APPLICATION NOTE

High-quality variant calling with the NovaSeq™ 6000Dx instrument

Access a validated workflow for highly accurate germline and somatic variant calling

illumina®

For *In Vitro* Diagnostic Use. Not available in all countries or regions.

Introduction

The NovaSeq 6000Dx instrument ushers in a new era for the clinical laboratory. Capable of generating up to 6 Tb of data in less than two days, it delivers powerful and scalable high-throughput next-generation sequencing (NGS) in a Food and Drug Administration (FDA)-regulated, Conformité Européenne (CE)-marked platform for *in vitro* diagnostic (IVD) applications (Figure 1). The NovaSeq 6000Dx instrument is compatible with the latest Illumina IVD library preparation solution for targeted sequencing, Illumina DNA Prep with Enrichment Dx. This kit features innovative on-bead tagmentation transposomes to mediate a normalized reaction. When combined with a simplified, single hybridization step, Illumina DNA Prep with Enrichment Dx provides a rapid library preparation and enrichment solution for IVD use.

This application note demonstrates a DNA-to-data solution that combines Illumina DNA Prep with Enrichment Dx with the NovaSeq 6000Dx instrument and secondary analysis on a paired DRAGEN™ Server with the Illumina DNA Prep with Enrichment Dx app. This solution produces high-quality sequencing data and sensitive detection of germline and somatic variants.

Methods

Sample preparation

For evaluation of germline variant detection, genomic DNA (gDNA) was extracted from CEPH/UTAH Pedigree 1463 NA12877-NA12880 (Coriell Institute).

For evaluation of somatic variant detection, DNA was extracted from FFPE treatments of NA12877 (Coriell Institute, Catalog no. GM12877) and NA12878 (Coriell Institute, Catalog no. GM12878) prepared by Horizon Diagnostics.

Library preparation

Libraries were prepared using Illumina DNA Prep with Enrichment Dx from 50 ng of input DNA for both germline and somatic variant detection.

Sequencing

Prepared libraries were sequenced on the NovaSeq 6000Dx instrument in IVD Mode using a read length of 2 × 151 bp. Data analysis with the DRAGEN for Illumina DNA Prep with Enrichment Dx app was configured as part of sequencing run planning with Illumina Run Manager.

Data analysis

Sequencing data were analyzed using the DRAGEN for Illumina DNA Prep with Enrichment Dx application on the NovaSeq 6000Dx instrument. Data analysis was configured as part of the sequencing run set up with Illumina Run Manager and was launched automatically after the sequencing run completed, eliminating the need for manual touchpoints to initiate analysis after sequencing (Figure 2).

Analysis run times were calculated for both S2 and S4 IVD flow cells and included BCL Conversion, mapping/aligning, and variant calling all within the DRAGEN for Illumina DNA Prep with Enrichment Dx analysis module (Table 1).

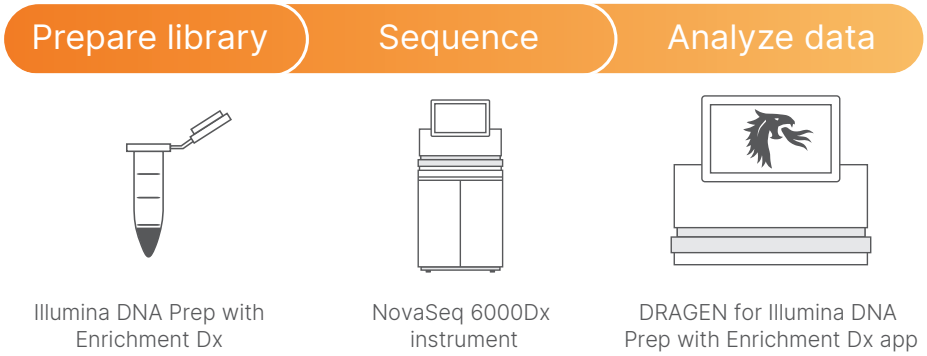

Figure 1: Variant calling workflow—The NovaSeq 6000Dx instrument is part of an integrated, three-step workflow that includes library preparation, high-throughput sequencing in either IVD or RUO Mode, and accelerated secondary data analysis with a paired DRAGEN server.

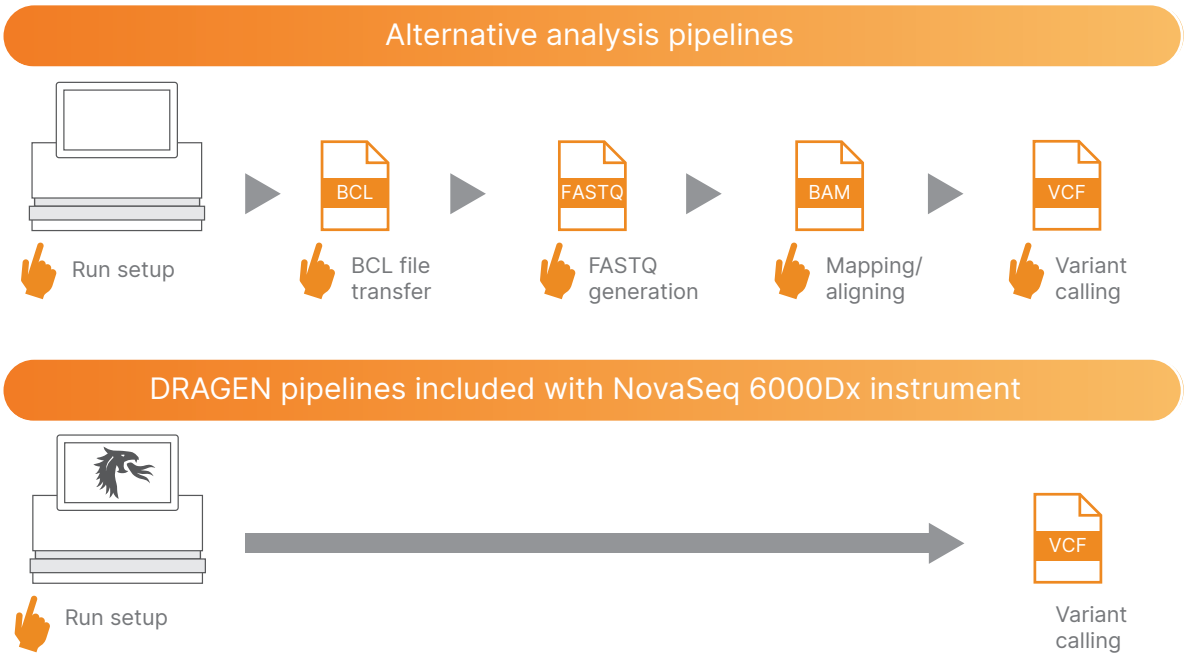

Figure 2: Fewer touchpoints with the DRAGEN platform—Alternative analysis pipelines require multiple manual touchpoints to execute analysis after sequencing. The DRAGEN platform automates many of these steps, reducing manual touchpoints to initial sequencing run setup and downstream tertiary analysis.

FASTQ files were compressed with DRAGEN ORA compression, an option that can be enabled within the DRAGEN for Illumina DNA Prep with Enrichment Dx analysis module, resulting in ~5× smaller file sizes, faster file transfers, and reduced data storage costs (Figure 3).

Table 1: Analysis times on the DRAGEN Server

Parameter	S2 flow cell	S4 flow cell
Workflow	Germline	Somatic
No. of samples per flow cell[a]	47	47
Analysis time[b]	4.3 hours	9.8 hours

a. Human enrichment samples with a 1.9 Mb enrichment panel.
b. Average across 8 sequencing runs.

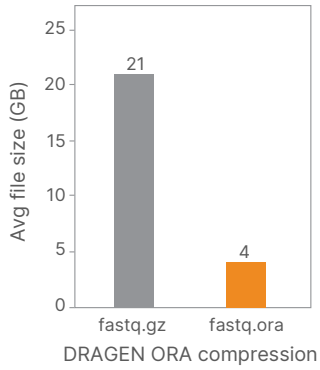

Figure 3: DRAGEN compression of FASTQ files—DRAGEN ORA compression reduces FASTQ file size by ~5.2 times. Average of 98 FASTQ files run on a S4 flow cell in somatic workflow.

Results

Highly accurate, precise variant calling

Accuracy of variant calling was measured by various metrics that evaluated the accuracy of variant calling against a known reference:

- **Positive percent agreement (PPA)**: calculated as the proportion of loci classified as variants by a reference method that the assay correctly reports
- **Negative percent agreement (NPA)**: calculated as the proportion of loci classified as wild type by a reference method that the assay correctly reports
- **Overall percent agreement (OPA)**: calculated as the proportion of loci correctly reported by the assay relative to a reference method.
- **Percent negative calls**: calculated as the number of observations passing reference as the outcome at a position divided by the total number of observations tested, excluding any invalid observations or those filtered as low depth
- **Percent positive calls**: calculated as the number of observations with the variant detected divided by the total number of observations tested, excluding any invalid observations or those filtered as low depth

High-quality sequencing data was obtained, enabling high accuracy and precision for germline and somatic variant calling (Table 2, Table 3).

Table 3: Precision metrics for variant detection

Workflow	Metric	Result
Germline	Percent negative calls	100%
	Percent positive calls	99.7%
Somatic	Percent negative calls	100%
	Percent positive calls	99.2%

Table 2: Accuracy metrics for germline and somatic variant detection

Criteria	No. of observations[a]	Result of observations[b]	Result by run[c]
PPA for germline SNVs	846	99.8	99.9
PPA for germline insertions	846	97.9	> 99.9
PPA for germline deletions	846	96.9	99.9
NPA for germline variant detection	846	> 99.9	> 99.9
OPA for germline variant detection	846	> 99.9	> 99.9
PPA for somatic SNVs	846	99.8	98.9
PPA for somatic insertions	846	100	100
PPA for somatic deletions	846	100	100
NPA for somatic variant detection	846	> 99.9	> 99.9
OPA for somatic variant detection	846	> 99.9	> 99.9

a. Calculated as number of samples per run (47) × number of runs (18) = 846.
b. Lowest observed value by sample replicate across all 18 runs.
c. Lowest value when data from each run are analyzed in aggregate.

Summary

This application note demonstrates the exceptional accuracy and efficiency of a validated, DNA-to-data workflow using Illumina DNA Prep with Enrichment Dx, the NovaSeq 6000Dx instrument, and a paired DRAGEN Server for germline and somatic variant calling.

Learn more

NovaSeq 6000Dx instrument, illumina.com/systems/sequencing-platforms/novaseq-6000dx

Illumina DNA Prep with Enrichment Dx, illumina.com/products/by-type/ivd-products/dna-prep-enrichment-dx

Intended use statements

NovaSeq 6000Dx instrument intended use (European Union/other)

The NovaSeq 6000Dx instrument is intended for sequencing of DNA libraries when used with *in vitro* diagnostic (IVD) assays. The NovaSeq 6000Dx instrument is intended for use with specific registered, certified, or approved IVD reagents and analytical software.

NovaSeq 6000Dx instrument intended use (United States)

The NovaSeq 6000Dx instrument is intended for targeted sequencing of DNA libraries from human genomic DNA extracted from peripheral whole blood or formalin-fixed, paraffin-embedded (FFPE) tissue when used with *in vitro* (IVD) diagnostic assays. The NovaSeq 6000Dx instrument is not intended for whole-genome or *de novo* sequencing. The NovaSeq 6000Dx instrument is intended for use with specific registered, certified, or approved IVD reagents and analytical software.

Illumina DNA Prep with Enrichment Dx (EU/other)

The Illumina DNA Prep with Enrichment Dx is a set of reagents and consumables used to prepare sample libraries from genomic DNA derived from human cells and tissue. User-supplied probe panels are required for the preparation of libraries targeting specific genomic regions of interest. The generated sample libraries are intended for use on Illumina sequencing systems.

Illumina DNA Prep with Enrichment Dx (United States)

The Illumina DNA Prep with Enrichment Dx Kit is a set of reagents and consumables used to prepare sample libraries from DNA extracted from peripheral whole blood and formalin-fixed, paraffin-embedded tissue. User-supplied probe panels are required for the preparation of libraries targeting specific genomic regions of interest. The generated sample libraries are intended for use on Illumina sequencing systems.

illumina®

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

M-GL-00770 v1.0