Replacing External Reference Materials with Internal Standards for Next Generation Sequencing

Sophie J. Deharvengt[1], Donald C. Green[1], Sarah B. Spracklin[1], Brad Austermiller[2], Tom Morrison[2], Gregory J. Tsongalis[1], Parth S. Shah[1].

1. Laboratory for Clinical Genomics and Advanced Technology, Department of Pathology and Laboratory Medicine, Geisel School of Medicine at Dartmouth, Dartmouth Hitchcock Medical Center, Lebanon, NH
2. AccuGenomics, Wilmington, NC

INTRODUCTION

- Identification of actionable mutations in circulating tumor DNA (ctDNA) enables gene-targeted therapy for solid tumors based on a simple blood test.
- Liquid biopsy comes with technical challenges:
  - ➢ The use of VAF to describe variants is not standardized due to the >100-fold variability of normal cell-free DNA (cfDNA) in liquid biopsy samples.
  - ➢ The deep sequencing required to detect low variant levels can reduce sample throughput.
- We validated **ctDNA-Seq,** which allows cfDNA tumor profiling from blood in our laboratory enabling sensitive detection of genomic variants.
- SNAQ™-SEQ spike-in standards are highly multiplexed mixtures of synthetic internal standard (IS) constructs to targeted regions of highest clinical importance. They are customizable and formulated to biochemically mimic the form and function of cfDNA, acting as a per sample limit control and to support variant reporting standardized to plasma concentration.

Aim: Testing the IS control as quality metrics for variants in 11 exonic regions in the context of cfDNA testing.

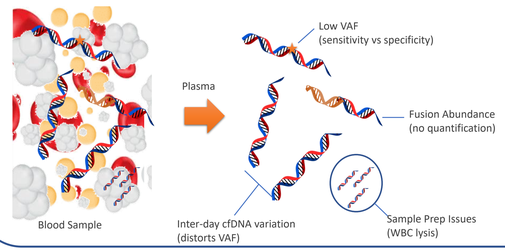
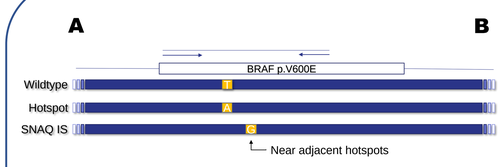


Figure 1: SNAQ™-SEQ Internal Standard (IS). SNAQ™-SEQ IS address issues with NGS testing of liquid biopsies: Variable normal background and low VAF targets.



KRAS	p.G12D/C
EGFR	p.L858R
EGFR	p.T790M
BRAF	p.V600E
AKT1	p.E17K
ERBB2	p.A775_G776insYVMA
KIT	p.D816V
NRAS	p.Q61R
GNAS	p.R201H
BRCA2	p.G3076E/V
ESR1	p.D538G; p.Y537S

Figure 2: SNAQ™-SEQ Internal Standard (IS).
A. SNAQ™ IS are synthetic constructs, each with unique variants positioned near known hotspot positions. The IS variants can be non-synonymous like those listed in 2B, or synonymous, requiring their addition to the pipeline's whitelist.
B. The Accukit CRC/NSCLC IS was pre-fragmented dsDNA to a mean length of 170±30 nucleotides to the 11 exonic regions containing the listed variants.

METHODS

Workflow

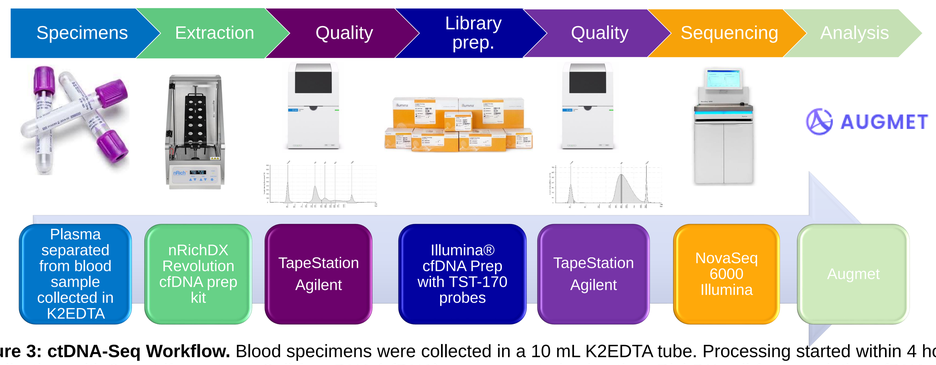


Figure 3: ctDNA-Seq Workflow. Blood specimens were collected in a 10 mL K2EDTA tube. Processing started within 4 hours of collection, with 2 centrifuge steps. Cell-free DNA (cfDNA) were extracted using the nRichDX cell-free revolution cfDNA prep kit. The Accukit CRC/NSCLC IS was spiked just before the library preparation. Seraseq® ctDNA complete material and human plasma were used. The libraries were prepared using 10 to 40ng of cfDNA, following the manufacturer's protocol. The libraries were qualified and quantified using the TapeStation, then sequenced on the NovaSeq6000. Dry lab informatics, including sequence alignment to reference genome (hg19) annotation, and interpretation was performed using AUGMET.

Filtering

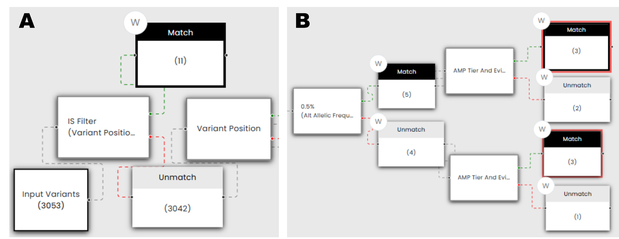


Figure 4: Filtering Workflow. Raw sequencing data and patient information was uploaded into the AUGMET platform, where it is processed and assessed for quality. The input variants are first filtered for the specific IS variants (A). The unmatched variants go through our custom filtering chain (B). Once the sequencing metrics pass, variants in the red highlighted filter chain nodes (TERT, VAF 0.5% and AMP Tier 1A-2D, VAF 0.5-0.3% AMP Tier 1A-2D) will be reviewed.

RESULTS

IS Effect on Variant Calling

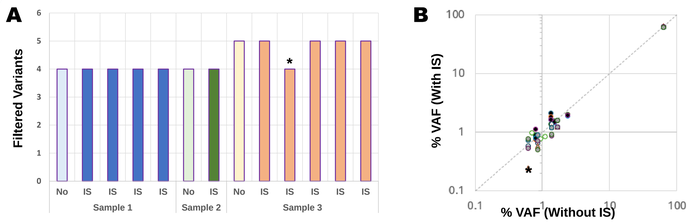


Figure 5: **Effect of IS spiked-in on Variant Calling.** Three different clinical samples were tested in different replicates with and without the IS control. **A.** The graphic shows the number of reportable pathogenic variants per sample. **B.** The plot shows the VAF of these samples without IS (black, white, grey filled symbols, x-axis), and with IS spike-ins with replicates.
In one sample (*), not all the IS variants were detected. After review, the quality metrics were suboptimal.

- ➢ The addition of the IS control at the first step of the library preparation did not interfere with variant detection or introduce false negative or false positive variants.
- ➢ There was no significant bias in the somatic VAFs when compared with non-IS spiked samples.
- ➢ A lower detection of the IS control can alert of bad quality metrics in this sample.

IS Spiked into Purified cfDNA

Synthetic randomly fragmented IS with unique non-synonymous variants were spiked into previously tested cfDNA patient samples at a low 0.5% VAF but at a constant input.

A

Sample	NRAS p.M67V	KRAS p.V8G	BRCA2 p.I3075V	AKT1 p.T21N	ERBB2 p.G776_V777insFPYT	GNAS p.S146T	KIT p.P812I	ESR1 p.N532H	BRAF p.R603G	EGFR p.L788R	EGFR p.I853F
1	0.42%	0.45%	0.35%	0.48%	0.25%	0.46%	0.31%	0.52%	0.99%	0.35%	0.33%
2	0.52%	0.50%	0.46%	1.20%	0.35%	0.23%	0.32%	0.22%	0.49%	0.56%	0.73%
3	0.49%	0.68%	0.60%	0.57%	0.20%	0.37%	0.34%	0.48%	0.39%	0.65%	0.69%
4	0.27%	0.35%	0.42%	0.86%	0.55%	0.64%	0.35%	0.64%	0.29%	0.28%	0.71%
5	0.53%	0.49%	0.21%	0.92%	0.34%	0.43%	0.44%	0.65%	0.42%	0.39%	0.66%
6	0.52%	0.55%	0.34%	0.59%	0.44%	0.32%	0.36%	0.24%	0.45%	0.56%	0.58%
7	0.56%	0.39%	0.88%	0.28%	0.21%	0.38%	0.33%	0.39%	0.43%	0.67%	0.37%

B

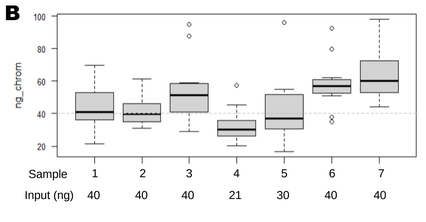


Figure 6: **IS Control Usage**. **A.** The table shows the detection of each of the 11 IS controls for 7 different samples. The non-detection can suggest a problem in the library preparation. **B.** By adding a constant input of IS controls, we can calculate the median of a sample's ng cfDNA input using the formula $(\frac{1}{VAF} - 1) * 88$ IS copies * 0.69 conversion rate * 0.0033 ng/genome.

- ➢ At a constant input of 88 copies per reaction, every IS variant was detected in all the samples at a VAF within stochastic sampling noise.
- ➢ Having a constant copy input allows us to estimate the capturable cfDNA levels, that could detect cfDNA quantification error.

CONCLUSIONS

- ➢ Synthetic randomly fragmented IS with unique non-synonymous variants were spiked into previously tested cfDNA patient samples at a low 0.5% VAF. The ability to detect the IS variants at the expected low VAF, provides a direct sensitivity QC for each sample. With no significant alteration of somatic variant detection, the replacement of an external positive run control with internal standards provides a better sensitivity QC than current NGS testing while also removing a costly external per run control sample. Further, IS VAF could be used to provide an independent QC of sample input by enabling direct calculation of ng of cfDNA input.