Oxford NANOPORE Technologies

WORKFLOW: STRUCTURAL VARIATION

Resolving structural variants with long nanopore sequencing reads

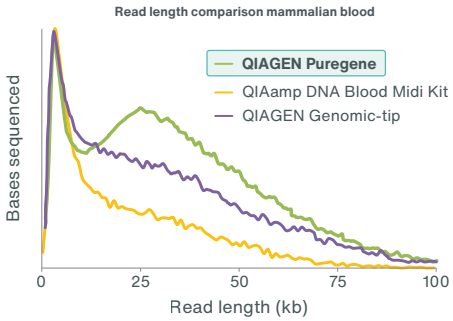
Structural variants (SVs) — genomic aberrations >50 bp — have known causative effects in an extensive range of both normal and aberrant phenotypes. The need to comprehensively characterise SVs is becoming increasingly clear; however, they are difficult to detect using short-read sequencing as the DNA must be both fragmented and amplified, limiting detection according to their size, complexity, and position in the genome. Using PCR-free nanopore sequencing, there is no limit to read length — long nanopore reads can span entire SVs in a single read, including within repetitive or GC-rich regions. High-output PromethION™ devices enable accurate resolution of even highly complex variants in any genomic context.

Here, we present a simple workflow for an effective whole-genome SV survey from a human clinical research blood sample, using the PromethION sequencing device range.

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 a suitable extraction method is often a trade-off between input requirements, expected fragment lengths, lab experience and hands-on time. To maximise the output of long reads in sequencing, it is important to select an extraction method that preserves high molecular-weight DNA. When starting from clinical research blood samples, we recommend the **QIAGEN Puregene Blood Kit***.

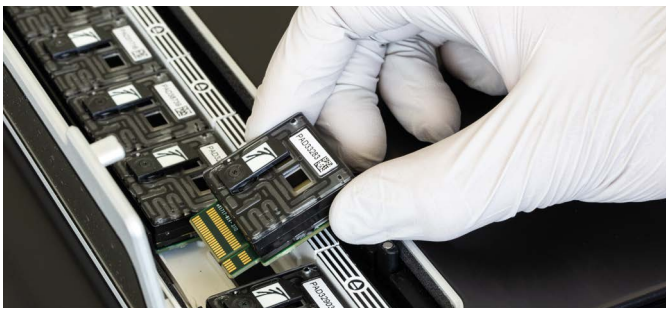


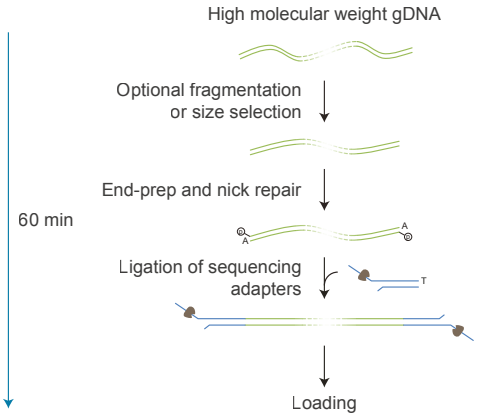
*Previously named QIAGEN Gentra Puregene Blood Kit

LIBRARY PREPARATION: selecting a kit

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

There is no upper read length limit in nanopore sequencing, with native DNA sequencing reads routinely spanning tens or hundreds of kilobases and the current record spanning over 4 megabases[1]. Fragmentation is optional: unfragmented DNA offers a simple workflow, but shearing and size selection can improve read N50. We recommend the Oxford Nanopore **Short Fragment Eliminator Expansion** to size select for fragments >25 kb, and the **Diagenode Megaruptor 3** for shearing.





To prepare native gDNA for sequencing, we recommend the **Ligation Sequencing Kit**, providing the greatest output and control over read lengths.

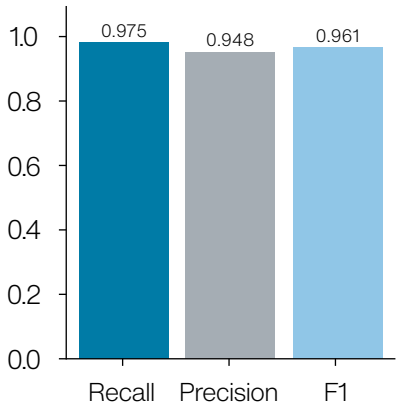
Find out more about library prep, including rapid and multiplexing options: **nanoporetech.com/prepare**

SEQUENCING: generating high outputs on the PromethION

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

To generate high coverage across the human genome, we recommend sequencing on a high-output PromethION Flow Cell. The PromethION device range features the powerful, benchtop PromethION 24 and PromethION 48 – configured for sequencing up to 24 or 48 flow cells – whilst the compact PromethION 2 devices provide the flexibility of two independent flow cells for lower sample throughput requirements. We recommend basecalling using high accuracy (HAC) mode using MinKNOW™ — the software onboard nanopore sequencing devices. For SV detection, we recommend sequencing to 30x depth of coverage on a single PromethION Flow Cell; output can be maximised by washing the flow cell using the **Flow Cell Wash Kit**.

Detection of SVs in the human genome at 30x depth of coverage



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

ANALYSIS: accurate SV calling for all levels of expertise

View the Genome in a Bottle open dataset:
labs.epi2me.io/giab-2023.05/

To call SVs and other variants in your nanopore sequence data, we recommend **wf-human-variation**, a comprehensive analysis solution which can be run as a fully automated, point-and-click EPI2ME™ workflow, or using the command line.

Using files produced by onboard basecalling, the workflow aligns to a provided reference and calls SVs using the analysis tool Sniffles2[2], outputting a VCF file of called sequence variants and an intuitive HTML report. Epigenetic variants — 5mC and 5hmC — can also be detected using the Modkit tool within this workflow, outputting a bedMethyl file, allowing the simultaneous analysis of SVs and their methylation state.

This analysis workflow also enables simultaneous calling of single nucleotide variants, copy number variants, and short tandem repeats. For best-practice workflows on the analysis of these variants across the human genome, visit **nanoporetech.com/resource-centre**

POD5 or BAM input file

wf-human-variation
Variant and methylation calling

VCF and bedMethyl files plus HTML report

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

Find out more at: **nanoporetech.com/structural-variation**

Oxford NANOPORE Technologies

Twitter: @nanopore
www.nanoporetech.com

References:
1. Oxford Nanopore Technologies. Ultra-Long DNA Sequencing Kit. Available at: https://store.nanoporetech.com/ultra-long-dna-sequencing-kit-v14.html [Accessed: 03 November 2023]
2. Smolka, M. et al. *BioRxiv* 2022.04.04.487055. DOI: https://doi.org/10.1101/2022.04.04.487055

WF_1042(EN)_V5_07Nov2023