TruSight™ Oncology 500 ctDNA provides robust performance for detecting variants at low allele frequency

Optimized chemistry combined with high sequencing depth and powerful analysis software.

Introduction

Using next-generation sequencing (NGS) to detect rare variants, such as those found in circulating cell-free DNA (cfDNA), requires high sequencing depth and the ability to distinguish true low-frequency variants from noise. Several factors impact the accuracy of low-frequency variant calling when using TruSight Oncology 500 ctDNA. This technical note reviews variables such as error correction, depth of sequencing, and cfDNA input that influence analytical sensitivity and specificity.

UMI-based error correction

TruSight Oncology 500 ctDNA combines unique molecular identifiers (UMIs) with error correction software to enable error rate reduction from 0.5% to ≤ 0.007%.[1] Noise reduction is achieved by filtering false variants based on alignment of UMI barcodes and subsequent read collapsing (Figure 1). This process removes false positives from the collapsed reads, enabling more accurate calling at low variant allele frequency (VAF), ≤ 0.5%.

cfDNA input

In some instances, due to sample scarcity, users must create libraries with less than optimal cfDNA input amounts. Decreasing cfDNA input results in a lower number of genome equivalents and reduced library diversity. Illumina recommends using 30 ng of cfDNA quantified by using electrophoresis-based methods; not methods that measure total DNA, which represents approximately 9000 genome equivalents. Reducing the input to 10 ng reduces the genome equivalents present to ~ 3000. This has a marked impact on the ability to detect low-frequency variants accurately because a mutation present at 0.5% VAF would be represented by only 15 starting copies at 10 ng versus 45 copies from 30 ng input.

Sequencing depth

When looking for low-frequency variants, increasing the sequencing depth maximizes the probability of their detection. To ensure accurate variant calling, Illumina recommends sequencing to 35,000× minimum raw sequencing depth. With 30 ng DNA input and 400 M reads, this would result in a median exon coverage (MEC) of 2500×. MEC is the median number of read families spanning the exon region (Figure 2). This means that after read collapsing, the median collapsed fragment coverage for all bases in the panel is 2500×, and typically > 90% of reads have coverage of ≥ 1000×, which is recommended for calling variants at a frequency of 0.5%.

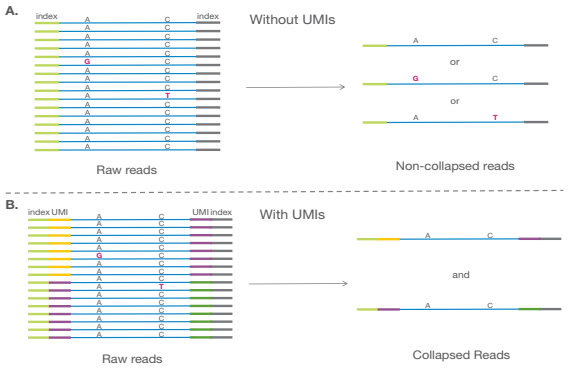

Figure 1: Error correction with DRAGEN™ TruSight Oncology 500 ctDNA Analysis Software—(A) 16 reads with two variants that could be true rare variants or introduced errors. Without error correction, it is impossible to distinguish between true variants and false positives. (B) Integration of UMIs enables recognition of multiple reads from the same starting molecule and collapses them into a single read. Each set of reads contains one error. After error correction, only one correct sequence remains.

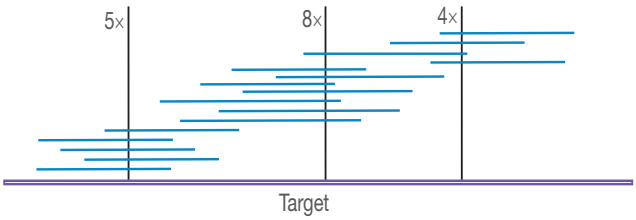

Figure 2: Median exon coverage—Horizontal lines represent sequencing reads for the targeted region. Vertical lines indicate which reads cover three specific bases. Sequencing depths and positions shown are for illustration only. Every base across a target is considered. The MEC for the region shown is 5×. Recommended MEC for 0.5% LOD following error correction is ≥ 1000×.

Analytical performance as a function of cfDNA input and sequencing depth

Assay performance is directly related to the analytical sensitivity and specificity across different limits of detection. Analytical sensitivity is defined as the ability to identify a variant correctly that is present (the true positive rate). Analytical specificity is defined as the ability not to call a variant when it is absent (the true negative). When sequencing depth decreases, the accuracy of low-frequency variant calling is also reduced. Increasing sequencing depth has a significant impact on the analytical sensitivity of the assay. For example, with 30 ng input, the analytical sensitivity at 0.2% VAF is 82.59% at 15,000× raw coverage depth.

Table 1: Expected analytical sensitivity and coverage depths for hotspot small variants with specificity of 99.99%

cfDNA input	Sequencing depth	Analytical sensitivity on hotspot at noted LOD								
		0.20%	0.30%	0.40%	0.50%	0.60%	0.70%	0.80%	0.90%	1.00%
10 ng	15,000×	34.97	55.25	70.67	81.39	88.47	92.99	95.79	97.50	98.53
	25,000×	36.33	56.89	72.22	82.69	89.47	93.71	96.29	97.84	98.76
	35,000×	36.97	57.65	72.93	83.27	89.91	94.03	96.51	97.99	98.85
	45,000×	37.26	57.99	73.25	83.53	90.10	94.16	96.61	98.05	98.89
30 ng	15,000×	82.59	95.11	98.73	99.69	99.93	99.98	100.00	100.00	100.00
	25,000×	86.25	96.66	99.25	99.84	99.97	99.99	100.00	100.00	100.00
	35,000×	87.67	97.20	99.42	99.88	99.98	100.00	100.00	100.00	100.00
	45,000×	88.42	97.47	99.49	99.90	99.98	100.00	100.00	100.00	100.00
50 ng	15,000×	94.12	99.15	99.89	99.99	100.00	100.00	100.00	100.00	100.00
	25,000×	96.85	99.68	99.97	100.00	100.00	100.00	100.00	100.00	100.00
	35,000×	97.68	99.81	99.99	100.00	100.00	100.00	100.00	100.00	100.00
	45,000×	98.06	99.85	99.99	100.00	100.00	100.00	100.00	100.00	100.00
70 ng	15,000×	97.26	99.75	99.98	100.00	100.00	100.00	100.00	100.00	100.00
	25,000×	99.11	99.96	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	35,000×	99.49	99.98	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	45,000×	99.64	99.99	100.00	100.00	100.00	100.00	100.00	100.00	100.00
100 ng	15,000×	98.54	99.91	99.99	100.00	100.00	100.00	100.00	100.00	100.00
	25,000×	99.79	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	35,000×	99.93	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	45,000×	99.96	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

The listed values for analytical sensitivity and specificity are simulated calculations and include UMI error correction. Analytical specificity assumes a fixed cutoff of two supporting fragments to call a variants. Cells shaded in green indicate highest sensitivity, in white indicate good sensitivity, and in pink are not recommended.

Table 2: Sequencing depth of 35,000× delivers the best sensitivity at 30 ng of cfDNA input for 0.5% LOD

cfDNA input	Analytical sensitivity on hotspot at noted LOD								
	0.20%	0.30%	0.40%	0.50%	0.60%	0.70%	0.80%	0.90%	1.00%
10 ng	36.97	57.65	72.93	83.27	89.91	94.03	96.51	97.99	98.85
30 ng	87.67	97.20	99.42	99.88	99.98	100.00	100.00	100.00	100.00
50 ng	97.68	99.81	99.99	100.00	100.00	100.00	100.00	100.00	100.00
70 ng	99.49	99.98	100.00	100.00	100.00	100.00	100.00	100.00	100.00
100 ng	99.93	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

The listed values for analytical sensitivity are expected for hotspot small variants with a specificity of 99.99% (theoretical data). Cells shaded in green indicate highest sensitivity, in white indicate good sensitivity, and in pink are not recommended.

Increasing raw depth to 45,000× raises the analytical sensitivity to 88.42%. Similarly, at 0.5% LOD, the analytical sensitivity increases from 99.69% at 15,000× raw depth to 99.9% at 45,000× raw depth (Table 1). Sequencing depth can be increased by loading fewer samples per run. A tradeoff is lower throughput per run. Users can consult the provided simulated calculations as a guide to assess the potential outcomes when suboptimal DNA inputs or coverages are obtained (Table 1).

In contrast to analytical sensitivity, analytical specificity decreases with higher coverage depth because increasing the number of reads also increases the possibility of false positives. However, analytical specificity decreases at a significantly lower rate and scale than the associated increase in analytical sensitivity. For example, for 30 ng input at 0.2% LOD, increasing the raw coverage from 10,000× to 40,000× raises analytical sensitivity from 54.77% to 80.02%, while analytical specificity only drops from 99.99% to 99.98%. For TruSight Oncology 500 ctDNA, the recommended minimum depth of sequencing is 35,000× in order to deliver a theoretical sensitivity above 95% (99.88%) for a limit of detection of 0.5% for hotspot variants (Table 2).

Summary

Accuracy is an important consideration with variant detection. Error correction methods, such as those used by the TruSight Oncology 500 ctDNA Analysis Software, can help to remove noise, thereby increasing the accuracy of variant calling at low VAF. Analytical sensitivity and analytical specificity vary according to factors such as cfDNA input and sequencing depth. Use this technical note as a guide to consider inputs that do not meet recommended guidelines when making decisions regarding experimental designs.

Reference

1. Illumina (2019) TruSight Oncology 500 ctDNA Data Sheet. (www.illumina.com/content/dam/illumina-marketing/documents/products/datasheets/trusight-oncology-500-ctdna-data-sheet-1170-2019-006.pdf). Accessed January 23, 2020.

Illumina • 1.800.809.4566 toll-free (US) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com