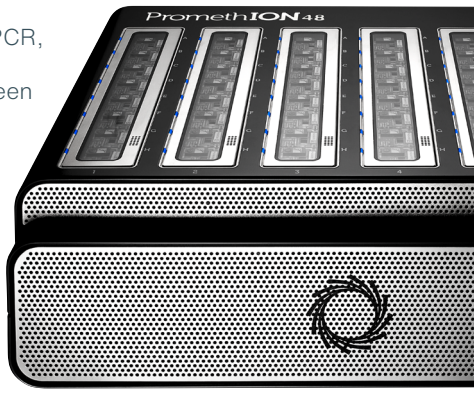
Oxford NANOPORE Technologies

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.

Using traditional short-read sequencing technology, epigenetic modifications are erased during PCR, so must instead be inferred via chemical treatment of DNA such as bisulfite conversion. However, this process can give variable results due to incomplete conversion, and cannot distinguish between methylation variants 5mC and 5-hydroxymethylcytosine (5hmC).



With PCR-free nanopore sequencing of native DNA, methylation can be directly detected at single-nucleotide resolution. Base modifications can be accurately called alongside the canonical nucleotide sequence, with no additional sample preparation, including in genomic regions that may be inaccessible to traditional sequencing methods; furthermore, both 5mC and 5hmC modifications can be called.

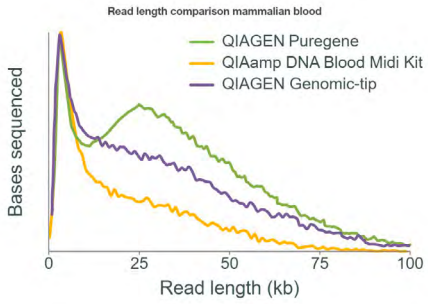
Here we present a simple workflow for genome-wide methylation calling from a human blood sample, using the PromethION™ platform.

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 Puregene Blood Kit***; which we have found produces high sequencing output and long read lengths, the latter being particularly important if you plan to phase your methylation calls.

*Previously named QIAGEN Gentra Puregene Blood Kit



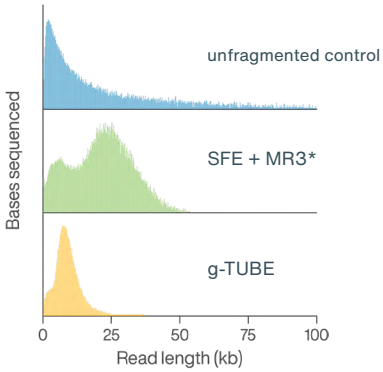
LIBRARY PREPARATION: selecting an approach

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 output, by shearing your extracted DNA with a **Covaris g-TUBE**, for a read length N50 of up to ~10 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

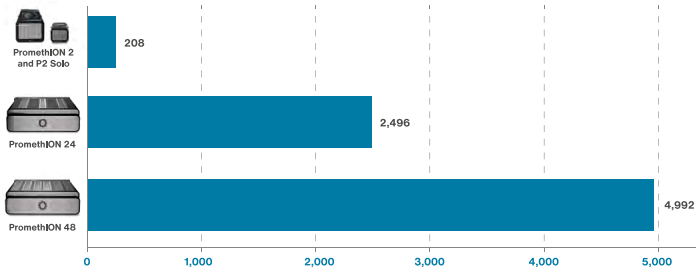
Find out more about library prep, including rapid and multiplexing options: nanoporetech.com/products/kits

SEQUENCING: generating high data outputs on the PromethION

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

For genome-wide calling of 5mC and 5hmC, 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.



Flow cell capacity per year
Flow cells run for 72 hours, two full device runs per week

For high-coverage nanopore sequencing with low sample processing requirements, we recommend the PromethION 2 devices, which support 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/data-analysis

For best performance in calling 5mC methylation in the human genome, we recommend the algorithm **Remora**, which is integrated into 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].

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 allows for the analysis of single nucleotide variants, copy number variants, structural variants, and short tandem repeats. This workflow, an EPI2ME™ solution, can be run with simple point-and-click implementation, or using the command line.

BAM input file

wf-human-variation
Variant and methylation calling

VCF and bedMethyl output files
HTML reports

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: 15 Mar 2023].

WF_1087(EN)_V4_31Mar2023