

MinION Mk1D IT requirements

Overview

The MinION™ is a small, self-contained device for nanopore sequencing. It plugs directly into a USB Type-C port with a recommended adapter (5 Gbps). The MinION device is controlled by the MinKNOW™ software. MinKNOW carries out several core tasks. These include data acquisition, real-time analysis and feedback, basecalling, data streaming, device control including selecting the run parameters, sample identification and tracking. Furthermore, it ensures that the platform chemistry is performing correctly to run the samples.

Despite its small size, the MinION Mk1D can sequence many gigabases (Gbases) in a single sequencing experiment. As such, we recommend that you purchase a powerful computer to ensure all the technology features are accessible. The default data analysis workflow when starting to use the MinION is below:



Configuring a new host computer

From MinKNOW v23.07 onwards, our new basecaller Dorado is integrated into MinKNOW. This enables accelerated basecalling on both NVIDIA GPUs and Apple Silicon. We strongly recommend using an NVIDIA GPU or Apple Silicon Mac if you want basecalling to keep up with the rate of data generation.

Component	Windows, Linux	macOS
Operating system	Windows 10/11, Ubuntu 20.04/22.04 LTS	macOS
Peripheral	USB Type-C (USB 2.0 speeds or greater)	USB Type-C (USB 2.0 speeds or greater)
Memory	16 GB or higher	16 GB or higher
GPU	NVIDIA RTX 4070 or higher	Apple M3 Max
CPU	Intel or AMD Processor with at least 4 cores	Apple M3 Max
Storage	1 TB SSD or greater	1 TB SSD or greater

We recommend internal solid-state storage for MinKNOW installation as well as data output/acquisition. Solid-state drives are much faster than traditional hard drives and are able to keep up with the flow of data generated during a sequencing run.

Example computers

Below is a non-exhaustive list of example models/workstations that meet or can be configured to meet the recommended specification for the MinION Mk1D. We have done basic testing but have not extensively tested any of these computers, so purchasing is at your

own risk. Ensure that you have configured the computer to meet the specification in the table above. Other manufacturers are available; check the availability for your region.

- Apple MacBook Pro 14” M3 Max
- Razer Blade 18 (RTX 4070 or greater)

The list above is accurate as of April 2024. Newer or equivalent models may have replaced the models above at time of reading.

IT requirements

The below requirements are to ensure correct operation of the MinION during a sequencing experiment:

Component	Minimum requirement
User account privilege level	Local Administrator privileges are needed for MinKNOW installation and updates. Local Administrator privileges are not required for running sequencing experiments.
Internet connection	An internet connection is required at all times for software updates and telemetry. Offline configurations can be made available for field use and expeditions. Please contact support@nanoporetech.com if you are planning to use your MinION offline.
Antivirus settings	Antivirus software scanning the entire hard drive uses significant amounts of resources and may cause performance issues. For this reason, we recommend that you switch the antivirus scans to manual setting rather than to automatic, so that they can be performed when the MinION is not in use.
OS update settings	Switch the OS system updates to manual mode, as downloading them during a sequencing experiment may impact performance. Updates that require restart will fatally halt a run.
Telemetry feedback	HTTPS/port 443 to 52.17.110.146, 52.31.111.95, 79.125.100.3 (outbound-only access) or DNS rule for ping.oxfordnanoportal.com
EPI2ME analysis	Ethernet: HTTPS/port: 443 TCP access to AWS eu-west-1 IP ranges: http://docs.aws.amazon.com/general/latest/gr/aws-ip-ranges.html
Software updates	HTTPS/port 443 to 178.79.175.200 and 96.126.99.215 (outbound-only access) or DNS rule for cdn.oxfordnanoportal.com

Telemetry

MinKNOW collects telemetry information during sequencing runs as per the Terms and Conditions to allow monitoring of device performance and enable remote troubleshooting. Some of this information comes from free-form text entry fields, therefore no personally-identifiable information should be included. We do not collect any sequence data.

The EPI2ME platform is hosted within AWS and provides cloud-based analysis solutions for multiple applications. Users upload sequence data in FASTQ format via the EPI2ME Agent, which processes the data through defined pipelines within the EPI2ME Portal. Downloads from EPI2ME are either in Data+Telemetry or Telemetry form. The EPI2ME portal uses telemetry information to populate reports.

Software updates

Depending on your geographical region, only one of 178.79.175.200 or 96.126.99.215 will be used for provision of updates to device

software. The Updates are triggered as pull requests, therefore outbound-only access is required.

File types

Nanopore sequencing data is stored in three file types: POD5, FASTQ and BAM. Basecalling summary information is stored in a sequencing_summary.txt file:

- POD5 is an Oxford Nanopore-developed file format which stores nanopore data in an accessible way and replaces the legacy .fast5 format. This output also reads and writes data faster, uses less compute and has smaller raw data file size than .fast5.
- .fast5 is a legacy file format based upon the .hdf5 file type, which contains all information needed for analysing nanopore sequencing data and tracking it back to its source. A .fast5 file contains data from multiple reads (4000 reads as default), and is several hundred Mb in size.
- FASTQ is a text-based sequence storage format, containing both the sequence of DNA/RNA and its quality scores. By default, nanopore sequencing experiments save up to 4000 DNA sequences in one FASTQ file. File size can vary from <1 Mb to tens of Mb depending on the number and length of sequences. Retaining only FASTQ files will allow use of standard downstream analysis tools using the DNA/RNA sequence, but no further sequence data can be generated when improvements in basecalling become available.
- BAM files are output if you perform alignment or modified base calling on the basecalled dataset.
- sequencing_summary.txt contains metadata about all basecalled reads from an individual run. Information includes read ID, sequence length, per-read q-score, duration etc. The size of a sequence summary file will depend on the number of reads sequenced.

Example file sizes below are based on different throughputs from an individual flow cell, with a run saving POD5, FASTQ, and BAM files with a read N50 of 23 kb.

Flow cell output (Gbases)	POD5 storage (Gbytes)	FASTQ.gz storage (Gbytes)	Unaligned BAM with modifications (Gbytes)
10	70	6.5	6
15	105	9.75	9
30	210	19.5	18

As an experiment progresses, POD5 files are produced for all reads. If you choose to basecall your data, these reads are used by the MinKNOW software to generate sequence data which is then stored in FASTQ files and/or BAM files.

Change log

Date	Version	Changes made
2024	V1	Initial document publication