APPLICATION NOTE

Tumor exome sequencing with Illumina DNA Prep with Enrichment and DRAGEN™ secondary analysis

Efficient and accurate variant analysis in FFPE tumor samples

illumina®

Introduction

During tumor progression, cells acquire mutations in their DNA that drive hallmarks of cancer, such as proliferation, survival, and immune escape. Understanding these mutations can help researchers understand the biology of the disease, pave the way for new biomarkers, and elucidate potential therapeutic targets. Because the exome represents less than 2% of the genome, cancer exome sequencing using next-generation sequencing (NGS) technology is a cost-effective method to investigate associated mutations.[1–3] Unfortunately, most exome sequencing solutions have 2-3 day workflows, require large amounts of DNA, and are not compatible with formalin-fixed, paraffin embedded (FFPE) tissue samples.

To help researchers take advantage of the information within the exome, Illumina offers Illumina DNA Prep with Enrichment (formerly Nextera™ Flex for Enrichment) and the DRAGEN secondary analysis. Illumina DNA Prep with Enrichment delivers high-quality exome sequencing libraries by combining bead-based library preparation chemistry with a simplified, single hybridization for enrichment. The resulting workflow has the fewest number of steps and the fastest turnaround time in the Illumina enrichment portfolio.[4] DRAGEN software provides rapid secondary analysis of NGS data from genomes, exomes, and transcriptomes. Without compromising accuracy, DRAGEN software delivers quickness, flexibility, and cost efficiency, enabling labs of all sizes and disciplines to do more with their NGS data.[5] The DRAGEN Enrichment App is the recommended method for analyzing enrichment data, including exomes. It delivers a full suite of enrichment-specific metrics and reporting with greater accuracy and in less time than previous analysis solutions.[5]

This application note demonstrates a streamlined, comprehensive exome sequencing workflow that integrates Illumina DNA Prep with Enrichment, proven Illumina sequencing, and highly accurate DRAGEN secondary analysis for identification of disease-associated variants in tumor samples (Figure 1).

Methods

Samples

Exome sequencing and variant analysis were performed using nine unique FFPE tumor–normal sample pairs (Table 1) and Quantitative Multiplex Reference Standard formalin-compromised DNA (fcDNA) (Horizon Discovery, Catalog no. HD799).

Library preparation

After extraction with the AllPrep DNA/RNA FFPE Kit (QIAGEN, Catalog no. 80234), sequencing libraries were prepared from either 50 ng or 10 ng of DNA input using Illumina DNA Prep with Enrichment and the Illumina Exome Panel. Single-plex enrichment reactions were performed with IDT for Illumina DNA Unique Dual Indexes using an overnight hybridization at 58°C.

Sequencing

All libraries were sequenced on a NovaSeq™ 6000 System with the appropriate flow cell and run parameters (Table 2). Samples can also be run on the NextSeq™ 1000 and NextSeq 2000 Sequencing Systems.

For a full list of kits and reagents for this workflow, see the Appendix.

Figure 1: Comprehensive exome sequencing workflow for FFPE tumor samples

Table 1: Tumor-normal sample pair information

Vendor	Source	Status	External ID	ΔCq
Horizon Discovery	N/A	Reference standard	HD799	N/A
Bio-Options	Uterus	Benign adjacent	NP-18449-1	0.26
Bio-Options	Uterus	Uterine cancer	NP-18449-4	0.38
Bio-Options	Colon	Benign adjacent	NP-18676-3	1.77
Bio-Options	Colon	Colon cancer	NP-18676-2	0.40
OutDo Biotech	Skin	Benign adjacent	15-21711-3	1.68
OutDo Biotech	Skin	Melanoma	15-21711-5	1.82
Bio-Options	Colon	Benign adjacent	NP-17874-3	1.35
Bio-Options	Colon	Colon cancer	NP-17874-2	2.01
OutDo Biotech	Skin	Benign adjacent	15-10691-12	1.62
OutDo Biotech	Skin	Melanoma	15-10691-7	2.07
Bio-Options	Stomach	Benign adjacent	NP-16695-4	2.22
Bio-Options	Stomach	Gastric cancer	NP-16695-1	1.79
Bio-Options	Uterus	Benign adjacent	NP-12844-3	2.27
Bio-Options	Uterus	Uterine cancer	NP-12844-2	2.77
Bio-Options	Pancreas	Benign adjacent	NP-18638-6	3.62
Bio-Options	Pancreas	Pancreatic cancer	NP-18638-5	3.74
Bio-Options	Uterus	Benign adjacent	NP-13587-4	4.34
Bio-Options	Uterus	Uterine cancer	NP-13587-3	4.24

Table 2: Exome sequencing run parameters

Sequencing system	Flow cell	No. of samples per run
NextSeq 1000 or NextSeq 2000 System	P2	3
NovaSeq 6000 System	SP/S1/S2/S4	6/13/33/80

The number of samples per run is based on a single flow cell in which 250 million reads are generated per sample.

Data analysis

After sequencing was complete, data was securely streamed directly from the instrument into the cloud ecosystem for push-button analysis using DRAGEN applications available through BaseSpace™ Sequence Hub (Table 3). The DRAGEN Enrichment App was used to align sequences from the individual samples to the reference genome, producing BAM alignment files and quality control and analysis metrics for exomes. The alignment BAM files were input into the DRAGEN Somatic App, which performed variant calling and germline variant filtering, generating a single VCF file with somatic variants from the input tumor–normal pairs (Figure 2).

Table 3: DRAGEN analysis pipelines

Analysis tool	Application
DRAGEN Enrichment App	Sequence mapping and alignment to reference, includes enrichment metrics
DRAGEN Somatic App	Somatic variant detection in tumor samples, includes tumor-only and tumor-normal modes

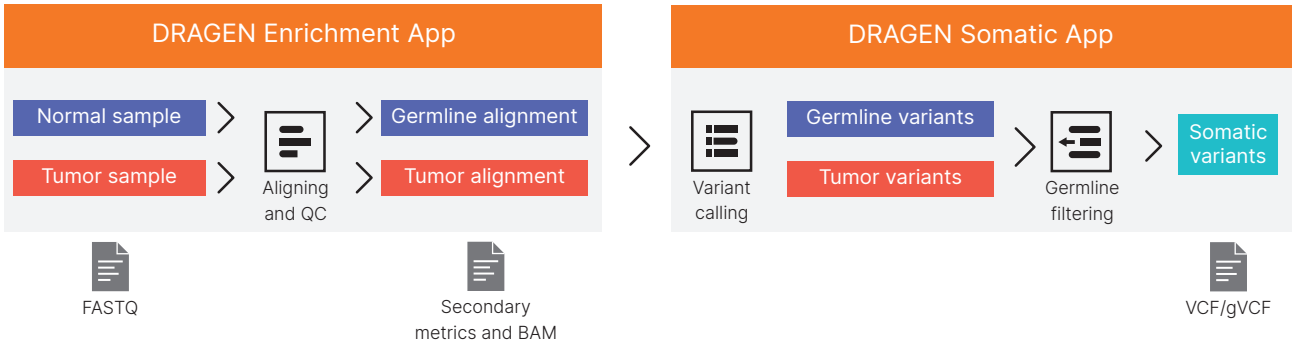

Figure 2: Tumor-normal variant analysis with DRAGEN apps— FASTQ files automatically generated after a sequencing run are processed by the DRAGEN Enrichment App to generate BAM alignments and secondary analysis metrics. The BAM files are input into the DRAGEN Somatic App for variant calling and germline filtering, which results in a single VCF file containing somatic variants.

Results

To demonstrate the performance of DRAGEN software for exome sequencing analysis, formalin-compromised and FFPE tumor–normal samples were evaluated.

Variant calling with formalin-compromised reference

Results with the formalin-compromised reference sample HD799 show that 100% of engineered small variants present (Table 4) were detected with 50 ng of input (Figure 3A), and all but one variant (95.8% analytical sensitivity) were detected with 10 ng of input (Figure 3B).

Table 4: HD799 small variants and VAF

Gene	Variant	Expected VAF
KRAS	G12D	6.0%
PIK3CA	E545K	9.0%
cKIT	D816V	10.0%
BRAF	V600E	10.5%
NRAS	Q61K	12.5%
KRAS	G13D	15.0%
PIK3CA	H1047R	17.5%
EGFR	G719S	24.5%

VAF, variant allele frequency

Variant calling with FFPE tumor samples

Nine FFPE tumor samples were evaluated with different input amounts. Results show significant correlation for variant allele frequencies between technical replicates for samples NP-18449-4 and NP-18676-2 with DNA inputs of 50 ng and 10 ng (Figure 4).

High-performing sequencing metrics

To further demonstrate the analytical performance of DRAGEN apps, various sequencing and enrichment metrics were evaluated. Highperforming metrics were achieved for all samples, including median fragment length (Figure 5A), percent aligned reads (Figure 5B), padded unique read enrichment (Figure 5C), mean target coverage depth (Figure 5D), percent target coverage at 50× (Figure 5E), and coverage uniformity (Figure 5F).

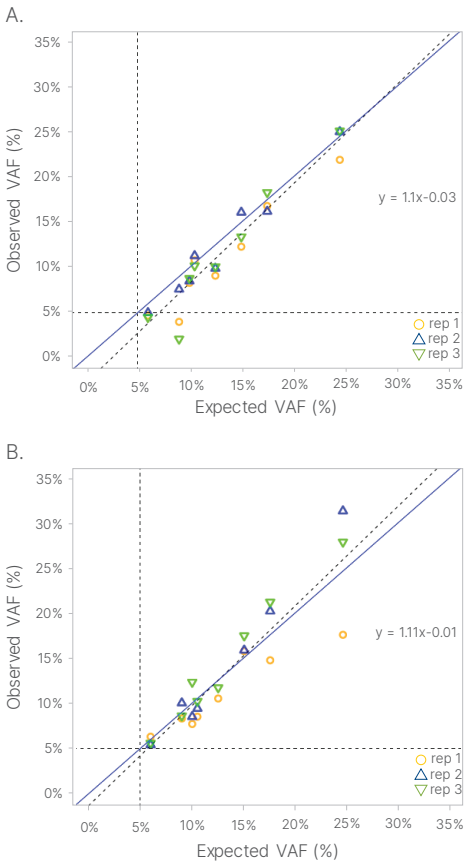

Figure 3: Detection of engineered variants in HD799 at different inputs—All engineered small variants in control sample HD799 were successfully detected with (A) 50 ng and (B) 10 ng of DNA input, except for the NRAS Q61K variant in replicate 2 (lack of blue triangle at 12.5% VAF).

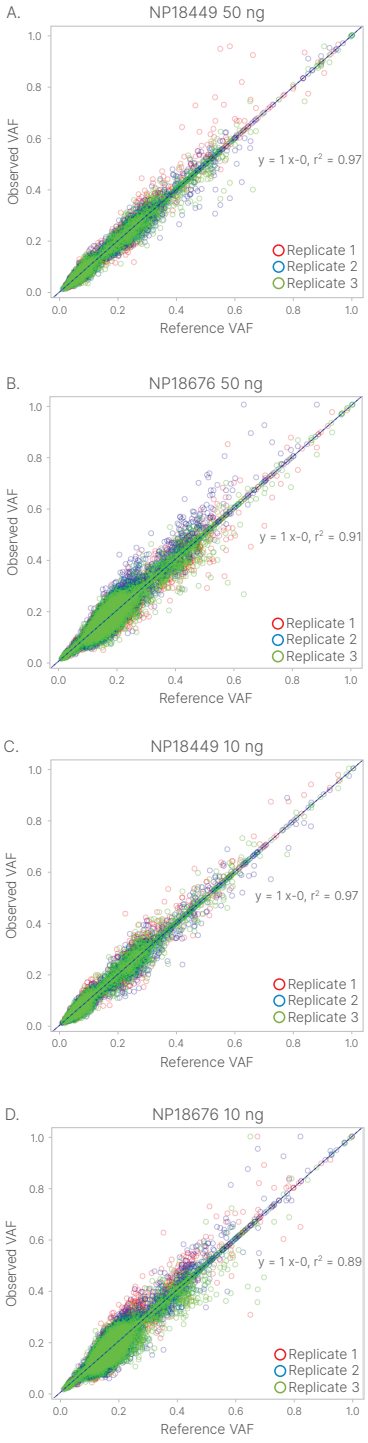

Figure 4: Variant detection in FFPE tumor samples at different inputs—Significant correlation was observed for variant allele frequencies between technical replicates for FFPE tumor samples at DNA inputs of 50 ng and 10 ng.

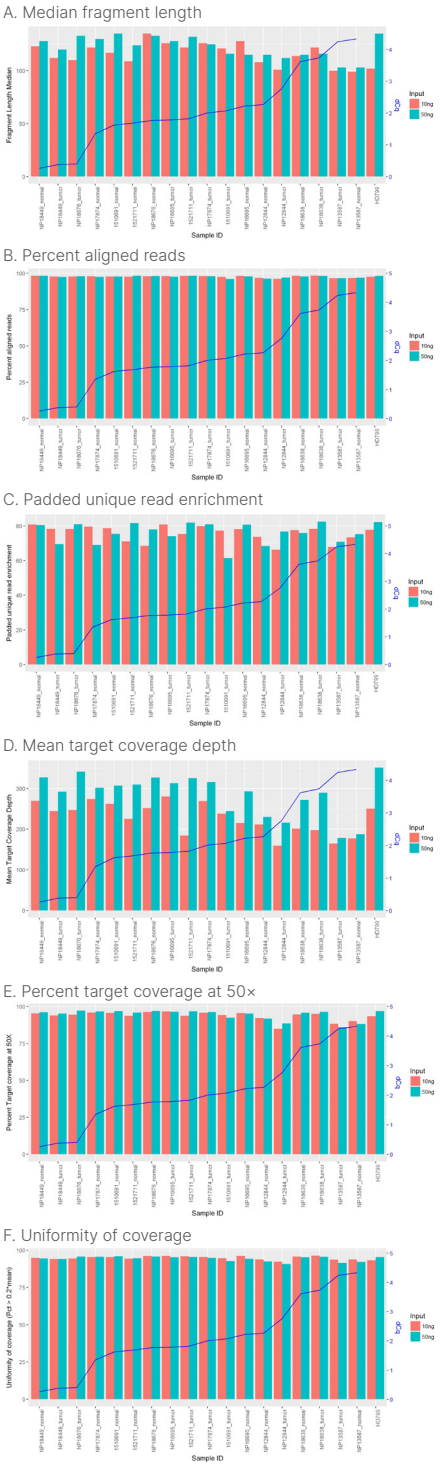

Figure 5: High-performing sequencing metrics—Samples are graphed left to right in order of decreasing quality, as indicated by ΔCq (blue line). High performing metrics were achieved for all samples at inputs of 10 ng (red) and 50 ng (blue).

Summary

This application note describes an exome sequencing workflow for FFPE tumor samples using Illumina DNA Prep with Enrichment for library preparation, the NovaSeq 6000 System for sequencing, and DRAGEN apps for data analysis. Results demonstrate the high analytical sensitivity and precision of DRAGEN secondary analysis for variant detection, particularly in FFPE tumor samples of varying quality and input amount.

Learn more

Illumina DNA Prep with Enrichment

Illumina sequencing systems

DRAGEN secondary analysis

References

1. Litchfield K, Summersgill B, Yost S, et al. Whole-exome sequencing reveals the mutational spectrum of testicular germ cell tumours. *Nat Commun*. 2015;6:5973.
2. Srivastava S, Cohen JS, Vernon H, et al. Clinical whole exome sequencing in child neurology practice. *Ann Neurol*. 2014;76:473–483.
3. Worthey EA, Mayer AN, Syverson GD, et al. Making a definitive diagnosis: successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. *Genet Med*. 2011;13:255–262.
4. Ilumina (2020). Illumina DNA Prep with Enrichment Data Sheet. Accessed January 24, 2024.
5. Illumina (2019). Illumina DRAGEN secondary analysis Data Sheet. Accessed January 24, 2024.

Appendix

Recommended library preparation, enrichment, and sequencing reagents for exome sequencing

Sequencing system	Sequencing reagents	Catalog no.
NextSeq 2000 System	NextSeq 1000/2000 P2 Reagents (200 cycles)	20040557
NovaSeq 6000 System	NovaSeq 6000 SP Reagent Kit (200 cycles)	20040326
	NovaSeq 6000 S1 Reagent Kit (200 cycles)	20012864
	NovaSeq 6000 S2 Reagent Kit (200 cycles)	20012861
	NovaSeq 6000 S4 Reagent Kit (200 cycles)	20027466
Product	**Description**	**Catalog no.**
Illumina DNA Prep with Enrichment, (S) Tagmentation, 96 samples	96 samples, 8 12-plex enrichment reactions	20025524
Illumina DNA Prep with Enrichment, (S) Tagmentation, 16 samples	16 samples, 16 1-plex enrichment reactions	20025523
Illumina DNA Prep, (S) Tagmentation, 96 samples	96 samples	20025520
Illumina DNA Prep, (S) Tagmentation, 16 samples	16 samples	20025519
Illumina Exome Panel	8 enrichment reactions	20020183
IDT for Illumina Nextera DNA Unique Dual Indexes Set A, Tagmentation (96 indexes, 96 samples)	96 indexes, 96 samples	20027213
IDT for Illumina Nextera DNA Unique Dual Indexes Set B, Tagmentation (96 indexes, 96 samples)	96 indexes, 96 samples	20027214
IDT for Illumina Nextera DNA Unique Dual Indexes Set C, Tagmentation (96 indexes, 96 samples)	96 indexes, 96 samples	20042666
IDT for Illumina Nextera DNA Unique Dual Indexes Set D, Tagmentation (96 indexes, 96 samples)	96 indexes, 96 samples	20042667

illumina®

1.800.809.4566 toll-free (US) | +1.858.202.4566 tel
techsupport@illumina.com | www.illumina.com

1170-2020-002-B