Sub-genic intolerance, ClinVar, and the epilepsies: A whole-exome sequencing study of 29,165 individuals

Epi25 Collaborative*

Summary

Both mild and severe epilepsies are influenced by variants in the same genes, yet an explanation for the resulting phenotypic variation is unknown. As part of the ongoing Epi25 Collaboration, we performed a whole-exome sequencing analysis of 13,487 epilepsy-affected individuals and 15,678 control individuals. While prior Epi25 studies focused on gene-based collapsing analyses, we asked how the pattern of variation within genes differs by epilepsy type. Specifically, we compared the genetic architectures of severe developmental and epileptic encephalopathies (DEEs) and two generally less severe epilepsies, genetic generalized epilepsy and non-acquired focal epilepsy (NAFE). Our gene-based rare variant collapsing analysis used geographic ancestry-based clustering that included broader ancestries than previously possible and revealed novel associations. Using the missense intolerance ratio (MTR), we found that variants in DEE-affected individuals are in significantly more intolerant genic sub-regions than those in NAFE-affected individuals. Only previously reported pathogenic variants absent in available genomic datasets showed a significant burden in epilepsy-affected individuals compared with control individuals, and the ultra-rare pathogenic variants associated with DEE were located in more intolerant genic sub-regions than variants associated with non-DEE epilepsies. MTR filtering improved the yield of ultra-rare pathogenic variants in affected individuals compared with control individuals. Finally, analysis of variants in genes without a disease association revealed a significant burden of loss-of-function variants in the genes most intolerant to such variation, indicating additional epilepsy-risk genes yet to be discovered. Taken together, our study suggests that genic and sub-genic intolerance are critical characteristics for interpreting the effects of variation in genes that influence epilepsy.

Introduction

Epilepsy is a clinical diagnosis in which the individual has an enduring predisposition to seizures. Although the most severe types most commonly begin in childhood with profound impact, epilepsies can begin at any age and have a cumulative incidence approaching 4%.[1–3] While the genetics of the epilepsies are complex, uncovering pathogenic variants can, in some cases, provide opportunities for targeted or precision medicines.[4,5] Whole-exome sequencing (WES) case-control studies have led to multiple insights into the epilepsies, such as the contribution of *de novo* variants in developmental and epileptic encephalopathy (DEE [MIM: 308350]), the role of the GABA pathway in genetic generalized epilepsy (GGE [MIM: 600669]), and the link between non-acquired focal epilepsy (NAFE [MIM: 604364, 245570]) and GATOR1 complex genes.[6–10] DEEs are a severe form of early onset, intractable epilepsy associated with developmental delay.[8,11–14] In contrast, GGE and NAFE, characterized by generalized seizures and focal seizures, respectively, are more common and generally less severe.[1,2,15–17] Yet, exome sequencing has revealed that a set of 43 genes typically associated with DEE also harbor ultra-rare variants in milder epilepsies.[7,9]

It is unknown how these variants cause such different epilepsy phenotypes despite being drawn from a set of shared genes and even from within the same gene. The likelihood of a gene's being associated with disease can be predicted *in silico*, in part, by a given gene's intolerance to functional variation in the general population.[18–20] Epilepsy-causing variants tend to be rare in the general population and located in the least tolerant genes.[7,9,18,21] While genic intolerance may help determine the likelihood of a gene-disease association, it does not clarify the differential impact of variants within the same gene.[22] Variants within the same gene may lead to widely different epilepsy phenotypes.[23–29] Predicting the differential effects of two variants within the same gene requires an understanding of sub-genic intolerance because different regions or domains will have varied importance for the protein's function and may therefore contribute differentially to disease phenotype or severity.[22] Consistent with this idea, distributions of disease mutations often cluster in specific genic sub-regions.[30] In general, epilepsy variants cluster in the most intolerant genic sub-regions.[22,31–33] The relationship between the severity of epilepsy caused by *SCN2A* variants and sub-genic intolerance has been explored,[32] but a more systematic study of the association of sub-genic intolerance and epilepsy severity has not been undertaken. Given that a single variant may lead to variable phenotypes,[34–37] we do not expect sub-genic intolerance to explain all severity variability, but a deeper investigation will add to our understanding of the complex sequelae of a single variant.

*Correspondence: jm4279@cumc.columbia.edu (Joshua E. Motelow), dg2875@cumc.columbia.edu (David B. Goldstein)
https://doi.org/10.1016/j.ajhg.2021.04.009.

The Epi25 Collaborative (Epi25) is the largest epilepsy exome analysis to date with more than 200 partners from 40 research cohorts contributing exome and phenotype data from more than 19,000 individuals with epilepsy (see web resources). The aspiration of the collaborative is that extensive exome data combined with accurate phenotypic data will allow for well-matched cohorts and clarify genotype-phenotype relationships in epilepsy, and Epi25 analyses have already yielded rich results for rare variants in the epilepsies.[9] A dataset of this magnitude and detail allows us to examine the presence of curated variants from a clinical database such as ClinVar.[38,39] Similarly, we are able to test for the burden of damaging variants in the ~15,000 genes not yet associated with Mendelian disease to detect the potential for epilepsy-gene discovery. Combining expansive genetic data from Epi25 and recently developed sub-genic intolerance metrics, we show that in a set of genes harboring missense variants in both milder and more severe epilepsies, variants in more severe epilepsies are preferentially located in less tolerant genic sub-regions. Furthermore, only ultra-rare (i.e., not found in a public database) pathogenic/likely pathogenic[40] ClinVar variants are increased in our cohort, and our sub-genic intolerance finding is replicated in these ultra-rare variants. Finally, there most likely remain undiscovered epilepsy-associated or epilepsy-risk genes among the genes most intolerant to loss-of-function variation.

Subjects and methods

Study design and participants

As described previously, we collected DNA and detailed phenotyping data on individuals with epilepsy from 40 sites in Europe, North America, Australasia, and Asia (Table S1).[9] Here, we analyzed individuals with DEEs (n = 2,007), GGE (also known as idiopathic generalized epilepsy; n = 5,771), and NAFE (n = 7,489), accounting for the first 3 years of enrollment in Epi25. A subset of the data is available on dbGaP: phs001489. Following sample quality control (QC), relatedness testing (see sample and variant QC), and clustering (see clustering), the combined epilepsy analysis included 13,171 affected individuals (1,782, 5,048, and 6,341 individuals with DEE, GGE, and NAFE, respectively) along with 14,100 control individuals (2,048 genomes and 12,052 exomes). In the included clusters in the individual epilepsy analyses, 1,835 individuals with DEE were compared to 13,978 control individuals, 5,303 individuals with GGE were compared to 15,677 control individuals, and 6,439 individuals with NAFE were compared to 15,678 control individuals. Control individuals were aggregated from local collections at the Institute for Genomic Medicine at Columbia University (IGM - Columbia University, New York, NY, USA). Control individuals who passed the same QC and who were not known to have phenotypes overlapping DEE, GGE, or NAFE or be related to a proband with epilepsy were analyzed following geographic ancestry clustering (Figure S1, Table S2).

Phenotyping procedures

As described previously, epilepsies were clinically diagnosed by epileptologists (see below for criteria DEEs, GGE, and NAFE) in accordance with the International League Against Epilepsy (ILAE) classification at the time of diagnosis and recruitment.[2,9] De-identified (non-PHI [protected health information]) phenotyping data were entered into the Epi25 data repository (hosted at the Luxembourg Centre for Systems Biomedicine) via online case record forms based on the RedCAP platform. De-identified data for subjects of previous coordinated efforts with phenotyping (e.g., the Epilepsy Phenome/Genome Project[41] and the EpiPGX Project, see web resources) that were already entered into a database were accessed and transferred to the new platform. Phenotyping data underwent review for uniformity among sites and QC, and inconsistencies were reviewed by the phenotyping committee.

Epilepsy definitions

Epilepsy diagnoses and classification for Epi25 have been described previously.[9] Briefly, DEE diagnosis required severe refractory epilepsy of unknown etiology with developmental plateau or regression and epileptiform features on electroencephalogram (EEG). Exclusion criteria included epileptogenic lesions on MRI. GGE diagnosis required a history of generalized seizure types with generalized epileptiform discharges on EEG. Exclusion criteria include focal seizures, moderate-to-severe intellectual disability, and epileptogenic lesions found on neuroimaging (when available). Diagnosis of NAFE required a history of focal seizures with either focal epileptiform discharges or normal findings on EEG. Exclusion criteria included neuroimaging lesions (except hippocampal sclerosis), a history of generalized seizures, and moderate-to-severe intellectual disability.

Informed consent

Adult subjects or the legal guardian for enrolled children signed informed consent at participating centers per the ethical requirements of the local rules at the time of enrollment.[9] The consent must not exclude data sharing to be included in the study. Consent forms for samples collected after January 25, 2015 required specific language according to the National Institutes of Health's Genomic Data Sharing Policy (see web resources). For control individuals, protocols were approved by Columbia University's institutional review board and participants provided informed consent for the use of DNA in genetic research.

Next-generation sequencing data generation

All Epi25 samples were sequenced at the Broad Institute of Harvard and the Massachusetts Institute of Technology (MIT) on the Illumina HiSeq X platform with the use of 151 bp paired-end reads. Exome capture was performed with Illumina Nextera Rapid Capture or TruSeq Rapid Exome enrichment kit (target size 38 Mb). FastQ files were transferred to the IGM.

Next-generation sequencing of control individuals was performed at the IGM or transfered to the IGM and was a mixture of whole-genome sequencing and whole-exome sequencing. Exomes were captured with multiple capture kits and sequenced according to standard protocols on Illumina's HiSeq 2000, HiSeq 2500, and NovaSeq 6000 (Illumina, San Diego, CA, USA) platform with 150 bp paired-end reads. Genomes were sequenced according to standard protocols on Illumina's HiSeq 2000, HiSeq 2500, and NovaSeq 6000 (Illumina, San Diego, CA, USA) platform.

Variant calling

Both affected individuals and control individuals were processed with the same IGM bioinformatic pipeline for variant calling.

Reads were aligned to human reference GRCh37 via DRAGEN (Edico Genome, San Diego, CA, USA)[42] and duplicates were marked with Picard (Broad Institute, Boston, MA, USA). Variants were called according to the Genome Analysis Toolkit (GATK - Broad Institute, Boston, MA, USA) Best Practices recommendations v.3.6.[43,44] Finally, variants were annotated with ClinEff[45] and custom annotations, including Genome Aggregation Database (gnomAD) v.2.1 frequencies,[20] regional-intolerance metrics,[31,32] *in silico* filters,[46] and ClinVar (as of 10/20/2020)[38,39] clinical annotation, were added via the IGM's in-house analysis tool for annotated variants (ATAV) platform.[47]

Sample and variant QC

Only samples with at least 90% of the consensus coding sequence (CCDS release 20)[48] covered at a minimum of 10×, ≤2% contamination levels according to VerifyBamID,[49] and single nucleotide variants (SNVs) and indels overlapping the Single Nucleotide Polymorphism database (dbSNP)[50] at least 85% and 80%, respectively, were included. We removed samples with a discordance between self-declared and sequence-derived gender to prevent phenotype-genotype mismatch. We used kinship-based inference for GWAS (KING) to detect related individuals and removed one of each pair that had an inferred relationship of second-degree or closer while favoring the inclusion of affected individuals over control individuals and well covered over poorly covered.[51]

We restricted analyses to variants within the CCDS inclusive of two base intronic extensions to accommodate canonical splice variants. All included variants had to fulfill the following criteria to be included: (1) at least 10× coverage of the site, (2) quality score (QUAL) ≥ 50, (3) genotype quality score (GQ) ≥ 20, (4) quality by depth score (QD) ≥ 5, (5) mapping quality score (MQ) ≥ 40, (6) read position rank sum score (RPRS) ≥ −3, (7) mapping quality rank sum score (MQRS) ≥ −10, (8) Fisher's strand bias score (FS) ≤ 60 (SNVs) or ≤ 200 (indels), (9) strand odds ratio (SOR) ≤ 3 (SNVs) or ≤ 10 (indels), (10) GATK Variant Quality Score Recalibration filter "PASS," and (11) alternate allele fraction for heterozygous calls ≥ 0.3. Known sequencing artifacts as described previously[52] as well as low-quality variants per Exome Aggregation Consortium,[53] gnomAD,[20] or the Exome Variant Server were excluded (see web resources).

Clustering

As previously described by Cameron-Christie and colleagues,[54] we performed principal-component analysis (PCA) for dimensionality reduction on a set of pre-defined variants to capture population structure. We applied the Louvain method of community detection with the first six principal components (PCs) as input to identify clusters within the data that reflect the geographic ancestry of the samples as previously described.[55,56] To check the quality of the clusters, we performed further dimensionality reduction by using the Uniform Manifold Approximation and Projection (UMAP)[57] on the first six PCs (Figures S1A–S1C) to disentangle geographic ancestry, which is then reflected in the cluster membership.[58,59] A neural-network pre-trained on samples with known geographic ancestry generated probability estimates for each of six groups (European, African, Latino, East Asian, South Asian, and Middle Eastern). We used a 95% probability cutoff to assign a geographic ancestry label to each sample. Samples that did not reach 95% for any of the ancestry groups were labeled "admixed" (Figure S1).

Clustering was performed on the combined epilepsies as previously described.[56] Clusters containing at least 20 affected individuals in each epilepsy type (DEE, GGE, and NAFE) and 20 controls were kept (Figure S1C, Table S3). Each epilepsy type/control group separately underwent clustering again to optimize ancestry matching for each epilepsy type (Figures S1D–S1L). The individual epilepsy clustering was used for individual epilepsy quantile-quantile plots (see quantile-quantile plots and genomic inflation factor λ, Figure 1), the analysis of common enrichment among DEE genes (Figure 2), and associated supplementary figures and tables. The combined epilepsy clustering was used for the combined epilepsy collapsing analysis, sub-genic comparisons, and ClinVar pathogenic/likely pathogenic analyses (Figures 3 and 4, control data in Figure 5) and associated supplementary figures and tables. The individual epilepsy clustering was also used to demonstrate potential for gene discovery (Figure 6) with associated supplementary figures and tables. All clusters underwent coverage harmonization (see coverage harmonization).

Coverage harmonization

As described previously,[52] coverage differences between affected individuals and control individuals introduce a bias because no variants can be called without sufficient coverage. To reduce the influence of coverage differences caused by different capture kits or sequencing depth in general, we used a site-based pruning approach and removed sites where the absolute difference in percentages of affected individuals compared to control individuals with at least 10× coverage was greater than 7.0%. Each cluster (see clustering) underwent independent coverage harmonization. This resulted in four sets of coverage maps (Figure S1).

Qualifying variant

In the context of collapsing analyses, qualifying variants have been defined in order to identify a set of variants that are enriched for real variant calls and variants with strong functional effects.[60] Here, we defined a qualifying variant (QV) as a variant passing both QC filters (see sample and variant QC) and model-specific filters (Table S4), such as variant effect filters, pathogenicity predictors, and internal and external minor allele frequency (MAF) filters. Variants could be drawn from three pools: (1) variants from Epi25 data and matched controls blinded to ClinVar status, (2) variants from Epi25 data and matched controls designated pathogenic/likely pathogenic (P/LP) in ClinVar as of 10/20/2020, or (3) all published P/LP ClinVar variants as of 10/20/2020. For analyses of variants in Epi25 data and matched controls blinded to ClinVar status (1) (Figures 1, 2, and 3, control data in Figures 5 and 6, Table 1), we applied the following filtering in addition to the variant QC filtering (see sample and variant QC): (1) all variants are "ultra-rare," meaning they are not found in any non-neuro gnomAD population; (2) we filtered all protein-truncating variants (PTVs) with loss-of-function transcript effect estimator (LOFTEE) to remove likely false-positive PTVs;[20] (3) we removed all variants located in regions with highly repetitive elements to reduce false-positive variants;[61] (4) we removed all variants in regions with a proportion expression across transcripts (pext) value less than 1/10 the maximum pext value for that gene because they are unlikely to affect translated mRNA;[62] and (5) we excluded variants with an internal allele frequency greater than 0.05% applied to the combined case-control call set by cluster excluding one allele to allow for clusters in which one allele might exceed that allele frequency threshold.[62] PTV effects included stop gain, frameshift, splice acceptor, and splice donor variants.

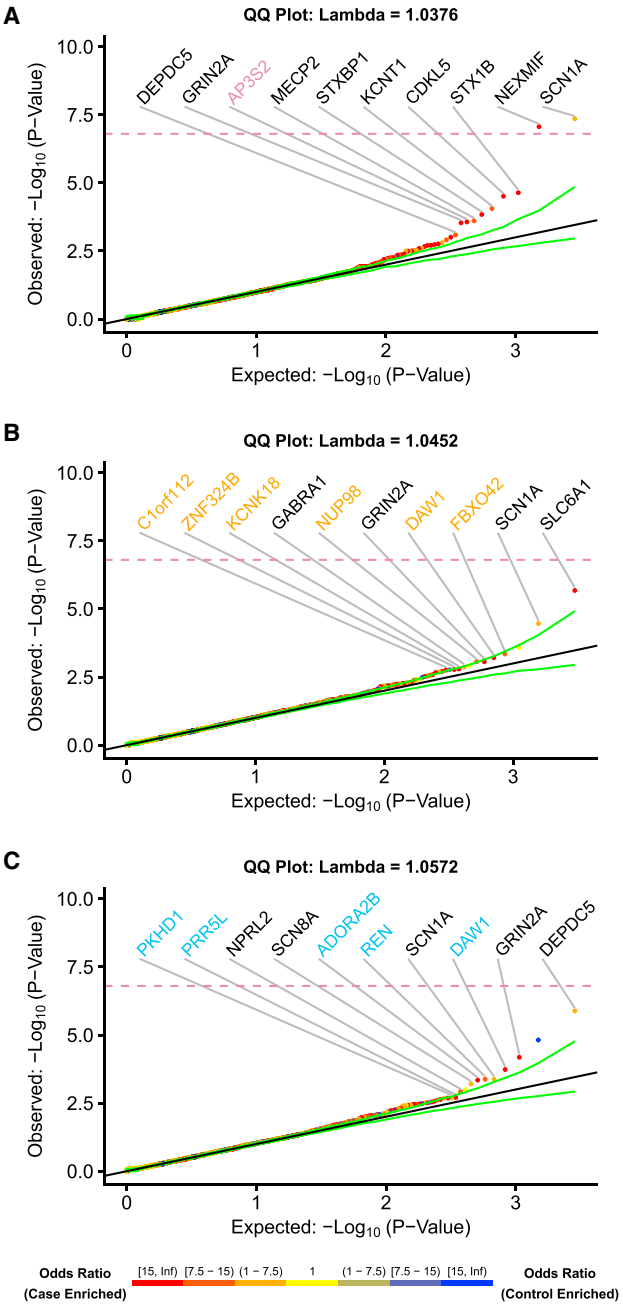


Figure 1. Quantile-quantile (QQ) plots for the protein-coding genes with at least one individual with epilepsy or control carrier
Qualifying variants were high-quality, ultra-rare variants with a predicted functional effect but restricting missense variants to REVEL $\geq$ 0.5 (when defined). We generated p values from the exact two-sided Cochran-Mantel-Haenszel (CMH) test by gene by cluster to indicate a different carrier status of affected individuals in comparison to control individuals. *SCN1A* ($p = 4.4 \times 10^{-8}$) and *NEXMIF* (previously known as *KIAA2022*, $p = 8.6 \times 10^{-8}$) achieved study-wide significance $p < 1.6 \times 10^{-7}$ after Bonferroni correction indicated by dashed line (see gene-based collapsing). (A) Developmental and epileptic encephalopathy (DEE)-affected individuals, (B) genetic generalized epilepsy (GEE)-affected individuals, and (C) non-acquired focal epilepsy (NAFE)-affected individuals. Top ten genes enriched among individuals with epilepsy are labeled. Point coloring determined by CMH odds ratio. Genes labeled in black are known epilepsy genes. Genes labeled in color are candidate epilepsy genes. The green lines represent the 95% confidence interval.

For P/LP variants found in Epi25 and matched controls (Figure 4) (2) and all published P/LP variants (Figure 5, non-control data) (3), no universal filtering was applied beyond variant QC. ClinVar variants could additionally be filtered by ClinVar "review status," which attempts to capture the level of review supporting the assertion of clinical significance for the variant with increasing number of "gold stars" from 0 to 4.[63–65]

In addition to the filtering applied above, we defined the following categories of missense variants to be utilized in the study. For "damaging" missense variants, REVEL[46] filter $\geq$ 0.5 (when defined) was applied. For "intolerant" missense variants, a missense tolerance ratio (MTR) filter $\leq$ 0.78 (when defined), which represents a variant in the most intolerant quartile of all regions in the exome to missense variation, was applied (see web resources).[32] To further enhance missense variants for those located in intolerant genic sub-regions, we utilized a separate model in which we added an exon-based localized intolerance model using Bayesian regression (LIMBR) percentile < 25. LIMBR is a sub-genic intolerance score previously shown to enhance selection for missense variants associated with DEEs.[31]

Gene-based collapsing

As described previously,[7,52,56] we performed gene-based collapsing to test whether there is a significant enrichment of affected individuals harboring a QV in a given gene compared to controls. For each gene within each cluster, we assigned an indicator variable (1/0 states) to each individual on the basis of the presence of at least one qualifying variant in the gene (state 1) or no qualifying variants in that gene (state 0) to create a gene-by-subject matrix for each cluster. From the collapsing matrices of the individual clusters, we extracted the number of affected individuals/control individuals with and without a QV per gene and used the exact two-sided Cochran-Mantel-Haenszel (CMH) test[66,67] to test for an association between disease status and QV status (Table S4) while controlling for cluster membership. Finally, we created quantile-quantile (QQ) plots (described below). We defined a study-wide Bonferroni multiplicity-adjusted significance threshold of $p < 1.6 \times 10^{-7}$ (0.05 / [18,650 CCDS genes × 17 non-synonymous models]).

The synonymous model was used as a putatively negative control (Figures S2 and S3A, Tables S4, S6–S8, and S16). Additional details for the 17 non-synonymous models can be found in Table S4. The top 200 ranked genes for each analysis can be found in the supplemental tables (Tables S6–S26). The membership of each gene in the following gene sets is also indicated: (D) 43 dominant genes associated with DEE in the Online Mendelian Inheritance in Man (OMIM, see web resources) (see gene set enrichment testing), (P) 101 dominant genes with epilepsy or related terms in its OMIM phenotype, (L) the 1,920 genes most intolerant to loss-of-function variation in the general population (see gene set enrichment testing), top 200 ranked genes in prior Epi25 DEE (D25), GGE (G25), or NAFE (N25) association analyses,[9] or top 300 ranked genes in prior GGE (G4K) or NAFE (N4K) Epi4K association analyses.[7] Epi4K was a large WES epilepsy project completed prior to Epi25.

Quantile-quantile plots and genomic inflation factor λ

We generated quantile-quantile (QQ) plots with empirical (permutation-based) expected probability distributions by using a previously described method.[7,52] For each collapsing model and cluster, the original case and control labels were randomly permuted,

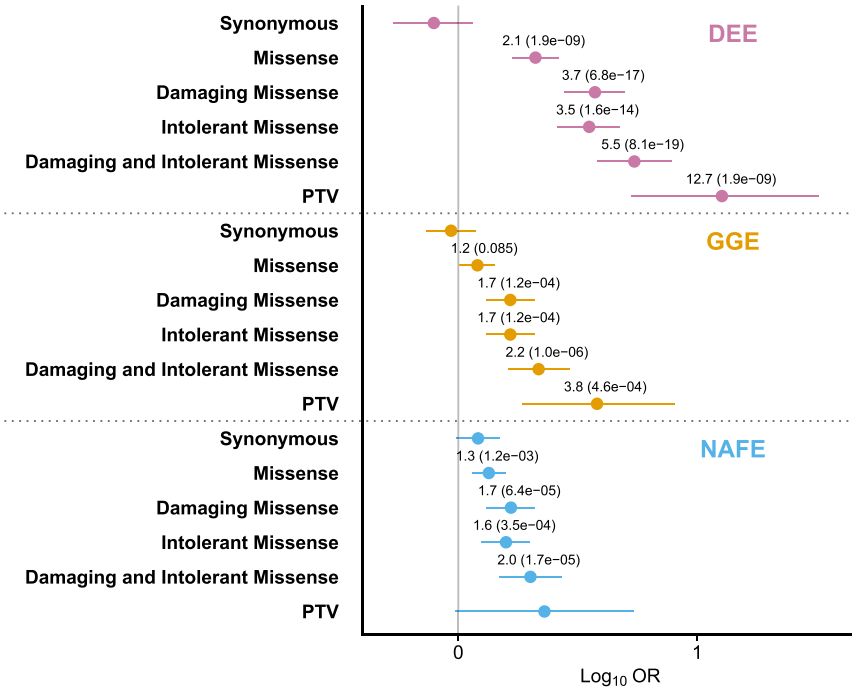


Figure 2. Gene set enrichment analysis shows indviduals with mild epilepsies enriched for rare variants in genes associated with severe epilepsies

Gene set burden testing with 24 genes drawn from the 43 OMIM epileptic encephalopathy phenotype series with dominant transmission by limiting to genes harboring damaging (REVEL ≥ 0.5) missense variants in all three epilepsies (see gene set enrichment testing, Table S5). All variants are ultra-rare (see subjects and methods). Pooled odds ratio, 95% confidence intervals, and FDR-corrected p value were generated from the exact two-sided Cochran-Mantel-Haenszel (CMH) test. Odds ratio and FDR-adjusted p values displayed for comparisons with unadjusted p value < 0.05. x axis displays the $\log_{10}$ of the odds ratio and confidence intervals. PTV, protein-truncating variants; "damaging," REVEL ≥ 0.5 (when defined); "intolerant," MTR ≤ 0.78 (when defined); DEE, developmental and epileptic encephalopathy; GGE, genetic generalized epilepsy; NAFE, non-acquired focal epilepsy.

while the rest of the gene-by-sample matrix was kept fixed. For each cluster, we extracted the number of newly sampled cases/controls with and without a QV per gene and used the CMH test to test for an association between case/control status (see gene-based collapsing) and QV status (see qualifying variant) while controlling for cluster membership. This process was repeated 1,000 times, and for each permutation, the p values were ordered. The mean of each rank-ordered estimate across the 1,000 permutations (i.e., the average 1st order statistic, the average 2nd order statistic, etc.) represents the empirical estimates of the expected ordered p values. We plotted the negative logarithm of the permutation-based expected distribution relative to the observed ordered statistic to get permutation-based QQ plots. We also used the permutation-based expected p values to estimate the genomic inflation factor λ on the basis of the regression method as described previously.[7,52] Genes labeled in black are known epilepsy-associated genes on the basis of manual review of the literature, while genes labeled in color are candidate epilepsy-associated genes.

Gene set enrichment testing

As described previously,[7] biologically informed gene sets can reveal important pathways or gene characteristics by aggregated signal across related genes (Table S5). We utilized the following gene sets (GS-1 to GS-6) informed by their OMIM disease associations, inheritance patterns, and genic intolerance.

(GS-1) 43 established dominant (e.g., autosomal dominant or x-linked dominant) DEE-associated genes drawn from OMIM Phenotypic Series PS308350 and PS617711 on 10/9/2020.
(GS-2) 24 genes drawn from the 43 genes in GS-1 for which in all three epilepsies have a damaging missense variant.
(GS-3) 101 established dominant genes associated with OMIM phenotypes containing epilepsy and epilepsy related terms on 02/16/2021.
(GS-4) 14 genes harboring ultra-rare missense variants associated with both DEE and with epilepsy but not DEE in ClinVar (*SZT2*, *SCN2A*, *SCN1A*, *HCN1*, *GABRA1*, *GABRG2*, *KCNQ3*, *SPTAN1*, *KCNT1*, *GRIN2B*, *GABRB3*, *CHD2*, *TBC1D24*, and *KCNQ2)* as of 10/20/2020.
(GS-5) 10 gene sets representing the genes without a confirmed disease phenotype in OMIM on 02/16/2021 (18,852 CCDS genes – 3,964 genes = 14,888 genes) distributed into 10 groups by their loss-of-function observed/expected upper bound fraction (LOEUF) decile were created.[20] LOEUF is the 90% upper bound of the confidence interval of the observed/expected ratio of predicted loss-of-function variants in gnomAD and can be used to bin genes into deciles of approximately 1,920 genes each.
(GS-6) 10 gene sets representing the genes without a confirmed phenotype in OMIM on 02/16/2021 (18,852 CCDS genes – 3,964 genes = 14,888 genes) distributed into 10 groups by their missense *Z* score were created.[19,20,68] Missense *Z* score captures the number of observed missense variants in a gene compared to the expected number of missense variants in the general population. The score was used to bin genes into deciles of approximately 1,920 genes each.

For a gene set analysis, we extracted the number of affected individuals/control individuals with and without at least one QV among any of the genes in the gene set and used the exact two-sided CMH test[66,67] to test for an association between disease status and QV status while controlling for cluster membership. To examine association with LOEUF deciles (Figure 6), we only used control individuals without a disease association in our database ("controls" and "healthy family members") (Table S2). We used a false discovery rate (FDR) correction for multiple comparisons. We performed 123 CMH tests to determine odds ratios for gene set enrichment testing and defined a significant enrichment at FDR < 0.05. For forest plots, odds ratios and p values were displayed for associations with an unadjusted p value < 0.05.

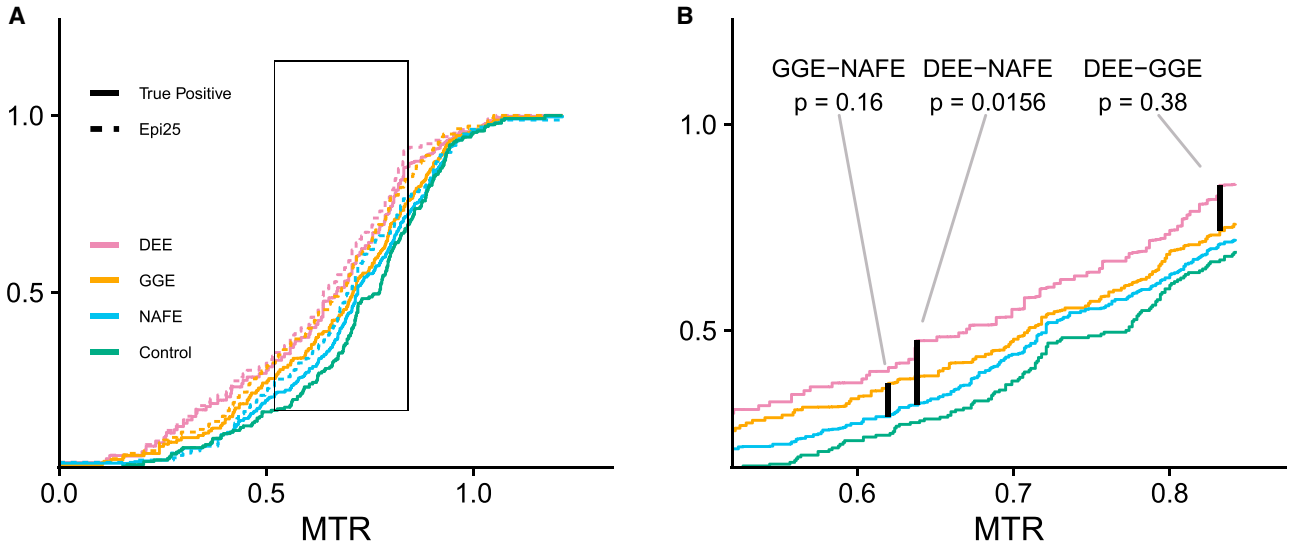


Figure 3. Sub-genic intolerance analysis reveals variants associated with DEE are located in more intolerant genic sub-regions
Comparison of cumulative distribution functions weighted by background control variant rate. Genes limited to 24 from OMIM epileptic encephalopathy phenotype series also containing damaging (REVEL ≥ 0.5) missense variants in all three epilepsies (see gene set enrichment testing, Table S5). DEE, developmental and epileptic encephalopathy; GGE, genetic generalized epilepsy; NAFE, non-acquired focal epilepsy.
(A) CDF drawn directly from Epi25 data (dashed line) and weighted by control CDF (solid lines) to estimate "true positive" distribution. (B) Enlarged box from (A) showing just "true positive" CDFs with control CDF. "True positive" median MTR DEE = 0.670, GGE = 0.710, and NAFE = 0.721. p values generated by 10,000 permutations of Kolmogorov–Smirnov test. Plots calculated from 614 missense variants (DEE = 100, GGE = 133, and NAFE = 153; control = 228).

Sub-genic intolerance comparison

We examined sub-genic intolerance scores (MTR) in multiple ways. We compared the raw MTR and MTR domain percentiles scores across epilepsy-affected and control individuals directly by using the Kruskal-Wallis test by rank. For groups with p value $<$ 0.05, we performed pairwise comparisons by using the Wilcoxon signed-rank test. This method may not be an adequate comparison because, despite enriching for damaging missense variants with REVEL, control individuals with qualifying variants (which are unlikely to be true positives) remain, indicating that some of the qualifying variants found in affected individuals may also be benign. Direct comparison of sub-genic intolerance scores among epilepsies is therefore difficult to interpret because the QV burden is different among epilepsies (see results) and the true positive rate among these QVs is unknown.

To compare MTR among epilepsies, it was necessary to estimate and compare the "true positive" distribution of scores for each epilepsy. To achieve this, we created a weighted average of the cumulative distribution function (CDF) of MTR scores for ultra-rare damaging missense variants in each epilepsy (CDF_{DEE}, CDF_{GGE}, and CDF_{NAFE}) and the CDF of ultra-rare damaging missense variants in our controls (CDF_{CTRL}) to obtain the "true positive" CDF for each epilepsy (CDF_{DEE_TP}, CDF_{GGE_TP}, and CDF_{NAFE_TP}). Only damaging missense variants with defined MTR scores were considered.

At a given MTR value, the "true positive" CDF is a weighted average of the epilepsy and control CDF with the weights determined by the QV rate of the control population at that MTR value. For example, if at an MTR score of 0.5, 4% of DEE-affected individuals have an ultra-rare damaging missense variant and 1% of control individuals of have an ultra-rare damaging missense variant, then $CDF_{DEE\_TP}(0.5) = 0.75 \times CDF_{DEE}(0.5) + 0.25 \times CDF_{CTRL}(0.5)$.[69] We then used a Kolmogorov–Smirnov test (statistic D) to compare the distribution of "true positive" MTR CDFs of each epilepsy pair. Given that we did not know the distribution of D, we performed a permutation test with 10,000 permutations for each comparison. We assessed significance at $p < 0.05$.

To compare sub-genic intolerance scores by gene, we compared the "true positive" mean MTR by gene for DEE compared to NAFE and compared to GGE. In a given gene, the "true positive" mean MTR is a weighted average of the epilepsy mean MTR and control mean MTR scores with the weights determined by the QV rate of the control population in that gene. For example, if in gene X, 4% of DEE-affected individuals have an ultra-rare damaging missense variant and 1% of control individuals have an ultra-rare damaging missense variant, then $Mean_{DEE\_TP}(X) = 0.75 \times Mean_{DEE}(X) + 0.25 \times Mean_{CTRL}(X)$. For those genes with no control variants, the means were calculated without weighting. We measured the number of genes where DEE had a lower weighted mean MTR and measured significance with a binomial test with the null hypothesis that DEE variants had a lower $Mean_{TP}$ in half of the genes in the tested gene set.

To compare the MTR values of published ClinVar variants (i.e., not drawn from our affected individuals or control individuals), we divided the variants into those associated with DEE and non-DEE epilepsy. ClinVar variants with phenotypes containing "epilepsy" or "epileptic" were considered associated with epilepsy. Those with phenotypes containing "West," "Dravet," "Lennox-Gastaut," "infantile spasm," "Ohtahara," "myoclonic," or "glut 1" were considered associated with DEE, while the remainder were classified as non-DEE epilepsy. There was an inadequate number of variants specifically associated with GEE and NAFE to further sub-divide them. For variants with multiple clinical associations, the most severe association was assigned. We looked at only ultra-rare variants with a defined MTR value. We limited our analysis to only those genes harboring variants in both

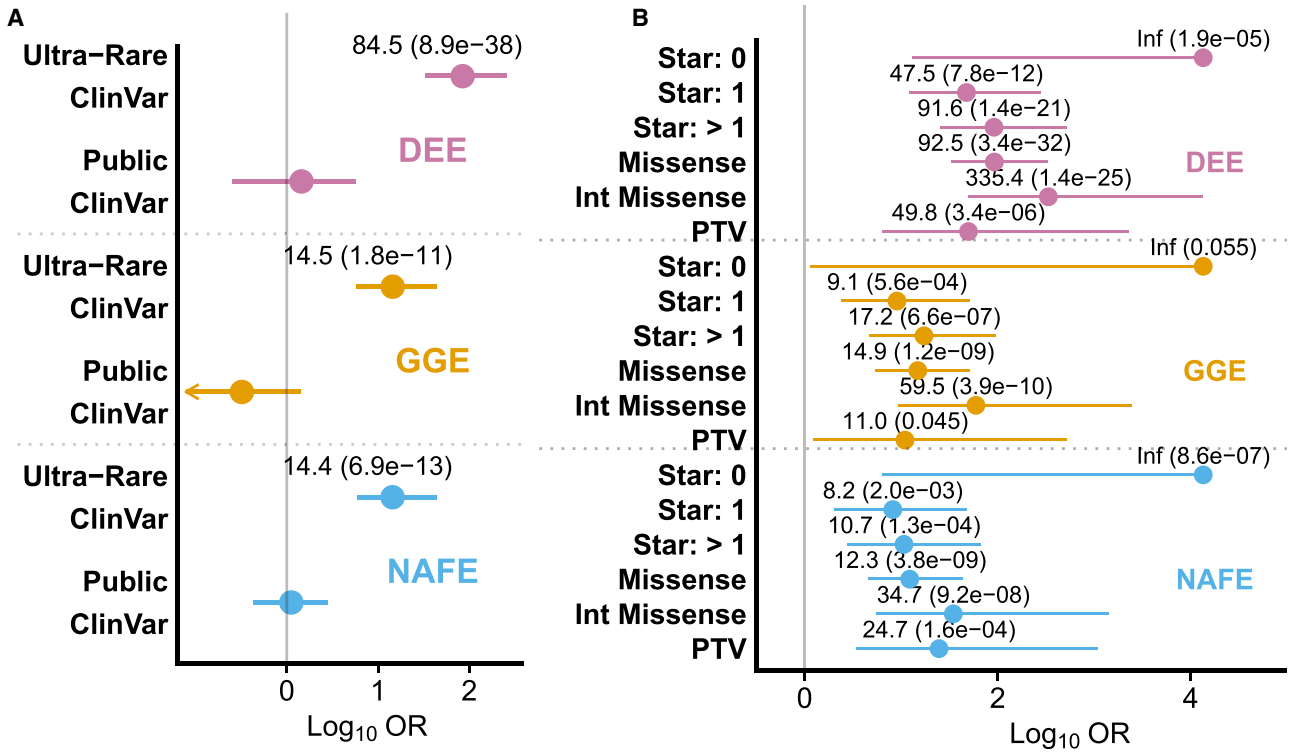


Figure 4. Burden of pathogenic/likely pathogenic (P/LP) variants in ClinVar found in Epi25 participants
Ultra-rare and intolerant P/LP variants are enriched in Epi25 participants with epilepsy compared to control individuals.
(A) Variants divided into ultra-rare (absent from non-neuro gnomAD populations) and public (present in non-neuro gnomAD populations) variants showing enrichment only among ultra-rare variants.
(B) Ultra-rare variants sub-divided to show drivers of enrichment. "Star" indicates the variant review status in ClinVar, which summarizes the level of review supporting the clinical significance of the variant with increasing number of "gold stars" from 0 to 4 (see qualifying variant). Pooled odds ratio, 95% confidence intervals, and FDR-corrected p value were generated from the exact two-sided Cochran-Mantel-Haenszel (CMH) test. Odds ratio and FDR-adjusted p values displayed for comparisons with unadjusted p value < 0.05. x axis displays the $\log_{10}$ of the odds ratio and confidence intervals. PTV, protein-truncating variants; "Int," "intolerant," MTR $\leq$ 0.78 (when defined); DEE, developmental and epileptic encephalopathy; GGE, genetic generalized epilepsy; NAFE, non-acquired focal epilepsy.

epilepsy groups (see gene set enrichment testing). The control variant set was drawn from the combined epilepsy analysis (Figures S1A–S1C). We used a two-sample Wilcoxon test to assess significance. We measured the number of genes where DEE had a lower mean MTR and measured significance with a binomial test with the null hypothesis that DEE variants had a lower mean MTR in half of the genes in the tested gene set.

Lollipop and MTR plots

Lollipop mutation diagrams were generated for the 24 genes analyzed for the sub-genic intolerance comparison (GS-2) via lollipops-v.1.5.3.[70] All 614 missense variants (DEE = 100, GGE = 133, NAFE = 153, and control = 228) were displayed across the linear gene structure of the associated gene. For each gene, the MTR distribution with missense variant locations plotted was juxtaposed against the lollipop mutation diagram. MTR data were downloaded from the MTR-Viewer website (see web resources).[71]

Comparison of evolutionary constrained regions

Evolutionary constraint for missense variants was assessed at three levels. For base-level scores, we used the GERP++ "rejected substitution" (RS) score in which higher scores correspond to greater constraint.[72,73] For exonic and domain constraint, we used exonic and domain subGERP scores, respectively.[22] We compared scores across epilepsies and controls directly by using the Kruskal-Wallis test by rank. No group reached statistical significance (p value $<$ 0.05), so no pairwise comparisons were performed.

Candidate non-OMIM epilepsy genes

To ascertain additional potential epilepsy-gene associations not found in OMIM, we highlighted genes that are (1) in the most intolerant decile to loss-of-function (LOF) variation in the general population by LOEUF rank, (2) not associated with a disease in OMIM, (3) harbor PTVs with LOFTEE filtering in more than one affected individual, and (4) harbor no control PTVs with LOFTEE filtering.

Data analysis and display

Unless otherwise noted in the methods, data analysis and visualization were performed with R (v.3.6.0).[74] Notches in boxplots indicate 1.58 * interquartile range / sqrt(n), which approximates the 95% confidence interval.[75]

Results

Gene-based collapsing in three types of epilepsies

The results of the gene-based collapsing should be viewed through the lens of prior rare-variant association analyses of epilepsy data and, specifically, Epi25 data. The data in

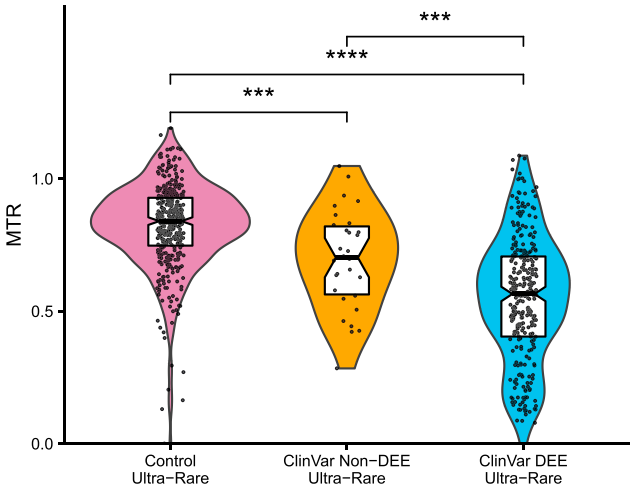


Figure 5. Comparison of median MTR scores of published ultra-rare P/LP ClinVar variants

Violin plots with boxplots showing distribution of MTR scores of published missense ClinVar P/LP variants divided into those associated with DEE (n = 302) and non-DEE (n = 29) epilepsies. We considered only those genes harboring missense variants in both groups (14 genes, see gene set enrichment testing, Table S2). Ultra-rare control variants (n = 335) drawn from Epi25 analysis (see sub-genic intolerance comparison). Comparisons by Wilcoxon signed-rank test. p values unadjusted. The middle horizontal line represents the median value and the lower and upper hinges represent the 1st and 3rd quartiles. The notches in the boxplot approximate the 95% confidence interval (see data analysis and display). MTR median ± standard deviation: DEE 0.57 ± 0.24, non-DEE 0.70 ± 0.18, control 0.83 ± 0.16. **p ≤ 0.01, *** p ≤ 0.001, ****p ≤ 0.0001.

this analysis are a superset of the data used in prior Epi25 analyses.[9] The cluster-based collapsing analysis allows for the inclusion of multiple ancestries because each geographic ancestry-matched cluster is analyzed separately (Figure S1). The results are then combined with the CMH test (see gene-based collapsing) accounting for population sub-structure.[56] The sample size increased in all three epilepsies (1,835 from 1,021 DEE-affected individuals, 5,303 from 3,108 GEE-affected individuals, and 6,349 from 3,597 NAFE-affected individuals) because of increased enrollment in Epi25 and the inclusion of affected individuals with non-European geographic ancestry. Other differences include a different control set and different *in silico* methods of indicating QV status. We ran gene-based collapsing (Tables S6–S26) for gene-discovery counting PTVs and damaging missense variants for all three epilepsies (Figure 1, Tables S9, S12, and S14) and all epilepsies combined (Figure S3B, Table S17). There was expected overlap among the top ranked genes from prior Epi25 analyses as well as the suggestion of candidate genes not previously associated with epilepsy (Tables S11–S26).

In the DEE collapsing analysis (Figure 1A, Table S9), the top two ranked genes were the same as in the prior Epi25 analysis, but now *SCN1A* ([MIM: 182389] OR = 7.1, p = 4.4×10^{-8}) and *NEXMIF* (previously known as *KIAA2022* [MIM: 300524] OR 26.5, p = 8.6×10^{-8}) both achieve study-wide significance. In contrast to prior Epi25 analyses, nine of the top ten ranked genes are known epilepsy genes,[76–87] demonstrating the strength of the increased sample size and clustering methodology. The remaining gene, *AP3S2* ([MIM: 602416] OR = 70.5, p = 2.7×10^{-4}), is a component of the AP3 complex, an adaptor-related complex with no prior association to epilepsy, although it was a top 200 hit in the prior Epi25 DEE analysis.[9,88] Hermansky-Pudlak syndrome 10 (MIM: 617050), which is notable for infantile onset of immunodeficiency and intractable seizures, is caused by bi-allelic mutations in *AP3D1* (MIM: 607246), a different component of the same AP3 complex.[89] To highlight candidate genes, we removed DEE-affected individuals in Figure 1A that harbored a qualifying variant in any of the 101 dominant genes with epilepsy or related terms in the OMIM phenotype and re-ran the collapsing analysis (Figure S4, Table S11). The 5th ranked gene, *SRCAP* ([MIM: 611421] OR = 6.8, p = 1.6×10^{-3}), is highly intolerant to loss-of-function variants (LOUEF = 0.1) and is associated with Floating-Harbor syndrome (MIM: 136140), which can include seizures.[90,91] In summary, this enlarged DEE analyses with affected individuals of non-European geographic ancestry produced results that more consistently elevated known epilepsy-associated genes and, importantly, proposed genes without prior epilepsy associations (*AP3S2* and *SRCAP*).

Four of the top ten ranked genes in the gene-based collapsing analysis for GGE (Figure 1B, Table S12) were previously associated with epilepsy (*SLC6A1* [MIM: 137165], *SCN1A*, *GRIN2A* [MIM: 138253], and *GABRA1* [MIM: 137160]).[92–95] The top hit is *SLC6A1* (OR = 16.6, p = 2.1×10^{-6}), which was a top 200 gene in the prior Epi25 GGE analysis but now approaches study-wide significance.[9] *SCL6A1* was initially implicated in DEE, but its role in generalized epilepsies has only been more recently revealed.[95,96] Among the remaining genes, there are two promising candidates: (1) *FBXO42* ([MIM: 609109] OR = 13.6, p = 4.5×10^{-4}), which is a highly intolerant gene (LOEUF = 0.27) important in the regulation of p53 and not yet implicated in disease but was a top 200 GGE-associated gene in the prior Epi25 analysis,[9] and (2) *KCNK18* ([MIM: 613655] OR = Inf, p = 1.6×10^{-3}), which is a potassium channel implicated in migraine pathology.[97,98] Promising candidate genes for GGE from the prior Epi25 analysis (*CACNA1G* [MIM: 604065] and *UNC79* [MIM: 616884]) were not among the top 200 associated genes, which may be related to the different method of filtering missense variants.[9] Further limiting missense variants to intolerant as well as damaging (Figure S5B, Table S13) elevated *CACNA1B* ([MIM: 601012], OR = 5.5, p = 3.5×10^{-4}). Bi-allelic LOF variants in *CACNA1B* cause severe epilepsy.[99] *CACNA1B* was the top gene associated with GGE in Epi4K,[7] a large WES epilepsy project prior to Epi25. No association was found in the prior Epi25 analysis and there is limited other literature linking *CACNA1B* to GGE. This new GGE Epi25 collapsing analysis did not confirm promising candidate genes from the prior Epi25 analysis but did provide additional support for the association between

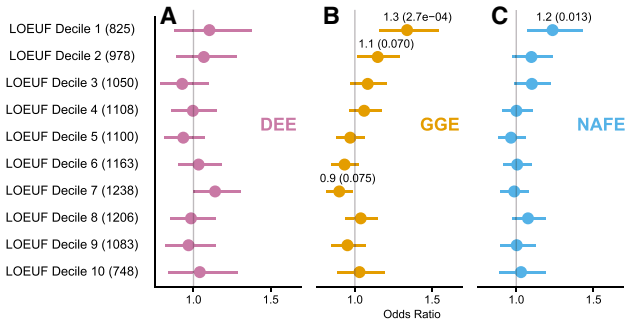


Figure 6. Burden of protein-truncating variants in intolerant non-OMIM genes

The burden of protein-truncating variants (PTVs) in genes not associated with a disease in OMIM in epilepsy-affected individuals in comparison to control individuals was assessed. We divided non-OMIM genes into 10 gene sets by their intersection with loss-of-function intolerance deciles defined by LOEUF (see gene set enrichment testing, Table S5). The number of genes in each gene set with at least one PTV in the case-control set is specified in the parenthesis. Pooled odds ratio, 95% confidence intervals, and FDR-corrected p value were generated from the exact two-sided Cochran-Mantel-Haenszel (CMH) test for (A) developmental and epileptic encephalopathies (DEE), (B) genetic generalized epilepsy (GGE), and (C) non-acquired focal epilepsy (NAFE). Odds ratio and FDR-adjusted p values are displayed in parentheses for comparisons with unadjusted p value < 0.05. x axis displays the odds ratio and confidence intervals.

CACNA1B and GGE and proposed candidate genes (*FBXO42* and *KCNK18*).

Gene-based collapsing analysis for NAFE (Figure 1C, Table S14) showed a familiar top hit, *DEPDC5* ([MIM: 614191] OR = 5.4, p = 1.3×10^{-6}), and four additional genes (*GRIN2A, SCN1A, SCN8A* [MIM: 600702], and *NPRL2* [MIM: 607072]), which have previously been implicated in NAFE.[7,9,80,84,92,100,101] Renin, the protein encoded by *REN* ([MIM: 179820] OR = 12.7, p = 4.2×10^{-4}), is produced by juxtaglomerular cells of the kidney but has been implicated as a target of adjuvant therapy for epilepsy.[102,103] *ADORA2B* ([MIM: 600446] OR = Inf, p = 4.5×10^{-4}), is a small gene encoding an adenosine receptor not associated with disease but being explored for its role in epileptogenesis.[104,105] *DAW1* (OR = 30.0, p = 1.8×10^{-4}), a little understood gene, supports cilia function.[106] The increased sample size did not further support promising genes from the prior Epi25 analysis, such as *TRIM3* (MIM: 605493), *PPFIA3* (MIM: 603144), and *KCNJ3* (MIM: 601534).[9] Further limiting missense variants to intolerant as well as damaging (Figure S5C, Table S15) removed all control-enriched genes from the top ten ranked genes and elevated known epilepsy genes. Interestingly, the 7th ranked gene, *TSC1* ([MIM: 605284], OR = 14, p = 1.7×10^{-3}), is typically associated with focal epilepsy in the context of tuberous sclerosis-1 (MIM: 191100) or focal cortical dysplasia, type II, somatic (MIM: 607341), although the individuals with focal epilepsy in this study do not have a lesion on MRI.[107,108] Like the GGE collapsing analysis, the NAFE collapsing analysis proposed different candidate genes rather than confirming those from prior Epi25 analyses.

Milder epilepsies remain enriched for ultra-rare variants in a limited gene set

Our group has previously observed that more mild epilepsies are enriched in genes also associated with severe phenotypes.[7,9] To limit the degree to which individual genes in the gene set drove that finding and facilitate comparisons of variants across epilepsies, we recapitulated that analysis but narrowed the gene set of dominant DEE-associated genes to include only those 24 genes containing at least one damaging missense variant in all three epilepsies (Figure 2, Tables S5 and S27). DEE (CMH pooled odds ratio [OR] = 2.1, FDR-adjusted p value [adj.p] = 1.9×10^{-9}) and NAFE (CMH pooled odds ratio [OR] = 1.3, FDR-adjusted p value [adj.p] = 1.2×10^{-3}) are enriched for all missense variants. All three epilepsies are enriched for damaging missense variants (DEE OR = 3.7, adj.p = 6.8×10^{-17}; GGE OR = 1.7, adj.p = 1.2×10^{-4}; NAFE OR = 1.7, adj.p = 6.4×10^{-5}), and removing the damaging filter, all three epilepsies are also enriched for variants in intolerant genic sub-regions (DEE OR = 3.5, adj.p = 1.6×10^{-14}; GGE OR = 1.7, adj.p = 1.2×10^{-4}; NAFE OR = 1.6, adj.p = 3.5×10^{-4}). Combining both improves enrichment in all three epilepsies (DEE OR = 5.5, adj.p = 8.1×10^{-19}; GGE OR = 2.2, adj.p = 1.0×10^{-6}; NAFE OR = 2.0, adj.p = 1.8×10^{-5}). Only DEE and GGE were enriched for loss-of-function variants (DEE OR = 12.7, adj.p = 1.9×10^{-9}; GGE OR = 3.8, adj.p = 4.6×10^{-4}), which is consistent with prior analyses.[9] In summary, despite restricting our DEE-associated gene set further to ensure that at least one affected individual per epilepsy harbored a damaging missense variant in each gene and enlarging our samples to include individuals of non-European ancestry, a familiar pattern of enrichment exists in the milder epilepsies.

Ultra-rare DEE variants in Epi25 are located in intolerant genic sub-regions

After demonstrating that more mild epilepsies (GGE, NAFE) were enriched for ultra-rare damaging missense variants in the same gene set as severe epilepsies (DEE) (Figure 2), we tested the hypothesis that variants associated with DEE were located in more intolerant sub-regions than those associated with GGE or NAFE. Despite filtering for pathogenicity with REVEL, there remains a background rate of enrichment of ultra-rare damaging missense variants in the control population (Figure 2, Table S29). This suggests that a portion of the ultra-rare damaging missense variants in our epilepsy-affected individuals are also benign, which makes direct comparison of the sub-genic intolerance score among epilepsy subtypes (Figure S6A) difficult to interpret because the burden of damaging missense variants in DEE-affected individuals is higher than those of GGE or NAFE (CMH; DEE-GGE OR = 2.2, adj.p = 7.8×10^{-7}; DEE-NAFE OR = 2.3, adj.p = 9.4×10^{-8}; Table S28). Instead, we estimated the distribution of MTR scores of "true positive" ultra-rare damaging missense variants in each epilepsy and made pairwise comparisons by using a Kolmogorov-Smirnov (K-S) test (see

Table 1. Non-OMIM genes intolerant to loss-of-function variants with multiple protein-truncating variants in genetic generalized epilepsy or non-acquired focal epilepsy

GGE-associated gene	Number of GGE-affected individuals in Epi25	GGE p value	NAFE-associated gene	Number of NAFE-affected individuals in Epi25	NAFE p value
NLGN2	3	8.6×10^{-3}	*WDR18*	4	0.01
HDLBP	4	8.9×10^{-3}	*SOCS7*	5	0.01
RC3H2	4	0.01	*TRIM9*	3	0.05
XPO5	3	0.02	*ENAH*	2	0.05

Genes listed are among the most intolerant decile to loss-of-function variation and harbor protein-truncating variants (PTVs) in more than one epilepsy-affected individual but harbor no PTVs in control individuals. Only the top four gene associations are shown per epilepsy. Full tables can be found in the supplemental information (Tables S37 and S38). p values drawn from ultra-rare protein-truncating variants collapsing analysis (Figure S9, Tables S19 and S20).

sub-genic intolerance comparison, Figure 3). Consistent with our hypothesis, the distribution of MTR scores for DEE variants was significantly different from NAFE ("true positive" median MTR DEE = 0.670 versus NAFE = 0.721, K-S, $p < 0.0156$), while the difference from GGE did not achieve statistical significance ("true positive" median MTR DEE = 0.670 versus GGE = 0.710, K-S, $p = 0.38$). On a per gene basis, the MTR scores of DEE variants are not uniformly more intolerant than GGE and NAFE (Figure S7). Although the above analysis demonstrates that DEE variants lay in more intolerant genic sub-regions than NAFE variants, it does not account for the possible differential contribution of specific genes to specific epilepsies among the 24 genes. To address this concern, we performed a second analysis that compared the weighted mean MTR of DEE compared to NAFE and to GGE (Table S29). The weighted mean MTR scores of the DEE variants was lower in 15 of the 24 genes compared to NAFE (binomial test, $p = 0.31$) and 15 of the 24 genes compared to GGE (binomial test, $p = 0.31$).

No clear relationship exists between gene, protein domain, and epilepsy type (Figure S8). Despite the large Epi25 dataset, we most likely remain underpowered to untangle the epilepsy by protein space relationship on an individual gene level.[33] MTR is calculated on a sliding window, making it independent of known gene structures. Domain-based MTR showed a smaller difference among the epilepsies (Figures S6A and S6B), suggesting that the sub-genic intolerance differences among the epilepsies is at least partially independent from gene structures.[32] We also examined whether missense variants associated with DEE were located in more evolutionary constrained bases, exons, or domains than milder epilepsies (Figures S6C–S6E). No comparison met statistical significance. This was true despite both evolutionary constrained and intolerant domains harboring pathogenic variants, although differences in domains may be difficult to assess given the limited number per gene.[22]

Only ultra-rare pathogenic/likely pathogenic ClinVar variants are enriched in Epi25

The sample size of Epi25 allows us to assess the representation of variants found in ClinVar, a heavily used clinical database of curated variants, in our three epilepsy subgroups and investigate whether sub-genic intolerance might add clinically useful information.[38,39] Using a set of 101 genes with epilepsy or related terms in their OMIM phenotypes (Table S5), we examined the burden of P/LP variants in our affected individuals compared to control individuals (Figure 4A, Table S30). Given the prior findings that epilepsy-affected individuals are enriched with ultra-rare variants but not more common variants,[7] we divided our ClinVar analysis into variants not found in the non-neuro gnomAD populations (ultra-rare) and variants seen in the general population (public). Consistent with prior reports, there was an increased burden of ultra-rare P/LP variants in our epilepsy-affected individuals compared to control individuals irrespective of epilepsy type (CMH; DEE OR = 84.5, adj.p = 8.9×10^{-38}; GGE OR = 14.5, adj.p = 1.8×10^{-11}; NAFE OR = 14.4, adj.p = 6.9×10^{-13}). There was no enrichment in public variants (Figure 4A). Epilepsy variants in ClinVar also found in gnomAD or future public datasets may require additional investigation to confirm pathogenicity.

Severe pathogenic/likely pathogenic ClinVar variants are located in intolerant genic sub-regions

Among ultra-rare ClinVar variants, we sought to determine whether we could further differentiate epilepsy variants from control variants (Figure 4B, Table S31). ClinVar "review status" attempts to capture the level of review supporting the assertion of clinical significance for the variant with increasing number of "gold stars" from zero to four.[63–65] Filtering ultra-rare P/LP ClinVar on the basis of review status did not improve discrimination in a dose-dependent fashion. In all three epilepsies, there were no zero star controls but the enrichment of variants with more than one star exceeded the enrichment of variants with one star (CMH; DEE OR = 47.5, adj.p = 7.8×10^{-12} → OR = 91.6, adj.p = 1.4×10^{-21}; GGE OR = 9.1, adj.p = 5.6×10^{-4} → OR = 17.2, adj.p = 6.6×10^{-7}; NAFE OR = 8.2, adj.p = 2.0×10^{-3} → OR = 10.7, adj.p = 1.3×10^{-4}). We next examined whether sub-genic intolerance filtering could further improve discrimination of affected individuals compared to control individuals. After filtering with MTR, the OR of ultra-rare missense variants increased in all three epilepsies (CMH; DEE OR = 92.5, adj.p = 3.4×10^{-32} → OR =

335.4, adj.p = 1.4×10^{-25}; GGE OR = 14.9, adj.p = 1.2×10^{-9} → OR = 59.6, adj.p = 3.9×10^{-10}; NAFE OR = 12.3, adj.p = 3.8×10^{-9} → OR = 34.7, adj.p = 9.2×10^{-8}). All three epilepsies were enriched with ultra-rare PTVs in ClinVar (DEE OR = 49.8, adj.p = 3.4×10^{-6}; GGE OR = 11.0, adj.p = 0.045; NAFE OR = 24.7, adj.p = 1.6×10^{-4}). Among the few public variants, only missense variants filtered with MTR were statistically enriched in NAFE-affected individuals, and overall, MTR filtering removed all 12 control missense variants but only four of ten epilepsy variants (Table S32). In summary, sub-genic intolerance filtering improved discrimination of both ultra-rare and public variants in ClinVar, suggesting sub-genic intolerance provides additive information to identify potential false-positive or variable penetrance variants in ClinVar.

Using ultra-rare P/LP ClinVar variants, we sought to confirm our Epi25 finding (Figure 3) that missense variants in severe epilepsies are located in more intolerant genic sub-regions than milder epilepsies. We compared median sub-genic intolerance scores between DEE and non-DEE epilepsies (see sub-genic intolerance comparison) in genes with missense variants in both epilepsy groups (Figure 5, Tables S5 and S33). The median MTR score was lower (more intolerant) for published ClinVar DEE variants compared to non-DEE epilepsy ClinVar variants (median DEE MTR = 0.57 versus median non-DEE MTR = 0.70, Wilcoxon signed-rank test, $p < 6.7 \times 10^{-3}$). When examined by gene, the mean MTR score for the DEE variants was lower than the non-DEE variants in 11 of 14 genes tested (binomial test, p = 0.057). Reassuringly, both DEE and non-DEE variants existed in more intolerant regions than ultra-rare control variants (median control MTR = 0.83, control set drawn from combined epilepsy clusters, see clustering).

Epilepsy genes remain to be discovered and are most likely loss-of-function intolerant

There are ~3,900 genes identified in OMIM as harboring variants that are causative or a risk factor for disease.[109] Analyzing likely damaging variants in non-OMIM genes may give a sense of as-yet to be discovered epilepsy genes (Figure 6, Tables S34 and S35). GGE and NAFE revealed a significant burden of PTVs in the intersection of non-OMIM genes with the decile of genes most intolerant to loss-of-function variation in the general population (GGE OR = 1.3, adj.p = 2.7×10^{-4}; NAFE OR = 1.2, adj.p = 0.013) (Figures 6B and 6C). We highlighted the top four genes in the most intolerant decile associated with GGE and NAFE that had more than one case PTV and no control PTVs (Table 1). The most significant GGE candidate gene, *NLGN2* (MIM: 606479, 3 cases), encodes neuroligin 2, which is a *trans*-synaptic adhesion molecule important in the synapse.[110] The most significant NAFE candidate gene was *WDR18* (4 cases), whose protein product forms the PELP1-TEX10-WDR18 complex important in ribosomal maturation.[111] Tables of potential DEE, GGE, and NAFE genes are included in the supplement (Tables S36–S38). Finally, to investigate additional candidate genes, we performed ultra-rare variant collapsing analysis with only PTVs (Figure S9, Tables S18–S20), only damaging missense variants (Figure S10, Tables S21–S23), and PTVs combined with damaging and intolerant missense variants further limited to intolerant LIMBR exons (see qualifying variant, Figure S11, Tables S24–S26).[31]

DEE-affected individuals also revealed a trend toward increased burden in the intersection of non-OMIM genes with the 7th most intolerant decile (DEE OR = 1.1, adj.p = 0.14, Figure 6A), which may reflect genes associated with recessive epilepsies.[20] None of the epilepsies revealed a significant burden of damaging and intolerant missense variants in missense intolerant genes (Table S35).

Discussion

In this, the largest Epi25 exome study of epilepsies to date including individuals of non-European geographic descent, we reaffirm that ultra-rare variants contribute to the three major epilepsy groups (Figure 1). Our collapsing analyses proposed epilepsy-associated genes (*AP3S2*, *SRCAP*, *FBXO42*, *KCNK18*, *REN*, and *ADORA2B*) requiring future confirmation. These associations reveal the power of increasing sample size with Epi25 and our clustering technique's inclusion of non-European populations. The p values in DEE analyses must be regarded in light of the smaller sample size of individuals with DEE (1,835 with DEE compared to 5,303 with GGE and 6,379 with NAFE). We were unable to confirm several promising candidate genes from the prior Epi25 analysis that may be secondary to different control groups, different *in silico* filters, or a larger sample size.[9] We confirmed enrichment of ultra-rare variants in GGE and NAFE in genes associated with DEE even when limited to genes in which all epilepsies have a damaging missense variant to limit single and distinct genes' driving associations with different epilepsies (Figure 2).

Sub-genic intolerance has broad implications. It has been shown to help improve discrimination between pathogenic and benign variants and confirm the pathogenicity of new variants.[22,32,112–115] Pathogenic variants may cluster in areas of regional intolerance,[31,32,116] and sub-genic intolerance scores may inform biochemical exploration, yielding novel insights into protein function.[117] To our knowledge, this is the broadest demonstration that sub-genic intolerance scores might not only be different between case and control but also affect disease severity (Figures 3 and 5).[32] This discrepancy may broadly inform the functional similarities of mutations leading to more severe disease across genes or, interestingly, across gene families.[118]

Using the large Epi25 dataset allowed us to assess variants documented in ClinVar (Figure 4). Allele frequency is known to be inversely associated with pathogenicity, and among Epi25 participants, only ultra-rare variants were enriched in affected individuals compared to control

individuals (Figure 4A). Previous analyses have used population-based MAFs to reclassify variants as benign.[32,68,119,120] The evolving nature of ClinVar classifications has been noted previously as more population-wide control data become available.[63,64,121] Within the ultra-rare MAF bin, review status did not provide additional enrichment in a dose-dependent manner in our data (Figure 4B), although it has indicated higher true positive value in other studies focused on more common variants.[63–65,122] One and two star ultra-rare pathogenic variants in ClinVar have been reported as possible false-positives,[122] although no study to our knowledge has systematically evaluated ultra-rare P/LP ClinVar variants for false-positivity or incomplete penetrance. Finally, four of the five ultra-rare and all 12 public missense P/LP variants harbored by control individuals were located in more tolerant regions of the exome (Figure 4B, Tables S31 and S32). The enrichment of ClinVar variants with MTR filtering suggests that regional intolerance may provide additional information to clinicians assessing ClinVar variants.

There most likely remain genes that will ultimately be associated with a disease, although the pace of discovery may be slowing.[109] In this Epi25 cohort, GGE and NAFE contained an increased burden of PTVs in the non-OMIM genes most intolerant to loss-of-function variation in the general population (Figure 6). No increase was seen for individuals with DEE, suggesting that gene discovery for DEE is advanced compared to the milder epilepsies. There are several genes with PTVs in multiple affected individuals but in no control individuals that are potential epilepsy or epilepsy-risk genes (Tables 1 and S37–S39). With increased sample size, these genes may become more prominent in future collapsing analyses.

Limitations of this study are that individuals with epilepsy were enrolled at variable ages, leaving open the possibility that a case may evolve from one epilepsy to another. While we posit that variant location determines the severity of the variant and therefore determines the phenotype, this does not address variants that have one autosomal dominant phenotype and a different autosomal recessive phenotype. The sub-genic intolerance score-by-gene interaction (Figures S6 and S7) may be secondary to different numbers of variants per gene, incomplete capture of all sub-genic intolerance information by MTR, or other factors that contribute to epilepsy severity. Examining the collective sub-genic intolerance scores of variants from multiple genes does not take into account within-gene comparisons (i.e., sub-genic intolerance distributions differ per gene, as do the epilepsy type-by-gene burdens). We attempted to address these confounds (Tables S29 and S33) but were under-powered. Future studies will be needed to understand the gene-by-intolerance score interaction. Finally, segregation analysis of variants in candidate epilepsy-associated genes (Table 1) could weaken or bolster the proposed relationships. Unfortunately, we do not have access to Epi25 family member data. As the Epi25 enrollment increases, we look forward to the increased power's allowing for the further elucidation of the genetic architectures of the epilepsies.

Data and code availability

The accession number for the Epi25 Year 1 whole-exome sequencing data reported in this paper is dbGaP: phs001489. Epi25 Year 2 will be available in the near future under the same accession number. Epi25 Year 3 is not yet publicly available.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.ajhg.2021.04.009.

Consortia

The members of the Epi25 Collaborative are Joshua E. Motelow, Gundula Povysil, Ryan S. Dhindsa, Kate E. Stanley, Andrew S. Allen, Yen-Chen Anne Feng, Daniel P. Howrigan, Liam E. Abbott, Katherine Tashman, Felecia Cerrato, Caroline Cusick, Tarjinder Singh, Henrike Heyne, Andrea E. Byrnes, Claire Churchhouse, Nick Watts, Matthew Solomonson, Dennis Lal, Namrata Gupta, Benjamin M. Neale, Gianpiero L. Cavalleri, Patrick Cossette, Chris Cotsapas, Peter De Jonghe, Tracy Dixon-Salazar, Renzo Guerrini, Hakon Hakonarson, Erin L. Heinzen, Ingo Helbig, Patrick Kwan, Anthony G. Marson, Slavé Petrovski, Sitharthan Kamalakaran, Sanjay M. Sisodiya, Randy Stewart, Sarah Weckhuysen, Chantal Depondt, Dennis J. Dlugos, Ingrid E. Scheffer, Pasquale Striano, Catharine Freyer, Roland Krause, Patrick May, Kevin McKenna, Brigid M. Regan, Caitlin A. Bennett, Costin Leu, Stephanie L. Leech, Terence J. O'Brien, Marian Todaro, Hannah Stamberger, Danielle M. Andrade, Quratulain Zulfiqar Ali, Tara R. Sadoway, Heinz Krestel, André Schaller, Savvas S. Papacostas, Ioanna Kousiappa, George A. Tanteles, Yiolanda Christou, Katalin Štěrbová, Markéta Vlčková, Lucie Sedlackova, Petra Laššuthová, Karl Martin Klein, Felix Rosenow, Philipp S. Reif, Susanne Knake, Bernd A. Neubauer, Friedrich Zimprich, Martha Feucht, Eva M. Reinthaler, Wolfram S. Kunz, Gábor Zsurka, Rainer Surges, Tobias Baumgartner, Randi von Wrede, Manuela Pendziwiat, Hiltrud Muhle, Annika Rademacher, Andreas van Baalen, Sarah von Spiczak, Ulrich Stephani, Zaid Afawi, Amos D. Korczyn, Moien Kanaan, Christina Canavati, Gerhard Kurlemann, Karen Müller-Schlüter, Gerhard Kluger, Martin Häusler, Ilan Blatt, Johannes R. Lemke, Ilona Krey, Yvonne G. Weber, Stefan Wolking, Felicitas Becker, Stephan Lauxmann, Christian Boßelmann, Josua Kegele, Christian Hengsbach, Sarah Rau, Bernhard J. Steinhoff, Andreas Schulze-Bonhage, Ingo Borggräfe, Christoph J. Schankin, Susanne Schubert-Bast, Herbert Schreiber, Thomas Mayer, Rudolf Korinthenberg, Knut Brockmann, Markus Wolff, Dieter Dennig, Rene Madeleyn, Reetta Kälviäinen, Anni Saarela, Oskari Timonen, Tarja Linnankivi, Anna-Elina Lehesjoki, Sylvain Rheims, Gaetan Lesca, Philippe Ryvlin, Louis Maillard, Luc Valton, Philippe Derambure, Fabrice Bartolomei, Edouard Hirsch, Véronique Michel, Francine Chassoux, Mark I. Rees, Seo-Kyung Chung, William O. Pickrell, Robert Powell, Mark D. Baker, Beata Fonferko-Shadrach, Charlotte Lawthom, Joseph Anderson, Natascha Schneider, Simona Balestrini, Sara Zagaglia, Vera Braatz, Michael R. Johnson, Pauls Auce, Graeme J. Sills, Larry W. Baum, Pak C. Sham, Stacey S. Cherny, Colin H.T. Lui, Norman Delanty, Colin P. Doherty, Arif Shukralla, Hany El-Naggar, Peter Widdess-Walsh, Nina Barišić, Laura

Canafoglia, Silvana Franceschetti, Barbara Castellotti, Tiziana Granata, Francesca Ragona, Federico Zara, Michele Iacomino, Antonella Riva, Francesca Madia, Maria Stella Vari, Vincenzo Salpietro, Marcello Scala, Maria Margherita Mancardi, Lino Nobili, Elisabetta Amadori, Thea Giacomini, Francesca Bisulli, Tommaso Pippucci, Laura Licchetta, Raffaella Minardi, Paolo Tinuper, Lorenzo Muccioli, Barbara Mostacci, Antonio Gambardella, Angelo Labate, Grazia Annesi, Lorella Manna, Monica Gagliardi, Elena Parrini, Davide Mei, Annalisa Vetro, Claudia Bianchini, Martino Montomoli, Viola Doccini, Carmen Barba, Shinichi Hirose, Atsushi Ishii, Toshimitsu Suzuki, Yushi Inoue, Kazuhiro Yamakawa, Ahmad Beydoun, Wassim Nasreddine, Nathalie Khoueiry Zgheib, Birute Tumiene, Algirdas Utkus, Lynette G. Sadleir, Chontelle King, S. Hande Caglayan, Mutluay Arslan, Zuhal Yapıcı, Pınar Topaloglu, Bulent Kara, Uluc Yis, Dilsad Turkdogan, Aslı Gundogdu-Eken, Nerses Bebek, Meng-Han Tsai, Chen-Jui Ho, Chih-Hsiang Lin, Kuang-Lin Lin, I-Jun Chou, Annapurna Poduri, Beth R. Shiedley, Catherine Shain, Jeffrey L. Noebels, Alicia Goldman, Robyn M. Busch, Lara Jehi, Imad M. Najm, Lisa Ferguson, Jean Khoury, Tracy A. Glauser, Peggy O. Clark, Russell J. Buono, Thomas N. Ferraro, Michael R. Sperling, Warren Lo, Michael Privitera, Jacqueline A. French, Steven Schachter, Ruben I. Kuzniecky, Orrin Devinsky, Manu Hegde, David A. Greenberg, Colin A. Ellis, Ethan Goldberg, Katherine L. Helbig, Mahgenn Cosico, Priya Vaidiswaran, Eryn Fitch, Samuel F. Berkovic, Holger Lerche, Daniel H. Lowenstein, and David B. Goldstein. See supplemental information for consortium member affiliations.

Acknowledgments

We thank the Epi25 principal investigators, local staff overseeing individual cohorts, and all of the individuals with epilepsy and their families who participated in Epi25 for their commitment to this international collaboration. J.E.M. is supported by the National Institutes of Health (TL1TR001875). This work is part of the Centers for Common Disease Genomics (CCDG) program, funded by the National Human Genome Research Institute (NHGRI) and the National Heart, Lung, and Blood Institute (NHLBI). CCDG-funded Epi25 research activities at the Broad Institute, including genomic data generation in the Broad Genomics Platform, are supported by NHGRI grant UM1 HG008895 (PIs: Eric Lander, Stacey Gabriel, Mark Daly, and Sekar Kathiresan). The Genome Sequencing Program efforts were also supported by NHGRI grant 5U01HG009088. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. A supplemental grant for Epi25 phenotyping was supported by "Epi25 Clinical Phenotyping R03," National Institutes of Health (1R03NS108145-01); D.H.L. and S.F.B. were the principal investigators. We also thank the Stanley Center for Psychiatric Research at the Broad Institute for supporting the genomic data generation. Additional funding sources and acknowledgment of individual cohorts are listed in the supplemental information.

Declaration of interests

B.M.N. is a member of the scientific advisory board at Deep Genomics and RBNC Therapeutics, a member of the scientific advisory committee at Milken, and a consultant for Camp4 Therapeutics, Takeda Pharmaceutical, and Biogen. R.S.D. is a consultant for AstraZeneca. D.B.G. is a founder and shareholder in Praxis Precision Medicines, a shareholder in and member of the scientific advisor board for Apostle Inc., a shareholder in Q State – Biosciences, and a consultant for Gilead Sciences, AstraZeneca, and GoldFinch Bio.

Received: February 9, 2021
Accepted: April 8, 2021
Published: April 30, 2021

Web resources

ATAV, https://github.com/igm-team/atav
ClinVar, https://www.ncbi.nlm.nih.gov/clinvar/
Consensus Coding Sequence, https://www.ncbi.nlm.nih.gov/CCDS/CcdsBrowse.cgi
Epi25 Collaborative, http://epi-25.org/
Epi25 WES results browser, https://epi25.broadinstitute.org/
EpiPGX project, http://www.epipgx.eu
Exome Aggregation Consortium (ExAC), http://exac.broadinstitute.org
Exome Variant Server, https://evs.gs.washington.edu/EVS/
Genome Aggregation Database (gnomAD), https://gnomad.broadinstitute.org
Genome Analysis Toolkit (GATK), https://gatk.broadinstitute.org/hc/en-us
lollipops-v.1.5.3, https://github.com/joiningdata/lollipops
MTR-Viewer, http://biosig.unimelb.edu.au/mtr-viewer/
NIH Genomic Data Sharing Policy, https://osp.od.nih.gov/scientific-sharing/policies/
OMIM, https://www.omim.org
Picard, http://broadinstitute.github.io/picard/
R, https://www.R-project.org/
Rare Exome Variant Ensemble Learner (REVEL), https://sites.google.com/site/revelgenomics/

References

1. Aaberg, K.M., Gunnes, N., Bakken, I.J., Lund Søraas, C., Berntsen, A., Magnus, P., Lossius, M.I., Stoltenberg, C., Chin, R., and Surén, P. (2017). Incidence and Prevalence of Childhood Epilepsy: A Nationwide Cohort Study. Pediatrics *139*, e20163908.
2. Fisher, R.S., Acevedo, C., Arzimanoglou, A., Bogacz, A., Cross, J.H., Elger, C.E., Engel, J., Jr., Forsgren, L., French, J.A., Glynn, M., et al. (2014). ILAE official report: a practical clinical definition of epilepsy. Epilepsia *55*, 475–482.
3. Hesdorffer, D.C., Logroscino, G., Benn, E.K., Katri, N., Cascino, G., and Hauser, W.A. (2011). Estimating risk for developing epilepsy: a population-based study in Rochester, Minnesota. Neurology *76*, 23–27.
4. EpiPM Consortium (2015). A roadmap for precision medicine in the epilepsies. Lancet Neurol. *14*, 1219–1228.
5. Ellis, C.A., Petrovski, S., and Berkovic, S.F. (2020). Epilepsy genetics: clinical impacts and biological insights. Lancet Neurol. *19*, 93–100.
6. May, P., Girard, S., Harrer, M., Bobbili, D.R., Schubert, J., Wolking, S., Becker, F., Lachance-Touchette, P., Meloche, C., Gravel, M., et al.; Epicure Consortium; EuroEPINOMICS CoGIE Consortium; and EpiPGX Consortium (2018). Rare coding variants in genes encoding $GABA_A$ receptors in genetic generalised epilepsies: an exome-based case-control study. Lancet Neurol. *17*, 699–708.

7. Epi4K consortium; and Epilepsy Phenome/Genome Project (2017). Ultra-rare genetic variation in common epilepsies: a case-control sequencing study. Lancet Neurol. *16*, 135–143.
8. Allen, A.S., Berkovic, S.F., Cossette, P., Delanty, N., Dlugos, D., Eichler, E.E., Epstein, M.P., Glauser, T., Goldstein, D.B., Han, Y., et al.; Epi4K Consortium; and Epilepsy Phenome/Genome Project (2013). De novo mutations in epileptic encephalopathies. Nature *501*, 217–221.
9. Epi25 Collaborative (2019). Ultra-Rare Genetic Variation in the Epilepsies: A Whole-Exome Sequencing Study of 17,606 Individuals. Am. J. Hum. Genet. *105*, 267–282.
10. Krenn, M., Wagner, M., Hotzy, C., Graf, E., Weber, S., Brunet, T., Lorenz-Depiereux, B., Kasprian, G., Aull-Watschinger, S., Pataraia, E., et al. (2020). Diagnostic exome sequencing in non-acquired focal epilepsies highlights a major role of GATOR1 complex genes. J. Med. Genet. *57*, 624–633.
11. Epi4K Consortium (2016). De Novo Mutations in SLC1A2 and CACNA1A Are Important Causes of Epileptic Encephalopathies. Am. J. Hum. Genet. *99*, 287–298.
12. EuroEPINOMICS-RES Consortium; Epilepsy Phenome/Genome Project; and Epi4K Consortium (2014). De novo mutations in synaptic transmission genes including DNM1 cause epileptic encephalopathies. Am. J. Hum. Genet. *95*, 360–370.
13. Heyne, H.O., Singh, T., Stamberger, H., Abou Jamra, R., Caglayan, H., Craiu, D., De Jonghe, P., Guerrini, R., Helbig, K.L., Koeleman, B.P.C., et al.; EuroEPINOMICS RES Consortium (2018). De novo variants in neurodevelopmental disorders with epilepsy. Nat. Genet. *50*, 1048–1053.
14. McTague, A., Howell, K.B., Cross, J.H., Kurian, M.A., and Scheffer, I.E. (2016). The genetic landscape of the epileptic encephalopathies of infancy and childhood. Lancet Neurol. *15*, 304–316.
15. Banerjee, P.N., Filippi, D., and Allen Hauser, W. (2009). The descriptive epidemiology of epilepsy-a review. Epilepsy Res. *85*, 31–45.
16. Jallon, P., Loiseau, P., and Loiseau, J. (2001). Newly diagnosed unprovoked epileptic seizures: presentation at diagnosis in CAROLE study. Coordination Active du Réseau Observatoire Longitudinal de l' Epilepsie. Epilepsia *42*, 464–475.
17. Jallon, P., and Latour, P. (2005). Epidemiology of idiopathic generalized epilepsies. Epilepsia *46* (*Suppl 9*), 10–14.
18. Petrovski, S., Wang, Q., Heinzen, E.L., Allen, A.S., and Goldstein, D.B. (2013). Genic intolerance to functional variation and the interpretation of personal genomes. PLoS Genet. *9*, e1003709.
19. Samocha, K.E., Robinson, E.B., Sanders, S.J., Stevens, C., Sabo, A., McGrath, L.M., Kosmicki, J.A., Rehnström, K., Mallick, S., Kirby, A., et al. (2014). A framework for the interpretation of de novo mutation in human disease. Nat. Genet. *46*, 944–950.
20. Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., et al.; Genome Aggregation Database Consortium (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. Nature *581*, 434–443.
21. Bennett, C.A., Petrovski, S., Oliver, K.L., and Berkovic, S.F. (2017). ExACtly zero or once: A clinically helpful guide to assessing genetic variants in mild epilepsies. Neurol. Genet. *3*, e163.
22. Gussow, A.B., Petrovski, S., Wang, Q., Allen, A.S., and Goldstein, D.B. (2016). The intolerance to functional genetic variation of protein domains predicts the localization of pathogenic mutations within genes. Genome Biol. *17*, 9.
23. Larsen, J., Carvill, G.L., Gardella, E., Kluger, G., Schmiedel, G., Barisic, N., Depienne, C., Brilstra, E., Mang, Y., Nielsen, J.E., et al.; EuroEPINOMICS RES Consortium CRP (2015). The phenotypic spectrum of SCN8A encephalopathy. Neurology *84*, 480–489.
24. Stamberger, H., Nikanorova, M., Willemsen, M.H., Accorsi, P., Angriman, M., Baier, H., Benkel-Herrenbrueck, I., Benoit, V., Budetta, M., Caliebe, A., et al. (2016). STXBP1 encephalopathy: A neurodevelopmental disorder including epilepsy. Neurology *86*, 954–962.
25. Heron, S.E., and Dibbens, L.M. (2013). Role of PRRT2 in common paroxysmal neurological disorders: a gene with remarkable pleiotropy. J. Med. Genet. *50*, 133–139.
26. Leen, W.G., Klepper, J., Verbeek, M.M., Leferink, M., Hofste, T., van Engelen, B.G., Wevers, R.A., Arthur, T., Bahi-Buisson, N., Ballhausen, D., et al. (2010). Glucose transporter-1 deficiency syndrome: the expanding clinical and genetic spectrum of a treatable disorder. Brain *133*, 655–670.
27. Wolff, M., Johannesen, K.M., Hedrich, U.B.S., Masnada, S., Rubboli, G., Gardella, E., Lesca, G., Ville, D., Milh, M., Villard, L., et al. (2017). Genetic and phenotypic heterogeneity suggest therapeutic implications in SCN2A-related disorders. Brain *140*, 1316–1336.
28. Blanchard, M.G., Willemsen, M.H., Walker, J.B., Dib-Hajj, S.D., Waxman, S.G., Jongmans, M.C., Kleefstra, T., van de Warrenburg, B.P., Praamstra, P., Nicolai, J., et al. (2015). De novo gain-of-function and loss-of-function mutations of SCN8A in patients with intellectual disabilities and epilepsy. J. Med. Genet. *52*, 330–337.
29. He, N., Lin, Z.J., Wang, J., Wei, F., Meng, H., Liu, X.R., Chen, Q., Su, T., Shi, Y.W., Yi, Y.H., and Liao, W.P. (2019). Evaluating the pathogenic potential of genes with de novo variants in epileptic encephalopathies. Genet. Med. *21*, 17–27.
30. Gelfman, S., Dugger, S., de Araujo Martins Moreno, C., Ren, Z., Wolock, C.J., Shneider, N.A., Phatnani, H., Cirulli, E.T., Lasseigne, B.N., Harris, T., et al. (2019). A new approach for rare variation collapsing on functional protein domains implicates specific genic regions in ALS. Genome Res. *29*, 809–818.
31. Hayeck, T.J., Stong, N., Wolock, C.J., Copeland, B., Kamalakaran, S., Goldstein, D.B., and Allen, A.S. (2019). Improved Pathogenic Variant Localization via a Hierarchical Model of Sub-regional Intolerance. Am. J. Hum. Genet. *104*, 299–309.
32. Traynelis, J., Silk, M., Wang, Q., Berkovic, S.F., Liu, L., Ascher, D.B., Balding, D.J., and Petrovski, S. (2017). Optimizing genomic medicine in epilepsy through a gene-customized approach to missense variant interpretation. Genome Res. *27*, 1715–1729.
33. Zhang, J., Kim, E.C., Chen, C., Procko, E., Pant, S., Lam, K., Patel, J., Choi, R., Hong, M., Joshi, D., et al. (2020). Identifying mutation hotspots reveals pathogenetic mechanisms of KCNQ2 epileptic encephalopathy. Sci. Rep. *10*, 4756.
34. Myers, C.T., Hollingsworth, G., Muir, A.M., Schneider, A.L., Thuesmunn, Z., Knupp, A., King, C., Lacroix, A., Mehaffey, M.G., Berkovic, S.F., et al. (2018). Parental Mosaicism in "De Novo" Epileptic Encephalopathies. N. Engl. J. Med. *378*, 1646–1648.
35. de Lange, I.M., Koudijs, M.J., van 't Slot, R., Gunning, B., Sonsma, A.C.M., van Gemert, L.J.J.M., Mulder, F., Carbo, E.C., van Kempen, M.J.A., Verbeek, N.E., et al. (2018).

Mosaicism of de novo pathogenic SCN1A variants in epilepsy is a frequent phenomenon that correlates with variable phenotypes. Epilepsia *59*, 690–703.
36. Winawer, M.R., Griffin, N.G., Samanamud, J., Baugh, E.H., Rathakrishnan, D., Ramalingam, S., Zagzag, D., Schevon, C.A., Dugan, P., Hegde, M., et al. (2018). Somatic SLC35A2 variants in the brain are associated with intractable neocortical epilepsy. Ann. Neurol. *83*, 1133–1146.
37. Kim, J.K., Cho, J., Kim, S.H., Kang, H.C., Kim, D.S., Kim, V.N., and Lee, J.H. (2019). Brain somatic mutations in MTOR reveal translational dysregulations underlying intractable focal epilepsy. J. Clin. Invest. *129*, 4207–4223.
38. Landrum, M.J., Lee, J.M., Benson, M., Brown, G.R., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Jang, W., et al. (2018). ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res. *46* (D1), D1062–D1067.
39. Landrum, M.J., Lee, J.M., Riley, G.R., Jang, W., Rubinstein, W.S., Church, D.M., and Maglott, D.R. (2014). ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res. *42*, D980–D985.
40. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al.; 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*, 405–424.
41. Abou-Khalil, B., Alldredge, B., Bautista, J., Berkovic, S., Bluvstein, J., Boro, A., Cascino, G., Consalvo, D., Cristofaro, S., Crumrine, P., et al.; EPGP Collaborative (2013). The epilepsy phenome/genome project. Clin. Trials *10*, 568–586.
42. Miller, N.A., Farrow, E.G., Gibson, M., Willig, L.K., Twist, G., Yoo, B., Marrs, T., Corder, S., Krivohlavek, L., Walter, A., et al. (2015). A 26-hour system of highly sensitive whole genome sequencing for emergency management of genetic diseases. Genome Med. *7*, 100.
43. McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., and DePristo, M.A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. *20*, 1297–1303.
44. Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr. Protoc. Bioinformatics *43*, 11.10.11–11.10.33.
45. Cingolani, P., Platts, A., Wang, L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) *6*, 80–92.
46. Ioannidis, N.M., Rothstein, J.H., Pejaver, V., Middha, S., McDonnell, S.K., Baheti, S., Musolf, A., Li, Q., Holzinger, E., Karyadi, D., et al. (2016). REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am. J. Hum. Genet. *99*, 877–885.
47. Ren, Z., Povysil, G., Hostyk, J.A., Cui, H., Bhardwaj, N., and Goldstein, D.B. (2021). ATAV: a comprehensive platform for population-scale genomic analyses. BMC Bioinformatics *22*, 149.
48. Pruitt, K.D., Harrow, J., Harte, R.A., Wallin, C., Diekhans, M., Maglott, D.R., Searle, S., Farrell, C.M., Loveland, J.E., Ruef, B.J., et al. (2009). The consensus coding sequence (CCDS) project: Identifying a common protein-coding gene set for the human and mouse genomes. Genome Res. *19*, 1316–1323.
49. Jun, G., Flickinger, M., Hetrick, K.N., Romm, J.M., Doheny, K.F., Abecasis, G.R., Boehnke, M., and Kang, H.M. (2012). Detecting and estimating contamination of human DNA samples in sequencing and array-based genotype data. Am. J. Hum. Genet. *91*, 839–848.
50. Sayers, E.W., Barrett, T., Benson, D.A., Bolton, E., Bryant, S.H., Canese, K., Chetvernin, V., Church, D.M., DiCuccio, M., Federhen, S., et al. (2011). Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. *39*, D38–D51.
51. Manichaikul, A., Mychaleckyj, J.C., Rich, S.S., Daly, K., Sale, M., and Chen, W.M. (2010). Robust relationship inference in genome-wide association studies. Bioinformatics *26*, 2867–2873.
52. Petrovski, S., Todd, J.L., Durheim, M.T., Wang, Q., Chien, J.W., Kelly, F.L., Frankel, C., Mebane, C.M., Ren, Z., Bridgers, J., et al. (2017). An Exome Sequencing Study to Assess the Role of Rare Genetic Variation in Pulmonary Fibrosis. Am. J. Respir. Crit. Care Med. *196*, 82–93.
53. Gravel, S., Henn, B.M., Gutenkunst, R.N., Indap, A.R., Marth, G.T., Clark, A.G., Yu, F., Gibbs, R.A., Bustamante, C.D.; and 1000 Genomes Project (2011). Demographic history and rare allele sharing among human populations. Proc. Natl. Acad. Sci. USA *108*, 11983–11988.
54. Cameron-Christie, S., Wolock, C.J., Groopman, E., Petrovski, S., Kamalakaran, S., Povysil, G., Vitsios, D., Zhang, M., Fleckner, J., March, R.E., et al. (2019). Exome-Based Rare-Variant Analyses in CKD. J. Am. Soc. Nephrol. *30*, 1109–1122.
55. Blondel, V.D., Guillaume, J.L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. J. Stat. Mech. *2008*, P10008.
56. Povysil, G., Chazara, O., Carss, K.J., Deevi, S.V.V., Wang, Q., Armisen, J., Paul, D.S., Granger, C.B., Kjekshus, J., Aggarwal, V., et al. (2021). Assessing the Role of Rare Genetic Variation in Patients With Heart Failure. JAMA Cardiol. *6*, 379–386.
57. McInnes, L., Healy, J., and Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. arXiv, 1802.03426. https://arxiv.org/abs/1802.03426.
58. Diaz-Papkovich, A., Anderson-Trocmé, L., Ben-Eghan, C., and Gravel, S. (2019). UMAP reveals cryptic population structure and phenotype heterogeneity in large genomic cohorts. PLoS Genet. *15*, e1008432.
59. Dai, C.L., Vazifeh, M.M., Yeang, C.H., Tachet, R., Wells, R.S., Vilar, M.G., Daly, M.J., Ratti, C., and Martin, A.R. (2020). Population Histories of the United States Revealed through Fine-Scale Migration and Haplotype Analysis. Am. J. Hum. Genet. *106*, 371–388.
60. Cirulli, E.T., and Goldstein, D.B. (2010). Uncovering the roles of rare variants in common disease through whole-genome sequencing. Nat. Rev. Genet. *11*, 415–425.
61. Krusche, P., Trigg, L., Boutros, P.C., Mason, C.E., De La Vega, F.M., Moore, B.L., Gonzalez-Porta, M., Eberle, M.A., Tezak, Z.,

Lababidi, S., et al.; Global Alliance for Genomics and Health Benchmarking Team (2019). Best practices for benchmarking germline small-variant calls in human genomes. Nat. Biotechnol. *37*, 555–560.
62. Cummings, B.B., Karczewski, K.J., Kosmicki, J.A., Seaby, E.G., Watts, N.A., Singer-Berk, M., Mudge, J.M., Karjalainen, J., Satterstrom, F.K., O'Donnell-Luria, A.H., et al.; Genome Aggregation Database Production Team; and Genome Aggregation Database Consortium (2020). Transcript expression-aware annotation improves rare variant interpretation. Nature *581*, 452–458.
63. Xiang, J., Yang, J., Chen, L., Chen, Q., Yang, H., Sun, C., Zhou, Q., and Peng, Z. (2020). Reinterpretation of common pathogenic variants in ClinVar revealed a high proportion of downgrades. Sci. Rep. *10*, 331.
64. Shah, N., Hou, Y.C., Yu, H.C., Sainger, R., Caskey, C.T., Venter, J.C., and Telenti, A. (2018). Identification of Misclassified ClinVar Variants via Disease Population Prevalence. Am. J. Hum. Genet. *102*, 609–619.
65. Rehm, H.L., Berg, J.S., Brooks, L.D., Bustamante, C.D., Evans, J.P., Landrum, M.J., Ledbetter, D.H., Maglott, D.R., Martin, C.L., Nussbaum, R.L., et al.; ClinGen (2015). ClinGen–the Clinical Genome Resource. N. Engl. J. Med. *372*, 2235–2242.
66. Mantel, N., and Haenszel, W. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. J. Natl. Cancer Inst. *22*, 719–748.
67. Cochran, W.G. (1954). Some Methods for Strengthening the Common X2 Tests. Biometrics *10*, 417–451.
68. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al.; Exome Aggregation Consortium (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285–291.
69. Hu, Y.J., Liao, P., Johnston, H.R., Allen, A.S., and Satten, G.A. (2016). Testing Rare-Variant Association without Calling Genotypes Allows for Systematic Differences in Sequencing between Cases and Controls. PLoS Genet. *12*, e1006040.
70. Jay, J.J., and Brouwer, C. (2016). Lollipops in the Clinic: Information Dense Mutation Plots for Precision Medicine. PLoS ONE *11*, e0160519.
71. Silk, M., Petrovski, S., and Ascher, D.B. (2019). MTR-Viewer: identifying regions within genes under purifying selection. Nucleic Acids Res. *47* (W1), W121–W126.
72. Goode, D.L., Cooper, G.M., Schmutz, J., Dickson, M., Gonzales, E., Tsai, M., Karra, K., Davydov, E., Batzoglou, S., Myers, R.M., and Sidow, A. (2010). Evolutionary constraint facilitates interpretation of genetic variation in resequenced human genomes. Genome Res. *20*, 301–310.
73. Davydov, E.V., Goode, D.L., Sirota, M., Cooper, G.M., Sidow, A., and Batzoglou, S. (2010). Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS Comput. Biol. *6*, e1001025.
74. R Core Team (2019). R: A Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing).
75. McGill, R., Tukey, J.W., and Larsen, W.A. (1978). Variations of box plots. Am. Stat. *32*, 12–16.
76. de Lange, I.M., Helbig, K.L., Weckhuysen, S., Møller, R.S., Velinov, M., Dolzhanskaya, N., Marsh, E., Helbig, I., Devinsky, O., Tang, S., et al.; EuroEPINOMICS-RES MAE working group (2016). De novo mutations of KIAA2022 in females cause intellectual disability and intractable epilepsy. J. Med. Genet. *53*, 850–858.
77. Fujiwara, T., Sugawara, T., Mazaki-Miyazaki, E., Takahashi, Y., Fukushima, K., Watanabe, M., Hara, K., Morikawa, T., Yagi, K., Yamakawa, K., and Inoue, Y. (2003). Mutations of sodium channel alpha subunit type 1 (SCN1A) in intractable childhood epilepsies with frequent generalized tonic-clonic seizures. Brain *126*, 531–546.
78. Claes, L., Ceulemans, B., Audenaert, D., Smets, K., Löfgren, A., Del-Favero, J., Ala-Mello, S., Basel-Vanagaite, L., Plecko, B., Raskin, S., et al. (2003). De novo SCN1A mutations are a major cause of severe myoclonic epilepsy of infancy. Hum. Mutat. *21*, 615–621.
79. Carvill, G.L., Weckhuysen, S., McMahon, J.M., Hartmann, C., Møller, R.S., Hjalgrim, H., Cook, J., Geraghty, E., O'Roak, B.J., Petrou, S., et al. (2014). GABRA1 and STXBP1: novel genetic causes of Dravet syndrome. Neurology *82*, 1245–1253.
80. Endele, S., Rosenberger, G., Geider, K., Popp, B., Tamer, C., Stefanova, I., Milh, M., Kortüm, F., Fritsch, A., Pientka, F.K., et al. (2010). Mutations in GRIN2A and GRIN2B encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes. Nat. Genet. *42*, 1021–1026.
81. Hoffbuhr, K., Devaney, J.M., LaFleur, B., Sirianni, N., Scacheri, C., Giron, J., Schuette, J., Innis, J., Marino, M., Philippart, M., et al. (2001). MeCP2 mutations in children with and without the phenotype of Rett syndrome. Neurology *56*, 1486–1495.
82. Schubert, J., Siekierska, A., Langlois, M., May, P., Huneau, C., Becker, F., Muhle, H., Suls, A., Lemke, J.R., de Kovel, C.G., et al.; EuroEPINOMICS RES Consortium (2014). Mutations in STX1B, encoding a presynaptic protein, cause fever-associated epilepsy syndromes. Nat. Genet. *46*, 1327–1332.
83. Krey, I., Krois-Neudenberger, J., Hentschel, J., Syrbe, S., Polster, T., Hanker, B., Fiedler, B., Kurlemann, G., and Lemke, J.R. (2020). Genotype-phenotype correlation on 45 individuals with West syndrome. Eur. J. Paediatr. Neurol. *25*, 134–138.
84. Dibbens, L.M., de Vries, B., Donatello, S., Heron, S.E., Hodgson, B.L., Chintawar, S., Crompton, D.E., Hughes, J.N., Bellows, S.T., Klein, K.M., et al. (2013). Mutations in DEPDC5 cause familial focal epilepsy with variable foci. Nat. Genet. *45*, 546–551.
85. Weaving, L.S., Christodoulou, J., Williamson, S.L., Friend, K.L., McKenzie, O.L., Archer, H., Evans, J., Clarke, A., Pelka, G.J., Tam, P.P., et al. (2004). Mutations of CDKL5 cause a severe neurodevelopmental disorder with infantile spasms and mental retardation. Am. J. Hum. Genet. *75*, 1079–1093.
86. Heron, S.E., Smith, K.R., Bahlo, M., Nobili, L., Kahana, E., Licchetta, L., Oliver, K.L., Mazarib, A., Afawi, Z., Korczyn, A., et al. (2012). Missense mutations in the sodium-gated potassium channel gene KCNT1 cause severe autosomal dominant nocturnal frontal lobe epilepsy. Nat. Genet. *44*, 1188–1190.
87. Barcia, G., Fleming, M.R., Deligniere, A., Gazula, V.R., Brown, M.R., Langouet, M., Chen, H., Kronengold, J., Abhyankar, A., Cilio, R., et al. (2012). De novo gain-of-function KCNT1 channel mutations cause malignant migrating partial seizures of infancy. Nat. Genet. *44*, 1255–1259.

88. Dell'Angelica, E.C., Ohno, H., Ooi, C.E., Rabinovich, E., Roche, K.W., and Bonifacino, J.S. (1997). AP-3: an adaptor-like protein complex with ubiquitous expression. EMBO J. *16*, 917–928.
89. Ammann, S., Schulz, A., Krägeloh-Mann, I., Dieckmann, N.M., Niethammer, K., Fuchs, S., Eckl, K.M., Plank, R., Werner, R., Altmüller, J., et al. (2016). Mutations in AP3D1 associated with immunodeficiency and seizures define a new type of Hermansky-Pudlak syndrome. Blood *127*, 997–1006.
90. Seifert, W., Meinecke, P., Krüger, G., Rossier, E., Heinritz, W., Wüsthof, A., and Horn, D. (2014). Expanded spectrum of exon 33 and 34 mutations in SRCAP and follow-up in patients with Floating-Harbor syndrome. BMC Med. Genet. *15*, 127.
91. Nikkel, S.M., Dauber, A., de Munnik, S., Connolly, M., Hood, R.L., Caluseriu, O., Hurst, J., Kini, U., Nowaczyk, M.J., Afenjar, A., et al.; FORGE Canada Consortium (2013). The phenotype of Floating-Harbor syndrome: clinical characterization of 52 individuals with mutations in exon 34 of SRCAP. Orphanet J. Rare Dis. *8*, 63.
92. Abou-Khalil, B., Ge, Q., Desai, R., Ryther, R., Bazyk, A., Bailey, R., Haines, J.L., Sutcliffe, J.S., and George, A.L., Jr. (2001). Partial and generalized epilepsy with febrile seizures plus and a novel SCN1A mutation. Neurology *57*, 2265–2272.
93. Cossette, P., Liu, L., Brisebois, K., Dong, H., Lortie, A., Vanasse, M., Saint-Hilaire, J.M., Carmant, L., Verner, A., Lu, W.Y., et al. (2002). Mutation of GABRA1 in an autosomal dominant form of juvenile myoclonic epilepsy. Nat. Genet. *31*, 184–189.
94. Strehlow, V., Heyne, H.O., Vlaskamp, D.R.M., Marwick, K.F.M., Rudolf, G., de Bellescize, J., Biskup, S., Brilstra, E.H., Brouwer, O.F., Callenbach, P.M.C., et al.; GRIN2A study group (2019). GRIN2A-related disorders: genotype and functional consequence predict phenotype. Brain *142*, 80–92.
95. Carvill, G.L., McMahon, J.M., Schneider, A., Zemel, M., Myers, C.T., Saykally, J., Nguyen, J., Robbiano, A., Zara, F., Specchio, N., et al.; EuroEPINOMICS Rare Epilepsy Syndrome Myoclonic-Astatic Epilepsy & Dravet working group (2015). Mutations in the GABA Transporter SLC6A1 Cause Epilepsy with Myoclonic-Atonic Seizures. Am. J. Hum. Genet. *96*, 808–815.
96. Johannesen, K.M., Gardella, E., Linnankivi, T., Courage, C., de Saint Martin, A., Lehesjoki, A.E., Mignot, C., Afenjar, A., Lesca, G., Abi-Warde, M.T., et al. (2018). Defining the phenotypic spectrum of SLC6A1 mutations. Epilepsia *59*, 389–402.
97. Lafrenière, R.G., Cader, M.Z., Poulin, J.F., Andres-Enguix, I., Simoneau, M., Gupta, N., Boisvert, K., Lafrenière, F., McLaughlan, S., Dubé, M.P., et al. (2010). A dominant-negative mutation in the TRESK potassium channel is linked to familial migraine with aura. Nat. Med. *16*, 1157–1160.
98. Sun, L., Shi, L., Li, W., Yu, W., Liang, J., Zhang, H., Yang, X., Wang, Y., Li, R., Yao, X., et al. (2009). JFK, a Kelch domain-containing F-box protein, links the SCF complex to p53 regulation. Proc. Natl. Acad. Sci. USA *106*, 10195–10200.
99. Gorman, K.M., Meyer, E., Grozeva, D., Spinelli, E., McTague, A., Sanchis-Juan, A., Carss, K.J., Bryant, E., Reich, A., Schneider, A.L., et al.; Deciphering Developmental Disorders Study; UK10K Consortium; and NIHR BioResource (2019). Bi-allelic Loss-of-Function CACNA1B Mutations in Progressive Epilepsy-Dyskinesia. Am. J. Hum. Genet. *104*, 948–956.
100. Gardella, E., and Møller, R.S. (2019). Phenotypic and genetic spectrum of SCN8A-related disorders, treatment options, and outcomes. Epilepsia *60* (*Suppl 3*), S77–S85.
101. Ricos, M.G., Hodgson, B.L., Pippucci, T., Saidin, A., Ong, Y.S., Heron, S.E., Licchetta, L., Bisulli, F., Bayly, M.A., Hughes, J., et al.; Epilepsy Electroclinical Study Group (2016). Mutations in the mammalian target of rapamycin pathway regulators NPRL2 and NPRL3 cause focal epilepsy. Ann. Neurol. *79*, 120–131.
102. Krasniqi, S., and Daci, A. (2019). Role of the Angiotensin Pathway and its Target Therapy in Epilepsy Management. Int. J. Mol. Sci. *20*, 726.
103. Gasparini, S., Ferlazzo, E., Sueri, C., Cianci, V., Ascoli, M., Cavalli, S.M., Beghi, E., Belcastro, V., Bianchi, A., Benna, P., et al.; Epilepsy Study Group of the Italian Neurological Society (2019). Hypertension, seizures, and epilepsy: a review on pathophysiology and management. Neurol. Sci. *40*, 1775–1783.
104. Liu, Y.J., Chen, J., Li, X., Zhou, X., Hu, Y.M., Chu, S.F., Peng, Y., and Chen, N.H. (2019). Research progress on adenosine in central nervous system diseases. CNS Neurosci. Ther. *25*, 899–910.
105. Chen, J.F., Eltzschig, H.K., and Fredholm, B.B. (2013). Adenosine receptors as drug targets–what are the challenges? Nat. Rev. Drug Discov. *12*, 265–286.
106. Lesko, S.L., and Rouhana, L. (2020). Dynein assembly factor with WD repeat domains 1 (DAW1) is required for the function of motile cilia in the planarian Schmidtea mediterranea. Dev. Growth Differ. *62*, 423–437.
107. Gupta, A., de Bruyn, G., Tousseyn, S., Krishnan, B., Lagae, L., Agarwal, N.; and TSC Natural History Database Consortium (2020). Epilepsy and Neurodevelopmental Comorbidities in Tuberous Sclerosis Complex: A Natural History Study. Pediatr. Neurol. *106*, 10–16.
108. Lim, J.S., Gopalappa, R., Kim, S.H., Ramakrishna, S., Lee, M., Kim, W.I., Kim, J., Park, S.M., Lee, J., Oh, J.H., et al. (2017). Somatic Mutations in TSC1 and TSC2 Cause Focal Cortical Dysplasia. Am. J. Hum. Genet. *100*, 454–472.
109. Bamshad, M.J., Nickerson, D.A., and Chong, J.X. (2019). Mendelian Gene Discovery: Fast and Furious with No End in Sight. Am. J. Hum. Genet. *105*, 448–455.
110. Chubykin, A.A., Atasoy, D., Etherton, M.R., Brose, N., Kavalali, E.T., Gibson, J.R., and Südhof, T.C. (2007). Activity-dependent validation of excitatory versus inhibitory synapses by neuroligin-1 versus neuroligin-2. Neuron *54*, 919–931.
111. Finkbeiner, E., Haindl, M., and Muller, S. (2011). The SUMO system controls nucleolar partitioning of a novel mammalian ribosome biogenesis complex. EMBO J. *30*, 1067–1078.
112. Kelly, M., Park, M., Mihalek, I., Rochtus, A., Gramm, M., Pérez-Palma, E., Axeen, E.T., Hung, C.Y., Olson, H., Swanson, L., et al.; Undiagnosed Diseases Network (2019). Spectrum of neurodevelopmental disease associated with the GNAO1 guanosine triphosphate-binding region. Epilepsia *60*, 406–418.
113. Szczałuba, K., Chmielewska, J.J., Sokolowska, O., Rydzanicz, M., Szymańska, K., Feleszko, W., Włodarski, P., Biernacka, A., Murcia Pienkowski, V., Walczak, A., et al. (2018). Neurodevelopmental phenotype caused by a de novo PTPN4 single nucleotide variant disrupting protein localization in neuronal dendritic spines. Clin. Genet. *94*, 581–585.

114. Havrilla, J.M., Pedersen, B.S., Layer, R.M., and Quinlan, A.R. (2019). A map of constrained coding regions in the human genome. Nat. Genet. *51*, 88–95.
115. Samocha, K.E., Kosmicki, J.A., Karczewski, K.J., O'Donnell-Luria, A.H., Pierce-Hoffman, E., MacArthur, D.G., Neale, B.M., and Daly, M.J. (2017). Regional missense constraint improves variant deleteriousness prediction. bioRxiv. https://doi.org/10.1101/148353.
116. Hemati, P., Revah-Politi, A., Bassan, H., Petrovski, S., Bilancia, C.G., Ramsey, K., Griffin, N.G., Bier, L., Cho, M.T., Rosello, M., et al.; C4RCD Research Group; and DDD study (2018). Refining the phenotype associated with GNB1 mutations: Clinical data on 18 newly identified patients and review of the literature. Am. J. Med. Genet. A. *176*, 2259–2275.
117. Ogden, K.K., Chen, W., Swanger, S.A., McDaniel, M.J., Fan, L.Z., Hu, C., Tankovic, A., Kusumoto, H., Kosobucki, G.J., Schulien, A.J., et al. (2017). Molecular Mechanism of Disease-Associated Mutations in the Pre-M1 Helix of NMDA Receptors and Potential Rescue Pharmacology. PLoS Genet. *13*, e1006536.
118. Pérez-Palma, E., May, P., Iqbal, S., Niestroj, L.M., Du, J., Heyne, H.O., Castrillon, J.A., O'Donnell-Luria, A., Nürnberg, P., Palotie, A., et al. (2020). Identification of pathogenic variant enriched regions across genes and gene families. Genome Res. *30*, 62–71.
119. Shearer, A.E., Eppsteiner, R.W., Booth, K.T., Ephraim, S.S., Gurrola, J., 2nd, Simpson, A., Black-Ziegelbein, E.A., Joshi, S., Ravi, H., Giuffre, A.C., et al. (2014). Utilizing ethnic-specific differences in minor allele frequency to recategorize reported pathogenic deafness variants. Am. J. Hum. Genet. *95*, 445–453.
120. Whiffin, N., Minikel, E., Walsh, R., O'Donnell-Luria, A.H., Karczewski, K., Ing, A.Y., Barton, P.J.R., Funke, B., Cook, S.A., MacArthur, D., and Ware, J.S. (2017). Using high-resolution variant frequencies to empower clinical genome interpretation. Genet. Med. *19*, 1151–1158.
121. Yang, S., Lincoln, S.E., Kobayashi, Y., Nykamp, K., Nussbaum, R.L., and Topper, S. (2017). Sources of discordance among germ-line variant classifications in ClinVar. Genet. Med. *19*, 1118–1126.
122. Wright, C.F., Eberhardt, R.Y., Constantinou, P., Hurles, M.E., FitzPatrick, D.R., Firth, H.V.; and DDD Study (2021). Evaluating variants classified as pathogenic in ClinVar in the DDD Study. Genet. Med. *23*, 571–575.

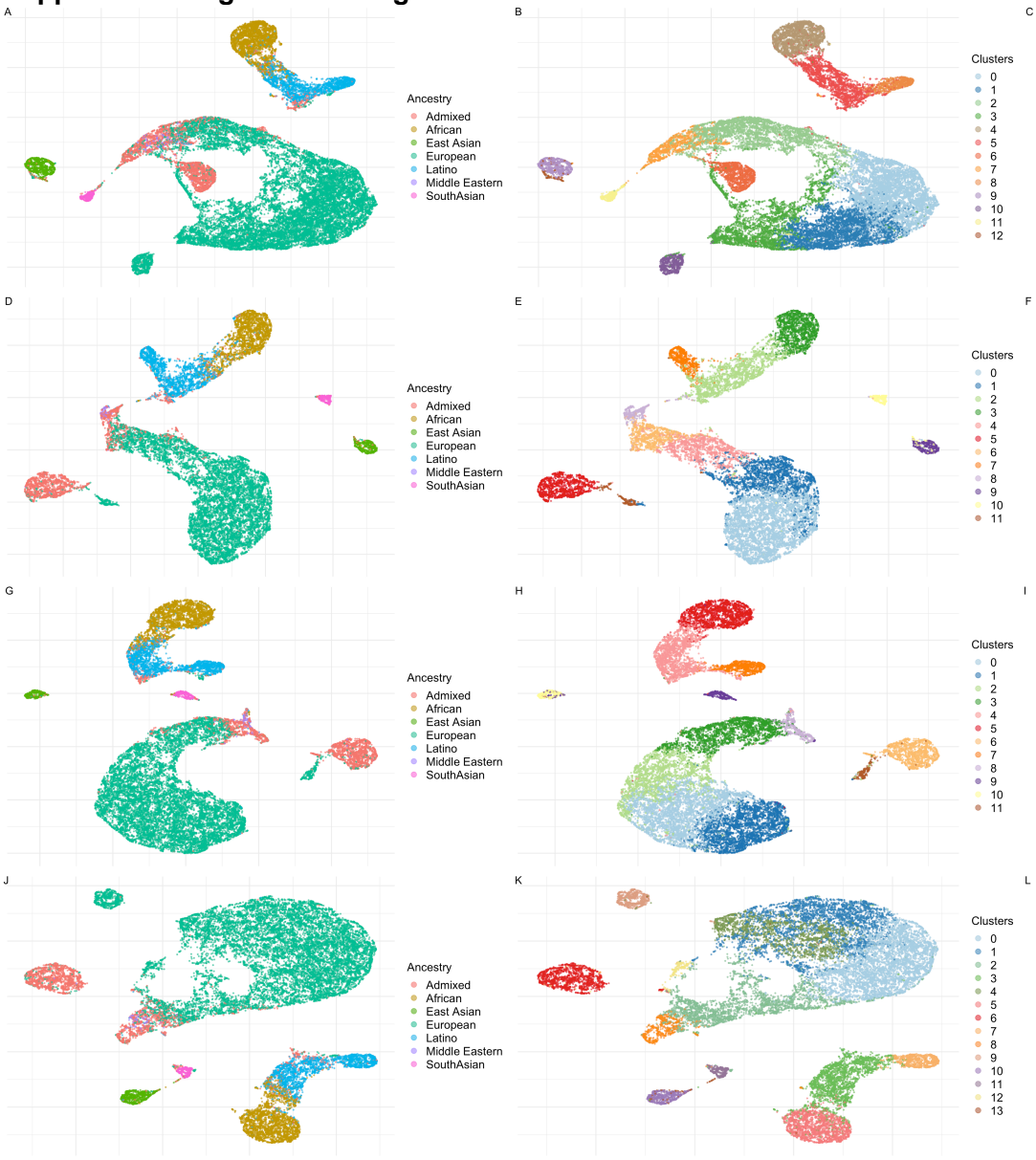
The American Journal of Human Genetics, Volume *108*

Supplemental information

Sub-genic intolerance, ClinVar, and the epilepsies: A whole-exome sequencing study of 29,165 individuals

Epi25 Collaborative

Supplemental Figures and Legends



C

Cluster	Control	DEE	GGE	NAFE
0	3155	199	1257	1756
1	2264	179	935	1275
2	1570	675	976	964
3	1595	198	769	1126
4	1783	21	361	319
5	2206	29	126	69
6	1355	14	102	65
7	302	209	303	348
8	753	23	90	50
9	235	211	125	334
10	8	42	104	613
11	237	38	106	100
12	44	17	33	31

F

Cluster	Cases	Controls
0	403	4410
1	187	2236
2	26	2174
3	21	1760
4	295	1137
5	12	1294
6	412	688
7	24	765
8	205	291
9	211	236
10	51	281
11	12	250

I

Cluster	Cases	Controls
0	1199	2883
1	1063	2349
2	795	1780
3	954	1561
4	133	2213
5	364	1756
6	104	1352
7	88	770
8	291	291
9	133	292
10	125	239
11	54	191

L

Cluster	Cases	Controls
0	1848	3369
1	1383	2106
2	1075	1810
3	100	2322
4	783	1291
5	295	1664
6	66	1360
7	49	756
8	347	304
9	623	12
10	334	238
11	85	234
12	43	175
13	31	49

Figure S1. Clustering and Geographic Ancestry.

UMAP and cluster assignments showing ancestry of case-control cohort for (A-C) all epilepsies combined, (D-F) developmental and epileptic encephalopathies (DEE), (G-I) genetic generalized epilepsy (GGE), and (J-L) non-acquired focal epilepsy (NAFE). Clusters shaded in red were excluded from the analysis.

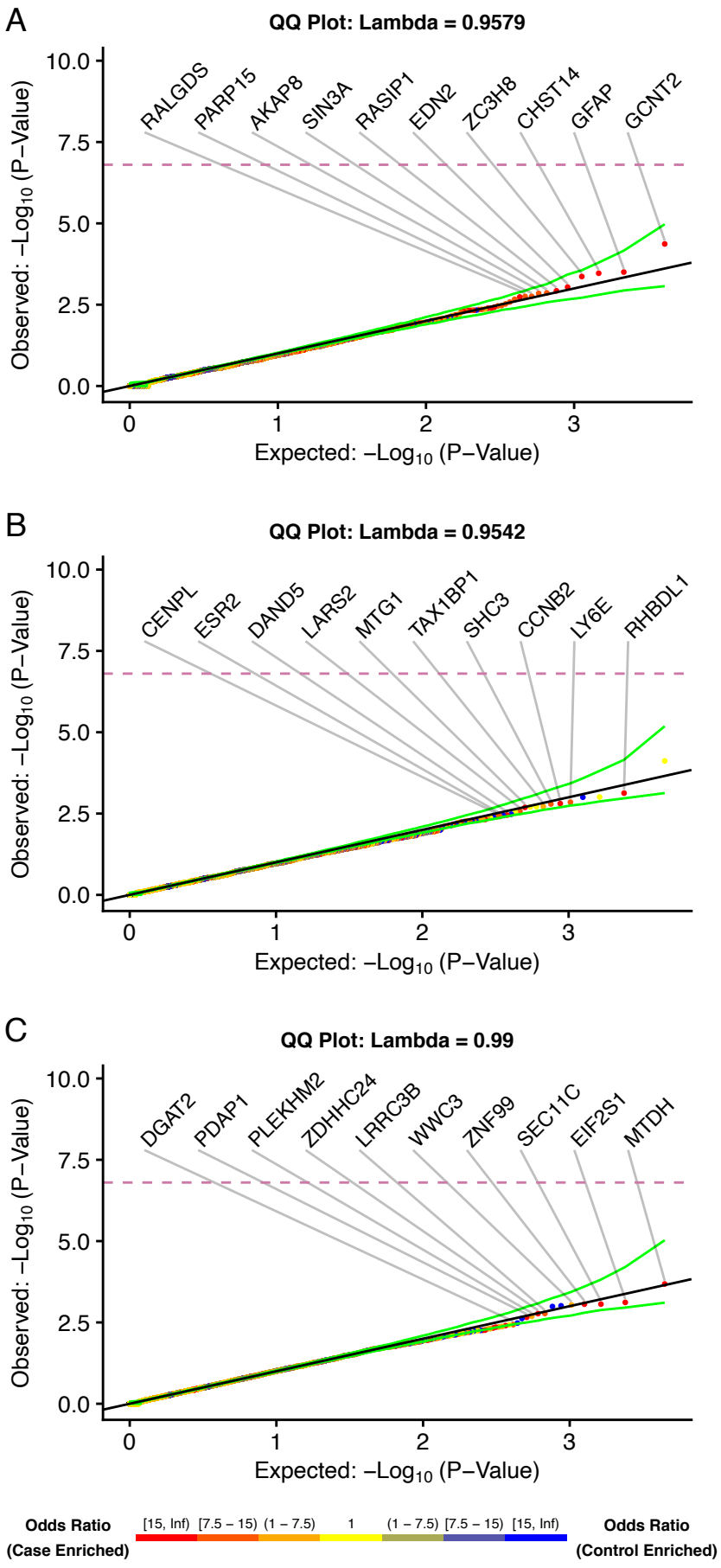

A
QQ Plot: Lambda = 0.9579
RALGDS
PARP15
AKAP8
SIN3A
RASIP1
EDN2
ZC3H8
CHST14
GFAP
GCNT2
Observed: −Log10 (P−Value)
Expected: −Log10 (P−Value)
B
QQ Plot: Lambda = 0.9542
CENPL
ESR2
DAND5
LARS2
MTG1
TAX1BP1
SHC3
CCNB2
LY6E
RHBDL1
Observed: −Log10 (P−Value)
Expected: −Log10 (P−Value)
C
QQ Plot: Lambda = 0.99
DGAT2
PDAP1
PLEKHM2
ZDHHC24
LRRC3B
WWC3
ZNF99
SEC11C
EIF2S1
MTDH
Observed: −Log10 (P−Value)
Expected: −Log10 (P−Value)
Odds Ratio
(Case Enriched)
[15, Inf)
[7.5 – 15)
(1 – 7.5)
1
(1 – 7.5)
[7.5 – 15)
[15, Inf)
Odds Ratio
(Control Enriched)

Figure S2. Synonymous Quantile-Quantile Plots

The quantile-quantile plots for the protein-coding genes with at least one case or control carrier of a synonymous variant. All variants are ultra-rare (qualifying variants were defined as a minor allele frequency of less than 0.05% in internal case and control by cluster, and absent in external reference cohorts). *P*-values were generated from the exact two-sided Cochran-Mantel-Haenszel (CMH) test by gene by cluster to indicate a different carrier status of cases in comparison to controls. Study-wide significance $p < 1.6 \times 10^{-7}$ after Bonferroni correction indicated by dashed line (see Statistical Analyses in Methods). (A) Developmental and epileptic encephalopathy (DEE), (B) genetic generalized epilepsy (GEE), and (C) non-acquired focal epilepsy (NAFE). Top ten case enriched genes are labeled. Point coloring determined by CMH odds ratio. The green lines represent the 95% confidence interval.

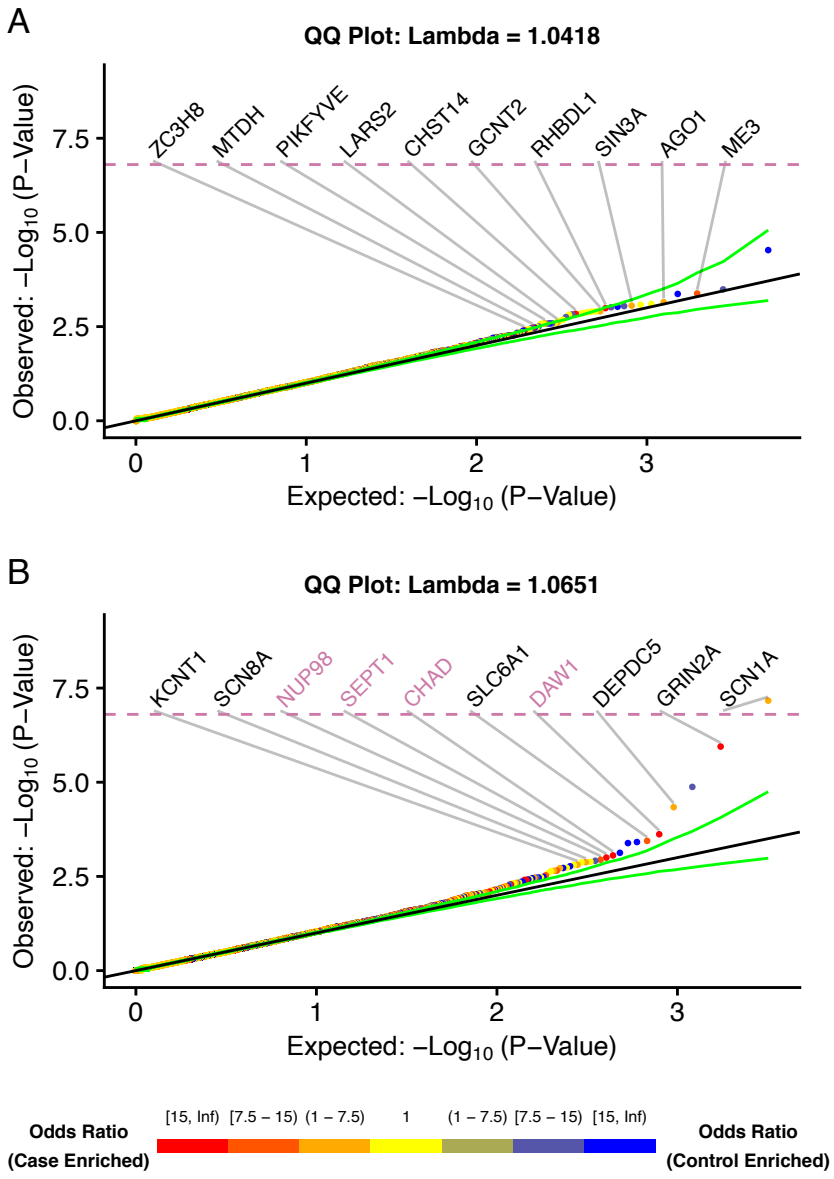


Figure S3. Quantile-Quantile Plots for the Protein-Coding Genes with at least One Case or Control Carrier for Epilepsies Combined

Qualifying variants were high quality, ultra-rare variants. *P*-values were generated from the exact two-sided Cochran-Mantel-Haenszel (CMH) test by gene by cluster to indicate a different carrier status of cases in comparison to controls. (A) Synonymous variants and (B) variants with a predicted functional effect but restricting missense variants to REVEL ≥ 0.5 (when defined). *SCN1A* ($p = 6.8 \times 10^{-8}$) achieved study-wide significance $p < 1.6 \times 10^{-7}$ after Bonferroni correction (see Statistical Analyses in Methods). Top ten case enriched genes are labeled. Point coloring determined by CMH odds ratio. Genes labeled in black are known epilepsy genes. Genes labeled in color are candidate epilepsy genes. The green lines represent the 95% confidence interval.

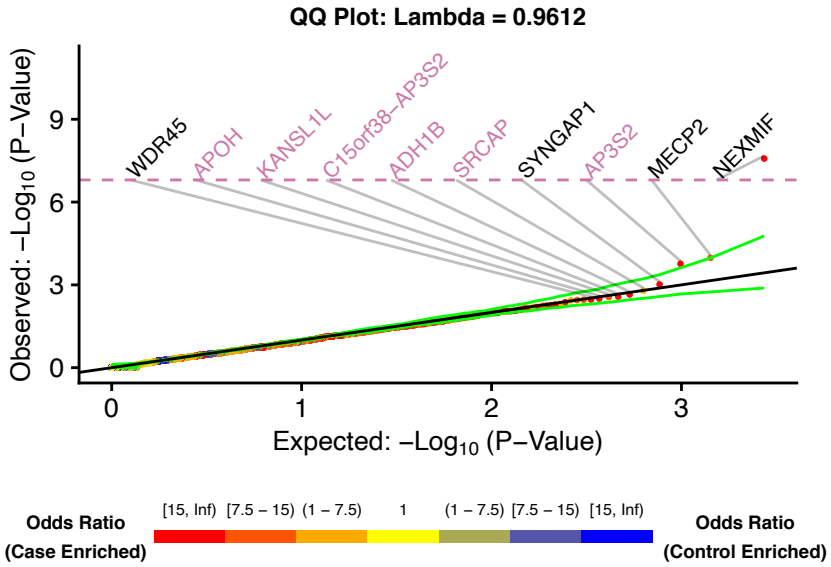


Figure S4. Quantile-Quantile Plots for the Protein-Coding Genes with at least One Case or Control Carrier Limiting Cases to DEE Cases without a Damaging Missense or Protein Truncating Variant in a OMIM Epilepsy Gene

Qualifying variants were high quality, ultra-rare variants with a predicted functional effect but restricting missense variants to REVEL ≥ 0.5 (when defined) and MTR ≤ 0.78 (when defined). Cases were Epi25 DEE cases without those that had a variant with a likely functional effect in any of the 101 genes associated with an epilepsy phenotype and dominant inheritance in OMIM (see Gene-Set Enrichment Testing in Methods, Table S5). This analysis removed 236 DEE cases from the original analysis set of 1,782 DEE. *P*-values were generated from the exact two-sided Cochran-Mantel-Haenszel (CMH) test by gene by cluster to indicate a different carrier status of cases in comparison to controls. No novel genes achieved study-wide significance $p < 1.6 \times 10^{-7}$ after Bonferroni correction (see Statistical Analyses in Methods). Point coloring determined by CMH odds ratio. Genes labeled in black are known epilepsy genes. Genes labeled in color are candidate epilepsy genes. The green lines represent the 95% confidence interval.

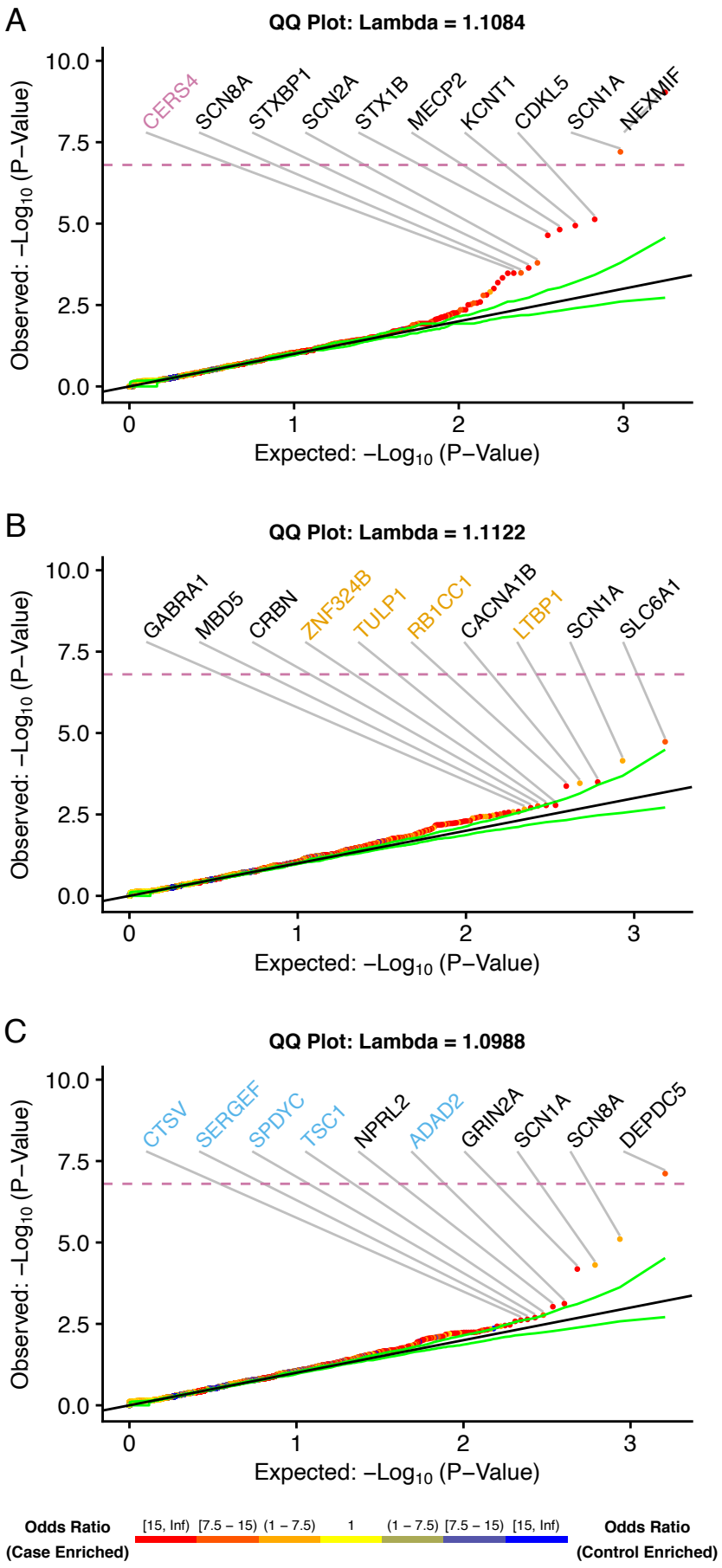

A
QQ Plot: Lambda = 1.1084
CERS4
SCN8A
STXBP1
SCN2A
STX1B
MECP2
KCNT1
CDKL5
SCN1A
NEXMIF
Observed: −Log10 (P−Value)
Expected: −Log10 (P−Value)
B
QQ Plot: Lambda = 1.1122
GABRA1
MBD5
CRBN
ZNF324B
TULP1
RB1CC1
CACNA1B
LTBP1
SCN1A
SLC6A1
Observed: −Log10 (P−Value)
Expected: −Log10 (P−Value)
C
QQ Plot: Lambda = 1.0988
CTSV
SERGEF
SPDYC
TSC1
NPRL2
ADAD2
GRIN2A
SCN1A
SCN8A
DEPDC5
Observed: −Log10 (P−Value)
Expected: −Log10 (P−Value)
Odds Ratio
(Case Enriched)
[15, Inf)
[7.5 – 15)
(1 – 7.5)
1
(1 – 7.5)
[7.5 – 15)
[15, Inf)
Odds Ratio
(Control Enriched)

Figure S5. Quantile-Quantile Plots for the Protein-Coding Genes with at least One Case or Control Carrier Limiting Missense Variants to Damaging and Intolerant

Qualifying variants were high quality, ultra-rare protein truncating variants (PTV) or missense variants with REVEL ≥ 0.5 (when defined) and MTR ≤ 0.78 (when defined). *P*-values were generated from the exact two-sided Cochran-Mantel-Haenszel (CMH) test by gene by cluster to indicate a different carrier status of cases in comparison to controls. (A) Developmental and epileptic encephalopathy (DEE) cases, (B) genetic generalized epilepsy (GEE) cases, and (C) non-acquired focal epilepsy (NAFE) cases. In the DEE analysis, *NEXMIF* ($p = 9.3 \times 10^{-10}$) and *SCN1A* ($p = 6.3 \times 10^{-8}$) achieved study-wide significance $p < 1.6 \times 10^{-7}$ after Bonferroni correction (see Statistical Analyses in Methods). In the NAFE analysis, *DEPDC5* ($p = 7.7 \times 10^{-8}$) achieved study-wide significance. Point coloring determined by CMH odds ratio. Genes labeled in black are known epilepsy genes. Genes labeled in color are candidate epilepsy genes. The green lines represent the 95% confidence interval.

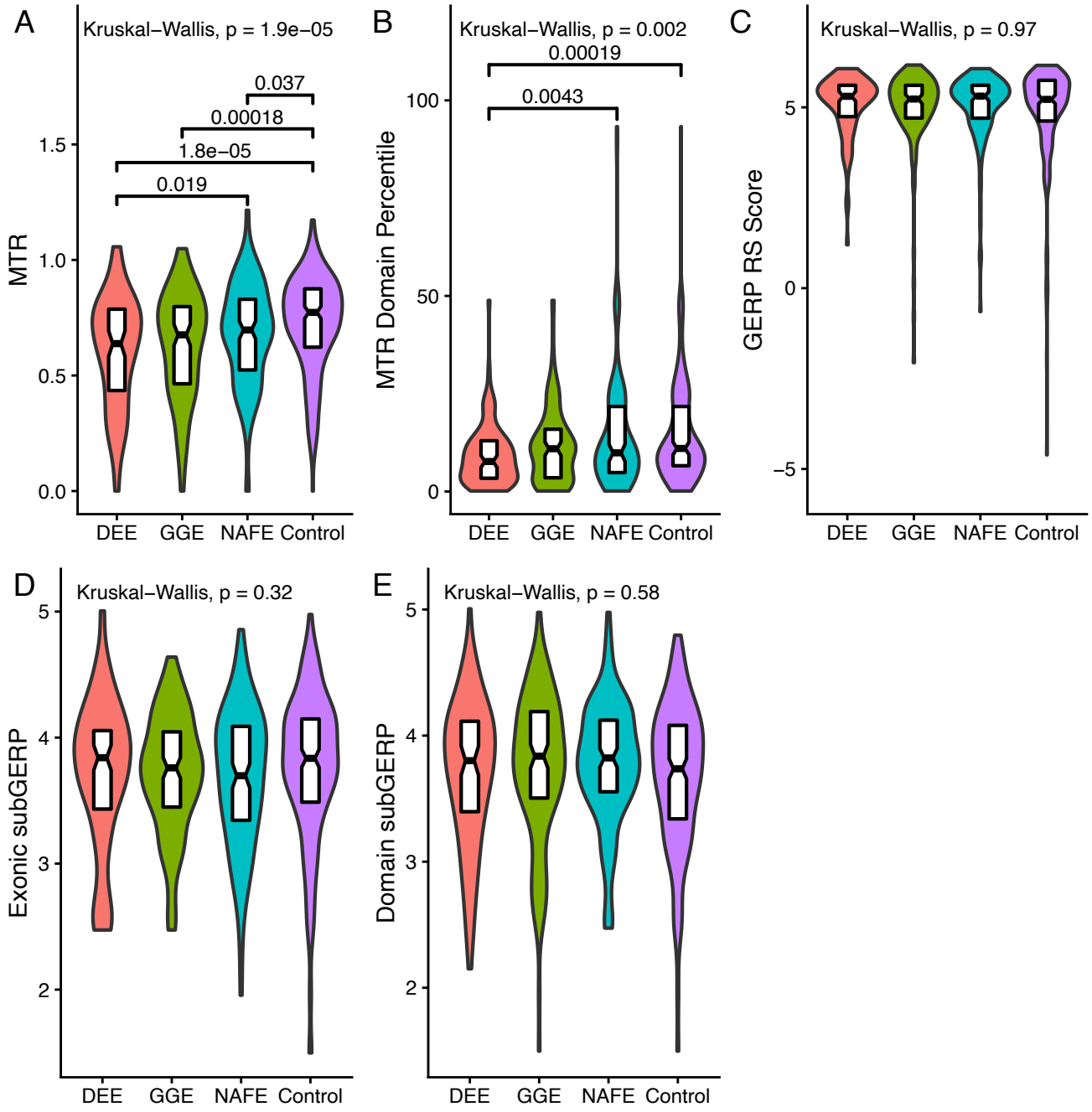


Figure S6. Direct Comparison of Intolerance and Evolutionary Constraint across the Epilepsies

Sub-genic intolerance but not sub-genic evolutionary constraint differs among the epilepsies. (A) Unweighted MTR scores of variants from individuals with epilepsy subtypes are significantly different compared to the variants from controls and DEE variants are significantly different compared to NAFE variants. Statistically insignificant comparisons ($p \geq 0.05$) are not shown. (B) Domain MTR percentile of genomic regions harboring missense variants are significantly different between DEE compared to control and DEE compared to NAFE but all other comparisons are not significant. (C) Base level evolutionary constraint score (rejected substitution score), (D) exonic subGERP and (E) domain subGERP of genomic regions harboring missense variants across epilepsies and controls do not derive from different distributions. The middle horizontal line represents the median value and the lower and upper hinges represent the 1st and 3rd quartiles. The notches in the boxplot extend 1.58 * IQR / sqrt(n), which gives approximately 95% confidence interval. Group differences determined by Kruskal-Wallis test by rank. Individual comparisons determined by Wilcoxon signed-rank test and only displayed if significant ($p < 0.05$). No multiple comparison correction performed. Plots calculated from 614 missense variants (DEE = 100, GGE = 133, and NAFE = 153, Control = 228). DEE = developmental and epileptic encephalopathy, GGE = genetic generalized epilepsy, NAFE = non-acquired focal epilepsy.

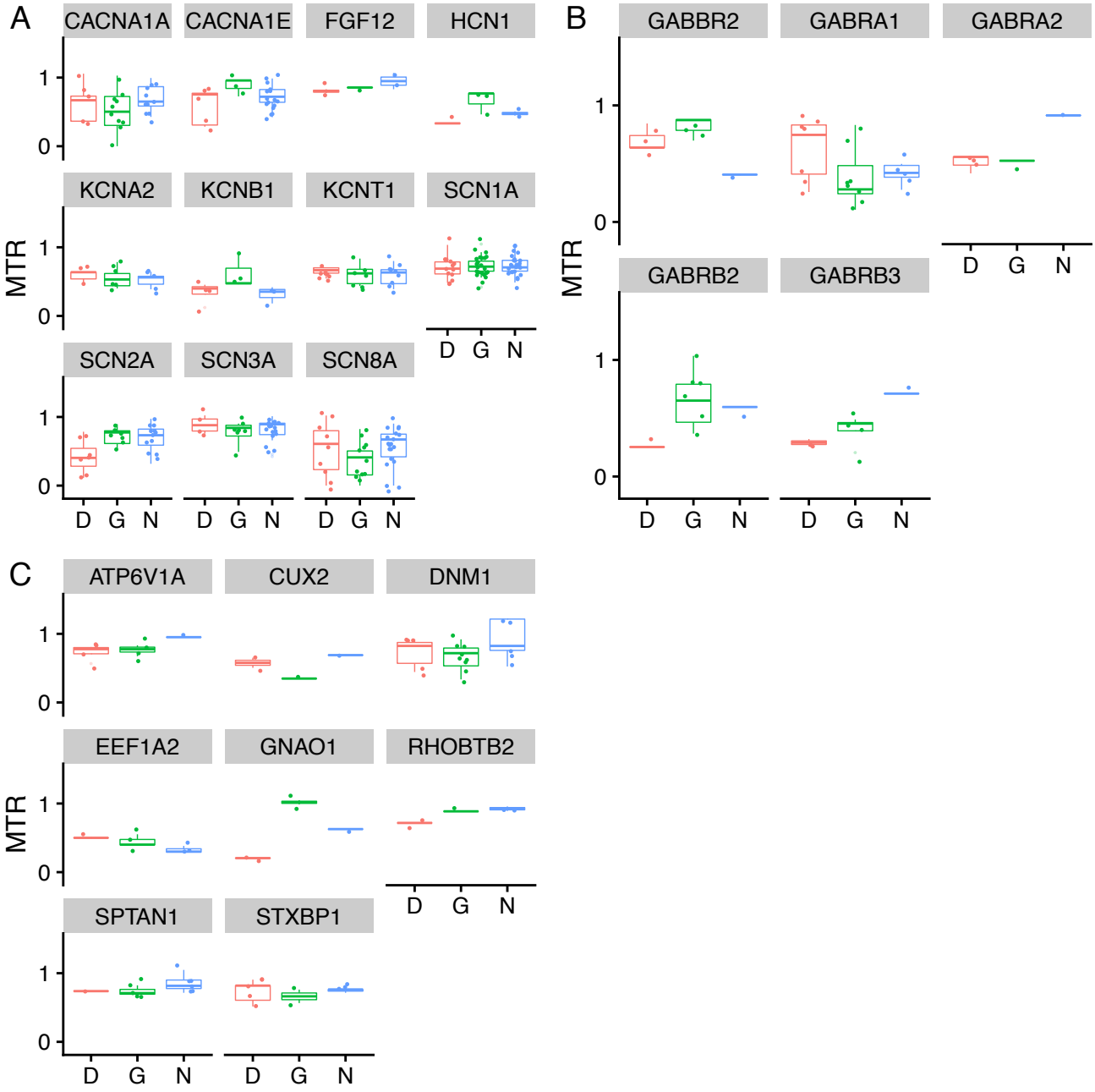
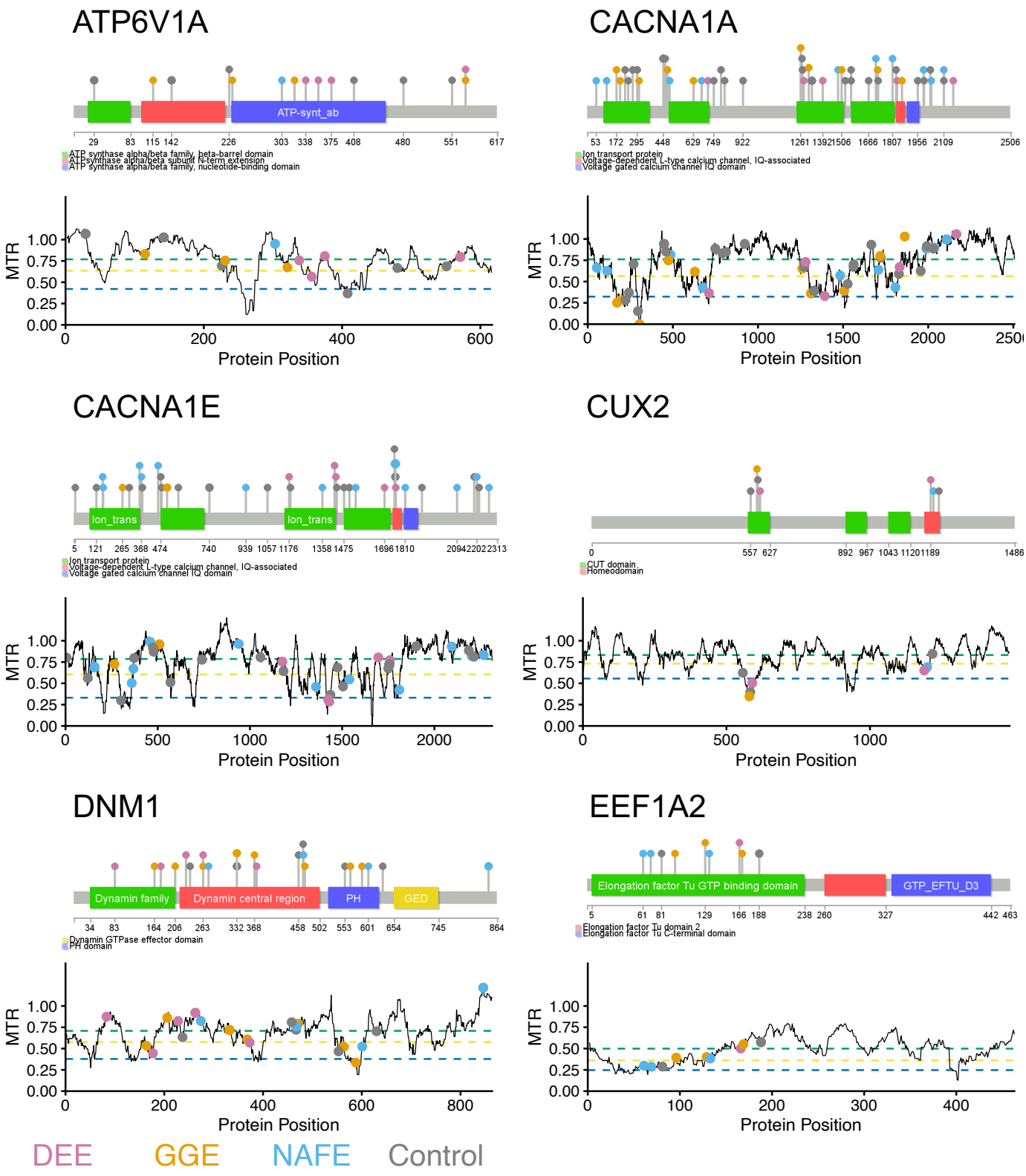
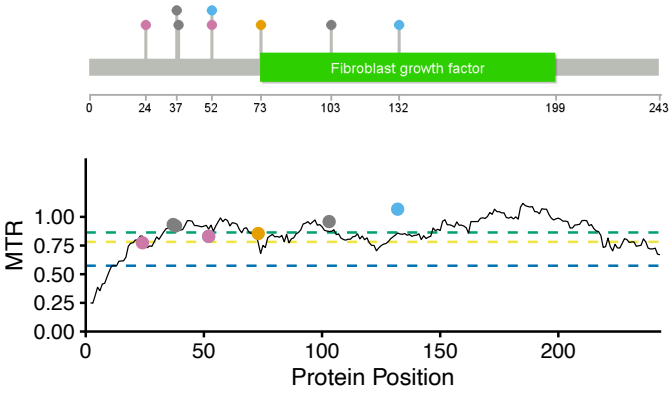


Figure S7. Sub-Genic Intolerance by Gene

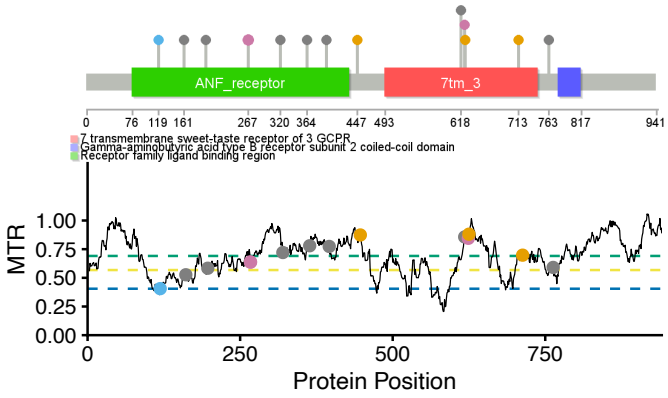
Scatter plots with box and whiskers plots showing the distribution of MTR scores of ultra-rare and damaging (REVEL ≥ 0.5) missense variants in Epi25 by gene and by epilepsy type. Gene-set created from 24 genes drawn from the 43 OMIM epileptic encephalopathy phenotype series with dominant transmission by limiting to genes harboring damaging (REVEL ≥ 0.5) missense variants in all 3 epilepsies (see Gene-Set Enrichment Testing in Methods, Table S5). The genes are divided down into (A) cation channels, (B) GABA-associated genes, and (C) other. Horizontal line represents the median value and the lower and upper hinges represent the 25th and 75th percentiles, respectively. D = developmental and epileptic encephalopathy, G = genetic generalized epilepsy, N = non-acquired focal epilepsy.


ATP6V1A
ATP-synt_ab
29 83 115 142 226 303 338 375 408 480 551 617
ATP synthase alpha/beta family, beta-barrel domain
ATPsynthase alpha/beta subunit N-term extension
ATP synthase alpha/beta family, nucleotide-binding domain
MTR
1.00
0.75
0.50
0.25
0.00
0 200 400 600
Protein Position
CACNA1A
53 172 295 448 629 749 922 1261 13921506 1666 1807 1956 2109 2506
Ion transport protein
Voltage-dependent L-type calcium channel, IQ-associated
Voltage gated calcium channel IQ domain
0 500 1000 1500 2000 250
Protein Position
CACNA1E
Ion_trans
Ion_trans
5 121 265 368 474 740 939 1057 1176 1358 1475 16961810 209422022313
Ion transport protein
Voltage-dependent L-type calcium channel, IQ-associated
Voltage gated calcium channel IQ domain
0 500 1000 1500 2000
Protein Position
CUX2
0 557 627 892 967 1043 11201189 1486
CUT domain
Homeodomain
0 500 1000
Protein Position
DNM1
Dynamin family
Dynamin central region
PH
GED
34 83 164 206 263 332 368 458 502 553 601 654 745 864
Dynamin GTPase effector domain
PH domain
0 200 400 600 800
Protein Position
EEF1A2
Elongation factor Tu GTP binding domain
GTP_EFTU_D3
5 61 81 129 166 188 238 260 327 442 463
Elongation factor Tu domain 2
Elongation factor Tu C-terminal domain
0 100 200 300 400
Protein Position
DEE
GGE
NAFE
Control

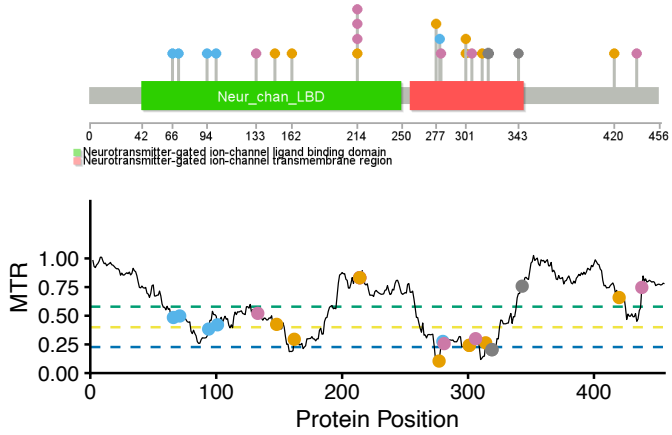
FGF12



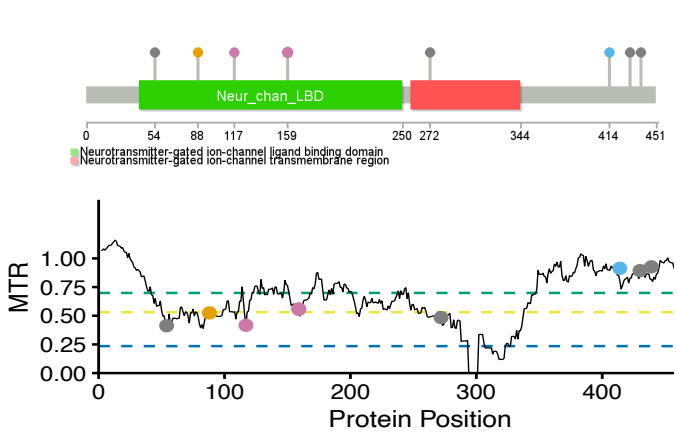
GABBR2



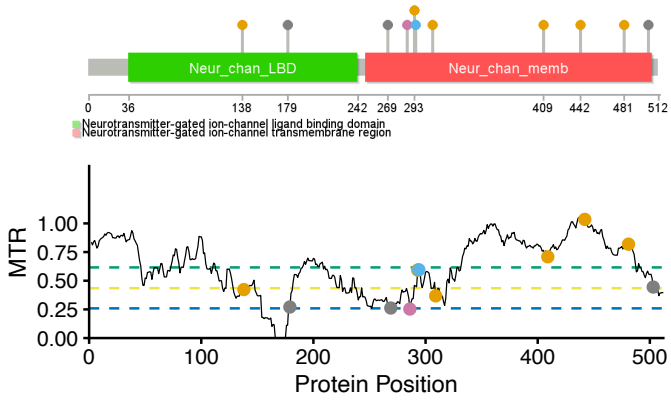
GABRA1



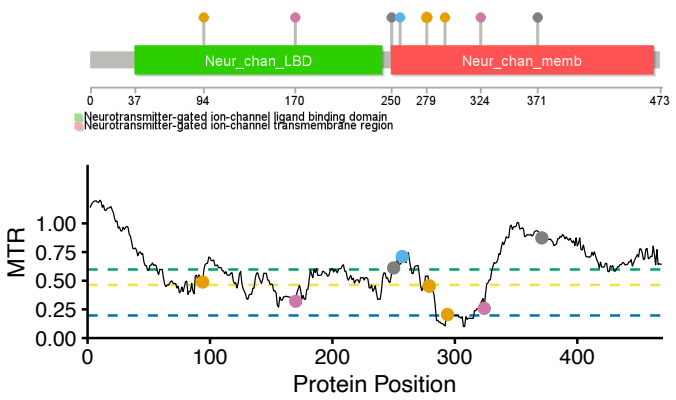
GABRA2



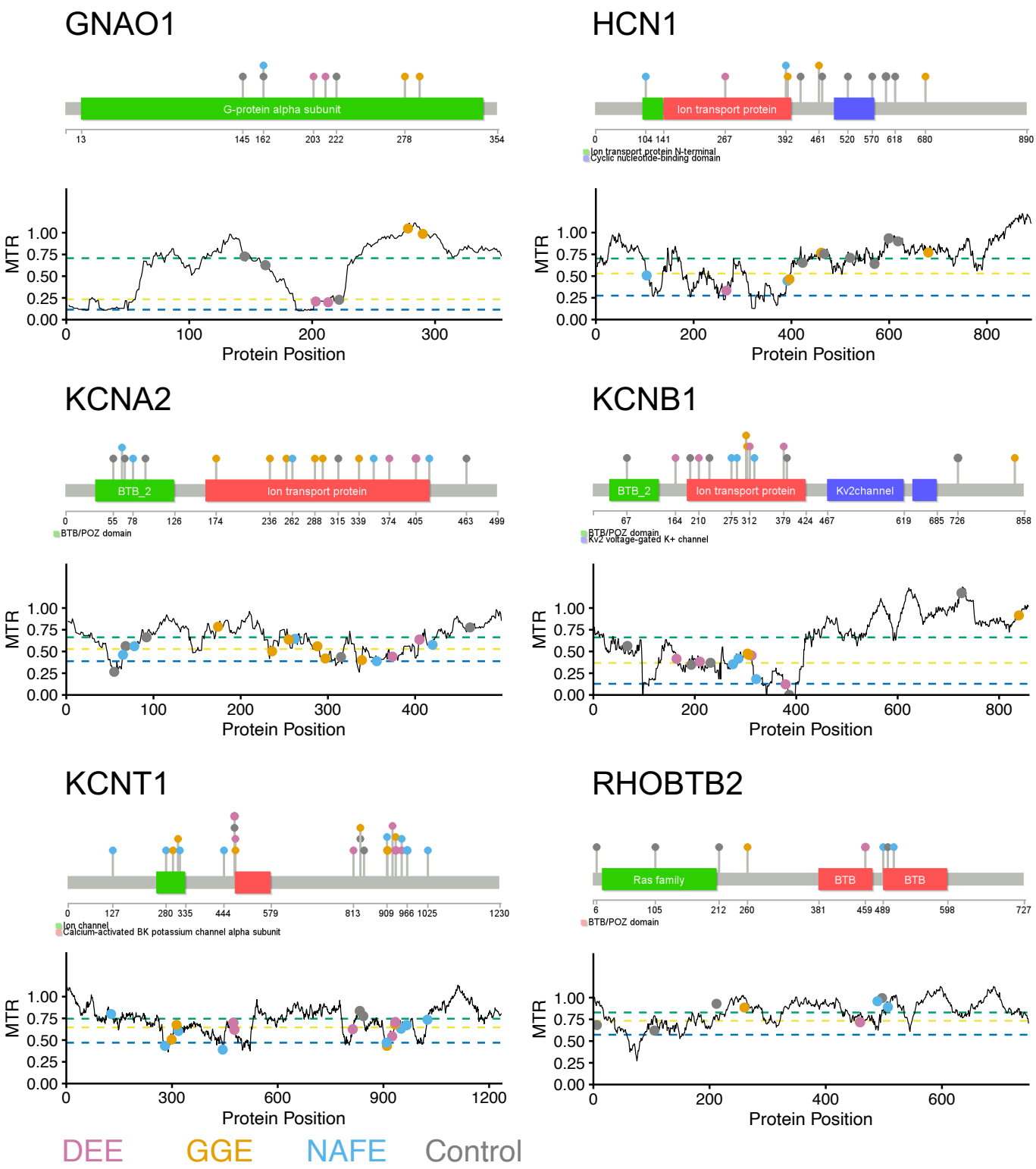
GABRB2



GABRB3



DEE GGE NAFE Control


GNAO1
G-protein alpha subunit
13
145 162
203 222
278
354
MTR
1.00
0.75
0.50
0.25
0.00
0
100
200
300
Protein Position
HCN1
Ion transport protein
0
104 141
267
392
461
520 570
618
680
890
Ion transport protein N-terminal
Cyclic nucleotide-binding domain
MTR
1.00
0.75
0.50
0.25
0.00
0
200
400
600
800
Protein Position
KCNA2
BTB_2
Ion transport protein
0
55 78
126
174
236 262 288 315 339
374
405
463
499
BTB/POZ domain
MTR
1.00
0.75
0.50
0.25
0.00
0
100
200
300
400
Protein Position
KCNB1
BTB_2
Ion transport protein
Kv2channel
67
164
210
275 312
379 424 467
619
685 726
858
BTB/POZ domain
Kv2 voltage-gated K+ channel
MTR
1.00
0.75
0.50
0.25
0.00
0
200
400
600
800
Protein Position
KCNT1
0
127
280 335
444
579
813
909 966 1025
1230
Ion channel
Calcium-activated BK potassium channel alpha subunit
MTR
1.00
0.75
0.50
0.25
0.00
0
300
600
900
1200
Protein Position
RHOBTB2
Ras family
BTB
BTB
6
105
212
260
381
459 489
598
727
BTB/POZ domain
MTR
1.00
0.75
0.50
0.25
0.00
0
200
400
600
Protein Position
DEE
GGE
NAFE
Control

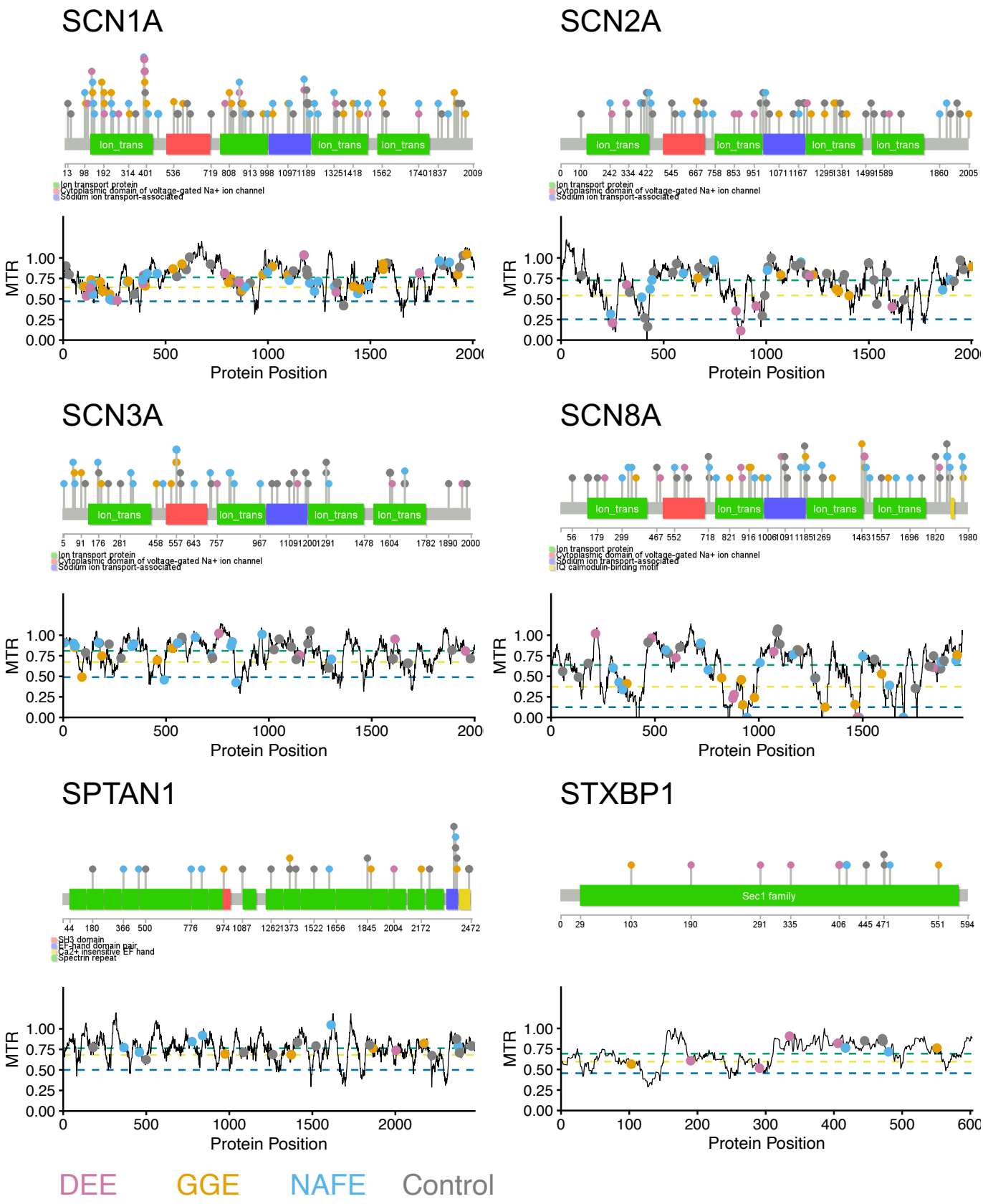

SCN1A
SCN2A
SCN3A
SCN8A
SPTAN1
STXBP1
Ion_trans
Sec1 family
Ion transport protein
Cytoplasmic domain of voltage-gated Na+ ion channel
Sodium ion transport-associated
IQ calmodulin-binding motif
SH3 domain
EF-hand domain pair
Ca2+ insensitive EF hand
Spectrin repeat
MTR
Protein Position
DEE
GGE
NAFE
Control

Figure S8. Lollipop Plots Juxtaposed with MTR Distributions

Juxtaposition of lollipop mutation diagrams and MTR distributions shows no clear relationship between variant location and MTR score. Lollipop plots with MTR distributions for each of the 24 genes drawn from the OMIM epileptic encephalopathy phenotype series which were included in the sub-genic intolerance comparisons (see Lollipop and MTR Plots in Methods). Both plots are annotated with variant locations for epilepsy subtypes and controls. Some variants in CACNA1A, FGF12, RHOBTB2 were not called from the canonical transcript and are therefore not aligned with the displayed canonical transcript MTR plot. In each MTR plot, the top dashed line (green) represents the median MTR value, the middle dashed line (yellow) represents the 25^{th} percentile, and the bottom dashed line (blue) represents the 5^{th} percentile.

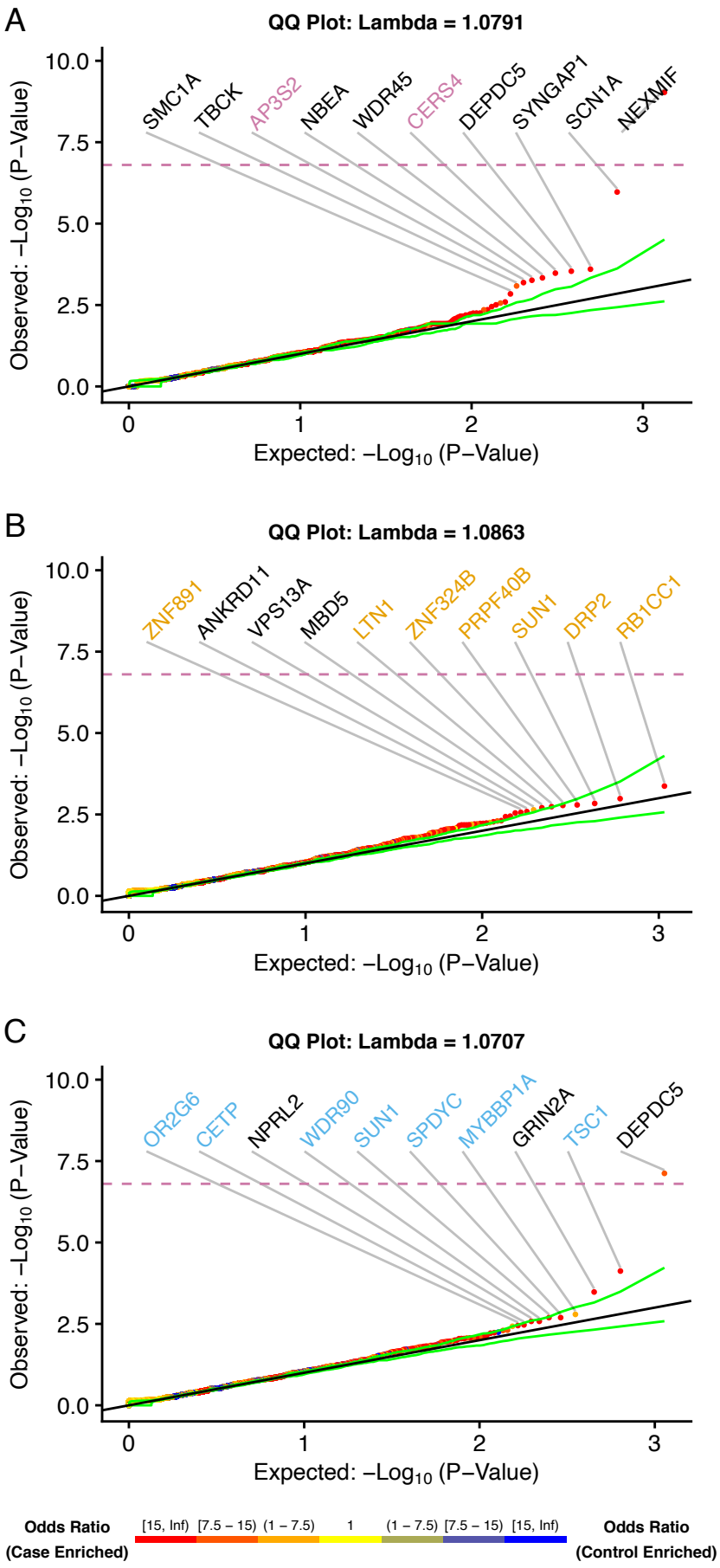

A
QQ Plot: Lambda = 1.0791
SMC1A
TBCK
AP3S2
NBEA
WDR45
CERS4
DEPDC5
SYNGAP1
SCN1A
NEXMIF
Observed: −Log10 (P−Value)
Expected: −Log10 (P−Value)
B
QQ Plot: Lambda = 1.0863
ZNF891
ANKRD11
VPS13A
MBD5
LTN1
ZNF324B
PRPF40B
SUN1
DRP2
RB1CC1
C
QQ Plot: Lambda = 1.0707
OR2G6
CETP
NPRL2
WDR90
SUN1
SPDYC
MYBBP1A
GRIN2A
TSC1
DEPDC5
Odds Ratio
(Case Enriched)
[15, Inf)
[7.5 – 15)
(1 – 7.5)
1
(1 – 7.5)
[7.5 – 15)
[15, Inf)
Odds Ratio
(Control Enriched)

Figure S9. Quantile-Quantile Plots for the Protein-Coding Genes with at least One Case or Control Carrier with a Protein Truncating Variant

Qualifying variants were high quality, ultra-rare variants with a protein truncating variant. *P*-values were generated from the exact two-sided Cochran-Mantel-Haenszel (CMH) test by gene by cluster to indicate a different carrier status of cases in comparison to controls. (A) Developmental and epileptic encephalopathy (DEE) cases, (B) genetic generalized epilepsy (GEE) cases, and (C) non-acquired focal epilepsy (NAFE) cases. In the DEE analysis, *NEXMIF* ($p = 9.3 \times 10^{-10}$) achieved study-wide significance $p < 1.6 \times 10^{-7}$ after Bonferroni correction (see Statistical Analyses in Methods). In the NAFE analysis, *DEPDC5* ($p = 7.6 \times 10^{-8}$) achieved study-wide significance. Point coloring determined by CMH odds ratio. Genes labeled in black are known epilepsy genes. Genes labeled in color are candidate epilepsy genes. The green lines represent the 95% confidence interval.

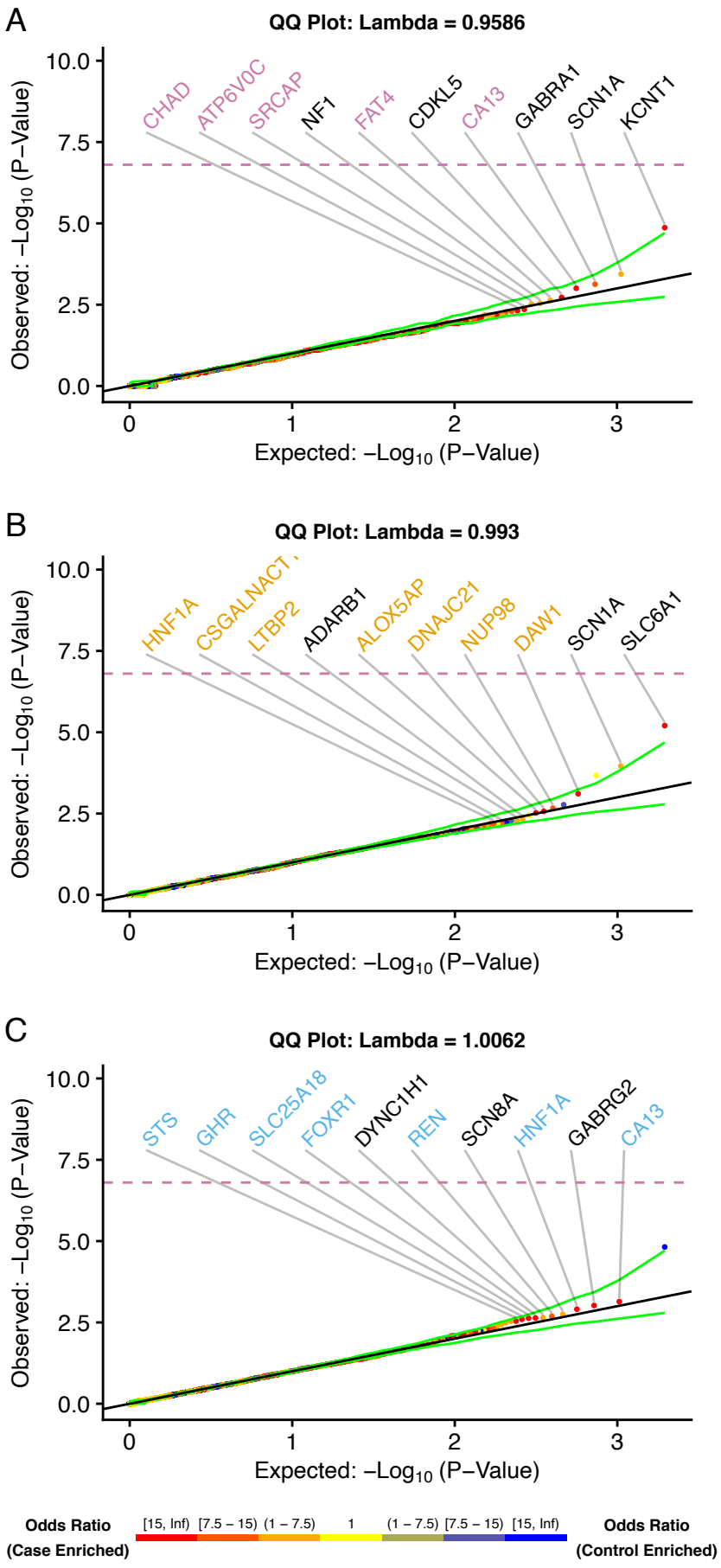

A
QQ Plot: Lambda = 0.9586
CHAD
ATP6V0C
SRCAP
NF1
FAT4
CDKL5
CA13
GABRA1
SCN1A
KCNT1
Observed: −Log10 (P−Value)
Expected: −Log10 (P−Value)
B
QQ Plot: Lambda = 0.993
HNF1A
CSGALNACT1
LTBP2
ADARB1
ALOX5AP
DNAJC21
NUP98
DAW1
SCN1A
SLC6A1
Observed: −Log10 (P−Value)
Expected: −Log10 (P−Value)
C
QQ Plot: Lambda = 1.0062
STS
GHR
SLC25A18
FOXR1
DYNC1H1
REN
SCN8A
HNF1A
GABRG2
CA13
Observed: −Log10 (P−Value)
Expected: −Log10 (P−Value)
Odds Ratio
(Case Enriched)
[15, Inf)
[7.5 – 15)
(1 – 7.5)
1
(1 – 7.5)
[7.5 – 15)
[15, Inf)
Odds Ratio
(Control Enriched)

Figure S10. Quantile-Quantile Plots for the Protein-Coding Genes with at least One Case or Control Carrier with a Damaging Missense Variant

Qualifying variants were high quality, ultra-rare variants with missense with REVEL ≥ 0.5 (when defined). *P*-values were generated from the exact two-sided Cochran-Mantel-Haenszel (CMH) test by gene by cluster to indicate a different carrier status of cases in comparison to controls. (A) Developmental and epileptic encephalopathy (DEE) cases, (B) genetic generalized epilepsy (GEE) cases, and (C) non-acquired focal epilepsy (NAFE) cases. No genes achieved study-wide significance $p < 1.6 \times 10^{-7}$ after Bonferroni correction (see Statistical Analyses in Methods). Point coloring determined by CMH odds ratio. Genes labeled in black are known epilepsy genes. Genes labeled in color are candidate epilepsy genes. The green lines represent the 95% confidence interval.

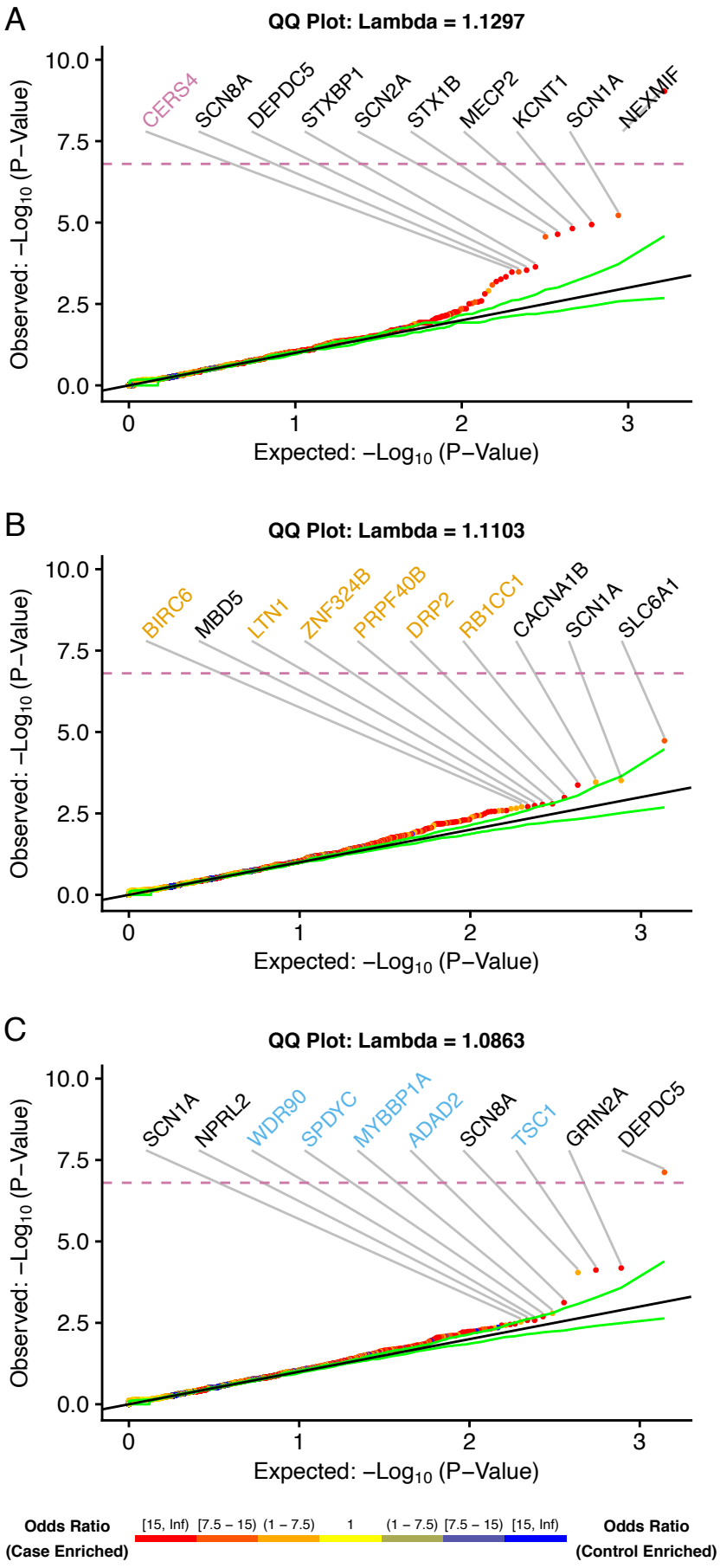

A
QQ Plot: Lambda = 1.1297
CERS4
SCN8A
DEPDC5
STXBP1
SCN2A
STX1B
MECP2
KCNT1
SCN1A
NEXMIF
Observed: −Log10 (P−Value)
Expected: −Log10 (P−Value)
B
QQ Plot: Lambda = 1.1103
BIRC6
MBD5
LTN1
ZNF324B
PRPF40B
DRP2
RB1CC1
CACNA1B
SCN1A
SLC6A1
C
QQ Plot: Lambda = 1.0863
SCN1A
NPRL2
WDR90
SPDYC
MYBBP1A
ADAD2
SCN8A
TSC1
GRIN2A
DEPDC5
Odds Ratio
(Case Enriched)
[15, Inf)
[7.5 – 15)
(1 – 7.5)
1
(1 – 7.5)
[7.5 – 15)
[15, Inf)
Odds Ratio
(Control Enriched)

Figure S11. Quantile-Quantile Plots for the Protein-Coding Genes with at least One Case or Control Carrier with either a Protein Truncating Variant or a Damaging and Intolerant Missense Variant in an Intolerant LIMBR Exon

Qualifying variants were high quality, ultra-rare variants with a predicted functional effect but restricting missense variants to REVEL $\geq$ 0.5 (when defined), MTR $\leq$ 0.78 (when defined) and LIMBR Exon percentile < 25%. *P*-values were generated from the exact two-sided Cochran-Mantel-Haenszel (CMH) test by gene by cluster to indicate a different carrier status of cases in comparison to controls. In the DEE analysis, *NEXMIF* ($p = 9.3 \times 10^{-10}$) achieved study-wide significance $p < 1.6 \times 10^{-7}$ after Bonferroni correction (see Statistical Analyses in Methods). In the NAFE analysis, *DEPDC5* ($p = 7.6 \times 10^{-8}$) achieved study-wide significance. (A) Developmental and epileptic encephalopathy (DEE) cases, (B) genetic generalized epilepsy (GEE) cases, and (C) non-acquired focal epilepsy (NAFE) cases. Top ten case enriched genes are labeled. Point coloring determined by CMH odds ratio. Genes labeled in black are known epilepsy genes. Genes labeled in color are candidate epilepsy genes. The green lines represent the 95% confidence interval.

Supplemental Tables

Site Name	*Site Code*	*DEE*	*GGE*	*NAFE*	*Other*	*Total*
Australia: Melbourne	AUSAUS	164	452	426	355	1397
Australia: Royal Melbourne	AUSRMB	0	89	154	61	304
Belgium: Antwerp	BELATW	96	39	22	1	158
Belgium: Brussels	BELULB	6	68	186	123	383
Canada: Andrade	CANUTN	41	43	8	3	95
Switzerland: Bern	CHEUBB	0	0	6	3	9
Cyprus	CYPCYP	8	57	52	33	150
Czech Republic: Prague	CZEMTH	16	0	0	0	16
Germany: Frankfurt/Marburg	DEUPUM	4	87	155	100	346
Germany: Giessen	DEUUGS	0	0	389	0	389
Germany: Bonn	DEUUKB	0	279	1138	569	1986
Germany: Kiel	DEUUKL	53	90	27	7	177
Germany: Leipzig	DEUULG	0	0	105	0	105
Germany: Tuebingen	DEUUTB	80	370	262	389	1101
Finland: Kuopio	FINKPH	20	57	628	22	727
Finland: Helsinki	FINUVH	27	52	23	1	103
France: Lyon	FRALYU	0	0	228	16	244
Wales: Swansea	GBRSWU	0	63	84	21	168
UK: UCL	GBRUCL	5	353	310	40	708
UK: Imperial/Liverpool	GBRUNL	0	191	344	1	536
Hong Kong	HKGHKK	0	21	24	0	45
Croatia	HRVUZG	19	6	2	3	30
Ireland: Dublin	IRLRCI	12	138	445	187	782
Italy: Milan	ITAICB	49	97	14	91	251
Italy: Genova	ITAIGI	146	263	15	404	828

Italy: Bologna	ITAUBG	112	73	146	58	389
Italy: Catanzaro	ITAUMC	5	72	260	36	373
Italy: Florence	ITAUMR	410	212	155	224	1001
Japan: Fukuoka	JPNFKA	171	0	3	194	368
Japan: RIKEN Institute	JPNRKI	30	67	0	2	99
Lebanon: Beirut	LEBABM	37	119	118	42	316
Lithuania	LTUUHK	58	118	95	24	295
New Zealand: Otago	NZLUTO	52	47	62	11	172
Turkey: Bogazici	TURBZU	125	15	14	53	207
Turkey: Istanbul	TURIBU	5	47	51	7	110
Taiwan	TWNCGM	3	12	273	178	466
USA: BCH	USABCH	90	31	15	39	175
USA: Baylor	USABLC	0	0	0	223	223
USA: Cleveland Clinic	USACCF	3	22	51	35	111
USA: Cincinatti	USACCH	0	358	0	0	358
USA: Philadelphia/CHOP	USACHP	0	991	475	0	1466
USA: Philadelphia/Rowan	USACRW	0	326	238	0	564
USA: EPGP	USAEGP	126	2	1	1	130
USA: FEBSTAT	USAFEB	0	0	0	31	31
USA: NYU HEP	USAHEP	0	0	205	0	205
USA: MONEAD	USAMON	0	30	72	16	118
USA: Nationwide	USANCH	0	310	0	5	315
USA: Penn/CHOP	USAUPN	34	104	207	116	461

Table S1. Summary of Epi25 Sites Contributing Cases in Analysis.

Summary of sites of Epi25 enrollment by epilepsy. DEE = developmental and epileptic encephalopathy, GGE genetic generalized epilepsy, NAFE = non-acquired focal epilepsy. See http://epi-25.org for more details.

Control Phenotype	*N*
Healthy Family Member	5209
Amyotrophic Lateral Sclerosis	3571
Control	2892
Dementia	735
Obsessive Compulsive Disorder	655
Pulmonary Disease	391
Ophthalmic Disease	264
Cardiovascular Disease	158
Infectious Disease	102
Liver Disease	79
Primary Immune Deficiency	26
Control Mild Neuropsychiatric Disease	10
Neurodegenerative	7
Hematological Disease	1

Table S2. Summary of Phenotypes of Included Controls.

Summary of phenotypes and sample sizes of controls.

Cluster	*Admixed*	*African*	*European*	*East Asian*	*Latino*	*Middle Eastern*	*South Asian*
0	4	0	3208	0	0	0	0
1	0	0	2389	0	0	0	0
2	487	0	2122	0	0	6	0
3	5	0	2088	0	0	0	0
4	6	682	0	0	13	0	0
5	55	26	28	0	115	0	0
7	660	14	66	0	35	85	0
8	2	0	8	1	152	0	0
9	0	0	0	670	0	0	0
11	35	1	37	1	0	0	170

Table S3. Table Summarizing Geographic Ancestry of Cases by Cluster.

Distribution of geographic ancestries for cases among clusters in the combined epilepsy analysis

Model Name	*Inheritance*	*Included Effects*	*REVEL*	*MTR*	*LIMBR Exon Percentile*	*Analyzed Groups*
Ultra-Rare Synonymous	Dominant	Synonymous	N/A	N/A	N/A	DEE, GGE, NAFE, and Combined
Ultra-Rare Dominant Deleterious	Dominant	Missense and PTVs	≥ 0.5	N/A	N/A	DEE, GGE, NAFE, and Combined
Ultra-Rare Dominant Deleterious and Intolerant	Dominant	Missense and PTVs	≥ 0.5	≤ 0.78	N/A	DEE, GGE, and NAFE
Ultra-Rare Dominant Deleterious and Intolerant Excluding Known Epilepsy Genes	Dominant	Missense and PTVs	≥ 0.5	≤ 0.78	N/A	DEE minus Cases with Deleterious Variant in Known Epilepsy Gene
Ultra-Rare Dominant Deleterious and Intolerant Plus LIMBR	Dominant	Missense and PTVs	≥ 0.5	≤ 0.78	< 25	DEE, GGE, and NAFE
Ultra-Rare Damaging Missense	Dominant	Missense	≥ 0.5	N/A	N/A	DEE, GGE, and NAFE
Ultra-Rare Protein Truncating Variants	Dominant	PTVs	N/A	N/A	N/A	DEE, GGE, and NAFE

Table S4. Collapsing Model Qualifying Variant Definition

Summary of collapsing models. PTV = protein truncating variant. PTV effects include stop gained, frameshift, splice acceptor, and splice donor variants.

In a separate Excel file (large tables):
Table S5. OMIM Gene Sets

Gene-sets utilized in analyses. “Gene associated with Dominant Epilepsy Phenotype” refers to GS-3, “Gene associated with Dominant Phenotype of DEE” refers to GS-1, “DEE Gene with Variants in all Epilepsies” refers to GS-2, and “Genes in ClinVar Intolerance Analysis” refer to GS-4. See Gene-Set Enrichment Testing in Methods for gene set abbreviations.

Table S6. Top 200 Genes with Burden of Ultra-Rare Synonymous Variants in Developmental and Epileptic Encephalopathy

Table S7. Top 200 Genes with Burden of Ultra-Rare Synonymous Variants in Genetic Generalized Epilepsy

Table S8. Top 200 Genes with Burden of Ultra-Rare Synonymous Variants in Non-Acquired Focal Epilepsy

Table S9. Top 200 Genes with Burden of Ultra-Rare Deleterious Variants in Developmental and Epileptic Encephalopathy

Table S10. Top 200 Genes with Burden of Ultra-Rare Deleterious and Intolerant Variants in Developmental and Epileptic Encephalopathy

Table S11. Top 200 Genes with Burden of Ultra-Rare Deleterious and Intolerant Variants in Developmental and Epileptic Encephalopathy without Cases with Variants in Known DEE genes

Table S12. Top 200 Genes with Burden of Ultra-Rare Deleterious Variants in Genetic Generalized Epilepsy

Table S13. Top 200 Genes with Burden of Ultra-Rare Deleterious and intolerant Variants in Genetic Generalized Epilepsy

Table S14. Top 200 Genes with Burden of Ultra-Rare Deleterious Variants in Non-Acquired Focal Epilepsy

Table S15. Top 200 Genes with Burden of Ultra-Rare Deleterious and Intolerant Variants in Non-Acquired Focal Epilepsy

Table S16. Top 200 Genes with Burden of Ultra-Rare Synonymous Variants in Combined Epilepsies

Table S17. Top 200 Genes with Burden of Ultra-Rare Deleterious Variants in Combined Epilepsies

Table S18. Top 200 Genes with Burden of Ultra-Rare Protein Truncating Variants in Developmental and Epileptic Encephalopathy

Table S19. Top 200 Genes with Burden of Ultra-Rare Protein Truncating Variants in Genetic Generalized Epilepsy

Table S20. Top 200 Genes with Burden of Ultra-Rare Protein Truncating Variants in Non-Acquired Focal Epilepsy

Table S21. Top 200 Genes with Burden of Ultra-Rare Damaging Missense Variants in Developmental and Epileptic Encephalopathy

Table S22. Top 200 Genes with Burden of Ultra-Rare Damaging Missense Variants in Genetic Generalized Epilepsy

Table S23. Top 200 Genes with Burden of Ultra-Rare Damaging Missense Variants in Non-Acquired Focal Epilepsy

Table S24. Top 200 Genes with Burden of Ultra-Rare Deleterious and Intolerant Variants in Intolerant LIMBR Exons in Developmental and Epileptic Encephalopathy

Table S25. Top 200 Genes with Burden of Ultra-Rare Deleterious and Intolerant Variants in Intolerant LIMBR Exons in Genetic Generalized Epilepsy

Table S26. Top 200 Genes with Burden of Ultra-Rare Deleterious and Intolerant Variants in Intolerant LIMBR Exons in Non-Acquired Focal Epilepsy

For Tables S6 – S26, summary of top 200 genes in collapsing analysis indicated in the table title. The last seven columns indicate gene group membership. D = 43 established dominant (e.g. autosomal dominant or x-linked dominant) DEE genes drawn from OMIM Phenotypic Series, P = 101 established dominant genes associated with OMIM phenotypes containing epilepsy and epilepsy, L = 1,920 genes most intolerant to loss-of-function variation in the general population, D25 = rank, if present, in top 200 genes in prior Epi25 DEE AC $\leq$ 3 collapsing analysis, G25 = rank, if present, in top 200 genes in prior Epi25 GGE AC $\leq$ 3 collapsing analysis, N25 = rank, if present, in top 200 genes in prior Epi25 NAFE AC $\leq$ 3 collapsing analysis, G4K = rank, if present, in top 200 genes in Epi4K GGE collapsing analysis, N4K = rank, if present, in top 200 genes in Epi4K NAFE collapsing analysis.[1; 2]

Table S27. DEE Gene Burden

Gene set burden of ultra-rare variants for each epilepsy type versus controls within a limited set of developmental and epileptic encephalopathy genes. Table shows data for Figure 2 (see figure legend for details).

Table S28. Deleterious Missense Burden in Epi25

Comparison of damaging (REVEL $\geq$ 0.5) missense variant burden among epilepsies in the 24 genes drawn from the 43 OMIM epileptic encephalopathy phenotype series with dominant transmission (see Gene-Set Enrichment Testing in Methods). Pooled odds ratio, confidence intervals and FDR corrected p-value were generated from the exact two-sided Cochran-Mantel-Haenszel test. DEE = developmental and epileptic encephalopathy, GGE = genetic generalized epilepsy, NAFE = non-acquired focal epilepsy.

Gene	*DEE Weighted Mean MTR*	*GGE Weighted Mean MTR*	*NAFE Weighted Mean MTR*	*DEE More Intolerant than NAFE*	*DEE More Intolerant than GGE*
'KCNT1'	0.66	0.63	0.63	FALSE	FALSE
'SPTAN1'	0.76	0.76	0.75	FALSE	FALSE
'DNM1'	0.72	0.67	0.76	TRUE	FALSE
'RHOBTB2'	0.74	0.77	0.82	TRUE	TRUE
'GABRA1'	0.60	0.38	0.41	FALSE	FALSE
'GABRB2'	0.28	0.60	0.23	FALSE	TRUE
'GABRA2'	0.55	0.75	0.49	FALSE	TRUE
'EEF1A2'	0.49	0.46	0.39	FALSE	FALSE
'SCN1A'	0.72	0.75	0.75	TRUE	TRUE
'SCN2A'	0.58	0.74	0.75	TRUE	TRUE
'SCN3A'	0.83	0.80	0.81	FALSE	FALSE
'CACNA1A'	0.67	0.67	0.70	TRUE	FALSE
'GABRB3'	0.35	0.48	0.74	TRUE	TRUE
'SCN8A'	0.57	0.60	0.63	TRUE	TRUE
'CACNA1E'	0.64	0.27	0.71	TRUE	FALSE
'KCNA2'	0.57	0.55	0.53	FALSE	FALSE
'GABBR2'	0.70	0.71	1.30	TRUE	TRUE
'ATP6V1A'	0.75	0.80	0.61	FALSE	TRUE
'KCNB1'	0.39	0.61	0.58	TRUE	TRUE
'STXBP1'	0.74	0.77	0.82	TRUE	TRUE
'FGF12'	0.83	0.94	0.94	TRUE	TRUE
'GNAO1'	0.27	0.75	0.49	TRUE	TRUE
'HCN1'	0.74	0.77	0.97	TRUE	TRUE
'CUX2'	0.59	0.64	0.60	TRUE	TRUE

Table S29. Weighted Mean MTR Score Comparison by Gene

Table showing the weighted mean MTR score for each gene in the subset of dominant DEE genes. See Sub-Genic Intolerance Comparison in Methods for calculation of weighted mean MTR. Gene-set is 24 genes drawn from the 43 OMIM epileptic encephalopathy phenotype series with dominant transmission by limiting to genes harboring damaging (REVEL $\geq$ 0.5) missense variants in all 3 epilepsies (see Gene-Set Enrichment Testing in Methods, Table S5). DEE = developmental and epileptic encephalopathy, GGE = genetic generalized epilepsy, NAFE = non-acquired focal epilepsy.

In a separate Excel file (large tables):

Table S30. ClinVar by MAF Bin

Gene set burden of ClinVar P/LP variants for each epilepsy type versus controls. Table shows data for Figure 4A (see figure legend for details).

Table S31. Ultra-Rare Variant ClinVar Burden

Gene set burden of ClinVar P/LP variants for each epilepsy type versus controls. Table shows data for Figure 4B (see figure legend for details).

Table S32. Public Variant ClinVar Burden

Public (present in non-neuro gnomAD populations) variants in Epi25 cases are limited. Significant enrichment is limited to intolerant missense variants associated non-acquired focal epilepsy (NAFE). PTV denotes protein-truncating variants, “Damaging” denotes REVEL $\geq$ 0.5 (when defined), “Intolerant” denotes MTR $\leq$ 0.78 (when defined), “Star” denotes ClinVar review status (see Qualifying Variant in Methods). Pooled odds ratio, confidence intervals and FDR corrected p-value were generated from the exact two-sided Cochran-Mantel-Haenszel (CMH).

Gene	*ClinVar DEE Mean MTR*	*ClinVar Non-DEE Mean MTR*	*DEE More Intolerant*
'CHD2'	0.70	0.80	TRUE
'GABRA1'	0.46	0.83	TRUE
'GABRB3'	0.51	0.46	FALSE
'GABRG2'	0.51	0.72	TRUE
'GRIN2B'	0.12	0.51	TRUE
'HCN1'	0.53	0.62	TRUE
'KCNQ2'	0.33	0.59	TRUE
'KCNQ3'	0.60	0.69	TRUE
'KCNT1'	0.61	0.73	TRUE
'SCN1A'	0.66	0.60	FALSE
'SCN2A'	0.48	0.72	TRUE
'SPTAN1'	0.40	0.84	TRUE
'SZT2'	0.83	0.90	TRUE
'TBC1D24'	0.88	0.87	FALSE

Table S33. Mean MTR Score Comparison by Gene in Published ClinVar Variants

Table showing the mean MTR score by gene for variants drawn from ClinVar. Gene-set is 14 genes drawn from ClinVar harboring ultra-rare missense variants associated with both DEE and with epilepsy but not DEE in ClinVar (see Gene-Set Enrichment Testing in Methods, Table S5). DEE = developmental and epileptic encephalopathy.

In a separate Excel file (large tables):

Table S34. Non-OMIM Genes Burden by LOF Int

Gene set burden of ultra-rare protein truncating variants in non-OMIM genes stratified by loss-of-function intolerance for each epilepsy type versus controls. Table shows data for Figure 6 (see figure legend for details).

Table S35. Non-OMIM Genes Burden by Mis Int

The burden of missense variants in genes not associated with a disease in OMIM in epilepsy cases in comparison to controls was assessed. Non-OMIM genes were divided into 10 gene-sets by their intersection with missense intolerance deciles defined by missense Z score (see Gene-Set Enrichment Testing in Methods). Number of genes in each gene-set is specified in the parenthesis. Pooled odds ratio, 95% confidence intervals and FDR corrected p-value were generated from the exact two-sided Cochran-Mantel-Haenszel (CMH) test.

Gene	*Number of DEE Cases in Epi25*	*P-Value*
PPARGC1A	2	0.10

Table S36. Non-OMIM Genes Intolerant to Loss-of-Function Variants with Multiple Protein Truncating Variants in Developmental and Epileptic Encephalopathy Cases

Genes in the most intolerant decile of genes with protein truncating variants (PTV) in more than one developmental and epileptic encephalopathy (DEE) case but no PTVs in internal and external controls. *P*-values drawn from Ultra-Rare Protein Truncating Variants collapsing analysis (Figure S10A, Tables S4 and S18).

Gene	*Number of GGE Cases in Epi25*	*P-Value*
NLGN2	3	8.6e-03
HDLBP	4	8.9e-03
RC3H2	4	0.01
XPO5	3	0.02
FAM120C	2	0.02
HNRNPH1	3	0.03
SEC24C	3	0.05
CLCN3	2	0.05
HELZ	2	0.05
RBM15B	2	0.05
SCAF4	2	0.05
PIK3AP1	2	0.09
SCAF8	2	0.09
CHMP6	2	0.10
PAXBP1	2	0.10
ZNF638	2	0.10
ZBTB21	2	0.10
CAMSAP2	2	0.10
ZFR	2	0.10
JADE2	2	0.11
STK39	2	0.12
SETDB1	2	0.12
TNRC6C	2	0.12

Table S37. Non-OMIM Genes Intolerant to Loss-of-Function Variants with Multiple Protein Truncating Variants in Genetic Generalized Epilepsy Cases

Genes in the most intolerant decile of genes with protein truncating variants (PTV) in more than one genetic generalized epilepsy (GGE) case but no PTVs in internal and external controls. *P*-values drawn from Ultra-Rare Protein Truncating Variants collapsing analysis (Figure S10B, Tables S4 and S19).

Gene	*Number of NAFE Cases in Epi25*	*P-Value*
WDR18	4	0.01
SOCS7	5	0.01
TRIM9	3	0.05
SORT1	2	0.05
ENAH	2	0.05
CLCN3	2	0.06
RC3H2	3	0.08
CUL2	2	0.13
IGF2BP3	2	0.13
ITPKB	2	0.13
MARK2	2	0.13
PAXIP1	2	0.14
PIK3AP1	2	0.14
CLK2	2	0.14
MAP3K12	2	0.14
SCAF1	2	0.14
PSMD1	2	0.14
VPS54	2	0.15
ZNF638	2	0.15
AJAP1	2	0.16
ZNF541	2	0.21
DNAJC14	2	0.21

Table S38. Non-OMIM Genes Intolerant to Loss-of-Function Variants with Multiple Protein Truncating Variants in Non-Acquired Focal Epilepsy Cases

Genes in the most intolerant decile of genes with protein truncating variants (PTV) in more than one non-acquired focal epilepsy (NAFE) case but no PTVs in internal and external controls. *P*-values drawn from Ultra-Rare Protein Truncating Variants collapsing analysis (Figure S10C, Tables S4 and S20).

Supplemental Subjects and Methods

A full list of authors and affiliations

Columbia University Institute for Genomic Medicine analysis group
Joshua E. Motelow[1], Gundula Povysil[1], Ryan S. Dhindsa[1], Kate E. Stanley[1], Andrew S. Allen[2], David B. Goldstein[1]

Epi25 sequencing, analysis, project management, and browser development at the Broad Institute
Yen-Chen Anne Feng[3-6], Daniel P. Howrigan[3, 5, 6], Liam E. Abbott[3, 5, 6], Katherine Tashman[3, 5, 6], Felecia Cerrato[6], Caroline Cusick[6], Tarjinder Singh[3, 5, 6], Henrike Heyne[3, 5, 6], Andrea E. Byrnes[3, 5, 6], Claire Churchhouse[3, 5, 6], Nick Watts[3, 5, 6], Matthew Solomonson[3, 5, 6], Dennis Lal[6-8], Namrata Gupta[5], Benjamin M. Neale[3, 5, 6]

Epi25 executive committee
Samuel F. Berkovic[9], Holger Lerche[10], David B. Goldstein[1], Daniel H. Lowenstein[11]

Epi25 strategy, phenotyping, analysis, informatics, and project management committees
Samuel F. Berkovic[9], Holger Lerche[10], David B. Goldstein[1], Daniel H. Lowenstein[11], Gianpiero L. Cavalleri[12, 13], Patrick Cossette[14], Chris Cotsapas[15], Peter De Jonghe[16-18], Tracy Dixon-Salazar[19], Renzo Guerrini[20], Hakon Hakonarson[21], Erin L. Heinzen[1], Ryan S. Dhindsa[1], Kate E. Stanley[1], Ingo Helbig[21-23], Patrick Kwan[24, 25], Anthony G. Marson[26], Slavé Petrovski[25, 27], Sitharthan Kamalakaran[1], Sanjay M. Sisodiya[28], Randy Stewart[29], Sarah Weckhuysen[16-18], Chantal Depondt[30], Dennis J. Dlugos[21, 22], Ingrid E. Scheffer[9], Pasquale Striano[31, 32], Catharine Freyer[11], Roland Krause[33], Patrick May[33], Kevin McKenna[11], Brigid M. Regan[9], Caitlin A. Bennett[9], Costin Leu[6, 7, 28]

Authors from individual Epi25 cohorts:

Australia: Melbourne (AUSAUS)
Samuel F. Berkovic[9], Ingrid E. Scheffer[9], Brigid M. Regan[9], Caitlin A. Bennett[9], Stephanie L. Leech[9]

Australia: Royal Melbourne (AUSRMB)
Terence J. O'Brien[24, 25], Slavé Petrovski[25, 27], Marian Todaro[24, 25], Patrick Kwan[24, 25]

Belgium: Antwerp (BELATW)
Sarah Weckhuysen[16-18], Peter De Jonghe[16-18], Hannah Stamberger[16-18]

Belgium: Brussels (BELULB)
Chantal Depondt[30]

Canada: Andrade (CANUTN)
Danielle M. Andrade[34, 35], Quratulain Zulfiqar Ali[34], Tara R. Sadoway[35]

Switzerland: Bern (CHEUBB)
Heinz Krestel[15], André Schaller[36]

Cyprus (CYPCYP)
Savvas S. Papacostas[37], Ioanna Kousiappa[37], George A. Tanteles[38], Christou Yiolanda[37]

Czech Republic: Prague (CZEMTH)
Katalin Štěrbová[39], Markéta Vlčková[39], Lucie Sedláčková[39], Petra Laššuthová[39]

Germany: Frankfurt/Marburg (DEUPUM)
Karl Martin Klein[40-43], Felix Rosenow[40-42], Philipp S. Reif[40-42], Susanne Knake[41, 42]

Germany: Giessen (DEUUGS)
Bernd A. Neubauer[44], Friedrich Zimprich[45], Martha Feucht[44], Eva Reinthaler[45]

Germany: Bonn (DEUUKB)
Wolfram S. Kunz[46, 47], Gábor Zsurka[46, 47], Rainer Surges[47], Tobias H. Baumgartner[47], Randi von Wrede[47]

Germany: Kiel (DEUUKL)
Ingo Helbig[21, 48, 49], Karl Martin Klein[40-43], Manuela Pendziwiat[48, 49], Hiltrud Muhle[49], Annika Rademacher[49], Andreas van Baalen[49], Sarah von Spiczak[49, 50], Ulrich Stephani[49], Zaid Afawi[51], Amos D. Korczyn[52], Moien Kanaan[53], Christina Canavati[53], Gerhard Kurlemann[54], Karen Müller-Schlüter[55], Gerhard Kluger[56, 57], Martin Häusler[58], Ilan Blatt[52, 59]

Germany: Leipzig (DEUULG)
Johannes R. Lemke[60], Ilona Krey[60]

Germany: Tuebingen (DEUUTB)
Holger Lerche[10], Yvonne G. Weber[10, 61], Stefan Wolking[10], Felicitas Becker[10, 62], Stephan Lauxmann[10], Christian Bosselmann[10], Josua Kegele[10], Christian Hengsbach[10],Sarah Rau[10], Bernhard J. Steinhoff[63], Andreas Schulze-Bonhage[64], Ingo Borggräfe[65, 66], Christoph J. Schankin[67, 68], Susanne Schubert-Bast[64], Herbert Schreiber[69], Thomas Mayer[70], Rudolf Korinthenberg[71], Knut Brockmann[72], Markus Wolff[73], Gerhard Kurlemann[74],
Dieter Dennig[75], Rene Madeleyn[76], Christian Bosselmann[10], Josua Kegele[10]

Finland: Kuopio (FINKPH)
Reetta Kälviäinen[77, 78], Anni Saarela[77], Oskari Timonen[78]

Finland: Helsinki (FINUVH)
Tarja Linnankivi[79], Anna-Elina Lehesjoki[80]

France: Lyon (FRALYU)
Sylvain Rheims[81, 82], Gaetan Lesca[83], Philippe Ryvlin[84], Louis Maillard[85], Luc Valton[86], Philippe Derambure[87], Fabrice Bartolomei[88], Edouard Hirsch[89], Véronique Michel[90], Francine Chassoux[91]

Wales: Swansea (GBRSWU)
Mark I. Rees[92], Seo-Kyung Chung[92, 93], William O. Pickrell[94], Robert H.W. Powell[95], Mark D. Baker[94], Beata Fonferko-Shadrach[94], Charlotte Lawthom[96], Joe Anderson[96]

UK: UCL (GBRUCL)
Sanjay M. Sisodiya[28], Natascha Schneider[28], Simona Balestrini[28], Sara Zagaglia[28], Vera Braatz[28]

UK: Imperial/Liverpool (GBRUNL)
Anthony G. Marson[26], Michael R. Johnson[97], Pauls Auce[98], Graeme J. Sills[99]

Hong Kong (HKGHKK)
Patrick Kwan[24, 25, 100], Larry W. Baum[101], Pak C. Sham[101], Stacey S. Cherny[102], Colin H.T. Lui[103]

Ireland: Dublin (IRLRCI)
Gianpiero L. Cavalleri[12, 13], Norman Delanty[12, 13, 104], Colin P. Doherty[13, 105], Arif Shukralla[13, 104], Hany El-Naggar[12, 13, 104], Peter Widdess-Walsh[12, 13, 104]

Croatia (HRVUZG)
Nina Barišić[106]

Italy: Milan (ITAICB)
Laura Canafoglia[107], Silvana Franceschetti[108], Barbara Castellotti[109], Tiziana Granata[110], Francesca Ragona[110]

Italy: Genova (ITAIGI)
Pasquale Striano[31, 32], Federico Zara[31, 32], Michele Iacomino[31], Antonella Riva[31, 32], Francesca Madia[31, 32], Maria Stella Vari[31], Vincenzo Salpietro[31], Marcello Scala[31], Maria Margherita Mancardi[31], Nobili Lino[31], Elisa Amadori[31], Thea Giacomini[31]

Italy: Bologna (ITAUBG)
Francesca Bisulli[111, 112], Tommaso Pippucci[113], Laura Licchetta[111, 112], Raffaella Minardi[111],Paolo Tinuper[111, 112], Lorenzo Muccioli[111, 112], Barbara Mostacci[111]

Italy: Catanzaro (ITAUMC)
Antonio Gambardella[114], Angelo Labate[114], Grazia Annesi[115], Lorella Manna[115], Monica Gagliardi[115]

Italy: Florence (ITAUMR)
Renzo Guerrini[20], Elena Parrini[20], Davide Mei[20], Annalisa Vetro[20], Claudia Bianchini[20], Martino Montomoli[20], Viola Doccini[20], Carmen Barba[20]

Japan: Fukuoka (JPNFKA)
Shinichi Hirose[116], Atsushi Ishii[116]

Japan: RIKEN Institute (JPNRKI)
Toshimitsu Suzuki[117, 118], Yushi Inoue[119], Kazuhiro Yamakawa[117, 118]

Lebanon: Beirut (LEBABM)
Ahmad Beydoun[120], Wassim Nasreddine[120], Nathalie Khoueiry-Zgheib[121]

Lithuania (LTUUHK)
Birute Tumiene[122, 123], Algirdas Utkus[122, 123]

New Zealand: Otago (NZLUTO)
Lynette G. Sadleir[124], Chontelle King[124]

Turkey: Bogazici (TURBZU)
S. Hande Caglayan[125], Mutluay Arslan[126], Zuhal Yapıcı[127], Pınar Topaloglu[127], Bulent Kara[128], Uluc Yis[128], Dilsad Turkdogan[128], Aslı Gundogdu-Eken[125]

Turkey: Istanbul (TURIBU)
Nerses Bebek[129]

Taiwan (TWNCGM)
Meng-Han Tsai[130], Chen-Jui Ho[130], Chih-Hsiang Lin[130], Kuang-Lin Lin[131], I-Jun Chou[131]

USA: BCH (USABCH)
Annapurna Poduri[132, 133], Beth R. Shiedley[132, 133], Catherine Shain[132, 133]

USA: Baylor College of Medicine (USABLC)
Jeffrey L. Noebels[134], Alicia Goldman[134]

USA: Cleveland Clinic (USACCF)
Robyn M. Busch[8, 135, 136], Lara Jehi[135, 136], Imad M. Najm[135, 136], Dennis Lal[8, 135], Lisa Ferguson[135], Jean Khoury[135, 136]

USA: Cincinnati Children's Hospital Medical Center (USACCH)
Tracy A. Glauser[137], Peggy O. Clark[137]

USA: Philadelphia/CHOP (USACHP) and Philadelphia/Rowan (USACRW)
Russell J. Buono[21, 138, 139], Thomas N. Ferraro[138], Michael R. Sperling[139], Dennis J. Dlugos[21, 140], Warren Lo[140], Michael Privitera[141], Jacqueline A. French[142], Patrick Cossette[14], Steven Schachter[143], Hakon Hakonarson[21]

USA: EPGP (USAEGP)
Daniel H. Lowenstein[11], Ruben I. Kuzniecky[144], Dennis J. Dlugos[21, 22], Orrin Devinsky[142]

USA: NYU HEP (USAHEP)
Daniel H. Lowenstein[11], Ruben I. Kuzniecky[144], Jacqueline A. French[142], Manu Hegde[11]

USA: Nationwide Children's Hospital (USANCH)
David A. Greenberg[145]

USA: Penn/CHOP (USAUPN)
Ingo Helbig[21-23], Colin A. Ellis[22], Ethan Goldberg[21, 146], Katherine L. Helbig[21, 146], Mahgenn Cosico[21, 146], Priya Vaidiswaran[21, 146], Eryn Fitch[21, 146]

Affiliations:

[1] Institute for Genomic Medicine, Columbia University, New York, NY 10032, USA
[2] Division of Integrative Genomics, Department of Biostatistics and Bioinformatics, Duke University, Durham, NC 27710, USA
[3] Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA
[4] Psychiatric & Neurodevelopmental Genetics Unit, Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114, USA
[5] Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, 7 Cambridge Center, Cambridge, MA 02142, USA
[6] Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
[7] Genomic Medicine Institute, Cleveland Clinic, Cleveland, OH 44195, USA
[8] Center for Genomic Medicine, Lerner Research Institute, Cleveland Clinic, Cleveland, OH 44195, USA
[9] Epilepsy Research Centre, Department of Medicine, University of Melbourne, Victoria, Australia
[10] Department of Neurology and Epileptology, Hertie Institute for Clinical Brain Research, University of Tübingen, 72076 Tübingen, Germany
[11] Department of Neurology, University of California, San Francisco, CA 94110, USA
[12] School of Pharmacy and Biomolecular Sciences, The Royal College of Surgeons in Ireland, Dublin, Ireland
[13] The FutureNeuro Research Centre, Ireland
[14] University of Montreal, Montreal, QC H3T 1J4, Canada
[15] School of Medicine, Yale University, New Haven, CT 06510, USA
[16] Neurogenetics Group, Center for Molecular Neurology, VIB, Antwerp, Belgium
[17] Laboratory of Neurogenetics, Institute Born-Bunge, University of Antwerp, Belgium
[18] Division of Neurology, Antwerp University Hospital, Antwerp, Belgium
[19] LGS Foundation, NY 11716, USA
[20] Pediatric Neurology, Neurogenetics and Neurobiology Unit and Laboratories, Children's Hospital A. Meyer, University of Florence, Italy
[21] Division of Neurology, Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA
[22] Perelman School of Medicine, University of Pennsylvania, PA 19104, USA
[23] Department of Biomedical and Health Informatics (DBHi), Children's Hospital of Philadelphia, Philadelphia, PA, USA
[24] Department of Neuroscience, Central Clinical School, Monash University, Alfred Hospital, Melbourne, Australia
[25] Departments of Medicine and Neurology, University of Melbourne, Royal Melbourne Hospital, Parkville, Australia
[26] Department of Pharmacology and Therapeutics, University of Liverpool, Liverpool, UK
[27] Centre for Genomics Research, Precision Medicine and Genomics, IMED Biotech Unit, AstraZeneca, Cambridge, UK
[28] Department of Clinical and Experimental Epilepsy, UCL Queen Square Institute of Neurology, London, UK and Chalfont Centre for Epilepsy, Chalfont St Peter, UK
[29] National Institute of Neurological Disorders and Stroke, MD 20852, USA
[30] Department of Neurology, Université Libre de Bruxelles, Brussels, Belgium
[31] IRCCS "G. Gaslini" Institute, Genova, Italy

[32] Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health, University of Genoa, Italy
[33] Luxembourg Centre for Systems Biomedicine, University Luxembourg, Esch-sur-Alzette, Luxembourg
[34] Department of Neurology, Toronto Western Hospital, Toronto, ON M5T 2S8, Canada
[35] University Health Network, University of Toronto, Toronto, ON, Canada
[36] Institute of Human Genetics, Bern University Hospital, Bern, Switzerland
[37] Department of Neurobiology and Clinical Neurophysiology (Epilepsy Center), The Cyprus Institute of Neurology and Genetics, 2370 Nicosia, Cyprus
[38] Department of Clinical Genetics, The Cyprus Institute of Neurology and Genetics, 2370 Nicosia, Cyprus
[39] Department of Paediatric Neurology, 2nd Faculty of Medicine, Charles University and Motol Hospital, Prague, Czech Republic
[40] Epilepsy Center Frankfurt Rhine-Main, Center of Neurology and Neurosurgery, Goethe University Frankfurt, Frankfurt, Germany
[41] Epilepsy Center Hessen-Marburg, Department of Neurology, Philipps University Marburg, Marburg, Germany
[42] LOEWE Center for Personalized Translational Epilepsy Research (CePTER), Goethe University Frankfurt, Germany
[43] Departments of Clinical Neurosciences, Medical Genetics and Community Health Sciences, Hotchkiss Brain Institute & Alberta Children's Hospital Research Institute, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada
[44] Pediatric Neurology, Children's Hospital, Giessen, Germany
[45] Department of Neurology, Medical University of Vienna, Vienna, Austria
[46] Institute of Experimental Epileptology and Cognition Research, University Bonn, 53127 Bonn, Germany
[47] Department of Epileptology, University Bonn, 53127 Bonn, Germany
[48] Institute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel, Kiel, Germany
[49] Department of Neuropediatrics, Christian-Albrechts-University of Kiel, 24105 Kiel, Germany
[50] DRK-Northern German Epilepsy Centre for Children and Adolescents, 24223 Schwentinental-Raisdorf, Germany
[51] Sackler School of Medicine, Tel-Aviv University, Ramat Aviv, Israel
[52] Tel-Aviv University Sackler Faculty of Medicine, Ramat Aviv 69978, Israel
[53] Hereditary Research Lab, Bethlehem University, Bethlehem, Palestine
[54] Bonifatius Hospital Lingen, Neuropediatrics Wilhelmstrasse 13, 49808 Lingen, Germany
[55] Epilepsy Center for Children, University Hospital Neuruppin, Brandenburg Medical School, Neuruppin, Germany
[56] Neuropediatric Clinic and Clinic for Neurorehabilitation, Epilepsy Center for Children and Adolescents, Vogtareuth, Germany
[57] Research Institute Rehabilitation / Transition / Palliation, PMU Salzburg, Austria
[58] Division of Neuropediatrics and Social Pediatrics, Department of Pediatrics, University Hospital, RWTH Aachen, Aachen, Germany
[59] Department of Neurology, Sheba Medical Center, Ramat Gan, Israel
[60] Institute of Human Genetics, Leipzig, Germany
[61] Department of Epileptology and Neurology, University of Aachen, Aachen, Germany
[62] Department of Neurology, University of Ulm, Ulm, Germany
[63] Epilepsy Center Kork, Kehl, Germany
[64] Department of Epileptology, University Hospital Freiburg, Freiburg, Germany
[65] Department of Pediatric Neurology, Ludwig Maximilians University, Munchen, German
[66] Dr von Hauner Childrens Hospital, Munich, Germany
[67] Department of Neurology, Ludwig Maximilians University, Munchen, Germany
[68] Department of Neurology, University Hospital Bern, Bern, Switzerland
[69] Neurological Practice Center & Neuropoint Patient Academy, Ulm, Germany
[70] Epilepsy Center Kleinwachau, Radeberg, Germany
[71] Department of Neuropediatrics and Muscular Disorders, University Hospital Freiburg, Freiburg, Germany
[72] University Children's Hospital, Göttingen, Germany
[73] Department of Pediatric Neurology & Developmental Medicine, University Childrens Hospital, Tubingen, Germany
[74] Department of Neuropediatrics, Westfalische Wilhelms University, Munster, Germany
[75] Private Neurological Practice, Stuttgart, Germany
[76] Department of Pediatrics, Filderklinik, Filderstadt, Germany
[77] Epilepsy Center, Kuopio University Hospital, Kuopio, Finland
[78] Institute of Clinical Medicine, University of Eastern Finland, Finland
[79] Child Neurology, New Children´s Hospital and Pediatric Research Center, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
[80] Medicum, University of Helsinki, Helsinki, Finland and Folkhälsan Research Center, Helsinki, Finland
[81] Department of Functional Neurology and Epileptology, Hospices Civils de Lyon and University of Lyon, France
[82] Lyon's Neuroscience Research Center, INSERM U1028 / CNRS UMR 5292, Lyon, France

[83] Department of Clinical Genetics, Hospices Civils de Lyon and University of Lyon, Lyon, France
[84] Department of Clinical Neurosciences, Centre Hospitalo-Universitaire Vaudois, Lausanne, Switzerland
[85] Neurology Department, University Hospital of Nancy, UMR 7039, CNRS, Lorraine University, Nancy, France
[86] Clinical Neurophysiology and Epilepsy Unit, Department of Neurology, University Hospital of Toulouse. Toulouse, France
[87] Department of Clinical Neurophysiology, Lille University Medical Center, EA 1046, University of Lille, France
[88] Clinical Neurophysiology and Epileptology Department. Timone Hospital, Marseille, France
[89] Department of Neurology, University Hospital of Strasbourg, Strasbourg, France
[90] Department of Neurology, Hôpital Pellegrin, Bordeaux, France
[91] Epilepsy Unit, Department of Neurosurgery, Centre Hospitalier Sainte-Anne and University Paris Descartes, Paris, France
[92] Faculty of Medicine and Health, University of Sydney, Sydney, Australia
[93] Kids Neuroscience Centre, Kids Research, Children Hospital at Westmead, Sydney, New South Wales, Australia
[94] Swansea University Medical School, Swansea University SA2 8PP, UK
[95] Department of Neurology, Morriston Hospital, Swansea Bay University Health Board, Swansea, Wales UK
[96] Neurology Department, Aneurin Bevan University Health Board, Newport, Wales, UK
[97] Division of Brain Sciences, Imperial College London, London, UK
[98] Department of Neurology, Walton Centre NHS Foundation Trust, Liverpool, UK
[99] School of Life Sciences, University of Glasgow, Glasgow, UK
[100] Department of Medicine and Therapeutics, Chinese University of Hong Kong, Hong Kong, China
[101] Department of Psychiatry, University of Hong Kong, Hong Kong, China
[102] Department of Epidemiology and Preventive Medicine and Department of Anatomy and Anthropology, Sackler Faculty of Medicine, Tel Aviv University, Israel
[103] Department of Medicine, Tseung Kwan O Hospital, Hong Kong, China
[104] The Department of Neurology, Beaumont Hospital, Dublin, Ireland
[105] Neurology Department, St. James Hospital, Dublin, Ireland
[106] Department of Pediatric University Hospital centre Zagreb, Croatia
[107] Integrated Diagnostics for Epilepsy, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy
[108] Neurophysiology, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy
[109] Unit of Genetics of Neurodegenerative and Metabolic Diseases, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy
[110] Department of Pediatric Neuroscience, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy
[111] IRCCS, Institute of Neurological Sciences of Bologna, Bologna, Italy
[112] Department of Biomedical and Neuromotor Sciences, University of Bologna, Bologna, Italy
[113] Medical Genetics Unit, Polyclinic Sant'Orsola-Malpighi University Hospital, Bologna, Italy
[114] Institute of Neurology, Department of Medical and Surgical Sciences, University “Magna Graecia”, Catanzaro, Italy
[115] Institute of Molecular Bioimaging and Physiology, CNR, Section of Germaneto, Catanzaro, Italy
[116] Department of Pediatrics, School of Medicine, Fukuoka University
[117] Laboratory for Neurogenetics, RIKEN Center for Brain Science, Saitama, Japan
[118] Department of Neurodevelopmental Disorder Genetics, Institute of Brain Science, Nagoya City University Graduate School of Medical Science, Nagoya, Aichi, Japan
[119] National Epilepsy Center, Shizuoka Institute of Epilepsy and Neurological Disorder, Shizuoka, Japan
[120] Department of Neurology, American University of Beirut Medical Center, Beirut, Lebanon
[121] Department of Pharmacology and Toxicology, American University of Beirut Faculty of Medicine, Beirut, Lebanon
[122] Institute of Biomedical Sciences, Faculty of Medicine, Vilnius University, Vilnius, Lithuania
[123] Centre for Medical Genetics, Vilnius University Hospital Santaros Klinikos, Vilnius, Lithuania
[124] Department of Paediatrics and Child Health, University of Otago, Wellington
[125] Department of Molecular Biology and Genetics, Bogaziçi University, Istanbul, Turkey
[126] Department of Child Neurology, Gulhane Education and Research Hospital, Health Sciences University, Ankara, Turkey
[127] Department of Child Neurology, Istanbul Faculty of Medicine, Istanbul University, Istanbul, Turkey
[128] Department of Child Neurology, Medical School, Marmara University, Istanbul, Turkey
[129] Department of Neurology, Istanbul Faculty of Medicine, Istanbul University, Istanbul, Turkey
[130] Department of Neurology, Kaohsiung Chang Gung Memorial Hospital, Kaohsiung, Taiwan
[131]
[132] Epilepsy Genetics Program, Department of Neurology, Boston Children's Hospital, Boston, MA 02115, USA
[133] Department of Neurology, Harvard Medical School, Boston, MA 02115, USA
[134] Neurology, Baylor College of Medicine
[135] Cleveland Clinic Epilepsy Center, Neurological Institute, Cleveland Clinic, Cleveland, OH 44195, USA
[136] Department of Neurology, Neurological Institute, Cleveland Clinic, Cleveland, OH 44195, USA
[137] Cincinnati Children's Hospital Medical Center, Cincinnati, Ohio, USA

[138] Cooper Medical School of Rowan University, Camden, NJ 08103, USA
[139] Thomas Jefferson University, Philadelphia, PA 19107, USA
[140] Nationwide Children's Hospital, Columbus, OH 43205, USA
[141] University of Cincinnati, Cincinnati, OH 45220, USA
[142] Department of Neurology, New York University, New York NY, USA
[143] Beth Israel Deaconess/Harvard, Boston, MA 02115, USA
[144] Department of Neurology, Hofstra-Northwell Medical School, New York, NY 11549, USA
[145] Department of Pediatrics, Ohio State Wexner Medical Center, Battelle Center for Mathematical Medicine. Columbus, Ohio, USA.
[146] The Epilepsy NeuroGenetics Initiative (ENGIN), Children’s Hospital of Philadelphia, Philadelphia, PA 19104, USA

Details of individual participating Epi25 cohorts

Australia: Austin Hospital, Melbourne (AUSAUS)
The Epilepsy Research Centre at the Austin Hospital in Melbourne, Australia, has been investigating the genetic basis of the epilepsies for over 20 years. The cohort in the Epi25 Collaborative were recruited to the epilepsy genetics research program over this period from the Austin Hospital, epilepsy clinics around Melbourne, and referrals from neurologists Australia-wide. Informed consent was obtained from individuals with epilepsy or their parent/guardian as appropriate. DNA was extracted from blood or saliva samples. A skilled team of researchers and clinicians conducted detailed clinical phenotyping which involved a systematic review of medical records, including EEG and MRI reports, and a validated epilepsy questionnaire. Information on family history of seizures and other neurological disorders has also been collected via interviews with the invdividuals with epilepsy and their families. Individuals with GGE, non-acquired focal epilepsy, or a DEE were included in Epi25. A very heterogenous collection of epilepsy syndromes are represented in the cohort, including individuals with EOAE, CAE, JME and late-onset GGE in the GGE cohort and individuals with TLE, FLE, and benign childhood focal epilepsies in the non-acquired focal cohort. The DEE cohort is particularly heterogenous and includes individuals with a range of DEE syndromes such as Ohtahara syndrome, Lennox-Gastaut syndrome, epilepsy with myoclonic-atonic seizures, and non-syndromic DEE. An additional subset of individuals with lesional focal epilepsy, such as malformations of cortical development or acquired epilepsy, were also included, as well as a selection of individuals with familial febrile seizures or FS+.

Most individuals in the cohort are of European geographic descent ('Anglo-Australian') although there is a diverse range of ethnic backgrounds including Asian, Middle Eastern, Indigenous Australian and mixed ethnicities. There is a known family history of seizures in 44% of the cohort (57% in the subset with GGE). The majority of the individuals have had some previous genetic testing, including CNV testing and single gene testing. The DEE cohort have been extensively investigated with multiple iterations of a research panel of known, novel and putative genes for epilepsy. In addition, many individuals with focal epilepsy have had a panel of known genes.

Australia: Royal Melbourne Hospital, Melbourne (AUSRMB)
The Royal Melbourne Hospital Cohort, Melbourne, Australia, was prospectively recruited from the Epilepsy and First Seizure Clinics of The Epilepsy Program of the Royal Melbourne Hospital. Phenotypic information was obtained by direct interview, and review of the medical records, for the enrolled individuals. Blood for DNA extraction was obtained on all participants and stored in the Biobank of the RMH Epilepsy Program in the Department of Medicine, The Royal Melbourne Hospital, The University of Melbourne. Written informed consent was obtained for all participants, and the recruitment and study procedures were approved by the Human Research and Ethics Committee of Melbourne Health (The Royal Melbourne Hospital) - HREC #2002.232 & 2017.450. Recruitment and phenotyping of the individuals from the Royal Melbourne Hospital, Victoria, Australia.[3-8] Funding of this work. Participation in the Phenotyping Committee of Epi25K. Critical review of the manuscript.

Belgium: Antwerp (BELATW)
Individuals were recruited by the VIB-Applied&Translational Neurogenomics Group of the University of Antwerp through epilepsy clinics at the different university hospitals in Belgium. All individuals were diagnosed with a (so far) unexplained presumed genetic epilepsy, and should have had at least 1 MRI of the brain excluding acquired causal lesions.The study was approved by the ethics committee of the University of Antwerp, and parents or the legal guardian of each proband signed an informed consent form for participation in the study. Genomic DNA of individuals was extracted from peripheral blood according to standard procedures. Clinical information was extracted from clinical files, as reported by their treating (paediatric) neurologists, and a subset was reviewed independently by two research team clinicians to ensure data quality and consistency.

Belgium: Brussels (BELULB)
Adult individuals with epilepsy were recruited consecutively through outpatient clinics and hospitalizations at Hôpital Erasme, Brussels, Belgium (between October 2004 and June 2017) and UZ Gasthuisberg, Leuven, Belgium (between October 2004 and June 2009). The study was approved by the Institutions' Review Boards. All individuals provided written informed consent for data collection; individuals with learning disability were included after consent from a parent or guardian. DNA was extracted from peripheral blood lymphocytes.[9] Clinical information was collected from medical records and stored in a secured, web-based database.

Canada: Andrade (CANUTN)
There are 86 individuals (37 DEE, 41 GGE, 8 NAFE) in the Andrade cohort, typically from the Greater Toronto region of Ontario, Canada. They are mostly of European geographic ancestry, but also African, South Asian, East Asian, Latino, Middle Eastern, Jewish and Indigenous. Individuals were recruited to each group through an REB protocol allowing for the collection of blood or saliva and data collaboration. Individuals that were previously consented were re-consented to allow for Whole Exome Sequencing and data sharing with the EPI25 group. After collection, the sample was de-identified, and

then extracted and stored at the Hospital for Sick Children, Toronto, Canada. Select samples of this cohort have been published in a few studies.[10-14]

Switzerland: Bern University Hospital and University of Bern, Bern (CHEUBB)
In the recruitment of our cohort, the Departments of Neurology and BioMedical Research, Bern University Hospital and University of Bern, Bern, Switzerland, and the Institute of Human Genetics, Bern University Hospital, Bern, Switzerland, were involved. The Swiss study population encompasses > 90 individuals (28 in year 2, 70 in year 4) with epilepsy between 2 and 63 years of age. All individuals have been de-identified for the Epi25 Study. The individual ascertainment protocol was according to Epi25 phenotyping requirements. Phenotyping information was taken from medical records, stored in the hospital's database, and entered in de-identified form into a RedCap database provided by Ep25. DNA source was individuals' venous blood. DNA was extracted with standard kits at the Institutes of Human Genetics or Clinical Chemistry of Bern University Hospital and stored there at -80 degrees C. Informed consent declarations are available from all individuals and have been approved by the American Institutional Review Board involved in the Epi25 Study. The Cantonal Ethics Committee Bern, Switzerland, granted permission for participation of Bern University Hospital and University of Bern in the Epi25 Study including all steps described above.

Cyprus: The Cyprus Institute of Neurology and Genetics (CYPCYP)
Epilepsy-affected subjects of the Cyprus cohort were largely recruited and enrolled in the Epi25 Consortium by physicians during routine clinical visits in the Cyprus Institute of Neurology and Genetics. Phenotypic data were collected at the time of enrollment and submitted into the Epi25 RedCap database in a de-identified manner. There are 123 unrelated individuals of Southern European ancestry in the Cyprus cohort, 59 GGE subjects, 53 NAFE and 11 DEE. All subjects selected for this study had clinical, neuroimaging and EEG or video-EEG characteristics meeting the International League against Epilepsy (ILAE) 2017 Seizure Classification. The controls cohort consisted of a group of 32 individuals of Southern European ancestry and were not diagnosed with epilepsy or other neuropsychiatric phenotypes. Genomic DNA samples were extracted from whole blood with the Gentra Puregene Blood Kit (Qiagen, Hilden, Germany) according to the manufacturer's guidelines. This study was carried out in compliance with the Cyprus National Bioethics Committee (ΕΕΒΚ/ΕΠ/2015/22). Written informed consent was obtained from all study participants or their legal guardians at the Cyprus Institute of Neurology and Genetics.

Czech Republic: University Hospital Motol, Prague (CZEMTH)
Individuals in our cohort have been diagnosed with West syndrome, myoclonic-astatic epilepsy or developmental and epileptic encephalopathy of unknown aetiology. Brain magnetic resonance imaging and metabolic screening excluded any underlying pathology. Individuals were collected at the Department of Child Neurology of the 2nd Medical Faculty and University Hospital Motol. Legal guardians of individuals signed an informed consent. The study was approved by the local ethics committee.

Germany: Epilepsy Center Frankfurt Rhine-Main, Goethe University, Frankfurt, and Epilepsy Center Hessen, Philipps University, Marburg (DEUPUM)
KMK, FR, SK and PSR contributed 348 samples. The individuals were recruited from the outpatient clinics and the video EEG monitoring units at the Epilepsy Centers Frankfurt Rhine-Main and Hessen-Marburg. All individuals were phenotyped in detail by epilepsy specialists (KMK, FR, SK, PSR) within the EpimiRNA project (European Union's 'Seventh Framework' Programme (FP7) under Grant Agreement no. 602130). EEGs and MRIs were performed as part of the clinical workup. Phenotypic classification and data entry for the biobank for paroxysmal neurological disorders was performed by KMK and PSR. Individual selection and data export from the biobank for the study was performed by KMK. DNA was extracted from peripheral blood or saliva. All individuals provided written informed consent.

Germany: University of Giessen, Giessen (DEUUGS)
Diagnosis of Rolandic epilepsy was performed according to the International Classification of Seizures and Epilepsies as described. Sleep activation, characteristic shape, and classification by two independent individuals were required for classification of the EEG trait. Atypical benign partial epilepsy of childhood (ABPE) was diagnosed employing the following criteria: Characteristic EEG trait of CTS, however, with trains of continuous generalized nocturnal discharges as a prerequisite of diagnosis in all ABPE cases. In addition, at least one of the following two features needed to be present: (1) seizures and EEG trait compatible with BECTS plus one or more additive seizure types like astatic seizures, atypical absences ("dreamy states") or myoclonic seizures as reported. (2) seizures compatible with BECTS plus a significant mental handicap, and/or severe developmental speech disorder.[15]

Germany: University of Bonn, Bonn (DEUUKB)
The sample recruitment site is the Department of Epileptology at the University of Bonn. The collection of 2036 blood DNA samples from individuals with epilepsy which were includes in the present study was conducted from 2007 till 2015 within the projects Epicure (Functional Genomics in Neurobiology of Epilepsy: A Basis for New Therapeutic Strategies) and NGEN-Plus (Genetic basis of Levetiracetam pharmacoresistance and side effects in human epilepsy) and has been

approved by the Ethics committee of University Bonn Medical Center (040/07). Genomic DNA was isolated from 10 ml aliquots of EDTA-anticoagulated blood by a salting-out technique.[16] From selected samples of this cohort GWAS data have been published in several studies.[17-20]

Germany: University Hospital Schleswig-Holstein, Kiel (DEUUKL)
Individuals were recruited by the Neuropedriatics Group of the University Hospital of Schleswig-Holstein and through the Israeli-Palestinian Family Consortium. The recruitment and analysis of these samples is covered by the Kiel IRB. Individuals with epilepsy, their parents or the legal guardian of each proband signed an informed consent form for participation in the study. Clinical data was collected from clinical files and a subset of individuals from Israel or Palestine was interviewed by a research team of clinicians to provide their clinical data. Genomic DNA of individuals was extracted from peripheral blood according to standard procedures.

Germany: TLE Leipzig (DEUULG)
Individuals were recruited by the Swiss Epilepsy Center in Zurich, Switzerland and samples were transferred for research and storage to the Institute of Human Genetics at the University of Leipzig, Germany. All individuals were diagnosed with temporal lobe epilepsy due to an indicative EEG. Most individuals had at least 1 MRI of the brain with focus on focal abnormalities, especially of the temporal lobe / hippocampal structures. The study was approved by the “Kantonale Ethikkommission Zürich”. Parents or the legal guardian of each proband signed an informed consent form for participation in research studies including whole genome analyses. Genomic DNA of individuals was extracted from peripheral blood according to standard procedures. Clinical information was extracted from clinical files, as reported by their treating neurologists.

Germany: University of Tübingen, Tübingen (DEUUTB)
Our study cohort consists of more than 1000 samples with mainly European geographic ancestry. These samples were recruited at Tubingen and 38 other cooperating departments of neurology from university clinics and outpatient clinics in Germany. The Ethics / informed consent was approved by the ethics committee of the Medical Faculty of the Eberhard-Karls University and at the University Hospital Tubingen. Individuals with idiopathic generalized epilepsies, epileptic encephalopathies, non-acquired and acquired focal epilepsies were systematically recruited in outpatient clinics of university and other hospitals, and from neurological practices by a letter of invitation sent to individuals. Retrospective data from medical reports of epileptologists were used. If deemed necessary, personal interviews of individuals were undertaken. The DNA source was blood. There is no single publication describing the whole sample. Typical publications including part of these samples are from the following consortia: Epicure, EuroEPINOMICS, EpiPGX, ILAE consortium on the genetics of complex epilepsies, Epi25.

Finland: Kuopio University Hospital, Kuopio (FINKPH)
Individuals diagnosed with epilepsy and visiting Epilepsy Center, Kuopio University Hospital (KUH), Finland have given their written informed consent to record their clinical data to a epilepsy research registry of KUH and University of Eastern Finland and collect a blood sample for DNA analysis. Consent was collected from the legal quardian, if applicable. The ethics committee of KUH has approved the study.

Finland: University of Helsinki, Helsinki (FINUVH)
Individuals were recruited at the University of Helsinki through pediatric epilepsy clinics in Helsinki and Tampere University Hospitals in Finland. All individuals were diagnosed with a presumed genetic epilepsy, the etiology remaining unknown. All individuals had an MRI done to exclude acquired causal lesions. This cohort included 95 individuals with epilepsy: 26 individuals with developmental and epileptic encephalopathy (DEE), 51 individuals with genetic generalized epilepsy (GGE) and 18 individuals with non-acquired focal epilepsy (NAFE). Clinical phenotyping involved a systematic review of medical records, including EEG and MRI reports. The study was approved by an ethics committee of The Hospital District of Helsinki and Uusimaa, Finland. The parents or the legal guardian of each proband signed an informed consent form for participation in the study. Genomic DNA of the individuals was extracted from peripheral blood or saliva according to standard procedures.

France: REPOMSE Cohort (FRALYU)
The REPO2MSE study is a multicenter prospective study, based on the French National Research Network on SUDEP predictors, which was approved by ethics committee (CPP Sud Est II n°2010-006-AM6) and competent authority (ANSM n° B100108-40).[21; 22] Its primary objective is to individualize risk factors of SUDEP in individuals suffering from drug-resistant focal epilepsy. 1069 Adult individuals (age ≥16 years) with drug-resistant focal epilepsy according to ILAE classifications who underwent long term monitoring using either video scalp EEG or intracranial EEG recordings and who gave written informed consent were recruited in 16 French epilepsy monitoring units. For all included individuals, we collected demographic and detailed clinical data, MRI data, inter-ictal EEG data, results of non-systematic complementary investigations performed to better localize the epileptogenic zone (i.e 18FDG PET, ictal SPECT) and raw data of all recorded seizures, which include EEG, video, pulse oximetry and EKG. For individuals who gave specific consent, we

also collected blood samples for genetic analyses which were centrally stored in the Department of Clinical Genetics at Hospices Civils de Lyon. All individuals then received a specific information about the collaboration between the REPOMSE study and the EPI25 project. Overall, 810 individuals from twelve participating centers confirmed their consent for transmission of their blood samples to the Broad Institute.

Wales: Swansea (GBRSWU)
The samples from Wales: Swansea are part of the Swansea Neurology Biobank (SNB). The SNB has been approved by the Welsh Research Ethics Committee (REC 17/WA/0290). Participants are recruited into the biobank, with written informed consent or assent, from regional National Health Service (NHS) neurology and epilepsy clinics. Participants provide written consent to share their clinical and genetic information anonymously with ethical research collaborations. Participants' medical records (including EEG and MRI results and epilepsy clinic letters) are reviewed by the research and clinical team (which include experienced epileptologists) to confirm diagnosis. SNB participants blood samples are sent to the UK Porton Down ECACC facility for DNA extraction and the DNA is then returned to be stored securely at Swansea University. For epi25 we have submitted approximately 310 bio-samples and individual records in Yrs. 1-5.

UK: University College London, London (GBRUCL)
Participant recruitment took place at the National Hospital for Neurology and Neurosurgery (United Kingdom). Written informed consent or assent was obtained between 10/01/2000 and 01/25/2015 from all participants according to local and national requirements and blood samples were collected for DNA extraction. 709 epilepsy cases were submitted for analysis.[7; 23] Allocation to the following groups was based on the clinical diagnosis and the specific inclusion and exclusion criteria of the Epi25 consortium: generalized genetic epilepsy (n=393, 145 male), non-acquired focal epilepsy (n=313, 146 male), developmental and epileptic encephalopathy (n=3, 2 male). Additionally, relatives were included, where samples were available (n=3, 2 male). Phenotypic information was obtained from local medical records by clinical or trained non-clinical researchers.

UK: University of Liverpool, Liverpool and Imperial College London, London (GBRUNL)
GBRUNL samples are derived from four separate, UK-wide, ethically approved studies coordinated by the University of Liverpool (UK) and Imperial College London (UK). The SANAD and MESS linked DNA Bank and Relational Database study recruited individuals with newly-diagnosed focal, generalised or unclassified epilepsy from out-patient neurology clinics between 2003-2006.[7; 23] The Pharmacogenetics of GABAergic Mechanisms of Benefit and Harm in Epilepsy study recruited individuals with refractory focal epilepsy, previously or prospectively exposed to adjunctive treatment with clobazam or vigabatrin, from out-patient neurology clinics between 2005-2009. The Refractory Juvenile Myoclonic Epilepsy Cohort (ReJuMEC) study recruited individuals with valproic acid resistant juvenile myoclonic epilepsy from out-patient neurology clinics between 2009 and 2010. The ongoing Standard and New Antiepileptic Drugs (SANAD-II) study, which is recruiting individuals with newly-diagnosed focal, generalised or unclassified epilepsy from out-patient neurology clinics between 2013-2019. In all cases, study participants provided written informed consent to the collection (via blood or saliva sampling) and analysis of their DNA for use in genetic and pharmacogenetic research related to epilepsy and its treatment. All studies were approved by research ethics committees in operation at the relevant time (SANAD DNA bank, North West MREC ref 02/8/45; GABAergic mechanisms, UCLH REC ref 04/Q0505/95; ReJuMEC, Cheshire REC ref 09/H1017/55; SANAD-II, North West REC ref 12/NW/0361). Assembly of the GBRUNL cohort was supported by generous funding from The Wellcome Trust, the Imperial College NIHR Biomedical Research Centre, the Department of Health (UK), the Medical Research Council (UK), and the National Institute of Health Research (UK).

Hong Kong: Chinese University of Hong Kong (HKGHKK)
Epilepsy individuals of Han Chinese ethnicity aged between 2 and 91 years were recruited from neurology clinics of five regional hospitals in Hong Kong covering a combined catchment population of approximately 3 million. Syndromic classification was adapted from the revised international organization of phenotypes in epilepsy. DNA was extracted from venous blood. The study was approved by ethics committees of the participating hospitals, and all individuals or their legal guardians gave written informed consent. The sample collection methodology has been described prevoiusly.[24]

Croatia: University Clinical Centre Zagreb, Zagreb (HRVUZG)
Pediatric individuals were recruited from University Medical Centre Zagreb and 2 individuals from 2 other epilepsy clinics in Croatia. All individuals were diagnosed as possible genetic epilepsy not yet explained. All individuals underwent MR brain imaging at least once, the acquired epilepsy causes were excluded. DNA was extracted from peripheral blood according to the accepted protocol. The study was approved by Hospital ethical committee and all parents or legal guardian of probands signed informed consent for participation in the study. Clinical information was extracted from clinical files. The cohort was also reviewed by reviewed by collaborative research team clinicians from University of Antwerp to ensure data quality and consistency.

Ireland: Dublin (IRLRCI)

Individuals were all adults and recruited from a specialized epilepsy clinic at Beaumont Hospital, Dublin, Ireland. Individuals were mostly of Irish ethnicity. This study was approved by the Beaumont Hospital Ethics Committee.[20]

Italy: Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan (ITAICB)
Our cohort included the DNA samples of 303 individuals with epilepsy and 62 controls (healthy subjects not related with the individuals with epilepsy, and without epilepsy history). The population included individuals with generalized epilepsy (GGE), individuals with developmental and epileptic encephalopathy (DEE), individuals with focal epilepsy due to cerebral malformations (mainly nodular heterotopia), and individuals with nonacquired focal epilepsy (NAFE). All of the individuals were diagnosed and followed at our Institute. Diagnosis of epilepsy was based on clinical, EEG and neurophysiological data, neuroimaging (MRI). Metabolic screening, karyotype, CGH array, analyses of single genes and customized panels were performed in some cases, when appropriate. The individuals did not undergo to exome sequencing analysis (before the Epi25 collection). The DNA of the individuals was extracted from peripheral blood, according with standard procedures, after signature of an Informed Consent form. The genetic study was approved by The Ethic Committee of our Institute. No publications have described genetic findings pertaining to the collected individuals until now. Clinical information was extracted from clinical files, as reported by their treating (paediatric and adult) neurologists.

Italy: Gaslini Institute, Genova (ITAIGI)
Individuals with generalized and focal epilepsy or developmental epileptic encephalopathy referred for to 'IRCCS G. Gaslini Institute'. The study was approved by the IRB and written informed consent was signed by the individuals/parents. Clinical information, including data on EEG and antiepileptic therapy, were recorded on data collection forms. Genomic DNA isolation and genetic analysis was carried out with the Nimblegen-SeqCapEZ-V244M enrichment kit on the Illumina HiSeq2000 system.[25-28]

Italy: IRCCS Institute of Neurological Science of Bologna, Bologna (ITAUBG)
323 unrelated individuals were consecutively recruited by the Adult and Pediatric Neurologists of the IRCCS Institute of Neurological Sciences, Bellaria Hospital, Bologna. Individuals with DEE (n=110), GGE (n=68) and NAFE (with or without brain lesions) (n=145) were referred by epilepsy clinics. All individuals were diagnosed with a (so far) epilepsy of uncertain aetiology and underwent neuro-radiological imaging (CT or MRI) and EEG. The local ethical committee approved the study. Specific consent was obtained from all individual participants. Genomic DNA of individuals was extracted from peripheral blood according to standard procedures. Clinical information was collected from medical records, as reported by their treating neurologists.[29; 30]

Italy: University Magna Graecia, Catanzaro (ITAUMC)
Individuals were recruited by the Epilepsy Group of the University Magna Graecia of Catanzaro (Italy) that includes a Pediatric and Adult Neurologic Unit with a specific focus on genetic epilepsy. In each individual, the diagnosis of epilepsy syndrome is based on comprehensive clinical, neuropsychological, electroencephalographic, and MR evaluations. Clinical data are stored into a database. The study was approved by the ethics committee of the University of Catanzaro Italy, and parents or the legal guardian of each proband signed an informed consent form for participation in the study. Genomic DNA of individuals was extracted from peripheral blood according to standard procedures.

Italy: Meyer Hospital, Florence (ITAUMR)
Individuals were studied at the Clinical Neurology Unit and Neurogenetics lab of the Neuroscience Department of Meyer Children's Hospital-University of Florence. All individuals were diagnosed with a unexplained presumed genetic epilepsy. The study was approved by the Pediatric Ethics Committee of the Tuscany Region. Parents or the legal guardians of each proband signed an informed consent form for participation in the study. Genomic DNA of individuals was extracted from peripheral blood according to standard procedures. Clinical information was extracted from clinical files, as reported by the treating neurologists.

Japan: RIKEN Institute, Tokyo (JPNRKI)
Japanese individuals with epilepsies were recruited by National Epilepsy Center, Shizuoka Institute of Epilepsy and Neurological Disorder. Epileptic seizures and epilepsy syndrome diagnoses were performed according to the International League Against Epilepsy classification of epileptic syndromes. Genomic DNA was extracted from peripheral venous blood samples using QIAamp DNA Blood Midi Kit according to manufacturer's protocol (Qiagen). The experimental protocols were approved by the Ethical Committee of RIKEN Institution and National Epilepsy Center, Shizuoka Institute of Epilepsy and Neurological Disorder. Written informed consent was obtained from all individuals and/or their families in compliance with the relevant Japanese regulations.

Lebanon: American University of Beirut Medical Center, Beirut (LEBABM)
Individuals with epilepsy were recruited at the American University of Beirut Medical Center. All recruited individuals participated in an ongoing centralized study evaluating the electroclinical syndromes of children and adults with new onset unprovoked seizures or epilepsy in Lebanon, which so far enrolled more than 4,000 individuals. As part of the protocol, all

individuals were evaluated with a dedicated seizure protocol MRI and a 3-hour sleep deprived video/EEG. The study was approved by the Institutional Review Board of the American University of Beirut Medical Center and all individuals or parents/guardians signed an informed consent form. Genomic DNA of the individuals was extracted from peripheral blood according to standard procedures. Clinical information for phenotyping was extracted from the case report files.[31; 32]

Lithuania: Vilnius University Hospital Santaros Klinikos, Vilnius (LTUUHK)
Individuals with epilepsy were recruited in Vilnius University Hospital Santaros Klinikos by a clinical geneticist through a referral of a neurologist or a child neurologist and according to inclusion/ exclusion criteria. The study was approved by Institution's Research Ethics Committee, and each proband or parents/ legal guardians of a proband signed an informed consent. Samples of genomic DNA were obtained during the routine procedure for blood sampling for genetic testing done in a clinical testing and the majority of individuals had chromosomal microarray, metabolic testing and/or gene/gene panel testing prior to the inclusion into the study. Clinical information was extracted from clinical files and obtained during the clinical genetic consultation.

New Zealand: University of Otago, Wellington (NZLUTO)
Cases were recruited as part of a larger study from neurology, paediatric and genetic outpatient services throughout New Zealand. Participants were between 1 month and 63 years of age from the following ethnic groups: New Zealand European, Māori, Pacific Peoples, Asian, Hispanic, Ethiopian, Middle Eastern. Using a structured interview and review of medical records, diagnosis was based on the International League of Epilepsy (ILAE) classification and made by a paediatric neurologist. The study protocol was approved by the New Zealand Health and Disability Ethics Committee. Participants gave written informed consent for clinical and genetic analysis. DNA was extracted from blood or saliva.

Turkey: Bogazici University, Istanbul (TURBZU)
Individuals with epilepsy were recruited by the Child Neurology and Neurology clinics at the different university hospitals in Turkey. The study was approved by the Institutional Review Board for Research with Human Subjects (INAREK) of Boğaziçi University, and parents or the legal guardian of each proband signed an informed consent form for participation in the study. Genomic DNA of individuals was extracted from peripheral blood according to standard procedures. Clinical information was reported by their treating (pediatric) neurologists. The cohort included a total of 171 individuals (128 EE, 28 GGE and 15 Focal epilepsy individuals) and 39 healthy controls. All epileptic encephalopathy individuals had severe epilepsy, with developmental delay and regression, normal neuroimaging and epileptiform activity on EEG. Healthy control group included individuals with no symptoms of any neurological disorder.

Turkey: Istanbul University, Istanbul (TURIBU)
Epilepsy individuals were recruited from Epilepsy Clinic, Department of Neurology, Istanbul Faculty of Medicine, Istanbul University. The study population consisted of individuals with idiopathic/genetic generalized epilepsies, lesional or non-lesional focal epilepsies and epileptic encephalopathies, including sporadic and familial cases. All individuals were long-term follow-up. Seizure types, age of onset, neurological examinations, past and family history, prognosis and response to treatment, features of electroencephalography and neuroimaging were evaluated. Ethics committee approval was obtained. Peripheral blood samples were collected from all individuals following written informed consent. DNA isolation was performed in Department of Genetics, Aziz Sancar Institute of Experimental Medicine, Istanbul University.

Taiwan: Kaohsiung Chang Gung Memorial Hospital, Kaohsiung (TWNCGM)
Recruitment site(s)/institution(s): Kaohsiung and LinKo Chang Gung Memorial Hospital, Taiwan; Study Population: Taiwanese; The study is approved by local institutional review board at Kaohsiung Chang Gung Memorial Hospital, Taiwan.

USA: Boston Children's Hospital, Boston (USABCH)
Cases from Boston Children's Hospital (BCH) were ascertained from 3 local repositories. All repository protocols are approved by the BCH Institutional Review Board and participants were consented under one (or more) of the following protocols. The Genetics of Epilepsy and Related Disorders protocol, led by Dr. Annapurna Poduri, enrolls individuals with a clinical epilepsy diagnosis for genotype/phenotype correlation. Samples are obtained from BCH and non-BCH individuals and biological samples collected for genetic sequencing. Individual medical records (BCH and outside records) are reviewed for phenotyping purposes.[33] The Phenotyping and Banking Repository of Neurological Disorders is a local repository led by Dr. Mustafa Sahin. BCH individuals with any neurological phenotype, including epilepsy, are enrolled and biological samples collected. Boston Children's Biobank for Health Discovery is a local repository led by Dr. Kenneth Mandl that enrolls any individual of BCH, regardless of diagnosis or phenotype. Samples from these two broader repositories are available to BCH researchers through an application process, including a supporting IRB-approved protocol. Individuals with a clinical diagnosis of epilepsy were reviewed for Epi25 eligibility using their BCH medical records.[33]

USA: Baylor College of Medicine (USABLC)

Healthy controls and individuals with genetic epilepsies

USA: Cleveland Clinic (USACCF)

Individuals with epilepsy were recruited through the Cleveland Clinic Epilepsy Center. All individuals had routine EEG and/or video-EEG monitoring and had been diagnosed with epilepsy. The study was approved by the Cleveland Clinic Institutional Review Board, and all participants (or their guardian/legally authorized representative) provided informed consent for study participation. Genomic DNA of individuals was extracted from peripheral blood according to standard procedures. Clinical information was extracted from electronic health records.

USA: Cincinnati Children's Hospital Medical Center (USACCH)

The samples were from subjects involved with the NIH funded 32 center Childhood Absence Epilepsy clinical trial (ClinicalTrials.gov Identifier: NCT00088452).[34] The subjects were children between 2.5 and 13 years old with newly diagnosed, EEG proven absence seizures who met ILAE criteria for Childhood Absence Epilepsy. Blood was obtained at the first treatment visit for DNA isolation. All subjects (or their parents/guardians) signed written informed consent permitting DNA isolation, storage, and pharmacogenetic analysis. Those informed consents allowing for sharing and broader genetic analysis were shared with Epi25K.

USA: Philadelphia/CHOP (USACHP) and Philadelphia/Rowan (USACRW)

The Philadelphia Cohort began in 1997 and collected blood, saliva and brain tissues from individuals with common forms of idiopathic human epilepsy, mostly genetic generalized epilepsy (GGE) and non acquired focal epilepsy (NAFE). The collection began at Thomas Jefferson University Hospital in Philadelphia and expanded to include six other sites: The Children's Hospital of Philadelphia, The University of Pennsylvania, The University of Cincinnati, Nationwide Children's Hospital, Beth Israel Deaconess and The University of Montreal. The cohort consists of 2615 samples from epilepsy individuals collected and supported during two periods of NIH funding (R01NS493060, 2001-2007 RJ Buono PI and R01NS06415401, 2009-2012 RJ Buono and H Hakonarson Co- PI). Over 1000 additional samples from first degree relatives of the individuals with epilepsy were also collected. Many of these samples are available to the research community via the NINDS sample repository at the Coriell Institute in Camden NJ. All studies were approved by Institutional Review Boards at each participating site. All individuals were identified and recruited by trained epileptologists at tertiary care centers using inclusion and exclusion criteria previously published.[20; 35] Diagnostic methods applied included EEG, MRI, and collection of deep phenotypic information on family history, medications, risk factors, age of onset, and other information. For the Epi25K project, blood and saliva were used as the source of DNA.

USA: Epilepsy Phenome/Genome Project (USAEGP)

Individuals with Infantile spasms (IS), Lennox–Gastaut syndrome (LGS), genetic generalized epilepsy (GGE), and non-acquired focal epilepsy (NAFE) were collected through the Epilepsy Phenome/Genome Project[36] (EPGP, http://www.epgp.org). More than 4,000 participants in EPGP were enrolled across 27 clinical sites from around the world. The subset of samples included in Epi25 were enrolled from 20 sites across the USA and in Australia. IS individuals were required to have hypsarrhythmia or a hypsarrhythmia variant on EEG. LGS individuals were required to have EEG background slowing or disorganization for age and generalized spike and wave activity of any frequency or generalized paroxysmal fast activity (GPFA). IS and LGS cases were enrolled as trios with both biological parents. Participants with NAFE and GGE were required to have a first degree relative who also had NAFE or GGE (did not have to be concordant). All individuals had no confirmed genetic or metabolic diagnosis, and no history of congenital TORCH infection, premature birth (before 32 weeks gestation), neonatal hypoxic-ischaemic encephalopathy or neonatal seizures, meningitis/encephalitis, stroke, intracranial haemorrhage, significant head trauma, or evidence of acquired epilepsy. Enrollment required detailed confirmation of detailed phenotypic data including medical record review and abstraction, individual interviews, EEG and MRI, and comprehensive review by expert scientific cores for EEG, MRI, and clinical final diagnosis.[36]

USA: NYU Human Epilepsy Project (USAHEP)

Participants were recruited for the Human Epilepsy Project at 33 different medical centers located in US, Canada, Australia, Austria, Finland, and Ireland. All participants were between 12 and 60 at the age of enrollment, and had a clinical history consistent with a diagnosis of focal epilepsy, as determined by an eligibility panel of epilepsy specialists. Participants were required to have two or more spontaneous seizures with clinically observable features in the past 12 months, and 4 or fewer months of anticonvulsant treatment. Those with major medical comorbidities, intellectual disability, or significant psychiatric disease were excluded, as were those with progressive neurological lesions on imaging or known neurodegenerative disease. Participants completed daily electronic diaries tracking seizures, medication adherence, and mood. Mood and cognition were assessed periodically via standardized instruments, and brain MRIs and EEGs were obtained for all participants. Blood was collected and banked annually, allowing for study of DNA, RNA and protein. HEP was approved by the IRBs at all participating sites, and all participants or their parent/legal guardian gave written informed consent. Minors also gave written assents. HEP was funded by the Epilepsy Study Consortium.

USA: MONEAD (USAMON)
The Maternal Outcomes and Neurodevelopmental Effects of Antiepileptic Drugs (MONEAD) study is an NIH-funded, prospective, observational, multi-center investigation of pregnancy outcomes for both pregnant women with epilepsy (PWWE) and their children. Enrollment occurred from December 2012 – January 2016. Twenty clinical sites at tertiary USA epilepsy centers were selected that specialize in management of women with epilepsy during childbearing years. The 20 MONEAD sites included Augusta University, Columbia University, Emory University, Geisinger Clinic, Brigham and Women's Hospital of Harvard University, Henry Ford Health System, Johns Hopkins University, Minnesota Epilepsy Group, New York University, Northwell Health, Northwestern University, Stanford University, University of Alabama, University of Arizona, University of Cincinnati, University of Miami, University of Pittsburgh, University of Southern California, University of Washington, and Wake Forest University. PWWE were recruited primarily from the 20 epilepsy practices, but also via referral from obstetricians and other physicians. PWWE could also self refer. Inclusion criteria for PWWE were ages 14–45 years and <20 weeks gestational age. Exclusion criteria included history of psychogenic non-epileptic spells, expected IQ<70, other major medical illness, progressive cerebral disease, and switching AEDs in pregnancy prior to enrollment. Unlike the NEAD study which enrolled only PWWE on the most common monotherapies, MONEAD was specifically designed to enroll all PWWE regardless of treatment regimen in order to obtain a representative sample of PWWE and their AED treatments. Data were collected from subjects and their medical records. Cohort described more fully previously.[37]

USA: Nationwide Children's Hospital (USANCH)
Individuals with epilepsy were recruited at a number of sites one the East Coast and Midwest of the USA including: (in New York) Mt. Sinai Medical Center; Columbia-Presbyterian Medical Center; Long Island Jewish Hospital; Montefiore Medical Center; Albert Einstein Medical Center; St. Luke's Hospital; New York University Medical Center. Children's Hospital, Cincinnati. Beth Israel Hospital and Brigham and Women's Hospital, Boston. JFK Medical Center, Edison, New Jersey. Temple University Hospital, Philadelphia, Pennsylvania, as well as referrals from physicians. The study population was of mixed ethnicity. Ascertainment was through probands with strictly defined JME: myoclonic jerks on or shortly after awakening, AOO 8-25 years, generalized EEG, generalized tonic-clonic seizures and/or absence seizures were not exclusions (childhood absence was exclusionary). All individuals were given a clinical interview, as were family members to confirm the clinical picture. DNA was usually from blood and/or transformed white cells, sometimes saliva. Consent was obtained from all subjects.[38]

USA: Penn/CHOP, Philadelphia (USAUPN)
The Penn/CHOP cohort included adult and pediatric individuals with epilepsy who were seen in our inpatient or outpatient clinical epilepsy programs and enrolled in our ongoing epilepsy research protocols. Biospecimens stored in our institutional biobanks at the time of enrollment were contributed to Epi25. Phenotypic information was reviewed prior to specimen submission to confirm eligibility for Epi25.

Supplemental acknowledgments

Additional funding sources and acknowledgment of individual cohorts are as follows: The **AUSAUS** cohort was supported by NHMRC Program Grant (ID: 1091593) and supplemental grant for Epi25 phenotyping: "Epi25 Clinical Phenotyping R03". National Institutes of Health (1R03NS108145-01). Lowenstein DH (PI), Berkovic SF. Effort at **AUSRMB** was supported by the the RMH Neuroscience Foundation, NHMRC Project Grant (#1059858) and NHMRC Program Grant (#1091593). S.W. r from the **BELATW** cohort eceives funding support from FWO-FKM (1861419N) and GSKE. Work conducted at **CANUTN** was supported by funding from Ontario Brain Institute–EpLink and McLaughlin Foundation. All individuals from **DEUPUM** were phenotyped in detail by epilepsy specialists (KMK, FR, SK, PSR) within the EpimiRNA project (European Union's 'Seventh Framework' Programme (FP7) under Grant Agreement no. 602130). KMK, FR and PSR were supported by the Detlev-Wrobel-Fonds for Epilepsy Research Frankfurt. The **DEUUGS** cohort is supported by DFG 416/5-1. The **DEUUKB** cohort is supported by grants from the DFG - Deutsche Forschungsgemeinschaft (KU 911/21-2 and KU 911/22-1 to W.S.K.; ZS 99/3-2 and ZS 99/4-1 to GZ). Funding for the **DEUUTB** cohort was obtained from the German Research Foundation (DFG) and the Fond Nationale de la Recherche (FNR) in Luxembourg in the frame of the Research Unit FOR-2715 (grants Le1030/14-1, Le1030/16-1, Kr5093/2-1, We4896/4-1, He5415/7-1), an additional DFG grant (DFG: WE4896/3-1), by the rare disease program of the Federal Ministry of Education and Research (BMBF, Treat-ION, grant 01GM1907), by the German Society for Epileptology, and by the foundation no-epilep. Effort at **CZEMTH** was supported by Ministry of Health of Czech Republic AZV 15–33041 and DRO 00064203. **FINKPH** is funded by Saastamoinen Foundation, Vaajasalo Foundation and Governmental Funding. **FINUVH** is funded by Foundation for Pediatric Research, Finland and Folkhälsan Research Foundation. The **FRALYU** site (REPO2MSE study) was funded by the French Ministry of Health (PHRC National 2009). The Swansea Neurology Biobank (**GBRSWU**) is hosted / sponsored by Swansea University and is supported by the Wales Gene Park and BRAIN (Brain Repair and Intracranial Neurotherapeutics) which are funded by Health & Care Research Wales. The **GBRUCL** cohort thanks the Epilepsy Society, Muir Maxwell Trust and Wellcome Trust Strategic Award (WT104033AIA) for funding support. The work conducted at GBRUCL was undertaken at University College London Hospitals, which received a

proportion of funding fromthe NIHR Biomedical Research Centres funding scheme. Assembly of the **GBRUNL** cohort was supported by generous funding from The Wellcome Trust, the Imperial College NIHR Biomedical Research Centre, the Department of Health (UK), the Medical Research Council (UK), and the National Institute of Health Research (UK). **HKGHKK** is supported by the Research Grants Council of the Hong Kong Special Administrative Region, China (project number CUHK4466/06M). **HRVUZG** acknowledges Sarah Weckhuysen and Hannah Stamberger (for help with sample shipping). **IRLRCI** Funded by research grants from Science Foundation Ireland (SFI) (16/RC/3948 and 13/CDA/2223) and co-funded under the European Regional Development Fund and by FutureNeuro industry partners. The **ITAIGI** cohort thanks the Genetic Commission of Italian League Against Epilepsy (LICE); this work was developed within the framework of the DINOGMI Department of Excellence of MIUR 2018-2022 (legge 232 del 2016). **ITAUBG** is supported by FIE, Federazione Italiana Epilessie. **LEBABM** site was funded in part by the Lebanese National Council for Scientific Research Award number: LCR 114110 522214 and American University of Beirut (Grant number: 113161352480 115372099990000). The **NZLUTO** site was supported by Health Research Council of New Zealand; Cure Kids. **TWNCGM** site was supported by Chang Gung Memorial Hospital; Ministry of Science and Technology, Taiwan; National Heath Research Institute, Taiwan. **USACHP** and **USACRW** are supported by NIH funding (R01NS493060, 2001-2007 RJ Buono PI and R01NS06415401, 2009-2012 RJ Buono and H Hakonarson Co- PI). U01HG006830 (PI Hakonarson) and The Children’s Hospital of Philadelphia Endowed Chair in Genomic Research (PI Hakonarson).

Supplemental References

1. Epi4K consortium; Epilepsy Phenome/Genome Project. (2017). Ultra-rare genetic variation in common epilepsies: a case-control sequencing study. Lancet Neurol16, 135-143.
2. Epi25 Collaborative. Electronic address: s.berkovic@unimelb.edu.au; Epi25 Collaborative. (2019). Ultra-Rare Genetic Variation in the Epilepsies: A Whole-Exome Sequencing Study of 17,606 Individuals. Am J Hum Genet105, 267-282.
3. Petrovski, S., Szoeke, C.E., Sheffield, L.J., D'Souza, W., Huggins, R.M., and O'Brien T, J. (2009). Multi-SNP pharmacogenomic classifier is superior to single-SNP models for predicting drug outcome in complex diseases. Pharmacogenet Genomics 19, 147-152.
4. Szoeke, C., Sills, G.J., Kwan, P., Petrovski, S., Newton, M., Hitiris, N., Baum, L., Berkovic, S.F., Brodie, M.J., Sheffield, L.J., et al. (2009). Multidrug-resistant genotype (ABCB1) and seizure recurrence in newly treated epilepsy: data from international pharmacogenetic cohorts. Epilepsia 50, 1689-1696.
5. Petrovski, S., Scheffer, I.E., Sisodiya, S.M., O'Brien, T.J., Berkovic, S.F., and Consortium, E. (2009). Lack of replication of association between scn1a SNP and febrile seizures. Neurology 73, 1928-1930.
6. Shazadi, K., Petrovski, S., Roten, A., Miller, H., Huggins, R.M., Brodie, M.J., Pirmohamed, M., Johnson, M.R., Marson, A.G., O'Brien, T.J., et al. (2014). Validation of a multigenic model to predict seizure control in newly treated epilepsy. Epilepsy Res 108, 1797-1805.
7. Speed, D., Hoggart, C., Petrovski, S., Tachmazidou, I., Coffey, A., Jorgensen, A., Eleftherohorinou, H., De Iorio, M., Todaro, M., De, T., et al. (2014). A genome-wide association study and biological pathway analysis of epilepsy prognosis in a prospective cohort of newly treated epilepsy. Hum Mol Genet 23, 247-258.
8. Petrovski, S., Szoeke, C.E., Jones, N.C., Salzberg, M.R., Sheffield, L.J., Huggins, R.M., and O'Brien, T.J. (2010). Neuropsychiatric symptomatology predicts seizure recurrence in newly treated patients. Neurology 75, 1015-1021.
9. Conte, F., Legros, B., Van Paesschen, W., Avbersek, A., Muglia, P., and Depondt, C. (2018). Long-term seizure outcomes in patients with drug resistant epilepsy. Seizure 62, 74-78.
10. Rezazadeh, A., Borlot, F., Faghfoury, H., and Andrade, D.M. (2017). Genetic generalized epilepsy in three siblings with 8q21.13-q22.2 duplication. Seizure 48, 57-61.
11. Hamdan, F.F., Myers, C.T., Cossette, P., Lemay, P., Spiegelman, D., Laporte, A.D., Nassif, C., Diallo, O., Monlong, J., Cadieux-Dion, M., et al. (2017). High Rate of Recurrent De Novo Mutations in Developmental and Epileptic Encephalopathies. Am J Hum Genet 101, 664-685.
12. Borlot, F., Regan, B.M., Bassett, A.S., Stavropoulos, D.J., and Andrade, D.M. (2017). Prevalence of Pathogenic Copy Number Variation in Adults With Pediatric-Onset Epilepsy and Intellectual Disability. JAMA Neurol 74, 1301-1311.
13. Aljaafari, D., Fasano, A., Nascimento, F.A., Lang, A.E., and Andrade, D.M. (2017). Adult motor phenotype differentiates Dravet syndrome from Lennox-Gastaut syndrome and links SCN1A to early onset parkinsonian features. Epilepsia 58, e44-e48.
14. Marques, P.T., Zulfiqar Ali, Q., Selvarajah, A., Faghfoury, H., Wennberg, R.A., and Andrade, D.M. (2020). Hyperammonemic Encephalopathy Associated with Perampanel: Case Report and Discussion. Can J Neurol Sci, 1-2.
15. Bobbili, D.R., Lal, D., May, P., Reinthaler, E.M., Jabbari, K., Thiele, H., Nothnagel, M., Jurkowski, W., Feucht, M., Nurnberg, P., et al. (2018). Exome-wide analysis of mutational burden in patients with typical and atypical Rolandic epilepsy. Eur J Hum Genet 26, 258-264.
16. Miller, S.A., Dykes, D.D., and Polesky, H.F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16, 1215.
17. Lal, D., Ruppert, A.K., Trucks, H., Schulz, H., de Kovel, C.G., Kasteleijn-Nolst Trenite, D., Sonsma, A.C., Koeleman, B.P., Lindhout, D., Weber, Y.G., et al. (2015). Burden analysis of rare microdeletions suggests a strong impact of neurodevelopmental genes in genetic generalised epilepsies. PLoS Genet 11, e1005226.
18. McCormack, M., Gui, H., Ingason, A., Speed, D., Wright, G.E.B., Zhang, E.J., Secolin, R., Yasuda, C., Kwok, M., Wolking, S., et al. (2018). Genetic variation in CFH predicts phenytoin-induced maculopapular exanthema in European-descent patients. Neurology 90, e332-e341.
19. Brainstorm, C., Anttila, V., Bulik-Sullivan, B., Finucane, H.K., Walters, R.K., Bras, J., Duncan, L., Escott-Price, V., Falcone, G.J., Gormley, P., et al. (2018). Analysis of shared heritability in common disorders of the brain. Science 360.
20. International League Against Epilepsy Consortium on Complex Epilepsies. Electronic address, e.-a.u.e.a. (2014). Genetic determinants of common epilepsies: a meta-analysis of genome-wide association studies. Lancet Neurol 13, 893-903.
21. Rheims, S., Alvarez, B.M., Alexandre, V., Curot, J., Maillard, L., Bartolomei, F., Derambure, P., Hirsch, E., Michel, V., Chassoux, F., et al. (2019). Hypoxemia following generalized convulsive seizures: Risk factors and effect of oxygen therapy. Neurology 92, e183-e193.

22. Alexandre, V., Mercedes, B., Valton, L., Maillard, L., Bartolomei, F., Szurhaj, W., Hirsch, E., Marchal, C., Chassoux, F., Petit, J., et al. (2015). Risk factors of postictal generalized EEG suppression in generalized convulsive seizures. Neurology 85, 1598-1603.
23. Leschziner, G., Jorgensen, A.L., Andrew, T., Pirmohamed, M., Williamson, P.R., Marson, A.G., Coffey, A.J., Middleditch, C., Rogers, J., Bentley, D.R., et al. (2006). Clinical factors and ABCB1 polymorphisms in prediction of antiepileptic drug response: a prospective cohort study. Lancet Neurol 5, 668-676.
24. Guo, Y., Baum, L.W., Sham, P.C., Wong, V., Ng, P.W., Lui, C.H., Sin, N.C., Tsoi, T.H., Tang, C.S., Kwan, J.S., et al. (2012). Two-stage genome-wide association study identifies variants in CAMSAP1L1 as susceptibility loci for epilepsy in Chinese. Hum Mol Genet 21, 1184-1189.
25. Muir, A.M., Myers, C.T., Nguyen, N.T., Saykally, J., Craiu, D., De Jonghe, P., Helbig, I., Hoffman-Zacharska, D., Guerrini, R., Lehesjoki, A.E., et al. (2019). Genetic heterogeneity in infantile spasms. Epilepsy Res 156, 106181.
26. Leu, C., Stevelink, R., Smith, A.W., Goleva, S.B., Kanai, M., Ferguson, L., Campbell, C., Kamatani, Y., Okada, Y., Sisodiya, S.M., et al. (2019). Polygenic burden in focal and generalized epilepsies. Brain 142, 3473-3481.
27. Wolking, S., Moreau, C., Nies, A.T., Schaeffeler, E., McCormack, M., Auce, P., Avbersek, A., Becker, F., Krenn, M., Moller, R.S., et al. (2020). Testing association of rare genetic variants with resistance to three common antiseizure medications. Epilepsia 61, 657-666.
28. Accogli, A., Severino, M., Riva, A., Madia, F., Balagura, G., Iacomino, M., Carlini, B., Baldassari, S., Giacomini, T., Croci, C., et al. (2020). Targeted re-sequencing in malformations of cortical development: genotype-phenotype correlations. Seizure 80, 145-152.
29. Licchetta, L., Bisulli, F., Vignatelli, L., Zenesini, C., Di Vito, L., Mostacci, B., Rinaldi, C., Trippi, I., Naldi, I., Plazzi, G., et al. (2017). Sleep-related hypermotor epilepsy: Long-term outcome in a large cohort. Neurology 88, 70-77.
30. Bisulli, F., Menghi, V., Vignatelli, L., Licchetta, L., Zenesini, C., Stipa, C., Morigi, F., Gizzi, M., Avoni, P., Provini, F., et al. (2018). Epilepsy with auditory features: Long-term outcome and predictors of terminal remission. Epilepsia 59, 834-843.
31. Nasreddine, W., Fakhredin, M., Makke, Y., Hmaimess, G., Sabbagh, S., Beaini, S., El Tourjuman, O., and Beydoun, A. (2020). Hyperventilation-induced high-amplitude rhythmic slowing: A mimicker of absence seizures in children. Epilepsy Behav 103, 106510.
32. Arabi, M., Dirani, M., Hourani, R., Nasreddine, W., Wazne, J., Atweh, S., Samara, H., Shatila, A.R., and Beydoun, A. (2018). Frequency and Stratification of Epileptogenic Lesions in Elderly With New Onset Seizures. Front Neurol 9, 995.
33. Smith, L.A., Ullmann, J.F., Olson, H.E., Achkar, C.M., Truglio, G., Kelly, M., Rosen-Sheidley, B., and Poduri, A. (2017). A Model Program for Translational Medicine in Epilepsy Genetics. J Child Neurol 32, 429-436.
34. Glauser, T.A., Holland, K., O'Brien, V.P., Keddache, M., Martin, L.J., Clark, P.O., Cnaan, A., Dlugos, D., Hirtz, D.G., Shinnar, S., et al. (2017). Pharmacogenetics of antiepileptic drug efficacy in childhood absence epilepsy. Ann Neurol 81, 444-453.
35. Buono, R.J., Lohoff, F.W., Sander, T., Sperling, M.R., O'Connor, M.J., Dlugos, D.J., Ryan, S.G., Golden, G.T., Zhao, H., Scattergood, T.M., et al. (2004). Association between variation in the human KCNJ10 potassium ion channel gene and seizure susceptibility. Epilepsy Res 58, 175-183.
36. EPGP Collaborative, Abou-Khalil B, Alldredge B, Bautista J, Berkovic S, Bluvstein J, Boro A, Cascino G, Consalvo D, Cristofaro S, et al. (2013). The epilepsy phenome/genome project. Clin Trials10, 568-586.
37. Meador, K.J., Pennell, P.B., May, R.C., Gerard, E., Kalayjian, L., Velez-Ruiz, N., Penovich, P., Cavitt, J., French, J., Hwang, S., et al. (2018). Changes in antiepileptic drug-prescribing patterns in pregnant women with epilepsy. Epilepsy Behav 84, 10-14.
38. Greenberg, D.A., Durner, M., Keddache, M., Shinnar, S., Resor, S.R., Moshe, S.L., Rosenbaum, D., Cohen, J., Harden, C., Kang, H., et al. (2000). Reproducibility and complications in gene searches: linkage on chromosome 6, heterogeneity, association, and maternal inheritance in juvenile myoclonic epilepsy. Am J Hum Genet 66, 508-516.