illumına®

Title: DRAGEN Enrichment v4.2.7 for NextSeq 1000/2000
Customer Release Notes
Part Number: N/A
Document Number: 200052831, Version: 00
Effective Date: 27-FEB-2024
Page 1 of 13

Customer Release Notes

DRAGEN Enrichment v4.2.7

for NextSeq 1000/2000

Title: DRAGEN Enrichment v4.2.7 for NextSeq 1000/2000
Customer Release Notes
Part Number: N/A
Document Number: 200052831, Version: 00
Effective Date: 27-FEB-2024
Page 2 of 13

Table of Contents

Title: DRAGEN Enrichment v4.2.7 for NextSeq 1000/2000
Customer Release Notes
Part Number: N/A
Document Number: 200052831, Version: 00
Effective Date: 27-FEB-2024
Page 3 of 13

INTRODUCTION

These Release Notes detail the key changes to the DRAGEN Enrichment Workflow onboard the NextSeq 1000/2000 instrument, since the v3.10.12 release. If you are updating from an older version, please see all intermediate release notes.

NEW FEATURES

- DRAGEN™ v4.2 offers significant improvements in accuracy, added features for a more comprehensive solution, and efficiency improvements. For full extensive details on each feature of pipeline, please consult the latest Illumina DRAGEN™ Bio-IT Platform User Guide available on the support website at https://support.illumina.com/downloads/illumina-dragen-bio-it-platform-user-guide.html
- **Reference Genome**
  - Enhanced multigenome (graph) reference v3 and Machine Learning (ML) models improves small variant calling accuracy.
    - DRAGEN v4.2 workflows require the installation of new reference genome Hash Tables (v9).
    - Pre-built instrument reference genomes can be downloaded from the NextSeq 1000/2000 Sequencing System Software Downloads page, for offline installation.
    - Please reference the NextSeq 1000/2000 Compatible Products page on the Illumina support site, for the list of available reference genomes and recommended genomes for different workflows.
  - New reference updates for hg38 improves variant calling in the Challenging, Medically Relevant Genes (CMRG) regions.
    - GRCh38 reference includes 34 sequences from chm13 and hs37d5 as decoys.
    - 29 decoys identified as missing segmental duplications.
    - 5 decoys identified in acrocentric arms of chromosomes 13, 14, 15, and 22 of CHM13.
    - Yield accuracy improvements in the CMRG genes: FANCD2, MAP2K3, KCNJ18, and KMT2C, as well as in the Y chromosome.
  - Support for the CHM13 v2.0 reference is available for research purposes. Only small variants have been fully validated.
    - Introducing support for the Telomere-to-Telomere CHM13v2.0 reference.
    - Accuracy has been validated only for WGS samples and small variant ML calls.
    - Accuracy tested on CMRG truth set.
- **Mapper/Aligner**
  - DNA and RNA Alignment updated.
    - DRAGEN mapper implemented a much more rigorous method for determining split-read alignments and influencing primary alignments and MAPQs with split-read analysis.
    - This sophisticated new method can support up to a maximum of 4095 secondary alignments and 4095 supplementary alignments per read.

illumina®

Title: DRAGEN Enrichment v4.2.7 for NextSeq 1000/2000
Customer Release Notes
Part Number: N/A
Document Number: 200052831, Version: 00
Effective Date: 27-FEB-2024
Page 4 of 13

- o New skew normal insert size model and pairing penalty function
  - Skew normal insert model better fits observed real-world asymmetric insert size distributions.
  - New PDF-based pairing penalty to avoid excessively penalizing larger insert size proper pairs.
  - Enabled by default for DNA pipelines, resulting in increased proper pairing rate.
  - RNA insert model unchanged.
- **Germline Small Variant Caller**
  - o ML model updates further improves small variant calling accuracy.
    - ML is enabled by default for human samples.
    - ML is supported on hg38, hg19, hs37d5, CHM13 references.
    - Supported for Germline and Enrichment workflows. Not supported for Somatic, RNA and Amplicon workflows.
    - ML will automatically be enabled/disabled for supported/unsupported workflows and reference types.
    - Accuracy improvements at high depths (> 100x) WGS.
    - DGT, DGQ and DQUAL for the non-ML variant caller are no longer included in the VCF when ML is enabled.
- **Somatic Small Variant Caller**
  - o Better T/O accuracy with improved Systematic Noise Filter + FFPE Blocklists.
    - New noise files result in approximate 30% fewer FP calls.
    - The default noise files are updated, and settings for custom generated noise files are also improved.
    - See notes below for importing and using noise baseline files.
- Improved **CNV and Structural Variant** calling accuracy.

NOTES

Reference Genome Usage Recommendations

Germline, Enrichment Germline, Amplicon Germline Workflows

- Use ALT-Masked mapping.
- Use Graph genomes.
- Use the Homo sapiens [1000 Genomes] hg38 Alt Masked Graph v3 reference genome.
- If not using hg38, the Homo sapiens [UCSC] hg19 Alt Masked Graph v3 or Homo sapiens [NCBI] hs37d5 v3 Graph are recommended.
- Non-graph is supported, but has reduced accuracy

Enrichment Somatic, Amplicon Somatic, RNA and scRNA Workflows

- Use ALT-Masked mapping.
- Use non-Graph genomes.

Title: DRAGEN Enrichment v4.2.7 for NextSeq 1000/2000
Customer Release Notes
Part Number: N/A
Document Number: 200052831, Version: 00
Effective Date: 27-FEB-2024
Page 5 of 13

- Use the Homo sapiens [1000 Genomes] hg38 Alt Masked v3 reference genome or Homo sapiens [UCSC] hg19 Alt Masked v3 reference genome.

Table 1 v4.2 Reference Support and Recommended Use for Human Data

Human		hg19	hs37d5	hg38	chm13	Recommended Reference Type
Germline	SNV	Yes	Yes	Yes	Yes	Graph
	CNV	Yes	Yes	Yes	Yes*	Graph
	SV	Yes	Yes	Yes	Yes*	Graph
	Expansion Hunter	Yes	Yes	Yes	No	Graph
	Targeted Callers	Yes	Yes	Yes	No	Graph
	RNA	Yes	Yes	Yes	Yes*	Non-Graph
	De Novo	Yes	Yes	Yes	Yes*	Graph
	Joint Genotyping	Yes	Yes	Yes	Yes*	Graph
	Biomarkers (HLA)	Yes	Yes	Yes	Yes*	Graph
	Gvcf Genotyper	Yes	Yes	Yes	Yes*	Graph
Somatic	SNV	Yes	Yes	Yes	Yes*	Non-Graph
	UMI SNV	Yes	Yes	Yes	Yes*	Non-Graph
	CNV	Yes	Yes	Yes	Yes*	Non-Graph
	SV	Yes	Yes	Yes	Yes*	Non-Graph
Methylation	Methylation	Yes	Yes	Yes	No	Non-Graph
Annotation	Illumina Annotation Engine (Nirvana)	Yes	Yes	Yes	No	n/a

(*) DRAGEN™ supports the component execution; however, the component's accuracy has not been established.

Table 2 v4.2 Reference Support and Recommended Use for Non-Human Data

Non-Human		Supported	Recommended Reference Type
Germline	SNV	Yes	Non-Graph
	CNV	No	n/a
	SV	Yes	Non-Graph
	Expansion Hunter	No	n/a
	Targeted Callers	No	n/a
	RNA	Yes	Non-Graph
	De Novo	Yes	Non-Graph
	Joint Genotyping	Yes	Non-Graph
	Biomarkers (HLA)	No	n/a
	Gvcf Genotyper	Yes	Non-Graph
Somatic	SNV	No	n/a
	UMI SNV	No	n/a
	CNV	No	n/a
	SV	No	n/a
Methylation	Methylation	No	n/a
Annotation	Illumina Annotation Engine (Nirvana)	Yes	n/a

Title: DRAGEN Enrichment v4.2.7 for NextSeq 1000/2000
Customer Release Notes
Part Number: N/A
Document Number: 200052831, Version: 00
Effective Date: 27-FEB-2024
Page 6 of 13

How to Update Instrument Reference Genomes

You can only import new reference genomes from the administrator account.
Steps:

1. Download the desired genome package tar.gz from the NextSeq 1000/2000 Software Downloads page, OR
2. Create a reference genome using the Reference Builder for Illumina Instruments BaseSpace Sequence Hub app. For more information, refer to Reference Builder for Illumina Instruments v1.0.0 App Online Help.
3. Select the control software menu, and then select **Process Management**. Make sure that there are no sequencing runs or on-instrument secondary analyses in progress.
4. From the control software menu, select **Minimize Application**.
5. Log into ilmnadmin.
6. Select the control software menu, and then select **DRAGEN**.
7. In the Genome section, select **View Installed Genomes** to view a list of all currently installed genomes.
8. Close the window.
9. Under Import New Reference Genomes, select Choose.

Navigate to the reference genome file (*.tar.gz) on the portable or mounted network drive, and then select **Open**.

Resource Files

Available Resource Files

DRAGEN requires resource files to function correctly and achieve the optimum performance for certain workflows.

All resource files are available for download at the Illumina DRAGEN Product Files support site here: https://support.illumina.com/sequencing/sequencing_software/dragen-bio-it-platform/product_files.html

These resources can be added as Aux file to the instrument.

Workflow	DRAGEN Version	Resource	Description
Enrichment - Somatic tumor-only mode	4.2.7	SNV Systematic Noise Baseline collection v1.1.0	A collection of noise baseline BED files for hg19, hs37d5, hg38 and for WGS and WES respectively. systematic-noise-baseline-collection-1.1.0.tar Extract the .tar archive and select the appropriate noise bed.gz file to download to the instrument. Then add the noise baseline argument to the sample sheet. If a SNV noise baseline is not used for somatic tumor only analysis, the accuracy may be sub-optimal. The tar archive contains the following files.

Title: DRAGEN Enrichment v4.2.7 for NextSeq 1000/2000
Customer Release Notes
Part Number: N/A
Document Number: 200052831, Version: 00
Effective Date: 27-FEB-2024
Page 7 of 13

Workflow	DRAGEN Version	Resource	Description
			snv_wes_nextera_hg19_max_v1.1_systematic_noise.bed.gz snv_wes_nextera_hg19_mean_v1.1_systematic_noise.bed.gz snv_wes_nextera_hg38_max_v1.1_systematic_noise.bed.gz snv_wes_nextera_hg38_mean_v1.1_systematic_noise.bed.gz snv_wes_nextera_hs37d5_max_v1.1_systematic_noise.bed.gz snv_wes_nextera_hs37d5_mean_v1.1_systematic_noise.bed.gz snv_wes_truseq_hg19_max_v1.1_systematic_noise.bed.gz snv_wes_truseq_hg19_mean_v1.1_systematic_noise.bed.gz snv_wes_truseq_hg38_max_v1.1_systematic_noise.bed.gz snv_wes_truseq_hg38_mean_v1.1_systematic_noise.bed.gz snv_wes_truseq_hs37d5_max_v1.1_systematic_noise.bed.gz snv_wes_truseq_hs37d5_mean_v1.1_systematic_noise.bed.gz snv_wgs_hg19_max_v1.1_systematic_noise.bed.gz snv_wgs_hg19_mean_v1.1_systematic_noise.bed.gz snv_wgs_hg38_max_v1.1_systematic_noise.bed.gz snv_wgs_hg38_mean_v1.1_systematic_noise.bed.gz snv_wgs_hs37d5_max_v1.1_systematic_noise.bed.gz snv_wgs_hs37d5_mean_v1.1_systematic_noise.bed.gz
	3.10.12 3.8.4	SNV Systematic Noise Baseline collection v1.0.0	A collection of noise baseline BED files for hg19, hs37d5, hg38 and for WGS and WES respectively systematic-noise-baseline-collection-1.0.0.bin The bin file is a self-extracting binary. Execute the bin file on a Linux system, and select the appropriate noise bed file to download to the instrument, and add the noise baseline argument to the sample sheet. If a SNV noise baseline is not used for somatic tumor only analysis, the accuracy may be sub-optimal. $ chmod +x systematic-noise-baseline-collection-1.0.0.bin $ ./systematic-noise-baseline-collection-1.0.0.bin WES_Nextera_IDT_hg19_v1.0_systematic_noise.bed.gz WES_Nextera_IDT_hg38_v1.0_systematic_noise.bed.gz WES_Nextera_IDT_hs37d5_v1.0_systematic_noise.bed.gz WES_TruSeq_IDT_hg19_v1.0_systematic_noise.bed.gz WES_TruSeq_IDT_hg38_v1.0_systematic_noise.bed.gz WES_TruSeq_IDT_hs37d5_v1.0_systematic_noise.bed.gz WGS_hg19_v1.0_systematic_noise.bed.gz WGS_hg38_v1.0_systematic_noise.bed.gz WGS_hs37d5_v1.0_systematic_noise.bed.gz
Enrichment - Germline CNV Calling	4.2.7 3.10.12	CNV Panel of Normals for Twist Bioscience for Illumina Exome 2.5 Panel	Pre-built TruSight PON for enrichment germline CNV, for hg19, hg38, hs37d5 The PON is a zip file. The file contains panels for different genomes and DRAGEN versions. Download the zip file and extract it to get *.combined.counts.gz files. Import the applicable .combined.counts.gz file to the instrument and add the PON argument to the sample sheet. If a PON is not provided, Enrichment CNV analysis will be disabled, but the run will succeed with other callers. The following PON files can be used on NextSeq1k2k: "NextSeq 2000 / DRAGEN 3.10 / <ref> / *combined.counts.gz" - can be used with DRAGEN v3.10.12 "NovaSeq 6000 / DRAGEN 4.2 / <ref> / *combined.counts.gz" - can be used with DRAGEN v4.2.7. It is recommended to generate a PON using DRAGEN v4.2.7 and NextSeq 2000 data.

Title: DRAGEN Enrichment v4.2.7 for NextSeq 1000/2000
Customer Release Notes
Part Number: N/A
Document Number: 200052831, Version: 00
Effective Date: 27-FEB-2024
Page 8 of 13

How to Import Noise Baseline Files

If using the DRAGEN Enrichment workflow in somatic mode, a noise baseline file should be used to filter out sequencing or systematic noise.

Import Baseline Files Using the User Interface

After importing the baseline file, a sequencing run can be set up using the DRAGEN Enrichment workflow in somatic mode.

Steps:

1. Download a standard baseline file from the Illumina Support Site, or download the custom baseline file from the DRAGEN server or DRAGEN Baseline Builder App.
2. From the control software menu, select **Minimize Application**.
3. Log in to ilmnadmin.
4. Select **Places**, and then select **Computer**.
5. Double-click **usr**, and then **local**.
6. Double-click **illumina**, and then **aux_files**.
7. Drag the noise baseline file to aux_files.

Import Baseline Files Using Terminal

After importing the baseline file, a sequencing run can be set up using the DRAGEN Enrichment workflow in somatic mode.

Steps:

1. Download a standard baseline file from the Illumina Support Site, or download the custom baseline file from the DRAGEN server or DRAGEN Baseline Builder App.
2. From the control software menu, select **Minimize Application**.
3. Log in to ilmnadmin.
4. Select **Applications**.
5. Under Favorites, select **Terminal**.
6. Enter the following command.
cp [/path/to/baselinefile] /usr/local/illumina/aux_files

Generate a Custom Noise Baseline File

The noise baseline file is built using normal samples that do not match to the subject that the samples are from. The recommended number of normal samples is 50.

To generate a custom noise baseline file, use one of the following methods:

• Use the DRAGEN Bio-IT Platform server. Refer to the DRAGEN Bio-IT Platform Online Help for instructions.
• Use DRAGEN Baseline Builder App on BaseSpace Sequence Hub.

Title: DRAGEN Enrichment v4.2.7 for NextSeq 1000/2000
Customer Release Notes
Part Number: N/A
Document Number: 200052831, Version: 00
Effective Date: 27-FEB-2024
Page 9 of 13

How to Import Panel of Normals for CNV Calling

A panel of normals (PON) is needed if enabling CNV calling in the DRAGEN Enrichment pipeline. The CNV caller is a reference-based normalization algorithm that uses additional externally supplied matched normal samples to determine a baseline level from which to call CNV events. These matched normal samples should be derived from the same library prep and sequencing workflow that is used for the case sample. The algorithm subtracts system level biases that are not sample specific.

A custom panel can be built, or a panel supplied by Illumina can be used.

Import the Panel of Normals Using the User Interface

> Steps:
> 1. Download a Panel of Normals from the Illumina Support Site, or download the custom panel of normals from the DRAGEN Baseline Builder App.
> 2. From the control software menu, select **Minimize Application**.
> 3. Log in to ilmnadmin.
> 4. Select **Applications**, and then select **Favorites**.
> 5. Select **+Other Locations**, and then select **Computer**.
> 6. Double-click **usr**, and then **local**.
> 7. Double-click **illumina**, and then **aux_files**.
> 8. Drag the Panel of Normals file (*.combined.counts.txt.gz file or *.gc-corrected.<ref>.gz file) to aux_files

Import Panel of Normals Using Terminal

> Steps:
> 1. Download a provided Panel of Normals from the Illumina Support Site, or download the custom panel of normals from the DRAGEN Baseline Builder App.
> 2. From the control software menu, select **Minimize Application**.
> 3. Log in to ilmnadmin.
> 4. Select **Applications**.
> 5. Under Favorites, select **Terminal**.
> 6. Enter the following command.
> cp [/path/to/baselinefile] /usr/local/illumina/aux_file

Generate a Panel of Normals

Use the DRAGEN Baseline Builder BaseSpace App, which accepts FASTQ, BAM, or CRAM and generates a CNV Baseline file (*.combined.counts.txt.gz file).

It is recommended to use ~50 samples to build the baseline.

KNOWN ISSUES

Known issues of the DRAGEN™ v4.2.7 Enrichment Workflow

Title: DRAGEN Enrichment v4.2.7 for NextSeq 1000/2000
Customer Release Notes
Part Number: N/A
Document Number: 200052831, Version: 00
Effective Date: 27-FEB-2024
Page 10 of 13

Comp	ID	Summary	Resolution/Workaround
CNV VC	DRAGEN-28695	Up to v4.0, if we cannot detect a purity or ploidy (DEGENERATE_DIPLOID) then we would FAIL all records in our VCF. In v4.2, we changed this behavior for WGS, but not for WES.	No workaround for WES. A fix has been made to future version.
Multi genome reference	DRAGEN-26308	WGS runtime increased with multigenome vs legacy genome	5% longer run time. For information only.
SNV Germline	DRAGEN-23801	Joint Calling in Mito is not giving proper VAF's, when one or more samples have a variant, but other samples have a homref call at the same position.	For some alleles, the AD values in the joint VCF are not accurate. Looking at the corresponding single sample gVCF can resolve the inconsistency.
SNV Germline	DRAGEN-26359	Small regression in INDEL sensitivity in v4.2 compared to v4.0	Single sample SNV has major accuracy improvements due to graph and reference and machine learning updates. In some cases, such as INDEL, there are very minor changes where precision may improve at cost of recall, and vice versa. For information only.
SNV VC	DRAGEN-25905	Hang observed on high depth samples, when target BED is used to run the SNV caller over regions which are close to the end of a chromosome.	Have more BED regions throughout the chromosome or increase bin memory.
SNV VC	DRAGEN-25933	VCF GQ values may not match VCF specification	In most positions, the probability that the position is a variant is very close to 1 and the impact is negligible. In corner cases where p(0/0) is not negligible, we have the wrong value in the GQ field. For information only.
SNV Somatic	DRAGEN-22241	Some T/N and T/O samples have >5% runtime regression relative to v4.0	No workaround
SNV Somatic	DRAGEN-24060	Small regression in INDEL FP across most T/O WGS and WES test datasets	Single sample SNV has major accuracy improvements in v4.2 due to graph mapper and reference updates. In some cases, such as INDEL, there are very minor changes where precision may improve at cost of recall, and vice versa. For information only.
SNV Somatic	DRAGEN-29580	Padding an interval does not always produce the same results as running with larger interval, for somatic VC.	Specific to Somatic VC. Use larger regions instead of padding. A fix has been made to future version.

Title: DRAGEN Enrichment v4.2.7 for NextSeq 1000/2000
Customer Release Notes
Part Number: N/A
Document Number: 200052831, Version: 00
Effective Date: 27-FEB-2024
Page 11 of 13

DRAGEN v4.2.7 workflows are compatible with Control Software 1.7.x. For information regarding compatibility with other Control Software versions, please reference the NextSeq 1000/2000 Compatible Products page on the Illumina support site. If you would like to update the control software, please follow the steps detailed in the NextSeq 1000/2000 Control Software Suite v1.7.x Release Notes on the Illumina Support Site.

Online Installation of Workflows

If the instrument is connected to the internet, you can install DRAGEN workflows directly from the Control Software. Online installation of workflows is available since Control Software v1.3 or later.

Steps:
1. Make sure that you have the password to the *ilmnadmin* account.
2. Log in to the *ilmnadmin* account:
   a. If you are logged in as *ilmnuser,* and in control software, select the control software menu, and then select **Exit Application** to access the desktop.
   b. Select the power button icon in the upper right corner and log out of *ilmnuser*.
   c. After you are on the login screen, select *ilmnadmin*, and then enter the password to log in.
   d. The control software automatically launches once you are logged in.
3. Make sure that there are no sequencing runs or on-instrument secondary analysis in progress.
4. On the control software menu, select **DRAGEN**. Under Version, the Available Workflows section lists the workflows currently installed on the system.
5. Select **Check Online**. Not all DRAGEN versions and workflows are compatible with online installation. Use offline installation for additional workflows.
6. Select the checkbox for the workflows that you would like to install. **NOTE: For Online Updates, DRAGEN BCL Convert must be installed before or with other workflows of the same DRAGEN Version.** You can view information about the latest version of a workflow in the release notes.
7. Select Install to start installation.
8. Enter *ilmnadmin* for the system password, and then select Authenticate.
9. After installation is complete, you will be navigated back to the DRAGEN screen and can view the updated list of installed DRAGEN workflows.

Offline Installation of Workflows

Steps:
1. When a DRAGEN workflow update is available, download the installer (*.tar.gz) from the NextSeq 1000/2000 Sequencing System support page. Save the installer to a local or portable drive.
2. If you saved the installer to a portable drive, plug the drive into a USB 3.0 port, located on both the side and back of the instrument. Gently move the instrument as needed to access the back.
3. Follow steps 1-3 above to log in to *ilmnadmin*.
4. Select the control software menu, and then select **DRAGEN**.
5. Under Version, select **Browse for New Version** to navigate to the installer.
6. Select Install to start installation.

Title: DRAGEN Enrichment v4.2.7 for NextSeq 1000/2000
Customer Release Notes
Part Number: N/A
Document Number: 200052831, Version: 00
Effective Date: 27-FEB-2024
Page 12 of 13

7. Enter *ilmnadmin* for the system password, and then select Authenticate.
8. After installation is complete, you will be navigated back to the DRAGEN screen and can view the updated list of installed DRAGEN workflows.

Note that starting with control software v1.5, it is possible to uninstall previous versions of DRAGEN workflows.

Title: DRAGEN Enrichment v4.2.7 for NextSeq 1000/2000
Customer Release Notes
Part Number: N/A
Document Number: 200052831, Version: 00
Effective Date: 27-FEB-2024
Page 13 of 13

RELEASE HISTORY

Revision	Release Reference	Originator	Description of Change
00	CN 1103441	Yi Lian	Initial release