DRAGEN TSO 500 Analysis Software Release Notes

V2.1.0

For TruSight Oncology 500, TruSight Oncology 500 HRD, and TruSight Oncology 500 High-Throughput

Introduction

These Release Notes detail the key changes to software components for the DRAGEN TSO 500 v2.1.0 Analysis Software.

This software is intended for use with the TruSight Oncology 500 Assay and the TruSight Oncology 500 HRD Assay.

- Software Version: 2.1.0
- Docker Image ID: 03033500cb0d
- DRAGEN version: 3.10.9

The software installer script, install_DRAGEN_TSO500-2.1.0.run, includes the following:

- dragen_tso500_2.1.0.tar – a tar file of the DRAGEN TSO500 docker image.
- uninstall_DRAGEN_TSO500-2.1.0.sh – a script for uninstalling DRAGEN TSO500.
- check_DRAGEN_TSO500-2.1.0.sh – a script for self-testing DRAGEN TSO500.
- build-hashtable_DRAGEN_TSO500-2.1.0.sh – a script for building the hash table.
- install.sh – a script used to install DRAGEN TSO500 based on the contents listed.
- resources/ – a directory containing all resources files necessary for DRAGEN TSO 500 Analysis Software.
- dragen-3.10.0-2.el7.x86_64.run – the DRAGEN installer.

NEW FEATURES (COMPARE TO DRAGEN TSO500 ANALYSIS SOFTWARE V1.1.1):

- The pipeline now includes DRAGEN variant callers for SNV, MNV, CNV, TMB, MSI, and gene fusion variant calling.
- Performance improvements:
  - Improvements in contamination QC and CNV QC to better handle samples with highly rearranged genomes
  - Improvements in MSI detection to better handle low quality samples and decrease the rate of sample QC failure
  - Improvements in small variant specificity
  - Improvements in RNA fusion sensitivity
  - Capability of calling insertions/deletions > 25bp and complex variants >3bp panel-wide
  - Improved sensitivity of EGFR complex variants
- BRCA 1/2 large rearrangements (exon-level CNVs) called at 43% or higher VAF for 3 or more exons, 50% or higher VAF for less than 3 exons.
- Genomic instability scoring powered by HRD technology from Myriad Genetics (not available in Japan; must be run with the TSO 500 HRD kit (MN 20076480) and an HRD license (MN 20073738))

DEFECT REPAIRS (COMPARE TO DRAGEN TSO500 ANALYSIS SOFTWARE V1.1.1):

- 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
- Fusion caller can now call fusions when breakpoint(s) are located in region(s) with high homology.

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. Blank rows in the [Data] section may cause a run failure.
- The sample sheet should not have blank rows after samples in the [Data] section. The workflow will fail if blank rows are present after the [Data] section.
- The TSO 500 RNA workflow is unstranded. Fusions or splice variants could involve antisense transcripts instead of the reported genes.
- An erroneous warning is displayed when running the software - "Access to undefined parameter `report_folder`"
- The TSO 500 analysis pipeline crashes when Illumina Annotation Engine 3.2.5 (used for TMB calculation) encounters rare annotation issues related to using Ensemble transcripts for a limited set of variants.
- If a certain step of the RNA sub-workflow (FASTQ Downsample, RNA QC, Splicegirl, or Fusion Merge) fails for at least one sample, the step is erroneously listed in the "Steps Not Completed" section of the Metrics Output file for all samples of the run. For the samples where the step did not fail, the Metrics Output and the Combined Variant Output files still report results.

PRODUCT LIMITATIONS:

- The sample sheet must be configured as described in the User Guide.
- This software version is only compatible with DRAGEN version 3.10.9.

- Performance not verified using reads other than 2 x 101.
- GIS analysis has not been verified using libraries with UDP indexes.
- 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.
- Unmapped long insertions are not likely to occur on shorter indels because there is sufficient reference-matching sequence in the reads.
- Incorrect calculation of variant allele frequency can occur in variants near the start and end of genomic reads, but there is a low probability of incorrect variant allele frequency in called variants due to sufficient variation in read start and end positions.
- Germline estimation uses 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.
- Germline estimation is difficult when tumor purity is >85% causing expected variant allele frequency for somatic and germline variants to converge.
- Poor quality wild type reads may align as chimeric and be miscalled during RNA analysis.
- BRCA 1/2 exon-level CNVs with two segments that diverge equidistant from baseline in opposite directions in highly rearranged genomes would occasionally report a "GAIN" due to variation in the calculated distance from baseline. These samples are expected to have high genomic instability and will be filtered as "undetermined".
- The contamination score threshold will fail approximately 1% of HRD samples due to the variant allele frequency (VAF) shifts of highly rearranged genomes and not true contamination of foreign human DNA. Visually investigation of VAFs across the genome can be performed to determine if a shift of VAFs is due true contamination.
- 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%.