Whole Exome Sequencing

Focus on the detection of disease-related genomic mutations with transcriptional impact

Introduction

Whole exome sequencing (WES) or exome sequencing has emerged as a routine resequencing technique for all protein-coding genes in a genome (the exome). Despite many protocols that differ in detail, the principle remains the same. In a first step, the genomic DNA is reduced to only the coding regions of the genes, known as exons (about 180,000 exons for humans). In a second step, the resulting exonic DNA is sequenced using next generation sequencing (NGS) technology. By focusing on the protein coding regions of the DNA (approximately 1% of the human genome), the identification of the functionally relevant genetic variants that alter protein coding are detected at much lower costs compared to whole genome sequencing (WGS). Therefore, WES is widely used for cancer exome sequencing, causal variant studies (mono- and polygenic diseases) or translational research. The affordable sequencing depth makes exome sequencing well suited to several applications that are based on reliable variant calls such as clinical diagnostics and rare variant mapping in complex disorders.

Microsynth Competences and Services

Experimental Design

Unlike other NGS projects, the setup for a whole exome sequencing project seems rather straightforward. Even replicates are not mandatory in numerous settings. Nevertheless, a thorough consulting by our experts helps you to get the best out of your study.

DNA Isolation

You may either perform the DNA extraction yourself or outsource this critical step to Microsynth. We have long-standing experience in DNA and RNA isolation from demanding matrices including isolation from fixed samples and serum.

Enrichment & Sequencing

The Agilent SureSelect in-solution capture is applied for targeting the human exons. Key benefits of the Human Whole Exome V6 + COSMIC panel are the comparable deep and complete target coverage, the high "on target" coverage, inclusion of more exons on the hard-to-capture targets and additional coverage of cancer relevant targets with the COSMIC-baits. The RNA-driven DNA capture is conducted by RNA oligonucleotides (or "baits") that are biotinylated for easy capture onto streptavidin-labeled magnetic beads (see Figure 1). To perform the capture, genomic DNA is sheared and assembled into a library format. Size
selection is performed on the library prior to capture. Size-selected libraries are then incubated with the baits, and RNA bait-DNA hybrids are then “fished” out of the complex mixture by incubation with the magnetic beads. RNA baits are digested such that the remaining nucleic acid is the targeted DNA of interest. Captured DNA is amplified and the targeted samples are sequenced (2x75 bp), resulting in at least a 100-fold coverage per sample.

**Bioinformatics Analysis**

Our WES Service allows several entry points and processing grades for our deliverables (see Figure 2). The Whole Exome Sequencing analysis consists of reads mapping to the human reference genome, calling and annotation of single nucleotide variations as well as small insertions and deletions (see Figure 3). Results are compared to dbSNP databases (e.g. ClinVar-NCBI-NIH). Additionally, a comprehensive coverage analysis of all captured exons is provided.

**Example Results of the Whole Exome Sequencing Analysis**

**Table 1:** This is a cutout of the full table detailing read coverage for every target in the sequenced exome.

**Table 2:** This excerpt of a results table shows detected variations (alternative compared to the reference), their annotation and whether they and their impact are already known (e.g. as ClinVar - Variations).

<table>
<thead>
<tr>
<th>CHR</th>
<th>POS</th>
<th>Gene ID</th>
<th>Transcript ID</th>
<th>Ref</th>
<th>Alt</th>
<th>Qual</th>
<th>DB</th>
<th>FEATURE</th>
<th>CONDITION(S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>chr 4</td>
<td>110765379</td>
<td>LRIT3</td>
<td>NM_198506.2</td>
<td>C</td>
<td>A</td>
<td>1207.77</td>
<td>ref</td>
<td>exon</td>
<td>-</td>
</tr>
<tr>
<td>chr 4</td>
<td>110765932</td>
<td>LRIT3</td>
<td>NM_198506.2</td>
<td>C</td>
<td>A</td>
<td>1207.77</td>
<td>ref</td>
<td>CDS</td>
<td>-</td>
</tr>
<tr>
<td>chr 4</td>
<td>110773184</td>
<td>LRIT3</td>
<td>NM_198506.2</td>
<td>C</td>
<td>A</td>
<td>1207.77</td>
<td>ref</td>
<td>exon</td>
<td>-</td>
</tr>
<tr>
<td>chr 4</td>
<td>11078458</td>
<td>LRIT3</td>
<td>NM_198506.2</td>
<td>G</td>
<td>A</td>
<td>2680.77</td>
<td>ref</td>
<td>exon</td>
<td>-</td>
</tr>
<tr>
<td>chr 4</td>
<td>110790792</td>
<td>LRIT3</td>
<td>NM_198506.2</td>
<td>A</td>
<td>T</td>
<td>1691.77</td>
<td>ref</td>
<td>CDS</td>
<td>-</td>
</tr>
</tbody>
</table>

**References**


https://www.nature.com/articles/gim2015142

http://cancer.sanger.ac.uk/cosmic