Supporting Neoantigen Discovery and Monitoring in Plasma Through Analytical Validation of a Deep Augmented Content Enhanced (ACE) Exome

#5385

Ravi Alla, Robin Li, Sean Michael Boyle, Shujun Luo, Rena McClory, John West, and Richard Chen
Personalis, Inc. | 1330 O'Brien Dr., Menlo Park, CA 94025

Contact: ravi.alla@personalis.com

Introduction

Neoantigens are increasingly critical in immuno-oncology as a therapeutic target for neoantigen-based personalized cancer vaccines and as a potential biomarker for immunotherapy response. An important step in identifying neoantigens is comprehensive exome and transcriptome sequencing of a tumor biopsy sample and the matched normal to enable identification of putative neoantigens derived from mutations in any gene in the genome.

However, as tumor biopsy samples cannot always be obtained, and because tumor heterogeneity can result in an incomplete set of neoantigens from a single biopsy, we developed our Accuracy and Content Enhanced (ACE) circulating tumor DNA (ctDNA) Exome to (1) identify neoantigens in cell free DNA (cfDNA) as a complement to tumor biopsy derived neoantigens and (2) track neoantigens in the cfDNA post immuno-therapy treatment. Our ACE ctDNA Exome covers ~ 20,000 genes as neoantigens can occur in any gene across the genome. This is in contrast to most current tumor cfDNA tests which are designed to assess a small number of variants or genes (often <100), and as such will miss many putative neoantigens. Our ACE ctDNA Exome is performed at very high sequencing depth to accurately identify lower allele frequency variants that are candidate neoantigens.

Here we want to assess the limit of detection (LOD) of our ACE ctDNA Exome assay for small nucleotide variants (SNV) and demonstrate the utility of the assay for both monitoring and de-novo identification of neoantigens directly from cfDNA.

Methods

Seracare ctDNA reference standards

We first wanted to estimate ACE ctDNA exome assay's ability to detect clinically relevant SNVs. To do this we used Seracare ctDNA Standards that harbor 25 SNVs across 21 cancer relevant genes (Table 1) at allele frequencies (AF) of 2%, 1% and 0.5%. To mimic plasma samples with varying levels of cfDNA, we used either 25ng, 50ng or 100ng of standards as input material for our assay. We performed exome enrichment and sequencing using our ACE platform. Our assay and analyses do not include any UMI or duplex sequencing steps. We sequenced to an average depth of 1000X across the whole exome. We then analyzed data using our somatic DNA variant calling pipeline using the cell line GM24385 (background for the seracare standards) as the matched normal sample. We intersected the identified somatic variants with the 25 Seracare reference variants to obtain sensitivity at various AFs and input cfDNA amounts.

Allele Frequency dilution using normal donor plasma

The Seracare reference standards only have 25 gold SNVs and are not designed to be used for specificity analysis. To address these issues, we designed a study to generate a larger set of gold SNVs and to assess specificity in addition to sensitivity of our ACE ctDNA exome assay. To do this, we used cfDNA from two healthy donors (D1 & D2) and gold set SNVs were identified by performing somatic variant calling using D1 as tumor and D2 as matched normal. Gold set variants are defined as heterozygous (~90% of identified variants) private germline variants in D1. We then constructed two AF dilutions where varying amounts of cfDNA from D1 were spiked into background cfDNA from D2 (Figure 2). We performed ACE ctDNA Exome assay and sequenced to an depth of ~1200X on these mixes. We then performed somatic variant calling in these D1+D2 dilution samples using D2 as matched normal. These somatic variants were then intersected with gold set variants identified above to generate sensitivity and specificity statistics.

SNV detection assessment Colorectal Cancer Tumor and Plasma

Finally, we obtained tumor (FFPE), matched normal (PBMC) and matched plasma (double spun within 4 hrs of collection in EDTA tubes) from 8 late stage (stages III/IV) colorectal cancer (CRC) cases. We chose late stage CRC samples because of high ctDNA shedding in plasma. Our goal with these samples was to identify somatic variants that are shared between the tumor and cfDNA tissues. We also assessed "targeted interrogation" of tumor variants in plasma as a monitoring approach. To do this, we interrogated plasma pileups to "look" for evidence of tumor variants. To reduce sequencing biases, errors and cfDNA related systematic mutations, we constructed a panel of normal from 20 healthy plasma donors (PON20). This PON20 data was used to "polish" or correct the pileup level evidence seen in the CRC plasma samples. This "polishing" has been shown to greatly decrease false positive calls.

Gene ID	c.HGVS	p.HGVS	Gene ID	c.HGVS	p.HGVS
AKT1	c.49G>A	p.E17K	IDH1	c.394C>T	p.R132C
APC	c.4348C>T	p.R1450*	JAK2	c.1849G>T	p.V617F
BRAF	c.1799T>A	p.V600E	KIT	c.2447A>T	p.D816V
CTNNB1	c.121A>G	p.T41A	KRAS	c.35G>A	p.G12D
EGFR	c.2573T>G	p.L858R	MPL	c.1544G>T	p.W515L
EGFR	c.2369C>T	p.T790M	NRAS	c.182A>G	p.Q61R
FGFR3	c.746C>G	p.S249C	PDGFRA	c.2525A>T	p.D842V
FLT3	c.2503G>T	p.D835Y	PIK3CA	c.1633G>A	p.E545K
FOXL2	c.402C>G	p.C134W	PIK3CA	c.3140A>G	p.H1047R
GNA11	c.626A>T	p.Q209L	RET	c.2753T>C	p.M918T
GNAQ	c.626A>C	p.Q209P	TP53	c.524G>A	p.R175H
GNAS	c.601C>T	p.R201C	TP53	c.818G>A	p.R273H
TP53	c.743G>A	p.R248Q			

Table 1: Seracare ctDNA Standards, SNVs only

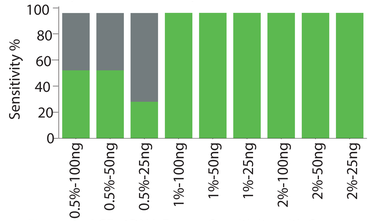
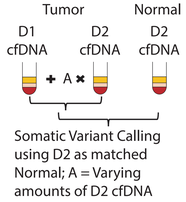


Figure 1: ACE ctDNA Exome Sensitivity with Seracare ctDNA Standards

Results

LOD Assessment using ctDNA reference standards

Our analysis of somatic variants from the Seracare standards revealed that the ACE ctDNA exome was successful in accurately identifying 24/25 (96%) gold set SNVs in 1% and 2% AF samples at 100ng, 50ng and 25ng starting DNA amounts (Figure 1). At 0.5% AF we see a considerable loss in sensitivity for gold set variants, using our somatic variant calling pipeline (Figure 1 green bars). By interrogating the pileups for the 0.5% AF samples and requiring >=3 reads of evidence supporting the gold set variants, we get 96% sensitivity (Fig 1 grey bars).



AF of Het Vars	D1:D2 mix Ratio
50% (Gold SNVs)	1:0
1.25%	1:39
0.625%	1:79

Figure 2: cfDNA Dilution scheme for Donors in LOD Study

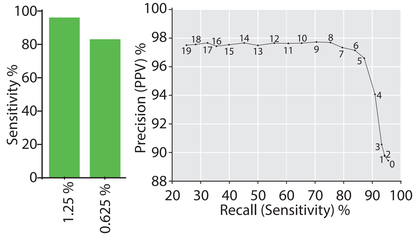


Figure 3: Sensitivity (L) and Precison-Recall curve (R) for AF mixes, nums on P-R curve indicate read evidence

LOD Assessment using normal donor plasma

We found ~5200 heterozygous variants that are private mutations in D1 compared to D2. We designated these as "Gold" variants. Our somatic variant calling assay had a sensitivity of 95.5% identifying gold variants at AF 1.25% and 82.5% at AF 0.625% (Figure 3L). To assess specificity, we used a PON20 polished read evidence cutoff for all somatic variants called in the mixes. Figure 3R shows precision recall curve for the 1.25% mix. Each point (and adjacent number) on the plot refers to the minimum number of polished reads required to confirm a plasma variant. For a variant monitoring case, where sensitivity is paramount, a more lenient cutoff of 2 reads of evidence yields ~94% sensitivity with ~90% PPV. For a denovo variant identification scenario, precision is more important, so a more stringent cutoff of 5-6 reads of evidence yields ~97% precision with a ~85% sensitivity.

CRC tumor vs ctDNA SNV comparison

We analyzed 8 late stage CRC tumor samples and identified between ~120-900 somatic variants with AF >10%. We then queried the matched plasma from these same patients for the same variants. Requiring at least 2 polished reads supporting a variant in cfDNA, we observed between 43% and 90% of tumor variants in matched plasma (Figure 4, grey bars indicate variants unique to tumor). We also called somatic variants directly in plasma and set a cutoff of 5 polished reads of evidence. Even with such a stringent cutoff we observed variants that are unique to plasma (dashed circle Figure 5). From the normal LOD study we observed a PPV of ~97% with a 5 read cutoff, so the variants seen in plasma are very likely to be real variants. These observations show that a tumor biopsy does not full represent the mutational load of a cancer and there could be a cohort of plasma derived variants (some of which maybe neoantigens) that could be missed.

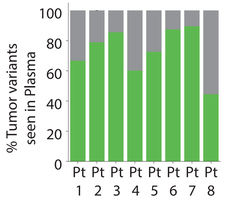


Figure 4: Variant concordance between Tumor and Plasma usinga cutoff of 2 reads of evidence (monitoring)

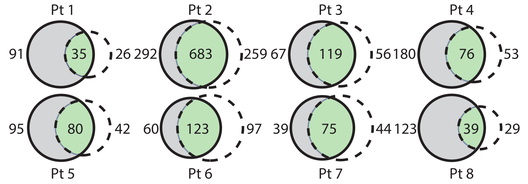


Figure 5: Variant concordance between Tumor (solid) and Plasma (dashed) using a cutoff of 5 reads of evidence (denovo plasma variant identification)

Conclusion

ACE ctDNA Exome has a sensitivity of > 95% to monitor and PPV of >97% to denovo detect SNVs in plasma, down to an AF of 1.25%. The assay's high sensitivity is important in cases when monitoring of tumor variants in plasma is desired. The high PPV is useful when identifying variants derived solely from plasma without a tumor biopsy. We believe the assay can be used a complement to the results from sequencing of the tumor biopsy alone.

Personalis®