DRAGEN TSO 500 Analysis Software Release Notes

V2.6.0

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

October 25, 2024

Introduction

These Release Notes detail the key features and known limitations for the DRAGEN TSO500 v2.6.0 Analysis Software. For full details, please consult the DRAGEN TSO 500 v2.6.0 Analysis Software User Guide available on the support website.

This software is intended for use with the TruSight Oncology 500, TruSight Oncology 500 High-Throughput, and TruSight Oncology 500 HRD assays.

DRAGEN TSO500 v2.6.0 Analysis Software is available:

- On DRAGEN server
- On Illumina Connected Analytics (ICA)
- As NovaSeq 6000Dx Analysis Applications:
    - Illumina DRAGEN TSO 500 (HRD) v2.6.0 Analysis Application on NovaSeq 6000Dx (RUO) v2.6.0-2v12
    - Illumina DRAGEN TSO 500 v2.6.0 Analysis Application on NovaSeq 6000Dx (RUO) v2.6.0-2v12 (for Japan only). This application does not support TruSight Oncology 500 HRD assay.

These Release Notes also cover features of BaseSpace Run Planning tool related to the run set-up and sample sheet generation for DRAGEN TSO 500 Analysis Software.

This document describes features and limitations for all abovementioned platforms. If an item is specific to a software platform, it will be specified using a tag: "ICA", "Server", "NovaSeq 6000Dx app", "BaseSpace Run Planning Tool".

NEW FEATURES:

- DRAGEN version was upgraded to 3.10.17. Installation on the standalone DRAGEN server requires DRAGEN version 3.10.17 or above.
- LibraryPrepKit field in the Sequencing Settings section of the sample sheet is now required
- Added ability to have spaces in sample sheet names without failing sample sheet validation
- Sample sheet templates are removed from the Resource Bundle due to the increased number of potential configurations to support the expanded set of instruments. Users are encouraged to use BaseSpace Run Planning Tool to generate sample sheets or use a sample sheet template library in the DRAGEN TSO 500 Analysis Software user guide.
- Per-sample gene- and exon-level coverage reports are provided in files {Sample_ID}.exon_coverage_report.tsv and {Sample_ID}.gene_coverage_report.tsv to facilitate downstream analysis and visualizations
- Added a new DRAGEN FASTQ QC metric - percentage of bases above Q30 (PCT_Q30_BASES) - to the MetricsOutput.tsv and <PREFIX>.mapping_metrics.csv

- Added a new QC metric - percent of soft-clipped reads (PCT_SOFT_CLIPPED_BASES (%)) - to the MetricsOutput.tsv
- Exon-Level CNVs now listed as Large Rearrangements in CombinedVariantOutput.tsv
- Added chromosome "stop" column to the "_logRatio.tsv" file.
- Illumina Connected Annotations (formerly, Illumina Annotation Engine, Nirvana) was updated to v3.2.7.
- (Server) Two versions of DRAGEN TSO 500 Analysis Software can now be installed on the same DRAGEN server if one of them is v2.6.0 and another is v2.x. During the installation, the version with the highest 3rd digit should be installed last.
- (ICA) Auto-launch of DRAGEN TSO 500 v2.6.0 Analysis Software on ICA when sequencing is performed on NovaSeq X, NextSeq 1000, and NextSeq 2000.
- (ICA) Auto-launch of DRAGEN TSO 500 v2.6.0 Analysis Software on ICA when starting the analysis with FASTQ files. The functionality can be used together with DRAGEN BCL Convert (that can be auto-launched before auto-launching DRAGEN TSO 500 v2.6.0 Analysis Software on ICA) to enable fully automated workflow when mixing different assays on one flow cell.
- (BaseSpace Run Planning Tool) Support for run planning and sample sheet generation for NovaSeq X, NextSeq 1000, and NextSeq 2000.
- (BaseSpace Run Planning Tool) Support for the new sample sheet parameter to start analysis with DRAGEN TSO 500 v2.6.0 Analysis Software on ICA with FASTQ files.
- (BaseSpace Run Planning Tool) Ability to handle flexible reads in BaseSpace Run Planning Tool. This enables the ability to re-queue analysis with modified reads in BaseSpace.

DEFECT REPAIRS:

- Fixed an issue causing an EGFR exon20 inframe_insertion variant be reported as intron variant and splice region variant when the insertion is a duplicate sequence of the intron/exon boundary.
- Fixed an issue causing a splice donor/acceptor site variant consequence to be reported incorrectly when there is no actual sequence change.
- Fixed issues causing RNA fusion false positive calls leading to improved specificity. ILMN Ref. 31642.
- Fixed issues causing incorrect and inconsistent RNA fusion calls between analysis runs. ILMN Ref. 31663.
- Fixed issues causing certain SNV false positive calls leading to improved specificity. ILMN Ref. 31962.
- Removed MNVs from TMB calculation as per algorithm design. ILMN Ref. 32574.
- Introduced improvements to reduce false negative calls for BRCA1 and BRCA2 one exon deletions contributing to improved sensitivity. ILMN Ref. 28747.
- Fixed an issue causing GIS score not to appear in CombinedVariantOutput.tsv in an HRD run with DNA and RNA samples.

KNOWN ISSUES:

- If Sample Feature is labeled as an HRD sample, but HRD probes fail, the CNV output will be marked as "NA" in the CombinedVariantOutput.tsv. The CNV results can be recovered by removing the HRD label in the sample sheet. This issue occurs in previous versions with absolute copy number enabled (v2.5.0 and above).
- MetricsOutput file is not always generated when input BCL files are corrupt or missing. Users are recommended to check the Errors.tsv file for analysis status. ILMN Ref. A34561.
- (NovaSeq 6000Dx app) On the Results page, the UI displays a a section header titled Sample ID, but no results are displayed in the table below. ILMN Ref. A25859.
- (NovaSeq 6000Dx app) The application will launch the analysis when the storage on the server is less than 4TB, which may lead to running out of space.
- (NovaSeq 6000Dx app) Requeue analysis with no changes throws an error. It works when the second option "Edit run settings and requeue analysis" is selected.
- (NovaSeq 6000Dx app) Analysis Run Results not displayed in run details for runs started by users with the role Sequencer Operator User.

PRODUCT LIMITATIONS:

- (ICA) Sample sheets generated for analysis on a standalone DRAGEN server cannot be used (without changes) to run DRAGEN TSO500 Analysis Software on ICA with auto-launch.
- (ICA) ICA Run Time depends on ICA instance availability, it will be affected by region and traffic.
- Performance not verified using reads other than 2 x 101.
- 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.
- The TSO 500 RNA workflow is unstranded. Fusions or splice variants could involve antisense transcripts instead of the reported genes.
- A high number of chimeric reads due to poor quality RNA libraries can lead to false positive RNA fusions reported.
- TMB number may be inflated in samples with >5% supplementary (chimeric) alignments due to the larger number of false positive indels. ILMN Ref. 8507.
- Germline estimation that is used for TMB calculation 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 the TMB estimation.
- Germline estimation is difficult when tumor purity is > 85% causing expected variant allele frequency for somatic and germline variants to converge.
- Some regions are known to be difficult to sequence. One example region is the TERT promoter region. Although sequencing can occur at the TERT promoter region, this location might result

in low coverage due to the GC rich content of the sequenced region. Another example region is the PMS2 gene which has high homology to pseudogenes and reads may not align properly. In general, the TSO 500 panel is designed to target unique regions, and the software accounts for background noise during small variant calling for each genomic position. This design is meant to prevent false positive calls. Analytical performance of the assay is evaluated panel-wide rather than for each gene or exon. However, due to these challenges certain regions covered in the product manifest are excluded from analysis due to high background noise. All excluded variants are identified in the VCF using a flag. This block list includes the following genes: HLA-A, HLA-B, HLA-C, KMT2B, KMT2C, KMT2D, chrY and positions with VAF > 1% occurred in six or more of the 60 baseline samples. The block list of excluded sites can be obtained on request from your local Illumina representative.

- Lower sensitivity and specificity may be seen in CNV amplifications and deletions with less than 20 probes and higher noise profiles. Contact your local Illumina representative for more details. The following genes are excluded from CNV calling due to high homology: HLA-A, HLA-B, HLA-C, KMT2B, KMT2C, KMT2D, HIST2H3A, HIST2H3C. These genes are excluded from CNV calling due to insufficient probe coverage (1 probe): DNAJB1, FANCF, FOXL2, HIST1H3A, HIST1H3C, HIST1H3D, HIST1H3E, HIST1H3F, HIST1H3G, HIST1H3H, HIST1H3I, HIST1H3J, HIST2H3D, TERC, and TERT (only covers promoter region); results for these genes are included in the VCF but are not included in the CombinedVariantOutput.tsv. ILMN Ref. 34500.
- Illumina Connected Annotations (formerly, Illumina Annotation Engine, 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%.
- When a single nucleotide insertion introduces only a stop codon, it should be annotated as stop_gained, but instead Illumina Connected Annotations (formerly, Illumina Annotation Engine, Nirvana) annotates the event as a frameshift_variant. ILMN Ref. 33596.
- BRCA1 and BRCA2 large rearrangements (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”.
- False negatives for BRCA1 and BRCA2 large rearrangements (exon-level CNVs) with a single or partial exon loss or gain and VAF lower than 61% are observed at higher rate than presented in product specification (sensitivity of 95% at VAF 50% or higher for fewer than 3 exons) due to the higher amount of noise associated with the smaller segment size. Pathogenic variants with single or partial exon CNVs are expected to have a prevalence of 0.17% in ovarian cancer samples (Jones et al., Genes Chromosomes Cancer.2023;62:589–596). The current implementation was designed to reduce false positives and has shown to have a high gene-level specificity (100%) with internal testing. ILMN Ref. 28747
- Genomic Instability Score and BRCA large rearrangements (exon-level CNVs) have not been verified with input over 80ng of FFPE.
- GIS analysis has not been verified using libraries with UDP indexes.
- RNA DRAGEN mapping in lower quality samples has been found to have a high number of duplicates that are not marked, leading to SpliceGirl to incorrectly call RNA fusions and splice variants. ILMN Ref. 35248.

- A lower call rate of RNA fusions can occur due to germline variants near a breakpoint which penalizes the alignment score leading to lower number of supplementary alignments. This prevents supplementary alignments from being counted as supporting reads. This is an RNA DRAGEN mapper limitation. ILMN Ref. 32445.
- The estimates for tumor fraction and ploidy may be less reliable for samples with lower Genomic Instability Score as they will have fewer genome rearrangements.
- 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. Visual investigation of VAFs across the genome can be performed to determine if a shift of VAFs is due to true contamination.
- DRAGEN small variant caller performs realignment of reads to reconstructed haplotypes, if a realigned read does not have the same length of CIGAR string as the original alignment a sample will fail small variant calling. This is a very rare occurrence and resequencing of the library has been found to remove the read causing this error. ILMN Ref. 34810.
- Several bioinformatics features have *beta* status. Beta features have not been verified by Illumina due to limited access to samples *or* lack of an appropriate orthogonal method to perform testing, and the use of *in silico* testing alone is not sufficient for verification purposes. Beta features are only available with the TSO 500 HRD kit and include:
- Tumor fraction
- Ploidy
- Absolute copy numbers
- Gene-level loss of heterozygosity (LOH) events

Release History

Revision	Release Reference	Originator	Description of Change
00	CN 1114722	Svetlana Bureeva	Initial Release