Oxford NANOPORE Technologies

WORKFLOW: STRUCTURAL VARIATION

Resolving structural variants with long nanopore sequencing reads

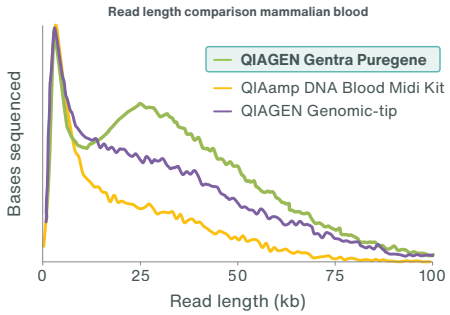
Structural variants (SVs), defined as variants spanning 50 bp or more, account for ten times as many variant bases as single nucleotide polymorphisms (SNPs) in the human genome[1]. With known causative effects in an extensive range of both normal and aberrant phenotypes, the need to comprehensively characterise SVs is becoming increasingly clear. Long native DNA reads produced by nanopore sequencing devices greatly improve the accuracy of detection of even the largest of SVs, including those regions inaccessible to other technologies[2].

Here, we present a simple workflow for an effective whole-genome SV survey from a human 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**

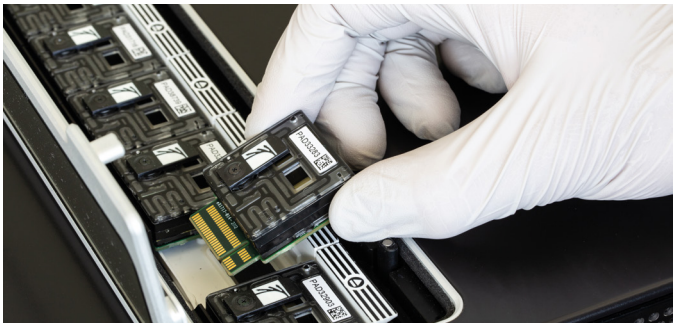
Selecting a suitable extraction method is often a trade-off between input requirements, expected fragment lengths, lab experience and hands-on time. To maximise the volume of data at long read lengths, we recommend the **QIAGEN Gentra Puregene Blood Kit**.

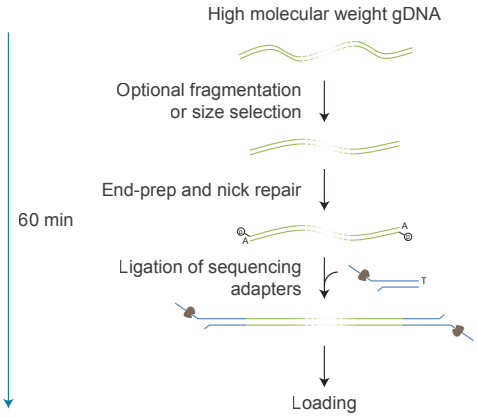


LIBRARY PREPARATION: selecting a kit

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 reads routinely spanning tens or hundreds of kilobases and the current record spanning over 4 megabases. 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.



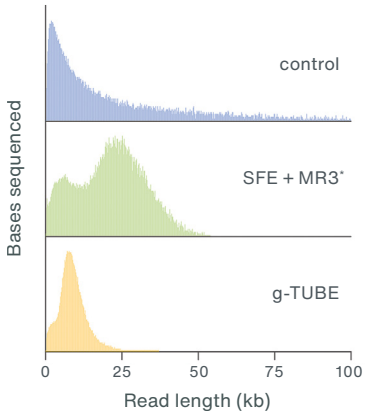


To prepare gDNA for sequencing, we recommend the **Ligation Sequencing Kit**, providing the greatest throughput and control over read lengths.

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

SEQUENCING: generating high yields on the PromethION

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



*Short Fragment Eliminator Kit + Diagenode Megaruptor 3

To maximise throughput, we recommend sequencing on a PromethION Flow Cell. The PromethION device range features the powerful, benchtop P24 and P48 – configured for sequencing up to 24 or 48 PromethION Flow Cells – whilst the compact P2 devices provide the flexibility of two independent, high-yield PromethION Flow Cells for lower sample throughput requirements. 30x coverage can be obtained by sequencing on a single PromethION Flow Cell for 72 hours; throughput is maximised by a nuclease flush using the **Flow Cell Wash Kit** and addition of fresh library every 24 hrs.

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

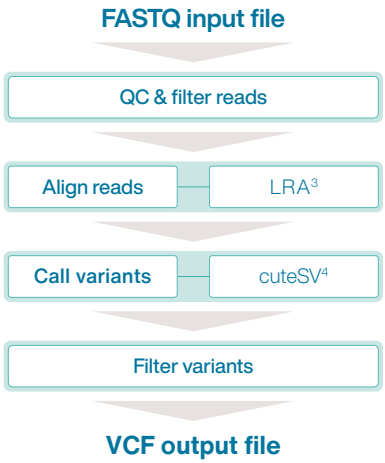
ANALYSIS: calling SVs without command line

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

To call variants in your nanopore sequence data, we recommend **wf-human-sv**, available on GitHub. The pipeline takes the FASTQ files produced by onboard basecalling, aligns to a provided FASTA reference genome, calls insertions, deletions and duplications >50 bp and outputs a VCF file of called sequence variants and a QC report.

The analysis workflow is also available in an interactive, guided tutorial format via EPI2ME Labs. **View the tutorial here: labs.epi2me.io/nbindex/**

For those wishing to avoid command line, our cloud-based EPI2ME™ SV workflow provides fully automated human whole-genome SV analysis generating a VCF file.



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

Oxford NANOPORE Technologies

Twitter: @nanopore
www.nanoporetech.com

References:
1. Eichler, E.E. Genetic Variation, Comparative Genomics, and the Diagnosis of Disease. N. Engl. J. Med. 4;381(1):64-74 (2019).
2. Ebbert, M.T.W. et al. Systematic analysis of dark and camouflaged genes reveals disease-relevant genes hiding in plain sight. Genome Biol. 20(1):97 (2019).
3. Chaisson, M. et al. GitHub: LRA [Online]. Available at: github.com/ChaissonLab/LRA [Accessed: 12 September 2022].
4. Jiang, T. et al. Long-read-based human genomic structural variation detection with cuteSV. Genome Biol. 21:189 (2020).