

DRAGEN v4.1.5 Software Release Notes

Introduction

These release notes detail the key changes to software components for the Illumina® DRAGEN™ Bio-IT Platform v4.1.5.

Changes are relative to DRAGEN™ v4.0.3. If you are upgrading from a version prior to DRAGEN™ v4.0.3, please review the release notes for a list of features and bug fixes introduced in subsequent versions.

DRAGEN™ Installers, User Guide and Release Notes are available here:

https://support.illumina.com/sequencing/sequencing_software/dragen-bio-it-platform.html

The v4.1.5 software package includes installers for the on-site server:

- DRAGEN™ SW Intel Centos 7 - dragen-4.1.5-9.el7.x86_64.run
- DRAGEN™ SW Intel Oracle 8 - dragen-4.1.5-9.el8.x86_64.run

The following configurations are also available on request:

- Amazon Machine Image (AMI)
- Microsoft Azure Image (VM)
- RPM packages for Centos 7 for Amazon Web Services (AWS)

Deprecated platforms:

- Support for DRAGEN Server v1 FPGA cards have been deprecated since DRAGEN™ v3.10

Contents

Overview	3
What's different	3
Interface Changes	3
Issues Resolved	5
Known Issues	9
SW Installation Procedure	13

Overview

DRAGEN™ v4.1 is the version that is released with Illumina's NovaSeq-X series sequencer with on-board and cloud pipelines. This goal of this software release is to make the same DRAGEN™ version available to on-server and cloud customers.

The bioinformatics features available in DRAGEN™ v4.1 is the same as DRAGEN™ v4.0.3. The Illumina DRAGEN™ Bio-IT Platform User Guide for v4.0 is applicable to v4.1 also, and available on the support website at <https://support.illumina.com/downloads/illumina-dragen-bio-it-platform-user-guide.html>

What's different

DRAGEN™ v4.1.5 retains the same bioinformatics features of v4.0.3. The software includes the following changes:

Change compared to DRAGEN™ v4.0	Impact on DRAGEN™ server and cloud platforms
Support for the FPGA hardware of the NovaSeq-X series sequencer	Not applicable
Support for a new FPGA and software architecture that allows multi-process execution on the NovaSeq-X series sequencer	Not available
Support for new modes of BCL de-multiplexing, and new multi-sample DRAGEN™ pipelines used on NovaSeq-X	Not applicable
Minor changes to select components to accommodate NovaSeq-X features	Minor changes to some component software
Minor accuracy adjustments, calibrated for NovaSeq-X sequenced data	Minor changes in accuracy of the SNV VC. Best performance with NovaSeq-X data
Run time optimizations targeted to NovaSeq-X on-instrument pipelines	Minor run time changes
Bug fixes as discovered on v4.0.3 and during v4.1 development	Improved overall stability and accuracy

Interface Changes

DRAGEN v4.1 introduces minor interface changes in some components:

- Downsampler
 - Some command line options have been updated for clarity

- `--enable-down-sampler-output` has been changed to `--enable-down-sampler-fastq`, to reflect that the setting enables only down sampled FASTQ output
- `--down-sampler-reads` has been changed to `--down-sampler-fragments`, to set the target number of sequence fragments instead of reads.
- Variant Caller
 - Added command line options, only used by NovaSeq-X pipelines
 - `--vc-ntd-error-params`, `--vc-estimate-ntd-error`, `--vc-mapping-metrics` allows pre-computed values to be passed into the variant caller
- The SW installer file name is changed
 - Note that the installer and `.rpm` now has "arch2" added in the name. No user actions needed
 - From `dragen-4.0.3-8.e17.x86_64.run` to `dragen-arch2-4.1.5-9.e17.x86_64.run`

Issues Resolved

Issues found on DRAGEN™ v4.0 or older that are fixed in v4.1.5

Component	Defect ID	Issue Description
Azure cloud	DRAGEN-21794	Fixed a race condition in XRT FPGA driver causing incorrect 64-bit reads, leading to incorrect FASTQC metrics on Azure cloud platforms
BCL	DRAGEN-20513	Improve error messages for per-sample settings errors, specify Sample_ID
BCL	DRAGEN-20064	Fix for BCL convert aborting with an empty [BCLConvert_Settings] section in v2 sample sheets
BCL	DRAGEN-16992	Fixed an incorrect error message when a value is missing for combinatorial index samples in BarcodeMismatchesIndex1
BCL	DRAGEN-19376	A crash could occur when using per-sample-settings with higher sample counts in a lane due to a hash-table pre-size using a signed integer as input that is overflowed. The fix is to use a larger integer size in the pre-size calculation.
BCL	DRAGEN-19292	Fix BCL FASTQ file paths in fastq_list.csv when ORA-interleaved compression format is used. Previously the fastq_list.csv file contained two files (instead of one single interleaved file) under the "Read1File" and "Read2File" columns and the files were not named correctly.
BCL	DRAGEN-21443	Some versions of RTA3 outputs cbcl with 0 qual + non zero base for 0/"N#", in other words masked nibbles (0 qual) does not zero out the base. bcl2fastq2 has a masking step, but DRAGEN and bcl-convert does not. This fix adds masking so that bcl-convert matches bcl2fastq2 for those RTA3 outputs.
BCL	DRAGEN-21718	Fix for BCL crash with threading error on very large sample datasets (150k)
CNV VC	DRAGEN-18956	In CNV VAF modeling, fix estimation of scale factor in presence of outliers
CNV VC	DRAGEN-17796	Fixed a bug in target interval split overlap method, if there is 1 bp overlap in input bed file, leading to crash
DNA Alignment	DRAGEN-16308	Fix for SIGABRT in read trimmer adapter trimming during RecomputeTags::computeTags, which happens when back-to-back read trimmer jobs are run, due to FPGA registers being left in undefined state.

Down sampling	DRAGEN-20563	Fixed a watchdog hang in the down-sampler, caused by unhandled single ended BAM input case
Driver	DRAGEN-22634	Fix for EL8 DKMS 3.0 breaking networking on boot, when upgrading to RHEL 8.6
Hash Table Builder	DRAGEN-20212	Fixed a crash in multi-genome hash table builder, when no isolated SNPs are kept
Infra	DRAGEN-20550	Fix for excessive watchdog logs filling up /var/ partition
Imputation	DRAGEN-19177	Improve error message when the input VCF is not bgzipped
Imputation	DRAGEN-19174	Improve error message when input VCF variants do not overlap with reference panel
Inputs	DRAGEN-19361	Fixed a potential crash in read_group_list when using fastq_list.csv, caused by a timing race condition
Metrics, HLA	DRAGEN-21921	Fix for mapper metrics being double counted when HLA is enabled
Paralog Caller	DRAGEN-19104	Fixed a hang in multi-caller use case, with targeted callers enabled, run when starting from BAM input
QC	DRAGEN-21857	An invalid check for 10 required columns for the --qc-cross-cont-vcf header resulted in an exception. Fixed the check to require 8 columns. Also improved error handling for invalid qc-cross-cont-vcf input, with clearer messages.
RNA Gene Fusion	DRAGEN-21241	Fix for Gene Fusion VCF qual score of "inf"
RNA Gene Fusion	DRAGEN-20991	Add mitochondrial genes filter to Gene Fusion VCF header
RNA Gene Fusion	DRAGEN-20998	Make Gene fusion VCF report PR:SR reads, like DRAGEN SV
SNV Somatic	DRAGEN-18195	Improve detection of MNVs for overlapping variants
SNV Somatic	DRAGEN-16319	Address a regression in SNP and INDEL FP on ICGC datasets that was introduced in DRAGEN v3.10

SNV Somatic	DRAGEN-19218	Improve elevated INDEL FP+FN seen in v4.0.3 compared to v3.10. SNP FP is rescued by up to 3% and INDEL FP is reduced by up to 7%.
SNV Somatic	DRAGEN-19052	Improve run times of Somatic SNV with NTD error estimation enabled
SNV Somatic	DRAGEN-22828	Improve run times in the mapper phase of Somatic SNV previously impacted by snperror estimation
SNV Somatic	DRAGEN-18660	Fix an issue where joint calling output for mito positions are all PASS
SNV VC	DRAGEN-16398	Fixed a LowGQ call accuracy issue. Any base that was adjacent to an insertion or deletion or soft clipped base was considered as a "mismatch" regardless of what the actual base is.
SNV VC	DRAGEN-20271	Robustness improvement. Add additional checks for reads with cigar length of 0
SNV VC, ML	DRAGEN-18342	Fixed a discordance bug in output between FASTQ vs BAM input to variant caller with ML. It resulted from mapq=0 reads that have 'N' bases, which had a difference when comparing to the ALT allele in BAM v FASTQ.
SNV VC, ML	DRAGEN-18807	Add ML files for the hs37d5 reference. ML accuracy improvements for hs37d5/hg19
SNV VC, ML	DRAGEN-19829	Fix various run-to-run variations in SNV VCF/gVCF, due to ML.
	DRAGEN-20268	
	DRAGEN-19829	
SNV VC, ML	DRAGEN-20235	Fix a potential crash in SNV VC with ML during the handling of disqualified reads
SV	DRAGEN-20389 SET-3913	Fix crash in SV caller, caused by a tiny candidate contig size that is generated in SV assembly stage
SV	DRAGEN-19631	Fix for some FFPE samples where the sv.vcf.gz file contains only 1 entry for a MantaBND. All MantaBND events should have 2 entries.
SV	DRAGEN-20267	After the introduction of T/O scoring model and filters, it was possible that regions marked as "hot spots" are still being filtered out. Fixed the bug

UMI	DRAGEN-19092	New option "umi-parse-only" added to enable the UMI parser for regular Map/Align without UMI (enable-umi=false). If user specifies the "umi-source", it is saved to the output bam with RX tag.
-----	--------------	---

Known Issues

Known issues of the DRAGEN™ v4.1.5 release

Component	Issue ID	Summary	Resolution / Workaround
BCL	DRAGEN-23388	BCL will crash when "--no-sample-sheet true" & 0 indexes supplied	No workaround. Fix planned for next version
BCL	DRAGEN-23363	No-sample-sheet setting omits index sequences from fastq headers	No workaround. Fix planned for next version
BCL	DRAGEN-22935	BCL Sample_Name + Sample_Project behavior differs from bcl2fastq2	No workaround. Improvement planned for next version
BCL	DRAGEN-22820	BCL outputs extra metrics when running with Per-Sample Settings, as compared to Global Settings. Demultiplex_Tile_Stats.csv and Quality_Tile_Metrics.csv contain metrics for tiles with 0 reads processed when Per-Sample Settings are used, but not when Global Settings are used.	No workaround. Fix planned for next version
BCL	DRAGEN-22480	Customers with high CPU core count systems have reduced BCL performance due to a thread limit, since v3.10	No workaround. Fix planned for next version
BCL	DRAGEN-20663	BCL does not abort when Combined Index Collision Check is enabled on a dual index run with one index removed	Uncaught user input error. Operation proceeds normally.
BCL	DRAGEN-19157	Filenames for interleaved FASTQs that are Ora compressed, are not the same as the original file names. For original filenames ending in "R1_001.fastq", "R2_001.fastq" the decompressed file names are "R_1.fastq", "R_2.fastq", dropping the identifier "001". This could potentially lead to duplicate file name conflicts	No workaround. Fix planned for next version
BCL	DRAGEN-19103	BCL crashes in Robust mode when *.filter file is missing for single lane dataset	No workaround. Fix planned for next version
BCL	DRAGEN-18920	bcl-convert outputs different PF cluster YieldQ30 and QualityScoreSum stats in the legacy stats file ConversionStats.xml as compared to bcl2fastq2.	No workaround. Fix planned for next version
BCL	DRAGEN-13771	A crash during bcl error checks can lead to hang, due to timing race condition	No workaround. Fix planned for next version

DNA Alignment	DRAGEN-23631	A specific mapper mode can produce reads with invalid CIGAR match length of 1023 (-1), leading to a crash in the Variant Caller	No workaround. Fix planned for next version
Gvcf Genotyper	DRAGEN-21775 DRAGEN-23457	Remaining alleles after AC=0 removal, are not trimmed in some cases	No workaround. Fix planned for next version
Gvcf Genotyper	DRAGEN-22446	Some unnormalized variants in msVCF output of Lettuce samples	No workaround. Fix planned for next version
Gvcf Genotyper	DRAGEN-21922	Some incorrect LPL and LAA values in msVCF	No workaround. Fix planned for next version
Gvcf Genotyper	DRAGEN-21458	RAM usage in some use cases is higher than expected, due to a memory leak.	No workaround. Fix planned for next version
HLA	DRAGEN-19524	Incorrect genotyping output format when minor allele has insufficient read evidence support.	Minor accuracy issue.
Infra, SNV VC	DRAGEN-21518	Regression in run times on Azure cloud nodes for Germline SNV	No workaround. Fix planned for next version
HW GRAPH	DRAGEN-18402	A very rare error in hardware graph has been seen, leading to assertion.	Re-run the sample
Imputation	DRAGEN-22549	Imputation end to end pipeline adds only the first chromosome name to VCF the header, leading to problems with downstream tools	Re-header the VCF using bcftools
Infra	DRAGEN-19988	A crash on Azure can leave the system in a bad state that requires intervention and prevents subsequent jobs from succeeding. "ERROR: xclRegRW: can't map CU: 0"	Known issue for which a solution is not available
Joint Genotyping	DRAGEN-21909	Accuracy on denovo WGS joint genotyping changed, due to an ML qual adjustment made to improve NovaSeq-X indel performance	Planned FP/FN accuracy tradeoff for improved performance on NovaSeq-X data
Joint Genotyping	DRAGEN-19844	Joint genotyping is up to 30% slower compared to v4.0	No workaround. Fix planned for next version
Ora Compress	DRAGEN-19279	File names are not preserved exactly as they were, for the interleaved decompression mode.	No workaround. Fix planned for next version

RNA Gene Fusion	DRAGEN-15168	Missed fusion in 1st exon of gene TLC1--TRBC2	No workaround. Fix planned for next version
scATAC	DRAGEN-23486	scATAC with combinatorial barcode position results in empty results	No workaround. Fix planned for next version
scRNA	DRAGEN-22854 SET-4565	scRNA/scATAC output has a missing column, which may affect compatible with downstream tools	No workaround. Fix planned for next version
SNV VC	DRAGEN-23630	An invalid alignment used to build the graph genome, leads to an incorrect allele frequency. Only one such instance has been found.	No workaround. Fix planned for next version
SNV VC	DRAGEN-22841	In rare cases, MNVs are wrong when the merging distance is greater than graph TLEN	No workaround. Fix planned for next version
SNV VC	DRAGEN-19234 DRAGEN-22280 SET-4548	Some variants are not emitted when evidence BAM is enabled	No workaround. Fix planned for next version
SNV VC	DRAGEN-17705	When output VCFs are not compressed, the md5sums are not available.	No workaround.
SNV VC, ML	DRAGEN-23609	Joint Genotyping accuracy regression on NA12878 hg38 trio when using graph genome and ML, due to PL and GP for 0/0 calls	No workaround. Fix planned for next version
SNV VC, ML, Joint Calling	DRAGEN-23390	Some incorrect GP/PL values in SNV VCF/gVCF when ML is enabled, for HetHom cases and multi-allelic calls	No workaround. Fix planned for next version
SNV Somatic, MNV	DRAGEN-21798	Rare samples may have a MNV length that overflows a variable, leading to a corrupted TAG and a downstream component (germline filter) that asserts.	No workaround. Fix planned for next version
SNV Somatic	DRAGEN-23727	Potential out of memory on NovaSeq-X instrument for somatic enrichment workflow, when the samples are very large	No workaround. Fix planned for next version
SNV Somatic	DRAGEN-23644 DRAGEN-23494	Somatic calling with evidence BAM output may encounter Out-Of-Memory for high depth samples	Lower the binner memory using "--bin_memory" option

SV	DRAGEN-21912	Rare samples may produce very long assembly length in tumor only SV caller, causing a 32bit integer overflow, leading to a segmentation fault	No workaround. Fix planned for next version
SV	DRAGEN-18913	Any regions overlapping the hotspot BED files for DRAGEN-SV will be called, even with minimal support. This introduces of 1 FP across our FLT3-ITD suites	No workaround.
TMB, SNV	DRAGEN-18826	Tumor only analysis with GVCF mode + TMB + germline filtering aborts in BSSH with out of memory issue.	No workaround. Fix planned for next version

SW Installation Procedure

- Download the desired installer from the Illumina support website and unzip the package
- The archive integrity can be checked using: `./<DRAGEN 4.1.5 .run file> --check`
- Install the appropriate release based on your Linux OS with the command: `sudo sh <DRAGEN 4.1.5 .run file>`
- Please follow the installer instructions. Server power cycle may be required after installation, depending on the currently installed version. If an updated FPGA shell image needs to load from flash, this is only achieved with power cycle.
 - A power cycle is required when upgrading from v3.3.7 or older
 - A power cycle is required when downgrading to v3.3.7 or older
 - A power cycle is not required when upgrading from a release after v3.3.7
- Procedure to downgrade to v3.3.7 or older:
 - Requires the following three steps. The prior .mcs file needs to be flashed manually:
 - Install the prior release: `sudo sh <DRAGEN 3.3.7 .run file>`
 - `program_flash /opt/edico/bitstream/07*/*.mcs`
 - Power cycle

Release History

Revision	Release Reference	Originator	Description of Change
00	1082864	Cobus De Beer	Initial release