ImmunoID NeXT Platform®

Analytical Validation Study Summary

Introduction

ImmunoID NeXT® is a universal cancer immunogenomics platform that can help transform the development of next-generation therapies by providing more comprehensive molecular data about each patient's cancer and immune response. Via the deep interrogation and analysis of ~20,000 genes in both DNA and RNA, ImmunoID NeXT consolidates multiple biomarker assays into one; providing a multidimensional view of the tumor and the tumor microenvironment (TME) from a single sample. The platform is an end-to-end solution for immuno- and precision oncology biomarker discovery applications, simultaneously enabling the ultra-sensitive detection of somatic small and structural variants as well as the advanced analysis of tumor escape mechanisms, immune repertoire profiles, human leukocyte antigen (HLA) typing and loss of heterozygosity, neoantigen load, tumor mutational burden (TMB), microsatellite instability (MSI), oncoviral presence, gene expression signatures, and more.

In order to validate the variant detection performance of the ImmunoID NeXT DNA and RNA sequencing assays, sensitivity and specificity were assessed across multiple variant types with varying allele frequencies (AFs). One major challenge in performing this type of study is the lack of established, gold-standard, real-world reference tumor samples containing known genomic alterations that can be used as "ground truth" to calculate the lower limit of detection (LOD) of the different variant types on an exome scale. Therefore, our validation approach incorporated the use of well-characterized, tumor-derived cell lines to generate reference samples that represent variants at AFs ranging from ≤5% to >95%.

Experimental Approach

We procured 21 cell lines and 11 matched normal cell lines, as shown in **Table 1**. The following subtypes were used for the validation of the various variant types:

- 11 tumor-derived cell lines and corresponding matched normal samples for small variants (i.e. single nucleotide variants (SNVs) and insertion/deletions (indels)) in DNA and RNA
- 10 tumor-derived cell lines for known gene fusion events to be detected in RNA

TABLE 1. Tumor-Derived Cell Lines Used for Validation of Small Variant and Gene Fusion Detection Performance

Small variants			Gene fusions	
Cell line	**Matched normal**	**Tumor type**	**Cell line**	**Tumor type**
COLO 829	COLO 829 BL	Melanoma	A673	Ewing's sarcoma
HCC1187	HCC1187 BL	Breast	HCC78	Lung
HCC1395	HCC1395 BL	Breast	K562	Leukemia
HCC1599	HCC1599 BL	Breast	Kasumi-1	AML
HCC1954	HCC1954 BL	Breast	LC-2ad	Lung
HCC2157	HCC2157 BL	Breast	LNCaP	Prostate
NCI-H128	NCI-BL128	Lung	MCF7	Breast
NCI-H1770	NCI-BL1770	Lung	NCI-H2228	Lung
NCI-H2009	NCI-BL2009	Lung	SJCRH30	Rhabdomyosarcoma
NCI-H2122	NCI-BL2122	Lung	THP-1	Acute myeloid leukemia
NCI-H2126	NCI-BL2126	Lung		

Genomic DNA and RNA were isolated from the cell lines. We then created an indexed genomic library using our proprietary library preparation protocol. These libraries were pooled and enriched using our patented ACE enrichment technology. The resulting enriched pools were then sequenced on Illumina NovaSeq™ next-generation sequencing instruments with paired-end reads measuring 150 base pairs in length.

To ensure the quality of the library preparation and sequencing, the following acceptance criteria were implemented:

TABLE 2. Sequencing Quality Metrics

DNA Quality Metrics	RNA Quality Metrics
>167M read clusters (300X mean coverage)	>100M read clusters
>90% reads mapped	>75% read pairs post ribosomal removal
<0.5 read pair duplication	>70% read pairs mapping to exons
>Q30 average base quality	
>0.5 capture specificity	

The gold set of variants was defined as follows:

Small Variants

The gold set is a subset of the somatic small variants in 11 matched tumor/normal cell lines published in the Cancer Cell Line Encyclopedia (CCLE), Broad Institute database (see **Table 1** for list of cell lines) and procured from American Type Culture Collection (ATCC). CCLE variants found in both whole exome sequencing and transcriptome sequencing were chosen. In addition, only variants found within the ImmunoID NeXT target region were selected. The final ImmunoID NeXT somatic small variant gold set consisted of 1,378 SNVs and 40 indels.

Fusions

Well-characterized gene fusion events were corroborated by two independent studies. Across 10 tumor-derived cell lines, this gold set consisted of 16 gene fusion events.

Validation of Small Variant Detection Performance

Overall Sensitivity and Specificity

We first sequenced 11 tumor and matched normal cell line pairs to establish the overall sensitivity and specificity values for small variants in DNA and RNA.

Limits of Detection (LOD)

To validate the LOD of small variant detection, we mixed three cancer cell lines with their matched normal to generate dilutions with tumor purity ranging from 5%-80%. This also generated variant AFs ranging from ≤5% to >95%. The cell lines and mixing scheme are shown in **Figure 1**. These mixtures were sequenced and the sensitivity (positive predictive agreement (PPA)) and specificity (positive predictive value (PPV)) were determined for each AF range in the DNA.

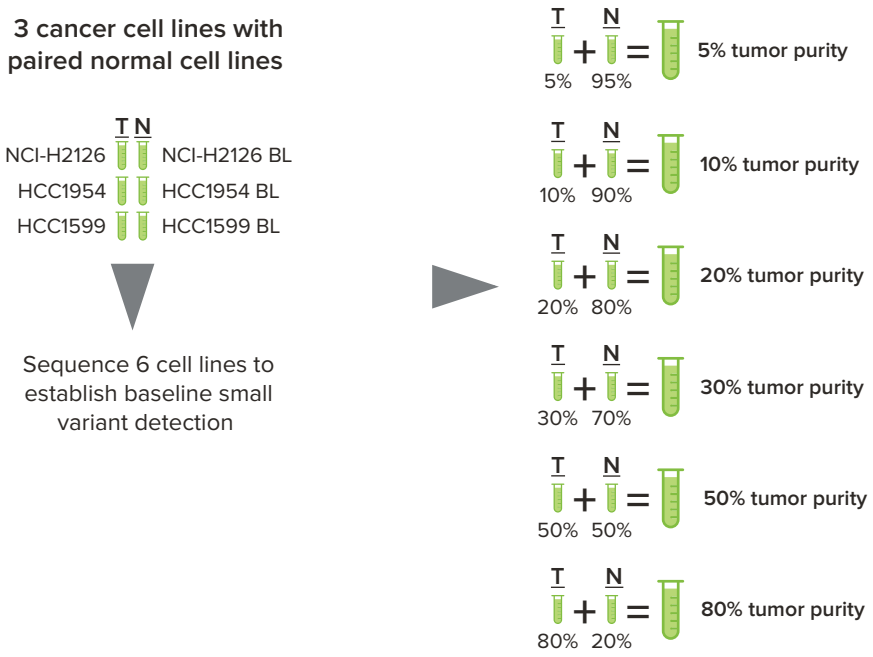

Figure 1: Dilution Scheme for LOD Experiment

Validation of Gene Fusion Detection Performance

Overall Sensitivity and Specificity

We sequenced 10 cancer cell lines to establish the overall sensitivity and specificity values for gene fusion events.

Validation of Gene Calculations

In order to analyze the experimental results, we employed the following statistical measurements:

- Analytical sensitivity is based on the number of true positives (TP) and false negatives (FN) and is calculated using the equation below:
  Analytical Sensitivity = TP*100/(TP+FN)
- Analytical specificity is based on the number of true negatives (TN) and false positives (FP) and is calculated using the equation below:
  Analytical Specificity = TN*100/(TN+FP)
- For the LOD calculations of sensitivity, we calculated the Positive Predictive Agreement (PPA) using the equation below:
  PPA = TP/(TP+FN)
- For the LOD calculations of specificity, we calculated the Positive Predictive Value (PPV) using the equation below:
  PPV = TP/(TP+FP)

Validation Results

DNA Small Variant Detection Performance

The analytical sensitivity for small variants was calculated as 99% for SNVs and 97% for indels (**Table 3**).

SNVs:

The number of SNVs found to be present in the LOD experiment for DNA were 8,139 at ≥10% MAF and 11,635 at ≥5% MAF. When we compared the SNVs detected by the ImmunoID NeXT exome to those known to be present, we found that the concordance was high (**Table 4**).

Indels

Indels were defined as mutations that result in a net change of 1 to 50 nucleotides in the tumor genome. The number of indels represented in the LOD experiment were 427 at ≥10% MAF and 618 at ≥5% MAF. When we compared the indels detected by the ImmunoID NeXT exome to those known to be present, we found that the concordance was also high (**Table 4**).

TABLE 3. Analytical Sensitivity for the ImmunoID NeXT Exome

	SNVs	Indels
Analytical sensitivity	99% (CI 99.0 - 100.0)	97% (CI 87.0 - 100.0)

TABLE 4. LOD Experiment Results for the ImmunoID NeXT Exome

	Sensitivity (PPA)		Specificity (PPV)	
MAF	≥5%	≥10%	≥5%	≥10%
Small variants	96% (CI 96.0 - 96.68)	98% (CI 98.16 - 98.7)	97% (CI 97.13 - 97.70)	99% (CI 98.46 - 98.95)
SNVs	97% (CI 96.48 - 97.12)	99% (CI 98.4 - 98.9)	98% (CI 97.63 - 98.16)	99% (CI 98.7 - 99.2)
Indels	88% (CI 84.85 - 90.19)	94% (CI 91.2 - 95.9)	89% (CI 85.91 - 91.01)	94% (91.31 - 96.03)

RNA Small Variant and Gene Fusion Detection Performance

Small variants

From a technical standpoint, the detection of somatic variants in RNA represents additional challenges above and beyond those of somatic variant detection in DNA. The widely varying expression levels of all genes (and particularly cancer-related genes), alternative splicing, and RNA editing are all potential incidences that make somatic variant calling in RNA uniquely challenging. While we are able to achieve very high sensitivity in RNA somatic variant detection, these challenges often lead to false positive calls and a reduction in performance relating to specificity.

To correct for this, we filter the small variants detected in the RNA with those detected in our highly-specific DNA somatic small variant detection pipeline. This results in the accurate detection of truly somatic, expressed small variants.

The analytical sensitivity for small variant detection was calculated to be 96% (**Table 5**).

Fusions

The set of ten cell lines used to determine the analytical sensitivity for fusion events contained a total of 16 unique fusion events. For this validation, the evidence threshold for fusion detection was ≥3 unique reads at a total assay sequencing yield of 200 million total reads.

The analytical sensitivity for gene fusion detection with the ImmunoID NeXT transcriptome was calculated to be 100% (**Table 5**). Analytical specificity for gene fusion detection was calculated to be >99%.

TABLE 5. Analytical Sensitivity for the ImmunoID NeXT Transcriptome

	Small Variants	Fusions
Analytical sensitivity	96% (CI 95.1 - 97.3)	96% (CI 95.1 - 97.3)

Conclusion

The results of this validation study demonstrate that the ImmunoID NeXT DNA and RNA sequencing assays achieve a greater level of performance, in terms of sensitivity and specificity to somatic small variants and gene fusions across all ~20,000 genes than other standard exome and transcriptome technologies that are currently available. Moreover, the sensitivity and specificity of the assays at an LOD as low as ≥5% is more in line with the level of performance that would be expected from a targeted (100-500-gene) next-generation sequencing (NGS) panel, an exceptional feat for a commercially-available broad, exome-scale platform.

Personalis, Inc.
U.S. Toll-free **1-855-GENOME4 (436-6634)**
Email **info@personalis.com**
Learn more at **www.personalis.com/products/immunoid-next/**

Personalis®

DOC-000560[D]