Swift BIOSCIENCES™

NGS Library Prep Methods to Achieve Comprehensive Coverage for WGS and Targeted Sequencing at Low DNA Input Quantities, Including FFPE

M. Hymes, L. Kurihara, J. Laliberte, C. Schumacher, S. Sandhu, J. Irish, S. Chupreta, and V. Makarov
Swift Biosciences, 58 Parkland Plaza, Suite 100, Ann Arbor, MI 48103, Tel: 734.330.2568

Digital Biology®

Abstract

When performing whole genome sequencing (WGS) or targeted sequencing using hybridization-based capture, unbiased, even coverage of the genome is required in order to conduct comprehensive analysis from the lowest possible sequence read depth. Highly efficient conversion of DNA fragments into library molecules is also necessary when DNA input quantity or quality is limited. We have developed a library preparation method to achieve these results with high quality DNA, as well formalin-fixed, paraffin-embedded (FFPE) samples.

This method uniquely repairs damage on both the 3' and 5' termini to enhance ligation efficiency to sheared DNA fragments. Combined with sequential ligation steps that optimize attachment to each terminus, this single tube 'with bead' method supports PCR-free sequencing from inputs as low as 10 ng circulating, cell-free DNA (cfDNA) or 100 ng physically sheared DNA. Input quantities down to 10 pg can be used with PCR amplification. Efficiency of library conversion is ~50% for physically sheared DNA and up to 90% for circulating cfDNA. Human WGS using this method demonstrates high complexity with exceptional coverage of GC-rich promoter regions compared to other methods. With inputs as low as 1 ng human DNA at 18X coverage, the genome was fully represented without loss of data or reduction in relative coverage of GC-rich promoter regions.

For hyb capture enrichment, our method demonstrates a reduction of duplicate reads from inputs at 10 ng and 1 ng as compared to other library preparation methods. Paired with an exome panel, we observed 5% and 26% duplicates with 10 ng and 1 ng inputs respectively. Coverage uniformity was also preserved at these low inputs. Performance was maintained with FFPE samples using the same input quantities from an AML cancer panel. This library preparation method creates a practice for analyzing low-input samples, including poor quality FFPE materials.

Accel-NGS® 2S Plus Library Kit

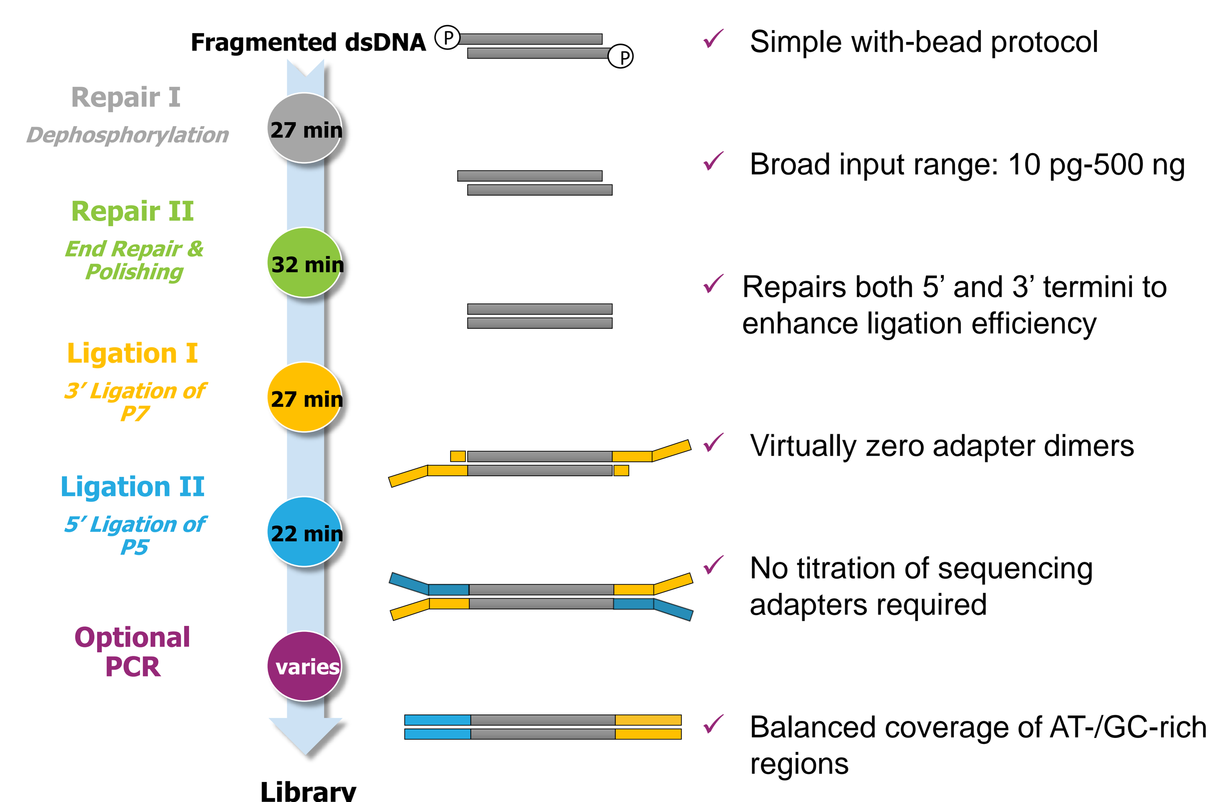


Figure 2. Accel-NGS 2S PCR-free workflow. Accel-NGS 2S PCR-free has 4 steps: Repair I dephosphorylates the 5' ends of the DNA, Repair II repairs and polishes DNA ends, Ligation I adds the P7 adapter to the 3' terminus, Ligation II adds the P5 adapter to the 5' terminus, when automated on the Sciclone® G3 NGSx liquid handling platform. Optional PCR amplifies the library when required.

PCR-Free Sequencing of cfDNA

15 ng of Circulating, Cell-free DNA

SEQUENCING METRICS	SAMPLE A	SAMPLE B
Total Reads	238,230,712	254,136,768
% Aligned	99.4%	99.4%
Average Genome Coverage	14.6 x	15.6 x
% Genome Missing	1.9%	2.0%
% Genome Covered ≥ 5X	99.6%	99.7%
% Genome Covered ≥ 10X	94.9%	95.8%
% Genome Covered ≥ 14X	92.6%	93.2%
% Duplication	0.04%	0.08%
Median Insert Size	172 bp	168 bp

Table 3. Sequencing metrics from Illumina® HiSeq® 2500. Two cfDNA samples were prepared PCR-free with the Accel-NGS 2S Plus DNA Library Kit and sequenced on a HiSeq 2500 Rapid Run with v3 chemistry. Analysis was performed using a custom pipeline using BWA as aligner and enrichment metrics were collected using Picard tools.

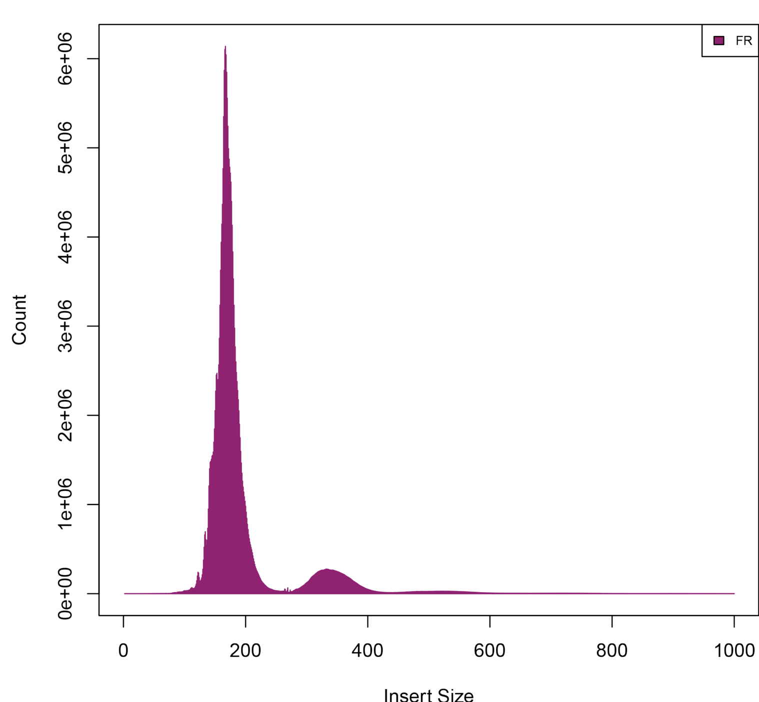
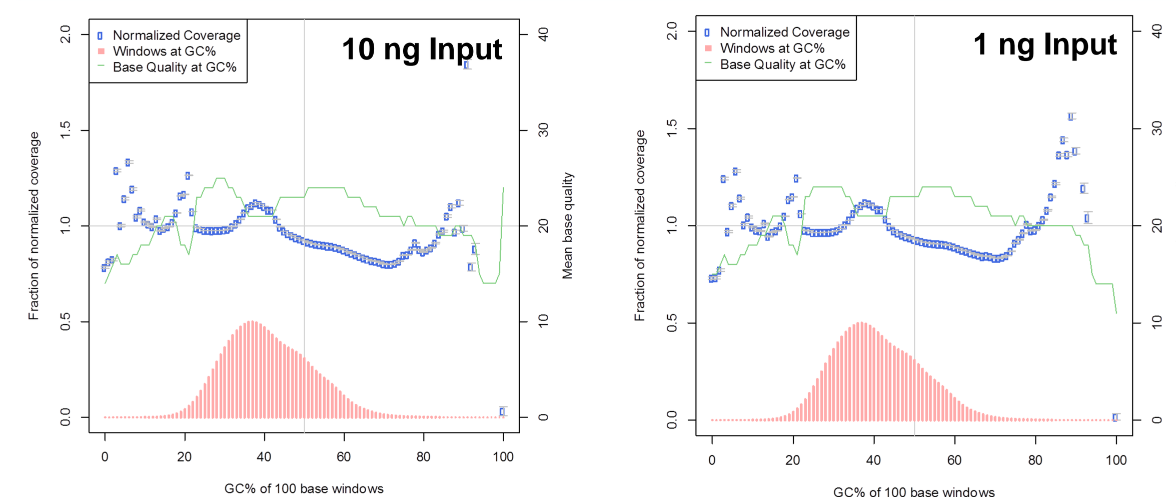


Figure 4. Aligned insert size of NGS libraries. The curve reflects the library size profile, where both peaks represent properly paired alignment. Chimeric reads were detected at 1.7 and 1.6% respectively for samples A and B.

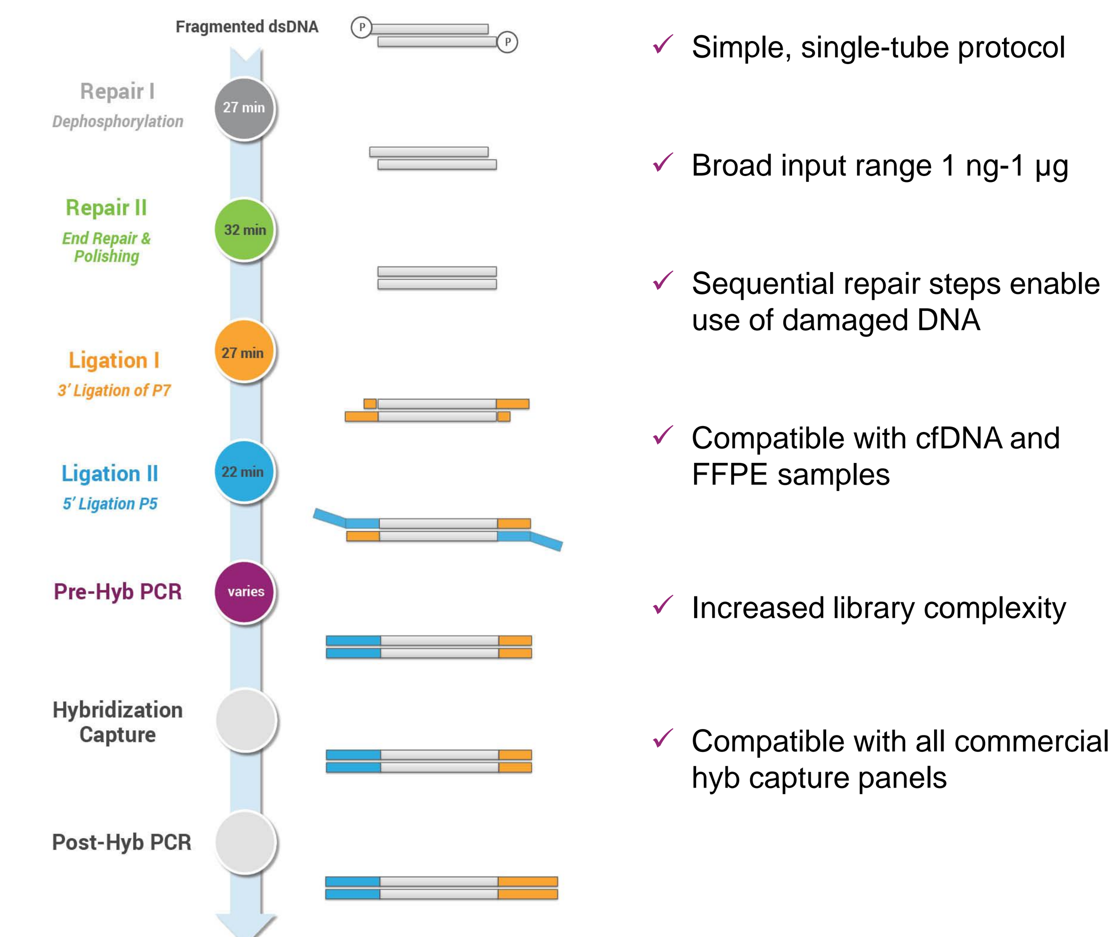
Uniform Coverage 1 ng Human DNA



INPUT	# READS	% ALIGNED	% DUP.	MEAN COVERAGE	% COVERED ≥10X	ESTIMATED LIBRARY SIZE
10 ng	524 M	98.6	4.0%	19X	93.3%	4,160,642,868
1 ng	530 M	98.8	17.8%	16X	89.8%	772,318,743

Human HapMap gDNA NA12878 was Covaris® sheared to 350 bp, libraries were constructed using Accel-NGS 2S and sequenced on a HiSeq 2500.

Accel-NGS 2S Hyb Library Kit



FFPE DNA Input

xGen® Pan-Cancer Panel

INPUT	SAMPLE TYPE	% ALIGNED	% DUP.	MEAN BAIT COV.	% COV. ≥ 1X	% COV. ≥ 20X	% ON TARGET
100 ng	Frozen	96	1	42X	99	91	80
	6 hr. Fix	96	1	43X	99	93	81
	24 hr. Fix	97	1	44X	99	93	82
	48 hr. Fix	97	1	45X	99	88	82
10 ng	Frozen	96	3	42X	99	90	80
	6 hr. Fix	96	5	41X	99	92	80
	24 hr. Fix	97	4	42X	99	93	81
	48 hr. Fix	97	8	42X	99	86	81
1 ng	Frozen	95	18	33X	99	85	77
	6 hr. Fix	94	32	26X	99	77	74
	24 hr. Fix	95	31	27X	100	79	76
	48 hr. Fix	95	44	22X	99	53	73

The Pan-Cancer panel is 0.9 Mb; libraries were normalized to 0.6M reads.

Table 1. Accel-NGS 2S Hyb performance metrics using the xGen® Pan-Cancer Panel for hybridization capture with FFPE DNA. To examine the effects of fixation on sequencing data, libraries were made with the Accel-NGS 2S Hyb kit from fresh-frozen kidney DNA and from the same sample fixed for 6-, 24-, or 48-hr., then paraffin embedded.

Agilent SureSelect^{XT} Q^2 Lab Solutions Comprehensive Cancer Panel

INPUT	SAMPLE TYPE	% ALIGNED	% DUP.	MEAN BAIT COV.	% COV. ≥ 1X	% COV. ≥ 20X	% ON TARGET
20 ng	FFPE	97	18	121X	99	99	89
10 ng	FFPE	96	30	96X	99	98	88

The QCCP panel is 1.3 Mb; libraries were normalized to 3M reads.

Table 2. Accel-NGS 2S Hyb performance metrics using the Agilent SureSelect^{XT} Q^2 Solutions Comprehensive Cancer Panel for hybridization capture with FFPE DNA. The Accel-NGS 2S Hyb kit was used to make libraries of three different inputs from the FFPE lung tumor sample. Many thanks to Q^2 Lab Solutions for generation and sequencing of libraries.

High Quality DNA Input

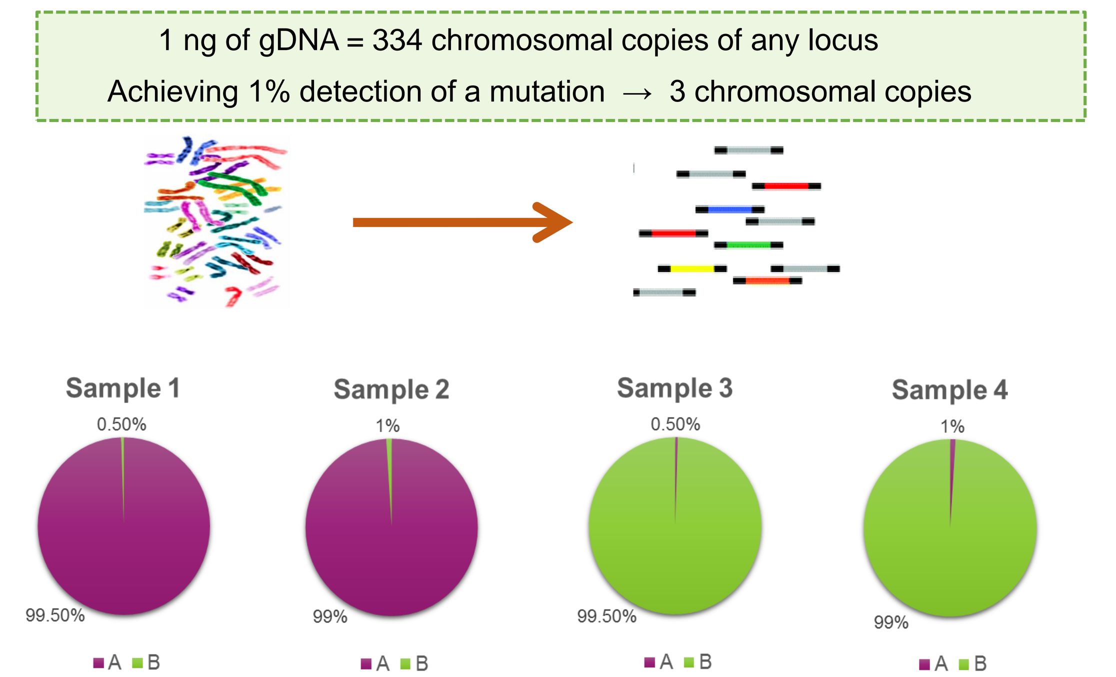
NimbleGen™ SeqCap™ EZ MedExome

INPUT	METHOD	% ALIGNED	EST. LIBRARY SIZE (M)	% DUP.	MEAN BAIT COV.
100 ng	Swift	93	1125	1	50X
	Kapa	93	240	6	51X
10 ng	Swift	93	275	5	52X
	Kapa	93	97	13	47X
1 ng	Swift	93	45	26	37X
	Kapa	90	7	71	10X

All libraries were normalized to 39M reads. The MedExome Panel is 64 Mb in size.

Table 5. Comparative performance metrics between Accel-NGS 2S Hyb and Kapa using SeqCap EZ MedExome hybridization capture. Libraries were made using HapMap DNA NA12878 (Coriell) with both the Swift Accel-NGS 2S Hyb Kit and the Kapa Library Preparation kit, followed by the NimbleGen SeqCap EZ MedExome Panel.

Limit of Detection (gDNA)

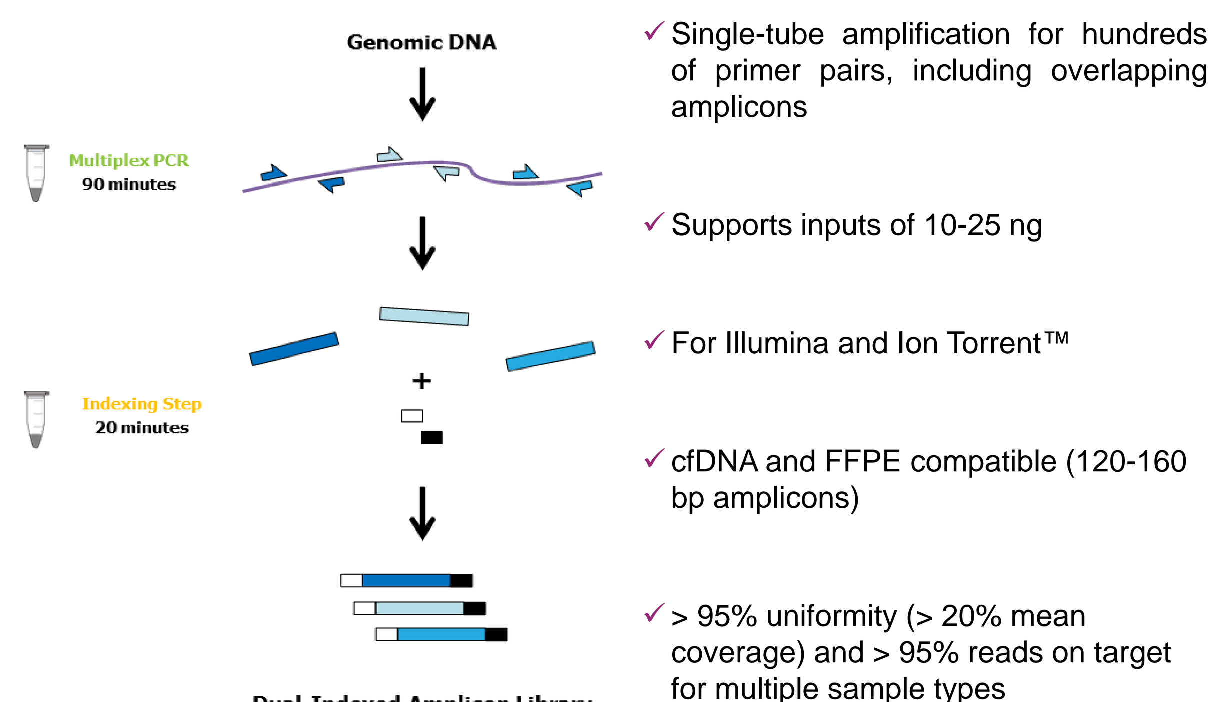


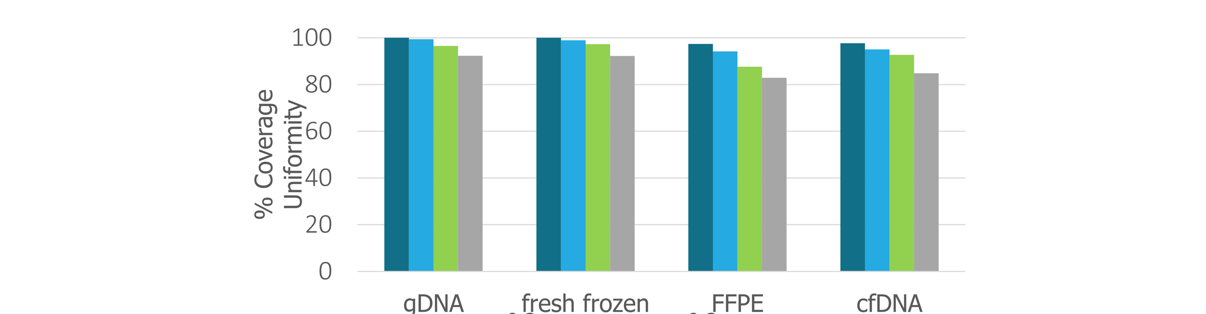
To assess the limit of detection of Accel-NGS 2S Hyb, DNA samples from two individuals (A and B) with different ethnic backgrounds were used to prepare libraries. 100 ng of DNA from one individual with a 0.5% or 1% spike-in of the DNA from the second individual was used as the input DNA. Once libraries were prepared, they were hybridized to xGen Pan-Cancer probes and SNPs were detected within this panel.

cfDNA Input	ALLELE FREQUENCIES				
CHR: POS	A Back-ground	B Back-ground	C Back-ground	1% A into 10 ng B	1% A into 10 ng C
2: 212244718	100%	0%	0%	0.6%	1.0%
12: 25361074	100%	0%	0%	1.6%	1.9%
12: 25361142	100%	0%	0%	1.1%	0.9%
12: 25361646	100%	0%	0%	1.9%	1.6%
12: 40688695	100%	0%	0%	0.5%	1.1%
12: 115108136	100%	0%	0%	0.7%	2.0%

To determine if SNPs present at 1% allele frequency could be detected, 1% of cfDNA sample (A) with a unique ethnic background was spiked into two 10 ng cfDNA samples (B and C) of different ethnic backgrounds. SNPs at 100% from sample A could be detected around 1% when those SNPs are not present in B and C backgrounds. Libraries were sequenced to an average coverage of 8700X.

Accel-Amplicon™ 56G Oncology Panel





Comprehensive & Hotspot Coverage of 56 Oncology-related Genes

ABL1	5	*CSF1R*	2	*FBXW7*	6	*GNAS*	2	*KIT*	14	*NPM1*	1	*STK11*	5
AKT1	2	*CTNNB1*	1	*FGFR1*	2	*HNF1A*	4	*KRAS*	3	*NRAS*	3	*SMAD4*	10
ALK	2	*DDR2*	1	*FGFR2*	4	*HRAS*	2	*MAP2K1*	5	*PDGFRA*	4	*SMARCB1*	4
APC	9	*DNMT3A*	1	*FGFR3*	6	*IDH1*	1	*MET*	6	*PIK3CA*	11	*SMO*	5
ATM	19	*EGFR*	9	*FLT3*	4	*IDH2*	2	*MLH1*	1	*PTEN*	14	*SRC*	1
BRAF	2	*ERBB2*	4	*FOXL2*	1	*JAK2*	2	*MPL*	1	*PTPN11*	2	*TP53*	21
CDH1	3	*ERBB4*	8	*GNA11*	2	*JAK3*	3	*MSH6*	4	*RB1*	12	*TSC1*	1
CDKN2A	2	*EZH2*	1	*GNAQ*	2	*KDR*	9	*NOTCH1*	3	*RET*	6	*VHL*	3

56G panel genes and number of amplicons per gene. Hotspot loci (white), contiguous, overlapping coverage (blue) and comprehensive coding exon coverage for TP53 (darker blue).

Conclusions

The Accel-NGS 2S Plus DNA Library Kit:
- Provides uniform, comprehensive genome coverage.
- Enables PCR-free sequencing of cfDNA from 10-15 ng input.
- Maintains comprehensive, uniform coverage down to 1 ng of DNA

The Accel-NGS 2S Hyb DNA Library Kit:
- Provides a greater depth of coverage for low input samples than other commercial kits.
- Is compatible with all commercially available hyb capture panels.
- Enables 1% allele detection from 10 ng of DNA.
- Maintains high genomic coverage from damaged FFPE samples.

Accel-Amplicon 56G Oncology Panel:
- Provides quality performance with > 95% on target and > 95% coverage uniformity.
- Enables a limit of mutation detection of 1% for liquid biopsy from 10 ng cfDNA.

Using a highly efficient library preparation method for WGS or WES, improves genomic coverage and detection of low frequency alleles .

We acknowledge and thank Q2 Lab Solutions for Agilent SureSelect^{XT} library preparation and sequencing.

www.swiftbiosci.com

© 2016, Swift Biosciences, Inc. The Swift logo and Accel-Amplicon are trademarks and Accel-NGS is a registered trademark of Swift Biosciences. Sciclone is a registered trademark of PerkinElmer Inc. Illumina and HiSeq are registered trademarks of Illumina, Inc. Covaris is a registered trademark of Covaris, Inc. xGen is a registered trademark of Integrated DNA Technologies, Inc. SureSelect^{XT} is a product of Agilent Technologies. NimbleGen and SeqCap are trademarks of Roche NimbleGen Inc. Ion Torrent is a trademark of Thermo Fisher Scientific Inc. 16-0798, 04/16 **Disclosure: All authors have a financial relationship with Swift Biosciences, Inc. as employees.**