DRAGEN v4.3.13 Software Release Notes

Introduction

These release notes detail the key changes to software components for the Illumina® DRAGEN™ Secondary Analysis Software v4.3.13.

Changes are relative to DRAGEN™ v4.3.6. If you are upgrading from a version prior to DRAGEN™ v4.3, please review the release notes for a list of features and bug fixes introduced in subsequent versions.

DRAGEN™ Installers, Resource Files, and Release Notes are available here:
https://support.illumina.com/sequencing/sequencing_software/dragen-bio-it-platform.html

DRAGEN™ User Guide is now available here:
https://help.dragen.illumina.com

The software package includes downloadable installers for Phase 3 and Phase 4 on-site servers:
- DRAGEN™ SW for x86 Oracle 8 - dragen-4.3.13-11.multi.el8.x86_64.run
- DRAGEN™ SW for x86 CentOS 7 - dragen-4.3.13-11.multi.el7.x86_64.run

The following configurations containing DRAGEN™ 4.3.13 are also available on request:
- AlmaLinux 8 and CentOS 7 based Amazon Machine Images (AMIs) for f1 instances, available in 12 regions
- CentOS 7 Microsoft Azure Image (VM) available in West US 2 for BYOL
- el8 and el7 compatible RPM packages for use with Amazon Web Services (AWS) f1 instances, for customer generated AMIs or customer generated docker images
- DRAGEN™ Kernel drivers for el8 and el7, for use with customer generated AMIs or QuickStart

DRAGEN™ v4.3.13 is also made available on:
- Illumina BaseSpace and ICA platforms
- AWS and Azure Marketplaces
  - On AWS see "DRAGEN Complete Suite"
  - On Azure see "DRAGEN Public VM Image - PAYG"

Deprecated platforms:
- Support for CentOS 7 ended on June 30, 2024. DRAGEN™ v4.3 is the final release with CentOS 7 installers. Future releases will support el8 and el9.
- Support for DRAGEN™ Server v1 FPGA cards have been deprecated since DRAGEN™ v3.10
- Support for Ubuntu has been deprecated since DRAGEN™ v3.9
- Support for CentOS 6 has been deprecated since DRAGEN™ v3.8

Contents

Overview

DRAGEN™ v4.3.13 is a patch release on v4.3 that is available for server, cloud and on-instrument. It contains:

- Various bug fixes
- Updates to support MiSeq i100 on-instrument analysis.
- Updates to Novaseq X on-instrument analysis.

For full extensive details on each feature of pipeline, please consult the latest Illumina DRAGEN™ Software User Guide available at https://help.dragen.illumina.com

Please review the section on Known Issues and limitations of the release.

Resource Files

Switching to DRAGEN™ v4.3 from a prior version requires updates to key resource files to function correctly and achieve the optimum performance. Additional resource files are made available for v4.3. 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

Recent changes to resource files:

- Added CNV panel of normals collection for v4.3
- Added ORA compression reference for duck

The following resource files are available:

Resource	Description	File name(s)
Hash Tables v10	Pre-built v10 pangenome and linear hash tables for hg38, hg19, hs37d5, CHM13.	Pangenome: hg38-alt_masked.cnv.graph.hla.rna-10-r4.0-1.tar.gz hg19-alt_masked.cnv.graph.hla.rna-10-r4.0-1.tar.gz hs37d5-cnv.graph.hla.rna-10-r4.0-1.tar.gz chm13_v2-cnv.graph.hla.rna-10-r4.0-1.tar.gz Linear: hg38-alt_masked.cnv.hla.methylated_combined.rna-10-r4.0-1.tar.gz hg19-alt_masked.cnv.hla.methylated_combined.rna-10-r4.0-1.tar.gz hs37d5-cnv.hla.methylated_combined.rna-10-r4.0-1.tar.gz chm13_v2-cnv.hla.methylated_combined.rna-10-r4.0-1.tar.gz
Multigenome Reference Builder Collection v4	HT mask BED, Graph BED, Graph exclusion BED, Graph msVCF and FASTA files for building hg38, hg19, hs37d5, chm13 references.	hg38-multigenome_reference_collection-v4.tar.gz hg19-multigenome_reference_collection-v4.tar.gz hs37d5-multigenome_reference_collection-v4.tar.gz chm13_v2-multigenome_reference_collection-v4.tar.gz
SNV Systematic Noise Baseline collection v2.0.0	A collection of Somatic noise baseline BED files for hg19, hs37d5, hg38 and for WGS and WES respectively. New files for Heme and FFPE WGS for hg38.	systematic-noise-baseline-collection-2.0.0.tar The tar archive contains the following files: IDPF_WGS_hg38_v.2.0.0_systematic_noise.snv.bed.gz FFPE_WGS_hg38_v2.0.0_systematic_noise.snv.bed.gz WGS_hg38_v2.0.0_systematic_noise.snv.bed.gz WGS_hg19_v2.0.0_systematic_noise.snv.bed.gz WGS_hs37d5_v2.0.0_systematic_noise.snv.bed.gz WES_hg38_v2.0.0_systematic_noise.snv.bed.gz WES_hg19_v2.0.0_systematic_noise.snv.bed.gz WES_hs37d5_v2.0.0_systematic_noise.snv.bed.gz
SV Systematic Noise Baseline collection v3.0.0	A collection of Somatic noise baseline BEDPE files for WGS hg19, hs37d5, hg38. New file for Heme WGS hg38.	sv-systematic-noise-baseline-collection-3.0.0.tar The tar archive contains the following files: IDPF_WGS_hg38_v3.0.0_systematic_noise.sv.bedpe.gz WGS_hg19_v3.0.0_systematic_noise.sv.bedpe.gz WGS_hg38_v3.0.0_systematic_noise.sv.bedpe.gz WGS_hs37d5_v3.0.0_systematic_noise.sv.bedpe.gz
CNV Population SNP VCF v1.0.0	Population SNP VCF for Somatic TO CNV for hg38, hg19, hs37d5 and chm13	Files from the GATK resource bundle uploaded for convenience: hg38_1000G_phase1.snps.high_confidence.vcf.gz hg19_1000G_phase1.snps.high_confidence.vcf.gz hs37d5_1000G_phase1.snps.high_confidence.vcf.gz chm13_1000G_phase1.snps.high_confidence.vcf.gz

CNV Panel of Normals for DRAGEN v4.3 v1.0	Collection of PON files (combined.counts.txt.gz) for exome	`CNV_PON_ExomeV2_DRAGEN_v4.3.tar.gz` The tar archive contains the following panels of normals: For each reference hg19, hg38, hs37d5, a panel generated from 54 samples, Illumina DNA Prep with Enrichment protocol, pooled by mass, overnight hybridization, sequencing on NovaSeq 6000 and NextSeq 2000, and generated for pangenome and linear reference usage.

SNV Exclusion BED collection v1.0.0	Somatic SNV ALU region exclusion BED files for hg38, hg19, hs37d5	`bed-file-collection-1.0.0.tar.gz`
		
		The tar archive contains the following files:
		`v1.0.0_hg38_Alu_regions.bed.gz`
		`v1.0.0_hg19_Alu_regions.bed.gz`
		`v1.0.0_hs37d5_Alu_regions.bed.gz`

Microsattelite Files v1.0.0	Microsattelite files for hg19, hs37d5, hg38 and for WGS and WES respectively	`microsatellite-files-1.0.0.tar.gz`
		
		The tar archive contains the following files:
		`WGS_v1.0.0_hg38_microsatellites.list`
		`WGS_v1.0.0_hg19_microsatellites.list`
		`WGS_v1.0.0_hs37d5_microsatellites.list`
		`WES_v1.0.0_hg38_microsatellites.list`
		`WES_v1.0.0_hg19_microsatellites.list`
		`WES_v1.0.0_hs37d5_microsatellites.list`

Imputation Reference Panel v2.1 and Genetic Map v2.0	Genetic map and reference panel for hg38	`genetic_maps-hg38-2.0.tar` `irp-hg38-2.1.2.0.tar`

ORA compression references	Compression references for human, methylated and non-human	Human:
		`oradataV2.tar.gz`
		
		Non-human and human methylated:
		Arabidopsis: `oradata_arabidopsis_thaliana.tar.gz`
		Cattle: `oradata_bos_taurus.tar.gz`
		Duck: `oradata_cairina_moschata.tar.gz`
		Roundworm: `oradata_caenorhabditis_elegans.tar.gz`
		Zebrafish: `oradata_danio_rerio.tar.gz`
		Chicken: `oradata_gallus_gallus.tar.gz`
		Soybean: `oradata_glycine_max.tar.gz`
		Human: `oradata_homo_sapiens.tar.gz`
		Human Bisulfite: `oradata_homo_sapiens_bisulfite.tar.gz`
		Mouse: `oradata_mus_musculus.tar.gz`
		Rice: `oradata_oryza_sativa.tar.gz`
		Rat: `oradata_rattus_norvegicus.tar.gz`
		Pig: `oradata_sus_scrofa.tar.gz`
		Wheat: `oradata_triticum_aestivum.tar.gz`
		Maize: `oradata_zea_mays.tar.gz`
		
		Combined all species:
		`oradata_all_species.tar.gz`

NOTES:

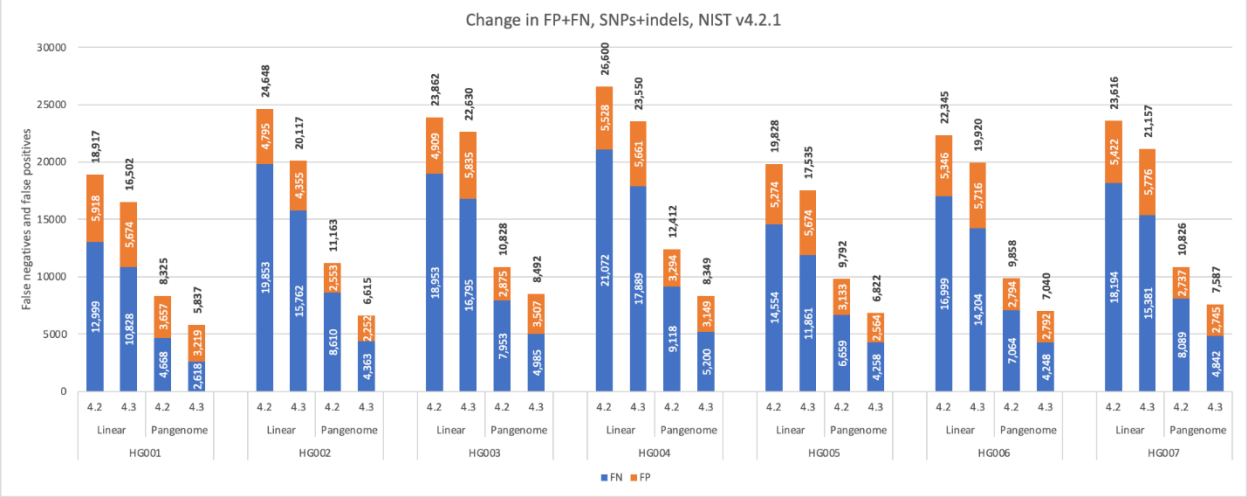
- ML Model files are included in the installer by default since v4.2 and does not need to be downloaded.

Reference Genome Support

DRAGEN supports the construction of reference hash tables for both human and non-human reference genomes. The reference autodetect feature of DRAGEN can recognize the reference hash tables build on the four Human reference genomes: hg19 (hg19), GRCh37/hs37d5 (hs37d5), GRCh38/hs38d1(hg38), and T2T-CHM13v2.0 (chm13).

DRAGEN supports pangenome reference hash tables which extend the reference genomes with alternative variant paths from a sample cohort used to construct the pangenome reference. A pangenome-based reference improves the mapping accuracy of Illumina reads in the "Difficult-to-Map Regions" of the genome and the downstream variant calling.

The pangenome is the highly recommended for Germline human analyses. The accuracy achieved with pangenome references are highlighted in the plot below.



In the following tables we summarize the reference support for each DRAGEN component and the recommended reference type for each component.

Table 1 v4.3 Reference Support and Recommended Use for Human Data

Human		hg19, hs37d5, hg38	chm13	Recommended Reference Type
Germline	SNV	Yes	Yes	Pangenome
	CNV	Yes	Yes*	Pangenome
	SV	Yes	Yes*	Pangenome
	Expansion Hunter	Yes	No	Pangenome
	Targeted Callers	Yes	No	Pangenome
	RNA	Yes	Yes*	Linear
	De Novo	Yes	Yes*	Pangenome
	Joint Genotyping	Yes	Yes*	Pangenome
	Biomarkers (HLA)	Yes	Yes*	Pangenome
	Gvcf Genotyper	Yes	Yes*	Pangenome
Somatic	SNV	Yes	Yes*	Linear
	UMI SNV	Yes	Yes*	Linear
	CNV	Yes	Yes*	Linear

	SV	Yes	Yes*	Linear
Methylation	Methylation	Yes	No	Linear
Annotation	Nirvana	Yes	No	n/a

Table 2 v4.3 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	Nirvana	Yes	n/a

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

Multi-version Installer

DRAGEN™ v4.3 has updated the installer for on-premises systems which results in a user interface break from v4.2 or earlier.

- v4.3.6, v4.3.13 and future versions can be installed on the server at the same time allowing for easy switching.
- Multi-version installers for prior supported versions are coming soon.

NOTE: Root privileges are still required for the installation.

Single Version Installation

Up to DRAGEN™ v4.2, only one version of the software can be installed at a time. Executing the `.run` file will remove any existing installed version and (re)install the new version.

After installation, the application and associated files are available at `/opt/edico`.
The single version installer will add `/opt/edico` to the Linux $PATH, so that the user can just call `dragen` without specifying the full path.

New Multi-version Installation

Starting with DRAGEN™ v4.3 and later, multiple compatible versions of the software can be installed at a time. Executing the `.run` file will add the new version to the system.

After installation, the application files are available at `/opt/dragen/{version}` and FPGA files are located at `/opt/bitstream/{bitstream version}`.
The multi-version installer will NOT add `/opt/dragen/{version}` to the Linux $PATH, since multiple versions can be present at a given time. User should manage the desired paths to the specific version they want to run.

Notes on multi-version installation:

- Installers released for DRAGEN™ v4.2 and earlier are single version packages.
- Single version packages and multi-version packages cannot be mixed.
    - Installation of a prior single version package will remove all the multi-version packages.
    - Installation of a multi-version package will remove any installed single version package.
- After installing a multi-version package, see a list of installed versions at any time by running `/usr/bin/dragen_versions`
- To remove any multi-version package, call `yum remove` on its Path.
- A multi-version installer can be identified by the presence of `multi` in the file name, e.g. `dragen-4.3.6-11.multi.el8.x86_64.run`

Example:

```
$ dragen_versions
The output format of this command may change. Use --json for machine readable output.

Dragen Version          Size (MB)  Install Date         Path
4.3.2                   1378.03    2024-03-10 18:26:17  /opt/dragen/4.3.2
4.4.3                   1381.41    2024-03-18 20:56:39  /opt/dragen/4.4.3
4.3.5                   1379.25    2024-03-11 15:20:24  /opt/dragen/4.3.5
```

```
Bitstream Version        Size (MB)  Install Date         Path
07.031.732 (0x18101306)  598.95     2024-03-10 18:26:03  /opt/bitstream/07.031.732
07.031.745 (0x18101306)  598.95     2024-03-18 20:56:18  /opt/bitstream/07.031.745

To remove a dragen version, call `yum remove` on its Path.
```

Location of dragen and resource files

DRAGEN Version	on-premises server	cloud instance
v4.3 and later	/opt/dragen/{version}	/opt/edico/
v4.2 and earlier	/opt/edico/	/opt/edico/

Issues Resolved

Changes made, and/or issues that are fixed in the DRAGEN™ v4.3.13 release:

Component	Description
Amplicon	Remove soft clips for germline pipeline, to fix accuracy on Pillar Myeloid panel
Amplicon	Fix for amplicon aligner incorrectly aligning reads causing FPs in Germline Mode
Amplicon	Fix for a missing 6bp deletion on MPN panel with the SeraSeq Myeloid sample
BCL	BCL large memory leak for missing cbcl files during psCBCL generation for on-instrument use cases
BCL	psCBCL workflow with parallel lanes errors out on lanes unlisted in the sample sheet, for on-instrument use cases
BCL	Compatibility with RTA when outputting all cycles with failed reads
BCL	Support custom fields in the sample sheet Data section
BCL	BCL convert sw only licensing fix. Removed a dongle check when use-hw=false to improve robustness
BCL	Remove superfluous warning messages regarding missing files when doing demux-map generation with incomplete data (early demux for stats, on-instrument). Which fixes a hang
BCL	Undetermined stats in Adapter_Cycle_Metrics.csv were missing for lane 1
BCL	Support for parallel no-lane-splitting on instruments
CNV	Skip ASCN if insufficient coverage found from case sample, to avoid crash
CNV	Block ASCN CNV workflow after b-allele counts generation if PON is not provided
CNV	Restrict models to those that use a reasonable number of states to avoid timeouts
CNV	Use thread safe function on WES ASCN CNV to resolve hang
CNV	WES ASCN Tumor Normal segfault fix
CNV	Add HRD soft fail if CNV model is failed
CNV	CNV filter update when MatchSv is not PASS
Compression	Fix a crash in AltContigTracker when CRAM decompression and map/align uses a different reference
Driver	Fix driver compilation failure with newer el9 kernels
DUX4	Fix a typo in DUX4 VCF output. The VCF header instead of "rearrangement" says "arrangement" in the FILTER description

DUX4	DUX4 update warning messages with option to skip sanity check
Explify	Improve file access speed and fix a UPIP Z_DATA error
Explify	Converted a major intermediate file ("dxsm files") from text-based JSON format to binary format to speed up read/write.
Explify	Fix non-deterministic results in VSP2 pipeline
Explify	Refactor threading in first pipeline stage to increase cpu utilization
Explify	Fixed uninitialized values when read QC is disabled by the user. These values were throwing off rpkm counts, among others.
Explify	Improve memory usage in various stages of the pipeline to keep the peak RAM approximately 100 GB max. Stress tested with wastewater samples.
Explify	Fix run time print for Explify
File IO	Fix handling for `|` and `.` in contig names in BED files. Fix unmatched regex in bed interval reader. Fix qscore evaluation fail on phix
File IO	Support bam-list as input to variant callers (new feature). Support BAM merging from a BAM list. If the input is of the bam_list.csv format and the user enabled map-align and/or sort, the system shall merge the BAM files in the list and output a single BAM file when the BAM output is enabled. Support a BAM list input option for tumor and normal inputs. The BAM list file has the following requirements: The file is a .csv formatted text file, where columns are comma separated The header always contain the following column fields: BamFile Each row shall contain a different BAM file Each BAM file may represent a different read group of a sample, or different sample.
GvcfGenotyper	Update error handling for the case where some samples do not have AD. Limit the number of warnings about missing AD in both logfile and on stdout
GvcfGenotyper	Improved and fixed sample renaming feature in the GvcfGenotyper
GvcfGenotyper	Improve run time in iGG step 3 when allelic balance metric is selected. Only calculate allelic balance p-values when requested.
NovaseqX on-instrument	Improve BCLConvert speeds for NoLaneSplitting=true, and support NoLaneSplitting across sub-lanes.
NovaseqX on-instrument	Enable LZ4 spill compression for all samples to avoid out-of-disk issues.
NovaseqX on-instrument	Improve single-ended run detection robustness.

NovaseqX on-instrument	Remove dupmark algorithm option
NovaseqX on-instrument	Copy germline tagging file and aux CNV pop B allele file to the output for somatic enrichment. These new aux files were added in Wave 3.
NovaseqX on-instrument	Fix for missing BCL convert report in DRAGEN Reports, when the Generate FastQC metrics default changed to false.
Infrastructure	DRAGEN can now handle multisample fastq lists with mixed single/paired read
Infrastructure	Implemented LZ4 compression on intermediate DBAM spill (Binner). Optional "--binner-compress-spill"
Infrastructure	Added diagnostic dumps for cases when zlib fails
Inputs	Fix various attempts to open files with empty filenames.Skip trying to open files with empty filenames.
Inputs, Outputs	Fix for Specifying RGPU in sample sheet creates duplicate entries in BAM file
Inputs, Outputs	BAM tags from fastq_list.csv file are made uppercase if they are not uppercase
Installer	Fix dragen sosplugin import. Our import conditional breaks on recent versions of the sos library in Oracle8
MRD	Remove indels from consideration. The current MRD statistics does not utilize indels, however INDEL counts were a) included in the count file b) incorrectly counted.
MSI, Somatic	Improve MSI memory usage. Fixes DRAGEN ICA somatic enrichment tumor-only with MSI watchdog low memory error
Paralog	Fix CYP21A2 POS for hg19/hs37d5. LPA SV VCF entry has wrong REF
Personalization	Fixes for personalization double free crashes
Ploidy	Fix crash with "ploidy estimator is not enabled", on somatic workflows
QC Metrics	Fix "ignore overlaps" logic for coverage metrics
RNA	Generating empty quant.sf file when no RNA quant stats were collected
RNA	Fix for missing intron motifs in SJ.out.tab
RNA Splice	Add RNA splice variant caller knowns file in the dragen package, that is autodetected and used. Fixes AR-v7 splice variant not called in v4.3.6
SNV Somatic	Fix a FN for a clinically actionable variant in EGFR, where the coverage is high due to EGFR amplification, resulting in a situation with 165 supporting reads out of 27k (VAF ~0.6, which is below the pruning threshold for the graph). The fix adds a second (absolute) pruning threshold in addition to the existing relative threshold, such that edges are pruned only when their weight is below both thresholds.

SNV VC germline	Fix for SNV VC germline calls tagged as MOSAIC when mosaic detection is enabled. This fix will allow active region boundaries to be the same regardless of whether mosaic detection is enabled or not.
SNV VC, Somatic	Report haploid genotype when ploidy=1. Fixes small WGS (non-human) VCF validation error. We now report haploid genotypes ('0' for normal sample and '1' for tumor sample) when --ploidy=1 is set in command line. We also output haploid GT for variant calls in non-PAR regions of chrX and chrY for male samples regardless of ploidy set in the command line.
Star Allele	Updated select.gvcf to be written only when internal option "--enable-select-gvcf" is set. This file is useful for testing/debugging, but should not be necessary as a default output.
SV	Mute unused SV caller logging messages, to clean up clutter in run logs and reduce file size
SV	Handle multiple assembly haplotypes for SV, to rescue the missing FLT3-ITD insertion in the SeraSeq sample and improves recall.
SV	Fix for Watchdog timeout when processing SV locus graph edges
SV	Add ability to disable tabix indexing in SV. Helps with references that have chromosomes longer than 512 Mbps, such as plants.
SV	Fix a bug in the assembly graph that only impacts SV in a specific edge case. May result in more haplotypes/variants when that edge case is hit (expected to be rare). The fix leads to proper handling of that edge case.
Targeted Caller	Fix to auto disable targeted caller for chm13
Targeted Caller	Fix crash "targeted caller likelihood must be a positive number", encountered with non-WGS or low-quality (i.e. noisy coverage) data.
Targeted Caller, HBA	Fix for HBA wrong REF in VCF. Fixes an allele mismatch bug for CNV variants in HBA targeted caller.
UMI	Fix for FASTQC+UMI reporting infeasibly long read lengths.
VC	Fix crash during read realigning in small VC
VC	Add support for reporting haploid VC metrics as ""Insertions (Hap)", to fix issue where thare are incorrect number of variants in vc_metric.csv

Known Issues

Known issues of the DRAGEN™ v4.3.13 release

Component/s	Summary	Resolution/Workaround
Amplicon	HD829 variant not detected but visible in IGV (Pillar Myeloid)	None.
Amplicon	Deletion called in amplicon where the deleted base is in the probe region	None.
Amplicon	Amplicon CNV exon level FN has been seen	None.
Amplicon	Amplicon read target coverage report can be incorrect for tiling amplicons in some cases. Target assignment and accuracy not impacted.	None.
Amplicon	FN at the edge of amplicon targets due to "vc-remove-all-soft-clips=false" in v4.3. The occurrence is rare and only affect the variants at the edge of targets.	None.
Amplicon	Missing FLT3 insertion variant in SV calling. Occurrence is rare	None.
BCL	Some combinations of indexing schemas have a significant run time regression compared to DRAGEN v3.7	A fix is not available, but a workaround for on-premises can be done of this specific indexing schema.
BCL	BCL conversion appends FASTQ files when using --force. FASTQ output may get concatenated if user uses the same output directory twice for BCL.	Do not run BCL conversion multiple times on the same output folder
BCL	BCL streaming over network sometimes hangs or takes hours, compared to minutes on local disk, when many files are opened and streamed.	Recent investigations has shown that network file systems may hit performance issue when streaming from large number of file handles, likely inducing severe caching issues on the network filer. A BCL setting exists that can reduce the number of simultaneous over-the-network accesses which has shown to alleviate the problem, with run time impact. The best remedy is to avoid letting the network file system get into a bad caching state, by using run modes with fewer open file handles, or run from local disk.
BCL	If a directory is specified as input to "--sample-sheet", BCL Convert will hang at the beginning of a run while trying to copy that path as a file to <outdir>/Reports/SampleSheet.csv	Specify the sample sheet file.

BCL	BCL Convert does not validate when "Logs" or "Reports" is provided for a Sample_Project, and the software will be unable to create the subdirectories if these strings are provided	None.
BCL	BCL Convert does not support the "--first-tile-only" option being specified for SP flow cells, but the new "--tiles" option can be used as a substitute	Use the new "—tiles" option
BCL	Does not error when no tile list exists in the RunInfo.xml file and "--tile" or "--exclude-tiles" is specified in the command line	Ensure that tiles are not specified when none exist.
BCL	When an index collision exists in a lane that has been excluded via the "ExcludeTiles" setting, the software will still error as it is a sample sheet validation error	None.
BCL	BCL Convert has differences in legacy stats compared to bcl2fastq2 (all of which are justified)	None. This is per design, please consult the BCL documentation for expected outputs
CNV VC	CNV VAF loci related to wrong purity/ploidy estimate when normal sample has problematic regions.	Rare occurrence. A possible workaround is to disable VAF-aware mode in CNV when needed.
Compression	CRAM decompress and map/align using different references, can falsely run into an alt contig error check, when two conditions are true: hash table used for cram decompression, and both hash tables combined contain alt contigs that exceed a threshold. The software counts contigs from both refs instead of only the map/align ref.	Supply the FASTA file to decompress the CRAM
DNA Alignment, Compression	DRAGEN "BAM to CRAM, back to BAM" is not lossless. There is a CIGAR mismatch for full soft clipped reads. The CRAM format does not store CIGAR and other fields for unmapped reads. This is the implementation per standard, and same in all CRAM tools. In v4.3 with graph ref and ga tags, fully clipped reads will be unmapped. But due to ga tags, the CIGAR is stored in the BAM record. When compressed with CRAM, and back to BAM, those CIGAR fields are lost, and the decompressed BAM mismatch with original BAM.	None.
Downsampling	Exome downsampling is not giving right coverage	For v4.3 a new fractional downsampler is available to handle this use case better.
Downsampling	Target coverage downsampler doesn't hit the correct coverages.	For v4.3 a new fractional downsampler is available to handle this use case better.
Downsampling	Crash during coverage downsampling.	For v4.3 a new fractional downsampler is available to handle this use case better.

Explify	Some cases report coverage greater than 1.0 for SARS-CoV-2	None.
Explify	The order of reads in input fastq can lead to small differences in output.	None. This does not meaningfully impact the results.
Gene Fusion	RNA pipeline has seen unexpectedly long run time and high cost, when very high depth of Globin levels are present in the sample. This leads to high processing on false events.	Disable Gene Fusion. Take extra care during sample prep.
GVCF Genotyper	Issues running non-human gvcf aggregation with IGG	None.
GVCF Genotyper	Sample rename feature for iGG does not work.	None.
GVCF Genotyper	iGG writes phased genotypes but does not write the phase group (FORMAT/PS). Vcf Validator may fail	None.
Hash Table Builder	Hang when building custom graph Hash Table with input msVCF containing mega-base long indels. In such case Hash Table builder runtime can be extremely long or hang completely.	Pre-filter the input msVCF to remove million-bases long indels.
HLA	There is a minor accuracy regression on HLA due to a miscall in the DRB1 gene after changing the reference from v3 graph to v4 graph, on hg38 reference.	None.
HW GRAPH	A rare segfault has been observed on cloud runs due to HW GRAPH error.	The rate of occurrence is < 1/1000 and not repeatable. Re-run the sample
Infrastructure	If an AWS node is configured to "IMDSv2 Required", S3 input file streaming does not work.	Typical configuration is "IMDSv2 Optional", in which case S3 input streaming works.
Personalization	Personalization feature crash has been observed on some Constellation data	None.
Q-score Evaluator	A small run to run variation in qscore per position stats has been observed on some samples.	None.
QC Metrics	Overlapping mates' calculation does not correctly handle the situation where a supplementary alignment overlaps with a primary alignment. The new mapper HW changes exposed this bug when the alignment of some reads changed. This should be an extremely rare bug limited to situations where a supplementary alignment overlaps the primary read, and where the primary read start and ends within the range of the supplementary.	None. Rare and impact on coverage metrics is very limited
RNA Quantification	Minor differences in RNA quant SJ.saturation.txt with different num-threads value.	Use a consistent number of threads across runs if desired

SNV Germline	There is a risk of overwriting files for the denovo filtering component, when "dn-output-vcf" is not used but "output-dir" is used and set to the path of the input file "dn-input-vcf". It is not doing a check.	Do not use "output-dir" for the denovo filtering feature. Follow the use guide instructions for specifying input and output paths
SNV Germline	Missing some expected mosaic variants in certain samples.	None. The mosaic caller is new. Continuous improvement of this new feature will happen in future releases.
SNV Somatic	DRAGEN MAF watchdog false error on an large sample while reading in JSON	Disable watchdog.
SNV Somatic	Accuracy (FP) regression on some FFPE/FF Normal/Normal datasets in somatic due to columnwise detection, as result of large difference in noise profiles between the FFPE tumor and FF normal sample.	There is no major FP regression in any other TN datasets, and accuracy improved across vast majority of samples. For these extreme cases, columnwise detection could be disabled if needed
SNV Somatic	A MNV FP has been identified on the TSO500 assay.	None.
SNV Somatic	Somatic SNV T/O MNV failing to merge two MNV calls in the edge case where we have a deletion upstream of another co-phased variant with and an out-of-phase SNP in between them that is covered by the REF allele of the upstream deletion. In this scenario, we will end up excluding haplotypes based on the haplotype max klen values being less than the distance between 2 variants in the group that are not actually in phase with one another.	None.
SNV Somatic	Somatic T/O WES Indel Sensitivity regression seen in some samples, due to a single new FN introduced by germline tagging with new Nirvana resources including gnomAD v4.0.	None.
SNV Somatic	Minor SNP FP regressions observed on some datasets due to hotspots VCF file updates. The net effect of the updated hotspots is positive over most test samples.	The FP in question are variants that were newly added to the hotspot file or variants that were originally present but represented in a non-normalized fashion.
SNV Somatic	Bam output runtime increases by 7% for Tumor Only mode	None.
SNV Somatic	3 new SNP FP introduced in one test sample due to updates to the somatic hotspots file to include a few thousand additional heme-specific hotspots.	Future updates to the hotspots file may resolve the FPs. Accuracy changes are expected in 4.3. While most of the accuracy metrics improve significantly across the majority of datasets, there are isolated cases of minor regressions on some.

SNV Somatic	Multiple FGT tags attached to 1 forcegt call, in the scenario where a variant is present multiple times in the input vc-forcegt-vcf file, once as a single variant and again as part of multiallelic records. In this scenario, DRAGEN will output a forced call with the INFO.FGT tag applied twice.	None.
SNV Somatic	FP regression observed on a NormalNormal HCC1395BL PCRFree on 4.3 compared to 4.2	None. Accuracy changes are expected in 4.3. While most of the accuracy metrics improve significantly across the majority of datasets, there are isolated cases of minor regressions on some.
SNV Somatic	Most somatic SNV datasets have >10% increase in total runtime relative to previous release v4.2.	New features, including updated mapper, columnwise detection and allele-specific systematic noise and hotspots adds extra compute complexity and are expected to increase runtime.
SNV Somatic	A fragment length filter affects HOM ALT calls in rare samples. The filter triggers when the median fragment size for alt-supporting reads is substantially higher than that for ref-supporting reads. If there are no ref-supporting reads, the median fragment size gets set to 0, which then triggers the filter.	None.
SV	Out of memory watchdog fault during SV graph edge processing on rare samples, where there are extremely high number of SV locus graph edges.	None.
SV	Very high coverage samples can take long time to complete on ICA.	None. v4.3.6 Germline BSSH/ICA apps now have more processing power and will improve the processing times.
SV	AWS S3 uplink streaming is not functional for many use cases of DRAGEN.	Recommended to use downling S3 streaming only
TMB	Tumor only TMB is not as reliable as T/N TMB.	None.
	DRAGEN Germline 4.3.6 App Timeout has been observed due to overflow or corrupted Y median coverage value.	None.

SW Installation Procedure

- Download the desired installer from the Illumina support website and unzip the package.
- The archive integrity can be checked using: `./<DRAGEN 4.3.13 .run file> --check`
- Install the appropriate release based on your Linux OS with the command: `sudo sh <DRAGEN 4.3.13 .run file>`

Release History

Revision	Release Reference	Originator	Description of Change
00	1115672	Cobus De Beer	Initial release