Report References

- Fountain MD, *et al*. (2019) *Genet. Med.* **21**(8):1797-1807. **PMID:**30679821
- Hao YH, *et al*. (2015) *Mol. Cell* **59**(6):956-69. **PMID:**26365382

Resources Used for Bioinformatics, Medical Review Filtering, and Reporting

- 1000 Genomes [Internet]: 1000 Genomes Project Consortium (2010) *Nature* **467**(7319):1061-1073. Available from: http://www.1000genomes.org.
- Ambry Clinical Validity Assessment: Smith ED, *et al.* (2017) *Hum Mutat.* **38**(5):600-608.
- BayesDel [Internet]: Feng BJ. (2017) *Hum Mutat* **38**(3):243-251.
- Berkeley Drosophila Genome Project [Internet]: Reese MG, *et al.* (1997) *J Comp Biol* **4**(3), 311-23. http://www.fruitfly.org/seq_tools/splice.html.
- ClinGen Clinical Validity Classifications [Internet]: https://www.clinicalgenome.org/knowledge-curation/gene-curation/clinical-validity-classifications; Rehm HL, *et al.* (2015) *N Engl J Med* **372**(23):2235-2242.
- Clinical Genomic Database [Internet]: Solomon BD, *et al.* (2013) *Proc Natl Acad Sci U S A.* **110**(24):9851-5. Available from: http://research.nhgri.nih.gov/CGD.
- Combined Annotation Dependent Depletion (CADD) [Internet]: Kircher M, *et al.* (2014) *Nat Genet.* **46**(3):310-5. Available from: http://cadd.gs.washington.edu.
- Database of Single Nucleotide Polymorphisms (dbSNP) [Internet]: Bethesda (MD): National Center for Biotechnology Information, National Library of Medicine. (dbSNP Build ID: 135). Available from: http://www.ncbi.nlm.nih.gov/projects/SNP.
- DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources. Firth, H.V. *et al* (2009). *Am J Hum Genet* **84:**524-533. https://decipher.sanger.ac.uk/
- ESEfinder [Internet]: Smith PJ, *et al.* (2006) *Hum Mol Genet* **15**(16):2490-2508 and Cartegni L, *et al.* (2003) *Nucleic Acid Res* **31**(13):3568-3571. Available from: http://cb.utdallas.edu/tools/ESE
- Exome Aggregation Consortium (ExAC) [Internet], Cambridge, MA (URL: http://exac.broadinstitute.org). (Lek M, *et al* 2016: see below)
- Exome Variant Server, NHLBI Exome Sequencing Project (ESP) [Internet]: Seattle, WA. Available from: http://evs.gs.washington.edu/EVS.
- Expression Atlas: Differential and Baseline Expression [Internet]: Petryszak, R. *et al.* (2013) *Nucleic Acids Res* **10**.1093/nar/gkt1270. Available from: http://www.ebi.ac.uk/gxa/home.
- Farwell Hagman KD, *et al.* (2016) *Genet Med* **19**(2):224-235.
- GeneMANIA [Internet]: Warde-Farley D, *et al.* (2010) *Nucleic Acids Res* **38**(Web Server issue):W214-20. Available from: http://genemania.org.
- GeneReviews [Internet]: Pagon RA, *et al.* editors. (1993-) Seattle, WA: University of Washington, Seattle. Available from: http://www.ncbi.nlm.nih.gov/books/NBK1116.
- Genome Aggregation Database (gnomAD) [Internet], Cambridge, MA. Available from: http://gnomad.broadinstitute.org/ (Lek M, *et al* 2016; Karczewski KJ, *et al.* 2020: see below)
- Grantham prediction: Grantham R. (1974) *Science* **185**(4151):862-864.
- Green RC, *et* al. (2013) *Genet Med* **15**(7):565-74.
- HGMD® [Internet]: Stenson PD, *et al.* (2014) *Hum Genet.* **133**(1):1-9. Available from: http://www.hgmd.cf.ac.uk.
- Integrative Genomics Viewer (IGV): Thorvaldsdóttir H, *et al.* (2012) *Brief Bioinform* **14**(2):178-192.
- Kalia SS, *et al.* (2016) *Genet Med* **19**(2):249-255.
- Karczewski KJ, *et al.* (2020) *Nature* **581**(7809):434-443
- Kyoto Encyclopedia of Genes and Genomes (KEGG) [Internet]: Kanehisa M, *et al.* (2014) *Nucleic Acids Res* **42**. http://www.genome.jp/kegg.
- Lek M, *et al* (2016) Nature 536(7616):285-91.
- Miller DT, *et al.* (2021) *Genet Med* **23**(8):1391-98.
- Miller DT, *et al.* (2022) *Genet Med* **24**(7):1407-14.
- Miller DT, *et al.* (2023) *Genet Med* 2023 Jun 15;100866. doi: 10.1016/j.gim.2023.100866.
- Mouse Gene Expression Database (GXD): Finger JH, *et al.* (2011): *Nucleic Acids Res* **39**(suppl 1):D835-D841. Available from: http://www.informatics.jax.org.
- Mouse Genome Database (MGD) [Internet]: Eppig JT, *et al.* (2012) *Nucleic Acids Res* **40**(1):D881-86 Available from: http://www.informatics.jax.org.
- Mutation Assessor (functional impact of protein mutations) [Internet]: Reva BA *et al.* (2011) *Nucleic Acids Res* **39**(17):e118. Available from: http://mutationassessor.org.
- NeXtProt [Internet]: Lane L, *et al.* (2012) neXtProt: a knowledge platform for human proteins. *Nucleic Acids Res* **40**(D1): D76-D83. Available from: http://www.nextprot.org.
- Maquat LE. Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nat Rev Mol Cell Biol* 2004 **5**(2):89-99.
- OMIM (Online Inheritance in Man) [Internet]: Copyright© 1966-2012 Johns Hopkins University. Available from: http://www.omim.org.
- PolyPhen [Internet]: Adzhubei IA, *et al.* (2010) *Nat Methods* **7**(4):248-249. Available from: http://genetics.bwh.harvard.edu/pph2.
- PROVEAN: Choi Y, *et al.* (2012) *PLoS One* **7**(10):e46688.
- RefSeq: The NCBI handbook [Internet]: Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2002 Oct. Chapter 18, The Reference Sequence (RefSeq) Project. Available from: http://www.ncbi.nlm.nih.gov/refseq.
- Richards, *et al.* On behalf of the ACMG Laboratory Quality Assurance Committee (2015) Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*, **17**(5), 405-424.
- SIFT [Internet]: Kumar P *et al.* (2009) *Nat Protoc.* **4**(7):1073-81. http://sift.jcvi.org.
- Splicing Prediction: Jaganathan K, *et al.* (2019) *Cell* **176**(3):535-548.e24.

ExomeNext® Assay Information

General Information: Ambry's ExomeNext® is a cost-effective, comprehensive, integrated exome sequencing assay designed to increase the diagnostic yield for genetic disorders that have eluded definitive delineation using traditional diagnostic approaches. The exome represents all the exons, which are the regions in the human genome that are translated into proteins. It is estimated that the protein-coding regions of the human genome contain about 85% of the disease-causing mutations. Whole-exome sequencing has been successfully applied to identify both inherited and de novo mutations in a diverse variety of autosomal dominant, recessive, and X-linked disorders. In addition to the primary analysis, which is performed with the purpose of uncovering the underlying genetic cause for a given clinical presentation, the clinical diagnostic exome may also be utilized to provide secondary findings, which are pathogenic or likely pathogenic alterations in genes that lead to diseases unrelated to the patient's present clinical presentation.

Result Reports: A primary clinical report will only be generated for the proband regardless of number of family members submitted. However, it may be possible to infer information about family member's results based on the proband's report. Pathogenic mutation(s) likely to factor into the patient's current clinical presentation are always reported. Since new scientific information becomes available on a regular basis, this could alter the interpretation of previously reported results. In the event of a change in interpretation, an unsolicited reclassification/amended report may be issued to the ordering clinician. Pathogenic mutations or likely pathogenic variants identified within the ACMG secondary findings gene list are reported separately unless opted out (Kalia, 2016; Miller, 2023). Expanded childhood onset secondary findings are available in a separate report for prenatal exome testing orders.

Test Limitations: This test was developed and its performance characteristics were determined by Ambry Genetics. It has not been cleared or approved by the US Food and Drug Administration. The FDA does not require this test to go through premarket FDA review. It should not be regarded as investigational or for research. This test should be interpreted in context with other clinical findings. This report does not represent medical advice. Any questions, suggestions, or concerns regarding interpretation of results should be referred to a genetic counselor, medical geneticist, or physician skilled in evaluating the relevant medical literature. This laboratory is certified under the Clinical Laboratory Improvement Amendments (CLIA) as qualified to perform high complexity clinical laboratory testing. The following types of mutations are detectable: nucleotide substitutions, small deletions, small insertions, small indels, and gross deletions and duplications. The overall coverage of each gene varies and each individual may have slightly different coverage yield. Accurate exon-level gross deletion and duplication detection by exome sequencing is dependent on several factors such as inherent sequence properties of the targeted regions, including shared homology and exon size, depth-of-coverage, efficiency of capture, and degree of read depth variation in the selected reference samples. Therefore, the specificity and sensitivity of gross deletion and duplication detection by exome sequencing may be reduced. Exome sequencing is not intended to analyze the following types of mutations: gross rearrangements, deep intronic variations, long repeat sequences, portions of genes with highly homologous pseudogenes, trinucleotide repeat sequences, mutations involved in tri-allelic inheritance, certain mitochondrial genome mutations, epigenetic effects, oligogenic inheritance, and X-linked recessive mutations in females who manifest disease due to skewed X-inactivation and other unknown abnormalities. A negative result from the analysis cannot rule out the possibility that the tested individual carries a rare unexamined mutation or mutations in an undetectable region. Ambry's ExomeNext® detection rate is 30% for positive or likely positive relevant findings identified in established disease-gene associations, and an additional 7% for candidate gene findings (Farwell, 2015; Farwell Hagman, 2017). Although molecular tests are highly accurate, rare diagnostic errors may occur. Possible diagnostic errors include sample mix-up, erroneous paternity identification, technical errors, clerical errors, and genotyping errors. Genotyping errors can result from trace contamination of PCR reactions, from rare genetic variants that may interfere with analysis, or from other sources.

Methodology: Genomic deoxyribonucleic acid (gDNA) is isolated from the patient's provided specimen. Samples are prepared using the IDT xGen Exome Research Panel V1.0 (IDT). Each DNA sample is sheared, adaptor ligated, PCR-amplified and incubated with the exome baits. Captured DNA is eluted and PCR amplified. Final quantified libraries are seeded onto an Illumina flow cell and sequenced using paired-end, 150 cycle chemistry on the Illumina NovaSeq, NextSeq or HiSeq. Initial data processing, base calling, alignments and variant calls are generated by various bioinformatics tools using genome assembly GRCh37/hg19. Data is annotated with the Ambry Variant Analyzer tool (AVA), including: nucleotide and amino acid conservation, biochemical nature of amino acid substitutions, population frequency, and predicted functional impact. Data analysis is focused on small insertions and deletions, canonical splice site alterations, and non-synonymous alterations. Gross deletion/duplication analysis is assessed for proband only for all genes within the targeted exome using a custom pipeline based on coverage and/or breakpoint analysis from NGS data and is followed by a confirmatory orthogonal method, as needed. The following sites are used to search for previously described gene mutations and polymorphisms: the Human Gene Mutation Database (HGMD), the Single Nucleotide Polymorphism database (dbSNP), gnomAD, ESP, 1000 genomes, and online search engines (e.g., PubMed). Variants are then filtered further based on applicable inheritance models. Co-segregation studies are performed if family members are available. All relevant findings undergo manual review by molecular geneticists using integrated genomics software (IGV) and/or undergo confirmation either by automated fluorescence dideoxy (aka "Sanger") sequencing or via coverage and alternate read ratios above established confidence thresholds. Additionally, gross deletions/duplications are confirmed by SNP Microarray (Affymetrix® CytoScan™ HD Array), in-house targeted array, MLPA, or Sanger sequencing. Co-segregation results may be confounded by many factors which cannot be completely ruled out including haploinsufficiency, reduced penetrance, age-of-onset, and/or variable expressivity. Relevant findings are evaluated from among the genes in Ambry's internal, dynamic gene database which classifies genes as characterized or uncharacterized Mendelian disease genes based on clinical validity (Smith, 2017). Characterized genes are those currently known to underlie at least one Mendelian genetic condition. Uncharacterized genes are those not currently known or with insufficient evidence to be associated with a Mendelian genetic condition. Characterized genes are analyzed first. If no positive findings are identified, reflex analysis of uncharacterized genes occurs for potential identification of a candidate gene finding. The analysis of candidate gene findings is only performed when an informative trio is received for testing and focuses on de novo, autosomal recessive, or X-linked inherited alterations. Each alteration remaining after inheritance model filtering is analyzed and/or reviewed by board certified molecular geneticists to identify the most likely causative alteration(s). Interpretation is based on the clinical, family, and test information provided by the referring provider and the current knowledge of genes and alterations at the time of reporting. Screening and analysis of known mtDNA mutations related to the proband's clinical phenotype is included if ordered. Amplification of the entire mitochondrial genome is carried out by long distance PCR and sequencing of mitochondrial DNA (mtDNA) is performed separately on Illumina MiSeq. If ordered, ribonucleic acid (RNA) **is** isolated from the patient's whole blood. RNA is converted to complementary DNA (cDNA) by reverse transcriptase polymerase chain reaction (RT-PCR).RNA analysis is performed for reportable germline DNA variants expected to impact splicing, provided such studies are likely to meaningfully inform variant classification. Variants in genes with limited expression in whole blood, limited gene-disease validity, or an inconsistent mechanism of disease do not qualify for RNA analysis. Additionally, variants reported as secondary findings do not qualify for RNA analysis. For eligible variants, primers are designed to amplify the relevant region of the pertinent gene from reverse transcribed cDNA. The splicing patterns in variant carriers are then compared to those in control individuals to identify aberrant splicing. The presence of aberrantly spliced RNA transcripts meeting quality thresholds is incorporated as evidence for the assessment and classification of the DNA variants.

Analysis of Alterations: The following lines of evidence are used to assess the pathogenic nature of an alteration: presence in affected and healthy populations, co-segregation information, functional studies, alteration type, conservation, *in silico* predictions, and presence in a functional protein domain. The absence of a particular line of evidence implies that no information was found or it does not apply for that alteration type (e.g. in silico for truncating alterations).

Analytical range: Approximately 75% of the bases are expected to have quality scores of Q30 or higher, which translates to an expected base-calling error rate of 1:1000, or an expected base-calling accuracy of 99.9%. Additionally, 90% and 95% of the exome will be covered at ≥20x and ≥10x respectively under current run conditions, generally sufficient for high quality heterozygous and homozygous variant calling for germline variants. For any given individual ~ 10% of the targeted exome is not sequenced well enough to make a confident call. Each individual may have slightly different coverage yield distributions within the exome. Exons plus at least 6 bases into the 5' and 3' ends of all the introns are analyzed and reported. The pipeline detects deletions and duplications >5 exons in size in sequences with sufficient resolution. The minimum depth of coverage for targeted mitochondrial bases is greater than 1,000X.