Supplementary Material to the Article

Koko M, Motelow JE, Stanley KE, Bobbili DR, Dhindsa RS, May P; Canadian Epilepsy Network; Epi4K Consortium; Epilepsy Phenome/Genome Project, EpiPGX Consortium; EuroEPINOMICS-CoGIE Consortium. Association of ultra-rare coding variants with genetic generalized epilepsy: a case-control whole exome sequencing study. *Epilepsia.* 2022. https://doi.org/10.1111/epi.17166

Table of Contents

SUPPLEMENTARY METHODS: 1

THE FIRST DATASET: 1

Samples: *1*

Sequence data generation: *1*

Sample quality control: *1*

Variant quality control and call rate harmonization: *2*

THE SECOND DATASET: 2

Samples: *2*

Sequence data generation: *2*

Sample quality control: *3*

Variant quality control and call rate harmonization: *4*

DUPLICATES AND ANCESTRY HARMONIZATION ACROSS COHORTS: 4

Ancestry matching between the case cohorts: *4*

Duplicates between the two case cohorts: *5*

OVERLAP WITH THE EPI25 COLLABORATIVE DATASETS: 5

QUALIFYING VARIANTS' DISTRIBUTION PLOTS: 5

QUANTILE-QUANTILE (QQ) PLOTS: 5

GENE SETS FOR GENE SET ASSOCIATION ANALYSIS: 6

Overrepresentation of gene sets among top ranked genes: *6*

REFERENCES: 7

SUPPLEMENTARY TABLES: 9

TABLE S1: NUMBERS OF ANALYZED SAMPLES FROM THE STUDY COHORTS. 9

TABLE S2: *GABRG2* VARIANTS IDENTIFIED IN INDIVIDUALS DIAGNOSED WITH GENETIC GENERALIZED EPILEPSY 10

TABLE S3: *GABRG2* VARIANTS IN INDIVIDUALS WITHOUT EPILEPSY (CONTROLS). 11

TABLE S4: TOP-RANKED GENES IN THE SECONDARY ANALYSES OF RARE FUNCTIONAL VARIATION. 12

TABLE S5: TOP-RANKED GENES IN THE SECONDARY ANALYSES OF PREDICTED LOSS OF FUNCTION (PLOF) VARIANTS. 13

TABLE S6: COMPARISONS OF TOP-RANKED GENES WITH THREE PREVIOUS LARGE-SCALE RARE VARIANT ASSOCIATION STUDIES OF GENETIC GENERALIZED EPILEPSY. 14
TABLE S7: ASSOCIATION OF GENES ENCODING $GABA_A$ RECEPTORS. 17
TABLE S8: ASSOCIATION OF OMIM GENES IMPLICATED IN SUSCEPTIBILITY TO GENERALIZED EPILEPSY....18
TABLE S9: ASSOCIATION OF OMIM GENES IMPLICATED AUTOSOMAL DOMINANT DEVELOPMENTAL AND EPILEPTIC ENCEPHALOPATHIES....20
TABLE S10: PPH2, REVEL AND MTR SCORES OF EPILEPSY RELATED *GABRG2* MISSENSE VARIANTS. 22
SUPPLEMENTARY FIGURES:....24
FIG. S1: FLOW CHART SUMMARIZING THE ANALYSIS STRATEGY USED IN THIS STUDY 24
FIG. S2: PRINCIPAL COMPONENT ANALYSIS FOR ANCESTRY MATCHING....25
FIG. S3: BALANCE OF ULTRA-RARE SYNONYMOUS QUALIFYING VARIANTS TALLIES BETWEEN CASES AND CONTROLS IN THE FIRST DATASET. 26
FIG. S4: BALANCE OF ULTRA-RARE SYNONYMOUS QUALIFYING VARIANTS TALLIES BETWEEN CASES AND CONTROLS IN THE SECOND DATASET. 27
FIG. S5: BALANCE OF ULTRA-RARE SYNONYMOUS QUALIFYING VARIANTS TALLIES BETWEEN CASES AND CONTROLS IN THE TOTAL DATASET....28
FIG. S6: ASSOCIATION OF ULTRA-RARE DELETERIOUS VARIANTS WITH GENETIC GENERALIZED EPILEPSY. 29
FIG. S7: ASSOCIATION OF ULTRA-RARE DELETERIOUS AND INTOLERANT VARIANTS WITH GENETIC GENERALIZED EPILEPSY....30
FIG. S8: ASSOCIATION OF RARE DELETERIOUS VARIANTS WITH GENETIC GENERALIZED EPILEPSY. 31
FIG. S9A: ASSOCIATION OF RARE PREDICTED LOSS OF FUNCTION VARIANTS (INCLUDING URVS) WITH GENETIC GENERALIZED EPILEPSY....32
FIG. S9B: ASSOCIATION OF RARE PREDICTED LOSS OF FUNCTION VARIANTS (EXCLUDING URVS) WITH GENETIC GENERALIZED EPILEPSY....33
FIG. S10: PEDIGREES OF FOUR FAMILIES WITH GENETIC GENERALIZED EPILEPSY 34
ACKNOWLEDGEMENTS....35
AFFILIATIONS....37
ADDITIONAL DATA 44

Supplementary Methods:

The first dataset:

Samples:

Participants whose sequencing data formed the first dataset included 1,214 GGE patients recruited by the Epi4K Consortium and Epilepsy Phenome/Genome Project as previously described,[1,2] and sequenced at the Institute for Genomic Medicine (IGM) at Columbia University (New York, USA). The diagnosis of a GGE syndrome required the patients to have generalized epilepsy with absence, myoclonic or tonic-clonic seizures and generalized spike-and-wave discharge on electroencephalography (EEG). To qualify for the familial analysis, patients were required to have at least one relative (up to the third degree) who had been diagnosed with epilepsy.[1] Ancestry matched controls (n = 14,100 before quality control) were selected from multiple collections of control cohorts at the IGM.[3]

Sequence data generation:

Whole exome sequencing of DNA samples from participants forming the first dataset was performed at IGM using Illumina's HiSeq 2000, HiSeq 2500 or NovaSeq 6000 platforms (Illumina, San Diego, CA, USA) following enrichment with Agilent All Exon Enrichment kits (Agilent Technologies, Santa Clara, CA, USA), NimbleGen SeqCap EZ Exome Enrichment kit (Roche NimbleGen, Madison, WI, USA), Twist Human Core Exome (Twist Bioscience, San Francisco, CA, USA) or IDT xGen Exome Research Panel (Integrated DNA Technologies, Coralville, IA, USA). The sequence data from all cases and controls were processed according to the IGM bioinformatics pipeline.[1,4] Sequencing reads were aligned to the human reference genome build 37 (GRCh37) using Illumina's Dynamic Read Analysis for GENomics (DRAGEN) Bio-IT Platform.[5,6] Picard (https://broadinstitute.github.io/picard/) and the Genome Analysis Tool Kit v3.6 (GATK) were used to perform duplicate read marking, base quality scores recalibration, indel realignment and haplotype calling. The samples were processed individually at different time points and the variants obtained from single sample calling were imported and integrated in the Analysis Tool for Annotated Variants (ATAV) Database.[4]

Sample quality control:

For the purpose of this study, samples with possible contamination (heterozygosity exceeding 2%) determined using VerifyBamID,[7] with discordance between self-declared and sequence-derived sex, or with low coverage (less than 85% of the consensus coding sequence release 20 (CCDS20) targets covered at a minimum of 10x) were removed. Related individuals were identified using Kinship-based Inference for GWAS[8] (KING). One of each pair that had an inferred relationship of third-degree or

closer was dropped, preferentially retaining affected over control individuals and samples with higher coverage for pairs with similar disease status. EIGENSTRAT[9] was then used to remove ethnicity outliers to minimize the effects of residual population stratification (Fig. S2).

Variant quality control and call rate harmonization:

The following variant-level parameters (hard filters) were enforced: Quality/Depth (QD) > 5, Quality (QUAL) > 50, Mapping Quality (MQ) > 40, SOR < 3 (SNVs) or < 10 (indels), Fisher's Strand bias score (FS) < 60 (SNVs) or < 200 (indels), Read Position Rank Sum score (RPRS) < -3, and Mapping Quality Rank Sum score (MQRS) < -10. Variants were required to pass GATK Variant Quality Score Recalibration (VQSR) filter. Known artifacts and variants failing quality filters in population databases (Exome Variant Server, ExAC, gnomAD) were excluded. Low quality genotype calls with total allelic depth (DP) < 10 or genotype quality (GQ) < 20 were filtered. Heterozygous calls had a minimum alternate allele fraction (AD/DP) of 0.3. As previously described, a coverage harmonization procedure was employed to remove the variants that are differentially covered across the cases and controls.[1] Briefly, this was based on plotted the cumulative difference in site coverage between cases and controls to identify a filtering cut-off that will minimize this difference while allowing the largest possible number of variants to be retained.

The second dataset:

Samples:

The individuals with generalized epilepsy analyzed here were selected from 2,524 individuals recruited by the EuroEPINOMICS-CoGIE Consortia, EpiPGX Consortium, and CENet as described previously.[10] For the purpose of this work, we used the sequence data of 989 individuals ascertained to have classical GGE phenotypes (childhood or juvenile absence epilepsy, juvenile myoclonic epilepsy, or epilepsy with generalized tonic-clonic seizures alone), early-onset absence epilepsy (age of onset < 3 years), epilepsy with myoclonic absences, or unclassified GGE. Familial cases had one or more, self-reported, first- or second-degree, affected relative. The controls for this dataset (n = 4,904 before quality control) were obtained from the database of Genotypes and Phenotypes[11] (dbGAP studies: MIGen Ottawa Heart Study controls, Rotterdam study controls and Alzheimer Disease Genetics Study controls) or from the Epi25 Collaborative.[12]

Sequence data generation:

Whole-exome sequencing of EuroEPINOMICS-CoGIE cases was done on Illumina's HiSeq 2000 platform using NimbleGen SeqCap EZ Human Exome Library (NimbleGen, Madison, WI, USA) at

Cologne Center for Genomics (Cologne, Germany). Whole-exome sequencing for the EpiPGX cohort was done at deCODE genetics (Reykjavik, Iceland) on Illumina's HiSeq 2500 platform with Nextera Rapid Capture Expanded Exome kit (Illumina, San Diego, CA, USA). Whole exome sequencing of individuals recruited by the Canadian Epilepsy Network (CENet) was performed by the McGill University and Génome Québec Innovation Center (MAGQUIC, Québec, Canada) on Illumina's HiSeq sequencing platforms using TruSeq or Roche Nimblegen EZ libraries. Controls from the Epi25 Collaborative were sequenced at the Broad Institute of Harvard and the Massachusetts Institute of Technology on Illumina's HiSeq platform using Illumina's Nextera Rapid Capture or TruSeq Rapid Exome enrichment kits. The Rotterdam Study controls were sequenced on Illumina's HiSeq 2000 platform using EZ Human Exome Library. The Alzheimer study controls were sequenced over multiple time points at the Broad Institute using different capture kits. Fastq files were aligned to GRCh37 as previously described[10,13] and jointly called using DRAGEN Bio-IT Platform.[5,6]

Sample quality control:

The sample-level call rate, autosomal and chrX inbreeding coefficients were collected using Plink[14] v1.9. and Picard (from GATK[15] v4.1.4.1). Samples with phenotypes other than GGEs or without appropriate permissions for inclusion and samples with extremely low variant counts (< 10,000 non-missing calls) were removed. Samples with genotyping rates lower than 80%, outlier samples on autosomal heterozygosity (> 4 median absolute deviations on autosomal inbreeding coefficient estimates), and samples with discordant or ambiguous sequencing sex based on chromosome X inbreeding co-efficient estimates ($F < 0.3$ for female and $F > 0.7$ for male predicted sequencing sex) were excluded. The remaining samples were scanned for relatedness (third degree) using KING.[8] For duplicates and pairs with matching phenotypes, the sample with the higher genotyping rate was retained. Otherwise, cases were preferentially retained. Next, multi-dimensional scaling (MDS) was used to project the major continental ancestry of the study samples on the MDS space of 1000 Genomes data (2,504 samples) using KING. The top principal components were visualized and used to classify the ancestry with a support vector machine using R package *e1071*.[16] Samples with predicted European ancestry were retained. Also, samples clustering with Finnish 1000 Genomes samples on PC1/2 were filtered. Following the baseline variant filtering steps outlined below, the variant calling metrics were re-examined to exclude any additional sample outliers. Here, all samples with SNV counts < 15,000 were filtered (this removed all Rotterdam Study controls and most Alzheimer Study controls). Outliers beyond 3 standard deviations per cohort on key variant calling metrics (Heterozygous-Homozygous calls ratio, Transitions-Transversions ratio, and Insertions-Deletions ratio) were filtered. To ensure adequate case control matching and the removal of ancestry outliers, PCA analysis using EIGENSTRAT[9] was employed (Fig. S2).

Variant quality control and call rate harmonization:

The variants were filtered for those located in the CCDS exonic coding regions (padded on each side to accommodate canonical splice sites and masked for low-complexity regions) using bcftools[17] v1.9. The variants were decomposed, normalized and sorted using bcftools and vt[18] v0.5. Low quality genotypes were filtered by setting calls with total allelic depth < 10 or genotype quality < 20 to missing. Heterozygous calls had a minimum alternate allele depth fraction (AD/DP) of 0.25. This genotype filtering was performed using bcftools. A combination of hard filtering and filtering based on recalibrated variant quality scores was employed to remove low quality variants. Variant calls with low quality were filtered (SNVs: QUAL < 10, QD < 2, MQ < 30, FS > 60, MQRankSum < -12.5, RPRS < -8; Indels: QUAL < 10, QD < 2, RPRS < -20, FS > 200). Variant Quality Score Recalibration (VQSR) was performed on the normalized and genotype-filtered call set using GATK based on these annotations: QD, FS, SOR, MQRankSum, and RPRS. SNVs and Indels failing VQSR Tranche 99.0 filter were removed. Since the datasets were sequenced using different capture kits, we performed additional harmonization steps to limit our analysis to the coding regions covered in all kits & to minimize the spurious effects caused differences in capture kits. Variants were retained only if they had genotyping rates ≥ 90% both in EpiPGX cases (largest case dataset; representing Illumina capture targets) and MIGen Ottawa controls (largest controls dataset; representing Agilent capture targets). After removal of sample outliers (see above), a final round of call rate harmonization was then performed where the variant call rate was calculated among the remaining cases and controls and variants with call rates below 95% in cases or controls were filtered. Also, the cumulative difference of call rate between cases and controls was plotted and 9.4% of the variants were removed to minimize this difference while retaining the largest possible number of variants.

Duplicates and ancestry harmonization across cohorts:

Ancestry matching between the case cohorts:

To maximize the ancestry matching between the two analyzed patient cohorts, the ancestry prediction among the cases was harmonized in our two cases datasets by using the same ancestry prediction model to ensure homogeneity in ancestry assignment. Principal components analysis was performed on genotypes of previously defined well covered exonic autosomal polymorphic markers.[1] A neural network model that uses the first five principal component axes as the independent variables, trained on more than two thousand individuals with pre-evaluated genetic ancestry from six ethnic groups (European, Middle Eastern, Hispanic, East Asian, South Asian, and African), was then used to predict the probability of a European ancestry. Cases with < 95% probability were excluded.

Duplicates between the two case cohorts:

To exclude likely duplicates between the two case cohorts, a genotype hashing approach adopted from the Gencrypt method[19] was used to avoid the need for genotype sharing across the two study sites. A group of variants with minor allele frequency > 0.1 and genotyping rate > 98% in both cohorts was identified. From this pool, 200 sets were created, each consisting of randomly selected non-overlapping 150 SNPs. For each sample, the genotypes over each set were concatenated keeping their order and converted to *sha256* cryptographic hashes. The hashes were exchanged and compared between cohorts. In total, 57 cases shared one or more hashes (according, likely to have identical genotypes in ≥ 150 randomly selected polymorphic markers) were considered possible duplicates. These were retained only in the first dataset and were removed from the second set.

Overlap with the Epi25 Collaborative datasets:

There is marginal overlap between the samples of individuals with GGE and the published analyses of the Epi25 Collaborative (before QC: less than 100 samples from the EuroEUROEPINOMICS-CoGIE and EpiPGX dataset). The controls used for the first dataset were similar to those used in the Epi25 Collaborative Y1-3 analysis (internal control datasets of the Institute for Genomics Medicine, NY). The controls used for the second dataset had substantial overlap with the Epi25 Collaborative Y1-2 analysis (978 individuals from MIGen Ottawa study and 332 individuals from Epi25 controls).

Qualifying Variants' distribution plots:

To ensure that we achieved an adequate case control matching and coverage/call rate harmonization in each dataset, we examined the distribution plots of qualifying variants tallies. Variant tallies were examined separately for each study dataset and collectively for the final merged dataset. The significance of the differences in the distribution density of ultra-rare synonymous variants was examined using Wilcoxon Rank Sum test with continuity correction as implemented in R (Fig. S3 – S5).

Quantile-Quantile (QQ) plots:

To obtain cohort-level QQ plots, the case-control labels were shuffled before running 1,000 iterations of gene collapsing analyses using Fisher's exact test. *P* values were ranked from these permutations and the negative $\log_{10}$ of the mean *p* values at each rank was plotted against the negative $\log_{10}$ of observed *p* values. Confidence intervals were obtained by plotting negative $\log_{10}$ of the expected *p* values corresponding the 2.5^{th} and 97^{th} percentiles from the permutations. The genomic inflation factors (λ) were calculated according to the previously described regression method implemented in the function *estlambda2()* from R package *QQperm*.[20]

The QQ plots for the combined analysis (Cochran-Mantel-Haenszel exact test) show the *p* values from those genes with at least one qualifying variant in the joint case cohort and the expected *p* values from a uniform distribution. The negative $\log_{10}$ of the observed *p* values was plotted against the negative $\log_{10}$ of an equal number of uniformly distributed *p* values ($-log_{10}((k-0.5)/n)$, where *k* is the gene rank and *n* the total genes). The confidence intervals for the expected *p* values were based on values drawn from a beta distribution ($-log_{10}(qbeta(\alpha/2, k, n-k$ and $-log_{10}(qbeta(1-\alpha/2, k, n-k)$, where $\alpha = 0.05$ for a 95% confidence interval) using the *stats* package[21] in R 3.3. The genomic inflation factors (λ) was calculated using *QQperm*.

Gene sets for gene set association analysis:

The list of genes composing the two gene sets of inhibitory signaling tested in this study were obtained from previously published work.[10,12]

Genes encoding $GABA_A$ receptors: *GABRA1, GABRA2, GABRA3, GABRA4, GABRA5, GABRA6, GABRB1, GABRB2, GABRB3, GABRD, GABRE, GABRG1, GABRG2, GABRG3, GABRP, GABRQ, GABRR1, GABRR2, GABRR3*.

GABAergic pathway genes: *ABAT, ADCY1, ADCY2, ADCY3, ADCY4, ADCY5, ADCY6, ADCY7, ADCY8, ADCY9, ANK2, ANK3, ARHGEF9, DISC1, DLC1, DNAI1, FGF13, GABARAP, GABARAPL1, GABARAPL2, GABBR1, GABBR2, GABRA1, GABRA2, GABRA3, GABRA4, GABRA5, GABRA6, GABRB1, GABRB2, GABRB3, GABRD, GABRE, GABRG1, GABRG2, GABRG3, GABRP, GABRQ, GABRR1, GABRR2, GABRR3, GAD1, GAD2, GLS, GLS2, GLUL, GNAI1, GNAI2, GNAI3, GNAO1, GNB1, GNB2, GNB3, GNB4, GNB5, GNG10, GNG11, GNG12, GNG13, GNG2, GNG3, GNG4, GNG5, GNG7, GNG8, GNGT1, GNGT2, GPHN, HAP1, KCNB2, KCNC1, KCNC2, KCNC3, KCNJ6, KIF5A, KIF5B, KIF5C, MAGI1, MKLN1, MTOR, MYO5A, NLGN2, NRXN1, NSF, PFN1, PLCL1, PRKACA, PRKACB, PRKACG, PRKCA, PRKCB, PRKCG, RDX, SCN1A, SCN1B, SCN2B, SCN3A, SCN8A, SEMA4D, SLC12A2, SLC12A5, SLC32A1, SLC38A1, SLC38A2, SLC38A3, SLC38A5, SLC6A1, SLC6A11, SLC6A13, SRC, STARD13, TRAK1, TRAK2*.

Overrepresentation of gene sets among top ranked genes:

A hypergeometric test was employed to examine the probability that *n* genes from a gene set of *N* genes appeared by chance among the top-ranked *k* genes when examining a total of 18,834 protein coding genes. The enrichment was tested at each rank *k* occupied by a gene from the gene set using a using *phyper* function from R *stats* package as follows: *phyper*(*n*-1, *N*, 18834-*N*, *k*, lower.tail= FALSE). This was limited to those genes with nominal *p* values < 0.05. As the change in the direction of effect was the main outcome we intended to investigate, the outcomes from these secondary analyses were not corrected for multiple testing. The Online Mendelian Inheritance in Man (OMIM, https://www.omim.org/) database was used to obtain a list of genes associated with susceptibility to

GGE (IGE, CAE & JME; phenotypic series: PS600669, PS254770 and PS600131) or causing Developmental and Epileptic Encephalopathies (DEE; phenotypic series: PS308350).

References:

1. Epi4K Consortium, Epilepsy Phenome/Genome Project. Ultra-rare genetic variation in common epilepsies: a case-control sequencing study. The Lancet Neurology. 2017;16(2):135–43. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1474442216303593
2. Epi4K Consortium. Epi4K: Gene discovery in 4,000 genomes. Epilepsia. 2012;53(8):1457–67. Available from: https://onlinelibrary.wiley.com/doi/10.1111/j.1528-1167.2012.03511.x
3. Epi25 Collaborative. Sub-genic intolerance, ClinVar, and the epilepsies: A whole-exome sequencing study of 29,165 individuals. The American Journal of Human Genetics. 202;108(6):965–82. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0002929721001403
4. Ren Z, Povysil G, Hostyk JA, Cui H, Bhardwaj N, Goldstein DB. ATAV: a comprehensive platform for population-scale genomic analyses. BMC Bioinformatics. 2021;22(1):149. Available from: https://bmcbioinformatics.biomedcentral.com/articles/10.1186/s12859-021-04071-1
5. Goyal A, Kwon HJ, Lee K, Garg R, Yun SY, Hee Kim Y, et al. Ultra-Fast Next Generation Human Genome Sequencing Data Processing Using DRAGEN™ Bio-IT Processor for Precision Medicine. OJGen. 2017;07(01):9–19. Available from: http://www.scirp.org/journal/doi.aspx?DOI=10.4236/ojgen.2017.71002
6. Zhao S, Agafonov O, Azab A, Stokowy T, Hovig E. Accuracy and efficiency of germline variant calling pipelines for human genome data. Sci Rep. 2020;10(1):20222. Available from: http://www.nature.com/articles/s41598-020-77218-4
7. Jun G, Flickinger M, Hetrick KN, Romm JM, Doheny KF, Abecasis GR, et al. Detecting and Estimating Contamination of Human DNA Samples in Sequencing and Array-Based Genotype Data. The American Journal of Human Genetics. 2012;91(5):839–48. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0002929712004788
8. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W-M. Robust relationship inference in genome-wide association studies. Bioinformatics. 2010;26(22):2867–73. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btq559
9. Patterson N, Price AL, Reich D. Population Structure and Eigenanalysis. PLoS Genet. 2006;2(12):e190. Available from: https://dx.plos.org/10.1371/journal.pgen.0020190
10. May P, Girard S, Harrer M, Bobbili DR, Schubert J, Wolking S, et al. Rare coding variants in genes encoding $GABA_A$ receptors in genetic generalised epilepsies: an exome-based case-control study. The Lancet Neurology. 2018;17(8):699–708. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1474442218302151

11. Tryka KA, Hao L, Sturcke A, Jin Y, Wang ZY, Ziyabari L, et al. NCBI's Database of Genotypes and Phenotypes: dbGaP. Nucl Acids Res. 2014;42(D1):D975–9. Available from: https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkt1211
12. Epi25 Collaborative. Ultra-Rare Genetic Variation in the Epilepsies: A Whole-Exome Sequencing Study of 17,606 Individuals. The American Journal of Human Genetics. 2019;105(2):267–82. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0002929719302071
13. Wolking S, Moreau C, McCormack M, Krause R, Krenn M, EpiPGx Consortium, et al. Assessing the role of rare genetic variants in drug-resistant, non-lesional focal epilepsy. Ann Clin Transl Neurol. 2021;8(7):1376–87. Available from: https://onlinelibrary.wiley.com/doi/10.1002/acn3.51374
14. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaSci. 2015;4(1):7. Available from: https://academic.oup.com/gigascience/article-lookup/doi/10.1186/s13742-015-0047-8
15. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research. 2010;20(9):1297–303. Available from: http://genome.cshlp.org/cgi/doi/10.1101/gr.107524.110
16. Meyer D, Dimitriadou E, Hornik K, Weingessel A, Leisch F [Internet]. e1071 (Formerly: E1071). 2019. Available from: https://CRAN.R-project.org/package=e1071.
17. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078–9. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btp352
18. Tan A, Abecasis GR, Kang HM. Unified representation of genetic variants. Bioinformatics. 2015;31(13):2202–4. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btv112
19. Turchin MC, Hirschhorn JN. Gencrypt: one-way cryptographic hashes to detect overlapping individuals across samples. Bioinformatics. 2012;28(6):886–8. Available from: https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/bts045
20. Petrovski S, Wang Q. QQperm: Permutation Based QQ Plot and Inflation Factor Estimation. 2016. Available from: https://CRAN.R-project.org/package=QQperm
21. R Core Team. R: A language and environment for statistical computing. 2017. Available from: https://www.R-project.org/

Supplementary Tables:

Table S1: Numbers of analyzed samples from the study cohorts.

Datasets		Analysis		
		All GGEs	Positive Family History	Negative Family History
First	Individuals with epilepsy	1,099	629	492
	Controls	6,814	6,862	6,857
Second	Individuals with epilepsy	829	316	513
	Controls	1,764	1,764	1,764
Total	Individuals with epilepsy	1,928	945	1,005
	Controls	8,578	8,626	8,621

Table S2: *GABRG2* variants identified in individuals diagnosed with genetic generalized epilepsy.

Group		Variant				In-silico predictions			GGE sub-syndrome	Previously reported phenotypes
		NM_000816	NM_198904	NM_198903	Location	PPh2	REVEL	MTR		
Familial GGEs	**Ultra-rare**	c.478G>T p.A160S	c.478G>T p.A160S	c.478G>T p.A160S	N-terminus	1	0.42	0.63	-	-
		c.530G>C p.R177P	c.530G>C p.R177P	c.530G>C p.R177P	N-terminus	1	0.69	0.32	EOAE	FS (p.R177G)
		c.595A>G p.M199V	c.595A>G p.M199V	c.595A>G p.M199V	N-terminus	1	0.91	0.56	-	GEFS (p.M199V); NAFE (p.M199V)
		c.639T>A p.Y213*	c.639T>A p.Y213*	c.759T>A p.Y253*	N-terminus	NA	NA	NA	-	-
		c.755T>C p.V252A	c.755T>C p.V252A	c.875T>C p.V292A	N-terminus	0.98	0.35	0.57	-	-
		c.1213G>A p.G405S	c.1237G>A p.G413S	c.1357G>A p.G453S	Cytoplasmic (TM3-TM4)	0.99	0.35	0.74	CAE	-
		c.1324G>T p.D442Y	c.1348G>T p.D450Y	c.1468G>T p.D490Y	Cytoplasmic (TM3-TM4)	1	0.94	0.76	CAE	-
		c.1370A>G p.N457S	c.1394A>G p.N465S	c.1514A>G p.N505S	TM4	1	0.84	0.66	-	-
Sporadic GGEs	**Ultra-rare**	c.259+2T>G IVS2SD	c.259+2T>G IVS2SD	c.259+2T>G IVS2SD	N-terminus	NA	NA	NA	-	-
		c.769+2T>G IVS6SD	c.769+2T>G IVS6SD	c.889+2T>G IVS7SD	N-terminus	NA	NA	NA	CAE	Familial CAE (IVS6SD)

The variants were analyzed up to an external minor allele frequency (MAF) of 0.1%. All those detected in the cases had an internal leave-one-out MAF and external MAF = 0 (i.e., ultra-rare variants). No rare variants were seen in individuals with epilepsy and all ultra-rare variants were found only once. CAE: Childhood Absence Epilepsy. GEFS: Generalized Epilepsy Febrile Seizures. FS: Febrile Seizures. NAFE: Non-Acquired Focal Epilepsy. TM: Transmembrane segment. PPh2, REVEL and MTR scores considered damaging/deleterious/intolerant are underlined.

Table S3: *GABRG2* variants in individuals without epilepsy (controls).

Group		Variant				In-silico predictions		
		NM_000816	NM_198904	NM_198903	Location	PPh2	REVEL	MTR
Controls	Rare	c.173C>A p.T58N	c.173C>A p.T58N	c.173C>A p.T58N	N-terminus	1	0.384	0.91
		c.530G>A p.R177Q	c.530G>A p.R177Q	c.530G>A p.R177Q	N-terminus	0.98	0.466	0.32
		c.691G>A p.D231N	c.691G>A p.D231N	c.811G>A p.D271N	N-terminus	1	0.485	0.73
		c.748G>A p.E250K	c.748G>A p.E250K	c.868G>A p.E290K	N-terminus	0.99	0.456	0.61
		c.1087C>T p.R363W	c.1087C>T p.R363W	c.1207C>T p.R403W	Cytoplasmic (TM3-TM4)	1	0.584	0.81
		c.1113_1115del p.K372del	c.1113_1115del p.K372del	c.1233_c.1235del p.K412del	Cytoplasmic (TM3-TM4)	NA	NA	NA
		c.1148G>A p.R383H	c.1172G>A p.R391H	c.1292G>A p.R431H	Cytoplasmic (TM3-TM4)	1	0.301	1.01
Controls	Ultra-rare	c.172A>G p.T58A	c.172A>G p.T58A	c.172A>G p.T58A	N-terminus	1.00	0.30	0.91
		c.571C>A p.Q191K	c.571C>A p.Q191K	c.571C>A p.Q191K	N-terminus	0.99	0.39	0.42
		NA	c.1130T>C p.L377P	c.1250T>C p.L417P	Cytoplasmic (TM3-TM4)	0.91	0.44	0.92
		c.1309C>T p.R437C	c.1333C>T p.R445C	c.1453C>T p.R485C	Cytoplasmic (TM3-TM4)	1.00	0.71	0.85

The variants were analyzed up to an external minor allele frequency (MAF) of 0.1%. Ultra-rare variants had an internal leave-one-out MAF and external MAF = 0. Two variants (p.T58A, p.D231N) were seen in two control individuals. The remaining variants were found only once. GGE: Genetic Generalized Epilepsy. FS: Febrile Seizures. TM: Transmembrane segment. PPh2, REVEL and MTR scores considered damaging/deleterious/intolerant are underlined.

Table S4: Top-ranked genes in the secondary analyses of rare functional variation.

Analysis	Variants		HGNC (OMIM gene)	Qualifying Cases			Qualifying Controls			OR (95% CI)	*P* value (homogeneity)
				1st Dataset	2nd Dataset	Both Datasets	1st Dataset	2nd Dataset	Both Datasets		
All GGEs	Rare	+ URVs	*ALDH8A1* (no)	14 (1.3%)	3 (0.4%)	17 (0.9%)	22 (0.3%)	2 (0.1%)	24 (0.3%)	3.9 (1.9 – 7.6)	7.5×10^{-5} (0.82)
		- URVs	*KIFC2* (no)	8 (0.7%)	1 (0.1%)	9 (0.5%)	5 (0.07%)	0 (0%)	5 (0.06%)	10.5 (3.1 – 40.4)	3.4×10^{-5} (0.64)
Familial GGEs	Rare	+ URVs	*ZIC3* (yes)	6 (1%)	0 (0%)	6 (1%)	3 (0.04%)	0 (0%)	3 (0.03%)	22.0 (4.7 – 136.7)	2.3×10^{-5} (0.53)
		- URVs	*ZIC3* (yes)	4 (0.6%)	0 (0%)	4 (0.5%)	0 (0%)	0 (0%)	0 (0%)	∞ (7.2 – ∞)	4.9×10^{-5} (0.23)
Sporadic GGEs	Rare	+ URVs	*COPA* (yes)	6 (1.4%)	7 (1.4%)	13 (1.3%)	19 (0.3%)	4 (0.2%)	23 (0.3%)	5.0 (2.2 – 10.7)	6.2×10^{-5} (0.69)
		- URVs	*FTHL17* (no)	5 (1.2%)	3 (0.6%)	8 (0.8%)	5 (0.07%)	2 (0.1%)	7 (0.08%)	10.4 (3.1 – 35.7)	6.5×10^{-5} (0.37)

Odds Ratio (OR) and *P* values are given from a Cochran-Mantel-Haenszel exact test. The accompanying homogeneity *p* value indicates the lowest *p* value from Breslow-Day & Woolf tests for homogeneity of odds, where *p* values < 0.05 indicate significantly different odds between the two analysis datasets. CI: Confidence Interval. GGE: Genetic Generalized Epilepsy. HGNC: Human Gene Nomenclature Consortium genes names. OMIM: Online Mendelian Inheritance in Man database. URVs: Ultra-rare variants. OMIM phenotypes: *ZIC3*: VACTERL, *COPA*: Autoimmune interstitial lung, joint, and kidney disease.

Table S5: Top-ranked genes in the secondary analyses of predicted Loss of Function (pLoF) variants.

Analysis	pLoF Variants		HGNC (OMIM gene)	Qualifying Cases			Qualifying Controls			OR (95% CI)	*P* value (homogeneity)
				1st Dataset	2nd Dataset	Both Datasets	1st Dataset	2nd Dataset	Both Datasets		
All GGEs	URVs only		*CEP350* (no)	5 (0.5%)	0 (0%)	5 (0.3%)	1 (0.01%)	0 (0%)	1 (0.01%)	31.1 (3.5 – 1460.3)	2.7×10^{-4} (0.28)
	Rare	+ URVs	*BEST3* (no)	8 (0.7%)	2 (0.2%)	10 (0.5%)	10 (0.1%)	0 (0%)	10 (0.1%)	5.6 (2.1 – 15.3)	3×10^{-4} (0.64)
		- URVs	*BEST3* (no)	7 (0.6%)	1 (0.1%)	8 (0.4%)	6 (0.09%)	0 (0%)	6 (0.07%)	7.7 (2.3 – 27.4)	3.1×10^{-4} (0.59)
Familial GGEs	URVs only		*CEP350* (no)	4 (0.6%)	0 (0%)	4 (0.5%)	1 (0.01%)	0 (0%)	1 (0.01%)	43.9 (4.3 – 2132.8)	2.3×10^{-4} (0.42)
	Rare	+ URVs	*CPA3* (no)	4 (0.6%)	0 (0%)	4 (0.5%)	1 (0.01%)	0 (0%)	1 (0.01%)	43.9 (4.3 – 2132.8)	2.3×10^{-4} (0.42)
		- URVs	*AIM1* (no)	6 (1%)	3 (1%)	9 (1%)	10 (0.1%)	5 (0.3%)	15 (0.2%)	5.2 (2.0 – 12.9)	4.6×10^{-4} (0.45)
Sporadic GGEs	URVs only		*GRIK5* (no)	3 (0.7%)	0 (0%)	3 (0.3%)	0 (0%)	0 (0%)	0 (0%)	∞ (5.8 – ∞)	3×10^{-4} (0.18)
	Rare	+ URVs	*GRIK5* (no)	3 (0.7%)	0 (0%)	3 (0.3%)	0 (0%)	0 (0%)	0 (0%)	∞ (5.8 – ∞)	3×10^{-4} (0.18)
		- URVs	*DSN1* (no)	4 (0.9%)	0 (0%)	4 (0.4%)	2 (0.02%)	0 (0%)	2 (0.02%)	28.1 (4.0 – 309.8)	2.7×10^{-4} (0.35)

Odds Ratio (OR) and *P* values are given from a Cochran-Mantel-Haenszel exact test. No gene reached the study wide significant *p* value of 2.9×10^{-7}. The accompanying homogeneity *p* value indicates the lowest *p* value from Breslow-Day & Woolf tests for homogeneity of odds, where *p* values < 0.05 indicate significantly different odds between the two analysis datasets. CI: Confidence Interval. GGE: Genetic Generalized Epilepsy. HGNC: Human Gene Nomenclature Consortium genes names. OMIM: Online Mendelian Inheritance in Man database. URVs: Ultra-rare variants.

Table S6: Comparisons of top-ranked genes with three previous large-scale rare variant association studies of genetic generalized epilepsy.

A. Association of top his from recent studies in the current analysis

Genes from recent studies				Outcomes in this study (All GGEs analysis)		
Study	**Rank**	**Gene**	***P* value**	**URVs PPh2**	**URVs REVEL**	**URVs MTR**
		Top-ranked genes		***P* values (Rank if ≤ 10) in this study**		
Epi4K & EP/GP	1	*CACNA1B*	0.000017	0.011	0.0028	0.0015 **(rank 6)**
	2	*KEAP1*	0.000056	0.0016	0.15	0.052
	3	*COPB1*	0.00022	0.039	0.089	0.0096
	4	*PHTF1*	0.00030	0.0071	0.0019	0.079
	5	*KCNQ2*	0.00040	0.61	1	1
	5	*SLC9A2*	0.00040	0.36	0.14	0.14
	7	*ATP1A3*	0.00092	0.035	0.016	0.025
	7	*GABRG2*	0.00092	0.000018 **(rank 1)**	0.00013 **(rank 1)**	0.000012 **(rank 1)**
	9	*ZNF100*	0.0010	0.041	0.039	0.039
	10	*CUX1*	0.0013	0.0066	0.017	0.0029
	10	*SCN1A*	0.0013	0.043	0.071	0.012
Epi25 Years 1&2	1	*CACNA1G*	0.00025	0.51	0.19	0.17
	2	*EEF1A2*	0.00038	0.21	0.57	0.57
	3	*GABRG2*	0.00062	0.000018	0.00013	0.000012 **(rank 1)**
	3	*UNC79*	0.00062	0.064	1	1
	5	*ALDH4A1*	0.0014	0.68	0.68	1
	6	*SLC6A1*	0.0020	1	0.57	0.57
	7	*RC3H2*	0.0020	0.75	1	1
	8	*GABRA1*	0.0022	0.0053	0.0023	0.0023 **(rank 9)**
	9	*LRRFIP1*	0.0052	0.76	0.59	1
	9	*DNAJC13*	0.0052	0.38	0.60	1
	9	*ZBTB2*	0.0052	1	1	1
Epi25 Years 1 – 3	1	*SLC6A1*	0.0000021	1	0.57	0.57
	2	*SCN1A*	0.000034	0.043	0.071	0.012
	3	*MYH8*	0.000262	0.31	0.33	0.065
	4	*FBXO42*	0.000447	1	1	1
	5	*DAW1*	0.000619	0.34	0.34	0.50

	6	*GRIN2A*	0.000862	1	1	1
	7	*NUP98*	0.000863	0.22	0.039	0.22
	8	*MYO5C*	0.001199	0.49	0.37	0.69
	9	*GABRA1*	0.001348	0.0053	0.0023	0.0023 **(rank 9)**
	10	*KCNK18*	0.001599	1	1	1

B. **Association of top his from the current analysis in recent studies**

Genes from this study (All GGEs analysis)				**Outcomes in previous studies**		
Analysis	**Rank**	**Gene**	***P* value**	**Epi4K & EP/GP**	**Epi25 Years 1&2**	**Epi25 Years 1 – 3**
		Top-ranked genes		***P* values (Rank ≤ 10) in recent studies**		
URVs PPh2	1	*GABRG2*	0.000018	0.00092 **(rank 7)**	0.00062 **(rank 3)**	0.0061
	2	*FAM13C*	0.000072	0.06	1	> 0.03
	3	*MRPL20*	0.000087	1	1	> 0.03
	4	*DHDH*	0.00012	0.05	1	> 0.03
	5	*PLCH2*	0.00017	1	1	> 0.03
	6	*ACSF2*	0.00056	0.26	1	> 0.03
	7	*KCNMA1*	0.00057	1	1	> 0.03
	8	*COL5A3*	0.00075	0.41	0.18	> 0.03
	9	*RLN3*	0.00085	1	0.27	> 0.03
	10	*TANC2*	0.0011	0.01	1	> 0.03
URVs REVEL	1	*GABRG2*	0.00013	0.00092 **(rank 7)**	0.00062 **(rank 3)**	0.0061
	2	*PDE1A*	0.00015	0.05	1	> 0.03
	3	*MDN1*	0.00023	0.66	0.18	> 0.03
	4	*WDR83*	0.00084	0.0098	1	> 0.03
	5	*RIOK2*	0.00085	0.0018	1	> 0.03
	6	*CEP350*	0.00090	0.045	0.27	> 0.03
	7	*TTC21B*	0.0014	0.46	1	> 0.03
	8	*PLEKHM3*	0.0016	0.023	0.57	> 0.03
	8	*FKBP10*	0.0016	0.26	1	> 0.03
	8	*SURF1*	0.0016	0.60	0.57	> 0.03
URVs MTR	1	*GABRG2*	0.000012	0.00092 **(rank 7)**	0.00062 **(rank 3)**	0.0061
	2	*CEP350*	0.00084	0.045	0.27	> 0.03
	3	*PRSS8*	0.00085	0.26	1	> 0.03

	4	*RELN*	0.0011	0.36	0.62	> 0.03
	5	*MDN1*	0.0013	0.66	0.18	> 0.03
	6	*CACNA1B*	0.0015	0.000017 **(rank 1)**	0.18	> 0.03
	7	*PRSS12*	0.0016	1	0.71	> 0.03
	8	*ZNF662*	0.0020	0.092	0.47	> 0.03
	9	*GABRA1*	0.0023	0.055	0.0022 **(rank 8)**	0.0013 **(rank 9)**
	9	*TCN2*	0.0023	0.26	0.66	> 0.03

Table S7: Association of genes encoding $GABA_A$ receptors.

Analysis	Gene	Qualifying Cases	Frequency in cases	Qualifying Controls	Frequency in controls	Direction of association	Association *P* value	Rank	Hypergeometric test *P* value
					All GGEs				
URVs PPh2	*GABRG2*	10	0.00519	4	0.00047	Cases	0.000018	1	1.0e-03
	GABRA1	5	0.00259	2	0.00023	Cases	0.0053	48	1.1e-03
URVs REVEL	*GABRG2*	7	0.00363	1	0.00012	Cases	0.00013	1	1.0e-03
	GABRA1	5	0.00259	1	0.00012	Cases	0.0023	19	1.6e-04
URVs MTR	*GABRG2*	7	0.00363	0	0.00000	Cases	0.000012	1	1.0e-03
	GABRA1	5	0.00259	1	0.00012	Cases	0.0023	9	3.5e-05
	GABRB3	2	0.00104	0	0.00000	Cases	0.044	216	1.3e-03
					Familial GGEs				
URVs PPh2	*GABRG2*	8	0.00847	4	0.00046	Cases	3.0e-06	1	1.0e-03
	GABRA1	3	0.00317	2	0.00023	Cases	0.0076	57	1.5e-03
	GABRB2	3	0.00317	3	0.00035	Cases	0.013	92	1.0e-04
URVs REVEL	*GABRG2*	5		1	0.00012	Cases	0.00010	1	1.0e-03
	GABRA1	3	0.00317	1	0.00012	Cases	0.0035	17	1.3e-04
	GABRB2	3	0.00317	3	0.00035	Cases	0.013	96	1.2e-04
URVs MTR	*GABRG2*	5	0.00529	0	0.00000	Cases	1.4e-05	1	1.0e-03
	GABRA1	3	0.00317	1	0.00012	Cases	0.0035	17	1.3e-04
					Sporadic GGEs				
URVs REVEL	*GABRB3*	2	0.00199	1	0.00012	Cases	0.033	188	1.7e-01
URVs MTR	*GABRB3*	2	0.00199	0	0.00000	Cases	0.015	71	6.9e-02
	GABRG2	2	0.00199	0	0.00000	Cases	0.015	72	2.4e-03

Results from the primary limited to genes with p values < 0.05.

Table S8: Association of OMIM genes implicated in susceptibility to generalized epilepsy.

Analysis	Gene	Qualifying Cases	Frequency in cases	Qualifying Controls	Frequency in controls	Direction of association	Association *P* value	Rank	Hypergeometric test *P* value
					All GGEs				
URVs PPh2	*GABRG2*	10	0.00519	4	0.00047	Cases	0.000018	1	7.4e-04
	KCNMA1	6	0.00311	2	0.00023	Cases	0.00057	7	1.1e-05
	GABRA1	5	0.00259	2	0.00023	Cases	0.00527	48	5.5e-06
	RORB	5	0.00259	2	0.00023	Cases	0.01538	133	2.3e-06
URVs REVEL	*GABRG2*	7	0.00363	1	0.00012	Cases	0.00013	1	7.4e-04
	GABRA1	5	0.00259	1	0.00012	Cases	0.00227	19	8.7e-05
	RORB	5	0.00259	2	0.00023	Cases	0.01538	83	2.9e-05
URVs MTR	*GABRG2*	7	0.00363	0	0.00000	Cases	0.000012	1	7.4e-04
	GABRA1	5	0.00259	1	0.00012	Cases	0.00227	9	1.8e-05
	RORB	4	0.00207	2	0.00023	Cases	0.03883	182	3.0e-04
	GABRB3	2	0.00104	0	0.00000	Cases	0.04440	216	1.5e-05
					Familial GGEs				
URVs PPh2	*GABRG2*	8	0.00847	4	0.00046	Cases	3.0e-06	1	7.4e-04
	RORB	5	0.00529	2	0.00023	Cases	0.00061	7	1.1e-05
	KCNMA1	4	0.00423	2	0.00023	Cases	0.00098	11	5.4e-08
	GABRA1	3	0.00317	2	0.00023	Cases	0.0076	57	7.4e-08
URVs REVEL	*GABRG2*	5	0.00529	1	0.00012	Cases	0.00010	1	7.4e-04
	RORB	5	0.00529	2	0.00023	Cases	0.00061	4	3.1e-06
	GABRA1	3	0.00317	1	0.00012	Cases	0.0035	17	2.2e-07
URVs MTR	*GABRG2*	5	0.00529	0	0.00000	Cases	1.4e-05	1	7.4e-04
	RORB	4	0.00423	2	0.00023	Cases	0.0031	13	4.0e-05
	GABRA1	3	0.00317	1	0.00012	Cases	0.0035	17	2.2e-07
					Sporadic GGEs				
	KCNMA1	4	0.00398	5	0.00058	Cases	0.0071	43	3.2e-02

URVs REVEL	*GABRB3*	2	0.00199	1	0.00012	Cases	0.033	188	8.3e-03
URVs	*GABRB3*	2	0.00199	0	0.00000	Cases	0.015	71	5.2e-02
MTR	*GABRG2*	2	0.00199	0	0.00000	Cases	0.015	72	1.3e-03

Results from the primary analyses limited to genes with *P* values < 0.05.

Table S9: Association of OMIM genes implicated autosomal dominant developmental and epileptic encephalopathies.

Analysis	Gene	Qualifying Cases	Frequency in cases	Qualifying Controls	Frequency in controls	Direction of association	Association *P* value	Rank	Hypergeometric test *P* value
					All GGEs				
URVs PPh2	*GABRG2*	10	0.00519	4	0.00047	Cases	0.000018	1	2.4e-03
	CACNA1A	11	0.00571	21	0.00245	Cases	0.00370	32	2.6e-03
	GABRA1	5	0.00259	2	0.00023	Cases	0.00527	48	2.0e-04
	KCNA2	5	0.00259	4	0.00047	Cases	0.00656	55	8.9e-06
	SCN1A	11	0.00571	25	0.00291	Cases	0.04277	376	2.0e-03
URVs REVEL	*GABRG2*	7	0.00363	1	0.00012	Cases	0.00013	1	2.4e-03
	GABRA1	5	0.00259	1	0.00012	Cases	0.00227	19	9.3e-04
	KCNA2	5	0.00259	4	0.00047	Cases	0.00656	43	1.5e-04
	GRIN2B	3	0.00156	1	0.00012	Cases	0.00958	58	1.1e-05
	HCN1	3	0.00156	1	0.00012	Cases	0.01860	103	4.6e-06
	KCNB1	4	0.00207	5	0.00058	Cases	0.03897	272	4.4e-05
	CACNA1A	8	0.00415	18	0.00210	Cases	0.04283	283	4.5e-06
URVs MTR	*GABRG2*	7	0.00363	0	0.00000	Cases	0.000012	1	2.4e-03
	GABRA1	5	0.00259	1	0.00012	Cases	0.00227	9	2.0e-04
	KCNA2	5	0.00259	4	0.00047	Cases	0.00656	32	6.0e-05
	GRIN2B	3	0.00156	1	0.00012	Cases	0.00958	42	3.0e-06
	SCN1A	9	0.00467	13	0.00152	Cases	0.01180	52	1.5e-07
	HCN1	3	0.00156	1	0.00012	Cases	0.01860	76	2.5e-08
	CHD2	3	0.00156	2	0.00023	Cases	0.02148	132	2.6e-08
	CACNA1A	6	0.00311	11	0.00128	Cases	0.03028	140	1.3e-09
	GABRB3	2	0.00104	0	0.00000	Cases	0.04440	216	1.8e-09
					Familial GGES				
URVs PPh2	*GABRG2*	8	0.00847	4	0.00046	Cases	3.0e-06	1	2.4e-03
	GABRA1	3	0.00317	2	0.00023	Cases	0.0076	57	8.2e-03
	GABRB2	3	0.00317	3	0.00035	Cases	0.013	92	1.4e-03
	SCN1A	7	0.00741	25	0.00290	Cases	0.030	202	1.4e-03
	HCN1	2	0.00212	2	0.00023	Cases	0.038	258	3.6e-04

URVs REVEL	*GABRG2*	5	0.00529	1	0.00012	Cases	0.00010	1	2.4e-03
	GABRA1	3	0.00317	1	0.00012	Cases	0.0035	17	7.4e-04
	GABRB2	3	0.00317	3	0.00035	Cases	0.013	96	1.6e-03
	CHD2	3	0.00317	4	0.00046	Cases	0.016	103	1.1e-04
	GRIN2B	2	0.00212	1	0.00012	Cases	0.020	123	1.1e-05
	HCN1	2	0.00212	1	0.00012	Cases	0.020	124	4.8e-07
	KCNB1	3	0.00317	5	0.00058	Cases	0.030	190	3.1e-07
	SCN1A	7	0.00741	25	0.00290	Cases	0.033	196	1.8e-08
	CACNA1A	5	0.00529	17	0.00197	Cases	0.037	201	9.5e-10
URVs MTR	*GABRG2*	5	0.00529	0	0.00000	Cases	1.4e-05	1	2.4e-03
	GABRA1	3	0.00317	1	0.00012	Cases	0.0035	17	7.4e-04
	CHD2	3	0.00317	2	0.00023	Cases	0.0052	21	1.6e-05
	SCN1A	6	0.00635	13	0.00151	Cases	0.0089	46	4.3e-06
	GRIN2B	2	0.00211	1	0.00012	Cases	0.020	92	2.6e-06
	HCN1	2	0.00211	1	0.00011	Cases	0.020	93	8.6e-08
	KCNB1	3	0.00317	4	0.00046	Cases	0.021	115	9.9e-09
	CACNA1A	4	0.00423	11	0.00128	Cases	0.037	151	2.4e-09
Sporadic GGEs									
URVs PPh2	*CACNA1A*	6	0.00597	18	0.00209	Cases	0.0062	48	1.1e-01
	KCNA2	3	0.00299	4	0.00046	Cases	0.017	141	4.5e-02
URVs REVEL	*KCNA2*	3	0.00299	4	0.00046	Cases	0.017	114	2.4e-01
	GABRB3	2	0.00199	1	0.00012	Cases	0.033	188	7.4e-02
URVs MTR	*GABRB3*	2	0.00199	0	0.00000	Cases	0.015	71	1.6e-01
	GABRG2	2	0.00199	0	0.00000	Cases	0.015	72	1.3e-02
	KCNA2	3	0.00299	4	0.00046	Cases	0.017	83	1.0e-03

Results from the primary analyses limited to genes with *P* values < 0.05.

Table S10: PPh2, REVEL and MTR scores of epilepsy related *GABRG2* missense variants.

A. Functionally characterized variants

Variant			Reported findings			In silico predictions			Functional validation	
NM_000816	NM_198904	NM_198903	Phenotypes	Segregation	PubMed ID	PPh2	REVEL	MTR	Effect	PubMed ID
N79S	N79S	N79S	GGE-GTCS	Unknown	20485450	0.55	0.34	**0.76**	(G-)LOF	23257655
			NAFE	Unknown	31327507				LOF	24798517
R82Q	R82Q	R82Q	FS+CAE	Inherited, incomplete penetrance	11326275	**1**	**0.9**	**0.66**	LOF	11326275
			GEFS+, NAFE, MAE	Inherited, incomplete penetrance	12919400					
P83S	P83S	P83S	GEFS+	Inherited, incomplete penetrance	21714819	**1**	**0.98**	**0.71**	LOF	24798517
			GGE	Unknown	30033060					
T90R	T90R	T90R	DS	*de novo*	34095830	**1**	**0.84**	**0.65**	LOF	34095830
A106T	A106T	A106T	DEE	*de novo*	27864268	0.79	0.31	**0.69**	(G-)LOF	27864268
			GGE	Unknown	31327507					
I107T	I107T	I107T	DEE	*de novo*	27864268	**1**	**0.52**	**0.57**	(G-)LOF	27864268
R177G	R177G	R177G	FS	Inherited, incomplete penetrance	16924025	**0.95**	**0.72**	**0.32**	LOF	16924025
G257R	G257R	G297R	FS+NAFE	Inherited, incomplete penetrance	25726841	**1**	**0.88**	**0.53**	LOF	25726841
P282S	P282S	P322S	DEE	*de novo*	27864268	**1**	**0.98**	**0.20**	(G-)LOF	27864268
P302L	P302L	P342L	DS	*de novo*	28197552	**0.79**	**0.86**	**0.36**	LOF	28197552
R323W	R323W	R363W	DEE	*de novo*	27864268	**1**	**0.82**	**0.35**	LOF	27864268
			FS+NAFE	Unknown	32086284					
R323Q	R323Q	R363Q	DEE	*de novo*	27864268	**1**	**0.93**	**0.35**	LOF	25726841
			LGS	Inherited, incomplete penetrance	30660939					
			NAFE	*de novo*	25726841				LOF	27864268
			GGE	Unknown	31327507					
K328M	K328M	K368M	FS+	Inherited, incomplete penetrance	11326274	**1**	**0.92**	**0.34**	LOF	11326274
F343L	F343L	F383L	DEE	*de novo*	27864268	**1**	**0.8**	**0.53**	(G-)LOF	27864268
I389V	I397V	I437V	NAFE	Inherited, incomplete penetrance	25726841	0.01	0.27	0.96	No effect	25726841

PPh2/REVEL/MTR scores considered damaging/deleterious/intolerant in this study are highlighted. CAE: Childhood Absence Epilepsy. DEE: Developmental and Epileptic Encephalopathy. DS: Dravet Syndrome. FS/+: Febrile Seizures/Plus. GEFS+: Generalized Epilepsy with Febrile Seizures Plus. GGE: Genetic Generalized Epilepsy. GGE-GTCS: Genetic Epilepsy with Generalized Tonic Clonic Seizures Only. (G-)LOF: Predominant loss-of-function with gain-of-function features. LGS: Lennox-Gastaut syndrome. LOF: Loss-of-function. MAE: Myoclonic Atonic/Astatic Epilepsy. NAFE: Non-acquired Focal Epilepsy.

B. Additional reported variants (not functionally characterized).

Variant			Report			In silico predictions		
NM_000816	NM_198904	NM_198903	PubMed ID	Phenotypes	Segregation	PPh2	REVEL	MTR
T94A	T94A	T94A	31327507	Lesional epilepsy	Unknown	0.86	**0.51**	**0.64**
T94K	T94K	T94K	28837158	NDD-E	*de novo* (mosaic)	**1**	**0.65**	**0.64**
T120S	T120S	T120S	31327507	GGE	Unknown	**0.97**	**0.61**	**0.52**
T120M	T120M	T120M	31327507	GGE	Unknown	**1**	**0.74**	**0.52**
R125C	R125C	R125C	31327507	NAFE	Unknown	**1**	**0.84**	**0.62**
R136Q	R136Q	R136Q	31327507	GGE	Unknown	**0.97**	0.42	**0.70**
D175V	D175V	D175V	31327507	NAFE	Unknown	**1**	**0.92**	**0.49**
M199V	M199V	M199V	27066572	GEFS+	Inherited, complete penetrance in a small pedigree	**1**	**0.91**	**0.56**
			31327507	NAFE	Unknown			
I218S	I218S	I258S	31327507	NAFE	Unknown	**0.99**	**0.92**	**0.67**
L237F	L237F	L277F	31327507	NAFE	Unknown	**1**	**0.9**	**0.67**
Y265C	Y305C	Y305C	31327507	GGE	Unknown	**0.99**	**0.63**	**0.47**
L313G	L313G	L353G	31344879	NDD-E	*de novo* (mosaic)	**1**	NA	**0.25**
S340T	S340T	S380T	31327507	NAFE	Unknown	0.13	0.29	**0.45**
K371T	K371T	K411T	30660939	EMAS	Unknown	0.15	0.43	0.95

GEFS+: Generalized Epilepsy with Febrile Seizures Plus. GGE: Genetic Generalized Epilepsy. EMAS: Epilepsy with Myoclonic Atonic Seizures. NAFE: Non-acquired Focal Epilepsy. NDD-E: Neurodevelopmental Disorder with Epilepsy.

C. Additional (likely) pathogenic variants in ClinVar (not functionally characterized).

Variant			ClinVar ID	In silico predictions		
NM_000816	NM_198904	NM_198903		PPh2	REVEL	MTR
L74V	L74V	L74V	VCV000205539	**1**	**0.75**	0.83
N167K	N167K	N167K	VCV000651395	**1**	**0.65**	**0.74**
P282T	P282T	P322T	VCV000205546	**1**	**0.98**	**0.20**
A300D	A300D	A340D	VCV000981261	**1**	**0.91**	**0.37**
L307V	L307V	L347V	VCV000205548	0.93	**0.79**	**0.23**
I321T	I321T	I361T	VCV000936276	**1**	**0.77**	**0.37**
G354V	G354V	G394V	VCV000408211	**1**	**0.71**	0.79

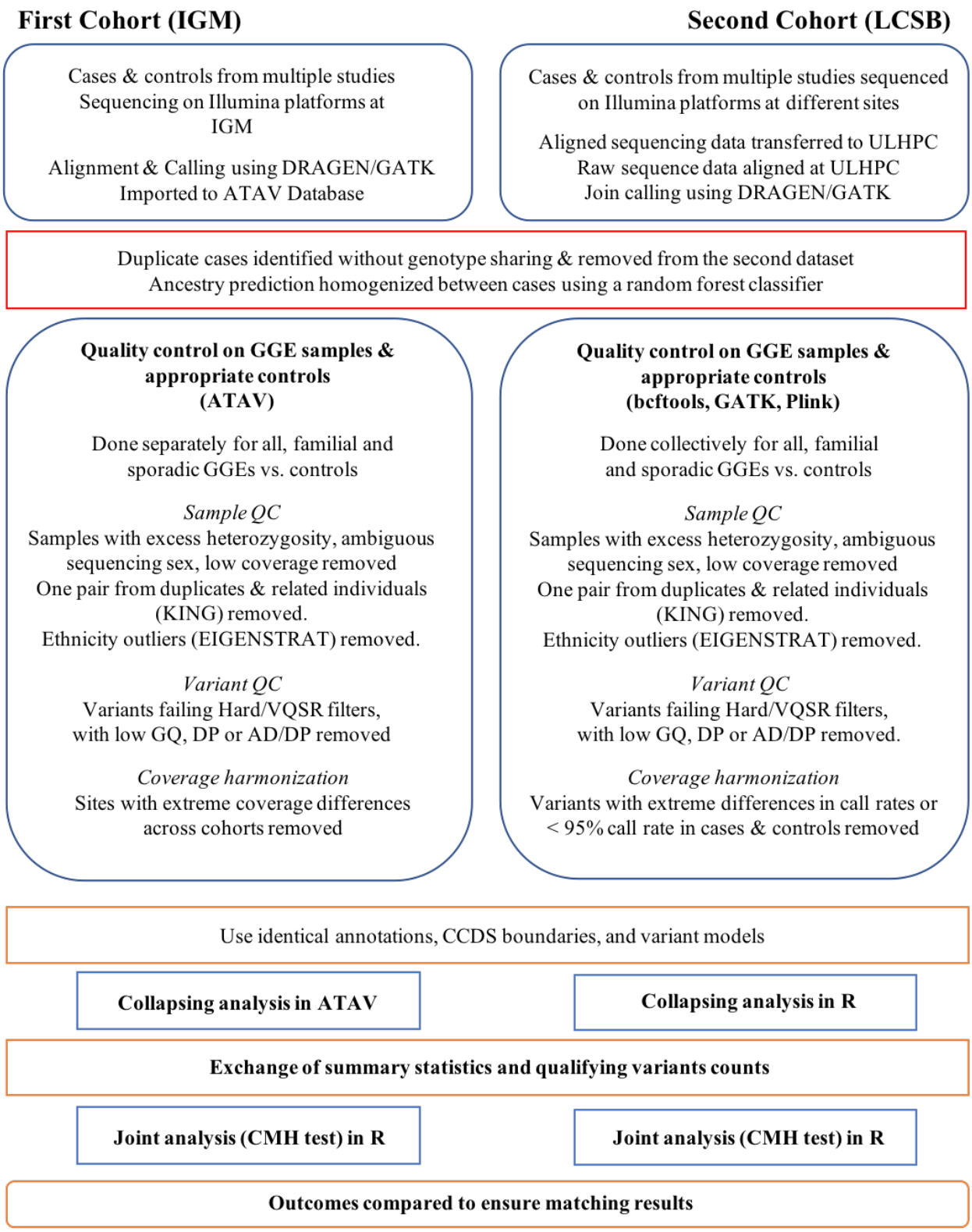

Fig. S1: Flow chart summarizing the analysis strategy used in this study. IGM: Institute of Genomic Medicine, New York, USA. LCSB: Luxembourg Centre for Systems Bioscience, Esch-sur-Alzette, Luxembourg. DRAGEN: Dynamic Read Analysis for Genomic platform. GATK: Genome Analysis Toolkit. ATAV: Analysis Tool for Annotated Variants. ULHPC: University of Luxembourg High Performance Computing Cluster. GGE: Genetic Generalized Epilepsy. QC: Quality control. GQ: Genotype Quality. AD: Allele Depth. DP: Depth. VQSR: Variant Quality Score Recalibration. CCDS: Consensus Coding Sequence. CMH: Cochran Mantel Haenszel test. Details on ATAV: https://github.com/igm-team/atav. Details on ULHPC: https://hpc.uni.lu/.

Fig. S2: Principal Component Analysis for ancestry matching. The plot shows eigenvectors on the first and second principal components from 1055 individuals with GGE vs. 6814 controls from the first dataset and 829 individuals with GGE vs. 1764 controls from the second dataset.

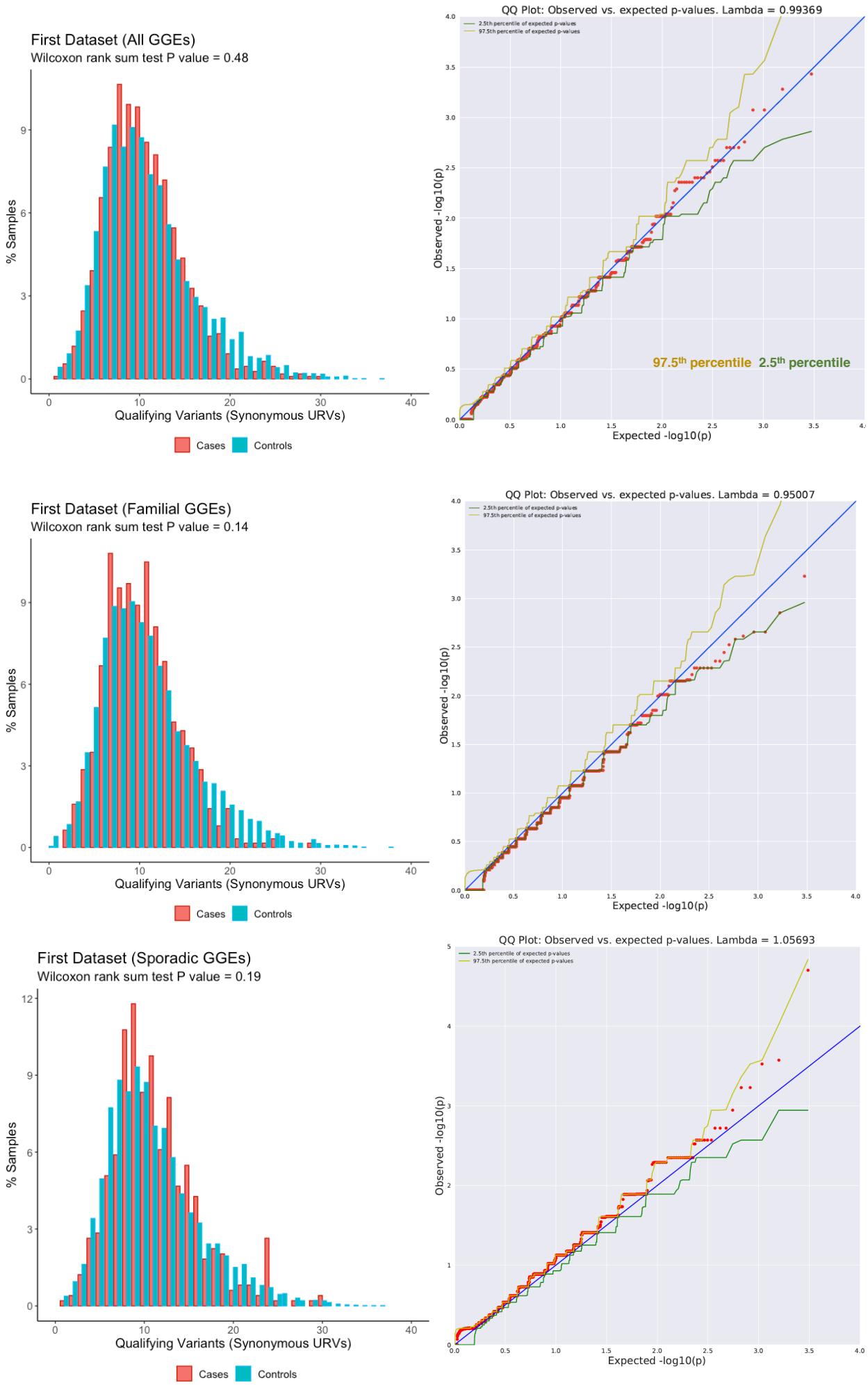

Fig. S3: Balance of ultra-rare synonymous qualifying variants tallies between cases and controls in the first dataset. There was no significant difference in the distribution of ultra-rare synonymous variants between the cases and controls. A single control sample in the first dataset with QVs tally exceeding 40 was not plotted but was included in significance testing. The QQ plots show the negative $\log_{10}$ of observed p values *vs.* the expected p values from 1000 permutations (mean, 2.5th and 97.5th centiles). P values were obtained from a two-sided Fisher's exact test of the association of ultra-rare synonymous qualifying variants

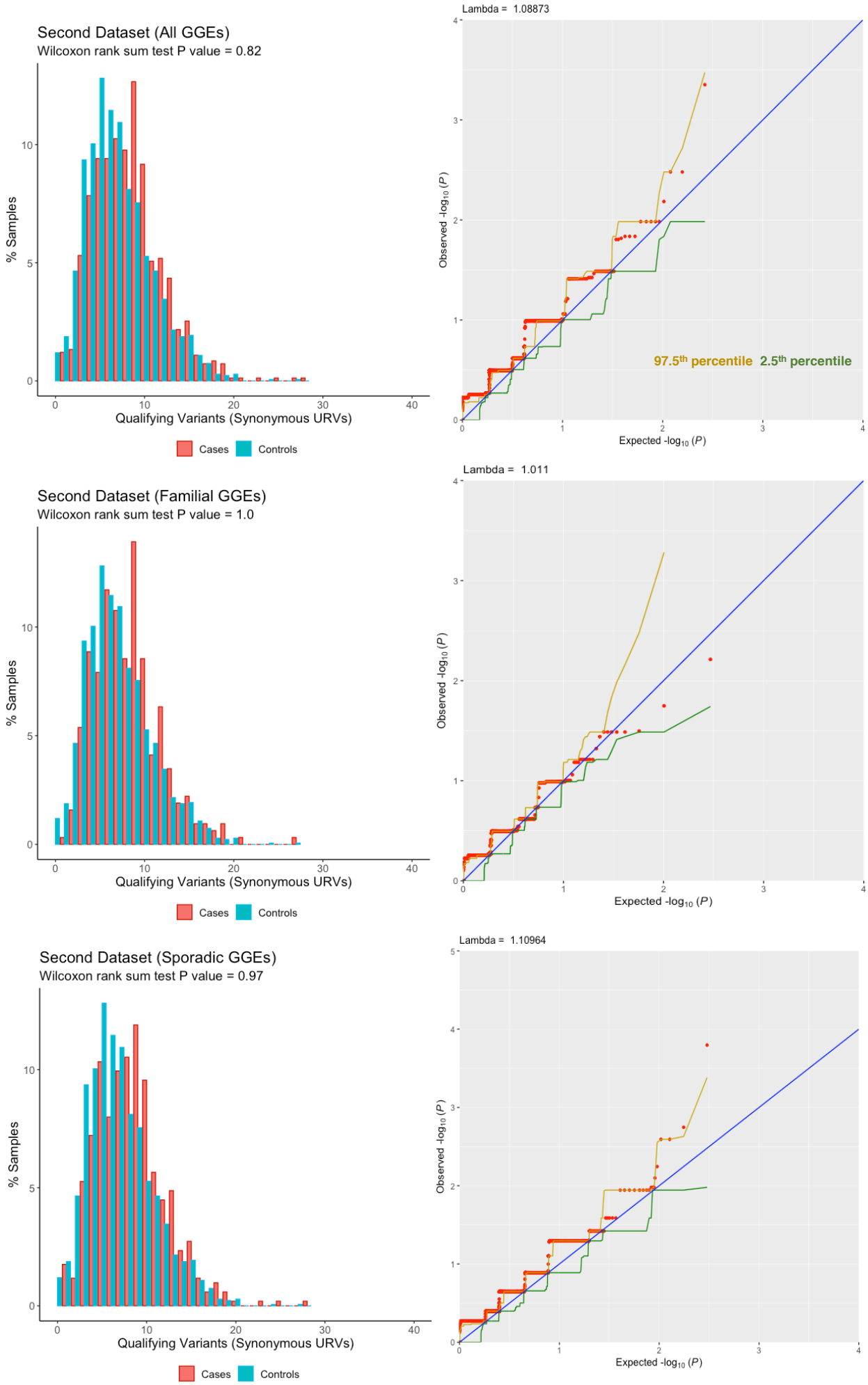

Fig. S4: Balance of ultra-rare synonymous qualifying variants tallies between cases and controls in the second dataset. There was no significant difference in the distribution of ultra-rare synonymous variants between the cases and controls. A single control sample in the first dataset with QVs tally exceeding 40 was not plotted but was included in significance testing. The QQ plots show the negative $\log_{10}$ of observed p values *vs.* the expected p values from 1000 permutations (mean, 2.5th and 97.5th centiles). P values were obtained from a two-sided Fisher's exact test of the association of ultra-rare synonymous qualifying variants

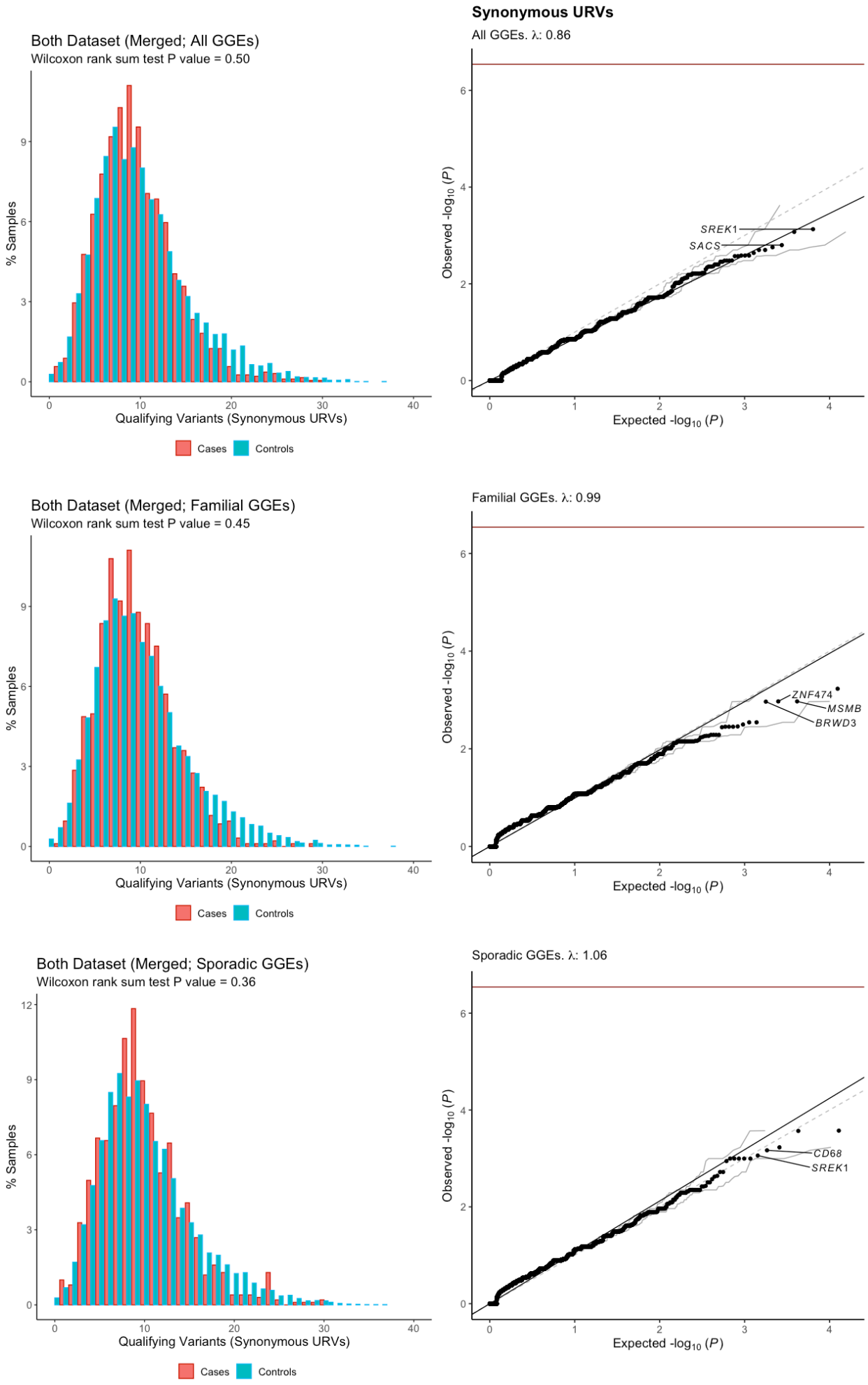

Fig. S5: Balance of ultra-rare synonymous qualifying variants tallies between cases and controls in the total dataset. The QQ plots show the negative $\log_{10}$ of observed p values *vs.* the expected p values from a uniform distribution. P values were obtained from a two-sided Cochran Mantel Haenszel exact test of the association of ultra-rare synonymous qualifying variants. The 95% confidence intervals are shown as grey solid lines. The slope of the solid black line indicates the genomic inflation factor, whereas the slope of the dotted line equals 1.

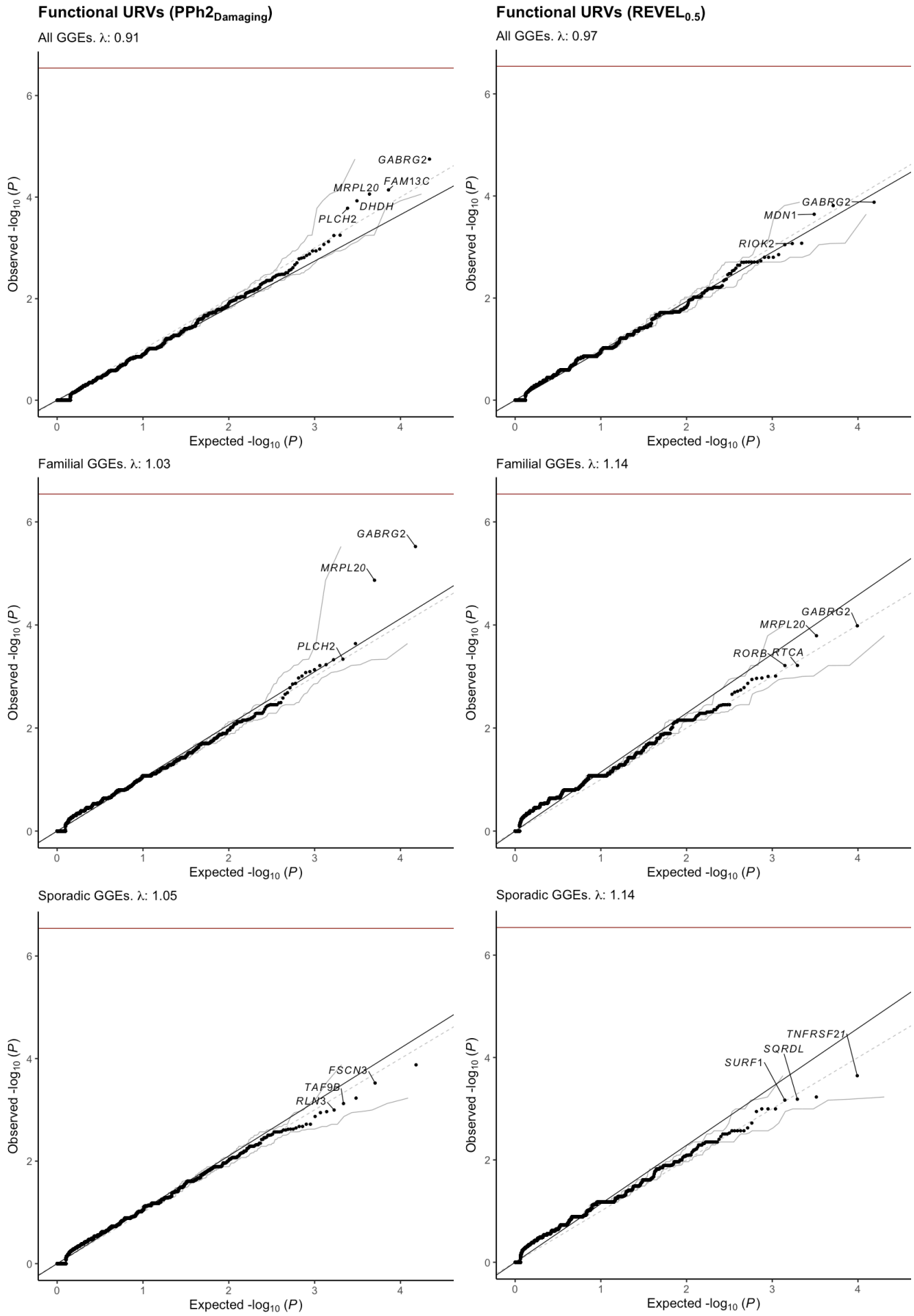

Fig. S6: Association of ultra-rare deleterious variants with genetic generalized epilepsy. The quantile-quantile plots compare observed *p* values (Cochran-Mantel-Haenszel exact test) and expected *p* values (drawn from a uniform distribution) in analyses of 1,928 individuals with genetic generalized epilepsy (GGEs) vs. 8,578 controls and subsets of familial GGEs (945 cases vs. 8,626 controls) or sporadic GGEs (1,005 cases vs. 8,621 controls). The 95% confidence intervals are shown as grey solid lines. The slope of the solid black line indicates the genomic inflation factor, whereas the slope of the dotted line equals 1. Labels: genes that are enriched in cases in both datasets among the five top-raking genes. Exome-wide significance after Bonferroni correction (dark red line) was defined by a *p* value < 2.9×10^{-7}.

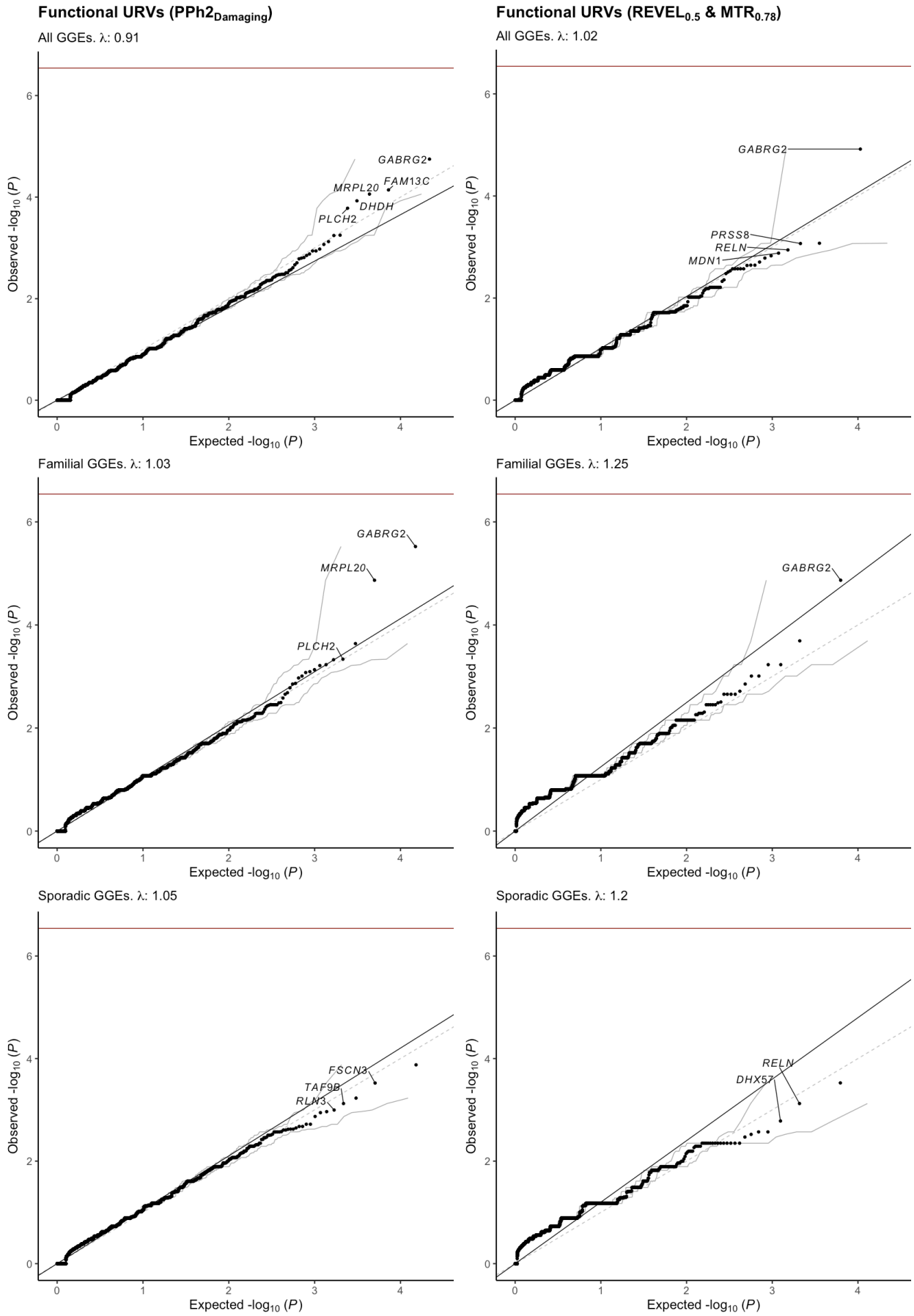

Fig. S7: Association of ultra-rare deleterious and intolerant variants with genetic generalized epilepsy. The quantile-quantile plots compare observed *p* values (Cochran-Mantel-Haenszel exact test) and expected *p* values (drawn from a uniform distribution) in analyses of 1,928 individuals with genetic generalized epilepsy (GGEs) vs. 8,578 controls and subsets of familial GGEs (945 cases vs. 8,626 controls) or sporadic GGEs (1,005 cases vs. 8,621 controls). The 95% confidence intervals are shown as grey solid lines. The slope of the solid black line indicates the genomic inflation factor, whereas the slope of the dotted line equals 1. Labels: genes that are enriched in cases in both datasets among the five top-raking genes. Exome-wide significance after Bonferroni correction (dark red line) was defined by a *p* value < 2.9×10^{-7}.

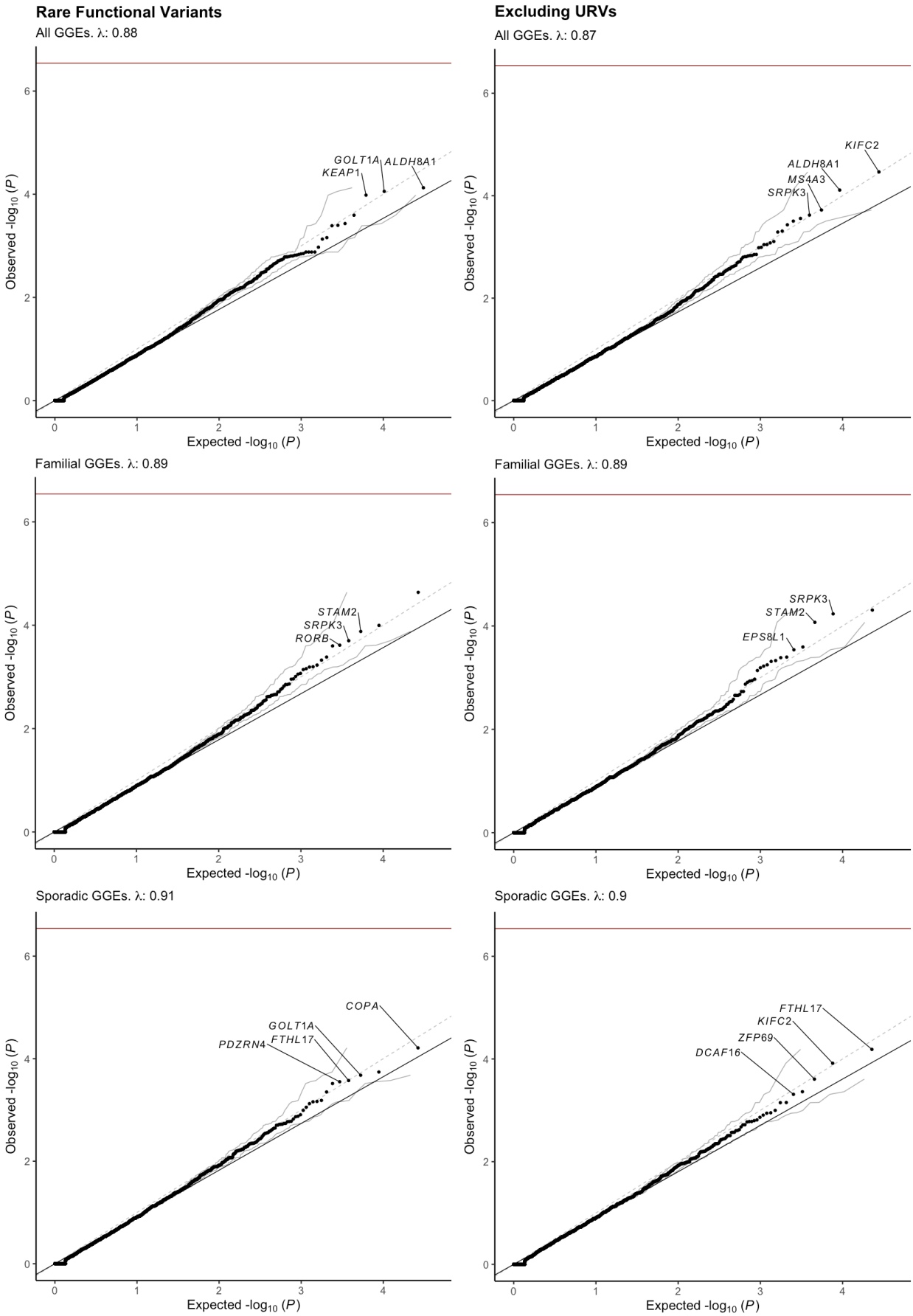

Fig. S8: Association of rare deleterious variants with genetic generalized epilepsy. The quantile-quantile plots compare observed *p* values (Cochran-Mantel-Haenszel exact test) and expected *p* values (drawn from a uniform distribution) in analyses of 1,928 individuals with genetic generalized epilepsy (GGEs) vs. 8,578 controls and subsets of familial GGEs (945 cases vs. 8,626 controls) or sporadic GGEs (1,005 cases vs. 8,621 controls). The 95% confidence intervals are shown as grey solid lines. The slope of the solid black line indicates the genomic inflation factor, whereas the slope of the dotted line equals 1. Labels: genes that are enriched in cases in both datasets among the five top-raking genes. Exome-wide significance after Bonferroni correction (dark red line) was defined by a *p* value < 2.9×10^{-7}.

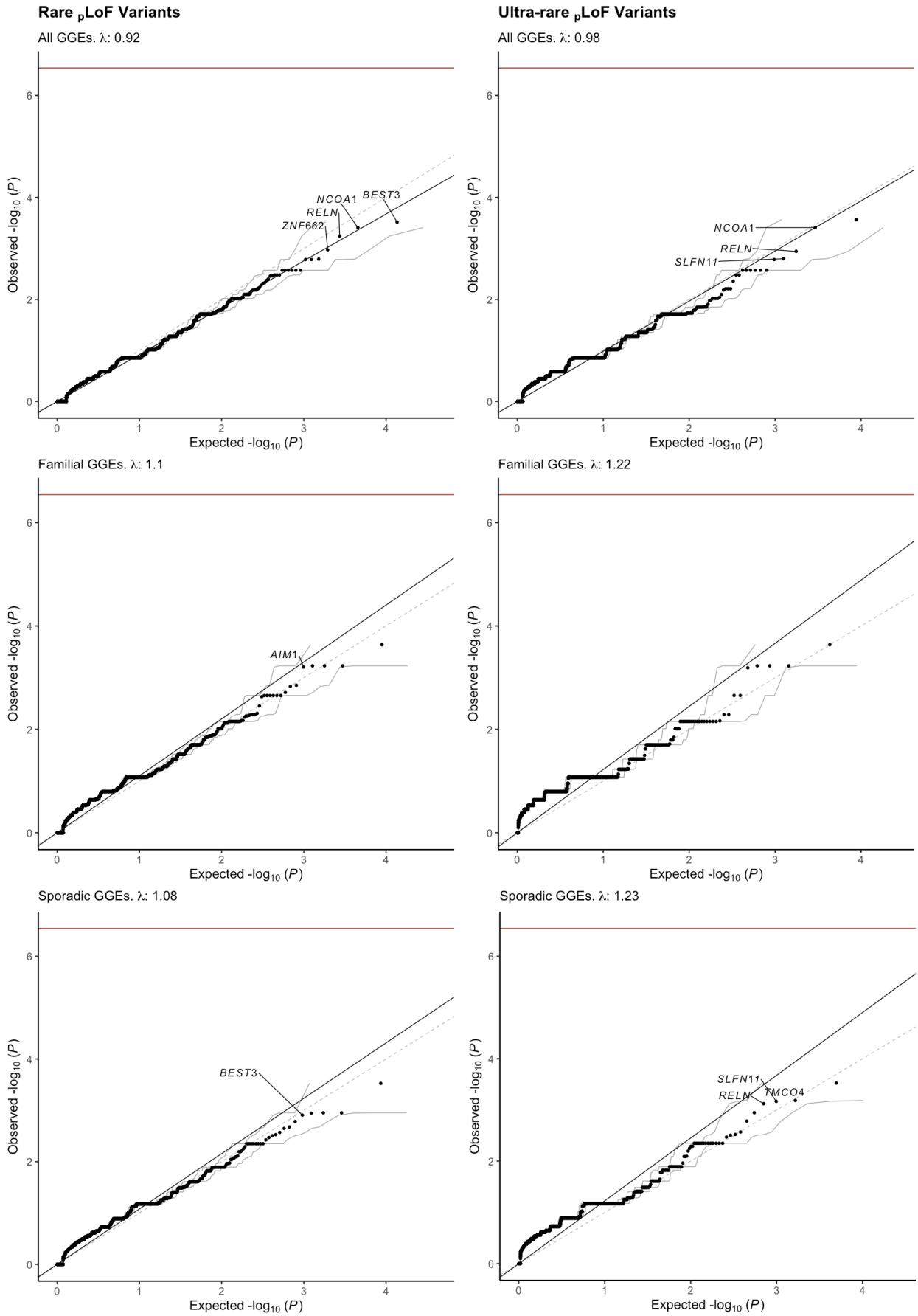

Fig. S9A: Association of rare predicted loss of function variants (including URVs) with genetic generalized epilepsy. The quantile-quantile plots compare observed *p* values (Cochran-Mantel-Haenszel exact test) and expected *p* values (drawn from a uniform distribution) in analyses of 1,928 individuals with genetic generalized epilepsy (GGEs) vs. 8,578 controls and subsets of familial GGEs (945 cases vs. 8,626 controls) or sporadic GGEs (1,005 cases vs. 8,621 controls). The 95% confidence intervals are shown as grey solid lines. The slope of the solid black line indicates the genomic inflation factor, whereas the slope of the dotted line equals 1. Labels: genes that are enriched in cases in both datasets among the five top-raking genes. Exome-wide significance after Bonferroni correction (dark red line) was defined by a *p* value < 2.9×10^{-7}.

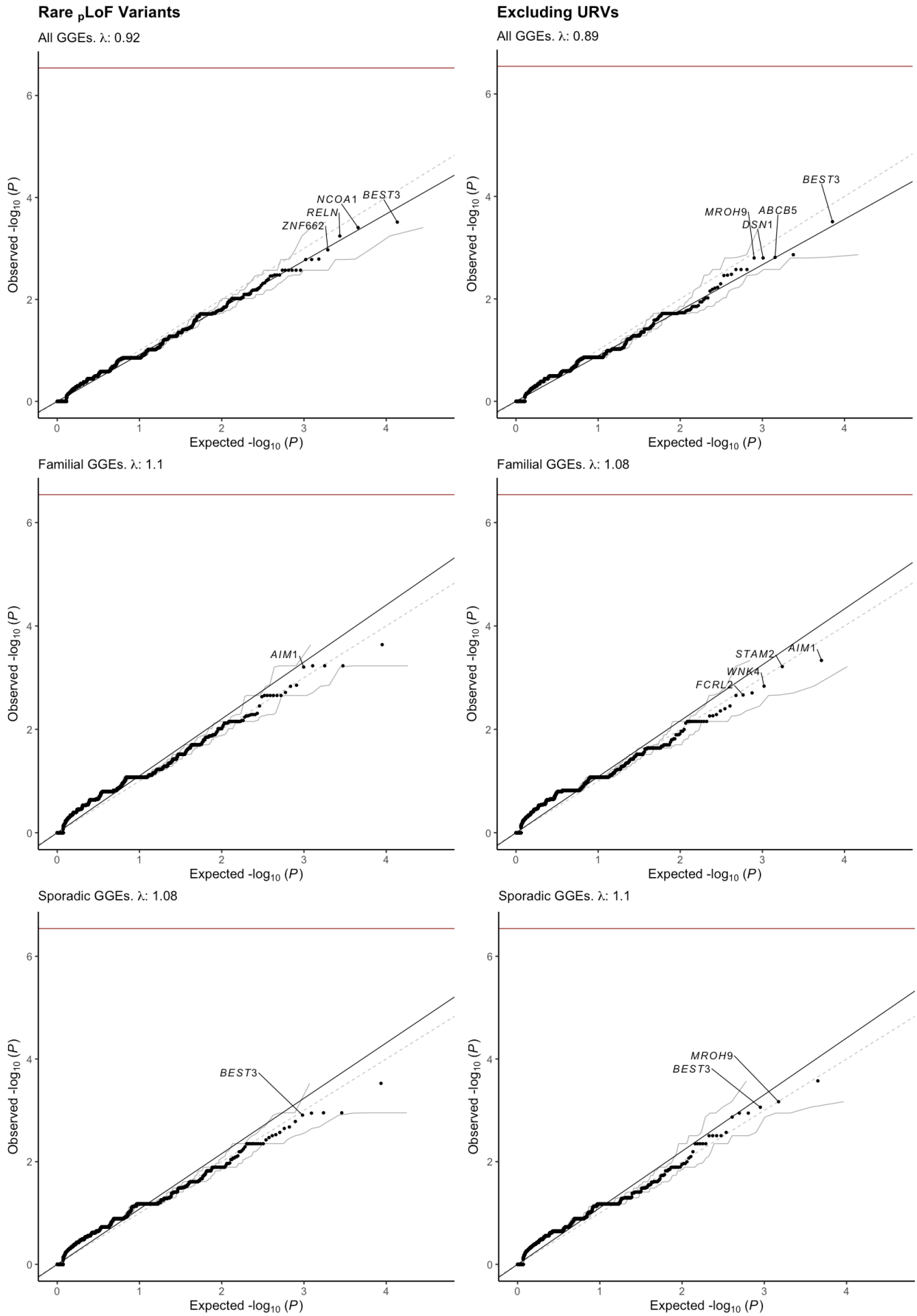

Fig. S9B: Association of rare predicted loss of function variants (excluding URVs) with genetic generalized epilepsy. The quantile-quantile plots compare observed *p* values (Cochran-Mantel-Haenszel exact test) and expected *p* values (drawn from a uniform distribution) in analyses of 1,928 individuals with genetic generalized epilepsy (GGEs) vs. 8,578 controls and subsets of familial GGEs (945 cases vs. 8,626 controls) or sporadic GGEs (1,005 cases vs. 8,621 controls). The 95% confidence intervals are shown as grey solid lines. The slope of the solid black line indicates the genomic inflation factor, whereas the slope of the dotted line equals 1. Labels: genes that are enriched in cases in both datasets among the five top-raking genes. Exome-wide significance after Bonferroni correction (dark red line) was defined by a *p* value < 2.9×10^{-7}.

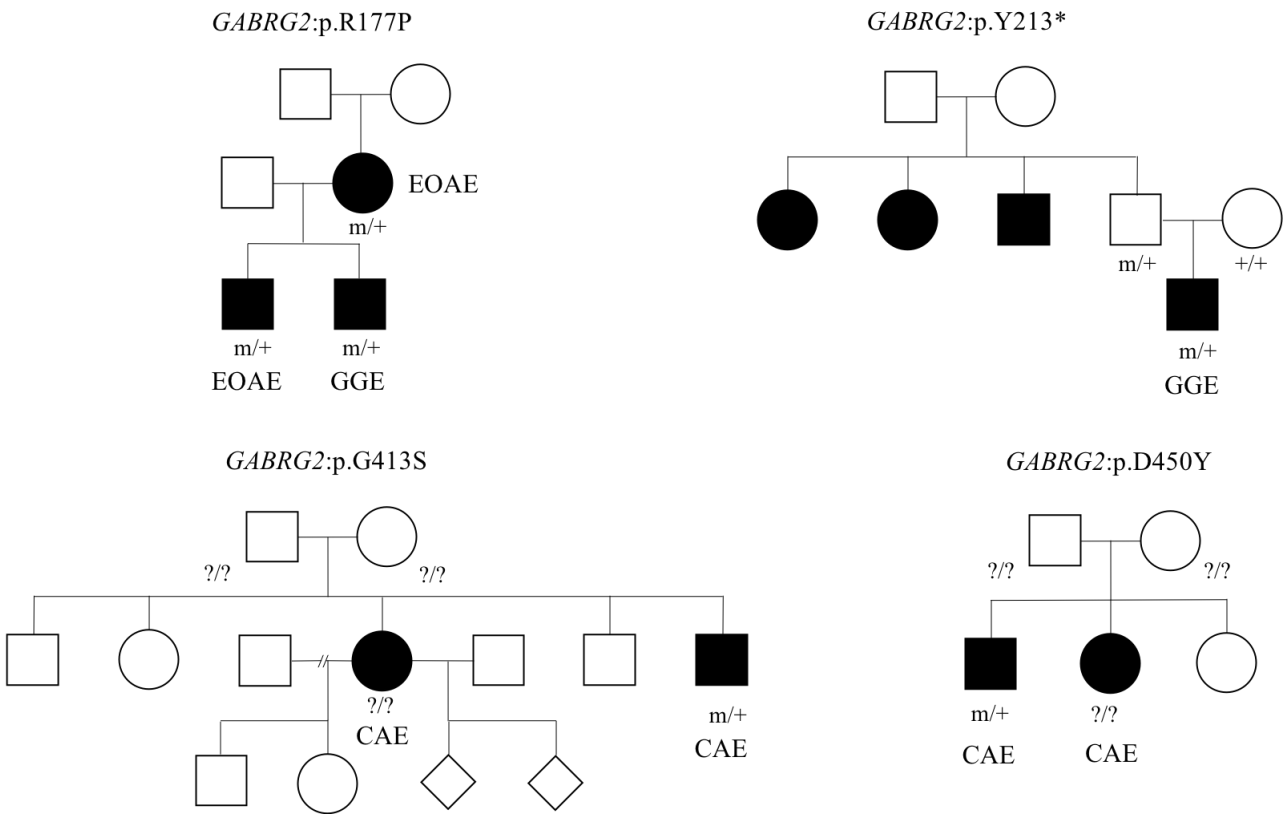

Fig. S10: Pedigrees of four families with genetic generalized epilepsy. Pedigree information was available for four carriers of ultra-rare variants in *GABRG2*. R177P, identified in a proband with Early Onset Absence Epilepsy (EOAE), was inherited from a parent with a similar phenotype (EOAE) and segregated in a sibling with Genetic Generalized Epilepsy (GGE) not further classified in a sub-syndrome. On the other hand, Y213* in a proband with GGE was inherited from a parent not diagnosed with epilepsy. Two other variants were identified in probands with Childhood Absence Epilepsy (CAE). These had siblings diagnosed with CAE, but segregation was not feasible at the time of this study. Pedigrees from six individuals carrying six other ultra-rare variants in *GABRG2* reported in this studt were not available.

Acknowledgements

- The Epi4K Consortium and Epilepsy Phenome/Genome Project were supported by grants from the National Institute of Neurological Disorders and Stroke (The Epilepsy Phenome/Genome Project U01 NS053998, Epi4K—Administrative Core U01 NS077274, Epi4K—Sequencing, Biostatistics and Bioinformatics Core U01 NS077303; Epi4K—Multiplex Families & Pairs Project U01 NS077367, and Epi4K—Phenotyping and Clinical Informatics Core U01 NS077276) and project grant (P1104) from Epilepsy-Research UK.
- The following individuals and/or groups contributed to control sample collection used by the Epi4K Consortium and EP/GP: D. Daskalakis; R Buckley; M.Hauser; J.Hoover-Fong, N. L. Sobreira and D. Valle; A. Poduri; T. Young and K. Whisenhunt; Z. Farfel, D. Lancet, and E. Pras; R. Gbadegesin and M. Winn; K. Schmader, S. McDonald, H. K. White and M. Yanamadala; R. Brown; S. H. Appel; E. Simpson; S. Halton, L. Lay; A. Holden; E. Behr; C. Moylan; A. M. Diehl and M. Abdelmalek; S. Palmer; G. Cavalleri; N. Delanty; G. Nestadt; D. Marchuk; V. Shashi; M. Carrington; R. Bedlack,; M. Harms; T. Miller; A. Pestronk; R. Bedlack; R. Brown; N. Shneider; S. Gibson; J. Ravits; A. Gilter; J. Glass; F. Baas; E. Simpson; and G. Rouleau; The ALS Sequencing Consortium; The Murdock Study Community Registry and Biorepository; the Carol Woods and Crosdaile Retirement Communities; Washington University Neuromuscular Genetics Project; the Utah Foundation for Biomedical Research; DUHS (Duke University Health System) Non-alcoholic Fatty Liver Disease Research Database and Specimen Repository, The Washington Heights, Inwood Columbia Aging Project.
- The collection of control samples and data was funded in part by: Biogen Idec.; Gilead Sciences, Inc.; New York-Presbyterian Hospital; The Columbia University College of Physicians and Surgeons; The Columbia University Medical Center; The Duke Chancellor's Discovery Program Research Fund 2014; Bill and Melinda Gates Foundation; The Stanley Institute for Cognitive Genomics at Cold Spring Harbor Laboratory; B57 SAIC-Fredrick Inc M11-074; The Ellison Medical Foundation New Scholar award AG-NS-0441-08; National Institute of Mental Health (K01 MH098126, R01 MH099216, R01 MH097993); National Institute of Allergy and Infectious Diseases (R56 AI098588); National Human Genome Research Institute (U01 HG007672); National Institute on Aging (R01 AG037212, P01 AG007232); and National Institute of Neurological Disorders and Stroke (U01 NS077303, U01 NS053998).
- The EuroEPINOMICS project was supported by the European Science Foundation through grants from national funding agencies (DFG: Le1030/11-1/2 to HL, Nu50/8-1 to PN, Sa434/5-1 to TS, He5415/3-1 to IH; FNR: INTER/ESF/10/02/CoGIE to RB, Academy of Finland: 141549 to AEL, Spanish Ministry of Economy and Competitiveness: EUIEURC-2011-4325 and SAF2010-18586 to JMS, TÜBİTAK grant 110S518 to HC).
- CENet was funded by joint funding from Genome Canada and Genome Quebec.

- We thank the Myocardial Infarction Genetics Exome Sequencing Consortium and the Broad Institute for generating the sequence data (supported by NHGRI funds; grant U54 HG003067 with Eric Lander as PI) of the Ottawa Heart Study (obtained from dbGaP through the accession number phs000806) used for the analysis of the EuroEPINOMICS, EpiPGX, and CENet.
- We utilized additional dbGaP data from the Alzheimer's Disease Sequencing Project obtained from dbGaP through the accession number phs000572 (PIs Richard A. Gibbs, Eric S. Lander, Richard K. Wilson, Gerard D. Schellenberg, and Sudha Seshadri were supported by grants: U24 AG041689, U01 AG032984, R01 AG033193, U24 AG021886; the sequence data generation was supported by NHGRI grants U54 HG003079, U54 HG003273, U54 HG003067).
- We are also grateful for the Epi25 Collaborative for providing access to their control cohort of Italian individuals that were of great help to match the southern European samples. The Epi25 project 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 (to Mark J. Daly). The Genome Sequencing Program efforts were also supported by NHGRI grant U01 HG009088-02. See the website of the Epi25 Collaborative for the full list of members (http://epi-25.org/epi25-members).
- The authors are also grateful to the participants of Rotterdam Study, the staff of the study, the participating general practitioners and pharmacists, members of the Genomics Lab and the ERGO support team. The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development, the Research Institute for Diseases in the Elderly, the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. Generation and management of genomics data for the Rotterdam Study were supported by the Netherlands Organization of Scientific Research NOW Investments (numbers 175.010.2005.011 and 911-03-012) and the Netherlands Genomics Initiative (NGI)/NOW project number 050-060-810 (Netherlands Consortium for Healthy Ageing).
- Parts of the computational analysis were done on the High-Performance Computing cluster of the University of Luxembourg (https://hpc.uni.lu/) and the ATAV platform at Columbia University (https://github.com/igm-team/atav). CENet sequences were stored or processed on Compute Canada cluster Beluga.

Affiliations

Albert Einstein College of Medicine: The Saul R. Korey Department of Neurology and the Department of Medicine, Bronx, New York 10461, USA.

Alex Boro, Sheryl R. Haut

Albert Einstein College of Medicine: The Saul R. Korey Department of Neurology and the Departments of Pediatrics and Epidemiology & Population Health, Bronx, New York 10461, USA.

Shlomo Shinnar

Antwerp University Hospital: Department of Neurology, Antwerp Edegem 2650, Belgium.

Sarah Weckhuysen

AstraZeneca: Centre for Genomics Research, Discovery Sciences, BioPharmaceuticals R&D, Waltham, Massachusetts 02451, USA.

Slavé Petrovski, Quanli Wang

Atrium Health/Levine Children's Hospital: Department of Pediatrics, Charlotte, North Carolina 28203, USA.

Rani K. Singh

Berlin Institute of Health at Charité – Universitätsmedizin and Max Delbrück Center for Molecular Medicine: Core Facility Genomics, 10115 Berlin, Germany.

Janine Altmüller

Boğaziçi University: Department of Molecular Biology and Genetics, Bebek, İstanbul 34342, Turkey.

Hande Caglayan

Boston Children's Hospital: Department of Neurology, Boston, Massachusetts 02115, USA.

Annapurna Poduri

Center for Personalized Translational Epilepsy Research (CePTER): Goethe University Frankfurt, Frankfurt 60528, Germany.

Karl M. Klein

Centre Hospitalier Universitaire de l'Université de Sherbrooke: Sherbrooke, Quebec J1H 5H3, Canada.

Cécile Cieuta-Walti

Children's Hospital of Philadelphia: Division of Neurology and Department of Biomedical and Health Informatics, Philadelphia, Philadelphia 19104 USA.

Ingo Helbig

Cincinnati Children's Hospital Medical Center: Division of Neurology, Cincinnati, Ohio 45229, USA.

Tracy Glauser

Cleveland Clinic: Department of Neurology, Cleveland, Ohio 44195, USA.

Jocelyn F. Bautista

Cologne Center for Genomics: University of Cologne, Cologne 50931, Germany.

Janine Altmüller, Peter Nürnberg, Thomas Sander

Columbia University: Columbia University Medical Center, Institute for Genomic Medicine, New York, New York 10032, USA.

Joshua Bridgers, Ryan S. Dhindsa, David B. Goldstein, Caroline Mebane, Joshua E. Motelow, Zhong Ren, Kate E. Stanley

Columbia University: Department of Neurology and the G.H. Sergievsky Center, New York, New York 10032, USA.

Ruth Ottman, Melodie R. Winawer

Comprehensive Epilepsy Center: Department of Neurology, NYU Grossman School of Medicine, New York, New York 10016, USA.

Judith Bluvstein, Orrin Devinsky, Jacqueline French, Daniel Friedman

Comprehensive Epilepsy Center: Montefiore Medical Center, Bronx, New York 10467, USA.

Alex Boro, Sheryl R. Haut, Shlomo Shinnar

Comprehensive Epilepsy Center: Oregon Health & Science University, Portland, Oregon 97239, USA.

Paul V. Motika

Danish Epilepsy Centre: Department of Epilepsy Genetics and Precision Medicine, Dianalund 4293, Denmark.

Rikke S. Møller

Danish Epilepsy Centre: Department of Pediatrics, Dianalund 4293, Denmark.

Marina Nikanorova

Duke University: Department of Biostatistics and Bioinformatics, Durham, North Carolina 27710, USA.

Andrew S. Allen

Duke University School of Medicine: Department of Neurology, Durham, North Carolina 27710, USA.

Rodney A. Radtke

Emory University School of Medicine: Department of Human Genetics, Atlanta, Georgia 30322, USA.

Michael P. Epstein

Epilepsy Center Frankfurt Rhein-Main: Department of Neurology, Center of Neurology and Neurosurgery, University Hospital Frankfurt, Goethe University Frankfurt, Frankfurt 60590, Germany.

Karl M. Klein

Epilepsy Research Centre: Department of Medicine, University of Melbourne (Austin Health), Heidelberg, Victoria 3052, Australia.

Susannah T. Bellows, Samuel F. Berkovic, Rosemary Burgess, Michael S. Hildebrand, Slavé Petrovski, Brigid M. Regan, Ingrid E. Scheffer

Florey Institute of Neuroscience and Mental Health: University of Melbourne, Parkville, Victoria 3052, Australia.

Steven Petrou

Folkhälsan Research Center and **University of Helsinki**: Helsinki 00250, Finland.

Anna-Elina Lehesjoki

Genetics Research Centre: St George's University of London, London SW17 0RE, UK.

Kate Everett

Hertie Institute for Clinical Brain Research: Department of Neurology and Epileptology, University of Tübingen, Tübingen 72076, Germany.

Felicitas Becker, Holger Lerche, Mahmoud Koko, Julian Schubert, Yvonne G. Weber, Stefan Wolking

Imperial College: Faculty of Medicine, Department of Brain Sciences, London W12 0NN, UK.

Michael R. Johnson, Sarah Langley

Institute of Experimental Epileptology and Cognition Research: Department of Epileptology, University of Bonn, Bonn 53105, Germany.

Wolfram S. Kunz

IRCCS 'G. Gaslini' Institute: Genova 16147, Italy.

Pasquale Striano, Federico Zara

Johns Hopkins Hospital: Department of Neurology, Baltimore, Maryland 21287, USA.

Eileen P.G. Vining

Johns Hopkins Hospital: Departments of Neurology and Pediatrics, Baltimore, Maryland 21287, USA.

Eric H. W. Kossoff

Kanta-Häme Central Hospital: Department of Pediatric Neurology, Hämeenlinna 13530, Finland.

Auli Siren

Lee Kong Chian School of Medicine: Nanyang Technological University Singapore, Singapore 636921, Singapore.

Sarah Langley

Mayo Clinic: Department of Neurology, Jacksonville, Florida 32224, USA.

Joseph Sirven

Mayo Clinic: Department of Neurology, Division of Epilepsy, Rochester, Minnesota 55905, USA.

Gregory D. Cascino

McGill University: Department of Neurology & Neurosurgery, Montreal, Québec QC H3A 2B4, Canada.

Dina Amrom

McGill University: Departments of Neurology & Neurosurgery and Human Genetics, Montreal, Québec QC H3A 2B4, Canada.

Eva Andermann

McGovern Medical School: University of Texas Health Science Center at Houston, Houston, Texas 77030, USA.

Gretchen K. von Allmen

Medical University of Vienna: Department of Neurology, Vienna 1090, Austria.

Martin Krenn, Fritz Zimprich

Monash University: Department of Neuroscience, The Central Clinical School, The Alfred Hospital, Melbourne, Victoria 3004, Australia.

Terence J. O'Brien

Montreal Neurological Hospital and Institute: Neurogenetics Unit and Epilepsy Research Group, Montreal, Québec QC H3A 2B4, Canada.

Eva Andermann

Montreal Neurological Hospital and Institute: Neurogenetics Unit, Montreal, Québec QC H3A 2B4, Canada.

Dina Amrom

Murdoch Children's Research Institute: Department of Medicine, University of Melbourne (Austin Health), Melbourne, Victoria 3084, Australia.

Michael S. Hildebrand

Newcastle University: Translational and Clinical Research Institute, Newcastle NE2 4HH, UK.

Rhys H. Thomas

Perelman School of Medicine at the University of Pennsylvania: Department of Neurology, Philadelphia, Philadelphia 19104, USA.

Ingo Helbig

Perelman School of Medicine at the University of Pennsylvania: Departments of Neurology and Pediatrics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania 19104, USA.

Dennis Dlugos

Ramos Mejía Hospital: Epilepsy Center, Neurology Division, Buenos Aires 1221, Argentina.

Damián E. Consalvo

Royal College of Surgeons in Ireland: School of Pharmacy and Biomolecular Sciences, Science Foundation Ireland FutureNeuro Research Centre, Dublin 2 D02YN77, Ireland.

Gianpiero Cavalleri, Norman Delanty

Rush University Medical Center: Department of Neurological Sciences, Rush Epilepsy Center, Chicago, Illinois 60612, USA.

Andres Kanner, Michael C. Smith

Saint Barnabas Medical Center: Division of Neurology, Livingston, New Jersey 07039, USA.

Eric B. Geller, Peter Widdess-Walsh

Sorbonne Université: Institut du Cerveau - Paris Brain Institute - ICM, Inserm, CNRS, AP-HP, Department of Medical Genetics, Paris 75013, France.

Stéphanie Baulac, Eric LeGuern

Stichting Epilepsie Instellingen Nederland (SEIN): Zwolle 8025 BV, Netherlands.

Bianca Berghuis

Swansea University Medical School: Institute of Life Science, Swansea University, Swansea SA2 8PP, UK.

Seo-Kyung Chung, Mark I. Rees

Swansea University Medical School: Neurology and Neuroscience, Swansea University, Swansea SA2 8PP, UK.

William O. Pickrell

The Children's Hospital at Saint Peter's University Hospital and Rutgers University: Department of Pediatrics, New Brunswick, New Jersey 08901, USA.

Anu Venkat

UCL Queen Square Institute of Neurology: Department of Clinical and Experimental Epilepsy, University College London, London WC1N 3BG, UK.

Andreja Avbersek, Costin Leu, Josemir W. Sander, Sanjay M. Sisodiya

Ulm University Hospital: Department of Neurology, University of Ulm, Ulm 89070, Germany.

Felicitas Becker

Université de Montreal: Centre de recherche du Centre hospitalier de l'Université de Montréal, Montréal, Quebec H2X 0A9, Canada.

Véronique Cloutier, Micheline Gravel, Caroline Meloche

Université de Montreal: Centre Hospitalier de l'Université de Montréal, Montréal, Quebec H2X 3E4, Canada.

Dang K. Nguyen

Université de Montreal: CHU Sainte-Justine, Montréal, Quebec H3T 1C5, Canada.

Anne Lortie

Université de Montreal: Départements de neurosciences, Faculté de médecine, Montréal, Québec H3C 3J7, Canada.

Patrick Cossette

Université de Paris: Reference Centre for Rare Epilepsies (CRéER), Department of Pediatric Neurology, Necker-Enfants Malades Hospital, Université de Paris, Paris 75743, France.

Rima Nabbout

Université du Québec à Chicoutimi: Centre Intersectoriel en Santé Durable, Chicoutimi, Saguenay, Quebec QC G7H 2B1, Canada.

Simon Girard, Claudia Moreau

Université du Québec à Chicoutimi: Département des sciences fondamentales, Chicoutimi, Saguenay, Quebec QC G7H 2B1, Canada.

Simon Girard

Université Laval: Centre Hospitalier Universitaire de l'Université Québec, Québec, Québec QC G1V 4G2, Canada.

Richard Desbiens

Université Libre de Bruxelles: Department of Neurology, Hôpital Erasme, Brussels 1070, Belgium.

Chantal Depondt

University "Magna Graecia" of Catanzaro: Department of Medical and Surgical Sciences, School of Medicine, Catanzaro 88100, Italy.

Antonio Gambardella

University Hospital RWTH Aachen: Department of Epileptology and Neurology, University of Aachen, Aachen 52074, Germany.

Yvonne G. Weber, Stefan Wolking

University Medical Center Schleswig-Holstein: Department of Neuropediatrics, Christian-Albrechts University of Kiel, Kiel 24105, Germany.

Hiltrud Muhle

University Medical Center Utrecht: Department of Genetics, Utrecht 3584 CX, Netherlands.

Bobby P.C. Koeleman, Anja Sonsma

University of California, San Diego: Department of Neurosciences, UC San Diego School of Medicine, La Jolla, California 92093, USA.

Jerry J. Shih

University of California, San Francisco: Department of Neurology, San Francisco, California 94143, USA.

Catharine Freyer, Heidi E. Kirsch, Daniel H. Lowenstein, Elliott H. Sherr, Joseph Sullivan

University of California, San Francisco: Departments of Neurology, Radiology and Neurosurgery and the UCSF Seizure Disorders Surgery Program, San Francisco, California 94143, USA.

Robert C. Knowlton

University of California, San Francisco: Department of Clinical Pharmacy, UCSF School of Pharmacy, San Francisco, California 94143, USA.

Brian K. Alldredge

University of Colorado School of Medicine: Departments of Pediatrics and Neurology, Denver, Colorado 80045, USA.

Kristen Park

University of Glasgow: School of Life Sciences, Glasgow G12 8QQ, UK.

Graeme J. Sills

University of Liverpool: Department of Biostatistics, Liverpool L69 3BX, UK.

Ben Francis

University of Liverpool: Department of Health Data Science, Liverpool L69 3BX, UK.

Andrea Jorgensen

University of Liverpool: Department of Molecular and Clinical Pharmacology, Clinical Sciences Centre, Liverpool L9 7LJ, UK.

Pauls Auce, Anthony G. Marson

University of Luxembourg: Luxembourg Centre for Systems Biomedicine, Belvaux 4367, Luxembourg.

Dheeraj Bobbili, Roland Krause, Patrick May, Sarah Peter

University of Michigan: Department of Neurology and Michigan Neuroscience Institute, Ann Arbor, Michigan 48109, USA.

Jack M. Parent

University of Michigan: Department of Neurology, Comprehensive Epilepsy Program, Ann Arbor, Michigan 48109, USA.

Simon Glynn

University of Michigan: Department of Pediatrics, Ann Arbor, Michigan 48109, USA.

Sucheta Joshi, Renée A. Shellhaas

University of Minnesota Medical School: Department of Neurology, Epilepsy Care Center, Minneapolis, Minnesota 55455, USA.

Miguel Fiol

University of North Carolina at Chapel Hill: Division of Pharmacotherapy and Experimental Therapeutics, Chapel Hill, North Carolina 27599, USA.

Erin L. Heinzen

University of Otago: Department of Paediatrics and Child Health, Newtown, Wellington 6021, New Zealand.

Lynette G. Sadleir

University of Pittsburgh School of Medicine: Department of Pediatrics, Pittsburg, Pennsylvania 15224, USA.

Patricia K. Crumrine

University of Southern Denmark: Institute of Regional Health Research, Odense 5230, Denmark.

Rikke S. Møller

University of Virginia: Comprehensive Epilepsy Program, Charlottesville, Virginia 22908, USA.

Nathan B. Fountain

University of Washington: Department of Pediatrics, Division of Genetic Medicine, Seattle, Washington 98115, USA.

Heather C. Mefford

University of Washington School of Medicine: Departments of Neurology and Pediatrics, Seattle Children's Hospital, Seattle, Washington 98105, USA.

Edward J. Novotny

Vanderbilt University Medical Center: Department of Neurology, Nashville, Tennessee 37232, USA.

Kevin Haas

VIB Center for Molecular Neurology: Applied &Translational Neurogenomics Group, VIB, University of Antwerp, Antwerp 2610, Belgium

Sarah Weckhuysen

Washington University School of Medicine: Department of Neurology, Division of Pediatric Neurology, St. Louis, Missouri 63110, USA.

Liu Lin Thio, Judith L. Weisenberg

Zucker School of Medicine at Hofstra/Northwell: Department of Neurology, Hofstra University, Hempstead, New York 11549, USA.

Ruben Kuzniecky, Juliann M. Paolicchi

Additional Data

The complete results of the presented gene-level association are publicly available as Additional Supporting Data: Koko *et al*. (2021), "Supplements to the article: Association of ultra-rare coding variants with genetic generalized epilepsy", Mendeley Data, V1, https://doi.org/10.17632/5p556jtwt9.1