MiSeq Reporter TruSight Tumor 15

Workflow Guide

For Research Use Only. Not for use in diagnostic procedures.

ILLUMINA PROPRIETARY
Document # 1000000001205 v02
September 2016

illumina®

Document # 1000000001205 v02

Introduction

The TruSight® Tumor 15 workflow is exclusively designed for the TruSight Tumor 15 targeted resequencing assay. Other sample preparation methods are not compatible with this workflow.

In the MiSeq Reporter Analyses tab, a run associated with the TruSight Tumor 15 workflow is represented with the letters **TT**. For more information about the MiSeq Reporter interface, see the *MiSeq Reporter Software Guide (document # 15042295)*.

This guide describes the analysis steps performed in the TruSight Tumor 15 workflow, and the types of information and analysis files generated by the workflow.

Workflow Requirements

- **Two manifest files**—The TruSight Tumor 15 workflow requires 2 assay-specific manifest files, TST_15-A-Manifest and TST_15-B-Manifest. Manifest files are included with the MiSeq Reporter software.
- **Reference genome**—In addition to the manifest files, the TruSight Tumor 15 workflow requires the hg19 reference genome for coordinates and chromosome mapping. By default, this reference is included with the MiSeq Reporter software.
- **MiSeq Reporter v2.6**, or later—Previous versions of MiSeq Reporter software are not compatible with the TruSight Tumor 15 workflow.
  MiSeq Reporter software is available from the MiSeq Reporter support page on the Illumina website.
- **Illumina Annotation Engine Support File**— Enables TruSight Tumor 15 workflow analysis. If upgrading from MiSeq Reporter v2.5 to v2.6, download the file from the MiSeq Reporter support page on the Illumina website.

NOTE
Settings and Reads sections of the TruSight Tumor 15 workflow sample sheet cannot be changed.

TruSight Tumor 15 Workflow Overview

Designed exclusively for the TruSight Tumor 15, the TruSight Tumor 15 workflow specifically reports on somatic variants of a set of reference panel genes associated with cancer. The genes queried are specified in the *TruSight Tumor 15 Reference Guide (document # 1000000001245)*.

The TruSight Tumor 15 workflow evaluates short regions of amplified DNA, or amplicons, for variants. Focused sequencing of amplicons enables high coverage of particular regions across many samples. This workflow is uniquely suited for detection of variants in formalin-fixed paraffin-embedded (FFPE) samples.

The workflow requires both Mix A and Mix B to analyze targeted regions.

The TruSight Tumor 15 workflow demultiplexes indexed reads, generates FASTQ files, aligns reads to a reference, identifies variants, and writes output files to the Alignment folder.

The workflow produces the following reports:

- **Run Metrics Report**— Shows run metrics and suggested values to determine if run quality results are within an acceptable range. For more information, see *RunMetricsReport.txt File* on page 1.
- **Sample Metrics Report**— Provides information on base coverage and read alignment for each sample. For more information, see *SampleMetricsReport.txt File* on page 1.
- **Filtered gVCF File Report**— Lists variant calls for a subset of variants specified in the TruSight Tumor 15 Report Definition File. For more information, see *Filtered gVCF File Report* on page 1.

Demultiplexing

For runs with multiple samples and index reads, demultiplexing compares each Index Read sequence to the index sequences specified in the sample sheet. No quality values are considered in this step.

Demultiplexing separates data from pooled samples based on short index sequences that tag samples from different libraries. Index reads are identified using the following steps:

- Samples are numbered starting from 1 based on the order they are listed in the sample sheet.
- Sample number 0 is reserved for clusters that were not successfully assigned to a sample.
- Clusters are assigned to a sample when the index sequence matches exactly or there is up to a single mismatch per Index Read.

NOTE
Illumina indexes are designed so that any index pair differs by ≥ 3 bases, allowing for a single mismatch in index recognition.

When demultiplexing is complete, 1 demultiplexing file named DemultiplexSummaryF1L1.txt is written to the Alignment folder with the following information:

- In the file name, **F1** represents the flow cell number.
- In the file name, **L1** represents the lane number, which is always L1 for MiSeq.
- A table of demultiplexing results with 1 row per tile and 1 column per sample, including sample 0.
- The most commonly occurring sequences for the index reads.

Other demultiplexing files are generated for each tile of the flow cell. For more information, see *Demultiplexing File Format* on page 1.

FASTQ File Generation

MiSeq Reporter generates intermediate analysis files in the FASTQ format, which is a text format used to represent sequences. FASTQ files contain reads for each sample and their quality scores, excluding clusters that did not pass filter.

FASTQ files are the primary input for alignment. The files are written to the BaseCalls folder (Data\Intensities\BaseCalls) in the MiSeqAnalysis folder, and then copied to the BaseCalls folder in the MiSeqOutput folder. Each FASTQ file contains reads for only 1 sample, and the name of that sample is included in the FASTQ file name. For more information, see *FASTQ File Naming* on page 1.

Alignment

During the alignment step, the banded Smith-Waterman algorithm aligns clusters from each sample against amplicon sequences specified in the manifest file.

The banded Smith-Waterman algorithm performs local sequence alignments to determine similar regions between 2 sequences. Instead of comparing the total sequence, the Smith-Waterman algorithm compares segments of all possible lengths. Local alignments are useful for dissimilar sequences that are suspected to contain regions of similarity within the larger sequence. This process allows alignment across small amplicon targets, often less than 10 bp.

Each paired-end read is evaluated in terms of its alignment to the relevant probe sequences for that read.

- Read 1 is evaluated against the reverse complement of the Downstream Locus-Specific Oligos (DLSO).
- Read 2 is evaluated against the Upstream Locus-Specific Oligos (ULSO).
- If the start of a read matches a probe sequence with no more than 1 mismatch, the full length of the read is aligned against the amplicon target for that sequence.

Alignments that include more than 3 indels are filtered from alignment results. Filtered alignments are written in alignment files as unaligned and are not used in variant calling.

Paired-End Evaluation

For paired-end runs, the top-scoring alignment for each read is considered. Reads are flagged as an unresolved pair under the following conditions:

- If either read did not align, or the paired reads aligned to different chromosomes.
- If 2 alignments come from different amplicons or different rows in the Targets section of the manifest.

Bin/Sort

The bin/sort step groups reads by sample and chromosome, and then sorts by chromosome position. Results are written to 1 BAM file per sample library.

Variant Calling

SNPs and short indels are identified using the somatic variant caller. Developed by Illumina, the somatic variant caller identifies variants present at low frequency in the DNA sample and minimizes false positives.

The somatic variant caller identifies SNPs in 3 steps:

- Considers each position in the reference genome separately
- Counts reference and variant bases at the given position for aligned reads that overlap the position
- Computes a variant score that measures the quality of the call

Variant scores are computed using a Poisson model that filters variants with a quality score below Q30. Also, the model only calls variants for bases that have a coverage depth of ≥ 500.

Variants are called for Mix A and Mix B separately. If a variant meets the following criteria, the variant is reported in the variant file:

- Has a depth of ≥ 500
- Has a variant frequency of 2.6% or greater as reported in the VCF files

Reasons why a locus for a mutation or a reference is classified as a "no call" include the following:

- The variant frequency is near the signal noise level, between 1% and 2.6%
- Has a depth of < 500
- Significant stand bias is detected, which can be a sequencing-specific error

Variant Annotation

The TruSight Tumor 15 workflow provides the following annotation for each variant in the VCF.

- Gene name
- Nucleotide change
- Transcript
- Amino acid change
- Consequence

Read Stitching

Read stitching is compatible only with the Amplicon DS workflow, GenerateFASTQ workflow, TruSight Tumor Panel (15 Genes) workflow, and TruSeq Amplicon workflow. Read stitching is not possible with any other Illumina alignment method or analysis workflows, but might be allowable input with some third-party analysis tools using the FASTQ files.

For each paired read, a minimum of 10 bases must overlap between Read 1 and Read 2 to be a candidate for read stitching. The minimum threshold of 10 bases minimizes the number of reads that are stitched incorrectly due to a chance match. Candidates for read stitching are scored as follows:

- For each possible overlap of 10 base pairs or more, a score of 1 – MismatchRate is calculated.
- Perfectly matched overlaps have a MismatchRate of 0, resulting in a score of 1.
- Random sequences have an expected score of 0.25.
- If the best overlap has a score of ≥ 0.9 ***and*** the score is ≥ 0.1 higher than any other candidate, then the reads are stitched together at this overlap.

Although the stitched reads are aligned as one, in the BAM file the stitched alignment is split into individual alignments.

During variant calling, stitched reads are processed together. A consensus read is generated by taking the base call and quality score of the read with the higher Q-score in the overlap region. When the Q-score is the same, but the base call differs, a “no call” is

used at that position. Sometimes read stitching can improve the accuracy of variant calling.

Paired-end reads that cannot be stitched are converted to 2 single reads in the FASTQ file.

Reports

Report Name	Description
Run Metrics Report	The run metrics report lists software versions, run ID, instrument ID, and sample information. The report shows run metrics and suggested values for quality of run results within an acceptable range. For Read 1 and Read 2, the report shows the percentage of reads with a quality score ≥ Q30.
Sample Metrics Report	The sample metrics report provides information for each sample in each library and includes calculations from the gVCF file that indicate the number of bases in the library that have ≥ 500 coverage. The report shows the percentage of reads in each library that aligned to the reference specified in the manifest, called on-target reads. The following information is listed for each sample ID: • The percentage of bases that have ≥ 500 Read 1 and Read 2 coverage for Mix A and Mix B • Out of the total number of reads, the percentage of reads on-target for Mix A and Mix B For amplicon coverage specifications, see the *TruSight Tumor 15 Data Sheet (document # 1170-2015-003)* on the Illumina website.
Filtered Output Report	The TruSight Tumor 15 workflow provides variant calls for all genes specified in the manifest files. The workflow also filters gVCF file information for a subset of variants. The subset is listed in the TruSight Tumor 15 Report Definition File and included with the MiSeq Reporter software installer. When a variant from this list is detected, it is added to a filtered report The filtered gVCF file report is provided in *.pdf, *.txt, and *.vcf formats. Filtered files are written to the Alignment folder and use the naming convention: SampleName_Report.vcf, *.pdf, or *.txt.

Analysis Output Files

The following analysis output files are generated for the TruSight Tumor 15 workflow and provide analysis results for alignment, variant calling, and coverage.

File Name	Description
*.bam files	Contains aligned reads for a given sample library. Located in the root level of the run folder in Alignment\Libraries.
*.vcf files	Contains information about variants found at specific positions in a reference genome. Per-library files are located in the root level of the run folder in Alignment\Libraries. Merged files are located root level of the run folder in Alignment.
gVCF files (*.genome.vcf)	Contains the genotype for each position, whether called as a variant or called as a reference. For more information, see *Genome VCF Files* on page 12. Per-library files are located in the root level of the run folder in Alignment\Libraries. Merged files are located root level of the run folder in Alignment. Merged files have the extension *.sample.genome.vcf.
AmpliconCoverage_M#.tsv	Contains details about the resulting coverage per amplicon per sample. M# represents the manifest number. Located in the root level of the run folder in Alignment.
*.ant files	Accompanying annotation files of merged variant call files and filtered VCF files. For import into VariantStudio software. Located in the root level of the run folder in Alignment.

Alignment Files

Alignment files contain the aligned read sequence and quality score. MiSeq Reporter generates alignment files in the BAM (*.bam) file format.

BAM File Format

A BAM file (*.bam) is the compressed binary version of a SAM file that is used to represent aligned sequences. SAM and BAM formats are described in detail at samtools.github.io/hts-specs/SAMv1.pdf.

BAM files are written to the root level of the run folder in Alignment\Libraries. BAM files use the file naming format of SampleName_S#.bam, where # is the sample number determined by the order that samples are listed in the sample sheet.

BAM files contain a header section and an alignments section:

- **Header**—Contains information about the entire file, such as sample name, sample length, and alignment method.
  Alignment methods include banded Smith-Waterman, Burrows-Wheeler Aligner

(BWA), and Bowtie. The term Isis indicates that an Illumina alignment method is in use, which is the banded Smith-Waterman method.

- **Alignments**—Contains read name, read sequence, read quality, alignment information, and custom tags.

```
GA23_40:8:1:10271:11781 64 chr22 17552189 8 35M * 0 0
TACAGACATCCACCACCACACCCAGCTAATTTTTG
IIIII>FA?C::B=:GGGB>GGGEGIIIHI3EEE#
BC:Z:ATCACG XD:Z:55 SM:I:8
```

The read **GA23_40:8:1:10271:11781** maps to the chromosome 22 (**chr22**) at position **17552189**, with alignment quality **8**, and match descriptor CIGAR string **35M**.

BAM files are suitable for viewing with an external viewer such as IGV or the UCSC Genome Browser.

BAM index files (*.bam.bai) provide an index of the corresponding BAM file.

Variant Call Files

Variant call files contain called variants. For the TruSight Tumor 15 workflow, MiSeq Reporter generates variant call files as VCF files and genome VCF files:

- VCF files contain information about variants found at specific positions.
- gVCF files contain information about all sites within the region of interest.

VCF File Format

VCF is a widely used file format developed by the genomics scientific community that contains information about variants found at specific positions in a reference genome.

VCF files use the file naming format SampleName_S#.vcf, where # is the sample number determined by the order that samples are listed in the sample sheet.

VCF File Header—Includes the VCF file format version and the variant caller version. The header lists the annotations used in the remainder of the file. The VCF header also contains the command line call used by MiSeq Reporter to run the variant caller. The command-line call specifies all parameters used by the variant caller, including the reference genome file and .bam file. The last line in the header is column headings for the data lines. For more information, see *VCF File Annotations* on page 1.

VCF File Data Lines—Contains information about a single variant. Data lines are listed under the column headings included in the header.

VCF File Headings

The VCF file format is flexible and extensible, so not all VCF files contain the same fields. The following tables describe VCF files generated by MiSeq Reporter.

Heading	Description
CHROM	The chromosome of the reference genome. Chromosomes appear in the same order as the reference FASTA file.
POS	The single-base position of the variant in the reference chromosome. For SNPs, this position is the reference base with the variant; for indels or deletions, this position is the reference base immediately before the variant.

Heading	Description
REF	The reference genotype. For example, a deletion of a single T is represented as reference TT and alternate T. An A to T single nucleotide variant is represented as reference A and alternate T.
ALT	The alleles that differ from the reference read. For example, an insertion of a single T is represented as reference A and alternate AT. An A to T single nucleotide variant is represented as reference A and alternate T.
QUAL	A Phred-scaled quality score assigned by the variant caller. Higher scores indicate higher confidence in the variant and lower probability of errors. For a quality score of Q, the estimated probability of an error is $10^{-(Q/10)}$. For example, the set of Q30 calls has a 0.1% error rate. Many variant callers assign quality scores based on their statistical models, which are high relative to the error rate observed.

VCF File Annotations

Heading	Description
FILTER	If all filters are passed, **PASS** is written in the filter column. • **LowDP**—Applied to sites with depth of coverage below a cutoff. • **LowGQ**—The genotyping quality (GQ) is below a cutoff. • **LowVariantFreq**—The variant frequency is less than the given threshold. • **R8**—For an indel, the number of adjacent repeats (1-base or 2-base) in the reference is greater than 8. • **SB**—The strand bias is more than the given threshold. For more information about sample sheet settings, see *MiSeq Sample Sheet Quick Reference Guide (document # 15028392)*.
INFO	Possible entries in the INFO column include: • **CSQ**—Consequence as predicted by Illumina Annotation Engine (IAE). • **DP**—The depth (number of base calls aligned to a position and used in variant calling).

Heading	Description
FORMAT	The format column lists fields separated by colons. For example, GT:GQ. Available fields include: • **AD**—Entry of the form X,Y, where X is the number of reference calls, and Y is the number of alternate calls. • **GQ**—Genotype quality. • **GQX**—Genotype quality. GQX is the minimum of the GQ value and the QUAL column. In general, these values are similar; taking the minimum makes GQX the more conservative measure of genotype quality. • **GT**—Genotype. 0 corresponds to the reference base, 1 corresponds to the first entry in the ALT column, and so on. The forward slash (/) indicates that no phasing information is available. • **NL**—Noise level; an estimate of base calling noise at this position. • **SB**—Strand bias at this position. Larger negative values indicate less bias; values near 0 indicate more bias. • **VF**—Variant frequency; the percentage of reads supporting the alternate allele.
SAMPLE	The sample column gives the values specified in the FORMAT column.

Per-Library and Merged VCF Files

The TruSight Tumor 15 workflow generates 2 sets of variant call files:

- Per-library VCF and gVCF files that are written to the Libraries folder
- Merged VCF and merged gVCF files that are written to the Alignments folder

Root Run Folder

Alignment—Contains merged gVCF (*.sample.genome.vcf) files.

Libraries—Contains per-library VCF (*.vcf) files, and per-library gVCF (*.genome.vcf) files.

Per-Library VCF Files

Using the somatic variant caller, variants are called in the Mix A library and the Mix B library to produce an independent set of VCF files for each library. The set of per-library VCF files include both VCF and gVCF files.

Per-library VCF files use the following naming convention, where S# represents the order the sample is listed in the sample sheet:

- SampleName_S#.genome.vcf—Reports all sites within the region of interest for a single library
- SampleName_S#.vcf—Reports variants only for a single library

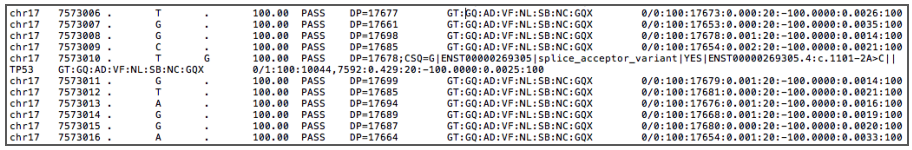
Genome VCF Files

Genome VCF (gVCF) files are VCF v4.1 files that follow a set of conventions for representing all sites within the genome in a reasonably compact format. The gVCF files generated in the TruSight Tumor 15 workflow include all sites within the region of interest specified in the manifest file.

For more information, see sites.google.com/site/gvcftools/home/about-gvcf.

The following example illustrates the convention for representing nonvariant and variant sites in a gVCF file.

Figure 1 Example gVCF File



```
chr17   7573006 .       T       .       100.00  PASS    DP=17677        GT:GQ:AD:VF:NL:SB:NC:GQX        0/0:100:17673:0.000:20:-100.0000:0.0026:100
chr17   7573007 .       G       .       100.00  PASS    DP=17661        GT:GQ:AD:VF:NL:SB:NC:GQX        0/0:100:17653:0.000:20:-100.0000:0.0035:100
chr17   7573008 .       G       .       100.00  PASS    DP=17698        GT:GQ:AD:VF:NL:SB:NC:GQX        0/0:100:17678:0.001:20:-100.0000:0.0014:100
chr17   7573009 .       C       .       100.00  PASS    DP=17685        GT:GQ:AD:VF:NL:SB:NC:GQX        0/0:100:17654:0.002:20:-100.0000:0.0021:100
chr17   7573010 .       T       G       100.00  PASS    DP=17678;CSQ=G|ENST00000269305|splice_acceptor_variant|YES|ENST00000269305.4:c.1101-2A>C||
TP53    GT:GQ:AD:VF:NL:SB:NC:GQX        0/1:100:10044,7592:0.429:20:-100.0000:0.0025:100
chr17   7573011 .       G       .       100.00  PASS    DP=17699        GT:GQ:AD:VF:NL:SB:NC:GQX        0/0:100:17679:0.001:20:-100.0000:0.0014:100
chr17   7573012 .       T       .       100.00  PASS    DP=17685        GT:GQ:AD:VF:NL:SB:NC:GQX        0/0:100:17681:0.000:20:-100.0000:0.0021:100
chr17   7573013 .       A       .       100.00  PASS    DP=17694        GT:GQ:AD:VF:NL:SB:NC:GQX        0/0:100:17676:0.001:20:-100.0000:0.0016:100
chr17   7573014 .       G       .       100.00  PASS    DP=17689        GT:GQ:AD:VF:NL:SB:NC:GQX        0/0:100:17668:0.001:20:-100.0000:0.0019:100
chr17   7573015 .       G       .       100.00  PASS    DP=17687        GT:GQ:AD:VF:NL:SB:NC:GQX        0/0:100:17680:0.000:20:-100.0000:0.0020:100
chr17   7573016 .       A       .       100.00  PASS    DP=17664        GT:GQ:AD:VF:NL:SB:NC:GQX        0/0:100:17654:0.001:20:-100.0000:0.0033:100
```

Use the OutputGenomeVCF sample sheet setting to generate gVCF files for the Enrichment workflow, PCR Amplicon workflow, and TruSeq Amplicon workflow. For more information, see the *MiSeq Sample Sheet Quick Reference Guide (document # 15028392)* or the workflow-specific reference guide.

Merged gVCF Files

MiSeq Reporter selects specific coordinates from the gVCF files generated for Mix A and Mix B to create a final merged VCF file for the sample.

Merged gVCF files are written to the Alignment folder. Merged gVCF files use the following naming convention:

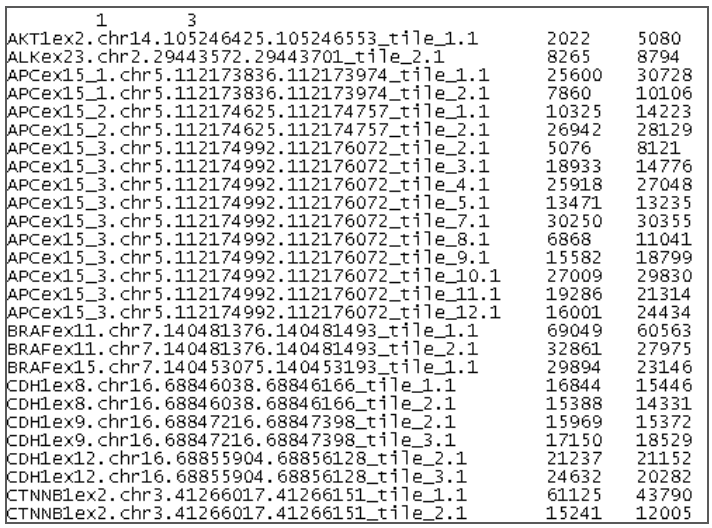
- SampleName.sample.genome.vcf—Reports for all sites

Amplicon Coverage File

One amplicon coverage file is generated for each manifest. The M# in the file name represents the manifest number as it is listed in the sample sheet.

Each file begins with a header row that contains the sample IDs associated with the manifest.

Figure 2 AmpliconCoverage_M1.tsv File



	1	3
AKT1ex2.chr14.105246425.105246553_tile_1.1	2022	5080
ALKex23.chr2.29443572.29443701_tile_2.1	8265	8794
APCex15_1.chr5.112173836.112173974_tile_1.1	25600	30728
APCex15_1.chr5.112173836.112173974_tile_2.1	7860	10106
APCex15_2.chr5.112174625.112174757_tile_1.1	10325	14223
APCex15_2.chr5.112174625.112174757_tile_2.1	26942	28129
APCex15_3.chr5.112174992.112176072_tile_2.1	5076	8121
APCex15_3.chr5.112174992.112176072_tile_3.1	18933	14776
APCex15_3.chr5.112174992.112176072_tile_4.1	25918	27048
APCex15_3.chr5.112174992.112176072_tile_5.1	13471	13235
APCex15_3.chr5.112174992.112176072_tile_7.1	30250	30355
APCex15_3.chr5.112174992.112176072_tile_8.1	6868	11041
APCex15_3.chr5.112174992.112176072_tile_9.1	15582	18799
APCex15_3.chr5.112174992.112176072_tile_10.1	27009	29830
APCex15_3.chr5.112174992.112176072_tile_11.1	19286	21314
APCex15_3.chr5.112174992.112176072_tile_12.1	16001	24434
BRAFex11.chr7.140481376.140481493_tile_1.1	69049	60563
BRAFex11.chr7.140481376.140481493_tile_2.1	32861	27975
BRAFex15.chr7.140453075.140453193_tile_1.1	29894	23146
CDH1ex8.chr16.68846038.68846166_tile_1.1	16844	15446
CDH1ex8.chr16.68846038.68846166_tile_2.1	15388	14331
CDH1ex9.chr16.68847216.68847398_tile_2.1	15969	15372
CDH1ex9.chr16.68847216.68847398_tile_3.1	17150	18529
CDH1ex12.chr16.68855904.68856128_tile_2.1	21237	21152
CDH1ex12.chr16.68855904.68856128_tile_3.1	24632	20282
CTNNB1ex2.chr3.41266017.41266151_tile_1.1	61125	43790
CTNNB1ex2.chr3.41266017.41266151_tile_2.1	15241	12005

Below the header rows are 3 columns:

- The first column is the Target ID as it is listed in the manifest.
- The second column is the coverage depth of reads passing filter.
- The third column is the total coverage depth.

Supplementary Output Files

The following output files provide supplementary information, or summarize run results and analysis errors. Although, these files are not required for assessing analysis results, they can be used for troubleshooting purposes.

File Name	Description
AnalysisLog.txt	Processing log that describes every step that occurred during analysis of the current run folder. This file does not contain error messages. Located in the root level of the run folder.
AnalysisError.txt	Processing log that lists any errors that occurred during analysis. This file is present only if errors occurred. Located in the root level of the run folder.
AmpliconRunStatistics.xml	Contains summary statistics specific to the run. Located in the root level of the run folder.
CompletedJobInfo.xml	Written after analysis is complete, contains information about the run, such as date, flow cell ID, software version, and other parameters. Located in the root level of the run folder.
DemultiplexSummaryF1L1.txt	Reports demultiplexing results in a table with 1 row per tile and 1 column per sample. Located in the root level of the run folder in Alignment.
ErrorsAndNoCallsByLaneTile ReadCycle.csv	A comma-separated values file that contains the percentage of errors and no-calls for each tile, read, and cycle. Located in the root level of the run folder in Alignment.
Mismatch.htm	Contains histograms of mismatches per cycle and no-calls per cycle for each tile. Located in the root level of the run folder in Alignment.
Summary.xml	Contains a summary of mismatch rates and other base calling results. Located in the root level of the run folder in Alignment.
Summary.htm	Contains a summary web page generated from Summary.xml. Located in the root level of the run folder in Alignment.

Manifest File Format

The TruSight Tumor 15 workflow requires 2 manifest files supplied by Illumina, 1 for Mix A and 1 for Mix B. The manifest files use a ***.txt** file format.

NOTE
There is no need to modify manifests files. The following manifest file description is provided for reference only.

The TruSight Tumor 15 manifest file contains a header section followed by 3 blocks of rows beginning with column headings: Probes, Targets, and Intervals.

- **Probes**—The Probes section has 1 entry for each pair of probes.

Column Heading	Description
Target ID	A unique identifier consisting of numbers and letters, and used as the display name of the amplicon.
ULSO Sequence	Sequence of the upstream primer, or Upstream Locus-Specific Oligo, which is sequenced during Read 2 of a paired-end run.
DLSO Sequence	Sequence of the downstream primer, or Downstream Locus-Specific Oligo. The reverse complement of this sequence forms the start of the first read. This sequence comes from the same strand as the ULSO sequence.
Chromosome	The chromosome of the amplicon that matches the reference chromosome.
Start Position, End Position	1-based chromosome endpoints of the entire amplicon including the sequence matching the probes. For example, if chromosome 1 started with **ACGTACACGT**, then a sequence with a Start Position of 2 and an End Position of 5 would be **CGTA**.

- **Targets**—The Targets section includes 1 entry for each amplicon amplified by a probe-pair. An expected off-target region is included in addition to the submitted genomic region.

Column Heading	Description
TargetA	Matches a target ID in the Probes section that corresponds to the ULSO probe sequence in Read 1.
TargetB	Matches a target ID in the Probes section that corresponds to the DLSO probe sequence in Read 2.
TargetNumber	Number of the targeted genomic region. The target region for a probe pair has index of 1. Any off-target amplicons have an index of 2, 3, and so on.
Chromosome	The chromosome of the amplicon that matches the reference chromosome.

Column Heading	Description
Start Position, End Position	1-based chromosome endpoints of the entire amplicon including the sequence matching the probes. For example, if chromosome 1 started with **ACGTACACGT**, then a sequence with a Start Position of 2 and an End Position of 5 would be **CGTA**.
Probe Strand	The strand of the amplicon indicated as a plus (+) or minus (-).
Sequence	Sequence of the amplified region between the ULSO and DLSO. This sequence comes from the forward strand if Probe Strand is plus (+) or from the reverse strand if Probe Strand is minus (–).

Revision History

Document	Date	Description of Change
Document # 1000000001205 v02	September 2016	Updated the merged gVCF file format from *.vcf to *.sample.genome.vcf. Changed references to the Filtered gVCF File Report to the Filtered Outputs Report because the filtered report is not a gVCF file. Updated the description of *.ant files to indicate that they are for import to VariantStudio software.
Document # 1000000001205 v01	October 2015	Updated the workflow name to TruSight Tumor 15.
Document # 1000000001205 v00	September 2015	Initial release.

Notes

Technical Assistance

For technical assistance, contact Illumina Technical Support.

Table 1 Illumina General Contact Information

Website	www.illumina.com
Email	techsupport@illumina.com

Table 2 Illumina Customer Support Telephone Numbers

Region	Contact Number	Region	Contact Number
North America	1.800.809.4566	Japan	0800.111.5011
Australia	1.800.775.688	Netherlands	0800.0223859
Austria	0800.296575	New Zealand	0800.451.650
Belgium	0800.81102	Norway	800.16836
China	400.635.9898	Singapore	1.800.579.2745
Denmark	80882346	Spain	900.812168
Finland	0800.918363	Sweden	020790181
France	0800.911850	Switzerland	0800.563118
Germany	0800.180.8994	Taiwan	00806651752
Hong Kong	800960230	United Kingdom	0800.917.0041
Ireland	1.800.812949	Other countries	+44.1799.534000
Italy	800.874909		

Safety data sheets (SDSs)—Available on the Illumina website at support.illumina.com/sds.html.

Product documentation—Available for download in PDF from the Illumina website. Go to support.illumina.com, select a product, then select **Documentation & Literature**.

Document # 1000000001205 v02

Illumina
5200 Illumina Way
San Diego, California 92122 U.S.A.
+1.800.809.ILMN (4566)
+1.858.202.4566 (outside North America)
techsupport@illumina.com
www.illumina.com