GenomeSeqDx: Clinical Genome Sequencing

Description:

GenomeSeqDx, or clinical genome sequencing (GS), can be used to identify the underlying molecular basis of a genetic disorder in an affected individual. The GenomeSeqDx test is different from other types of genetic diagnostic tests in terms of the number of genes that are sequenced simultaneously and the inclusion of non-coding regions of the genome. Genome sequencing can be used to identify the molecular basis of a genetic disorder in individuals:

- With a genetically heterogeneous disease, as pathogenic findings could be present in many different genes
- With a long list of differential diagnoses
- With an atypical presentation of a genetic disorder
- Who have exhausted other currently available genetic testing options

The GenomeSeqDx test includes both the protein-coding and non-coding regions of the human genome, allowing for the potential detection of characterized/pathogenic variants in regions that are not assessed by exome sequencing (ES). The protein-coding regions represent ~20,000 genes and account for approximately 2% of all human genetic material.[1] The non-coding regions include promoter, intronic, and untranslated regions. While much of the data generated from sequencing the genome is not well understood at this time, GS may provide more reliable coverage of the exonic regions.[2,3] GS has lower depth of coverage on average compared to ES, but more positions are covered to adequate depth for accurate variant calling than ES.

For GenomeSeqDx, an individual's genome sequence is compared to published reference sequences, other individuals from the affected individual's family and control individuals, and phenotype-driven reporting is performed using Human Phenotype Ontology and gene-phenotype associations. Additional resources such as GnomAD, OMIM, PubMed and ClinVar are used to evaluate genes and sequence changes of interest, which are then interpreted according to the American College of Medical Genetics and Genomics guidelines .[4,5] Genome testing is most effective when other family members, specifically both biological parents (if available), are included in the analysis of the affected individual's genome sequence.

The clinical sensitivity of genome sequencing depends, in part, on the proband's clinical phenotype. Studies have demonstrated that genome sequencing identifies a causal variant in approximately 20-68% of probands, with a higher yield for cases that specifically include other family members and have strict clinical inclusion criteria.[6-12]

Result Reporting:

GenomeSeqDx analysis is performed on the proband and parental samples, and/or additional relatives as needed, when submitted together for analysis. A single report will be issued for the affected individual in the family. A separate report will not be issued for unaffected parents or other family members who may also have submitted a specimen for the purpose of allowing better interpretation of the results from the affected individual. If additional reports are requested for other affected family members, additional fees will apply.

The GenomeSeqDx report issued for the affected individual in the family may contain variations in genes previously implicated in a human disease similar to the clinical presentation of the affected individual or in genes hypothesized to be related to the cause of the disease (candidate genes), based upon the function, tissue of expression, and phenotype of model organisms with alterations in the gene. Variants in candidate genes may also be reported based on internal data, such as observations of previous cases with similar phenotypes and types of variations in the same gene.

> **ACMG Secondary Findings:**
>
> The American College of Medical Genetics and Genomics (ACMG) recommends that secondary findings identified in a specific subset of genes associated with medically actionable, inherited disorders be reported for all probands undergoing exome or genome sequencing. Please refer to the latest version of the ACMG recommendations for reporting of secondary findings in clinical exome and genome sequencing for complete details of the genes and associated genetic disorders. Secondary findings will be included for all exome or genome sequencing reports, unless a family opts-out of receiving this information on the Informed Consent as part of the test requisition form. The status of any secondary finding(s) reported for the affected individual will be provided for all relatives included as part of the proband's test; **GeneDx does not conduct an independent evaluation of secondary findings in relatives.** Relatives have the ability to opt-out of receiving secondary findings. Secondary findings will be confirmed by an alternate test method when needed.

Test Methods:

An affected individual's clinical records and prior genetic testing results will be reviewed prior to analysis. Genomic DNA from the submitted specimen is sequenced with paired-end reads on an Illumina platform. Average mean sequencing coverage is >= 40x across the genome, with a minimum threshold of 30x for any individual sample. Bi-directional sequence reads are assembled and aligned to reference sequences based on NCBI RefSeq transcripts and human genome build GRCh37/UCSC hg19. A custom-developed analysis tool (XomeAnalyzer), is used to filter and analyze data for the identification of sequence variants and copy number variants.[13] Reported, clinically significant variants, are confirmed by an appropriate orthogonal method in the proband and, if submitted, in

selected relatives as necessary. Sequence variants are reported according to the Human Genome Variation Society (HGVS) guidelines. Copy number variants are reported based on the probe coordinates, the coordinates of the exons involved, or precise breakpoints when known. Reportable variants include pathogenic variants, likely pathogenic variants and variants of uncertain significance. Likely benign and benign variants, if present, are not routinely reported.

The analysis for the proband includes evaluation of variants that are identified to be *de novo* (when both parents submitted), compound heterozygous (when both parents submitted), homozygous, heterozygous and X-linked. In addition, analysis takes into consideration family structure, reported phenotype and provided clinical and/or differential diagnosis. This is a phenotype-driven test of a very large number of genes; therefore, reported results are focused on pathogenic and likely pathogenic variants in genes related to the clinical information provided.

Limitations:

The GenomeSeqDx test attempts to evaluate the complete coding and non-coding regions of the genome. However, it is not technically possible to uniquely resolve and align the entire genome at present due to homology and other structural complexities in some regions. It is anticipated that approximately 99% of the coding region of an affected individual's genome (i.e., the exome) will be assessed with the GenomeSeqDx test at 10x coverage.

The available scientific knowledge about the function of all genes in the human genome is incomplete at this time. It is possible that the GenomeSeqDx test may identify the presence of a genetic variant in the genomic sequence of an affected individual, but it will not be recognized as causative for the affected individual's disorder due to insufficient knowledge about the variant or the gene and its function. Reanalysis of the data is available upon request by the healthcare provider to incorporate updated clinical information and/or newly emerging gene and variant information. Updates to the classification of a sequence variant may be accessed through ClinVar (www.clinvar.com). Even if these tests identify the underlying genetic cause of a disorder in an affected individual, it is possible that such a diagnosis will not permit an accurate prediction of the prognosis or severity of the disease. While there is a possibility that identifying the genetic cause may help direct management and treatment of the disease, it is also possible that this knowledge will not change management or treatment.

References:

1. Bamshad et al. (2011) Nature Reviews. Genetics 12 (11):745-55 (PMID: 21946919)
2. Lelieveld et al. (2015) Hum. Mutat. 36 (8):815-22 (PMID: 25973577)
3. Belkadi et al. (2015) Proc. Natl. Acad. Sci. U.S.A. 112 (17):5473-8 (PMID: 25827230)
4. Retterer et al. (2016) Genet. Med. 18 (7):696-704 (PMID: 26633542)
5. Richards et al. (2015) Genetics In Medicine 17 (5):405-24 (PMID: 25741868)

6. Clark et al. (2018) NPJ Genom Med 3 :16 (PMID: 30002876)

7. Petrikin et al. (2018) NPJ Genom Med 3 :6 (PMID: 29449963)

8. Bick et al. (2017) J Pediatr Genet 6 (2):61-76 (PMID: 28496993)

9. Willig et al. (2015) Lancet Respir Med 3 (5):377-87 (PMID: 25937001)

10. Farnaes et al. (2018) NPJ Genom Med 3 :10 (PMID: 29644095)

11. Kingsmore et al. (2019) Am. J. Hum. Genet. 105 (4):719-733 (PMID: 31564432)

12. Scocchia et al. (2019) NPJ Genom Med 4 :5 (PMID: 30792901)

13. Retterer et al. (2015) Genet. Med. 17 (8):623-9 (PMID: 25356966)