DRAGEN TSO500 ctDNA Analysis Software

Customer Release Notes

V2.5.0

For TruSight Oncology 500 ctDNA Assay

March 6, 2024

Introduction

These Release Notes detail the key changes to software components for the DRAGEN TSO500 ctDNA v2.5.0 Analysis Software on DRAGEN server. For full details, please consult the DRAGEN TSO 500 ctDNA v2.5.0 Analysis Software User Guide available on the support website.

This software is intended for use with the TruSight Oncology 500 ctDNA and TruSight Oncology 500 ctDNA v2 assays.

- Software Version: 2.5.0
- DRAGEN software version 3.10.15

The software package includes:

- dragen_tso500_ctdna_2.5.0.tar – a tar file of the DRAGEN_TSO500_ctDNA_v2.5 docker image
- uninstall_DRAGEN_TruSight_Oncology_500_ctDNA-2.5.0.sh - bash script used to uninstall DRAGEN TSO500 ctDNA
- resources/ - workflow resource bundle for DRAGEN TSO500 ctDNA v2.5.0
- DRAGEN installer:
  - dragen-3.10.15-8.el8.x86_64.run - DRAGEN installer for CentOS 7
  - dragen-3.10.15-8.el7.x86_64.run - DRAGEN installer for Oracle Linux 8
- test_DRAGEN_TruSight_Oncology_500_ctDNA-2.5.0.sh - bash script used to validate DRAGEN TSO500 ctDNA installation is successful
- install_DRAGEN_TruSight_Oncology_500_ctDNA-2.5.0.run - script used to install TSO500 ctDNA
- DRAGEN_TruSight_Oncology_500_ctDNA.sh - bash script used to launch DRAGEN TSO 500 ctDNA

NEW FEATURES:

- All SNVs, insertions and deletions that are part of MNVs are now reported both individually and as merged variants in the final VCF and the Combined Output Variant file.
- DRAGEN sex prediction algorithm was added to the TSO500 pipeline. The sex is predicted based on the read count information in the sex chromosomes and the autosomal chromosomes.
- MSI JSON file containing MSI results was added to the Results folder.
- SV evidence BAMs were added to Logs_Intermediates folder.

FIXED ISSUES:

- Illumina Annotation Engine 3.2.6 (aka Nirvana) includes the following bug fix:
  - A fixed RefSeq version (105.20220307) was incorporated that fixed canonical transcript

assignments for some prominent genes and variants. For example, for variant BRAF NP_004324.2. V600E, the canonical transcript and HGVS notations are now fixed in the CombinedVariantOutput file (was presented as BRAF NP_001361187.1. V640E in the previous version).

- Fixed an error of CombinedVariantOutput file displaying more than one annotation for some variants.
- Fixed errors related to the sample sheet validator in BaseSpace Run Planning tool that can generate sample sheet covering analyses run on DRAGEN server.
- Fixed an error in the MetricsOutput file where NTC samples (No-Template Control samples with 0 reads) were indicated as "Pass" instead of "Fail".
- Fixed errors with MNV calling when phased variant merging distance is greater than 10bp. Ensure haplotype from sufficiently long k-mers are used for merging phased variants.
- Fixed an instance where DRAGEN small variant calling did not correctly handle overlapping mates, mistakenly detected strand bias and subsequently incorrectly filtered the variant. Strand bias was measured because the strongest support for the variant consistently came from one strand. This should however not be measured as strand bias since both mates were in agreement.
- Fixed an issue causing runs with samples with extreme copy number gains (e.g. fold change > 50, corresponding to ~250 copies when tumor fraction is ~40%) in a particular region or contrived samples take significantly longer than 20 hours.
- Updated SV caller for more efficient SV processing.

KNOWN ISSUES:

- Moving or modifying files during the analysis may cause the analysis to fail or provide incorrect results.
- Using control-c during a running analysis may cause an FPGA error. To recover from an FPGA error, shut down and restart the server.
- The sample sheet should not have blank rows between samples in the [Data] section, this may cause a run failure.
- Performance not verified using reads other than 2 x 151, paired end, dual index.
- The software does not notify the user when InterOp files for RunQC are missing or corrupted.
- Some contrived samples such as SeraCare Complete Mutation Mix, which have multiple structural variants (SVs) and high library conversion efficiencies, could generate a high number of chimeric reads and high number of candidate SVs. Occasionally, the SV caller may filter some of the reads and lead to occasionally missing fusions. In such cases downsampling the FASTQs can help recover those fusion calls. Contact your local support team for additional details and a workaround.
- Analysis fails when starting from V1 sample sheets due to missing adapter sequences in V1 sample sheet template. Users are recommended to start with V2 sample sheet template or add adapter sequences manually.

- Pipeline does not exit early and continues to the next DragenCaller step due to TSO500 ctDNA FASTQ validation failure if Fastq_list.csv is missing.
- FastqGeneration step fails when using S4 flow cells and the run folder size is up to 1.2 TB, and FASTQ files up to 4.2 TB. This is due to FASTQ files being duplicated in both the Nextflow works folder as well as the Logs_Intermediates/FastqGeneration folder causing the disk space to run out before the FastqGeneration step could be completed.
- ctDNA pipeline fails for an NTC sample (No-Template Control samples with 0 reads) due to absence of Evidence BAM File.
- In the V2 CNV cutoff bed file, gene "MYCL" should be listed instead of "MCYL1".

PRODUCT LIMITATIONS:

- The sample sheet must be configured as described in the provided templates, User Guide or by using BaseSpace Run Planning tool.
- Sample sheets generated for auto-launch on ICA are not compatible and cannot be reused without changes for DRAGEN TSO500 ctDNA Analysis Software on a Local DRAGEN server, and vice versa.
- The values in the Run Metrics section will be listed as 'NA' if the analysis was started from FASTQs or if the analysis was started from BCLs but the InterOp files are missing or corrupted.
- Germline estimation uses the latest publicly available population data and is estimated to be representative of targeted population, the impact of rare germline mutations is expected to be limited.
- The Illumina Annotation Engine (aka Nirvana) may report incorrect HGVS c. and HGVS p. notation for small variants occurring in RefSeq transcripts that exhibit transcript sequences differing from the genomic reference (i.e., RNA-edits). Currently the HGVS c. error rate is 0.00527% and the HGVS p. error rate is 0.00737%.
- The CNV caller has slightly higher noise for sample types that are not included in the baseline used for normalization (eg., cell lines). The baseline samples consist of mostly healthy donor clinical samples and SeraCare-contrived samples.
- MSAF output has had limited testing and needs to be used with caution. Updates to the small variant calling have led to an increased MSAF in samples with higher DNA input.

Release History

Revision	Release Reference	Originator	Description of Change
00	CN 1104006	Svetlana Bureeva	Initial Release