

BRIEF COMMUNICATION



The phenotypic spectrum associated with loss-of-function variants in monogenic epilepsy genes in the general population

Victoria Smuk ¹, Javier A. López-Rivera ^{1,2}, Costin Leu ^{1,3,4} and Dennis Lal^{1,4,5,6}✉

© The Author(s), under exclusive licence to European Society of Human Genetics 2022

Variants in monogenic epilepsy genes can cause phenotypes of varying severity. For example, pathogenic variants in the *SCN1A* gene can cause the severe, sporadic, and drug-resistant Dravet syndrome or the milder familiar GEFS + syndrome. We hypothesized that coding variants in epilepsy-associated genes could lead to other disease-related phenotypes in the general population. We selected 127 established monogenic epilepsy genes and explored rare loss-of-function (LoF) variant associations with 3700 phenotypes across 281,850 individuals from the UK Biobank with whole-exome sequencing data. For 5.5% of epilepsy genes, we found significant associations of LoF variants with non-epilepsy phenotypes, mostly related to mental health. These findings suggest that LoF variants in epilepsy genes are associated with neurological or psychiatric phenotypes in the general population. The evidence provided may warrant further research and genetic screening of patients with atypical presentation and inform clinical care of comorbid disorders in individuals with monogenic epilepsy forms.

European Journal of Human Genetics (2023) 31:243–247; <https://doi.org/10.1038/s41431-022-01211-w>

INTRODUCTION

Monogenic epilepsies are a heterogeneous collection of rare severe developmental epileptic encephalopathies characterized by a combination of frequent epileptic activity with developmental impairment [1]. Large-scale genetic screens helped delineate the phenotypic spectrum of monogenic epilepsies and identified >150 causal genes [1–3]. Many of these genes encode ion channel proteins and play a role in neuronal excitability and action potential generation [1]. Individuals carrying pathogenic variants in the same monogenic epilepsy gene can exhibit varying severity of their phenotype with potentially different comorbid disorders (e.g., autism, ataxia, intellectual disability, or non-neurological disorders) [1].

Several established monogenic epilepsy genes cause mild to severe epilepsy forms [4]. Particularly, loss-of-function (LoF) variants in *SCN1A* are the most frequent cause of developmental epileptic encephalopathies [4]. Concurrently, other variants in *SCN1A* confer clinically non-actionable risk for common forms of epilepsy with typically milder outcomes [5]. Many other epilepsy genes cause comorbid disorders that significantly impact the quality of life and life expectancy of affected individuals [6]. However, not every patient expresses the full spectrum of comorbidities, and for many well-established epilepsy genes, new phenotypes are discovered years after the first association with epilepsy [7, 8].

Genotype-phenotype associations for epilepsy genes are typically performed through genetic testing for people with a neurological disease, leading to a potential strong ascertainment bias. We hypothesized that coding variants in epilepsy-associated

genes may lead to other recognizable clinical traits in the general population not enriched for individuals with epilepsy. Therefore, we sought to identify associations between established monogenic epilepsy genes and non-epilepsy but disease-related phenotypes in the general population represented by the UK Biobank (UKB) [9].

METHODS

To select a high-confidence set of established monogenic epilepsy genes, we merged three studies [1–3] (Supplementary Table 1). Variant level genotype and association data for 3700 phenotypes were obtained for all selected epilepsy genes using GeneBass [9] (accessed in August 2021), a resource of rare variant association statistics from single-variant and gene-based tests in 281,850 individuals with whole-exome sequencing (WES) data from the UKB project. The UKB provides medical and genetic data for over 500,000 participants living in the UK [10]. We collected association statistics from SKAT-O, gene-based burden, and single-variant association tests of LoF variants (nonsense, frameshift, splice-site) in all 127 high-confidence epilepsy genes. SKAT-O, gene-based burden, and single-variant tests were performed using a generalized mixed-model as implemented in SAIGE-GENE that can account for population structure and sample relatedness in large cohorts [11]. We explored phenotype associations only for LoF variants, as their functional consequence is more robust to interpret than missense variants. The gene-based burden test provides the highest power when all variants are causal and have the same effect direction, while SKAT-O retains power when only a proportion of the tested variants are causal and in the presence of trait-increasing and trait-decreasing variants [12]. To identify significant associations, we filtered the dataset according to the Bonferroni-corrected thresholds for significance as defined by the developers ($\alpha = 2.5 \times 10^{-8}$ for SKAT-O;

¹Genomic Medicine Institute, Lerner Research Institute, Cleveland Clinic, Cleveland, OH, USA. ²Department of Molecular Medicine, Cleveland Clinic Lerner College of Medicine, Case Western Reserve University, Cleveland, OH, USA. ³Department of Clinical and Experimental Epilepsy, Institute of Neurology, University College London, London, UK. ⁴Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA. ⁵Epilepsy Center, Neurological Institute, Cleveland Clinic, Cleveland, OH, USA. ⁶Cologne Center for Genomics, University of Cologne, Cologne, NRW, Germany. ✉email: lald@ccf.org

Received: 17 February 2022 Revised: 17 August 2022 Accepted: 3 October 2022

Published online: 17 October 2022

$\alpha = 6.7 \times 10^{-7}$ for burden tests). The associations were categorized by trait type. We then explored all significant associations to identify LoF variants driving the signals.

RESULTS

We generated a list of 127 high-confidence epilepsy genes from the three selected studies [1–3]. This dataset of 127 genes had 786–2903 genotype-phenotype associations per gene, each with two test statistics, equaling 485,593 genotype-phenotype associations. After multiple testing corrections, for seven of 127 epilepsy genes (5.5%), we found exome-wide significant associations with non-epilepsy phenotypes (Fig. 1A, B). Altogether, there were ten significant associations involving continuous traits only (Table 1). Continuous phenotypes are known to provide more power for association studies than categorical phenotypes [13]. Five associations (50%) in four genes were related to mental health phenotypes based on responses from a subset of 157,366 UKB participants who completed a mental health questionnaire (detailed description of phenotypes in Supplementary Table 2) [9]. Variants in *SLC25A12*, *ADSL*, *LMNB2*, and *GABRA1* were associated with mental health phenotypes related to mood and alcohol usage disorders (Fig. 1C). The significant associations were between LoF variants in *SLC25A12* and “recent changes in speed or amount of moving or speaking” ($P_{\text{SKAT-O}} = 3.00 \times 10^{-15}$) and “recent lack of interest or pleasure in doing things” ($P_{\text{SKAT-O}} = 1.49 \times 10^{-10}$); *ADSL* and “frequency of failure to fulfill normal expectations due to drinking alcohol in last year” ($P_{\text{SKAT-O}} = 1.57 \times 10^{-13}$); *LMNB2* and “recent changes in speed or amount of moving or speaking” ($P_{\text{SKAT-O}} = 5.26 \times 10^{-12}$); and *GABRA1* and “frequency of memory loss due to drinking alcohol in last year” ($P_{\text{SKAT-O}} = 7.52 \times 10^{-9}$). We identified one association related to brain morphology; between LoF variants in *GRIN2D* and “volume of gray matter in the right frontal medial cortex” ($P_{\text{BURDEN}} = 2.63 \times 10^{-7}$). Four associations were related to blood assay results. *CASR* was associated with calcium blood levels and *CHD2* with lymphocyte percentage ($P_{\text{SKAT-O}} = 9.99 \times 10^{-9}$), neutrophil percentage ($P_{\text{BURDEN}} = 1.29 \times 10^{-7}$), and lymphocyte count ($P_{\text{BURDEN}} = 5.55 \times 10^{-9}$). Two of ten identified gene-phenotype associations (SKAT-O or burden tests) had only one LoF variant with a nominal P value in the single-variant-based tests (SAIGE single-variant tests). Variants contributing to each association are detailed in Fig. 1C and Supplementary Table 3.

DISCUSSION

There has been an increasing interest in the phenotypic variability associated with monogenic epilepsies [4]. We used gene-based association summary statistics for 3700 phenotypes in 281,850 individuals from the general population to screen 127 established monogenic epilepsy genes for association with non-epilepsy phenotypes. We identified seven genes associated with disease-related phenotypes in the general population (*GRIN2D* with brain morphology; *CHD2* with immune function; *SLC25A12*, *LMNB2*, *ADSL*, and *GABRA1* with mental health; *CASR* with calcium metabolism).

Gain-of-function (GoF) variants in *GRIN2D* cause severe developmental epileptic encephalopathy [14]. Our study provides evidence that haploinsufficiency of *GRIN2D* is associated with gray matter volume in a general population cohort. Several GRIN genes are reported as dosage-sensitive, and both GoF and LoF variants are implicated in GRIN-related disorders [15]. Previously, variants in *GRIN1* and *GRIN2B* were found to cause cortical development malformations (i.e., polymicrogyria) [7, 15]. Our study is in line with evidence that LoF variants in GRIN genes cause brain malformations in the rodent knockdown model; however, human malformations generally represent GoF pathophysiology [7].

We found three of the seven pleiotropic epilepsy gene-phenotype associations add to a body of evidence for heterogeneous clinical

effects of these genes. LoF variants in *CHD2* were significantly associated with types of white blood cell counts. Truncating mutations in *CHD2* cause a heterogeneous group of epilepsies varying in severity, from febrile seizures plus to more severe epileptic encephalopathies, including a Dravet-like syndrome and Lennox-Gastaut syndrome [8]. LoF variants in *CHD2* have also been found to cause non-neurological cancer phenotypes, including chronic lymphocytic leukemia [16]. The evidence provided by our study may indicate patients with *CHD2*-related epilepsy have higher risk for cancers like lymphomas. However, further research is needed on this topic. In line with previous findings, we found that LoF variant burden in *SLC25A12* and *LMNB2* were associated with non-epilepsy phenotypes. Variants in *SLC25A12* can cause mitochondrial dysfunction, resulting in neurological disease. Mitochondrial dysfunction has a negative impact on synaptic plasticity and neurogenesis associated with depression [17]. We found LoF variants in *SLC25A12* associated with the UKB phenotype “recent lack of interest or pleasure in doing things” or anhedonia, a core phenotype of major depressive disorder. The pathogenicity of *LMNB2* in disease development has not been studied extensively. However, dysfunction of *LMNB2* results in diseases with various inheritance patterns, namely autosomal recessive progressive myoclonus epilepsy and autosomal dominant microcephaly and lipodystrophy [18]. Individuals with *LMNB2*-related progressive myoclonus epilepsy suffer from progressive ataxia and poor gait [18]. We identified a significant association between variants in *LMNB2* and “recent changes in speed or amount of moving or speaking”, pointing to a deleterious effect of LoF variants in *LMNB2* potentially independent of the seizure phenotype.

We found for two of seven epilepsy genes an association with previously known phenotypes (*GABRA1* with alcohol usage [19]; and *CASR* with calcium metabolism that plays a vital role in brain development, regulation of neuronal excitability, and synaptic transmission [20]). Finally, the association of the LoF variant burden in *ADSL* with mental health issues related to alcohol usage is new.

One limitation of our study is the large sample size of the UKB cohort that enables the identification of small gene-level effects as statistically significant, especially for continuous phenotypes. Similarly, risk-increasing variants identified in few or single carriers will be of low clinical significance due to the very low frequency in the general population. However, such variants may be important for some carriers and place them on the extreme end of the trait distribution. Overall, a more detailed clinical follow-up of individual variant carriers is needed to understand the implications of our findings. In addition, the association analysis is based on WES data only from middle-aged individuals of European ancestry. Therefore, generalization of these findings to individuals of other ancestries and age groups is limited. Also, the 3700 phenotypes in the UKB are an extensive selection but may not encompass all relevant phenotypes for the investigated monogenic epilepsy genes. The mental health phenotypes are from questionnaires; thus, the symptoms are subjective and nonspecific. However, although health surveys are subject to biases, they can still be used for association estimates [10]. Exploration of WES data in additionally available biobanks is needed to validate our findings. Finally, our results should be interpreted considering the UKB doesn't represent a broad distribution across all health outcomes and is enriched with healthy volunteers [10]. The healthy volunteer selection bias is particularly strong for neurological and psychiatric disorders, where disorder status may influence participation in research [10].

Overall, our study adds to our understanding of the phenotypic spectrum of several monogenic epilepsy genes and shows they are likely to cause recognizable clinical traits rather than severe epilepsy forms in otherwise healthy individuals. The evidence provided by the present study may help us better elucidate phenotypic heterogeneity with monogenic epilepsy-associated genes, promote further research and genetic screening of patients

Genotype - phenotype associations

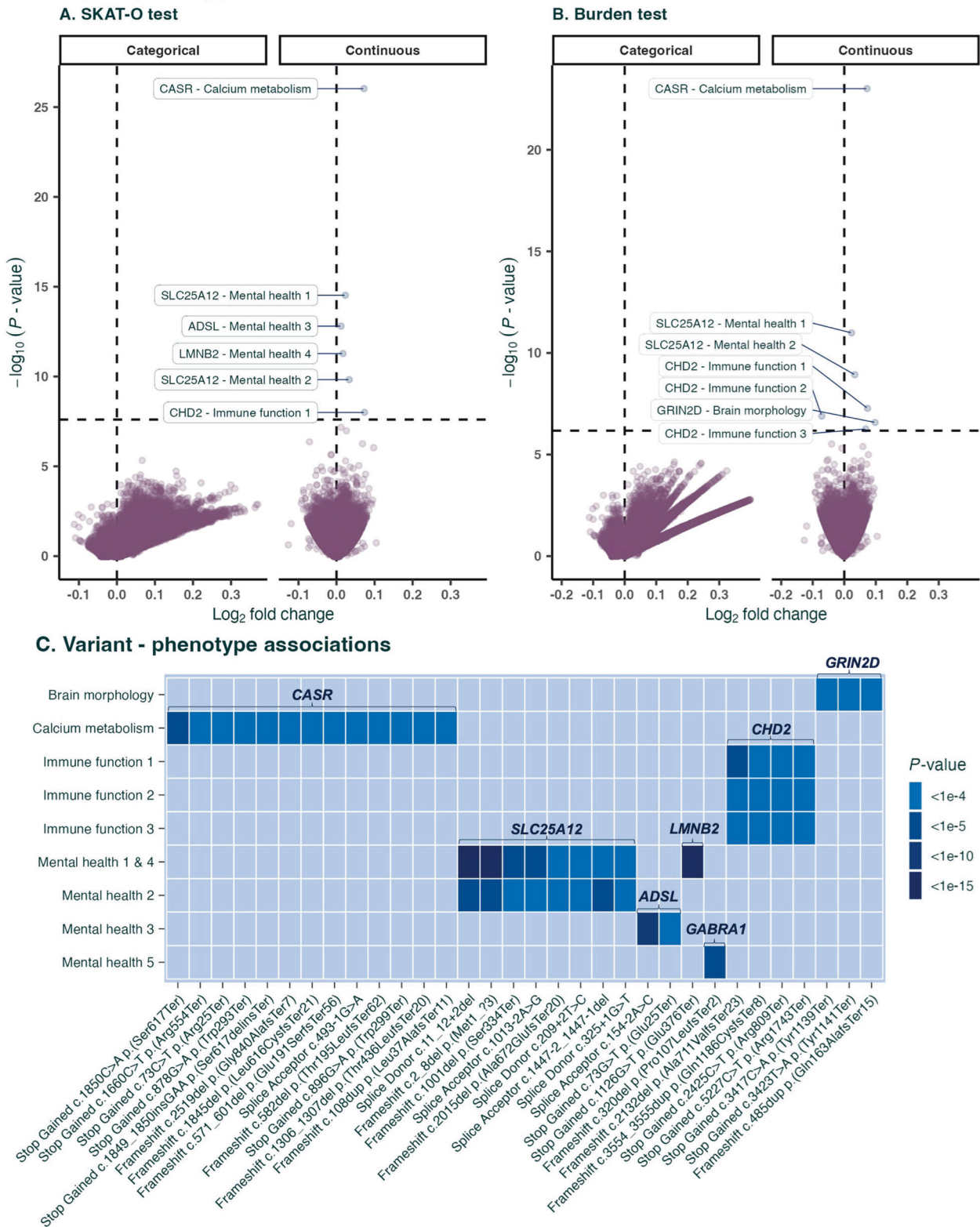


Fig. 1 Significant genotype-phenotype associations in epilepsy-associated genes. A Volcano plot illustrating the significant associations identified by the gene-based SKAT-O test. The horizontal dashed line indicates the threshold for a significant association ($\alpha = 2.5 \times 10^{-8}$). The significantly associated genes are labeled according to the phenotype category in GeneBass (details listed in Table 1). **B** Volcano plot illustrating the significant associations identified by the gene-based burden test. The horizontal dashed line indicates the threshold for a significant association ($\alpha = 6.7 \times 10^{-7}$). The significantly associated genes are labeled according to the phenotype category in GeneBass (as in Table 1). **C** Heatmap showing the association results of all LoF variants with single-variant P value ≤ 0.05 that were collapsed for each gene (human genome build GRCh38). **GABRA1* does not appear in (A, B) as a beta value was not available.

Table 1. Summary of significant genotype-phenotype associations.

Gene	GeneBass phenotype category	UKB phenotype description	P value (SKAT-O)	P value (burden)	Beta	Reported disorders in the literature
<i>CASR</i>	Calcium metabolism	Calcium blood biochemistry	9.36×10^{-27}	9.79×10^{-24}	0.072	Calcium Metabolism disorders (PMID: 32386559) Hyperparathyroidism (PMID: 26710757) Juvenile myoclonic epilepsy (PMID: 31611775)
<i>SLC25A12</i>	Mental health 1	Recent changes in speed/amount of moving or speaking	2.99×10^{-15}	1.03×10^{-11}	0.023	Early infantile epileptic encephalopathy-39, Neurometabolic epilepsy (PMID: 32700846, 31766059) Autism (PMID: 34115189)
	Mental health 2	Recent lack of interest or pleasure in doing things	1.49×10^{-10}	1.18×10^{-9}	0.033	
<i>ADSL</i>	Mental health 3	Frequency of failure to fulfill normal expectations due to drinking alcohol in last year	1.57×10^{-13}	2.20×10^{-3}	0.012	Adenylosuccinase deficiency (PMID: 33648541) Metabolic epilepsy (PMID: 32681428) Autism (PMID: 34121999)
<i>LMMB2</i>	Mental health 4	Recent changes in speed/amount of moving or speaking	5.26×10^{-12}	3.28×10^{-3}	0.017	Progressive myoclonus epilepsy (PMID: 34466237) Microcephaly (PMID: 33033404) Lipodystrophy (PMID: 17467691)
<i>GABRA1</i>	Mental health 5	Frequency of memory loss due to drinking alcohol in last year	7.52×10^{-9}	NA	NA	Juvenile myoclonic epilepsy (PMID: 26918889) Alcohol dependence (PMID: 29673739)
<i>CHD2</i>	Immune function 1	Lymphocyte percentage	9.99×10^{-9}	5.30×10^{-8}	0.074	Epilepsy of infancy with migrating focal seizures (PMID: 31677157) Chronic lymphocytic leukemia (PMID: 26031915)
	Immune function 2	Neutrophil percentage	4.35×10^{-7}	1.29×10^{-7}	-0.071	
	Immune function 3	Lymphocyte count	5.57×10^{-7}	5.55×10^{-7}	0.068	
<i>GRIN2D</i>	Brain morphology	Volume of gray matter in Frontal Medial Cortex (right)	9.18×10^{-7}	2.63×10^{-7}	0.098	Developmental and epileptic encephalopathy (PMID: 31918992)

Shown are all genes that surpassed the threshold for statistical significance after Bonferroni correction for the SKAT-O ($\alpha = 2.5 \times 10^{-8}$) or the burden tests ($\alpha = 6.7 \times 10^{-7}$). Positive beta values represent risk-increasing joint effects of the tested variants in one gene, while negative beta values represent risk-decreasing joint effects.

with atypical presentation, and inform clinical care of comorbid disorders in individuals with monogenic epilepsies.

DATA AVAILABILITY

The data analyzed during this study can be found on GeneBass (<https://genebass.org/>) [9].

REFERENCES

- Wang J, Lin ZJ, Liu L, Xu HQ, Shi YW, Yi YH, et al. Epilepsy-associated genes. *Seizure* 2017;44:11–20.
- Truty R, Patil N, Sankar R, Sullivan J, Millichap J, Carvill G, 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;4:397–408.
- Lindy AS, Stosser MB, Butler E, Downtain-Pickersgill C, Shanmugham A, Retterer K, et al. Diagnostic outcomes for genetic testing of 70 genes in 8565 patients with epilepsy and neurodevelopmental disorders. *Epilepsia* 2018;59:1062–71.
- Helbig I, Abou Tayoun AN. Understanding Genotypes and Phenotypes in Epileptic Encephalopathies. *Mol Syndromol*. 2016;7:172–81.
- The International League Against Epilepsy Consortium on Complex Epilepsies. Genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. *Nat Commun*. 2018;9:5269.
- Noebels JL. Single-Gene Determinants of Epilepsy Comorbidity. *Cold Spring Harb Perspect Med*. 2015;5:a022756.
- Smith RS, Walsh CA. Ion Channel Functions in Early Brain Development. *Trends Neurosci*. 2020;43:103–14.
- De Maria B, Balestrini S, Mei D, Melani F, Pellacani S, Pisano T, et al. Expanding the genetic and phenotypic spectrum of *CHD2* -related disease: from early neurodevelopmental disorders to adult-onset epilepsy. *Am J Med Genet*. 2021;188:522–33.
- Karczewski KJ, Solomonson M, Chao KR, Goodrich JK, Tiao G, Lu W, et al. Systematic single-variant and gene-based association testing of thousands of phenotypes in 394,841 UK Biobank exomes. *Cell Genom*. 2022;2:100168.
- Davis KAS, Coleman JRI, Adams M, Allen N, Breen G, Cullen B, et al. Mental health in UK Biobank – development, implementation and results from an online questionnaire completed by 157 366 participants: a reanalysis. *BJPsych Open*. 2020;6:e18.
- Zhou W, Zhao Z, Nielsen JB, Fritsche LG, LeFaive J, Gagliano Taliun SA, et al. Scalable generalized linear mixed model for region-based association tests in large biobanks and cohorts. *Nat Genet*. 2020;52:634–9.
- Lee S, Abecasis GR, Boehnke M, Lin X. Rare-variant association analysis: study designs and statistical tests. *Am J Hum Genet*. 2014;95:5–23.
- Cohen J. The Cost of Dichotomization. *Appl Psychol Meas*. 1983;7:249–53.
- Li D, Yuan H, Ortiz-Gonzalez XR, Marsh ED, Tian L, McCormick EM, et al. *GRIN2D* Recurrent De Novo Dominant Mutation Causes a Severe Epileptic Encephalopathy Treatable with NMDA Receptor Channel Blockers. *Am J Hum Genet*. 2016;99:802–16.
- Fry AE, Fawcett KA, Zelnik N, Yuan H, Thompson BAN, Shemer-Meir L, et al. De novo mutations in *GRIN1* cause extensive bilateral polymicrogyria. *Brain* 2018;141:698–712.
- Wilson MM, Henshall DC, Byrne SM, Brennan GP. *CHD2*-Related CNS Pathologies. *IJMS*. 2021;22:588.
- Babenko VN, Smagin DA, Galyamina AG, Kovalenko IL, Kudryavtseva NN. Altered *Slc25* family gene expression as markers of mitochondrial dysfunction in brain regions under experimental mixed anxiety/depression-like disorder. *BMC Neurosci*. 2018;19:79.
- Farajzadeh Valilou S, Karimzad Hagh J, Salimi Asl M, Abdi Rad I, Edizadeh M, Pooladi A. A novel biallelic *LMNB2* variant in a patient with progressive myoclonus epilepsy and ataxia: a case of laminopathy. *Clin Case Rep*. 2021;9.
- Sahni S, Tickoo M, Gupta R, Vaswani M, Ambekar A, Grover T, et al. Association of serotonin and GABA pathway gene polymorphisms with alcohol dependence: a preliminary study. *Asian J Psychiatry*. 2019;39:169–73.
- Gilsoul M, Grisar T, Delgado-Escueta AV, de Nijs L, Lakaye B. Subtle Brain Developmental Abnormalities in the Pathogenesis of Juvenile Myoclonic Epilepsy. *Front Cell Neurosci*. 2019;13:433.

ACKNOWLEDGEMENTS

This research has been conducted using summary statistics generated from the UK Biobank resource (under application 26041 and 48511).

AUTHOR CONTRIBUTIONS

CL and DL designed the study. VS analyzed the data. VS, JALR, CL, and DL wrote the paper. CL and DL supervised the study. All authors interpreted the data and revised the paper.

FUNDING

This research did not receive any grant from the public, commercial, or not-for-profit funding agencies.

COMPETING INTERESTS

The authors declare no competing interests.

ETHICS APPROVAL

No ethical approval was required. The study used de-identified gene-based association summary statistics.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41431-022-01211-w>.

Correspondence and requests for materials should be addressed to Dennis Lal.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.