An Augmented, Extended Cancer Gene Panel Integrating DNA and RNA Sequencing
Comprehensively Assesses Neoplastic Genetic Lesions

The Complexity of Cancer Mutations

Through next-generation sequencing, we have come to understand cancer as a genomic disorder affecting key genes in cell growth and survival pathways. Cancer is a disease of driver mutations, often in multiple genes, which come in a plethora of forms. Cancer mutations include small variations, copy number alterations (CNAs), gene fusions, alternative splicing, changes in gene expression, and more. We must assess the different types of cancer mutations to accurately determine prognosis and make therapeutic decisions. Tumor purity and heterogeneity reduce the variant allele frequency (VAF) of mutations in cancer samples to boot, so we must also be robust against low representation alleles.

Assay Dictates Sensitivity to Different Variant Types

<table>
<thead>
<tr>
<th>Assay</th>
<th>Typical Sequencing Amount</th>
<th>Variant Types Detectable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exome</td>
<td>12Gb</td>
<td>SNVs, Indels, CNAs, Fusion, Gene Expression, Splice isoforms, Allele Expression, Low Detection Limit, Cross-validation</td>
</tr>
<tr>
<td>Transcriptome</td>
<td>14Gb</td>
<td>SNVs, Indels, CNAs, Fusion, Gene Expression, Splice isoforms, Allele Expression, Low Detection Limit, Cross-validation</td>
</tr>
<tr>
<td>Whole Genome</td>
<td>&gt;100Gb</td>
<td>SNVs, Indels, CNAs, Fusion, Gene Expression, Splice isoforms, Allele Expression, Low Detection Limit, Cross-validation</td>
</tr>
<tr>
<td>Cancer Panel (DNA+RNA)</td>
<td>15Gb</td>
<td>SNVs, Indels, CNAs, Fusion, Gene Expression, Splice isoforms, Allele Expression, Low Detection Limit, Cross-validation</td>
</tr>
</tbody>
</table>

Different assays are capable of detecting different variant types. In cancer there are a wide array of different variant types that can act as driver mutations. In order to detect all of the major variant types, one must assess both DNA and RNA. Moreover, given the mixed purity and heterogeneity of neoplastic tissues, a limit of detection down into the single digit percentile range is vital.

A comprehensive, high depth cancer panel applied to both DNA and RNA from the same sample is well equipped to assess all of these features, and has the unique advantage of internal cross-validation.

Designing an Optimized Cancer Panel

Standard enrichment strategies typically suffer in regions with a molecular structure that does not match the majority of the genome’s sequence, such as regions of high GC content.

We developed a novel technology called ACE (Accuracy and Content Enhanced) that fills in the gaps that would typically be seen with a standard enrichment assay. ACE uses a custom-designed probe set and includes multiple enrichment steps.

In the figure to left, the exons of five major cancer genes are plotted along the horizontal axis, and the mean read depth across 16 samples is plotted along the vertical axis. In blue is standard exome coverage depth at 12Gb of sequencing. In green is the augmented coverage depth added by ACE.

Using ACE enrichment strategies allows us to more completely capture all targeted regions.

Choose Genes to Include on the Panel

We took a data-driven approach to generating a list of all cancer-relevant genes by referring to numerous databases and the literature, as well as key cancer pathways. We also scrutinized the list for any key cancer genes that may have been missed manually. We arrived at a panel design that includes over 1,300 cancer-related genes.

Panels can be sequenced to very high depth. While a panel of over 1,300 genes is considered quite large in general, it is still of a size that allows for sequencing to mean depths of over 500x with high sensitivity for alleles with VAF<5%.

One final advantage to an enrichment-based panel design is the ability to target both RNA and DNA, facilitating integrated analysis.

Assessing Copy Number Alterations and Gene Expression

Copy number state in DNA can correlate with expression in the RNA.

Genes present at one copy in a cell line are expressed at a lower level than genes present at four copies.

Similarly, hyper-amplified genes demonstrate substantially higher expression than non-amplified genes in the same cell line. These findings are not unexpected. It is known that copy number and gene expression can correlate with one another, and hyper-amplification and complete deletion are well-described mutational mechanisms in cancer samples.

However, without looking at the RNA in conjunction with the DNA, there is no way to be certain that any given CNAs will affect RNA expression.

Gene Fusions Correlate with Copy Number Alteration Boundaries

Another set of variant types that benefit from integrated analysis are gene fusions from the RNA with CNA calls from the DNA.

Gene fusions are mediated by chromosomal translocations. In the circle plots to the left, fusion genes are represented as orange lines connecting two genes together. Cross-referencing our findings from the fusion analysis in RNA with the CNA calling from our gene panel, we are able to see clearly that the fusion events are happening right at positions where we see a shift in copy number, demonstrating that the fusions are not the only consequence of the translocation events in the DNA. Rather, these translocations also flank and are potentially facilitating copy number alterations.

© 2015 Personalis, Inc. All rights reserved. Personalis®, ACE Exome®, ACE Clinical Exome®, and ACE Genome® are registered trademarks or trademarks of Personalis, Inc., (“Personalis”) in the United States and/or other countries.

www.personalis.com | info@personalis.com | +1 855-436-6634 | +1 650-752-1300

Personalis®
Pioneering Genome-Guided Medicine

AGBT 2015