Illumina Clinical Services Laboratory Variant Classification

Summary

Variant classification in the Illumina Clinical Services Laboratory (ICSL) is performed in accordance with established guidelines (Richards et al. 2015), and utilizes an integrated evaluation of knowledge regarding the gene and disease relationship, frequency information obtained from population databases, and other relevant evidence gleaned from the literature and other resources.

Gene Curation

- For a predefined list of genes associated with Mendelian disorders, information regarding the gene and associated disease is held in an internally curated list that details the disease name, gene symbol, transcript, inheritance mode, and penetrance and prevalence estimates. The information in the list is curated from four sources (GeneReviews, Genetics Home Reference, Orphanet, and Online Mendelian Inheritance in Man (OMIM)) and is updated regularly.
- For genes outside of the predefined list in which a variant of interest is selected for curation, a manual gene curation is pursued to establish the strength of evidence supporting a relationship with Mendelian disease. The information is curated using the sources listed above as well as Human Gene Mutation Database (HGMD) and ClinVar.

Variant Curation

- Population frequency information is gathered from:
    - 1000 Genomes Project
    - NHLBI GO Exome Sequencing Project (ESP)
    - Exome Aggregation Consortium (ExAC)

  Frequency information also may be interrogated from our internal database of whole genome sequences.
- Literature searches are performed in PubMed, PubMed Central (PMC), Google Scholar, Google, HGMD and ClinVar for each variant using the gene name, cDNA change, amino acid change, and rsID, utilizing alternative nomenclature as available.
- Additional resources are consulted as required depending on the variant. For example, the effect of potentially truncating or elongating variants is confirmed using Mutalyzer.
- Variants that are not found in HGMD are subject to classification through the use of an automated classification scoring system. An autoscore is calculated based on allele frequency, inheritance mode, and disease prevalence and penetrance estimates of the associated disease and is used to determine the potential of a variant to cause the associated disease. Autoscore classifies variants as benign or likely benign when they are found at a frequency that is much greater than the prevalence of the disease. These variants are not subjected to manual review.

Variants that cannot be ruled out on the basis of autoscore are subjected to a literature search as outlined above. If no literature is found, the variant is classified as a variant of unknown significance. If literature is found, the variant is subjected to the manual curation process outlined above.

Variant Classification

Variants are classified based on assessment and concordance of the available evidence. Each variant classification is subject to professional review. Our classification system has been developed from the recommendations of the American College of Medical Genetics (ACMG) for variant classification and reporting (Richards et al. 2015), with consideration that many variants are detected in ostensibly healthy adults and with the addition of a sixth category termed VUS-Suspicious. This category was developed for variants of unknown significance that have limited evidence for pathogenicity but are deemed noteworthy for reporting, bringing attention to variants that are on the border between unknown significance and likely pathogenic. Variants classified as VUS-Suspicious are submitted as VUS in ClinVar, with language distinguishing the VUS-Suspicious classification found in the evidence summary.

The classification criteria for each of the six categories are as follows:

Pathogenic

- The variant is reported in the literature in multiple unrelated cases, usually with control data.
- The variant segregates with disease in at least two generations of a family.
- The variant is located in a highly conserved region.
- Functional evidence or other evidence suggests a deleterious effect of the variant on gene expression or function consistent with the mechanism of disease.

Likely Pathogenic

- The variant is reported in the literature in a limited number of cases, with or without control data.
- The variant is located in a highly conserved region.
- Limited or no functional evidence available, but overall biological expectations suggestive of the variant having a deleterious effect on gene expression or function consistent with the mechanism of disease.

Variant of Unknown Significance-Suspicious (VUS-Suspicious)

- There is limited evidence that the variant could be causative of disease. The information available is insufficient to categorize the variant as likely pathogenic. This category was added to bring attention to variants that are on the border between unknown significance and likely pathogenic. For example, if the variant is reported in only a single homozygote or compound heterozygote for a recessive condition or in a very limited number of cases for a dominant condition, this evidence is limited and suggestive of pathogenicity but is not conclusive.
- Null variants, including nonsense, frameshift, canonical +/- 1 or 2 splice sites, and initiation codon variants, with no other supporting evidence are considered to be suspicious for pathogenicity and are classified in this category. Additional evidence is needed to classify these variants as likely pathogenic or pathogenic.

Variant of Unknown Significance

- Little or nothing has been reported regarding this variant, or the reported evidence in the literature is incomplete and/or contradictory.
- The evidence could be contradictory within the literature or between the literature and other available evidence (e.g., allele frequency).

Likely Benign

- The variant is reported in the literature in a similar number of cases and controls if control data are available.
- The variant does not segregate with disease within a family.
- Variant frequency is higher than expected in the general population based on inheritance mode and disease prevalence and penetrance estimates.
- The variant may be non-conserved and / or predicted to be well-tolerated.
- Functional evidence or other evidence suggests no deleterious effect of the variant on gene expression or function.

Benign

- The variant is not reported in the literature in cases or is reported in a similar number of cases and controls if control data are available.
- Established in the literature as a variant that is not associated with Mendelian disease.
- The variant does not segregate with disease within a family.
- Variant frequency is too high to be causative based on inheritance mode and disease prevalence and penetrance estimates.
- The variant may be non-conserved and / or predicted to be well-tolerated.
- Functional evidence or other evidence suggests no deleterious effect of the variant on gene expression or function.

Sources Used in Variant Classification

Frequency information:
1000 Genomes Project: http://browser.1000genomes.org
NHLBI GO Exome Sequencing Project (ESP): http://evs.gs.washington.edu/EVS/
Exome Aggregation Consortium (ExAC): http://exac.broadinstitute.org/

Literature searches:
PubMed: https://www.ncbi.nlm.nih.gov/pubmed
PubMed Central (PMC): https://www.ncbi.nlm.nih.gov/pmc
Google Scholar: https://scholar.google.com/
Google: https://www.google.com/

Gene and disease information:
GeneReviews: https://www.ncbi.nlm.nih.gov/books/NBK1116/
Genetics Home Reference: https://ghr.nlm.nih.gov/
Orphanet: http://www.orpha.net/consor/cgi-bin/index.php?lng=EN
Online Mendelian Inheritance in Man (OMIM): https://www.ncbi.nlm.nih.gov/omim

Disease-specific resources:
Cystic Fibrosis Mutation Database: http://www.genet.sickkids.on.ca/app
Clinical and Functional Translation of CFTR (CFTR2): http://cftr2.org/
Tuberous Sclerosis Project: http://tsc-project.partners.org/
Locus-Specific Mutation Databases: http://www.hgvs.org/locus-specific-mutation-databases
The International Society for Gastrointestinal Hereditary Tumours (InSiGHT): https://www.insight-group.org/variants/databases/
Locus Specific Database list: http://grenada.lumc.nl/LSDB_list/lsdbs

Additional resources:
ClinVar: https://www.ncbi.nlm.nih.gov/clinvar/
dbSNP: https://www.ncbi.nlm.nih.gov/SNP/
HUGO Genome Nomenclature Committee (HGNC): http://www.genenames.org/
Human Gene Mutation Database (HGMD): http://www.hgmd.cf.ac.uk/ac/index.php
Human Genome Variation Society: http://www.hgvs.org
Mutalyzer: https://mutalyzer.nl/
NCBI Protein database: https://www.ncbi.nlm.nih.gov/protein
UCSC Genome Browser: http://genome.ucsc.edu

References

Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL; ACMG Laboratory Quality Assurance Committee. 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. 2015 May; 17(5):405-24