Oxford NANOPORE Technologies

WORKFLOW: SNV CALLING AND PHASING

Calling and phasing single nucleotide variants in the human genome with long nanopore reads

Compared to short-read technologies, nanopore sequencing does not suffer from GC bias and does not require PCR, enabling wider access to the genome for variant calling, and thereby increasing our understanding of genetic variation in health and disease.

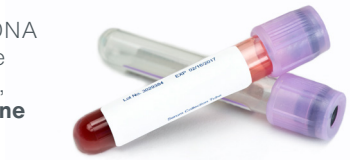
Assigning variants to the maternal or paternal chromosome (phasing) is important for understanding their inheritance and functional impact. Long and ultra-long nanopore sequencing reads greatly enhance phasing, especially of heterozygous variants, as an individual read is more likely to contain multiple single nucleotide variants (SNVs).

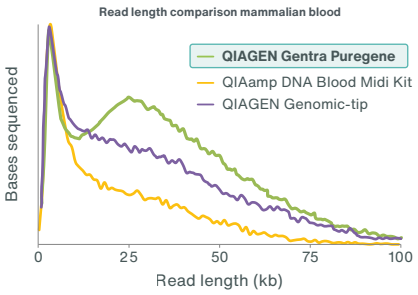
Here we present a simple workflow for calling and phasing SNVs in the human genome.

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**

Long sequencing reads are required for optimal phasing. Obtaining high molecular weight (HMW) DNA from your sample is therefore crucial. Selecting the most suitable HMW DNA extraction method depends on your sample type. For DNA extraction from whole blood, we recommend the **QIAGEN Gentra Puregene Blood Kit**, which we have found maximises the yield of long sequencing reads.



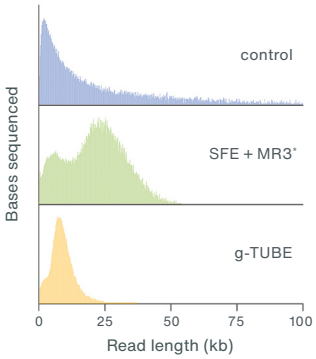


LIBRARY PREPARATION: selecting an approach

Find out more about size selecting for long fragments: **community.nanoporetech.com/extraction_method_groups/size-selection**

We have found that read length has a strong effect on phasing performance. To prepare your extracted gDNA, we advise light shearing and size selecting for >25 kb fragments. We recommend the **Oxford Nanopore Short Fragment Eliminator Expansion** for size selection (to >25 kb), and the **Diagenode Megaruptor 3** for shearing.

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



*Oxford Nanopore Short Fragment Eliminator Expansion + Diagenode Megaruptor 3

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

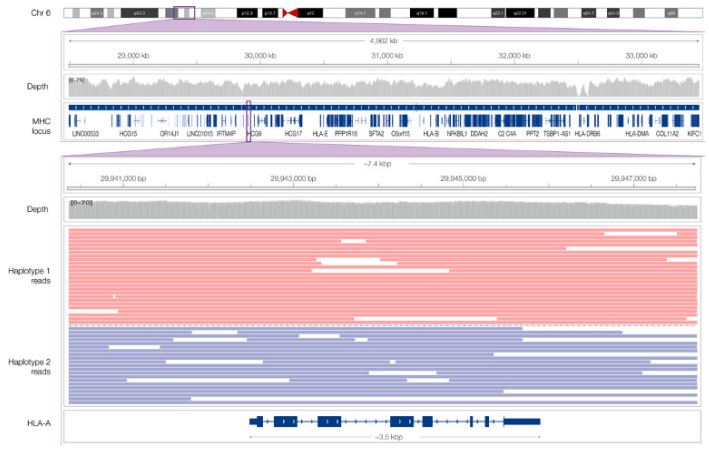
SEQUENCING: generating high data yields of long reads with PromethION

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

For SNV calling, we recommend sequencing a human genome to 40x–60x read depth; this can be achieved by sequencing a single genomic sample on one or two PromethION™ Flow Cells for 72 hours.

Whilst longer reads significantly improve phasing, read depth is less important, with good metrics achieved at 45x and diminishing returns for higher depth.

We recommend basecalling in high accuracy mode with Guppy version 3.6.1 or higher. Throughput can be maximised by using the **Flow Cell Wash Kit** and loading fresh library every 24 hours.



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

ANALYSIS: selecting tools for SNV calling & phasing

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

To call SNVs in the human genome we recommend **Medaka**, developed and supported by Oxford Nanopore, and with easy installation. Analysis runtime for this workflow is <48 hours using a single PromethION or GridION™ GPU and 8 CPU cores.

Alternatively, we have found that the third-party **PEPPER/DeepVariant**[1] workflow provides the best SNV calling metrics for datasets with read depths <60x, basecalled with Guppy 3.6.1 and above.

For Phasing, we advise using the Oxford Nanopore **wf-human-snp** pipeline and third-party tool **WhatsHap**[2]. WhatsHap is integrated into Medaka, so this additional step is unnecessary when using Medaka for SNV calling.

FASTQ input file

Alignment	minimap2
SNV calling	DeepVariant or Medaka
Phasing	WhatsHap + wf-human-snp

VCF output file

Find out more at: **nanoporetech.com/applications/investigation/snvs-phasing**

Oxford NANOPORE Technologies

Twitter: @nanopore
www.nanoporetech.com

References:
1. Shafin, K. PEPPER. Software. Available at: https://github.com/kishwarshafin/pepper/ [Accessed: 12 September 2022].
2. Martin, M. WhatsHap: fast and accurate read-based phasing. *bioRxiv*. DOI: https://doi.org/10.1101/085050 (2020).

WF_1085(EN)_V4_19Sept2022