Oxford NANOPORE Technologies

WORKFLOW: TUMOUR–NORMAL

Characterising genomic and epigenomic variation between tumour-normal research samples using long nanopore sequencing reads

Genomic instability is characteristic of most cancers. Paired tumour-normal whole-genome sequencing enables a deeper understanding of genomic and epigenomic variability in cancer, which will prompt the discovery of new cancer biomarkers and new insights into the genetics of treatment-resistant tumours.

Using traditional sequencing technology, short DNA fragments must undergo PCR — erasing all epigenetic information, introducing bias, and limiting variant detection only to regions amenable to amplification. Furthermore, complex structural variants (SVs), which can reach megabase scale, cannot be spanned by short reads and may be missed.

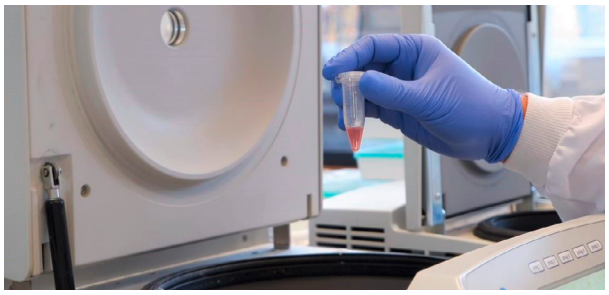
Capturing a wide range of tumour-specific variation within a single sequencing assay has the potential to enhance the classification accuracy of tumour types and the identification of driver mutations involved in cancer progression. With nanopore sequencing, long, native DNA reads capture single nucleotide variants (SNVs), SVs, copy number variants (CNVs), short tandem repeats (STRs), and epigenetic modifications including both 5mC and 5hmC in a single dataset. The long reads span complex and repetitive regions, thus simplifying phasing. Variants can be confidently assigned to the maternal or paternal chromosome, providing deeper resolution for the molecular characterisation of cancer research samples.

Here we present an end-to-end workflow to detect somatic variation between tumour-normal paired research samples using the PromethION™ sequencing device range.

EXTRACTION: obtaining high molecular-weight DNA

Find extraction protocol guidance:
community.nanoporetech.com/docs/prepare

Selecting a suitable extraction method to obtain high molecular-weight (HMW) DNA greatly depends on your sample type. There are a range of protocols available in the Community covering human clinical research samples, including brain tissue and blood. We highly recommend assessing DNA yield using a Qubit instrument, sample quality via a Nanodrop instrument, and DNA fragment length distribution using a fragment analyser before proceeding to library preparation.

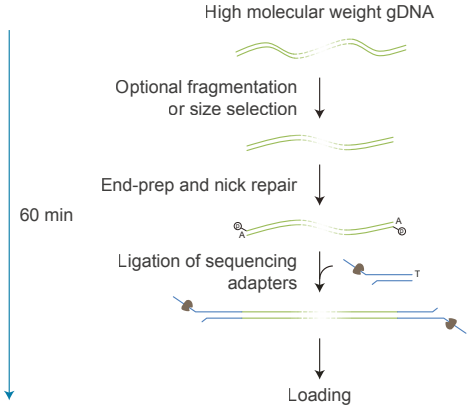


LIBRARY PREPARATION: selecting the right sequencing kit for your samples

Find out more about size selection methods:
community.nanoporetech.com/extraction_method_groups/size-selection

When starting with HMW DNA, shearing and size selection can improve read length N50. We recommend the Oxford Nanopore **Short Fragment Eliminator Expansion** to size select for longer fragments and the **Diagenode Megaruptor 3** for light shearing. However, if it is not possible to isolate HMW DNA, or sample input amounts are limited, you can instead proceed straight to library preparation without size selection or fragmentation.

To prepare HMW DNA libraries for sequencing, we recommend the **Ligation Sequencing Kit**. Offering the greatest control over read length and output, this PCR-free library preparation method also preserves base modifications in the native DNA.

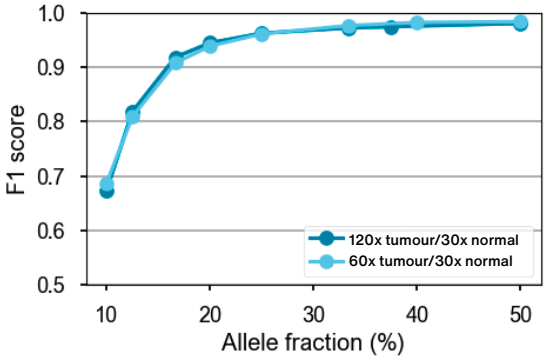


Find out more about library prep: nanoporetech.com/library-preparation

SEQUENCING: utilising high-output PromethION Flow Cells

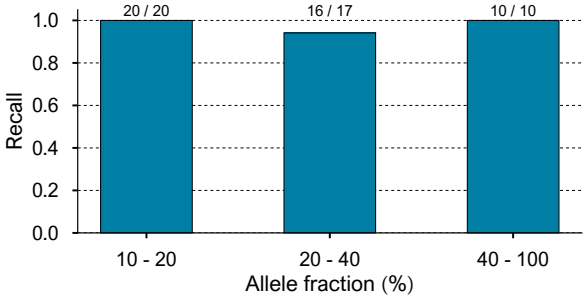
For high-output sequencing, we recommend the powerful PromethION device range. This features the benchtop PromethION 24 and PromethION 48 — configured for sequencing on up to 24 or 48 PromethION Flow Cells — and the compact PromethION 2 devices, which enable sequencing on up to two flow cells, for lower sample throughput requirements.

Detection of SNVs with different allele frequencies



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

Detection of SVs with different allele frequencies (60x tumour/30x normal)



To characterise SVs, SNVs, and methylation, we recommend sequencing normal tissue research samples to 30x depth of coverage and paired tumour research samples to 60x. We recommend basecalling using super accuracy mode.

Output can be maximised by washing the flow cell using the **Flow Cell Wash Kit** and loading fresh library every 24 hours.

You may wish to quality check your library by sequencing a small amount on a Flongle™ Flow Cell — 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.

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

ANALYSIS: detecting somatic variants between tumour-normal samples

To identify somatic variants in the cancer genome via tumour-normal sequencing, we recommend using the analysis workflow **wf-somatic-variation** — an EPI2ME™ solution, which can be run with simple point-and-click implementation or using the command line. This workflow takes the BAM file produced by onboard basecalling, aligns to a provided reference genome, then calls SNVs and small indels between paired samples using ClairS[1], and SVs (>50 bp) using nanomonsv[2]. Methylation analysis of 5mC and 5hmC is performed using Modkit. The total time taken for somatic variant data analysis is between 6.5–10 hours.

View the tumour-normal open dataset:
labs.epi2me.io/colo-2023.05/

BAM input file

wf-somatic variation
SNV, indel, and SV analysis
5mC and 5hmC
methylation analysis

VCF output files and HTML report
bedMethyl output file

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

To find out more about tumour-normal nanopore sequencing, visit: **nanoporetech.com/cancer-research**

Oxford NANOPORE Technologies

References:
1. Luo, R. et al. *Nat. Mach. Intell.* 2: 220-227 (2022). DOI: https://doi.org/10.1038/s42256-020-0167-4
   Shiraishi, Y. et al. *Nucleic Acids Res.* gkad526 (2023). DOI: https://doi.org/10.1093/nar/gkad526
2. GitHub. Modkit. Available at: https://github.com/nanoporetech/modkit [Accessed: 04 Aug 2023]

WF_1200(EN)_V1_04Aug2023

Twitter: @nanopore
www.nanoporetech.com