TECHNICAL NOTE

The NovaSeq™ X Series software ecosystem

Streamlined and comprehensive informatics with onboard secondary analysis and integration with Illumina Connected Software

illumina®

Introduction

The NovaSeq X and NovaSeq X Plus Sequencing Systems are built with breakthrough technological innovations to deliver the throughput and accuracy needed to answer the most complex questions in biology. In addition to advancements in chemistry and optics, the NovaSeq X Series is built with DRAGEN™ hardware onboard the instrument to accelerate and streamline secondary analysis, and compress data by 80% without loss. The NovaSeq X Series also sets a new standard in operational simplicity with a software ecosystem built specifically to support the Illumina next-generation sequencing (NGS) workflow (Figure 1); Illumina Connected Software includes:

- Lab operations—Clarity LIMS™ software
- Run planning and setup—Illumina Run Manager and BaseSpace™ Sequence Hub
- Secondary analysis—BaseSpace Sequence Hub, DRAGEN™ secondary analysis, Illumina Connected Analytics
- Tertiary analysis—Emedgene™ software, Illumina Connected Insights*

With flexible local and cloud-based options for lab operations, run planning, and data analysis, the NovaSeq X Series enables users to run high-throughput sequencing without creating a bioinformatics bottleneck.

This technical note presents the software solutions and platforms that integrate with the NovaSeq X series and provides an overview of the solutions at each stage of the NGS workflow.

Connected Software for every stage of the NGS workflow

Lab operations management

The NovaSeq X Series integrates with Clarity LIMS software leveraging cloud services. Clarity LIMS software is a laboratory information management system (LIMS) designed for efficient sample tracking and workflow management. With Clarity LIMS software, labs can streamline adoption of new workflows with an expanding menu of preconfigured protocols, customize and automate workflows, and efficiently track samples in real time with audit trails.

Learn more about Clarity LIMS software

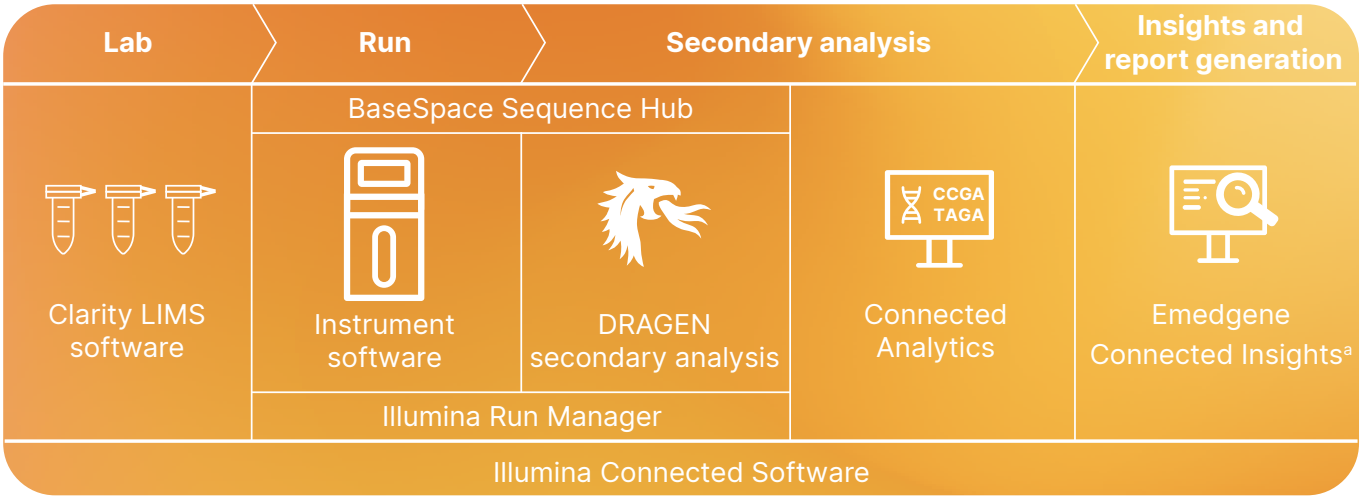

Figure 1: Illumina Connected Software supports the entire NGS workflow—The NovaSeq X Series fully integrates with the Illumina software and informatics portfolio throughout the NGS workflow to deliver a flexible, optimized, and customizable user experience.

a. Not available in all countries. Illumina Connected Insights supports user-defined tertiary analysis through API calls to third-party knowledge sources.

Run and secondary analysis setup

Illumina offers flexible run planning options to meet the varying needs of different labs (Table 1). Run planning includes sequencing run configuration and DRAGEN secondary analysis configuration. Secondary analysis can occur onboard the sequencing system or in the cloud. The benefit is a streamlined workflow, easier data management, and a reduced number of touchpoints.

Illumina Run Manager

The NovaSeq X Series features Illumina Run Manager. Accessed on-instrument or via networked devices, Illumina Run Manager enables control of user and instrument management, application configuration, run setup, and more. With an intuitive interface that provides guided prompts to reduce the chance of error, users can plan and start sequencing runs, track libraries with audit trails, and monitor run progress (Figure 2).

BaseSpace Sequence Hub

Users can plan sequencing runs on the NovaSeq X Series in BaseSpace Sequence Hub, a genomics cloud-computing platform designed to provide simplified data management and analytical sequencing tools in a user-friendly format. BaseSpace Sequence Hub provides remote run setup using an intuitive, graphical interface, and run monitoring with real-time sequencing metrics and history.

Learn more about BaseSpace Sequence Hub

Table 1: Comparison of run planning options

Parameter	Illumina Run Manager	BaseSpace Sequence Hub
User interface	On-instrument or on-premise computer	Web browser
Sample sheet generation	v2	v2
Autolaunch DRAGEN secondary analysis	Yes	Yes

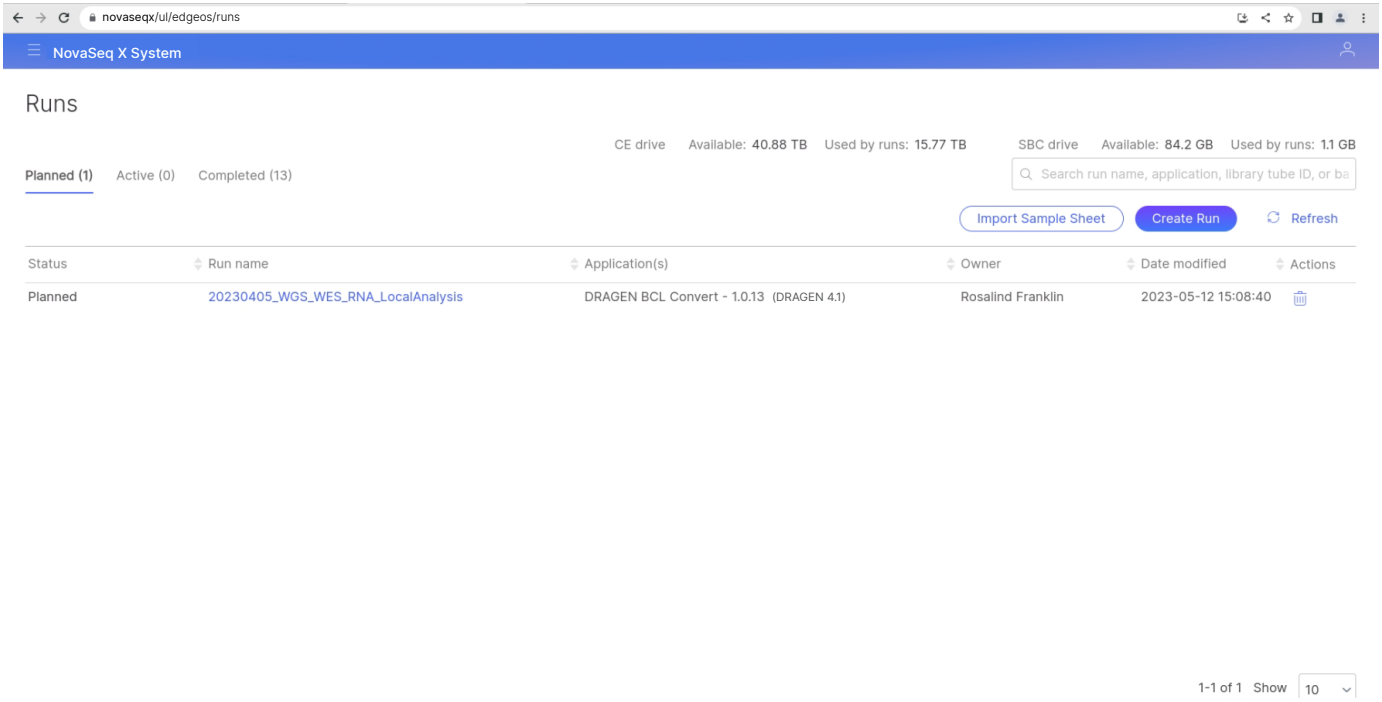

Figure 2: Illumina Run Manager interface—Illumina Run Manager features an intuitive user interface that enables instrument management, application configuration, run setup, and more.

Sequencing

The NovaSeq X and NovaSeq X Plus Systems incorporate thoughtful ergonomic design and usability innovations to simplify operations and optimize the user experience with an extra-large 4K resolution touch screen with an intuitive, informative display. Users can monitor run progress on-instrument or in BaseSpace Sequence Hub to track run QC metrics such as Q30 and yield (Figure 3).

Learn more about the NovaSeq X Series

Secondary analysis

The most common DRAGEN secondary analysis pipelines are available onboard the NovaSeq X Series. Additional pipelines are available via on-premise servers and cloud-based platforms. With faster data transfers, easier data management, and lossless compression that reduces storage costs up to 80%, DRAGEN software provides efficient secondary NGS data analysis. Technical innovations, including DRAGEN Multigenome (graph) and Machine Learning, deliver unprecedented, award-winning accuracy.[1]

Learn more about DRAGEN secondary analysis

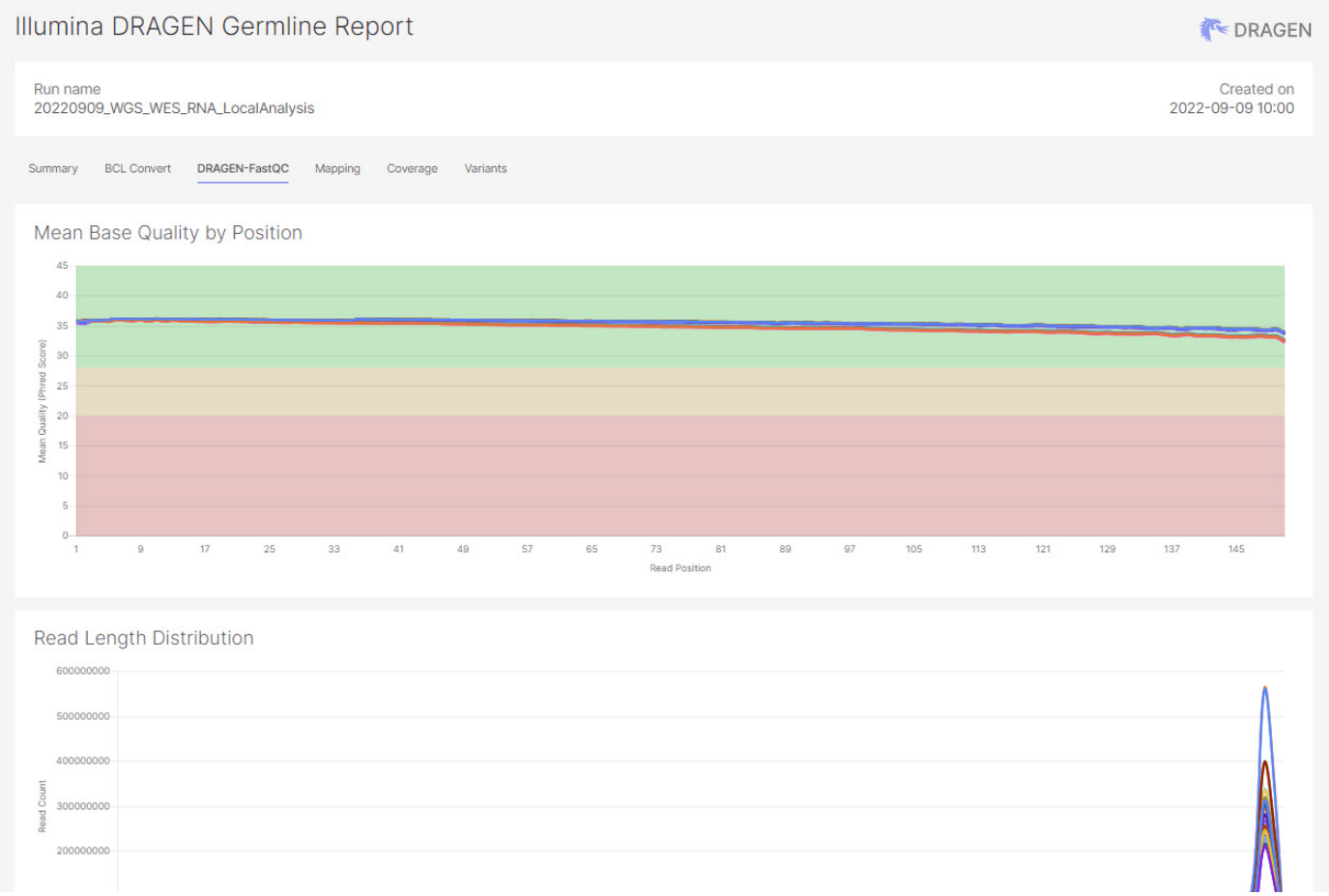

Figure 3 Sequencing run report—The NovaSeq X Series allows users to view run QC metrics and FASTQ reports.

Selecting a data analysis workflow

Flexible run modes

Users have multiple options when setting up a sequencing run on the NovaSeq X Series to launch data analysis automatically, including FASTQ file generation, alignment, and variant calling (Figure 4):

- **Manual**—User manually sets up a sequencing run, with the option to attach a v2 sample sheet to be used in downstream analysis
- **Local**—User sets up sequencing run locally in Illumina Run Manager and data analysis proceeds automatically on the instrument using DRAGEN onboard or on an on-premise computer on the same network
- **Cloud**—User sets up sequencing run in BaseSpace Sequence Hub, data analysis proceeds automatically in the cloud in BaseSpace Sequence Hub or Connected Analytics
- **Hybrid**—User sets up sequencing run in BaseSpace Sequence Hub, data analysis proceeds automatically on the instrument using DRAGEN onboard†

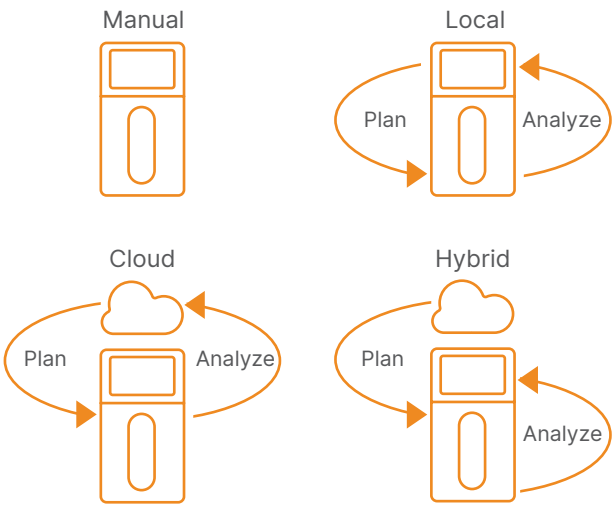

Figure 4: Run mode options—Users can choose from different options when setting up a run on the NovaSeq X Series.

† Hybrid mode will be available in a future software update.

DRAGEN onboard NovaSeq X Series

DRAGEN onboard NovaSeq X Series provides automated, accurate secondary analysis, speeds up computing, and enables lossless data compression for faster data transfers and easier data management. DRAGEN software can run multiple secondary analysis pipelines in parallel, either onboard or in the cloud.

Automated DRAGEN analysis pipelines available on the instrument include:

- DRAGEN BCL Convert
- DRAGEN Germline for whole-genome sequencing (WGS)
- DRAGEN Enrichment for whole-exome sequencing (WES)
- DRAGEN RNA for whole-transcriptome sequencing

DRAGEN onboard NovaSeq X analysis run times can vary given the workflows selected, features enabled, number of samples, sample quality, and yield (Table 2).

When considering back-to-back sequencing runs, there is a 6.5-7 hour "Latency Window" between runs that consists of a wash step, hands-on run set up, and cluster generation for the subsequent run. If DRAGEN analysis will complete within the Latency Window, the second run can be started immediately following the wash step without any impact to analysis of the first sequencing run. If DRAGEN analysis for the first sequencing run will take longer than 6.5 hours, users should wait before starting the second run to make sure that analysis is complete (Figure 5).

Additional access options for DRAGEN secondary data analysis

When assessing DRAGEN secondary analysis options, it is important to consider the needs of a lab and its users. Labs prioritizing ease of use and minimal touch points can select DRAGEN onboard NovaSeq X Series, while labs that are interested in scalability and flexibility can customize DRAGEN software within Connected Analytics or via a DRAGEN server. With the NovaSeq X Series, an expanded menu of DRAGEN pipelines can be accessed through available cloud-based or on-premise solutions, enabling labs to select an option that best suits their needs (Table 3).

Table 2: DRAGEN run times[a]

DRAGEN configuration	No. of samples[b]	Run time[c]
BCL convert (with ORA compression)	24,576	~1 hr 30 min
RNA (with differential expression)	1536 transcriptomes (mRNA)	~4 hr 30 min
Enrichment (Germline)	512 exomes	~4 hr 30 min to 5 hr 30 min
Germline (map align, no variant calling)	48 whole genomes (30×)	~3 hr 30 min
Germline (small variant calling)	48 whole genomes (30×)	~5 hr 30 min to 7 hr[d]
Germline (all variant callers)	48 whole genomes (30×)	~7 hr 30 min[d]

a. Run times may vary based on sample quality, analysis settings, and sample configuration. Data above are representative run times for data generated by Illumina. Customer run times may vary.
b. Number of samples for dual 10B flow cells is based on internal Illumina data.
c. Dual 10B flow cells of the same configuration started simultaneously or single 10B flow cell.
d. Onboard analysis time precludes immediate start of next sequencing run. Assumes 6.5-hour latency window during which DRAGEN analysis must complete to enable back-to-back sequencing.

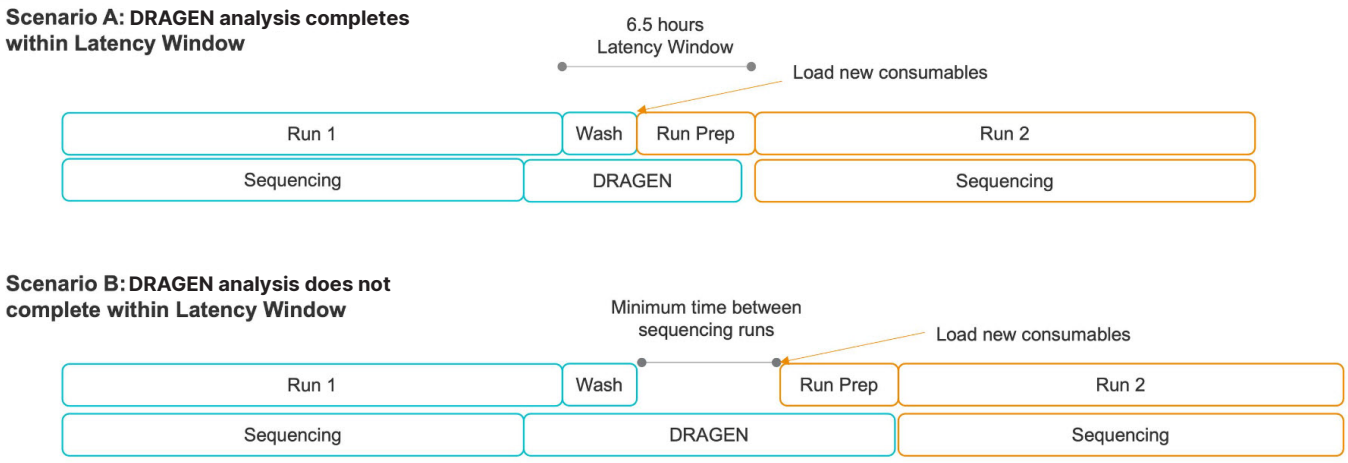

Figure 5: Planning back-to-back sequencing runs—When considering back-to-back sequencing runs on the same side of a NovaSeq X Series, users should consider if DRAGEN analysis will complete within the required Latency Window between runs. If it will, as in scenario A, then the second run can start immediately. If it will not, as in scenario B, users should wait for analysis to complete before starting the second run.

Table 3: DRAGEN deployment options

Deployment option	DRAGEN integrated with NovaSeq X Series	DRAGEN on BaseSpace Sequence Hub	DRAGEN on Connected Analytics	DRAGEN server
Description	Efficient automated onboard workflow	Push-button analysis with intuitive graphical interface	Customizable pipelines in a secure, compliant cloud	Local, customizable server with the latest pipelines
No. of touchpoints[a]	1 (sequencing setup)	2 (sequencing setup, secondary analysis)	1 (sequencing setup)	3 (sequencing setup, manual file transfer, secondary analysis)
File transfers required for analysis	0 (VCF files generated onboard)	1 (BCL files streamed to BaseSpace Sequence Hub)	1 (BCL files streamed to Connected Analytics–accessible cloud bucket)	2 (BCL, FASTQ, or BAM files transferred to network storage and then to DRAGEN server)
Supported applications	BCL convert, WGS (Germline), enrichment (Germline + Somatic), RNA	Full suite of DRAGEN pipelines: RNA, DNA amplicon, Single-cell RNA, Differential expression, RNA amplicon, Methylation, Metagenomics, RNA pathogen detection, COVID lineage, scATAC-Seq, and Imputation		

a. From sequencing setup to VCF generation.

DRAGEN on BaseSpace Sequence Hub

The cloud-based DRAGEN suite available on BaseSpace Sequence Hub combines accurate, efficient analysis with a secure ecosystem and versatile scalability. After encrypted data transfer from the instrument into BaseSpace Sequence Hub, DRAGEN software enables push-button secondary analysis with a curated set of applications.

DRAGEN on Connected Analytics

Connected Analytics is a secure and flexible bioinformatics platform that integrates with sequencing systems for data transfer in real time. Connected Analytics supports highly automated workflows and custom solutions for optimized high-throughput studies. It offers a highly secure environment with guaranteed data residency, single sign-on access, audit logs, and access control supporting international standards. The Connected Analytics Cohorts module enables better study design by allowing users to build and optimize studies with large cohorts of samples.

DRAGEN on-premise server

A DRAGEN server relies on a local repository to collect and store NGS data. After data are transferred from the instrument to the storage solution via a local network connection, data are then moved to the DRAGEN server to perform a selected analysis pipeline.

Tertiary analysis and interpretation

Featuring automated, explainable-AI (XAI)–powered variant interpretation, Emedgene is a comprehensive software platform designed to streamline tertiary analysis workflows for research use. Emedgene offers modular options for a single vendor solution for rare genetic and hereditary disease research applications.

Learn more about Emedgene

Illumina Connected Insights harnesses extensive knowledge sources via powerful API-integrations to support streamlined variant interpretation through report generation, including prioritization of relevant clinical trials, drug labels, and guidelines for somatic oncology applications, soon rare disease, and more.

Learn more about Connected Insights

Correlation Engine is an interactive omics knowledgebase that puts private omics data in biological context with highly curated public data. This software expedites the time to discovery for researchers.

Learn more about Correlation Engine

Summary

The NovaSeq X and NovaSeq X Plus Sequencing Systems provide extraordinary sequencing power to fuel data-intensive applications and large-scale genomics with exceptional accuracy. In addition to numerous technical innovations and advances, the NovaSeq X Series integrates with the Illumina bioinformatics suite of software tools throughout the NGS workflow to maximize operational simplicity, flexibility, and configurability. With local and cloud-based options at each stage that range from simplified push-button apps to customizable, command line interfaces, users can tailor NGS workflows on the NovaSeq X Series to their needs and run high-throughput sequencing without creating a bioinformatics bottleneck.

Learn more

NovaSeq X and NovaSeq X Plus Sequencing Systems

Illumina Connected Software

References

1. Mehio R, Ruehle M, Catreux S, et al. DRAGEN Wins at Precision-FDA Truth Challenge V2 Showcase Accuracy Gains from Alt-aware Mapping and Graph Reference Genomes. illumina.com/science/genomics-research/dragen-wins-precisionfda-challenge-showcase-accuracy-gains.html. Accessed January 12, 2023.

illumina®

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

M-GL-01587 v1.0