High sensitivity somatic variant detection from formalin-fixed, paraffin-embedded samples using a versatile custom enrichment cell-free DNA (cfDNA) Library Preparation assay

Alexandria Dabuet, William Jordan, Sohela de Rozieres, Lisa Watson, Anna Babayan and Mariela Cuadras[£], Illumina, 5200 Illumina Way, San Diego, CA 92122, USA

[£]Corresponding Author

INTRODUCTION

Molecular profiling of solid tumors by targeted Next Generation Sequencing (NGS) is an emerging practice for standard of care in oncology. The process of fixation and paraffin embedding tissues in formalin-fixed paraffin-embedded (FFPE) samples can damage and/or modify genomic DNA making tumor genetic analysis challenging, especially for detecting low abundant somatic mutations. We evaluated the performance of Illumina cfDNA Prep with Enrichment assay detecting low frequency variants in FFPE DNA. The assay is a novel and flexible custom enrichment library preparation method that leverages Unique Molecular Identifiers (UMI) for error correction.

MATERIALS & METHODS

The workflow is depicted on Figure 1. In brief, library preparation consisted of end-repair of fragmented FFPE DNA, followed by ligation of adapters containing UMIs, addition of indexes by PCR, and enrichment with a single hybridization step. A panel of 2000 kb that target cancer-related genes and Illumina Exome 2.5 panel were used for enrichment. Libraries were prepared from 10, 20 and 40 ng of commercial FFPE Cell lines or FFPE human tissues (Table 1). FFPE DNA was fragmented in a Covaris LE220 Model (Table 2).

Libraries were sequenced on Illumina NextSeq™ 550 or NovaSeq™ 6000 sequencing systems. BCL files were converted to FASTQ and subsampled to indicated number of reads in BaseSpace™ Sequence Hub. Alignment and variant calling were performed using DRAGEN™ Enrichment App v4.0.3.

Library Prep and Enrichment

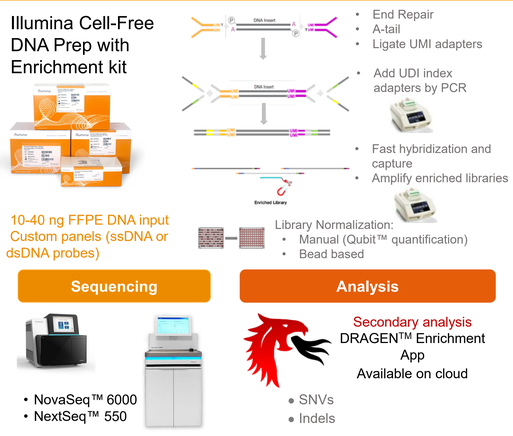


Figure 1. Illumina cell-free DNA Prep with Enrichment workflow for FFPE samples

Samples and upstream steps

Samples

Sample ID	Cancer type	ΔCq
8752	NSCLC	5.7
315	NSCLC	5.5
4075	CRC	4.2
679	Breast cancer	4.1
11883	CRC	2.9
3952	Breast cancer	2.2
448	NSCLC	2.1
449	NSCLC	1.5
FFPE TST-DNAv2[1,2]	N/A	0.87
HD653[1]	N/A	-1.24
HD301[1]	N/A	-1.87

[1]FFPE cell line
[2]Custom FFPE cell line, Seracare LGC Clinical Diagnostics

Table 1. Samples analyzed using Illumina cfDNA Prep with Enrichment

Covaris LE220 settings

Factor	Setting
Peak incident power	450 watts
Duty Factor	30%
Cycles per burst	200
Treatment time	250 seconds
Temperature	7°C

Table 2. Covaris settings

RESULTS

Library performance metrics for FFPE samples

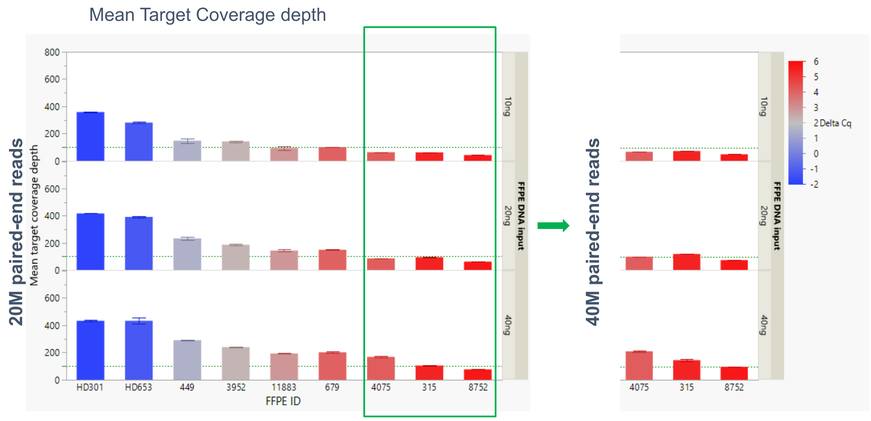


Figure 2. Libraries from high quality FFPE samples (ΔCq ≤2) enriched with an 80bp-ssDNA 2000 Kb panel and sequenced on NextSeq 550 (2 x 149bp) achieved ≥100x Mean Target Coverage depth (MTC) with as low as 10 ng input. MTC decreases with sample quality (higher ΔCq) but improves with input and/or deeper sequencing. Other library QC metrics are summarized in Table 4.

Analytical Sensitivity for SNVs and Indels

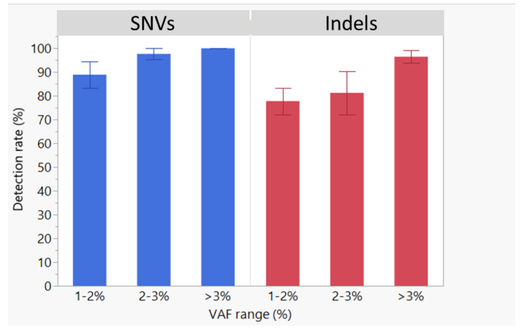
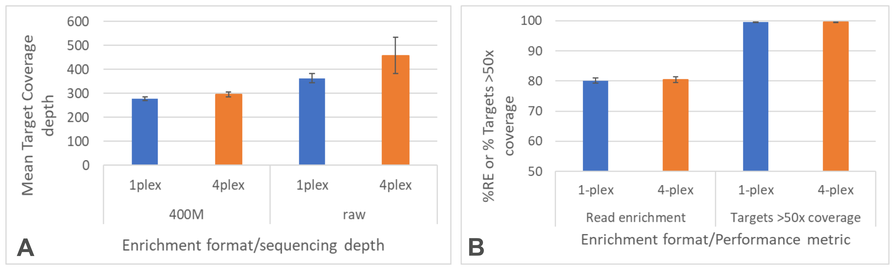


Figure 3, Table 3. Variants in a custom FFPE cell line were diluted to different allele frequency levels. Then six replicates of 10 ng input libraries were prepared. Libraries were 1-plex enriched with an 80bp-ssDNA 2000 Kb panel and sequenced on NextSeq 550 (2 x 149bp). Sequencing reads were subsampled to 40M paired-end reads, and the number of replicates where the indicated variants were detected was evaluated. SNVs ≥2% VAF and Indels ≥3% VAF were detected at 90% sensitivity.

COSMIC ID	Gene	Variant type	Detected VAF (%)*	No. of replicates detected
COSV51614178	MLH1	SNV	2.0	6
COSV57169334	MYD88	SNV	2.3	6
COSV58963463	CSF3R	SNV	2.6	6
COSV55545304	KRAS	SNV	2.9	6
COSV50630049	CBL	SNV	2.9	5
COSV61615239	IDH1	SNV	3.0	5
COSV52274101	MSH6	SNV	3.5	6
COSV59205440	SF3B1	SNV	3.9	6
COSV56057713	BRAF	MNV	2.1	6
COSV64288359	PTEN	Del	2.3	6
COSV56542602	VHL	Del	2.5	6
COSV52740986	TP53	Del	2.7	6
COSV55388067	KIT	Del	3.2	5
COSV61376874	ARID1A	Del	3.2	6
COSV62688630	CTNNB1	Del	4.7	6
COSV67575778	JAK2	Del	5.3	5
COSV56060749	BRAF	Ins	1.9	6
COSV55386625	KIT	Ins	2.8	6
COSV64290304	PTEN	Ins	3.1	6
COSV51766549	EGFR	Ins	3.5	6
COSV51772596	EGFR	Ins	3.6	6
COSV57195669	CEBPA	Ins	5.0	6

*Average of detected replicates

Compatible with Illumina Exome 2.5 Enrichment Panel



*Raw reads: ~638M – 1.08B paired-end reads.

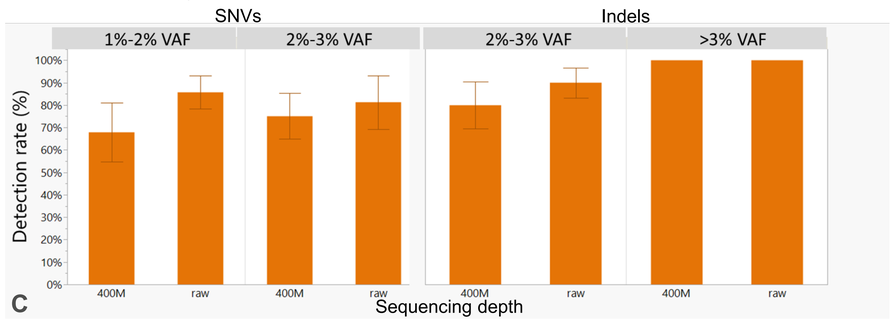


Figure 4. Library performance metrics and variant detection were evaluated in 20 ng libraries enriched with Illumina Exome 2.5 panel following 1-plex and 4-plex enrichment formats. Reads were sub-sampled to 400M paired-end reads (A, B, C). **A-B**. Performance metrics obtained when 8 libraries were sequenced on one S4 lane of NovaSeq 6000 platform. MTC increases with deeper sequencing (raw*). **C.** 4-plex-enriched libraries detected up to ~ 80% SNVs and ~90% Indels at 2%-3% VAF. Detection rate of low VAF variants improved with deeper sequencing (raw*).

For Research Use Only. Not for use in diagnostic procedures.

Performance comparison between 1-plex and 4-plex enrichment

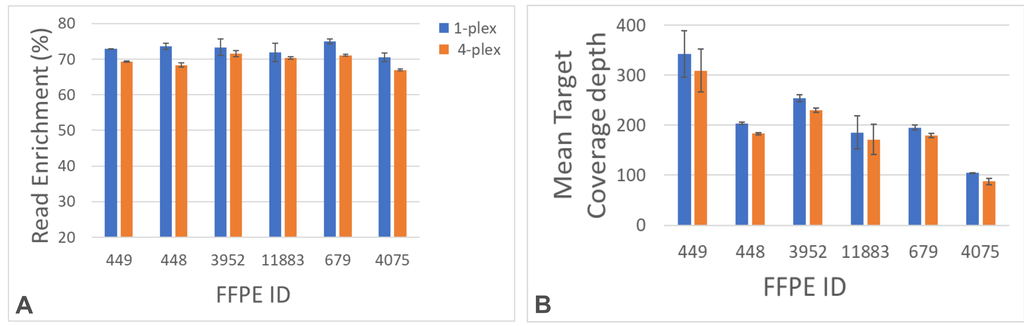


Figure 5. 20ng input libraries from selected FFPE samples were individually (1-plex) and 4-plex enriched with an 80bp-ssDNA 2000 Kb panel and sequenced on NextSeq 550 (2 x 149bp). After subsampling sequencing reads to 40M paired-end reads, the libraries showed similar read enrichment (A), Mean target coverage depth (B) and above 90% targets with ≥ 50x coverage (not shown), regardless of the enrichment format.

Other library QC metrics*

FFPE ID	% Read enrichment	Fragment length (bp)	% Targets ≥50x coverage
HD301	75.3	183	99.6
HD653	76.7	173	99.6
449	73	137	99.5
3952	73.3	133	99.5
11883	71.8	127	99.3
679	74.09	124	99.3
4075	70.5	122	95.3
315	42.3	111	98
8752	42.4	108	71.7

*Shown for 20ng, 40M paired-end reads, from Figure 2

Table 4. Libraries from medium to high quality FFPE samples (ΔCq ≤4) shown on figure 2, showed >120 bp length, achieved >70% read enrichment and >70% target coverage.

CONCLUSION

These results demonstrate performance of Illumina cfDNA Prep with Enrichment in preparing sequencing-ready libraries from FFPE DNA that produce accurate data for downstream analyses, including somatic variant calling from as little as 10 ng input. The assay is a versatile library preparation kit compatible with user-supplied enrichment panels allowing researchers to tailor their experimental design based on their research needs.

illumina®