TT015

Reference standards for the development and analytical validation of a next-generation lung cancer liquid biopsy assay

Signe Fransen[1], Fenella Raymond[1,2], Sally Grindstaff[1], Christina Zakarian[1], Thomas Persse[1], Pete Batzel[1], Kira Misura[1], and Chris Raymond[1,2]

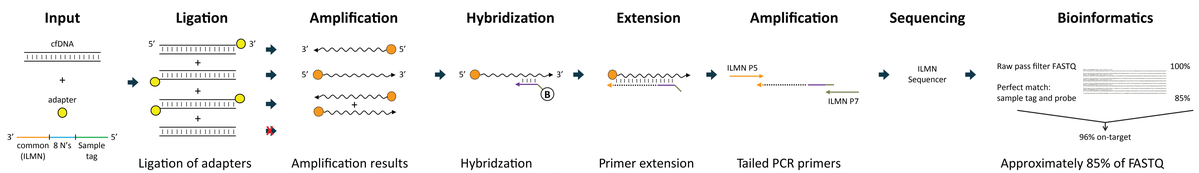
[1]Salish Bioscience, Inc., Seattle, WA; [2]Ripple Biosolutions, Seattle, WA

BACKGROUND

- We are developing a cfDNA-based next-generation sequencing assay for patients with non-small cell lung cancer (NSCLC) that has improved accuracy for the detection of low frequency variants
- Assay validation requires reference standards for optimization/familiarization (O&F) and analytical validation (AV) that conform to the intended use of the assay
- Here, we describe the results from our assay prototype and a collaboration to develop reference standards that enable O&F and AV

METHODS & RESULTS

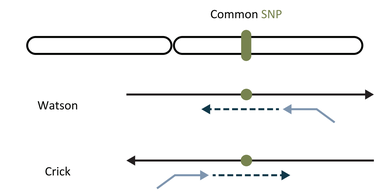
Figure 1. Salish Bioscience targeted hybrid capture platform



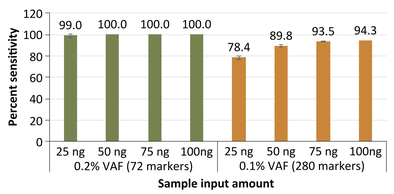
- Conventional cfDNA cloning methods rely on PCR, which requires the attachment of adapters to both strands of the cfDNA. This implies that only 16% to 25% of input cfDNA molecules are carried forward into sequencing.
- In the Salish method, every cfDNA strand ligated to an adapter is amplified and carried forward. This implies that at least one cfDNA strand from 64% to 75% of input cfDNA molecules is represented in the final sequencing data.

Figure 2. Sensitivity and performance of a prototype assay. A 1:500 blend of spike-in and background cfDNAs was genotyped at 352 known polymorphic SNP loci using a custom capture panel

A) Opposite strand probes were used to genotype common SNPs



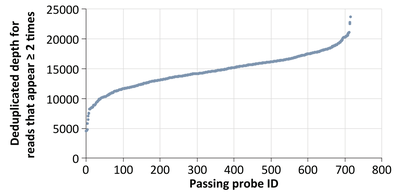
B) Sensitivity analysis of the prototype assay



Relative to the background sample, the spike-in sample had 72 opposite-allele, homozygous SNPs at a variant allele frequency (VAF) of 0.2% (green bars), and 280 heterozygous SNPs at a VAF of 0.1% (tan bars)

- The threshold for positive calls was that they a) must be observed on both chromosomal strands and b) must have a cumulative VAF >= 0.05%

C) Distribution of probe-by-probe coverage for 100ng input



- The initial genotyping panel had 716 probes covering 358 loci. Six of these probes failed to capture adequate sequence. The capture performance of the remaining 710 is shown.

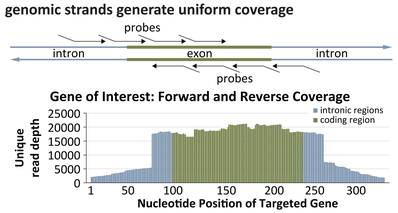
Figure 3. The validation of a highly sensitive panel for the detection of actionable NSCLC mutations required the development of custom reference samples

A) Salish NSCLC Targeted Panel Design

Gene	Region sequenced
ALK	fusions and kinase domain
BRAF	exon 11 - exon 15
EGFR	exon 18 - exon 21
ERBB2 (HER2)	exon 8, exon 17, exon 19, exon 20
KEAP1	entire CDS
KRAS	entire CDS
MET	intron 13, exon 14, intron 14
NTRK1	fusions and kinase domain
RET	fusions and kinase domain
ROS1	fusions and kinase domain
STK11	entire CDS
TP53	entire CDS

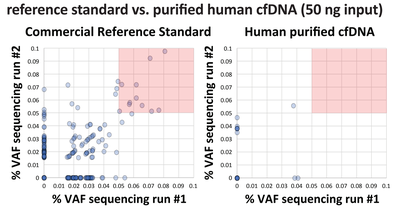
- The goal of this initial panel was to create a compact, cost-effective assay footprint that can identify tumor-derived DNA mutations that indicate patients may benefit from certain targeted therapies

B) Tiled probes that capture the forward and reverse genomic strands generate uniform coverage



- The graph shows the cumulative forward and reverse strand coverage for an EGFR exon from 100 ng of input cfDNA. Our goal is to achieve an LoD_{90} for sensitivity of 0.1% for actionable mutations.

C) Unanticipated background variants in a commercial reference standard vs. purified human cfDNA (50 ng input)

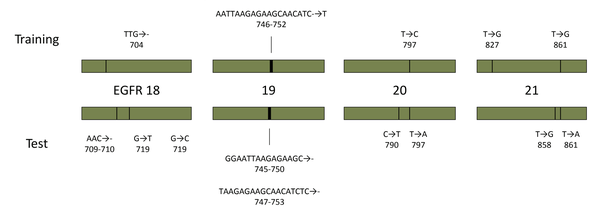


Variant calls observed in the 693 bp exon 2 region of the KEAP1 gene

- The high level of background variant calls in the commercial reference standard made it unsuitable for establishing the specificity of low VAF calls (red shaded area)

Figure 4. Training (O&F) and test (AV) reference standards were created

A) EGFR Variants in the training and test set

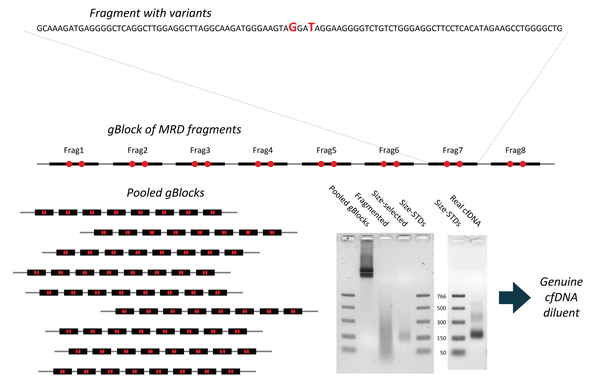


Most of the test and training set variants were identified as cancer mutations in clinOportal. For fusions, genomic breakpoints that predict known, canonical driver mutations were created *in silico*. Similar mutation types matched to each exon were chosen for both reference standards.

B) Reference mutations were "tagged" with an additional base change at ± 3 nt

- Phased base change "tags" unambiguously mark reference standard DNA and differentiate this from mutations anticipated in patient samples

C) Tagged reference mutations were introduced in the context of the surrounding reference sequence



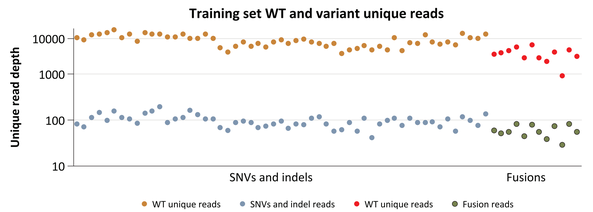
- These fragments were concatenated into several gBlocks that were subsequently pooled, fragmented, size selected, and introduced into purified cfDNA from a healthy human donor

Figure 5. The Salish NSCLC assay was used to show that all 67 mutations in the training reference set are present at uniform VAF

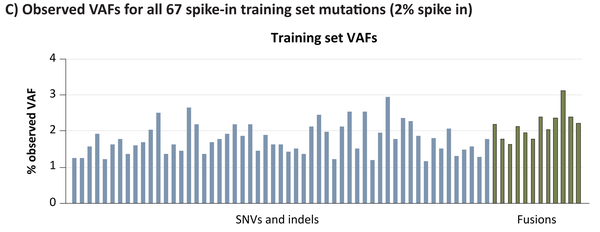
A) List of training set variants containing 34 SNVs, 9 single-base indels, 12 multi-base indels, and 12 fusions

Gene and exon	AA change	Nucleotide change
ALK exon 20	Y1078*	C->G
ALK exon 20	I1116M	C->G
ALK exon 21	G1125D	G->A
ALK exon 22	L1162R	T->G
ALK exon 23	F1196C	T->G
ALK exon 25	G1263R	G->A
ALK exon 27	E1321*	A->T
BRAF exon 11	G466E	G->A
BRAF exon 15	N581S	A->C
EGFR exon 18	L704del	TTG->-
EGFR exon 19	ELREATS746-752Vdelins	AATTAAGAGA AGCAACATC->T
EGFR exon 20	C797R	T->C
EGFR exon 21	Y827D	T->G
EGFR exon 21	L861R	T->G
ERBB2 exon 8	N319Y	A->T
ERBB2 exon 19	V750-	GTG -> - -
ERBB2 exon 19	I752insMS	- > GTC
ERBB2 exon 20	YVSR751-754dup	- -> TATGTCTCCCGC
ERBB2 exon 20	L823*	C->-
KEAP exon 2	P3Mfs	C -> -
KEAP exon 2	T60Tfs	- -> G
KEAP exon 2	H129Pfs	CACC -> -
KEAP exon 2	Q201Afs	CA -> -
KEAP exon 3	X214splice	G->C
KEAP exon 3	L267G fs	- -> GG
KEAP exon 3	Y426stop	- -> G
KEAP exon 3	C319Lfs	A -> C
KEAP exon 4	D448M fs	G -> -
KEAP exon 4	X511splice	G->A
KEAP exon 5	C518stop	T->A
KEAP exon 5	V568S	G -> -
KEAP exon 6	G570stop	- -> T
KEAP exon 6	stop625fs	TG -> -
KRAS exon 2	G12V	G->T
KRAS exon 3	Q61L	A->T
KRAS exon 4	A146P	G->C
KRAS exon 5	K175-K180 del	CTTCTTTTTCTTCTTTTT -> -
MET exon 14 region	intron 14 + exon 14 del	TTCTTTCTCTCTGTTTTAAGATCTG -> -
MET exon 14 region	X1010 splice	G -> T
STK11 exon 1	Q7stop	C->T
STK11 exon 1	E65Dfs	- -> C
STK11 exon 1	X97splice	G -> C
STK11 exon 2	Q112H fs	- -> T
STK11 exon 7	QHS305-307Lfs(splice)	AGCACAGG -> -
STK11 exon 8	W332stop	G->A
STK11 exon 8	G370D fs (splice)	GT -> AC
TP53 exon 2	W23stop	G->A
TP53 exon 4	X33splice	G->A
TP53 exon 4	W91stop	G->A
TP53 exon 5	C135F	G -> T
TP53 exon 5	H179R	A->G
TP53 exon 7	G245V	G->T
TP53 exon 8	R273C	C->T
TP53 exon 8	E298*	G->T
TP53 exon 9	Q331*	G->T
EML4 intron 6 - ALK intron 19		fusion
EML4 intron 13 - ALK intron 19		fusion
KIF5 exon15-RET exon12		fusion
KIF5 exon16-RET exon12		fusion
KIF5 exon22-RET exon12		fusion
KIF5 exon23-RET exon12		fusion
CDC74 exon6-ROS1 exon32		fusion
CDC74 exon6-ROS1 exon34		fusion
EZR exon10-ROS1 exon34		fusion
SLC34A2 exon12-ROS1 exon32		fusion
SLC34A2 exon12-ROS1 exon34		fusion
sdc4 exon2-ROS1 exon32		fusion

B) Unique read depth of WT and variant reads used for the NGS analysis of training mutation VAFs



C) Observed VAFs for all 67 spike-in training set mutations (2% spike in)



WT = wild type

CONCLUSIONS

- Promising results from the prototype assay prompted us to develop a precision medicine cfDNA assay for NSCLC
- We have collaborated to create NSCLC reference standards that consist of training and test samples for validation
- These will contain similar but independent mutations and are suitable for measuring sensitivity and specificity at low VAFs
- We believe this large number of mutations in different contexts lends statistical power to AV and anticipate these types of reference standards will be useful for liquid biopsy method validation

Presented at AMP 2022; November 1–5, 2022; Phoenix, AZ

RIPPLE BIOSOLUTIONS

SALISH BIOSCIENCE