DRAGEN TSO500 ctDNA Analysis Software on ICA

Release Notes

V2.1.1

For TruSight Oncology 500 ctDNA Assay

October 26, 2023

Introduction

These Release Notes detail the key features and known limitations of software components for the DRAGEN TSO500 ctDNA v2.1.1 Analysis Software on ICA. Below is a summary of the changes included in DRAGEN TSO 500 ctDNA v2.1.1 Software on ICA. For full details, please consult the DRAGEN TSO 500 ctDNA v2.1.1 on ICA 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.

- Software Version: 2.1.1
- DRAGEN software version 3.10.9

NEW FEATURES:

ICA functionality

- Sequencer integration including data streaming to ICA has been expanded and now supports NovaSeq 6000/6000Dx and DRAGEN TSO500 ctDNA v2.1.1 Analysis Software on ICA
- Auto-launch functionality now supports automated start of the DRAGEN TSO500 ctDNA v2.1.1 Analysis Software on ICA after completion of sequencing and data streaming utilizing settings provided in the sample sheet v2. The sample sheet can be generated manually or using BaseSpace Run Planning tool.
- DRAGEN TSO500 ctDNA v2.1.1 Analysis Software bundle on ICA has been updated with the specialized BCL Convert module (BCL Convert v3_10_9 for TSO500) improving experience for FASTQ -> VCF analysis
- Improvements to the ICA user interface: removed single tile option due to limited use, optimized fields for analysis input, removed requirement for user to specify locations of Resource folder and hashtable, added validation for the storage size selection
- Improvements to the content of error messages and logs
- Contextual help improvements

BaseSpace Sequence Hub (BaseSpace) functionality

- BaseSpace Run Planning tool now supports sample sheet v2 generation for DRAGEN TSO500 ctDNA v2.1.1 Analysis Software on ICA. Generated sample sheets can be uploaded to NovaSeq 6000/6000Dx to configure run, provide input for data streaming to ICA, auto-launch and running of the DRAGEN TSO500 ctDNA v2.1.1 Analysis Software on ICA.
- Analysis run folder and other analysis outputs (analysis error messages and logs) are now output in BaseSpace in addition to ICA
- DRAGEN TSO500 ctDNA Analysis Software users can now monitor analysis progress via BaseSpace interface in addition to monitoring sequencing status

Bioinformatics pipeline

- The following components were replaced with a DRAGEN version from a non-DRAGEN version for improved speed and accuracy:
  - MSI module
  - CNV caller
  - TMB module
  - Small Variant Caller
  - Sample QC module
- Performance improvements:
  - Improved MSI specificity
  - Enhanced bTMB accuracy due to improved filtering of CH and germline variants
  - Improved specificity for small variant calling
  - Ability to call complex variants panel-wide
  - Ability to call insertions and deletions > 25 bp
  - Improved Contamination QC to better handle samples with highly rearranged genomes
- Fusion directionality is now reported in the combined variant output
- The Illumina Annotation Engine (Nirvana) was updated to version 3.2.6
- Spoiler was updated for contamination calling
- The software now accepts FASTQ files generated by the stand-alone BCLconvert software
- The installer was updated to be compatible with Oracle Linux 8 and CentOS 7
- The software now provides a Metrics Output per sample as well as for the full run
- The Analysis Output has been updated
  - Now it generates an output folder in the specified location with the folder name "DRAGEN_TSO500_ctDNA_ Analysis_YYYMMDD-HHMMSS"
    - Within the analysis folder, each analysis step generates a subfolder within the Logs_ Intermediates folder.
  - Inputs to the running docker container are mapped from native locations on the server to the following locations in the container

Input	Running Docker Container Location
Run folder	/opt/illumina/run-folder Sample
Sample sheet	/opt/illumina/SampleSheet.csv
FASTQ folder	/opt/illumina/fastq-folder
Resources	/opt/illumina/resources
Analysis output folder	/opt/illumina/analysis-folder

- The pipeline was updated to reduce the time to failure in case of sample sheet errors. This was accomplished by running validation ahead of all other steps and allowing the step to be executed on a more widely available node. The current 'time to failure' now corresponds directly to the size of the input run or FASTQ folder, as this must first be copied into a scratch location to support SampleSheetValidation.

DEFECT REPAIRS:

- Illumina Annotation Engine 3.2.6 (aka Nirvana) includes the following enhancements

and bug fixes:

- Added genes and transcripts from the NCBI Homo sapiens Updated Annotation Release 105.20201022 to provide the latest RefSeq content for GRCh37
- Reduced the HGVS c. error rate by 54% and HGVS p. error rate by 20%. Currently the HGVS c. error rate is 0.00527% and the HGVS p. error rate is 0.00737%.
- Fixed issues related to incorrect CDS coordinates in some edge cases
- Improved detection of frameshifts when variants partially overlap the coding sequence

KNOWN ISSUES:

- 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 metrics output step module shows no error message when input file is missing.
- FastqGeneration issue: Missing .bcl files can cause FastqGeneration failure, but pipeline does not generate a MetricsOutput.tsv file with failed the steps.
- The software does not notify the user when InterOp files for RunQC are missing or corrupted.
- For 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, the runtime may take significantly longer than 20 hours
- ICA Run Time depends on ICA instance availability, it will be affected by region and traffic
- Some contrived samples such as SeraCare Complete Mutation Mix, which have multiple structural variants (SVs) and high 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.

PRODUCT LIMITATIONS:

- The sample sheet must be configured as described in the User Guide or by using BaseSpace Run Planning tool.
- Sample sheets generated for auto-launch are not compatible and cannot be reused without changes for DRAGEN TSO500 ctDNA Analysis Software v2.1.1 on a Local DRAGEN server, and vice versa.
- Added validation for the storage size selection generates an error if "Small" or "Medium" values are selected ("Large" is required as a minimum) but the error message appears with a delay.
- 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.
- For 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, the runtime may take significantly longer than 20 hours
- 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%.

Release History

Version	Workflow#	Author	Description of Change
00	CN 1086293	Svetlana Bureeva	Initial Release
01	CN 1095165	Svetlana Bureeva	Added support for the new assay version