Illumina Clinical Services Laboratory Copy Number Variant Classification for Rare and Undiagnosed Genetic Disease

Summary

Classification of copy number variants (CNVs) in nuclear genes for rare and undiagnosed genetic disease (RUGD) in the Illumina Clinical Services Laboratory (ICSL) is based on the established recommendations of the American College of Medical Genetics and Genomics (ACMG) (Kearney et al. 2011), and utilizes an integrated evaluation of knowledge regarding the gene content of the copy number event, curated gene-disease relationships (GDRs), evidence gleaned from the literature and other resources, and frequency of similar CNVs identified in control populations.

Gene Curation and Classification

- In instances where a CNV of interest overlaps or is contained within a single gene or where a single gene in the CNV interval is identified as a candidate based on overlap with the proband's phenotype, the strength and classification of the associated GDR is established through gene curation in accordance with the ClinGen framework for evaluating the clinical validity of a GDR (Strand et al. 2017).
    - Within this framework, GDRs may be classified as Definitive, Strong, Moderate, Limited, No Known Disease Relationship, or Conflicting Evidence Reported (Disputed or Refuted).
    - If a GDR has been previously curated by ClinGen, by Ceyhan-Birsoy et al. (2017), or by ICSL, and classified as Strong or Definitive, further gene curation is not required, and these data are utilized to support CNV classification and reporting.
    - If gene curation data are unavailable or if the curated GDR is classified as Moderate or below, gene curation is performed following the gene-disease validity standard operating procedure available from ClinGen at the time of curation.
    - Where applicable, professional judgement and expertise are utilized in determining the evidential strength and the final clinical validity classification.
- For CNVs with two or more overlapping genes, full assessment through the ClinGen framework of all associated GDRs generally may not be feasible. In these instances, potential GDRs are examined using information from resources such as OMIM, GeneReviews, ClinVar, and primary literature to determine whether any genes within the CNV are implicated in a disease with phenotypic features that overlap those reported in the proband.

Copy Number Variant Curation and Classification

- CNVs of interest are subjected to a manual curation. Evidence is collected to support the application of the guidelines as described by Kearney et al. (2011).
- CNV visualization is performed in the University of California Santa Cruz (UCSC) Genome Browser and the CNV of interest is compared with other CNVs in the region as identified in tracks from

ClinGen, ClinVar, DECIPHER (Firth et al. 2009), Database for Genomic Variants (DGV) (MacDonald et al. 2014), and Developmental Delay cases and controls (Cooper et al. 2011; Coe et al. 2014).

- Control population frequency information is gathered from the Genome Aggregation Database (gnomAD). Frequency information also may be interrogated from our internal database of whole genome sequences.
- Literature searches are performed in PubMed, PubMed Central, Google, and Google Scholar for each CNV using the gene name(s), event (gain/duplication or loss/deletion), chromosomal band, and or chromosome position.
- Additional resources are consulted as required.
- CNVs are classified based on assessment and concordance of the available evidence. Each CNV classification is subject to professional review.
- ICSL's classification system has been developed from the recommendations of the ACMG for CNV interpretation and reporting (Kearney et al. 2011).
- Variants may be classified as Pathogenic, Uncertain Clinical Significance – Likely Pathogenic, Uncertain Clinical Significance, Uncertain Clinical Significance – Likely Benign, or Benign, depending on the available evidence and application of the guidelines and internal criteria.
    - Pathogenic: Documented as clinically significant in multiple publications, even if penetrance and expressivity are variable. Includes large CNVs that may not be described in the literature of the same size but overlap with an interval with established clinical significance.
    - Uncertain Clinical Significance – Likely Pathogenic: 1) CNV described in a single case report but with well-defined breakpoints and phenotype that overlap with patient. 2) A gene within the CNV has a very compelling function relevant to patient phenotype.
    - Uncertain Clinical Significance: 1) CNV contains genes but unknown if genes are dose sensitive. 2) CNV is described in multiple contradictory publications or databases.
    - Uncertain Clinical Significance – Likely Benign: 1) CNV has no gene in interval but is reported because of size criteria in laboratory. 2) CNV is described in small number of cases in databases for the general population but does not represent a common polymorphism.
    - Benign: CNV has been reported in publications or curated databases as a benign variant. CNV should be documented in greater than 1% of the population. Remember to consider dosage, as benign duplications could have clinical relevance for deletions.

Sources Used in CNV and Gene Curation

Frequency information:
Genome Aggregation Database (gnomAD): http://gnomad.broadinstitute.org/about

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

Additional resources:
ClinGen: https://clinicalgenome.org/
ClinGen Dosage Sensitivity: https://www.clinicalgenome.org/curation-activities/dosage-sensitivity/
ClinVar: https://www.ncbi.nlm.nih.gov/clinvar/
DECIPHER: https://decipher.sanger.ac.uk/
The Human Protein Atlas: http://www.proteinatlas.org/
ISCA: http://dbsearch.clinicalgenome.org/ (archived)
NCBI Gene: http://www.ncbi.nlm.nih.gov/gene
NCBI Protein database: https://www.ncbi.nlm.nih.gov/protein
UCSC Genome Browser: http://genome.ucsc.edu
Uniprot: https://uniprot.org

References

Ceyhan-Birsoy O, Machini K, Lebo MS, Yu TW, Agrawal PB, Parad RB, Holm IA, McGuire A, Green RC, Beggs AH, Rehm HL. 2017 Jul. A curated list for reporting results of newborn genomic sequencing. Genet Med. 19 (7): 809-818.

Coe BP, Witherspoon K, Rosenfeld JA, van Bon BW, Vulto-van Silfhout AT, Bosco P, Friend KL, Baker C, Buono S, Vissers LE, Schuurs-Hoeijmakers JH, Hoischen A, Pfundt R, Krumm N, Carvill GL, Li D, Amaral D, Brown N, Lockhart PJ, Scheffer IE, Alberti A, Shaw M, Pettinato R, Tervo R, de Leeuw N, Reijnders MR, Torchia BS, Peeters H, O'Roak BJ, Fichera M, Hehir-Kwa JY, Shendure J, Mefford HC, Haan E, Gécz J, de Vries BB, Romano C, Eichler EE. 2014 Oct. Refining analyses of copy number variation identifies specific genes associated with developmental delay. Nat Genet. 46(10):1063-71.

Cooper GM, Coe BP, Girirajan S, Rosenfeld JA, Vu TH, Baker C, Williams C, Stalker H, Hamid R, Hannig V, Abdel-Hamid H, Bader P, McCracken E, Niyazov D, Leppig K, Thiese H, Hummel M, Alexander N, Gorski J, Kussmann J, Shashi V, Johnson K, Rehder C, Ballif BC, Shaffer LG, Eichler EE. 2011 Aug 14. A copy number variation morbidity map of developmental delay. Nat Genet. 43(9):838-46.

Firth HV, Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, Van Vooren S, Moreau Y, Pettett RM, Carter NP. 2009 Apr. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. Am J Hum Genet. 84(4):524-33.

Kearney HM, Thorland EC, Brown KK, Quintero-Rivera F, South ST. 2011 Jul. American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants. Genet Med. 13(7):680-5.

MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW. 2014 Jan. The Database of Genomic Variants: a curated collection of structural variation in the human genome. Nucleic Acids Res. 42(Database issue):D986-92.

Strande NT, Riggs ER, Buchanan AH, Ceyhan-Birsoy O, DiStefano M, Dwight SS, Goldstein J, Ghosh R, Seifert BA, Sneddon TP, Wright MW, Milko LV, Cherry JM, Giovanni MA, Murray MF, O'Daniel JM, Ramos EM, Santani AB, Scott AF, Plon SE, Rehm HL, Martin CL, Berg JS. 2017 Jun 1. Evaluating the Clinical Validity of Gene-Disease Associations: An Evidence-Based Framework Developed by the Clinical Genome Resource. Am J Hum Genet. 100(6):895-906.