Oxford Nanopore Technologies

Investigating methylation in the human genome with Oxford Nanopore

Introduction

Reveal more biology with direct sequencing of native DNA.

Epigenetics, the study of chemical modifications that can alter phenotype without altering nucleotide sequence, is an ever-growing field of research. The study of epigenetic modifications in humans — encompassing modifications to DNA, RNA, and histone proteins — is significant in fields ranging from developmental biology[1] to clinical research into diseases, including cancers[2] and neurological disorders[3].

The most well-characterised and widely studied epigenetic modification in mammalian genomes, and especially in humans, is 5-methylcytosine (5mC) DNA methylation — the addition of a methyl group to a cytosine nucleotide. Methylation plays an important role in regulating gene expression, with aberrant methylation in gene promoters being associated with disease.

There are many methods available for the analysis of methylation in human genomes. However, traditional techniques of detection present some limitations. For example, legacy short-read sequencing technology involves PCR, during which epigenetic modifications are lost; as methylation cannot be sequenced from PCR products, the presence of methylation is instead inferred via chemical treatment of DNA samples prior to sequencing.

Using Oxford Nanopore sequencing, it is possible to prepare and sequence native DNA and RNA molecules, without the need for PCR. This enables the direct detection of intact methylation alongside nucleotide sequence, without chemical conversion or additional library preparation steps.

Providing unprecedented resolution of methylation in targeted regions or across the whole human genome, the end-to-end nanopore workflows combine simple library preparation with flexible sequencing options to suit your experimental goals.

This guide provides an introduction to the direct sequencing of DNA methylation in human genomes. Oxford Nanopore sequencing is also the only technology that enables the direct sequencing and analysis of epigenetic modifications in native RNA. To find out more about native RNA sequencing, visit nanoporetech.com/rna-and-cdna-sequencing.

1. Greenberg, M.V.C. and Bourc'his, D. The diverse roles of DNA methylation in mammalian development and disease. *Nat. Rev. Cell Bio*. 20: 590–607 (2019). DOI: https://doi.org/10.1038/s41580-019-0159-6

2. Simpson, J.T. et al. Detecting DNA cytosine methylation using nanopore sequencing. *Nat. Methods* 14: 407–410 (2017). DOI: https://doi.org/10.1038/nmeth.4184

3. Weng, Y.L. et al. DNA modifications and neurological disorders. *Neurotherapeutics* 10(4): 556–567 (2013). DOI: https://doi.org/10.1007/s13311-013-0223-4

Gold-standard methylation calling with Oxford Nanopore: benchmarking performance

To benchmark nanopore methylation calling performance, we sequenced two replicate native DNA libraries of the well-characterised human genome HG002 on a PromethION™ device, generating 20x depth of coverage per sample (ONT_1 and ONT_2); these datasets were also merged to produce a third with 40x coverage (ONT_3). Basecalling and modification calling was performed using Remora: an algorithm that is now integrated into the basecaller Dorado, which can be used as a standalone tool or through the software onboard Oxford Nanopore sequencers. These methylation calls were compared with those from two short-read bisulfite sequencing datasets (BS-Seq_1[4] and BS-Seq_2[5,6]).

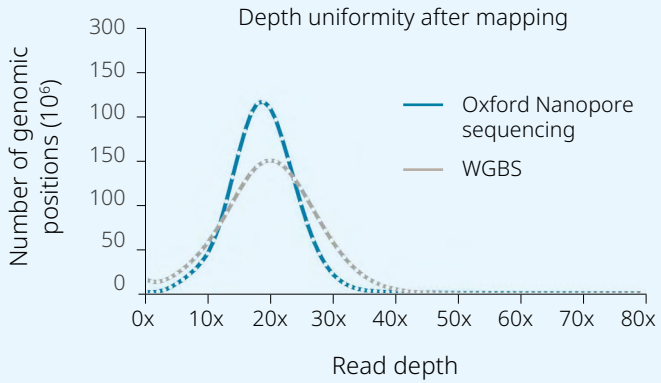


Figure 1. Oxford Nanopore methylation data shows greater evenness of coverage than whole-genome bisulfite sequencing (WGBS).

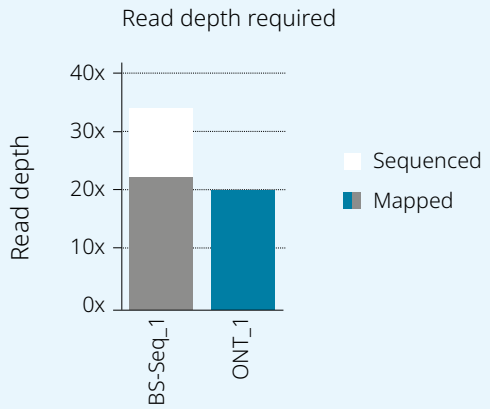


Figure 2. Whilst short-read bisulfite data shows incomplete mapping, Oxford Nanopore methylation reads map fully to the reference genome.

Bisulfite conversion followed by short-read sequencing is a widely used method of methylation detection. Bisulfite converts unmethylated cytosines to uracil, which is read as thymine during sequencing. 5mC is not converted, although the modification is lost during PCR. Comparison with a reference, or ideally a paired, untreated dataset, is then used to infer the presence of 5mC as cytosine. However, bisulfite treatment is a time-consuming process requiring the use of hazardous chemicals, and considerably fragments the sample, causing degradation of up to 90% of the DNA[7]. Chemical conversion can also be incomplete, with variability seen from sample to sample, which may impact the validity of results; furthermore, bisulfite treatment does not distinguish between 5mC and 5-hydroxymethyl cytosine (5hmC).

In 5mC benchmarking studies, the Oxford Nanopore sequencing data showed greater uniformity of coverage than the bisulfite data (Figure 1). Not all DNA sequences are amenable to PCR, meaning that areas of the genome, such as GC-rich regions, may be poorly represented in short-read sequencing datasets, or missed entirely. With PCR-free Oxford Nanopore sequencing, it is possible to generate more even coverage across the genome, including regions that are inaccessible to other sequencing methods[8]. Bisulfite conversion reduces sequence complexity which, combined with short-read sequencing, makes mapping to a reference genome challenging. By contrast, the long, PCR-free nanopore reads mapped to the reference genome (hg38) fully (Figure 2).

In the Oxford Nanopore datasets, the combined analysis time for both canonical and modified basecalling was over 2.5x faster than that for bisulfite modification calling alone. A high correlation (0.97) was seen between the methods for 5mC calling, and the nanopore datasets showed high reproducibility (0.95). With Oxford Nanopore methylation detection, there is also no need to sequence a second, paired sample to determine the presence or absence of methylation — or for a separate run to be performed to enable the analysis of other variants of interest. Instead, structural variants (SVs), single nucleotide variants (SNVs), copy number variants (CNVs), short tandem repeats (STRs), and methylation can all be called from a single dataset.

Long and ultra-long nanopore reads retain long-range information for methylation and genomic variants, and also enable effective phasing. An HG002 library prepared with the Ultra-Long DNA Sequencing Kit produced a read length N50 of 100 kb with a mapped depth of coverage of 40x. Single nucleotide polymorphism (SNP)-based phasing of the data enabled haplotype resolution of 90% of methylation calls, representing >80% of CpGs across the human genome.

The uniform sequencing coverage seen in Oxford Nanopore data allows for methylation detection across the human genome. 5mC is most commonly seen in CpG dinucleotides. Comparing CpG methylation calling between the datasets revealed that 20x bisulfite data enabled only ~75% of CpGs to be called in the human genome, whilst at the same depth of coverage,

Gold-standard methylation calling with Oxford Nanopore: benchmarking performance

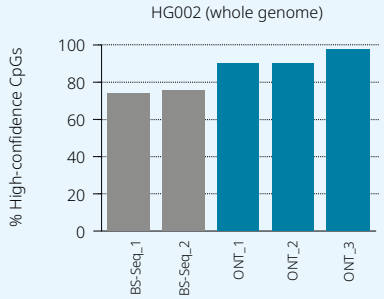


Figure 3. Percentage of called high-confidence CpGs in 20x bisulfite datasets compared with 20x and 40x nanopore methylation datasets.

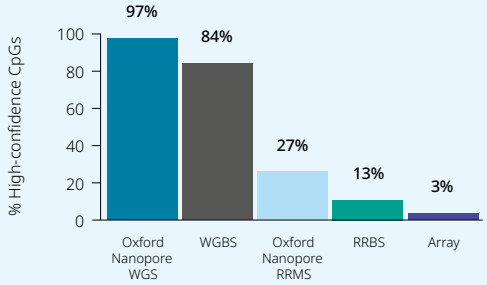


Figure 4. Percentage of CpGs detected in the human genome using different sequencing approaches (40x depth) and methylation microarray.

Oxford Nanopore sequencing enabled ~90% of CpGs to be called (**Figure 3**). Increasing nanopore sequencing depth to 40x enabled detection of >97% of CpGs. Though this dataset focuses on 5mC detection, a 5hmC Remora model is also available. We observe positive benchmarks for the 5hmC model validation metrics on a range of sample types. It is also possible to train your own Remora model to detect any epigenetic modification of interest.

Compared with WGBS, Oxford Nanopore methylation detection demonstrates:

- Strong correlation
- A higher number of CpG positions called
- Lower data requirements
- Faster analysis
- A simpler workflow, with no toxic components
- Better reproducibility and consistency run-to-run
- More even coverage, less effect of GC bias
- The option to phase methylation

Targeted methods offer a cost-efficient way to perform methylation analysis without the need to perform whole-genome sequencing (WGS). Methylation microarrays and reduced-representation bisulfite sequencing (RRBS) are frequently used to target methylated CpGs; however, both rely on bisulfite conversion. Methylation arrays have been used in large-scale 5mC profiling studies, despite being limited to only thousands of CpG sites, representing ~3% of the ~28 million CpGs present in the human genome (**Figure 4**). RRBS focuses on a larger subset of CpG islands; however, the method is expensive and time consuming.

To address these challenges, Oxford Nanopore has developed Reduced-Representation Methylation Sequencing (RRMS): a cost-effective method for characterisation of methylation patterns across regions of interest, such as CpG islands. The simple, end-to-end workflow utilises adaptive sampling (**page 7**), which requires no special library prep, offering flexible, real-time enrichment of native DNA during sequencing. Canonical basecalling is performed alongside the detection of base modifications, so that nucleotide sequence and epigenetic modifications, such as 5mC and 5hmC, can be detected simultaneously in a single sequencing experiment. RRMS displays more even coverage and higher reproducibility than RRBS, while capturing a significantly larger proportion of CpGs than RRBS and methylation arrays (**Figure 4**) — with near-perfect correlation (0.94) observed between the methods.

View the methylation benchmarking results in full: nanoporetech.com/methylation-benchmarking-poster

Read more about RRMS: nanoporetech.com/rrms-poster

4. Publicly available BS-Seq open dataset. Available at: https://labs.epi2me.io/category/data-releases/ [Accessed 11 Jan 2024]

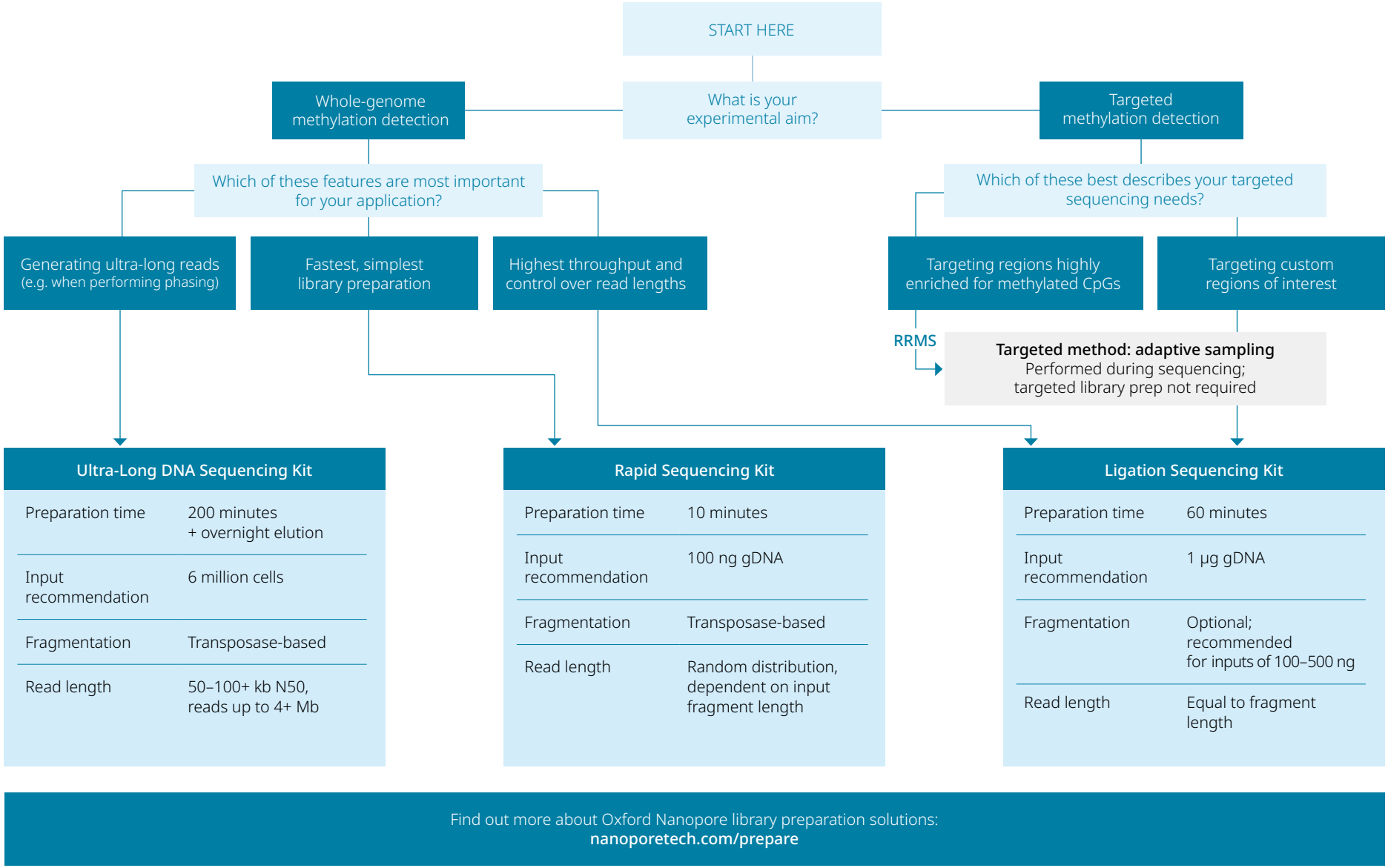
5. NCBI. Sample GSM5649436. https://www.ncbi.nlm.nih.gov/geo/query/acc. cgi?acc=GSM5649436 [Accessed 11 Jan 2024]

6. NCBI. Sample GSM5649437. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSM5649437 [Accessed 11 Jan 2024]

7. Grunau, C. et al. Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res*. 29(13):E65–5 (2001). DOI: https://doi.org/10.1093/nar/29.13.e65

8. Ebbert, M.T.W. et al. Systematic analysis of dark and camouflaged genes reveals disease-relevant genes hiding in plain sight. *Genome Biol*. 20,97 (2019). DOI: https://doi.org/10.1186/s13059-019-1707-2

Which approach do I choose ?



Find out more about Oxford Nanopore library preparation solutions:
nanoporetech.com/prepare

From sample to answer

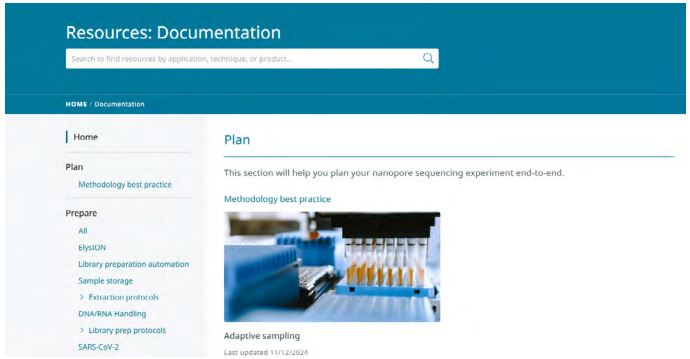
Extraction

How do I extract high-quality DNA from my sample?

The Documentation section of our website features recommended DNA extraction protocols for a wide range of sample types, including from cell lines, blood, and saliva. In this section you'll also find info sheets covering effective sample storage, DNA (and RNA) handling, size selection, and more.

View extraction protocols:
nanoporetech.com/extraction-methods

Documentation



Library prep

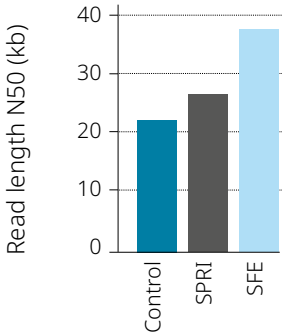
Should I perform fragmentation or size selection?

As read length is unrestricted with Oxford Nanopore sequencing, fragmentation of your DNA sample is optional. However, unless you plan to phase your data, maximising read length is generally not required for methylation detection. We have observed that performing light shearing can help improve sequencing output, to help maximise depth of coverage of methylated regions of interest.

If you do wish to perform phasing, you can optimise for long reads by avoiding fragmentation, whilst size selection can be used to minimise unwanted shorter fragments. For longer reads still — read length N50s over 50 kb and reads reaching the megabase scale — the Ultra-Long DNA Sequencing Kit can be used (**page 5**). Guidance on both fragmentation and size selection is available on the Documentation section of our website.

Find out more about optional fragmentation and size selection methods: nanoporetech.com/documentation/prepare

Read length N50s from three sequencing libraries, prepared without size selection (control), with SPRI size selection (SPRI), and with the Oxford Nanopore Short Fragment Eliminator Kit (SFE)



From sample to answer

Library prep

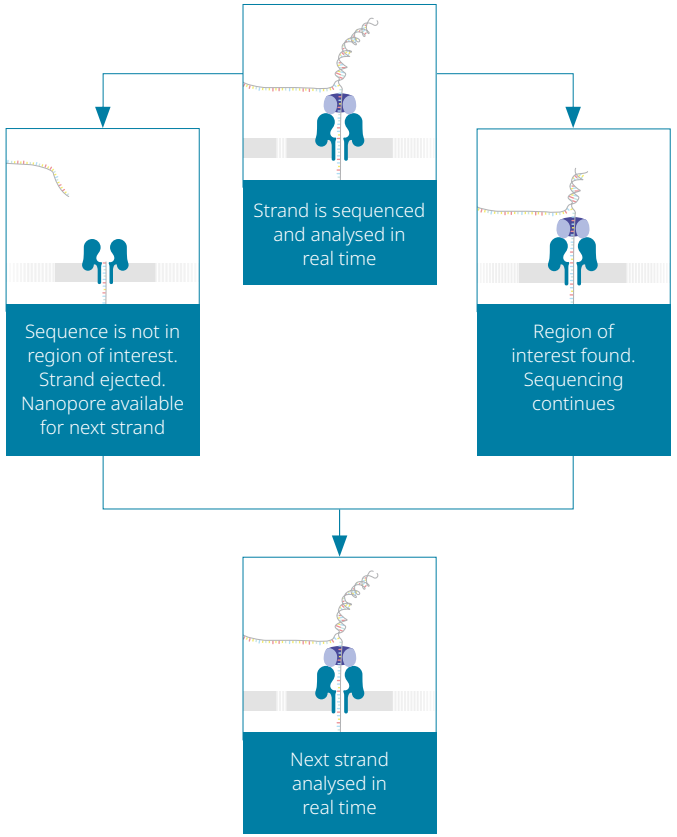
Can I perform targeted sequencing without losing methylation information?

Opting for a targeted approach is ideal for efficiently assessing methylation in specific regions of interest. To ensure that methylated bases are preserved through library preparation for detection in sequencing, PCR-free target enrichment is essential.

Adaptive sampling is an innovative, PCR-free method of target enrichment that is performed entirely during an Oxford Nanopore sequencing run, with no special library prep needed; it is also the method underpinning RRMS. Sequences to be enriched (or alternatively, depleted) are supplied to MinKNOW™, the software on board nanopore sequencing devices, in a BED file. DNA molecules representing regions of interest are allowed to pass through the nanopores and are sequenced in full. When an off-target sequence enters a nanopore, it is recognised in real time and ejected from the pore, freeing up more sequencing time for regions of interest. With no limit to the length of target regions, adaptive sampling is ideal for enrichment of very large regions of interest from native DNA samples, such as megabase-scale SVs or whole chromosomes.

View the targeted sequencing getting started guide:
nanoporetech.com/targeted-sequencing-guide

Enrichment/depletion of DNA strands during real-time sequencing with adaptive sampling



From sample to answer

Library prep

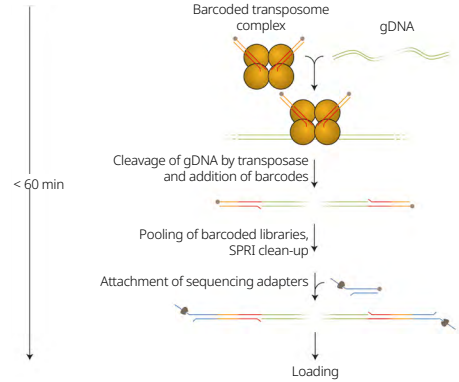
How can I sequence multiple samples at once?

Our Native Barcoding Kits enable sample barcoding without the use of PCR, for multiplexed sequencing of up to 96 native DNA libraries. This is achieved via the ligation of barcodes to each sample to be sequenced in the same run. These are then pooled and prepared for sequencing as a single sample.

Where rapid library preparation is key, the Rapid Barcoding Kits enable the preparation of up to 96 barcoded libraries in one hour. Here, a transposase-based method is used to shear the DNA and attach barcodes in a single step.

Find out more about Oxford Nanopore library prep kits:
nanoporetech.com/prepare

The rapid barcoding workflow



Sequencing

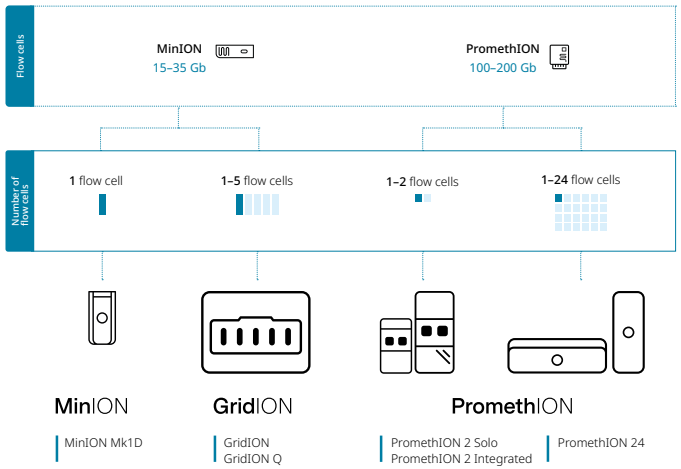
Which device should I choose?

A wide range of Oxford Nanopore sequencing devices is available, with each suiting different experimental requirements. For targeted analysis of methylation in specific regions, the MinION™, a portable device that can be operated from a laptop, is ideal. Providing the flexibility to scale up or down, the GridION™ allows sequencing on up to five MinION Flow Cells at once, for on-demand sequencing of target panels.

For analysis of methylation across the whole human genome, we recommend sequencing on a PromethION device. The ultra-high-throughput, benchtop PromethION 24 device, with powerful onboard compute, enables sequencing on up to 24 PromethION Flow Cells. The compact PromethION 2 Integrated, with integrated compute, and PromethION 2 Solo, which connects to a GridION or existing compute infrastructure, provide a simple plug-and-play setup, offering the flexibility of two independent, high-output PromethION Flow Cells for lower sample throughput requirements.

Find out more about Oxford Nanopore sequencing devices:
nanoporetech.com/sequence

High-output and scalable Oxford Nanopore sequencing devices



From sample to answer

Sequencing

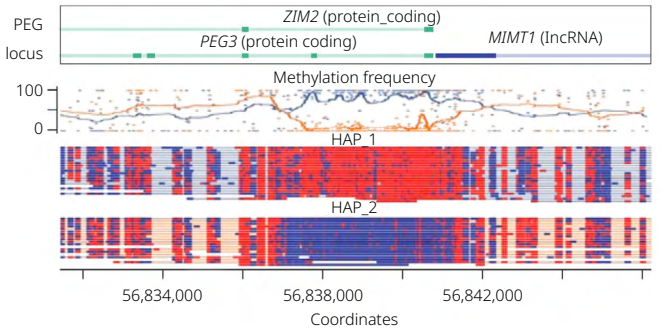
How much data do I need?

For good metrics in human genome-wide calling of 5mC and 5hmC methylation, we recommend sequencing to 20x depth of coverage; this was the read depth used in our Oxford Nanopore methylation calling performance benchmarking dataset (**pages 3–4**).

To phase your data, we recommend increasing read depth to around 30x. These depths can be achieved by sequencing on one PromethION Flow Cell for 72 hours. Long and ultra-long, native nanopore reads allow for easy and accurate phasing of data, with uniform coverage enabling haplotype-resolved methylation calling across large regions of interest.

View best practice guidance on human variant and methylation calling:
nanoporetech.com/resource-centre/workflow-human-variant-calling

A differentially methylated region overlapping two imprinted genes in the human genome was observed with haplotype-resolved nanopore methylation data

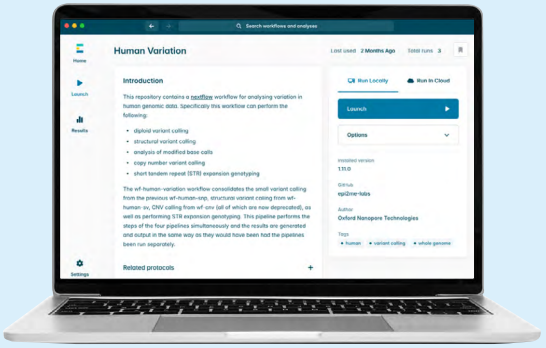


Data analysis

How can I detect methylation in my data?

For the highest accuracy canonical base and methylation calls from a single run, we recommend using the basecaller Dorado, which delivers gold-standard methylation calling alongside canonical basecalling, with minimal impact to basecalling speed. Dorado is integrated into MinKNOW — the software onboard nanopore sequencing devices — enabling live basecalling; it can also be used as a standalone tool. We recommend using the high accuracy (HAC) basecalling model. Using Dorado in HAC mode on a PromethION 24 device, live methylation calling can be performed in real-time in all positions concurrently.

For a deeper analysis of variants within the human genome, we recommend the workflow wf-human-variation. As well as providing methylation annotations and enabling haplotype phasing, this workflow — an EPI2ME™ solution — allows for the analysis of SNVs, CNVs, SVs, and STRs. The workflow can be run with simple, point-and-click implementation, or via the command line.



Find out more about nanopore data analysis:
nanoporetech.com/epi2me

From sample to answer

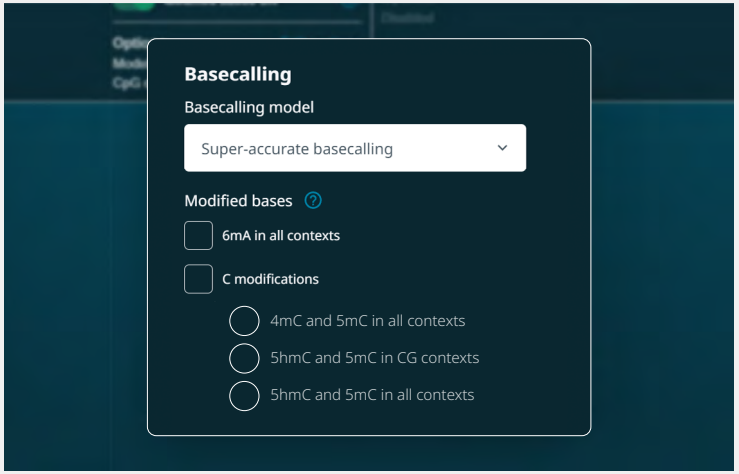
Data analysis

Can I distinguish between 5mC and 5hmC?

Unlike bisulfite sequencing and other available sequencing technologies, which are unable to distinguish between 5mC and 5hmC, Oxford Nanopore technology can accurately detect and distinguish between both modifications in the same dataset.

Ever-increasing modifications of both DNA and RNA are becoming available through the basecalling software Dorado and gradually integrated into the device software MinKNOW. DNA modifications currently supported are: 5mC, 5hmC, 6mA, and 4mC. RNA modifications include m^6A, pseU, m^5C, and inosine.

Read more about the latest modification models integrated into MinKNOW and Dorado: nanoporetech.com/platform/accuracy#base-modifications



Data analysis

Can I detect methylation at single-base resolution?

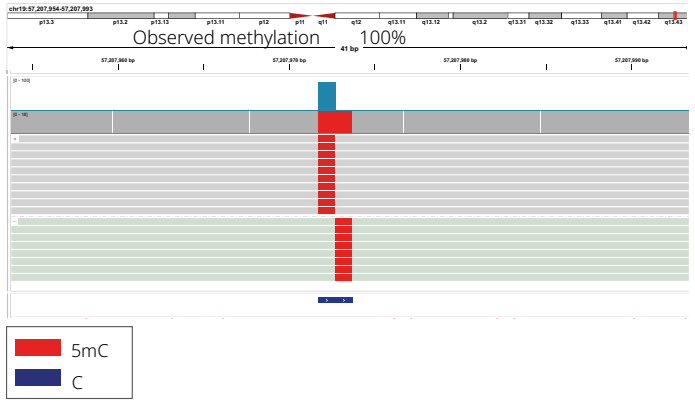
Traditional analysis of CpG methylation typically involves characterising averaged levels of CpG methylation at a site of interest; however, this precludes resolution of heterogeneity. With Oxford Nanopore sequencing, it is possible to perform high-accuracy methylation calling at the single-base level, allowing for exploration of epigenetic heterogeneity with high confidence.

Single-molecule DNA methylation patterns allow:

- Investigation of heterogenous cell populations
- Identification of epialleles
- Sample comparisons: for example, matched tumour-normal sequencing in cancer research

Find out more about epigenetics and methylation analysis using nanopore sequencing: nanoporetech.com/epigenetics

Nanopore sequencing enables single-nucleotide resolution of base modifications with high accuracy



Case studies

Case study 1: simultaneous structural variation and methylation analysis in hundreds of diverse brains with Oxford Nanopore sequencing

Neurological conditions like Alzheimer's disease and related dementias (ADRD) are driven by complex genetic and epigenetic mechanisms that remain poorly understood. In a recent study, Billingsley, Meredith, and Daida *et al*. used Oxford Nanopore sequencing to 'identify complex regulatory mechanisms in the brain that were inaccessible using previous approaches'[9]. Using long nanopore reads, the team produced high-quality genome assemblies from brain research samples of 351 neurologically normal individuals of European and African/mixed African ancestry, profiling both SVs and DNA methylation in a single assay.

The researchers identified approximately 234,905 SVs, including insertions and deletions. A subset of these SVs were found to impact gene expression in the brain, as demonstrated by quantitative trait locus (QTL) analyses linking SVs to gene expression changes in post-mortem frontal cortex tissue.

In parallel, the researchers generated haplotype-resolved DNA methylation profiles at millions of CpG sites. The authors highlighted that: 'ONT methylation data is on par with the quality and coverage of whole genome bisulfite sequencing ... and Illumina BeadChip Arrays. ONT, however, can simultaneously capture genomic variants and measure all modifications on native DNA molecules'. This integrative approach revealed *cis*-acting SVs that impact local methylation landscapes, offering new insight into how genetic variants can influence epigenetic regulation.

By resolving these complex genomic and epigenomic interactions at single-molecule resolution, Oxford Nanopore sequencing provides a more complete view of the molecular architecture underlying brain function. These insights open the door to more targeted biomarker discovery and therapeutic strategies for neurodegenerative diseases like Alzheimer's disease.

Read the publication (Dec 2024):
nanoporetech.com/diverse-brains-sequencing

Case study 2: integrated methylation and haplotype analysis in cancer with long Oxford Nanopore reads

Cancer is a multifactorial disease characterised by a complex interplay of genomic, epigenomic, and transcriptomic alterations that collectively drive tumorigenesis. Gaining a comprehensive, multiomic view that captures everything from simple mutations like SNVs to complex rearrangements and methylation patterns, with haplotype-level resolution, is essential. In a landmark study, O'Neill *et al*. used Oxford Nanopore sequencing to analyse 189 tumour research samples from 181 patients, including 41 matched tumour-normal pairs[10]. Samples represented 26 cancer types from the Personalized OncoGenomics (POG) cohort.

The researchers were able to phase somatic mutations and resolve complex SVs that had been previously missed by Illumina short-read sequencing. Notably, they also identified human papillomavirus (HPV) integration events; HPV infection is an established driver of cervical cancer and has been implicated in many head-and-neck and anogenital cancers.

As Oxford Nanopore technology can sequence native DNA and directly detect methylation, the team identified that promoter hypermethylation in *BRCA1* and *RAD51C* was a likely driver of homologous recombination deficiency in tumours with no detectable coding mutations — a key finding with potential clinical implications for therapy selection.

The team also found that methylation patterns generated from Oxford Nanopore reads clustered by the tumour tissue of origin regardless of metastatic site. The authors noted that: 'This finding suggests the potential utility of nanopore-derived DNA methylation for detecting or confirming tissue of origin in advanced and metastatic cancers'.

This research underscores the unique power of nanopore sequencing to deliver a comprehensive, multiomic view of cancer genomes — resolving SVs, phasing somatic mutations, and profiling DNA methylation in a single workflow. By uncovering key epigenetic drivers, preserving tissue-of-origin methylation signatures, and detecting viral integration events, the findings position Oxford Nanopore sequencing as a transformative technology both for advancing our mechanistic understanding of cancer and for potential clinical stratification.

Read the publication (Nov 2024):
nanoporetech.com/cancer-cohort-sequencing

9. Billingsley, K.J., Meredith, M., and Daida, K. et al. Long-read sequencing of hundreds of diverse brains provides insight into the impact of structural variation on gene expression and DNA methylation. *bioRxiv* 628723 (2024). DOI: https://doi.org/10.1101/2024.12.16.628723

10. O'Neill, K. et al. Long-read sequencing of an advanced cancer cohort resolves rearrangements, unravels haplotypes, and reveals methylation landscapes. *Cell Genom*. 4(11):100674 (2024). DOI: https://doi.org/10.1016/j.xgen.2024.100674

Oxford Nanopore Technologies

Find out more at
nanoporetech.com/epigenetics

Oxford Nanopore Technologies

phone +44 (0)845 034 7900

email support@nanoporetech.com

oxford-nanopore-technologies

@nanopore

@nanoporetech.com

www.nanoporetech.com

GS_1131(EN)_V5_30Apr2025