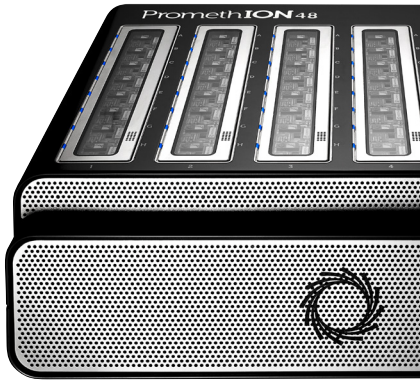
Oxford NANOPORE Technologies

WORKFLOW: EPIGENETICS

Calling methylation in the human genome with PCR-free nanopore sequencing

Methylation plays a fundamental role in regulating gene expression; aberrant methylation patterns are strongly associated with numerous diseases, such as cancer and developmental disorders. The 5-methylcytosine (5mC) modified nucleotide, for example, is an important transcriptional repressor and mediates genomic imprinting.



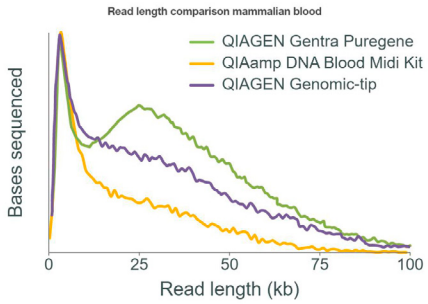
As PCR amplification is unnecessary for nanopore sequencing, base modifications can be called alongside the canonical nucleotide sequence, with no additional sample preparation. Nanopore sequencing also lacks GC bias, facilitating a consistent read depth, and access to genomic regions that may be inaccessible to traditional methods of sequencing and methylation calling.

Here we present a simple workflow for genome-wide methylation calling from a human blood sample, using the PromethION™ platform.

EXTRACTION: obtaining high molecular-weight DNA

Find more extraction protocol recommendations for your sample type, plus guidance on DNA storage and contaminants: **community.nanoporetech.com/docs/prepare**

Selecting the most suitable DNA extraction method depends on your sample type and experimental aim. For DNA extraction from whole blood, we recommend the **QIAGEN Gentra Puregene Blood Kit**, which we have found produces high sequencing throughput and long read lengths, the latter being particularly important if you plan to phase your methylation calls.



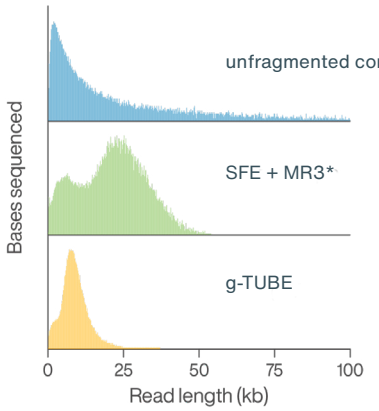
LIBRARY PREPARATION: selecting an approach

Find out more about size selection methods: **community.nanoporetech.com/extraction_method_groups/size-selection**

Read length has little impact on methylation calling alone, without phasing. We therefore recommend optimising library preparation for sequencing throughput, by shearing your extracted DNA with a **Covaris g-TUBE**, for a read length N50 of ~11 kb.

However, if your aim is to phase your data and determine allele-specific methylation, maximising read length is key. For this, we suggest using the Oxford Nanopore **Short Fragment Eliminator Expansion** to size-select for fragments >10 kb, and the **Diagenode Megaruptor 3** for light shearing, for a read length N50 of ≥20 kb.

To prepare gDNA for sequencing, we recommend the **Ligation Sequencing Kit**.



*Oxford Nanopore Short Fragment Eliminator Expansion + Diagenode Megaruptor 3

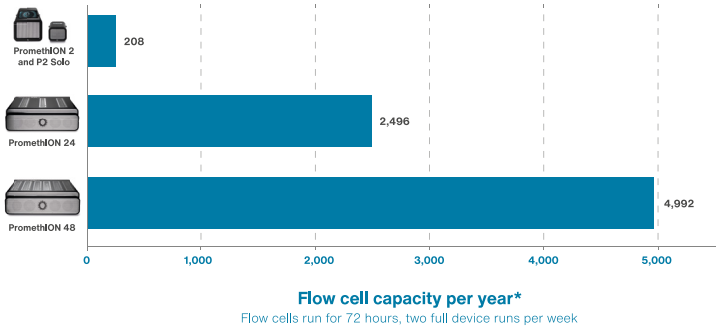
Find out more about library prep, including rapid and multiplexing options: nanoporetech.com/products/kits

SEQUENCING: generating high data yields on the PromethION

Find out more about the Flow Cell Wash Kit:
https://store.nanoporetech.com/flow-cell-wash.html

For genome-wide calling of 5mC, the most studied base modification, we recommend sequencing a human genome to a read depth of 20x. A slightly higher read depth of around 30x is advised for phasing your data. These depths can be achieved by sequencing a single genomic sample on one PromethION Flow Cell for 72 hours.

Throughput can be maximised by using the **Flow Cell Wash Kit** and loading fresh library every 24 hours. For high-coverage nanopore sequencing with low sample processing requirements, we recommend the PromethION 2, which supports two PromethION Flow Cells. For high-throughput sequencing projects, we recommend the higher capacity platforms, PromethION 24 and PromethION 48, which can simultaneously run up to 24 or 48 flow cells, respectively.



Find out more about PromethION: **nanoporetech.com/products/promethion**

ANALYSIS: calling methylation alongside canonical bases

Find out more about data analysis solutions:
nanoporetech.com/analyse

For best performance in calling 5mC methylation in the human genome, we recommend the algorithm **Remora**, which is integrated into Guppy and MinKNOW™ – the software onboard nanopore sequencing devices. Remora models separate canonical basecalling from methylation calling, thus enabling the highest quality canonical and methylation calls from a single run, with minimal computational overhead.

Remora models for the detection of both 5mC and 5hmC are available. For more advanced usage, such as if you wish to train your own models to detect further epigenetic modifications of interest, you can access the Remora repository on GitHub[1].

If you wish to perform phasing, we advise using the Oxford Nanopore **wf-human-snp** pipeline and third-party tool **WhatsHap**[2].

FAST5 input file

Canonical and modified base calling (Guppy/MinKNOW with Remora)

Mapping (minimap2)

Phasing (wf-human-snp + WhatsHap)

Aggregate modified bases (modbam2bed)

Mapped BAM and bedMethyl output files

Find out more at: **nanoporetech.com/epigenetics**

Oxford NANOPORE Technologies

Twitter: @nanopore
www.nanoporetech.com

References:
1. GitHub. Remora. Available at: https://github.com/nanoporetech/remora [accessed: 10 August 2022].
2. Martin, M. et al. WhatsHap: fast and accurate read-based phasing. *BioRxiv.* DOI: https://doi.org/10.1101/085050 (2016).