Genetics in Medicine (2022) 24, 729–735

ELSEVIER

Genetics in Medicine

www.journals.elsevier.com/genetics-in-medicine

BRIEF REPORT

Identification and quantification of oligogenic loss-of-function disorders

Arthur Stefanski[1,2], Eduardo Pérez-Palma[1,3], Marko Mrdjen[1], Megan McHugh[1], Costin Leu[1,4], Dennis Lal[1,2,4,5,*]

[1]Genomic Medicine Institute, Cleveland Clinic Lerner Research Institute, Cleveland, OH; [2]Epilepsy Center, Neurological Institute, Cleveland Clinic, Cleveland, OH; [3]Universidad del Desarrollo, Centro de Genética y Genómica, Facultad de Medicina Clínica Alemana, Santiago, Chile; [4]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA; [5]Cologne Center for Genomics (CCG), University of Cologne, Cologne, Germany

ARTICLE INFO

Article history:
Received 29 October 2021
Accepted 29 October 2021
Available online 3 December 2021

Keywords:
CNV
Epilepsy
Loss-of-function
Monogenic
Oligogenic

ABSTRACT

Purpose: Monogenic disorders can present clinically heterogeneous symptoms. We hypothesized that in patients with a monogenic disorder caused by a large deletion, frequently additional loss-of-function (LOF)-intolerant genes are affected, potentially contributing to the phenotype.
Methods: We investigated the LOF-intolerant gene distribution across the genome and its association with benign population and pathogenic classified deletions from individuals with presumably monogenic disorders. For people with presumably monogenic epilepsy, we compared Human Phenotype Ontology terms in people with large and small deletions.
Results: We identified LOF-intolerant gene dense regions that were enriched for ClinVar and depleted for population copy number variants. Analysis of data from >143,000 individuals with a suspected monogenic disorder showed that 2.5% of haploinsufficiency disorder–associated deletions can affect at least 1 other LOF-intolerant gene. Focusing on epilepsy, we observed that 13.1% of pathogenic and likely pathogenic ClinVar deletions <3 megabase pair, covering the diagnostically most relevant genes, affected at least 1 additional LOF-intolerant gene. Those patients have potentially more complex phenotypes with increasing deletion size.
Conclusion: We could systematically show that large deletions frequently affected admditional LOF-intolerant genes in addition to the established disease gene. Further research is needed to understand how additional potential disease-relevant genes influence monogenic disorders to improve clinical care and the efficacy of targeted therapies.

Introduction

In many rare genetic disorders, individuals with the same affected gene still present with widely heterogeneous symptoms or phenotypes.[1,2] The clinical heterogeneity across individuals with variants affecting the same gene is not fully understood. Nongenetic factors such as environmental stress (eg, trauma or smoking), epigenetic, and

*Correspondence and requests for materials should be addressed to Dennis Lal, Genomic Medicine Institute, Cleveland Clinic Lerner Research Institute, Cleveland, OH 44195. *E-mail address:* lald@ccf.org

doi: https://doi.org/10.1016/j.gim.2021.10.026

multiple hits in different disease-associated genes have been associated with clinical heterogeneity.[2]

Gene panels are currently the most widely adopted clinical genetic test for rare disorders. Gene panels focus on a small selection of established and strong candidate genes associated with a disease (usually <300 genes).[3,4] Even in cases where an exome sequencing was performed, clinical laboratories frequently reported only results of established disease genes related to the patient's predefined phenotype. As a result, variants overlapping genes that could explain a patient's heterogeneous symptoms may not be reported.

Similarly, for large deletions that affect many genes, clinical genetic tests typically report only the most likely phenotype-associated gene within the affected genomic region. Genes with currently no established human disease phenotype, including genes known to be intolerant for loss-of-function (LOF) variants (LOF-intolerant genes),[5] ie, likely to cause haploinsufficiency disorders, are typically ignored. Because such LOF-intolerant genes could contribute to the phenotype, the current practice may lead to incomplete genetic testing results in some individuals and exacerbate knowledge gaps. If the notion that, in theory, additional genes do play a role even in monogenic disorders proves to be accurate, understanding the magnitude of this problem and its potential role in disease is unexplored.

Large deletions are frequently the cause of rare disorders with haploinsufficiency as the disease mechanism.[6] Because of their size and genomic position, large deletions can affect (in contrast to small truncating variants) several LOF-intolerant genes, apart from the assumed causal gene. We hypothesized that the phenotypes of individuals affected by such large deletions are representations of oligogenic disorders, irrespective of the involvement of a dominant gene with modifiers or an equal contribution of several genes. To explore our hypothesis, we first investigated the distribution of LOF-intolerant genes across the human genome. Next, we assessed the frequency of potential digenic and oligogenic disorders in >143,000 individuals referred to genetic testing for a suspected monogenic disorder.[6] We finally investigated ClinVar,[7] focusing on epilepsy, an established disorder with many known disease-associated LOF-intolerant genes and reported causal deletions.

Materials and Methods

Genomic distribution of LOF-intolerant genes

All 18,197 human genes were retrieved from the UCSC Human Genome Browser in hg19 assembly. We defined 3 intervals (0.5 megabase pair [Mb], 1 Mb, 5 Mb) and investigated the distribution of LOF-intolerant genes following 2 approaches: (1) around each gene and (2) as a nonoverlapping sliding window across the genome. In approach (1), we added 2.5 Mb before and after the gene. In approach (2) we investigated fixed 5 Mb sliding windows that do not overlap (eg, 0-5 Mb, 5-10 Mb, etc.). Using Bedtools,[8] we annotated and counted all genes (N = 18,197) and all LOF-intolerant genes (n = 3230) to these artificial intervals. LOF-intolerant genes were defined by a probability of LOF intolerance (pLI) > 0.9 on the basis of variation observed in 60,706 exomes of the Exome Aggregation Consortium.[5]

Frequency of deletions affecting multiple LOF-intolerant genes in affected individuals

We made use of the largest clinical copy number variant (CNV) studies to date. Truty et al[6] investigated >143,000 individuals with a suspected monogenic disorder (ie, hereditary cancer syndromes and cardiovascular, neurological, or pediatric disorders) and identified 2844 CNVs overlapping 1507 candidate genes. Of the 2844 reported CNVs, we selected pathogenic and likely pathogenic (P/LP) deletions and affecting genes associated with haploinsufficiency disorders.[6] Next, for each deletion we annotated all known LOF-intolerant genes (pLI ≥ 0.9)[5] using Bedtools.

Next, we screened ClinVar, the largest patient's variant repository.[7] Of the approximate 851,000 variants available on the hg19 assembly in ClinVar (accessed on January 31, 2021), we selected all P/LP deletions < 3 Mb (n = 37,426). We restricted the variant size to exclude chromosomal abnormalities and focus on the disease with its potential modifiers. To select the clinically relevant epilepsy genes, we intersected the 20 most frequently mutated genes from the 2 largest epilepsy clinical genetic screens published to date.[9,10] The intersection of genes from both studies yielded 12 genes, namely *CDKL5*, *FOXG1*, *KCNQ2*, *MECP2*, *PCDH19*, *PRRT2*, *SCN1A*, *SCN2A*, *SLC2A1*, *STXBP1*, *TSC2*, and *UBE3A*. Those 12 epilepsy genes contained 2 genes that are not LOF-intolerant genes (*MECP2* and *PRRT2*) but are nevertheless among the most commonly reported epilepsy genes in almost 20,000 individuals screened.[9,10] The gene *FOXG1* did not show any pathogenic CNV that affected an additional LOF-intolerant gene and was not considered further in the analysis, leaving the gene set with a final count of 11 genes and 1326 ClinVar P/LP deletions.

In addition, for the 11 epilepsy genes, we extracted variant data from the Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources (DECIPHER),[11] which contains classified patient variants and clinical phenotype data. Of the 1280 variants available (accessed September 10, 2021), we selected P/LP variants with existing phenotype information in the form of Human Phenotype Ontology (HPO)[12] terms (n = 690). To focus on variants likely to affect multiple genes, deletions < 200 kb were removed from further analysis. We used the HPO terms to categorize deletions to determine which variants are

more likely to cause a complex phenotype. Variants causing complex phenotypes are defined as containing non-neurological phenotypes such as bone, heart, limb, or skin in addition to their brain phenotype (eg, seizures or developmental delay), and they were grouped as (1) Deletion-complex and (2) Single-nucleotide variant-complex. We then used Pearson's correlation analysis to investigate the correlation of complex phenotypes with the size of the selected DECIPHER deletions. We also tested deletions found in the general population (UK Biobank).[13] From a sample of 472,228 array genotyped individuals from the UK Biobank, 171,825 deletions were available. The correlation between the number of LOF-intolerant genes and deletions from ClinVar and the UK Biobank was tested using Pearson's correlation analysis for each 5 Mb window.

Statistical analysis

We used R version 4.0.3 for all the analyses.[14] To determine the distribution of adjacent genes per interval for all genes and LOF-intolerant genes across the human genome, we quantified all adjacent LOF-intolerant genes contained in each CNV. Odds ratios (ORs) for the distribution of adjacent genes across all genes were calculated using Fisher's exact test. To determine the correlation between the number of LOF-intolerant genes and the total number of genes across all genomic segments, we performed a Pearson's correlation analysis. For each CNV, we considered all genes partially or entirely overlapping with the CNV as affected. ORs for the distribution of adjacent genes across all genes were calculated using Fisher's exact test.

Results

Distribution of LOF-intolerant genes across the genome

To explore the genome-wide distribution of LOF-intolerant genes and identify regions of the genome that may be prone to oligogenic deletion disorders, we investigated the number of neighboring genes (1) around every human gene in each interval and (2) across the genome in nonoverlapping sliding windows (Figure 1A). For each interval size, we found that the top decile of genes with most neighboring LOF-intolerant genes had significant 8.5-fold to 14.3-fold enrichments of neighboring LOF-intolerant genes compared with the lowest decile (0.5 Mb around each gene, 8.5-fold; 1 Mb around each gene, 13.1-fold; 5 Mb around each gene, 14.3-fold; Figure 1B). This enrichment was driven by a nonuniform distribution of all genes across the genome (Supplemental Figure 1). The numbers of neighboring LOF-intolerant genes for each gene are detailed in Supplemental Tables 1 to 3. Figure 1C shows the distribution of LOF-intolerant genes across the genome. The number of LOF-intolerant genes within a genomic region across the genome is correlated with the total number of genes ($r = 0.76$, $P = 2.2 \times 10^{-16}$). We detected 40 genomic windows with a higher proportion of LOF-intolerant genes than all other genes (OR > 1.57-11.60; $P < .05$; Figure 1D, Supplemental Figure 2 for details). We found the highest number of LOF-intolerant genes per 5 Mb interval at 11q13.1 with 188 genes overall, of which 36 were LOF-intolerant. This locus contains 48 known disease-associated genes (Supplemental Table 3). In contrast, we also observed gene sparse regions. For example, *CDH8* is an LOF-intolerant gene at 16q21 without neighboring genes in close proximity. Finally, we investigated the correlation between LOF-intolerant genes and P/LP ClinVar deletions or deletions found in 171,825 individuals from the general population (UK Biobank).[13] We found that the number of P/LP ClinVar deletions was significantly but weakly correlated with number of LOF-intolerant genes across genomic segments ($r = 0.087$, $P = .031$). Similarly, the number of LOF-intolerant genes and deletions in the general population were anticorrelated ($r = –0.15$, $P = .0003$).

Pathogenic deletions affecting monogenic disease–associated genes may cause oligogenic phenotypes in some individuals

To empirically assess the frequency of potential digenic and oligogenic disorders in individuals referred to genetic testing for a suspected monogenic disorder, we evaluated whether deletions classified as pathogenic in the screening of >143,000 individuals[6] associated with monogenic disorders affect additional LOF-intolerant genes that are potentially crucial to the disease presentation. By reanalyzing the data,[6] we identified, in total, 2844 CNVs in 384 clinically tested genes. We found that pathogenic deletions in 12.9% of genes (11 of 85 genes) associated with haploinsufficiency disorders affected at least 1 patient with a deletion covering an additional LOF-intolerant gene (Table 1). Overall, 2.5% of deletions involving genes associated with haploinsufficiency disorders affected at least 1 other LOF-intolerant gene. For example, 1 of 6 individuals with a deletion encompassing *KCNQ2* had another potential genetic finding in an LOF-intolerant gene (*EEF1A2*, Table 1).

Next, we explored the frequency of additional LOF-intolerant genes affected by large deletions in epilepsy. This disorder has been associated with many haploinsufficient genes and pathogenic deletions. We investigated ClinVar, focusing on the 11 most commonly reported epilepsy genes (see methods). In total, 13.1% ($n = 174$) of all P/LP ClinVar deletions involving the 11 epilepsy genes affected additional LOF-intolerant genes (Table 1). The fraction of deletions affecting at least 1

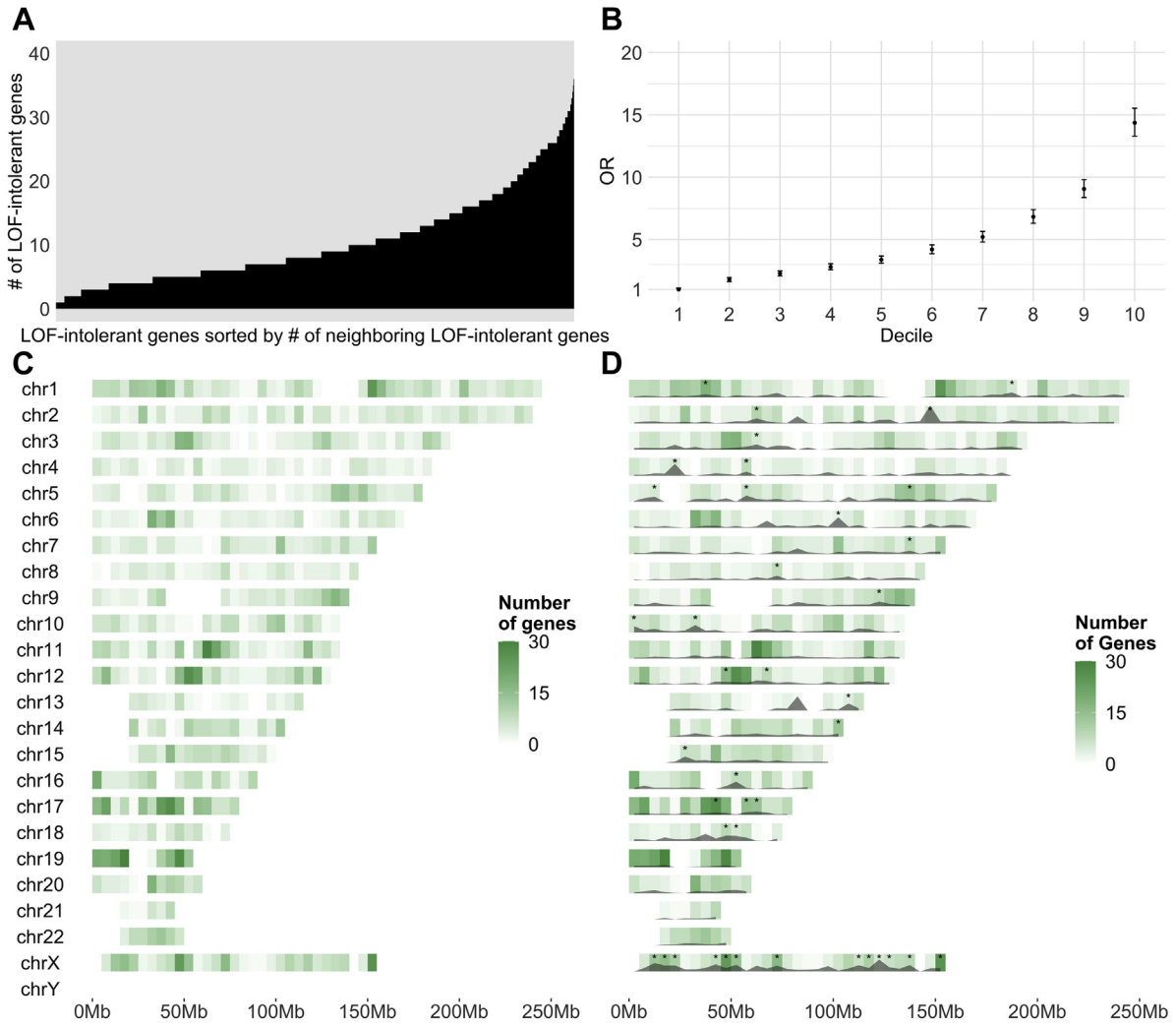

Figure 1 **Distribution and hotspots of LOF-intolerant genes across the genome.** A. Number of neighboring LOF-intolerant genes across LOF-intolerant genes within 5 megabase pair (Mb) and 2.5 Mb on each side. Top region 11q11-q13.4 with 36 LOF-intolerant genes within 5 Mb. B. ORs for the distribution of neighboring LOF-intolerant genes. On the basis of the number of neighboring LOF-intolerant genes, these neighboring LOF-intolerant genes were allocated to 10 deciles containing an increasing number of genes from deciles 1 to 10. Decile 1 contains the lowest number of neighboring genes and was used as a reference for deciles 2 to 10 with higher numbers of neighboring genes. ORs and *P*-values were calculated using a logistic regression model of deciles 2 to 10 against decile 1. The ORs with 95% CI are shown. C. Density of LOF-intolerant genes within 5 Mb intervals. The darker the green, the more genes within each 5 Mb window. Chr19 has the genomic segment with highest density of LOF-intolerant genes ($n = 30$). D. Density of LOF-intolerant genes within nonoverlapping 5 Mb sliding windows and fold-enrichment of LOF-intolerant genes in comparison with non-LOF–intolerant genes ($P < .05$). The gray areas depict the fold-enrichment of LOF-intolerant genes. # denotes number and * denotes significant enrichment. Chr, chromosome; LOF, loss-of-function; Mb, megabase pair; OR, odds ratio.

additional LOF-intolerant gene ranged from 1.5% to 77.7% for each epilepsy gene (Table 1). For example, 9.8% of individuals with a deletion encompassing the gene *SLC2A1* will likely have at least 1 other high pLI gene deleted that might contribute to the pathogenicity of the reported CNV.

Finally, we examined the impact of the size of multigenic deletions on phenotype complexity. The genetic and HPO[12] term data for the 11 epilepsy genes were collected from DECIPHER. We found that patient entries with more complex phenotypes, defined by reporting of non-neurological HPO terms, were correlated with a larger size of deletions in the DECIPHER database ($r = 0.18$, $P = 10^{-4}$) (Supplemental Table 4).

Discussion

Our study found that deletions involving genes associated with haploinsufficiency disorders affect at least 1 other

Table 1 Di/Oligogenic rare disorders due to large deletions might be common among patients thought to have a monogenic disorder

Data set	Gene	P/LP Deletions	Deletions Affecting Additional LOF-Intolerant Genes	Percentage of All P/LP Deletions Affecting Additional LOF-Intolerant Genes	All Additional LOF-Intolerant Genes (No Double Count of Genes)
Truty et al.[6]	*APC*	37	1	2.7	*LRRTM2, HDAC3, ARRDC3, CHD1, SEMA6A, LOX, LMNB1, SLC12A2, KIF3A, SEPT8, HSPA4, VDAC1, PPP2CA, PHF15, DDX46, SMAD5, FAM13B, CDC23, ETF1, HSPA9, MATR3, UBE2D2, ZMAT2, PCDHGC4, ARAP3, NDFIP1, YIPF5, JAKMIP2, FBXO38, CSF1R, PDGFRB, CAMK2A, NDST1, SYNPO, DCTN4, TNIP1, G3BP1, RASA1, ELL2, LNPEP, FBXL17, FER, YTHDC2, KCNN2, SNX2, CSNK1G3, ZNF608, MEGF10, FBN2, RAPGEF6, PDLIM4, FSTL4, AFF4, KLHL3, CTNNA1, KDM3B, NRG2, ANKHD1-EIF4EBP3, ANKHD1, PCDHAC2, ARHGAP26, RBM27, TCERG1, PPP2R2B, PPARGC1B, RBM22, DMXL1, CTC-432M15.3, FNIP1, ADAMTS19, KCTD16, DPYSL3, CSNK1A1, GRIA1*
	CHEK2	80	1	1.3	*EWSR1, NIPSNAP1, AP1B1, NF2, ZNRF3*
	KCNH2	3	1	33.3	*CDK5, SLC4A2, AGAP3, ABCF2, RHEB, PRKAG2*
	KCNQ2	6	1	16.7	*EEF1A2*
	NF2	5	1	20.0	*EWSR1, NIPSNAP1, AP1B1, ZNRF3*
	NFIX	1	1	100.0	*NACC1, CACNA1A*
	RPL26	1	1	100.0	*TP53, KDM6B, CHD3*
	TBX1	2	1	50.0	*UFD1L, SEPT5, HIRA*
	TCF4	1	1	100.0	*ST8SIA3, TXNL1, ONECUT2, NEDD4L, MALT1, WDR7, ZNF532*
	TSC2	16	2	12.5	*PKD1*
	UBE3A	5	5	100.0	*GABRB3*
	Total	636[a]	16	2.5	
ClinVar	*SCN2A*	47	1	2.1	*SCN3A*
	UBE3A	59	1	1.7	*MAGEL2*
	CDKL5	114	3	2.6	*SCML1, NHS, CDKL5, SCML2, PPEF1*
	PCDH19	71	3	4.2	*SRPX2*
	MECP2	263	4	1.5	*BGN, SLC6A8, IRAK1, RPL10, GDI1, FAM50A, HAUS7, IDH3G, ABCD1, RENBP, ARHGAP4, L1CAM, OPN1LW, FLNA, TKTL1, TAZ, PLXNA3, G6PD, GAB3, DKC1, ATP2B3, HCFC1, MPP1, PDZD4, F8*
	SLC2A1	41	4	9.8	*ATP6V0B, YBX1, KDM4A, IPO13, PTPRF, RNF220*
	STXBP1	66	6	9.1	*ENG, SLC25A25, DNM1, GOLGA2, SET, ZER1, LRRC8A, PPP2R4, CIZ1, SPTAN1, NUP188, RALGPS1*
	SCN1A	240	7	2.9	*SCN3A, SCN2A*
	TSC2	191	18	9.4	*FBXL16, UBE2I, CLCN7, TMEM204, MAPK8IP3, PKD1, CRAMP1L, CASKIN1*
	KCNQ2	104	26	25.0	*MRGBP, GID8, ZNF512B, TCFL5, DIDO1, YTHDF1, GMEB2, MYT1, EEF1A2, ZNF512B*
	PRRT2	130	101	77.7	*SH2B1, MAZ, TAOK2, ATXN2L, IL27, CORO1A, TBC1D10B, FBRS*
	Total	1326	174	13.1	

Clinical CNV data set: Considering 85 LOF-intolerant gene-associated disorders, at least 1 patient had 11 genes with a deletion covering an additional LOF-intolerant gene.[6] ClinVar data set: 174 (13.1%) of all P/LP ClinVar deletions affect at least 1 additional LOF-intolerant gene.

CNV, copy number variant; *LOF*, loss-of-function; *P/LP*, pathogenic/likely pathogenic.

[a]P/LP deletions from all 85 LOF-intolerant genes.

LOF-intolerant gene in 2.5% of all individuals. For presumably monogenic epilepsies, the number was even higher. For example, it has been estimated that disease-associated CNVs are found in 2% to 3% of people with schizophrenia, 10% of people with autism or epilepsy, and 25% or more of people with intellectual disability.[6,15-17] Possibly, several deleted LOF-intolerant genes contribute to the phenotypic spectrum of individuals with the same disorder. Advanced interpretation practices that incorporate this information into clinical practice should be established.

The frequency with which large deletions affected LOF-intolerant genes in addition to the putative causal gene is locus and disorder-specific as indicated by the range of 2.5% for haploinsufficiency disorders to 13.1% for monogenic epilepsies. Because *PRRT2* lies within a recurrently deleted region (16p11.2), it has a major contribution to the results. However, even when *PRRT2* is removed from the analysis, we still observe 6% of deletions in the ClinVar data set with additional LOF-intolerant genes affected. These results are in line with previous studies showing a second hit in about 2.5% of individuals.[18-20] Our study's limitation is that we cannot establish the exact frequency of additional affected LOF-intolerant genes in presumably monogenic disorders. For example, we extracted 6 P/LP deletions for the gene *KCNQ2* from Truty et al,[6] but 105 P/LP deletions were available from ClinVar, likely because of differences in the screening methodologies, ascertainment, and reporting. However, study by Truty et al[6] is a targeted sequencing study. It is possible that a subset of people with large chromosomal disorders were removed by a first-tier genetic test such as a microarray analysis. In addition, the CNV boundaries are restricted and limited to the genes assessed by the panel. In contrast, the ClinVar data set is more heterogeneous. It includes sequence-based CNV detection technologies and a large fraction of CNVs that have been identified in first-line genetic tests using traditional cytogenetic screening technologies. As an example, *UBE3A* ClinVar deletions are small compared with deletions in the study by Truty et al,[6] with most deletions being <100 base pair (bp). The mean ClinVar CNV size is 115,360 bp compared with 665,248 bp for CNVs from the study by Truty et al.[6] More disorder-specific research using different types of genetic screening methods is needed to establish better estimates. Although, we could show a significant correlation of the size of deletions with brain and non-neurological phenotypes, it was found to be weak. Our result should be interpreted in the context of its limitations owing to DECIPHER data being heterogeneous and incomplete. In addition, because defining complex phenotypes is challenging, we decided to inspect a small selection of epilepsy genes known to cause well-characterized brain phenotypes. More follow-up gene disorder–specific genotype–phenotype studies with deep phenotype data are needed. In addition, to what degree additional deleted LOF-intolerant genes affect the phenotype is not understood. Finally, it is unclear whether additional deleted LOF-intolerant genes should be reported, given our sparse knowledge of their effects because reporting them can lead to uncertainty and stress.

Despite the limitations, our study clearly showed that large deletions frequently affect additional LOF-intolerant genes along with the disease-associated genes. Current CNV interpretation guidelines integrate LOF-intolerant genes. However, the weight assigned to predicted haploinsufficient genes is relatively low, and only 1 such gene is considered even if additional predicted haploinsufficient genes are affected by the same CNV. In future, especially for patients with a haploinsufficiency disorder, a large deletion, and atypical phenotypic presentation, greater attention should be paid to all deleted LOF-intolerant genes within the interval of the deletion.

Data Availability

All relevant data and methods are reported in the article and the Supplemental Tables 1-4. The R script to reproduce the analyses is available at https://github.com/dlal-group/Oligogenic-LOF.

Acknowledgments

This work was supported by the National Institute of Health (grants: NIH 5U54NS108874, 1R01NS112499-01A1), the Center Without Walls on ion channel function in epilepsy (Channelopathy-associated Research Center, grant U54 NS108874), German Federal Ministry for Education and Research (Bundesministerium für Bildung und Forschung, Treat-ION, 01GM1907D), and Dravet Syndrome Foundation research grant to D.L.

Author Information

Conceptualization: A.S., D.L.; Data Curation: A.S.; Formal Analysis: A.S., E.P.-P., M.M., M.M.H.; Investigation: A.S., E.P.-P., M.M., M.M.H.; Methodology: A.S., D.L.; Project Administration: A.S., D.L.; Supervision: D.L.; Visualization: A.S.; Writing-original draft: A.S., C.L., D.L.; Writing-review and editing: A.S., E.P.-P., M.M., M.M.H., C.L., D.L.

Ethics Declaration

This study contains previously published data by Truty et al.[6] In their Ethics Declaration Truty et al[6] declare "data from genes unrelated to the presenting clinical phenotype were de-identified for analysis of baseline copy number variants under institutional review board approval (Western Institutional Review Board, 20161796)."

Conflict of Interest

D.L. consults for Encoded Therapeutics Inc. All other authors declare no conflicts of interest.

Additional Information

The online version of this article (https://doi.org/10.1016/j.gim.2021.10.026) contains supplemental material, which is available to authorized users.

References

1. United States Congress. Rare Diseases Act of 2002. Published November 6, 2002. https://www.govinfo.gov/content/pkg/PLAW-107publ280/html/PLAW-107publ280.htm. Accessed August 10, 2020.
2. Yehia L, Seyfi M, Niestroj LM, et al. Copy number variation and clinical outcomes in patients with germline PTEN mutations. *JAMA Netw Open*. 2020;3(1):e1920415. http://doi.org/10.1001/jamanetworkopen.2019.20415.
3. Liu Z, Zhu L, Roberts R, Tong W. Toward clinical implementation of next-generation sequencing-based genetic testing in rare diseases: where are we? *Trends Genet*. 2019;35(11):852–867. http://doi.org/10.1016/j.tig.2019.08.006.
4. Adams DR, Eng CM. Next-generation sequencing to diagnose suspected genetic disorders. *N Engl J Med*. 2018;379(14):1353–1362. http://doi.org/10.1056/NEJMra1711801.
5. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536(7616):285–291. http://doi.org/10.1038/nature19057.
6. Truty R, Paul J, Kennemer M, et al. Prevalence and properties of intragenic copy-number variation in Mendelian disease genes. *Genet Med*. 2019;21(1):114–123. http://doi.org/10.1038/s41436-018-0033-5.
7. Landrum MJ, Lee JM, Benson M, et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res*. 2018;46(D1):D1062–D1067. http://doi.org/10.1093/nar/gkx1153.
8. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26(6):841–842. http://doi.org/10.1093/bioinformatics/btq033.
9. Lindy AS, Stosser MB, Butler E, et al. Diagnostic outcomes for genetic testing of 70 genes in 8565 patients with epilepsy and neurodevelopmental disorders. *Epilepsia*. 2018;59(5):1062–1071. http://doi.org/10.1111/epi.14074.
10. Truty R, Patil N, Sankar R, et al. Possible precision medicine implications from genetic testing using combined detection of sequence and intragenic copy number variants in a large cohort with childhood epilepsy. *Epilepsia Open*. 2019;4(3):397–408. http://doi.org/10.1002/epi4.12348.
11. Firth HV, Richards SM, Bevan AP, et al. DECIPHER: database of chromosomal imbalance and phenotype in humans using Ensembl resources. *Am J Hum Genet*. 2009;84(4):524–533. http://doi.org/10.1016/j.ajhg.2009.03.010.
12. Köhler S, Carmody L, Vasilevsky N, et al. Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources. *Nucleic Acids Res*. 2019;47(D1):D1018–D1027. http://doi.org/10.1093/nar/gky1105.
13. Aguirre M, Rivas MA, Priest J. Phenome-wide burden of copy-number variation in the UK Biobank. *Am J Hum Genet*. 2019;105(2):373–383. http://doi.org/10.1016/j.ajhg.2019.07.001.
14. R Core Team. R: A language and environment for statistical computing. The R Foundation. Published October 10, 2020. https://www.r-project.org/. Accessed November 10, 2020.
15. Rees E, Kendall K, Pardiñas AF, et al. Analysis of intellectual disability copy number variants for association with schizophrenia. *JAMA Psychiatry*. 2016;73(9):963–969. http://doi.org/10.1001/jamapsychiatry.2016.1831.
16. Schaefer GB, Mendelsohn NJ, Professional Practice, Guidelines Committee. Clinical genetics evaluation in identifying the etiology of autism spectrum disorders: 2013 guideline revisions. *Genet Med*. 2013;15(5):399–407. Published correction appears in *Genet Med*. 2013;15(8):669. https://doi.org/10.1038/gim.2013.32.
17. Kessi M, Xiong J, Wu L, et al. Rare copy number variations and predictors in children with intellectual disability and epilepsy. *Front Neurol*. 2018;9:947. http://doi.org/10.3389/fneur.2018.00947.
18. Tang CS-M, Dattani S, So MT, et al. Actionable secondary findings from whole-genome sequencing of 954 East Asians. *Hum Genet*. 2018;137(1):31–37. http://doi.org/10.1007/s00439-017-1852-1.
19. Chen W, Li W, Ma Y, et al. Secondary findings in 421 whole exome-sequenced Chinese children. *Hum Genomics*. 2018;12(1):42. http://doi.org/10.1186/s40246-018-0174-2.
20. Haer-Wigman L, van der Schoot V, Feenstra I, et al. 1 in 38 individuals at risk of a dominant medically actionable disease. *Eur J Hum Genet*. 2019;27(2):325–330. http://doi.org/10.1038/s41431-018-0284-2.