TSO500 Local App Software Release Notes

V2.2.0

For TruSight Oncology 500 Assay

November 13, 2020

Introduction

These Release Notes detail the key features and known limitations to software components for the Local TSO500 App v2.2.0

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

The software package includes:

- trusight-oncology-500-ruo-2.2.0.12.zip
    - md5sum.txt
    - resources
    - trusight-oncology-500-ruo-dockerimage-ruo-2.2.0.12.tar
    - trusight-oncology-500-ruo.img
    - TruSight_Oncology_500_RUO.sh

NEW FEATURES:

- Support for new NovaSeq 6000 v1.5 Reagent Kits.
- Support for new "V2" samplesheet template.
- Customers may now analyze a subset of samples by designating Sample ID or Pair ID using the "—sampleOrPairIDs" command line argument (replaces the "–samplePairIDs" argument).
    - PairID now required when submitting the sample sheet.

DEFECT REPAIRS:

- BCL-Convert version updated to v3.5.7 to prevent analyses from crashing due to a rare tile drop out issue.
- The "Splice Support Reads" and the "Reference Reads Transcript" columns in the Combined Variant Output files have been corrected (previously the values for these were reversed).
- Read-stitching component (Gemini) updated to prevent an analysis from crashing if an indel was detected at the end of a chromosome.
- Read collapsing component (ReCo) updated to correct a metric that is used during read collapsing.
- Gene name for gene amplifications corrected from "MYCL1" to "MYCL".
- Duplicate content in the header of the MergedSmallVariants.genome.vcf has been removed.
- Variants detected outside of the target manifest regions are now removed.
- Manta RNA fusion calling step has been corrected to address a rare race condition which can cause analysis to fail.

- Illumina Annotation Engine updated to address over 99% of known situations where the software may report incorrect C-Dot and P-Dot notation values for DNA variants on affected RefSeq transcripts (see Product Limitations below).

KNOWN ISSUES:

- Local app may use additional available hardware/compute resources, recommendation is to not run multiple local app instances on a single node.
- When using NovaSeq XP workflow, sample sheets require lane information and may provide unclear error messages when lane information is excluded.
- An error will occur when trying to run more than 96 samples on a single node.
- Rare instances of low multiplexed runs (3 DNA samples in a run) cause delayed results when performing read collapsing.

PRODUCT LIMITATIONS:

- Performance not verified using read lengths other than 2 x 101.
- FASTQ Generation is not supported with read lengths other than 2 x 101 or 2 x 151.
- Unmapped long insertions are not likely to occur on shorter indels because there is sufficient reference-matching sequence in the reads. Product claims only indels up to 25 base pairs.
- Complex variants are specifically output only for a specific region of the EGFR gene, component and phased variants would both be contained in the output.
- Incorrect calculation of variant allele frequency can occur in variants near the start and end of genomic reads, but variation in read start and end positions in an enrichment assay is sufficient to make incorrect variant allele frequency in output variants a low-probability situation.
- Germline estimation using high tumor purity (>70%) can impact estimation, due to somatic and germline variants appearing with similar variant allele frequency.
- Germline estimation uses latest publicly available population data and estimated to be representative of targeted population, the impact of rare germline mutations is expected to be limited.
- Poor quality wild type reads may align as chimeric and be miscalled during RNA analysis.
- When filtering fusion variant candidates, strand consideration is not used in annotation, which allows contiguous sequences to be intragenic, when they are on different strands.
- Manta will not call fusions where both breakpoints map to the same gene transcript: FIP1L1-PDGFRA (when both breakpoints overlap Ensembl transcript ENST00000507166) and GOPC-ROS1 (when both breakpoints overlap Ensembl transcript ENST00000467125).
- Manta may not always call fusions in the following situations, which have been exclusively observed in synthetic commercial controls:
    - Multiple fusion breakpoints from a single fusion gene pair with breakpoints within approximately 150 base pairs of each other (observed with ETV6-ABL1 and ETV6-NTRK3 fusions).
    - Multiple fusions from two different gene pairs with breakpoints within approximately 150 base pairs of each other (observed with IRF2BP2-NTRK1, TFG-NTRK1, SQSTM1-NTRK1 fusions).
    - Breakpoint(s) are located in region(s) with high homology (observed with fusions with breakpoints on SEPT14 exon 10).

- Gathering output requires all samples scattered to be gathered at once, gathering a subset of results is not supported.
- The Illumina Annotation Engine (aka Nirvana) may report incorrect protein (P-Dot) and transcript (C-Dot) changes in HGVS nomenclature for small variants located on a RefSeq transcript where an RNA-edit has occurred. Most known variants on these transcripts are unaffected. A list of affected Canonical RefSeq transcripts and Cosmic Variants from those transcripts can be found below. A full explanation of this product limitation can be found in PQN2020-1090. [1]

Affected Canonical RefSeq Transcripts

Transcript ID	Gene Symbol
NM_002467.4	MYC
NM_003224.5	ARFRP1
NM_004119.2	FLT3
NM_006904.6	PRKDC
NM_198291.2	SRC
NM_021960.4	MCL1
NM_001025366.2	VEGFA

Affected Cosmic Variants from Canonical RefSeq Transcripts

The list of affected variants is based on an analysis of COSMIC database version 92 variants located along the Canonical RefSeq Transcripts listed above [2]. New variants are regularly submitted to COSMIC. This list of affected variants may change over time.

Chr:Position	REF*	ALT**	Gene Symbol	Transcript ID	COSMIC_ID
chr1:150548890	A	ATCTA	MCL1	NM_021960.4	COSV57189597
chr6: 43738444	C	T	VEGFA	NM_001025366.2	COSV104569261
chr8:48805817	G	GG	PRKDC	NM_006904.6	COSV58041377
chr8:128748839	GC	G	MYC	NM_002467.4	COSV104388447
chr8:128748840	C	A	MYC	NM_002467.4	COSV104388806
chr8:128748840	C	G	MYC	NM_002467.4	COSV104388204
chr8:128748841	T	C	MYC	NM_002467.4	COSV104388663
chr13:28608094	C	CACTTTTCCAAAAGCACCTGATCCTAGTACCTTCCCAAACTCTAAATTTTCTCTTGGAAACTCCCATTTGAGATCATATTCATATTCGTTCATC	FLT3	NM_004119.2	COSV54069050
chr13:28608124	C	CTTCCCAAACTCTACTGTTGCGTTCATCACTTTTCCAAAAGCACCTGATCCTAGTACC	FLT3	NM_004119.2	COSV54044227
chr13:28608129	C	CAAACTCAAAAGCACCTGATCCTAGTACCTTCCC	FLT3	NM_004119.2	COSV54054381
chr13:28608129	C	CAAACTCTAAATTTTCTCTTGGAAACTCCCATTATCCTAGTACCTTCCC	FLT3	NM_004119.2	COSV54043729
chr13:28608129	C	CAAACTCTAAATTTTCTCTTGGAAACTCCCATTTTCCAAAAGCACCTGATCCTAGTACCTTCCC	FLT3	NM_004119.2	COSV54075746
chr20:36030939	G	GTGGCC	SRC	NM_198291.2	COSV99050886
chr20:62331336	C	CC	ARFRP1	NM_003224.5	COSV53926174

*Reference base(s)
**Alternate base(s)

[1] TSO 500 uses the Canonical RefSeq transcript when annotating variants passed into the Combined Variant Output file. The Illumina Annotation Engine selects canonical transcripts based on the following rules…

- Order all overlapping transcripts by coding sequence length.
- Pick the longest transcript that has an associated Locus Reference Genome (LRG) sequence.
- If no LRGs exist for the set of transcripts, pick the longest transcript that is coding.
- If there is a tie, pick the transcript with the smaller accession id number.

[2] Released 27 August 2020.

Release History

Version	ER#	Author	Description of Change
00	1044258	Trey Howard	Initial Release