basecalling, while keeping up with one (MinION Mk1C), five (GridION) and up to 41 flow cells (on the PromethION 48, or 48 flow cells on PromethION 48 A100) running fast basecalling simultaneously. Basecalling is carried out directly on the device and uses the MinKNOW software. The basecalled reads are written to the device; users can then export the reads out for further analysis via the Network or removable storage.

Fast, High Accuracy and Super Accurate models and compatibilities

The Dorado and Guppy basecallers offer three different basecalling models: a Fast model, a High accuracy (HAC) model, and Super accurate (SUP) model.

The Fast model is designed to keep up with data generation on Oxford Nanopore devices (MinION Mk1C, GridION, PromethION). The HAC model provides a higher raw read accuracy than the Fast model and is more computationally-intensive. The Super accurate model has an even higher raw read accuracy, and is even more intensive than the HAC model.

For more information about basecalling accuracy, please consult the [Accuracy page on the Oxford Nanopore website](#)

A comparison of the speed of the models is provided in the table below:

Device	Fast model		HAC model		sup model	
	No. of keep-up flow cells	Basecalling only	No. of keep-up flow cells	Basecalling only	No. of keep-up flow cells	Remora 5mC calling
PromethION 48 (A100)	48	5.1	18.4	4.7	3.9	3.4
PromethION 48	41	3.6	8	3.8	1.4	1.3
GridION Mk1	5	3.9	5	3.4	1.8	1.6
MinION Mk1C	1	N/A	0.1	N/A	N/A	N/A

The number of keep-up flow cells was calculated assuming a 30 Gbase flow cell output in 72 hours for MinION and GridION, and 150 Gbase output in 72 hours for PromethION.

Note: Remora currently runs on CPU, which limits its speed. We will soon be integrating our new basecaller (Dorado), which will allow Remora to run on GPU with a considerable speed increase.

Sources of basecalling error

There are two main errors an event detection algorithm can make:

- **Insertions:** when an extra base is inserted where there should not be one - typically, this means a span of raw data points that corresponds to a single set of pore contents.
- **Deletions:** when a base is missed out.

Provided protocol scripts

Protocol scripts

The experiment protocols are a repository of Python scripts that are used to control the nanopore device during an experiment. The protocol selector in the MinNOW GUI requires input from the user at the start of the experiment, and will pick the script that fits the requirements in the options selected.

Once a protocol is selected, it controls the hardware settings such as temperature and voltage, as well as the duration of the run. During a run, the script uses MinNOW analysis tools to run and monitor the real-time elements of the experiment, for example channel flicking, voltage adjustments, pore scans, and selecting pore-containing channels.

Post-run basecalling

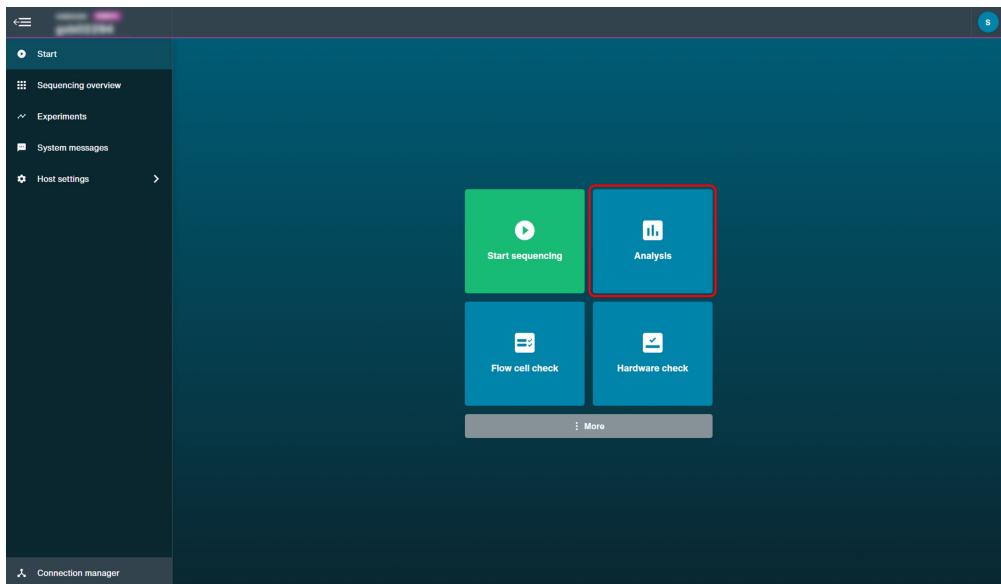
Basecalling overview

A user can basecall and demultiplex their data directly in MinNOW after a sequencing experiment has finished to avoid having to use command-line tools, or to re-analyse old data using the latest basecalling models. Reads can also be aligned against a reference post-run.

Note: Both barcoding and alignment can be run on .fastq, .pod5 or .fast5 reads when coupled with basecalling.

Accessing post-run basecalling

Post-run analysis can be accessed by selecting **Analysis** on the start page, then selecting **Basecalling**.



The options provided for basecalling data once a run has finished are described in the [MinKNOW protocol](#) under 'Post-run analysis'.

The user also has the option to use barcoding and alignment during basecalling analysis. However, both options can be carried out in independent post-run analyses.

Basecalling model selection

Guppy, the basecaller used by MinKNOW, provides multiple models for basecalling nanopore data in post-run basecalling. The models that can be selected are described in [Local basecalling](#) earlier in this document.

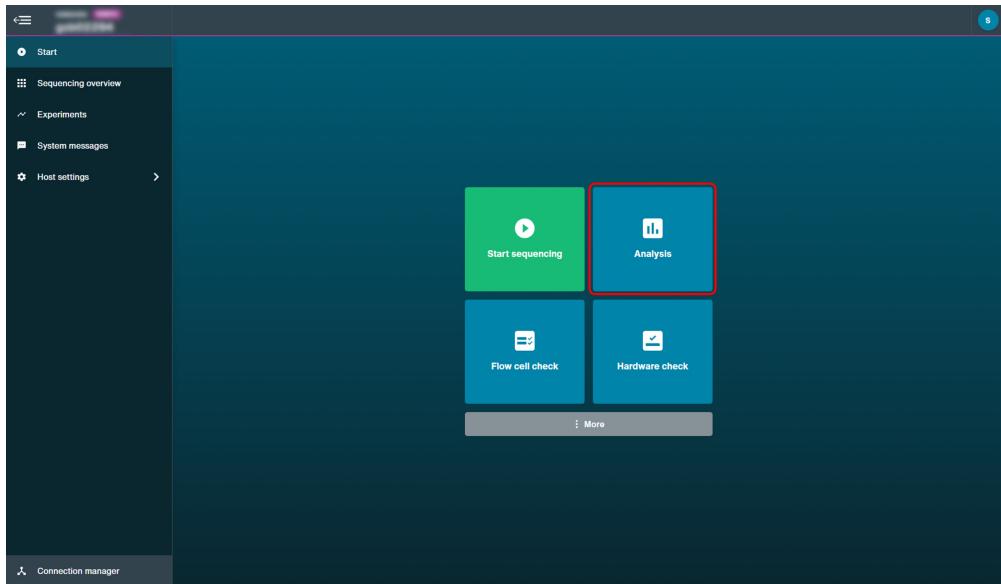
Post-run barcoding

Barcode overview

A user can barcode their data directly in MinKNOW after a sequencing experiment has finished, using .fastq data.

Accessing post-run barcoding

Post-run barcoding set-up can be found in the analysis menu.



The options provided for barcoding data once a run has finished are described in the [MinKNOW protocol](#) under 'Post-run analysis'.

Post-run alignment

Alignment overview

A user can align their data directly in MinKNOW after a sequencing experiment has finished, using a .fastq file from a previous run.

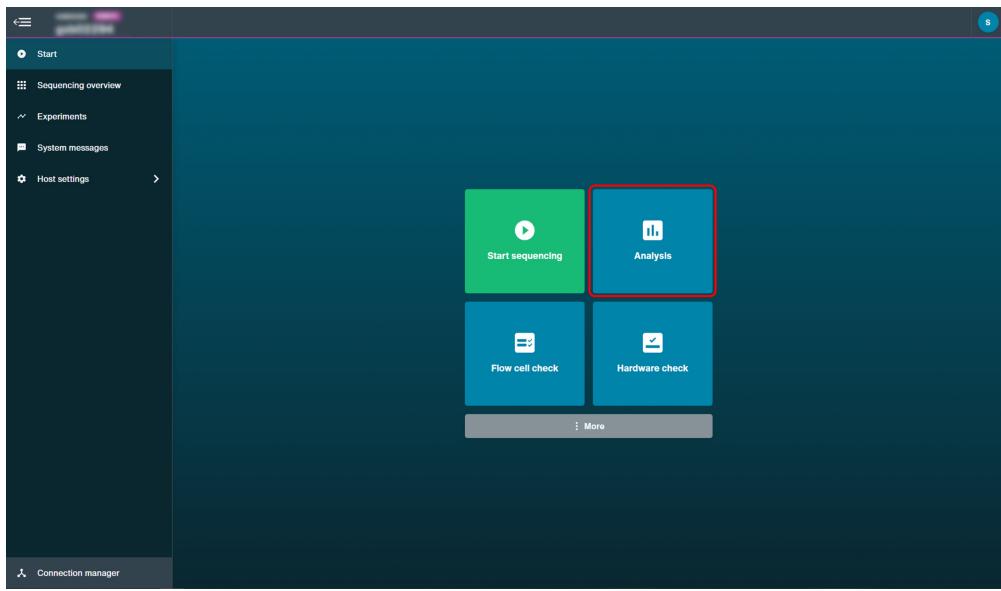
The reference file of bacterial-sized genomes must be uploaded locally as a .fasta file or minimap index file. Alignment hits from these files are used to populate the alignment graphs.

A .bed file may also be uploaded alongside the reference .fasta or minimap index file when the user is interested in a particular region of the reference (e.g. specific gene in the chromosome). The .bed file alignment hits will be highlighted in the sequencing .txt file generated in the data folder.

When alignment is run independently from basecalling, .bam files are generated.

Accessing post-run alignment

Post-run alignment can be accessed through alignment menu and selecting **Alignment**.



The options provided for alignment data once a run has finished are described in the [MinKNOW protocol](#) under 'Post-run analysis'.