Low-copy number variant detection using the xGen™ cfDNA & FFPE DNA Library Prep Kit and G4™ Platform

Targeted sequencing of DNA libraries generated with the IDT xGen cfDNA & FFPE DNA Library Prep Kit on Singular Genomics G4 Sequencing Platform enables highly precise variant detection across a range of allele frequencies

G4 Platform compatibility

Confident variant calling using UMIs

High-efficiency library prep and enrichment workflow

xGen cfDNA AND FFPE DNA LIBRARIES ARE COMPATIBLE WITH THE G4 SEQUENCING PLATFORM

The Singular Genomics G4 is an innovative benchtop sequencer that leverages a 4-color rapid sequencing by synthesis (SBS) chemistry with advanced optics and fluidics engineering to provide single-day turnaround times across all applications. Combining the G4 with the xGen cfDNA & FFPE DNA Library Prep Kit provides the combined benefits of a highly efficient library preparation workflow and fast, flexible sequencing. Prepared libraries contain Unique Molecular Identifiers (UMIs) and can be used in an array of applications, including low frequency variant detection, when paired with xGen Custom Universal Blocking* Oligos in an xGen Hybridization Capture workflow.

BENEFITS

1. Seamless and efficient library preparation using the xGen cfDNA & FFPE Library Prep Kit with Singular Genomics Dual Index PCR Primers.
2. xGen Custom Universal Blocking Oligos with xGen Hybridization Capture workflow help to ensure a high on-target rate to enable deeper coverage of targeted variants.
3. Fixed UMI sequences help to enable reliable variant detection with as few as 15 copies of a rare allele.

HIGH-QUALITY TARGET ENRICHMENT METRICS USING A RANGE OF INPUTS

The xGen cfDNA & FFPE DNA Library Prep Kit is designed specifically for use with degraded samples, such as cell-free DNA (cfDNA) or DNA extracted from formalin-fixed paraffin-embedded (FFPE) samples. The workflow features a proprietary single-strand ligation strategy that maximizes conversion, suppresses adapter-dimer formation, and reduces chimera rates, delivering higher conversion and library complexity than conventional TA ligation-based methods. By integrating Singular Genomics Index Primers into the xGen cfDNA & FFPE DNA Library Preparation workflow and using hybridization capture for target enrichment, researchers can generate deep coverage of high-complexity libraries from low input samples. This comprehensive workflow delivers mapping rates of ≥99.9% for all sample inputs tested and shows a high level of consistency between replicates and sample variations as seen by low standard deviations of ≤0.01% (data not shown).

Uniform, deep sequencing coverage was attained ranging from low hundreds to 5700X, with ≥95.1% of target bases being covered at 0.5–1.5X of the overall mean coverage (Table 1). Consistently high on-target percentages reflect the specificity of the hybridization capture workflow; fold-80 scores were always at, or below, 1.27 which further highlights the quality of the library and sequencing results generated during this workflow (Figure 1).

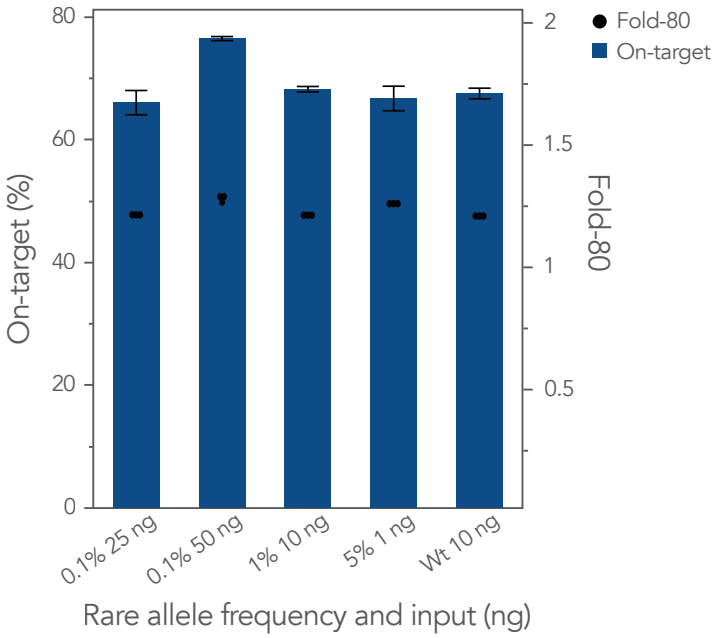

Figure 1. High-quality sequencing metrics from a range of sample inputs. cfDNA reference standards (Horizon, HD780) containing 8 SNVs/SNPs at 5%, 1%, 0.1%, and 0% (WT) allelic frequencies were used to generate xGen cfDNA & FFPE DNA libraries indexed with Singular Genomics Index Primers (*n* =3/sample). An Onco-Hotspot xGen Custom Hyb Cap Panel Accel was designed to capture known mutations in the Horizon standards along with other cancer-relevant somatic mutations. The design approach used a 2X tiling across each exon containing known mutations, resulting in 129 probes covering 9.5 kb of target space. Libraries were captured in 3-plex (*n* = 5) following the xGen Hybridization Capture protocol using xGen Custom Universal Blocking Oligos and amplification primers designed for use with Singular Genomics libraries. Sequencing was performed using 2 x 150 paired-end reads on the G4 Sequencing Platform, then data was subsampled l to 20 million reads per library. Analysis was performed using Picard with dupMarked and hg38 as a reference genome. (Values in the chart represent the mean of the three replicates and error bars represent the standard deviation).

Table 1. High-quality sequencing metrics from a range of sample inputs.

Sample	Mass into library prep (ng)	Estimated rare allele copy number in library prep	Target bases 0.5-1.5X of mean (%)	HS library size (unique molecules)	Mean target coverage
Wild Type (0%)	10	0	98.8	$1.22x10^5$	1400
5%	1	15	95.1	$2.34x10^4$	269
1%	10	30	98.6	$1.35x10^5$	1546
0.1%	50	15	98.3	$4.76x10^5$	5783
0.1%	25	7	98.6	$2.91x10^5$	3382

USING SINGLE-STRAND UMI CONSENSUS FOR VARIANT DETECTION

The addition of UMIs and hybridization capture to the xGen cfDNA & FFPE DNA Library Prep Kit workflow in combination with the G4 Sequencing Platform, enables detection of variants from a wide range of allele frequencies. Because the UMI sequences are fixed, it is even possible to identify and correct sequencing or PCR errors within the UMI, preventing artificial inflation of library complexity and incorrect variant calls.

Various approaches can be used for removing PCR duplicates before variant calling using the xGen cfDNA & FFPE DNA Library Kit, including deduplication methods and single strand error correction strategies. The most common method for removing PCR duplicates before variant calling is based on retaining only the highest quality read of all reads with the same start-stop coordinates (Start-Stop). To minimize the removal of reads originating from unique molecules, both the start-stop position and the UMI sequence can be used to deduplicate reads (Start-Stop UMI); this deduplication method chooses the read with the highest read quality which may or may not contain PCR errors or low frequency variants. A final method using UMIs can be used to correct errors in sequencing data at the same time as removing duplicate reads (ssConsensus); this method uses all reads sharing both the start-stop sites and UMI to choose the most likely base at each position, producing a collapsed single read family that can be used for variant calling.

Table 2. Comparison of analysis approaches to detect mutations and the number of replicates which correctly identify each expected mutation at 0.1% allele frequency.

		Alleles		0.1% alleles 50 ng input			0.1% alleles 25 ng input		
chr	pos	WT	Mutation	Start/ stop	UMI Start/ stop	ssConsensus	Start/ stop	UMI Start/ stop	ssConsensus
chr1	114713909	G	T	3	2	3	2	2	2
chr1	114713915	C	T	3	3	3	3	2	3
chr12	25245350	C	T	3	0	3	3	0	3
chr3	179218303	G	A	3	2	3	2	0	3
chr7	55174771	AGGAATTAAGAGAAGC	A	2	3	3	1	1	3
chr7	55181305	A	ATGGCCAGCG	1	2	3	1	1	3
chr7	55181378	C	T	3	1	3	3	0	3
chr7	55191822	T	G	3	3	3	1	3	3

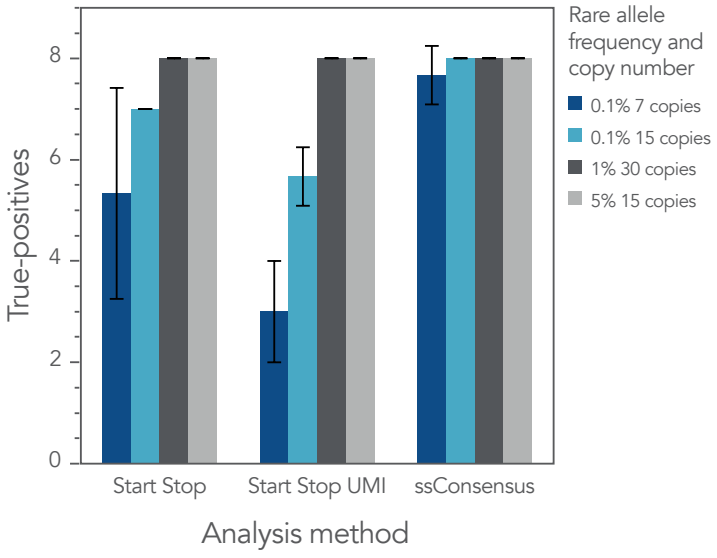

Figure 2. True-positive variant detection. True-positives between analysis methods, varying allele frequencies, and rare-allele copy numbers. Libraries were prepared and enriched as described above Figure 1. Sequencing was performed using 2 x 150 paired-end reads and subsampled to 20 million reads per library. Variant calling was performed using VarDict, using 2 as the minimum alt-allele depth parameter. The samples used contained 8 expected variants used to determine true-positives. Error bars represent standard deviations of the three replicates.

In libraries generated using samples containing 1% and 5% rare-allele frequencies (Figure 2, grey bars), each of the 8 expected variants were detected regardless of analysis method (Figure 2). Sequencing of libraries generated using the 0.1% rare-allele frequency samples with 2 different mass inputs resulted in an expectation of either 7 or 15 copies of the rare allele going into library preparation. For the sample containing 15 copies (Figure 2, light blue bars), all 8 expected variants were detected in each replicate when using the ssConsensus analysis method. For the samples containing only 7 copies (Figure 2, dark blue bars), the 8 expected variants were detected in 2 of 3 replicates, with the final replicate library missing only a single expected variant, (Figure 2, Table 2). Without using UMIs to create single read families from samples containing this low mutant allele fraction, several known variants across multiple replicates are missed, highlighting the importance of UMIs in analysis for sensitive applications. In summary, pairing xGen NGS library prep with Singular Genomics index primers and sequencing on the G4 Platform provides a seamless workflow that enables confident low copy number variant identification.

ORDERING INFORMATION

Product	Catalog #	URL
xGen cfDNA and FFPE DNA Library Prep Kit v2 MC	100010207	https://www.idtdna.com/pages/products/next-generation-sequencing/workflow/xgen-ngs-library-preparation/dna-library-preparation/cfdna-ffpe-prep-kit
xGen Hybridization and Wash Kit v2	10010352, 100110354	https://www.idtdna.com/pages/products/next-generation-sequencing/workflow/xgen-ngs-hybridization-capture/hybridization-capture-core-reagents
xGen Custom Hyb Panel—Accel		https://www.idtdna.com/pages/products/next-generation-sequencing/workflow/xgen-ngs-hybridization-capture/custom-hyb-panels#ordering
Custom amplification primers and xGen Custom Universal Blocking Oligos for Singular Genomics indexed libraries		https://go.idtdna.com/NGS-Request-Consult.html
Singular Genomics Dual Index PCR Primers (1-96)	700,110	https://singulargenomics.com/g4/reagents/
Singular Genomics F2 300 Cycle Kit	700,103	https://singulargenomics.com/g4/reagents/
Singular Genomics G4 Sequencing Platform	700,001	https://singulargenomics.com/g4/