Duplex Xdrop® enrichment: Enrichment of staggered long DNA fragments for complete deep sequencing of BRCA2

Background

To better elucidate the steps of the **Xdrop®** workflow and how these are connected, we enriched DNA to sequence a well-known gene and highlight the outcome expectations at various milestones from sample preparation to read mapping.

BRCA2 is a human tumor suppressor gene located on chromosome 13. The protein it codes for is responsible for DNA repair. BRCA2 is involved in preventing genomic rearrangements that can be oncogenic, and many of the mutations identified to date in this gene correlate with increased risk of cancer. For this reason, BRCA2 is extremely well characterized and numerous enrichment methods specifically target the gene for subsequent sequencing. Our goal in this Technical Note is not to compare performance, but to highlight features of the Xdrop® workflow, show data output for specific milestones, and facilitate creative applications of this simple but powerful technology.

The Xdrop® workflow

The Xdrop® technology combines high-resolution droplet PCR (dPCR) with droplet sorting and multiple displacement amplification in droplets (dMDA) to enrich specific long DNA fragments (~100 kb) in a sample.

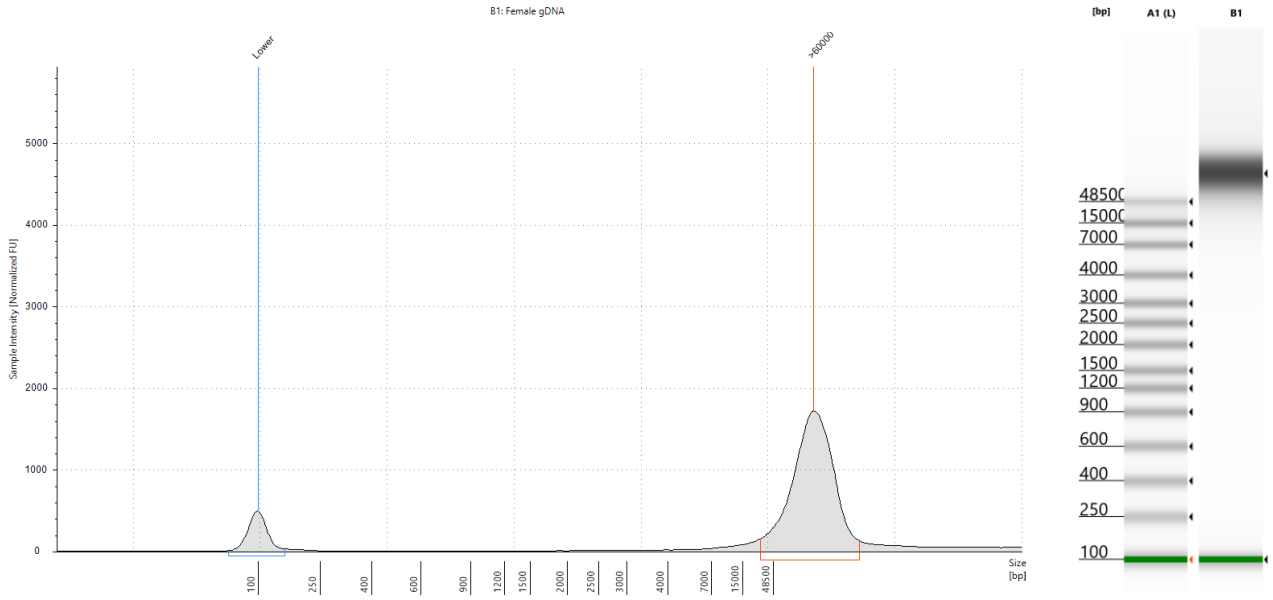
Xdrop® partitions the DNA into millions of double emulsion droplets. During the droplet PCR, droplets containing the target DNA molecules accumulate amplicons of a 120-160 bp Detection Sequence within or adjacent to a region of interest.

As a result of the amplified Detection Sequence, positive droplets fluoresce and can be sorted by standard cell sorter. The sorted long DNA fragments are released and then re-encapsulated individually for unbiased droplet multiple displacement amplification (dMDA).

The resulting enriched and amplified long fragments containing the region of interest are compatible with both long- and short-read library preparation and sequencing.

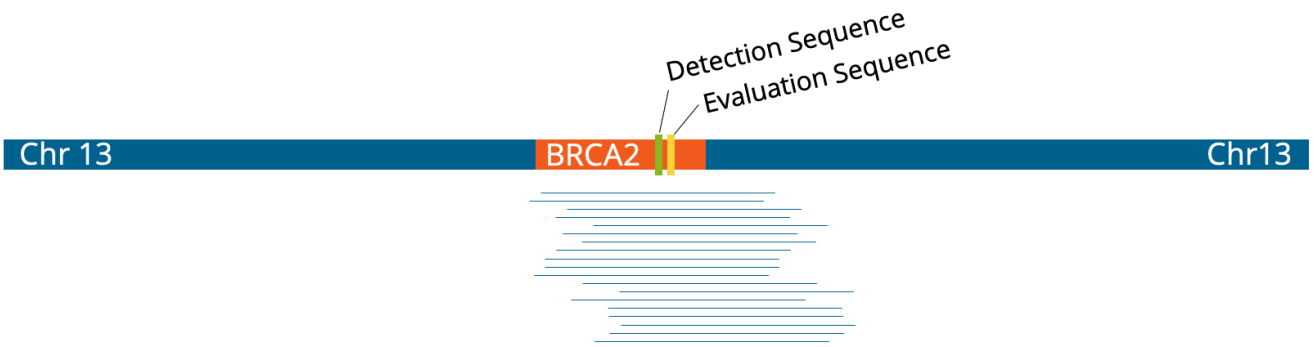
Milestone 1: DNA of high molecular weight

To enrich long fragments of DNA, it stands to reason that samples must include DNA molecules of high molecular weight. For our BRCA2 case study, we started with human female genomic DNA (Reg. G1521, Promega®) extracted from blood of 6–7 anonymous donors between the ages of 21 and 65. The bulk of the DNA molecules were >60 kb long at a concentration of 82.2 ng/µl, as measured on an Agilent® TapeStation system.



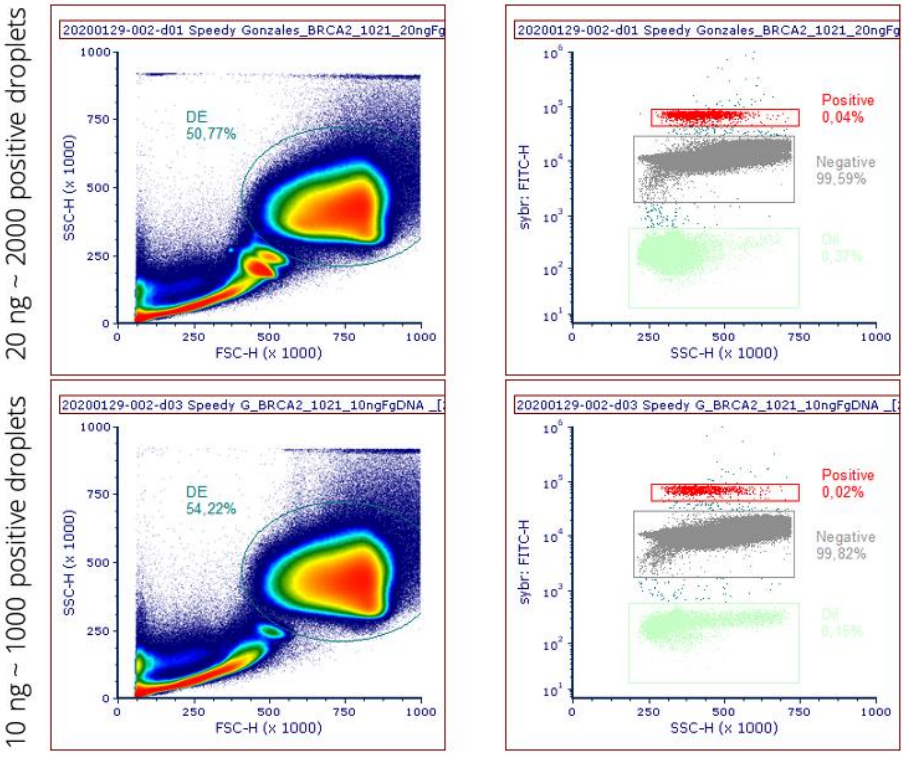
Milestone 2: Primer design and selection

The primers used in the Xdrop® workflow serve the purpose of identifying droplets containing DNA fragments that include the region of interest. They are not used to amplify targets. Design is straightforward with the Samplix primer design tool. The primer design was ready in a few minutes and we chose a primer set for the Detection Sequence (used in the droplet PCR) and another one for the Validation Sequence, to check the enrichment success by qPCR.



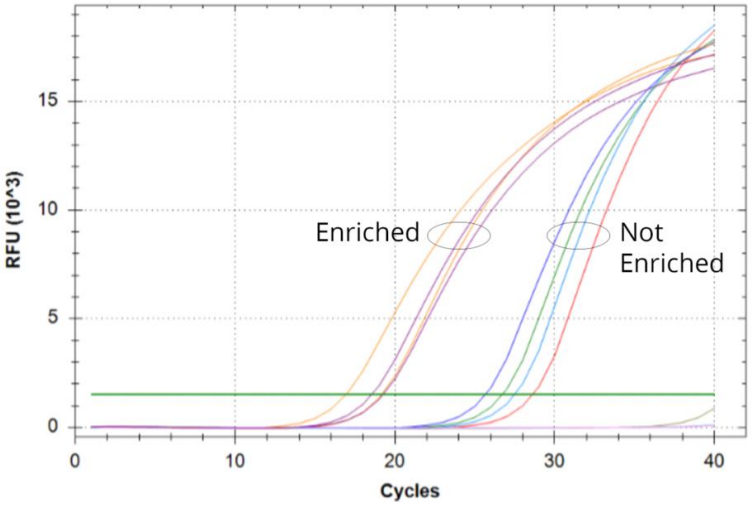
Milestone 3: Droplet sorting

We ran 2 input amounts in duplicate for Xdrop® dPCR: 10 and 20 ng of DNA. After running the dPCR and staining the droplets, we could easily distinguish and collect the positive droplets by gating on the higher fluorescence signal. The two samples with 20 ng input resulted in ~2000 positive droplets while the 10 ng samples gave half that amount (~1000), as expected. We released the target fragments from the sorted droplets and then re-encapsulated them for amplification via droplet MDA.



Milestone 4: Enrichment assessment

After droplet MDA, we estimated the enrichment by qPCR using the initially designed qPCR primers that were positioned approximately 50 bp upstream of the Detection Sequence. The enriched outputs of each input amount replicate (10 ng and 20 ng DNA) were run side-by-side with a standard curve consisting of diluted original unenriched sample (40, 20, 10 and 5 ng). We calculated enrichment from the results using the Samplix online enrichment predictor. Enrichment ranged roughly between 500 and 3500-fold, with an average of 1400-fold.

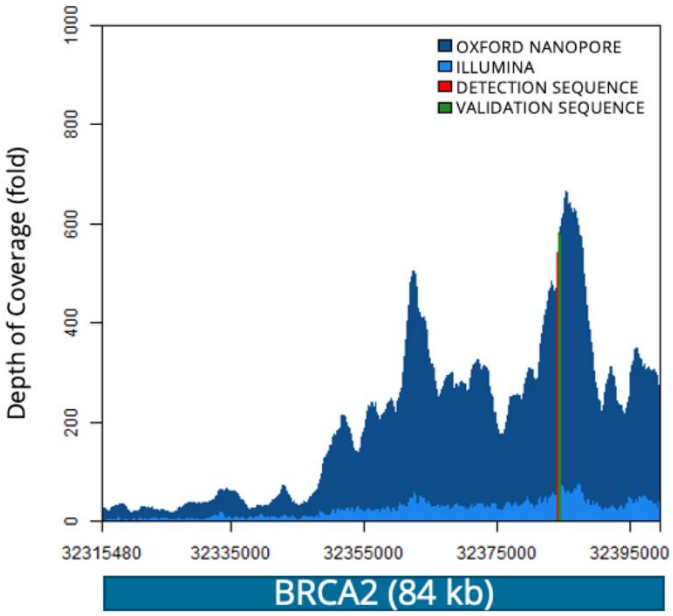


	Input HMW DNA	Enrichment estimation
	20 ng	554-fold
	20 ng	3586-fold
	10 ng	992-fold
	10 ng	521-fold

Milestone 5: Sequencing and target coverage

We pooled the 4 enriched samples and prepared libraries for short-read sequencing on an Illumina iSeq instrument and long-read sequencing on the MinION from Oxford Nanopore Technology. Read mapping showed high coverage of a 10 kb (RoT10) and a 100 kb (RoT100) region around the Detection Sequence with both sequencing methods. Using the sequencing data, we calculated enrichment as the ratio of the percent reads on target to the percent of the genome length attributed to the target region.

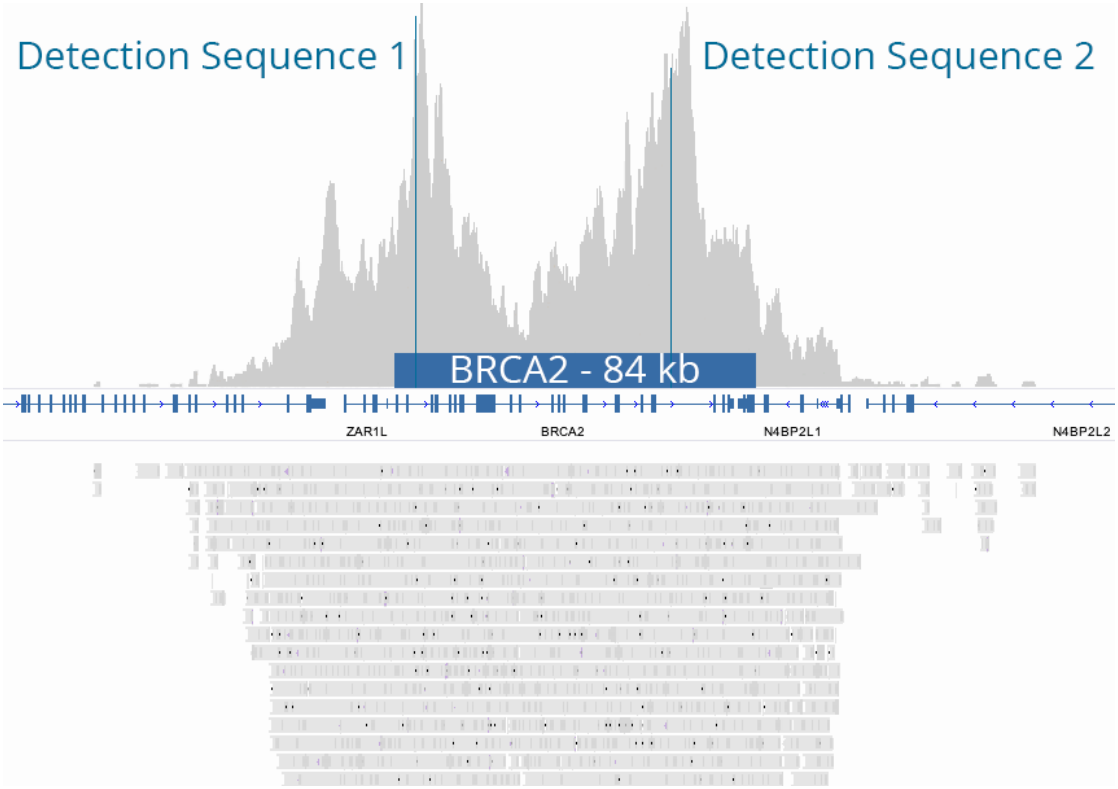
However, because of the length of BRCA2 (~84 kb) and the positioning of the Detection Sequence, the resulting enrichment had low coverage of the distant upstream portion of the gene.



Technology	Data	On-target reads	Enrichment
Oxford nanopore	4 Gb	6477	262x
Illumina	0.3 Gb	11267	240x

Milestone 6: Expanding enrichment range by multiplexing

To expand the range of Xdrop™ enrichment, we designed primers for a second Detection Sequence and repeated the workflow above with a duplex droplet PCR.The result was an expanded set of positive droplets with long DNA fragments containing Detection Sequence 1, Detection Sequence 2, or both. By enriching these staggered fragments, our sequencing reads spanned all of BRCA2.



Conclusions

Xdrop™ empowers genomic applications with a simple design. Workflow steps generate easy-to-interpret data and Indirect Sequence Capture features the flexibility to concatenate enrichment targets and explore regions beyond 100 kb. With creativity, Xdrop™ opens new avenues of investigation previously limited by complicated, time-consuming and expensive methods.