Characterising structural variants in FFPE cancer research samples

Formalin-fixed, paraffin-embedded (FFPE) sampling is a widely used, simple method to preserve important cancer research samples. Structural variation (SV) is a key mutational process in cancer, defined as genomic alterations >50 bp in length, and can impact whole genes or chromosomes. Formalin fixing damages DNA molecules and reduces their fragment length; however, sequences spanning several kilobases still remain. Using traditional short-read sequencing technology, these long sequences must be further fragmented and crucial SV information is therefore lost, resulting in incorrect or incomplete identification of SVs.

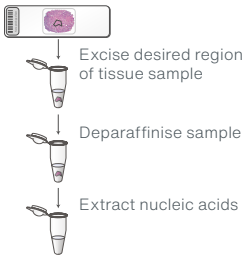
Oxford Nanopore has optimised a workflow specifically designed to extract and sequence DNA spanning several kilobases from FFPE samples, enabling accurate SV detection and analysis. With long nanopore sequencing reads, SVs can be comprehensively characterised across cancer genomes.

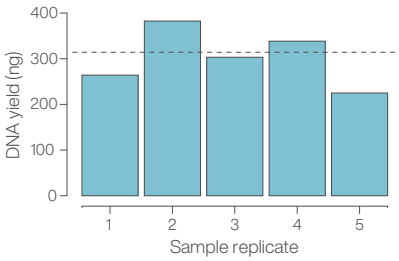
Here, we present a simple workflow for the analysis of SVs in DNA extracted from FFPE samples using a PromethION™ Flow Cell.

EXTRACTION: isolating, quantifying, and assessing sample integrity

Sample age and preparation variability will impact both the quality and degree of fragmentation of DNA extracted from FFPE samples; the Oxford Nanopore workflow is carefully designed to minimise additional fragmentation. Beginning with removal of the waxy paraffin layer from the sample, the workflow continues with extraction of the DNA using the QIAGEN DNA FFPE Tissue Kit. We recommend assessing DNA yield using a Qubit instrument and sample quality via a Nanodrop instrument before proceeding to library preparation.

View the FFPE human cell DNA extraction workflow for detailed guidance:
community.nanoporetech.com/extraction_methods





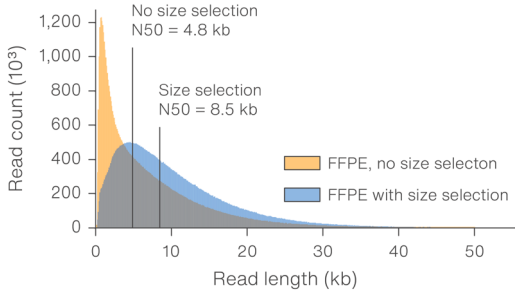
FFPE DNA extracted from a reference-grade sample

LIBRARY PREPARATION: selecting the right sequencing kit for your sample

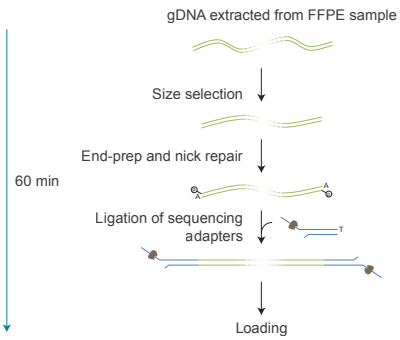
Find detailed guidance on SPRI size selection of DNA:
community.nanoporetech.com/extraction_methods/spri-size-selection

Further DNA fragmentation is not required when preparing the FFPE-extracted gDNA for nanopore sequencing. To ensure the long molecules containing SV information are enriched, we recommend selecting for DNA fragments above ~1.5–2 kb using SPRI bead-based size selection prior to library preparation.

The **Ligation Sequencing Kit** (see figure below), a PCR-free library preparation approach, can then be used to maximise sequencing yield. However, if the FPPE sample is very degraded, best results can be obtained by preparing samples using the **PCR Sequencing Kit**.



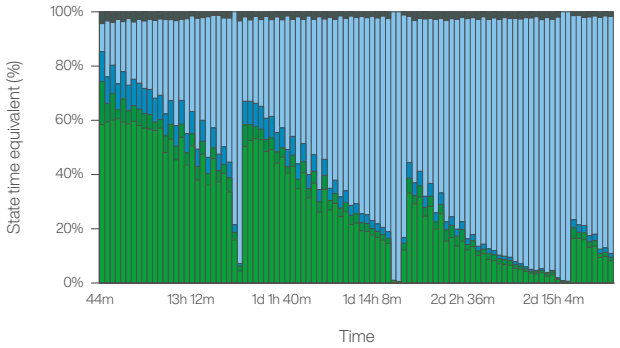
Read length distributions obtained when sequencing reference-grade FFPE samples, prepared with and without size selection



Find out more about library preparation kits, including low-input options: store.nanoporetech.com/sample-prep.html

SEQUENCING: utilising high-output PromethION Flow Cells

Learn more about the Flow Cell Wash Kit:
store.nanoporetech.com/flow-cell-wash-kit-r9.html



For best SV calling metrics across the human genome, we recommend sequencing to ≥30x depth of coverage, though this can be reduced to 15x for a lower-pass whole-genome survey. As the quality of the initial FFPE sample preservation can influence sequencing yield, you may wish to quality check your library by sequencing a small amount on a Flongle Flow Cell before proceeding. These smaller, low-cost flow cells can be used with a GridION or MinION device, and checked in real-time to indicate sequencing performance before moving to higher-depth sequencing.

To maximise output, 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, with high-performance integrated compute — whilst the compact P2 devices provide the flexibility of two independent, high-yield PromethION Flow Cells for lower sample throughput requirements.

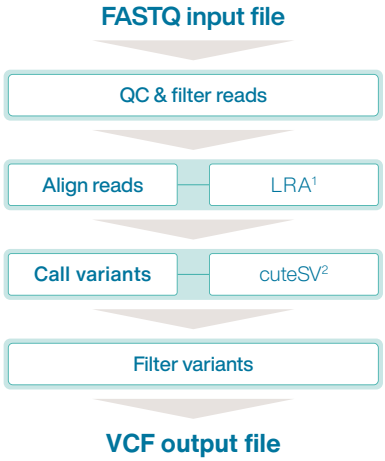
Find out more about PromethION and Flongle: nanoporetech.com/products

ANALYSIS: detecting SVs from FFPE sample sequencing data

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

To call germline SVs in your nanopore sequence data, we recommend *wf-human-sv*. Available on GitHub, this pipeline takes the FASTQ files generated during sequencing, aligns to a FASTA reference genome using LRA[1], and calls insertions, deletions, and duplications of >50 bp using cuteSV[2]. The pipeline outputs a file of called variants (VCF) and a QC report. This workflow is also available as an interactive tutorial via EPI2ME Labs and, for those wishing to avoid the command line, as a fully automated, cloud-based workflow in EPI2ME.

If you are studying somatic SVs in a cancer research sample, we recommend the tool Sniffles2[3], using 'somatic' mode. For paired tumour:normal clinical research sample datasets, somatic variants can be analysed using nanomonsv[4].



View nanopore data analysis tools and pipelines on GitHub: github.com/nanoporetech

Find out more at: **nanoporetech.com/cancer-research**

References:

1. Chaisson, M. et al. GitHub: LRA [Online]. Available at: github.com/ChaissonLab/LRA [Accessed: 21 June 2022].
2. Jiang, T. et al. Long-read-based human genomic structural variation detection with cuteSV. Genome Biol. 21:189 (2020).
3. Smolka, M. et al. Comprehensive structural variant detection: from mosaic to population-level. bioRxiv. doi: https://doi.org/10.1101/2022.04.04.487055 (2022).
4. Shiraishi, Y. et al. Precise characterization of somatic structural variations and mobile element insertions from paired long-read sequencing data with nanomonsv. bioRxiv. doi: https://doi.org/10.1101/2020.07.22.214262 (2021).

Oxford NANOPORE Technologies

Twitter: @nanopore
www.nanoporetech.com