GPDx
Genomic Precision Diagnostic

Igenomix®
WITH SCIENCE ON YOUR SIDE

Whole Exome Sequencing (WES)

1. Methodology

Whole Exome Sequencing (WES) is a Next Generation Sequencing (NGS) based approach used to identify genetic variants linked to a disorder. This methodology analyzes the part of an individual's DNA called exons; this region provides instructions to make proteins. Since most of the disease-causing variants are present in the exons, WES is thought to be an efficient technique to determine disease-causing variants that may lead to a particular disease.

Whole exome sequencing was performed on DNA extracted from the patient's blood, saliva, or tissue. Approximately 45 Mb of genomic sequence, corresponding to 99% of coding sequencies (Consensus Coding Sequence, CCDS) were enriched from fragmented DNA with probes that are designed against the human genome and are optimized for coverage of all medically relevant genes (Nextera DNA Flex Pre-Enrichment Library Prep, Illumina). The library generated was sequenced on the Novaseq 6000 Illumina platform. Raw sequencing data were processed by the Igenomix in-house bioinformatics pipeline (v1.0). In brief, the raw data was first demultiplexed to link molecular barcodes with the sample ID, followed by the trimming of adaptors and low-quality reads. The clean reads were mapped to the human genome reference and duplicated reads were removed before variant calling and annotation. Quality metrics were collected for each sample. Quality metrics cut-off for each sample is as follows: Q30 >80%, mean coverage depth >100x, and % bases with coverage above 20x >95%. Samples that do not pass the cut-off were resequenced.

Exome coverage

- Panel size: ~ 45Mb, corresponding to the 99% of the Consensus Coding Sequence (CCDS, RefSeq, GENCODE, ENSEMBL and ClinVar databases)
- Mean coverage: 157.96 (>100x)
- Highly uniform coverage of the entire exome (20,000 genes):
  - >97.0% of positions covered at >10x
  - >96.2% of positions covered at >20x
- Reads: paired-end 300 cycles (2x150 bp)

Enhanced coverage of clinically relevant regions

- 4,555 disease-associated genes (OMIM):
  - Mean coverage: 189 (>100x)
  - >97.0% of positions covered at >20x
- ClinVar Pathogenic and Likely pathogenic variants:
  - Mean coverage: 201.4 (>200x)
  - >99.4% of target regiones covered at >20x
- HGMD Pathogenic and Likely pathogenic variants:
  - Mean coverage: 214.0 (>200x)
  - >99.3% of target regiones covered at >20x

Variant types included in study

- NGS can detect with highly sensitivity and specificity:
  - SNVs
  - InDels
  - CNVs: ≥15 exons with a sensitivity >80%
    ≥0.5 Mb with a sensitivity >95%
- NGS cannot detect: homopolymer stretches, variants in pseudogene regions, gene fusions, balanced translocations, inversions, ploidy changes, uniparental disomy, and repeat expansion regions.

Variant interpretation was performed by clinical scientists. All variants with allele frequency less than 1% in human population database (ExAC, 1000G, gnomAD) were evaluated. Variants present in the coding region or in the conservative splice site, reported intronic and UTR variants were prioritized. Combined molecular biology, gene function annotation, genetic inheritance pattern of the disease (OMIM), information from disease databases (HGMD, ClinVar, ClinGen), and literature search (PubMed) were utilized to correlate the clinical phenotype with genetic variants. In addition to SNVs and small Indels, copy number variants (CNVs) are detected from targeted sequence data using the ExomeDepth (v1.1.15) method (PMID: 22942019). This algorithm detects rare CNVs based on comparison of the read-depths of the test data with the matched aggregate reference dataset. CNVs with an extension greater than 15 exons can be detected with a calculated sensitivity > 80 %, while CNVs above 0.5 Mb, a sensitivity value > 95% has been calculated. The validation of the analysis of CNVs has been carried out by in silico generation of different samples with positive events on autosomal chromosomes.

Pathogenicity of variants was determined using the recommendation from the American College of Medical Genetics (ACMG) (PMID: 25741868). Please note that the classification of variants can change over time. To determine if there are any changes to the classification of reported variants, please contact Igenomix.

The categories of variants set out by the ACMG are as follows: Pathogenic, Likely Pathogenic, Variant of Uncertain Significance, Likely Benign, Benign, and risk-associated SNP. Only Pathogenic, Likely Pathogenic, and Variants of Uncertain Significance (VUS) related to the clinical indication of the patient will be reported in this genetic test. In accordance to the ACMG guidelines for reporting secondary findings in clinical exome sequencing (PMID: 27854360), pathogenic and likely pathogenic variants in the following genes are reported if consent is indicated on the Test Request Form: ACTA2, ACTC1, APC, APOB, ATP7B, BMPR1A, BRCA1, BRCA2, CACNA1S, COL3A1, DSC2, DSG2, DSP, FBN1, GLA, KCNH2, KCNQ1, LDLR, LMNA, MEN1, MLH1, MSH2, MSH6, MUTYH, MYBPC3, MYH11, MYH7, MYL2, MYL3, NF2, OTC, PCSK9, PKP2, PMS2, PRKAG2, PTEN, RB1, RET, RYR1, RYR2, SCN5A, SDHAF2, SDHB, SDHC, SDHD, SMAD3, SMAD4, STK11, TGFBR1, TGFBR2, TMEM43, TNNI3, TNNT2, TP53, TPM1, TSC1, TSC2, VHL, and WT1.

Applications

WES is used in diagnosing or evaluating a genetic disorder where the results are expected to influence medical management and clinical outcomes of a patient or a family directly or indirectly. With the advent of technology, sequencing has become a routine process in clinical diagnosis. In situations where the clinical presentation is unclear and the condition in question is unknown, sequencing and analysis of a small number of genes at a time is costly and time-consuming process. This may further delay the diagnosis, which could have an impact on patient's quality of life.

WES is a cost-effective diagnostic solution that permits sequencing data from ~24,000 genes from a simple blood draw or other accepted biological samples. WES examines a wider range of genes and variants, which is especially worthwhile for couples looking to know if they are carriers of common recessive disorders and to diagnose genetic disease in an affected patient.

Limitations

The probes used for this test are designed to detect known genes in the human genome. Therefore, this test is unable to detect genes not defined by the NCBI reference genome GRCh37 or non-human genome sequences including viral sequences or non-nuclear DNA. In addition, due to the limitations of NGS technologies, the following variants cannot be readily detected: large deletions/duplications greater than 40 base pairs, copy number variations (unless applying specific analysis with their specific limitations for CNVs detection, as described above), homopolymer stretches, variants in pseudogene regions, gene fusions, balanced translocations, inversions, ploidy changes, uniparental disomy, and repeat expansion regions. Furthermore, variants present outside the exon (noncoding region) could be missed; these variants can affect gene activity and protein production that may lead to genetic disorders. This technique does not cover the entire exome, (the % of bases with coverage above 20x is approximately 97%). It may not be possible to resolve certain details about variants such as mosaicism, phasing, or mapping ambiguity. Analytical limitations may also occur due to the provided clinician information. Accurate and thorough clinical information of the patient(s) and family members is required as incomplete information may lead to false positive or negative results.

2. Sample requirements and TAT

The following sample types are accepted for Igenomix genetic tests. A thorough labelling of the tube with unique identifying information is suggested, incorrect labelling can lead to rejection of the sample. The minimum required information to identify and accept a sample is: Patient's full name, Date of birth, Gender, Medical Record Number, Clinical Indication, and Phenotype description (mainly through an updated clinical report).

Sample type	TAT	Container	Volume	Temperature
Peripheral Blood	3-4 weeks	EDTA tube	3mL	Room temperature
CVS	3-4 weeks	CVS sterile tube either transferred into a sterilized conical tube that contains (RPMI) 1640 media or into a saline solution with 1% antibiotic	300-500 mg of tissue obtained from routine CVS	Room temperature
Amnio	3-4 weeks	Sterilized conical tube sealed with parafilm	15-20 mL Amniotic fluid	Room temperature
Products of Conception	3-4 weeks	Tissue in sterile container in saline Cardiac or cord blood in Vacutainer	1 cm3 (sterile) fetal tissue and/or villi in tissue culture media or Preferred fetal tissue sample sites include buttocks or thigh. If fetal tissue is not available placental villi can be utilized	Room temperature
Extracted DNA	3-4 weeks	In a sealed eppendorf tube	A minimum 1 microgram of DNA at a concentration of 50-100 ng/microliters	Room temperature

*Maternal blood sample must be sent with all products of conception, CVS and Amnio samples