A Comprehensive, Targeted Next-Generation Sequencing Method That Rapidly and Accurately Detects Circulating Tumor DNA Variants at 0.1% Frequency in Plasma Samples

Abstract #5574

Jessica L Larson[1], Liangjing Chen[1], Lando Ringel[1], Blake Printy[1], Farol L Tomson[2], Yves Konigshofer[2], Dan Brudzewsky[2], Sarah Statt[1], Joseph Kaplan[1], Shobha Gokul[1], Jeffrey Shelton[1], Gary J Latham[1], and Brian C Haynes[1]

[1]Asuragen, Inc., Austin, TX; [2]SeraCare Life Sciences, Milford, MA

Summary

- A fast, efficient, and sensitive NGS panel approach that achieves specific and reliable detection of rare variants in liquid biopsy specimens has been developed.
- The workflow is both modular, which may spur the rapid development of new panel designs for clinical trial research and precision diagnostics, and amenable to a kitted format, which may improve access to liquid biopsy analyses.

Introduction

Mutation analysis of circulating tumor DNA (ctDNA) in blood-based liquid biopsies provides a minimally invasive approach to detect and monitor disease. Existing next-generation sequencing (NGS) liquid biopsy techniques have laborious and/or inefficient workflows, heuristic error-correction algorithms, and variable performance with clinical tumor-plasma samples. We present a method that combines an efficient wet-bench workflow with accurate dry-bench analytics to reduce costs and turnaround time and is relevant to clinical research and patient testing.

Materials and Methods

Input DNA molecules from Horizon cfDNA and Seraseq™ ctDNA v2 reference materials, cancer patient plasma samples, and healthy control and mutation-positive patient plasma admixtures (created taking into account mutation copy number) were uniquely tagged with a random molecular barcode (MBC), amplified in an efficient PCR protocol, and sequenced using a targeted panel covering >500 somatic mutation hotspots (Figure 1). Sequence-ready libraries were prepared from input DNA within ~9 hours. Data were analyzed with a custom bioinformatics pipeline to correct for multiple background errors. Variants were identified with a site-specific machine-learning model which effectively eliminated recurring non-biological aberrations that remained despite MBC-facilitated error-suppression. Variant calls in plasma were verified by Droplet Digital™ PCR (Bio-Rad), material permitting. Variants in corresponding FFPE samples from all cancer patient plasma samples were determined by the QuantideX® NGS DNA Hotspot 21 Kit†.

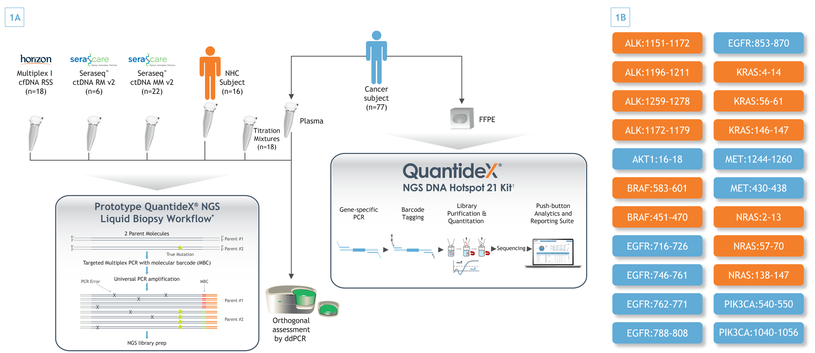

Figure 1. Overview of Approach. A) Six different cfDNA sample types were evaluated with a custom liquid biopsy NGS workflow: Horizon Multiplex I cfDNA Reference Standard Set (RSS), Seraseq ctDNA Reference Material (RM) v2, Seraseq ctDNA Mutation Mix (MM) v2, plasma from normal healthy controls (NHC), plasma from cancer patients, and patient plasma titration mixes. FFPE tissue from all cancer patients was also profiled with our QuantideX® NGS DNA Hotspot 21 Kit†. Plasma mutation calls were confirmed by BioRad ddPCR, material permitting. **B)** The Prototype QuantideX® NGS Liquid Biopsy Workflow panel covers over 500 DNA mutation hotspots. Amplicons range in size from 83 to 129 nt.

*Proof-of-concept data only. Future availability and performance cannot be ensured.

†This product is under development. Future availability and performance cannot be ensured.

Presented at AACR 2018

Results

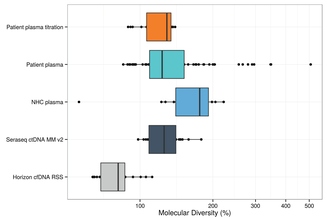

Figure 2. Seraseq v2 Reference Standards and Plasma Samples Have Comparable Post-Sequencing Molecular Diversity of Barcoded Molecules Relative to Mass Input by Qubit (330 cps/ng). Molecular diversity is plotted for all amplicons in all samples; the median number of unique MBCs across all amplicons is at least 100% of expectation based on input DNA quantities, indicating a highly efficient workflow.

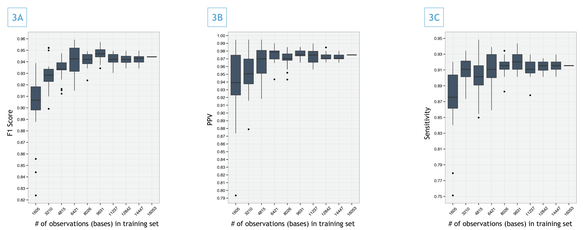

Figure 3. Variant Caller Accuracy Metrics are Consistently High Even with Small Training Data Cohorts at ≥0.1% VAF. A complete data cohort of all nucleotide base sequence observations ≥0.1% VAF at All COSMIC Sites with ≥1 non-reference ≥0.01% observation across all samples was generated. Data subsets consisting of 10% to 100% of the complete cohort (x-axis) were randomly generated 20 times and used to train a corresponding model. **A)** F1 score (harmonic mean of PPV and sensitivity), **B)** PPV, and **C)** Sensitivity for the complete cohort (all sites & samples) was calculated for each submodel (y-axis).

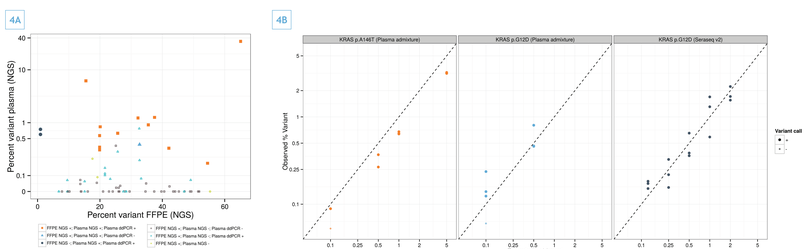

Figure 4. Analytical Concordance of A) FFPE NGS and Patient Plasma NGS and ddPCR, and B) Plasma Titration Mixes and Seraseq v2 Samples with Expected KRAS Variants. A) Somatic mutations were detected in 15/77 cancer cases by plasma NGS with a machine-learning based variant caller (orange and blue), including 2 variants present at trace levels in tissue (dark blue). There was a high level of concordance for positive calls by NGS and ddPCR in plasma; 14/15 mutations positive by NGS (and above its limit of detection) were also positive by ddPCR (≥3 mutant droplets, ≥10,000 accepted droplets, VAF ≥0.1%). **B)** In titration mixes, the observed and expected mutation frequencies were highly correlated (Spearman ρ = 0.930, p-value = 2.363x 10^{-8}). Only two of these variants were false negatives by our variant caller; both were observed below 0.1%. The observed KRAS p.G12D VAF in the Seraseq v2 samples was also correlated with its expected frequency (Spearman ρ = 0.938, p-value = 2.311x 10^{-7}).

	# ≥0.5% libraries	≥0.5% T+	≥0.5% F-	≥0.5% F+	≥0.5% Sn	≥0.5% PPV	# ≥0.1% libraries	≥0.1% T+	≥0.1% F-	≥0.1% F+	≥0.1% Sn	≥0.1% PPV
Horizon cfDNA RSS	10	60	0	0	100%	100%						
Seraseq™ ctDNA RM & MM v2	9	69	3	0	95.8%	100%	15	120	0	3	100%	97.6%
Patient plasma & titration mixes	26	24	2	1	92.3%	96.0%	41	31	11	3	73.8%	91.2%
Overall	45	153	5	1	96.8%	99.4%	56	151	11	6	93.2%	96.2%

Table 1. Performance Metrics for Machine-Learning Based Variant Caller at Expected VAFs of ≥0.5% (blue) and ≥0.1% (orange) in Horizon cfDNA RSS, Seraseq ctDNA RM & MM v2, and Patient Plasma and Titration Mix Libraries. The ≥0.5% overall metrics for the variant caller are calculated with the Horizon, Seraseq, and plasma samples with expected variants at ≥0.5%; the ≥0.1% overall metrics are calculated based on Seraseq and plasma samples with expected variants at ≥0.1% (including those with variants ≥0.5%). For patient plasma samples, a true variant expected at ≥0.1% was defined as having at least two of three conditions satisfied: (1) positive call & VAF ≥5% in FFPE, (2) positive call (see Figure 4) & VAF ≥0.1% in plasma ddPCR, and (3) VAF ≥0.1% in plasma NGS.

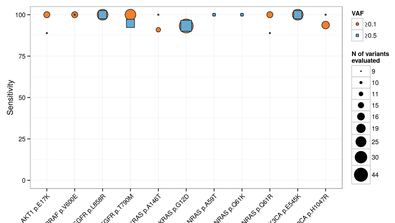

Figure 5. Analytic Sensitivity for All Expected Variants with >2 Observations Across All Sample Types at Expected VAFs of ≥0.5% (blue) and ≥0.1% (orange). For these variants ≥90% sensitivity at ≥0.1% VAF was achieved. The model failed to accurately call 5/8 variants with ≤2 observations ≥0.1% (Sn ≤ 50%); an additional 3 variants with ≤2 observations ≥0.1% and 6 with ≤2 observations ≥0.5% were identified with 100% sensitivity.

Conclusions

- This targeted NGS method is able to distinguish a low frequency ctDNA signal from background noise in plasma cell-free DNA using a streamlined PCR-based workflow and purpose-built bioinformatics pipeline.
- In both reference ctDNA standards and patient plasma samples and admixtures, mutations down to 0.1% AF were correctly identified while maintaining a low false-positive rate (sensitivity = 93.2% and PPV = 96.2%)*.
- This technology may provide an easy-to-use, high-performance, and adaptable NGS diagnostic framework for disease detection and therapeutic intervention monitoring.

seracare™ Passion. Innovation. Precision.

Asuragen®
www.asuragen.com