Genetic correlation between background EEG oscillations and genetic generalized epilepsy

Remi Stevelink[1,2,*], Jurjen J. Luykx[3,4,5,*,#], Bochao D. Lin[3,4,] Costin Leu[6], Dennis Lal[6], Alexander Smith[6], Dick Schijven[3,4], Johannes A. Carpay[7], Koen Rademaker[3], Roiza A. Rodrigues Baldez[8], Kees Braun[2], Floor Jansen[2], ILAE Consortium on Complex Epilepsies (see appendix for a list of members and affiliations), Epi25 Collaborative (see appendix for a list of members and affiliations), Dirk Smit[9,@] & Bobby Koeleman[1,@]

[1]Department of Genetics, UMC Utrecht Brain Center, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

[2]Department of Neurology, UMC Utrecht Brain Center, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands

[3]Department of Psychiatry, UMC Utrecht Brain Center, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands;

[4]Department of Translational Neuroscience, UMC Utrecht Brain Center, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

[5]GGNet Mental Health, Apeldoorn, the Netherlands

[6]Broad Institute of Harvard and MIT, Cambridge, Massachussets, USA.

[7]Department of Neurology, Tergooi Hospital, Hilversum, the Netherlands

[8] Clinical Research Laboratory on Neuroinfectious Diseases, Evandro Chagas Clinical Research Institute (IPEC), Fundação Oswaldo Cruz (FIOCRUZ), Rio de Janeiro, RJ, Brazil

[9]Psychiatry department, Amsterdam Neuroscience, Amsterdam University Medical Center, location AMC, The Netherlands.

[*] These authors contributed equally to this work;

[@] These authors jointly directed this work.

[#] To whom correspondence should be addressed at:

Jurjen Luykx, Department of Psychiatry, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands; Email: J.Luykx@umcutrecht.nl

Keywords: epilepsy, EEG, genetic, correlation, polygenic.

Words: Abstract: 328; Body: 3573

Running title: Genetic overlap between epilepsy and EEG

Number of figures: 4; number of tables: none.

Abstract

Paroxysmal epileptiform abnormalities on EEG are the hallmark of epilepsies, but it is uncertain to what extent epilepsy and background EEG oscillations share neurobiological underpinnings. Confounding factors, including the heterogeneous etiology of epilepsies and medication effects hamper studies on background brain activity in people with epilepsy. This limitation may be partly overcome by assessing the genetic overlap between epilepsy and EEG power spectra based on genome-wide association studies (GWASs). We used the summary statistics of the recently published EEG (n=8,425) and epilepsy (n=37,021) GWASs to compute estimates of genetic overlap between EEG power spectra and epilepsy subtypes using linkage disequilibrium score regression (LDSC) and polygenic risk scoring (PRS). We found significant genetic correlations between genetic generalized epilepsy (GGE) and EEG beta power (r_g=0.44; p=0.01) and between GGE and theta power (r_g =0.25; p=0.02), which were replicated in an independent GGE cohort. Supporting this finding, beta and theta powers' PRSs were both strongly and positively associated with generalized epilepsy (p-values ranging from $5.3x10^{-8}$ to $8.5x10^{-4}$), suggesting that genetic predisposition to GGE and to altered background brain activity partially overlap. All PRS directions of effect were replicated in an independent cohort of people with epilepsy and controls (n=25,303; p=0.023). Furthermore, subjects in the highest 10% quantile of beta-power PRS scores were 1.4 fold more likely to suffer from GGE compared to subjects in the lowest 10% quantile (OR: 1.40; 95%CI: 1.18-1.67; $p=1.5x10^{-4}$). Moreover, by meta-analyzing EEG and epilepsy GWASs data for gene-based tests and by imputing gene expression, we found increased enrichment of brain-expressed genes relative to both the epilepsy and the EEG GWASs individually. Finally, Mendelian Randomization analyses suggested a causal effect of GGE genetic liability on beta oscillations. In conclusion, our results point to shared biological mechanisms underlying background EEG and the susceptibility for seizures. Our findings thus open avenues to investigate the clinical utility of theta and beta oscillations in GGE. They may also inspire preclinical research into the links between background EEG oscillations and epilepsies.

Introduction

The power of oscillations in background EEG is a highly stable and heritable human trait (Smit *et al.*, 2005). It is easily acquired and can be automatically analyzed by software, rather than subjective interpretation. Epilepsy is fairly heritable and is characterized by altered brain excitability (Thomas and Berkovic, 2014; Devinsky *et al.*, 2018). Oscillatory activity is believed to serve an essential role in cortico-thalamic functioning, and can be measured as power of oscillations in background EEG at different broad-band frequencies (Steriade, 2000). Neurophysiological relationships between background EEG and generalized epileptiform discharges have been well described (Gloor, 1978, 1979; Gloor and Fariello, 1988; Steriade, 2005), providing a rationale to explore the shared neurobiological basis between epilepsy as a disorder of excitability and oscillatory power as a measure of cortico-thalamic function.

Here, we therefore assessed whether oscillatory background EEG is genetically correlated with focal and generalized epilepsy. The association between genetic variants and background brain activity was previously investigated in a sizeable genome-wide association study (GWAS) on subjects without epilepsy (Smit *et al.*, 2018). We combined this data with a recently published large GWAS of epilepsy (International League Against Epilepsy Consortium on Complex, 2018), to examine genetic correlations between several types of epilepsy and oscillatory brain activity across frequency bands (delta: 1-3.75Hz, theta: 4-7.75 Hz, alpha: 12.75 Hz and beta: 13-30Hz). Next, we utilized polygenic risk scoring to assess whether patients with GGE have a genetic predisposition towards altered background brain activity. We then replicated genetic correlation and polygenic analyses using an independent cohort. Furthermore, we meta-analyzed both EEG and epilepsy traits to assess pleiotropic signals. Finally, we performed Mendelian Randomization to gain insight into possible causal relationships between genetic variants associated with epilepsy and those associated with background EEG. We thus provide converging evidence for consistent cross-trait genetic overlap between epilepsy and background EEG.

Materials and methods

Study population: discovery dataset

The participants derived from the epilepsy GWASs (International League Against Epilepsy Consortium on Complex, 2018) for the current analyses were Caucasian subjects. Case/control ascertainment and diagnostic criteria were previously reported (International League Against Epilepsy Consortium on Complex, 2018). As described previously (International League Against Epilepsy Consortium on Complex, 2018), epilepsy specialists diagnosed patients and ascertained phenotypic subtypes. Population-based datasets, some of which had been screened to exclude neurological disorders, were used as controls.Summary statistics from the recent epilepsy GWAS conducted by the ILAE Consortium on Complex Epilepsies GWAS were available for n=12,803 cases (suffering from either focal or generalized epilepsy) and 24,218 controls (International League Against Epilepsy Consortium on Complex, 2018). From those participants the following subjects were excluded for those analyses requiring individual-level genotype data: Finnish ancestry (none had GGE) and the subset of the EPICURE-SP1 cohort that lacked informed consent for the current analyses, resulting in subject-level genotype data being available for 11,446 patients with epilepsy and 22,078 controls. Subjects with epilepsy were stratified into GGE (n=3,122) and focal epilepsy (n=8,324); GGE was further subdivided into childhood absence epilepsy (CAE; n=561), juvenile absence epilepsy (JAE; n=311), juvenile myoclonic epilepsy (JME; n=1,000) and generalized tonic-clonic seizures only (GTCS only; n=195). GGE subtype information was not available for 1055 patients.

We downloaded summary statistics of the ENIGMA-EEG GWAS of resting state oscillation power in the delta (1-3 Hz), theta (4-8 Hz), alpha (8-12 Hz), and beta (13-30 Hz) bands (Smit et al, 2018) at the vertex (Cz) electrode (n=8,425 participants). In this study, subjects with a history of neurological illness, including epilepsy, were excluded (Supplemental methods).

Approval for the source studies was obtained by all relevant institutional review boards and all study participants provided written informed consent according to the Declaration of Helsinki.

Replication dataset

To replicate our findings, we used data from the Epi25 collaborative (http://epi-25.org/). This

cohort currently constitutes 4,851 people with epilepsy, of whom 2,612 have focal epilepsy and 2,239 have GGE (no data on GGE subtypes were available). The cases were matched to a total of 20,428 controls from the Partners Healthcare Biobank (n=14,857), the Epi25 collaborative (n=210), the Genetics and Personality consortium (n=456), and an in-house project on inflammatory bowel disease (n=4,905). The cohorts were genotyped on the Illumina Global Screening Array, with the exception of the Partners Healthcare Biobank participants, who were genotyped on the Illumina Multi-Ethnic Screening Array. Approval was obtained by all relevant institutional review boards and all study participants provided written informed consent according to the Declaration of Helsinki.

Genetic correlation analyses

Genetic correlations between epilepsy subtypes and oscillatory brain activity were computed using bivariate linkage-disequilibrium score regression (LDSC) (Bulik-Sullivan *et al.*, 2015). For these analyses, as no individual-level genotype data were available from the EEG dataset, we used published summary statistics of the EEG frequency bands (alpha, beta, delta and theta; n=8,425 participants) and the epilepsy subtypes (focal, GGE, CAE, JAE, JME and GTCS only; n=12,803 cases suffering from either focal or generalized epilepsy and 24,218 controls) from the ILAE consortium as a discovery dataset (International League Against Epilepsy Consortium on Complex, 2018). For LDSC replication analyses, we used unpublished data from the Epi25 collaborative (http://epi-25.org/; n=4,851 people with epilepsy and 20,428 controls). For discovery and replication LDSC analyses, default settings of LDSC were used, with pre-computed LD score weights derived from the European subset of the 1000 Genomes project (Genomes Project *et al.*, 2015). The significance threshold was Bonferroni-corrected for the two epilepsy subtypes studied (GGE and focal) but not for the EEG power spectra, since these were all highly correlated at $p<10^{-17}$ (Supplemental Table 1), resulting in a significance threshold of p=0.05/2=0.025.

Polygenic risk score (PRS) analyses

For PRS analyses we used individual-level genotype data derived from the epilepsy GWAS (n=11,446 patients with epilepsy and 22,078 controls, International League Against Epilepsy Consortium on Complex, 2018) and summary statistics from the EEG GWAS (n=8,425; Smit *et al.*, 2018). Quality control was performed as reported in the published epilepsy GWAS (International League Against Epilepsy Consortium on Complex, 2018). We then added a genotype filter for call rate >0.99 and the exclusion of genetically related subjects to allow for

highly conservative PRS estimates. Genetic inter-relatedness was calculated with KING (Manichaikul *et al.*, 2010) and one subject from each pair with third-degree or higher relatedness (kinship coefficient >0.0442) was excluded. PRSice (Euesden *et al.*, 2015) was used with default settings to assess whether subjects with epilepsy have different EEG frequency power PRSs compared to controls. In brief, to each SNP we assigned a weight proportional to its association in the four EEG GWASs (alpha, beta, delta and theta). Next, individual PRSs were calculated as the sum of weighted effect alleles for every subject from the epilepsy cohort. These PRSs were standardized with a Z-score transformation: $\frac{PRS-mean(PRS)}{SD(PRS)}$. SNPs were pruned to a subset of genetically uncorrelated SNPs (linkage disequilibrium $R^2<0.1$) and PRS values were calculated using a number of different P-value thresholds (P_ts) from 0.0001 to 0.5. Next, logistic regression analyses, corrected for sex and 10 genetic ancestry principal components, were performed to assess the association of these PRS scores with GGE. The PRSs with the highest association with GGE was chosen as the 'best-fit', after which logistic regression analyses were repeated to assess the association of this PRS with the other epilepsy subtypes. We used a conservative $p<0.001$ significance threshold to correct for multiple comparisons, as recommended for PRSice (Euesden *et al.*, 2015). Δ Explained variance represented by the Nagelkerke R^2 was computed using a logistic regression of the PRS, subtracted by the baseline model (covariates only: sex and 4 PCs). We then repeated the analyses in the independent Epi25 cohort. This dataset contained approximately one-third fewer GGE cases than the discovery cohort, providing insufficient power to exactly replicate our discovery PRS findings. We therefore performed quasi-replication using a one-sample test of the proportion to assess concordance effect directions between discovery and replication PRS analyses, computing z-scores that were converted into p-values:

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$$

where the p = the sample proportion; H_0 represents the null hypothesis: $p = p_0$; and the alternative hypothesis H_1: $p \neq p_0$.

Meta-analysis, Gene-based testing and Gene expression analysis

We found the strongest genetic correlation between GGE and beta oscillations, leading us to meta-analyze the beta oscillation GWAS (Smit *et al.*, 2018) and the GGE discovery GWAS (International League Against Epilepsy Consortium on Complex, 2018). We thus seeked to establish whether this approach would implicate novel genes associated with either phenotype and whether tissue enrichment results would change by including the EEG GWAS relative to the epilepsy GWAS (International League Against Epilepsy Consortium on Complex, 2018). To that end, we used a novel LD-based method for combined analysis of the case-control and quantitative traits (Demontis *et al.*, 2019) of the discovery GWAS and the EEG GWAS. The Epi25 collaborative dataset was not available for meta-analysis at the time of publication. We then used MAGMA positional gene-based analysis (https://ctg.cncr.nl/software/magma) for both the original GGE GWAS and the enhanced GGE + beta oscillation meta-analysis GWASs. Next, we performed a tissue enrichment analysis as implemented in FUMA (http://fuma.ctglab.nl/) by selecting the top 250 genes from these gene-based analyses and comparing them to gene sets. These gene sets consisted of differentially expressed genes (DEGs) based on the GTEx v7 database (https://gtexportal.org), where genes are considered differentially expressed in a tissue when they show a significant up or down regulation compared to all other GTEx tissues in a simple t-test ($p<0.05$). Enrichment analysis for presence of these DEGs in the top 250 genes from the gene-based analysis was then performed using the hypergeometric test. We selected 8 brain tissues (excluding cerebellar tissues) with the addition of whole blood and spleen gene expression for comparison purposes.

To functionally annotate the GWAS meta-analysis results, we used S-Predixcan, seeking to elucidate the tissue-specific expression pathways of the genes using the all GTEx cortical brain tissues (Gamazon *et al.*, 2015; Barbeira *et al.*, 2018). S-PrediXcan may have increased power to detect significant gene - phenotype associations. S-PrediXcan combines genetic variants in a sparse elastic net prediction model and is based on the GTEx genotype -- expression database and is available online (http://https://github.com/hakyimlab/MetaXcan). Using the precomputed models, PrediXcan imputes gene expression in a specific-tissue manner in subjects with genomic SNP data available. These gene-expression levels are then associated with the trait of interest. S-PrediXcan is the extension of Predixcan that uses only summary statistics, thus avoiding the need for individual level genomic data.

Mendelian Randomization

To explore possible causal effects of GGE genome-wide loci on EEG background oscillations, we conducted Mendelian Randomization analyses using GGE and EEGs summary statistics data. 228 SNPs significantly associated with GGE ($p<5x10^{-8}$) were extracted from both the GGE and EEG GWASs. The summary statistics of 228 SNPs were harmonized to ensure the SNP effect direction corresponded with equal effect alleles across GGE and EEG. We used the "TwosampleMR" package (Hemani *et al.*, 2018) in R to perform fixed effects inverse variance-weighted (IVW), weighted median, and MR Egger models. We then performed sensitivity analyses, including horizontal pleiotropic effects estimated by the intercept of MR Egger, residual heterogeneity due to pleiotropy estimated by Cochran's Q test (Burgess *et al.*, 2015), and leave-one-out analyses (for the IVW-fixed effect model) to evaluate if any single instrumental variable was driving the results. GSMR (generalized summary-data based mendelian randomisation) analyses were performed using the "GSMR" (Zhu *et al.*, 2018) package in R. To that end, first the LD matrix of the selected SNPs was calculated using PLINK (Purcell *et al.*, 2007) and GCTA (Yang *et al.*, 2011) within 1000 Genomes Phase 3 data (Genomes Project *et al.*, 2015). The minimum number of instrumental variables in the GSMR model was loosened from 10 to 5 as there were only 8 independent ($r^2<0.01$, LD window =10Mb), significant loci identified in the GWAS of GGE (and none in the EEG GWAS). We used default options in GSMR with Heterogeneity in Dependent Instruments (HEIDI) testing for instrumental outliers' detection. At the end, we repeated GSMR with loosened LD prune thresholds (namely r2<0.1, r2<0.15 and r2<0.2) because GSMR takes LD structure into account by adding the LD matrix. The significance threshold was Bonferroni corrected for all these seven mendelian randomisation models (p=0.05/7=0.007).

Data availability

GWASs summary statistics used for the current analyses are available online: http://enigma.ini.usc.edu/research/download-enigma-gwas-results/; http://www.epigad.org/gwas_ilae2018_16loci.html.

Results

Genetic correlations between epilepsy and oscillatory brain activity

In a total study population of 45,446 subjects (n=8,425 from the EEG and n=37,021 from the epilepsy GWASs), we computed genetic correlations between alpha, beta, delta and theta oscillatory brain activity with focal epilepsy and GGE. We found significant correlations between GGE and beta-power (Rg=0.44 ± s.e. of 0.18; p=0.01) and theta-power (Rg=0.25 ± 0.11; p=0.02; **Figure 1**, upper panel, Supplemental table 2)). This was further supported by the correlations between beta and theta-power with the GGE subtypes CAE, JAE and JME: all had similarly high correlation coefficients (**Figure 1**, upper panel, Supplemental table 2). We found no genetic correlations between focal epilepsy and any of the EEG phenotypes (**Figure 1**, upper panel, Supplemental table 2). We then attempted to replicate the genetic correlations using the unpublished Epi25 dataset and found similar genetic correlations (both in sign and effect size) to the discovery analyses (**Figure 1**, lower panel): GGE correlated with beta power (Rg=0.52 ± 0.21, p=0.01), while the genetic correlation between theta power and GGE paralleled the discovery cohort (albeit not reaching significance: Rg=0.16 ± 0.12; p=0.18). All genetic correlation estimates with focal epilepsy were again non-significant.

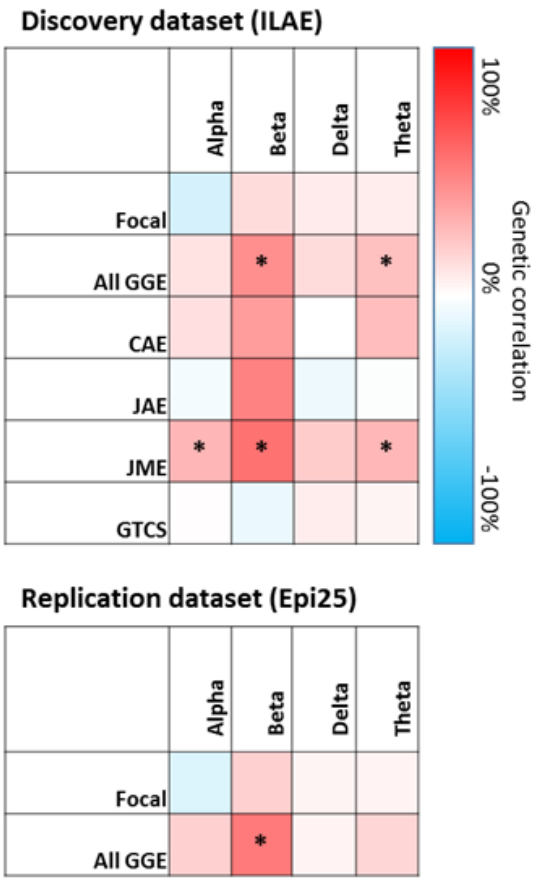


*Figure 1: Genetic correlations between EEG frequency bands and epilepsy subtypes. Genetic correlations were calculated by comparing the EEG frequency band GWAS with the ILAE GWAS (upper panel, discovery dataset) and the Epi25 GWAS (lower panel, replication dataset). * P<0.05.*

Oscillatory brain activity polygenic scores are associated with generalized epilepsy

We used polygenic scoring to utilize the full distribution of EEG oscillation-associated SNPs and assess whether people with epilepsy have a different polygenic score for specific frequency bands compared to controls. We observed significant positive associations between beta and theta-power PRSs with GGE, in line with the LDSC results (**Figure 2A)**. In particular beta power PRSs were strongly associated with GGE (beta=0.11; $p=5.3x10^{-8}$; explained variance 0.21%; **Figure 2A&B**), which was further supported by significant associations of beta power PRS with its subtypes CAE (beta=0.15; $p=8.5x10^{-4}$) and JME (beta=0.12; $p=3.6x10^{-4}$). Furthermore, of the participants in the GGE case-control cohort, those in the highest 10% quantile of beta-power PRS scores were 1.4 fold more likely to have GGE compared to the people in the lowest 10% PRS quantile (OR: 1.40; 95%CI: 1.18-1.67; $p=1.5x10^{-4}$; see **Figure 2C**). When using the independent Epi25 cohort as a replication dataset we found that the directions of effect agreed with the discovery analyses for all associations between EEG PRSs and GGE ($p_{one\text{-}sided}$= 0.023; $p_{two\text{-}sided}$= 0.046; supplemental Figure 1). In

line with our LDSC analyses, EEG PRSs could not explain the occurrence of *focal* epilepsy in our study population (all standardized beta's were non-significant and around zero).

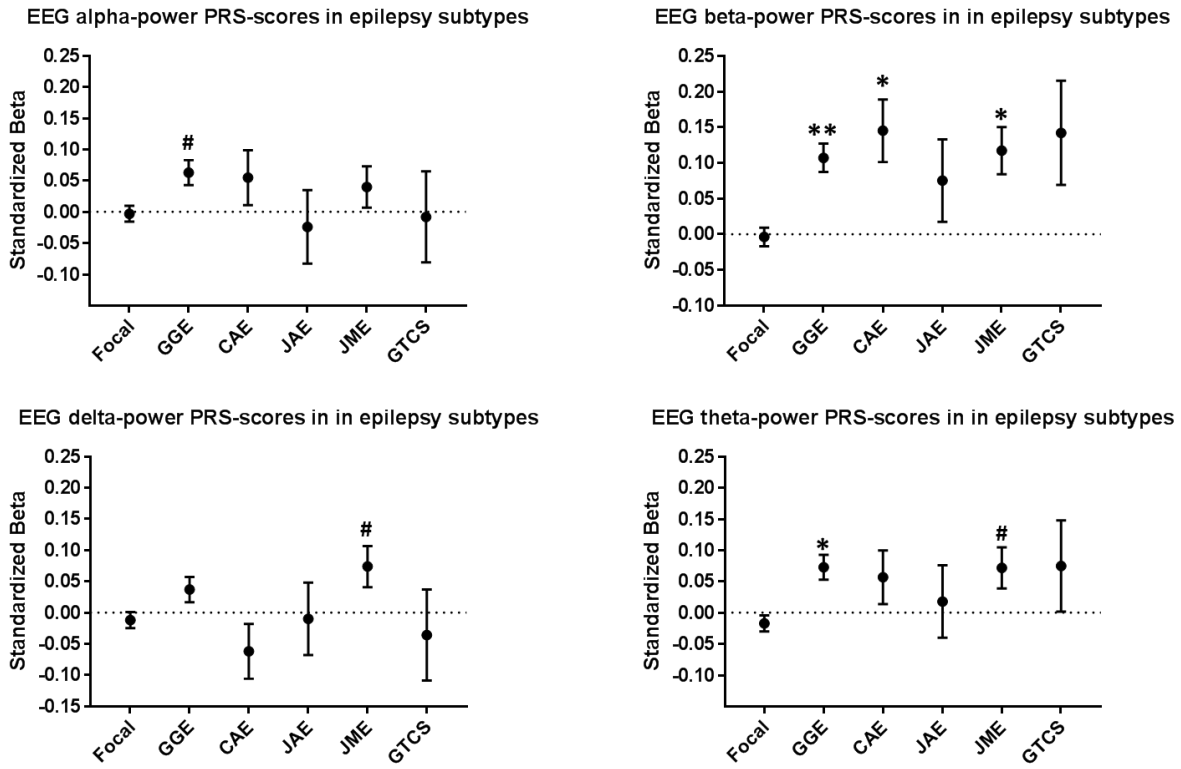


*Figure 2 A: Beta and theta power EEG oscillation polygenic risk scores (PRSs) are associated with generalized epilepsy but not with focal epilepsy. The 'best-fit' P-value threshold (P_t) was chosen based on the most significant association with GGE, which was then applied to all other epilepsy subtypes. The number of SNPs included in each model were: 2,670 for Alpha power (P_t =0.0105), 10,861 for Beta power (P_t=0.06245), 8,182 for Delta power (P_t =0.0446), and 3,833 for Theta power (P_t=0.01665). Logistic regression analyses were performed to assess the association between the PRSs and the different epilepsy subtypes, corrected for sex and 10 principal components. #: P<0.05; *: P<0.001; **: P<10^{-7}. Focal = focal epilepsy; GGE = genetic generalized epilepsy; CAE (childhood absence epilepsy), JAE (juvenile absence epilepsy), JME (juvenile myoclonic epilepsy) and GTCS (generalized tonic-clonic seizures only) are GGE subtypes.*

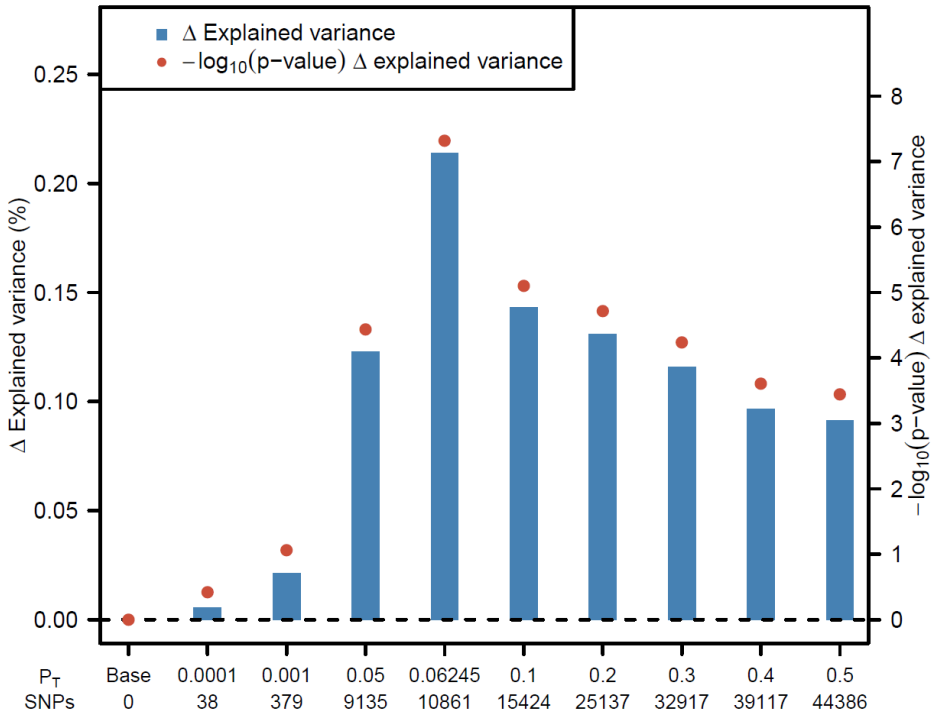


Figure 2B: Explained variances of beta-power PRSs at several P-value thresholds (P_t) in 3,122 GGE cases and 22,078 controls. P-value thresholds are displayed on the X axis, where the number of included SNPs increases with more lenient P_ts. P_t=0.06245 was the 'best-fit' P_t (most significant P_t), as assessed with PRSice. Δ Explained variance represents the Nagelkerke R^2 (shown as %) of a logistic regression of the beta-power PRS score, subtracted by the baseline model (covariates only: sex and 4 PCs). -Log10 P-values of Δ explained variance per P_t (red dots) represent P-values from the same logistic regression, with sex and 4 PCs included as covariates.

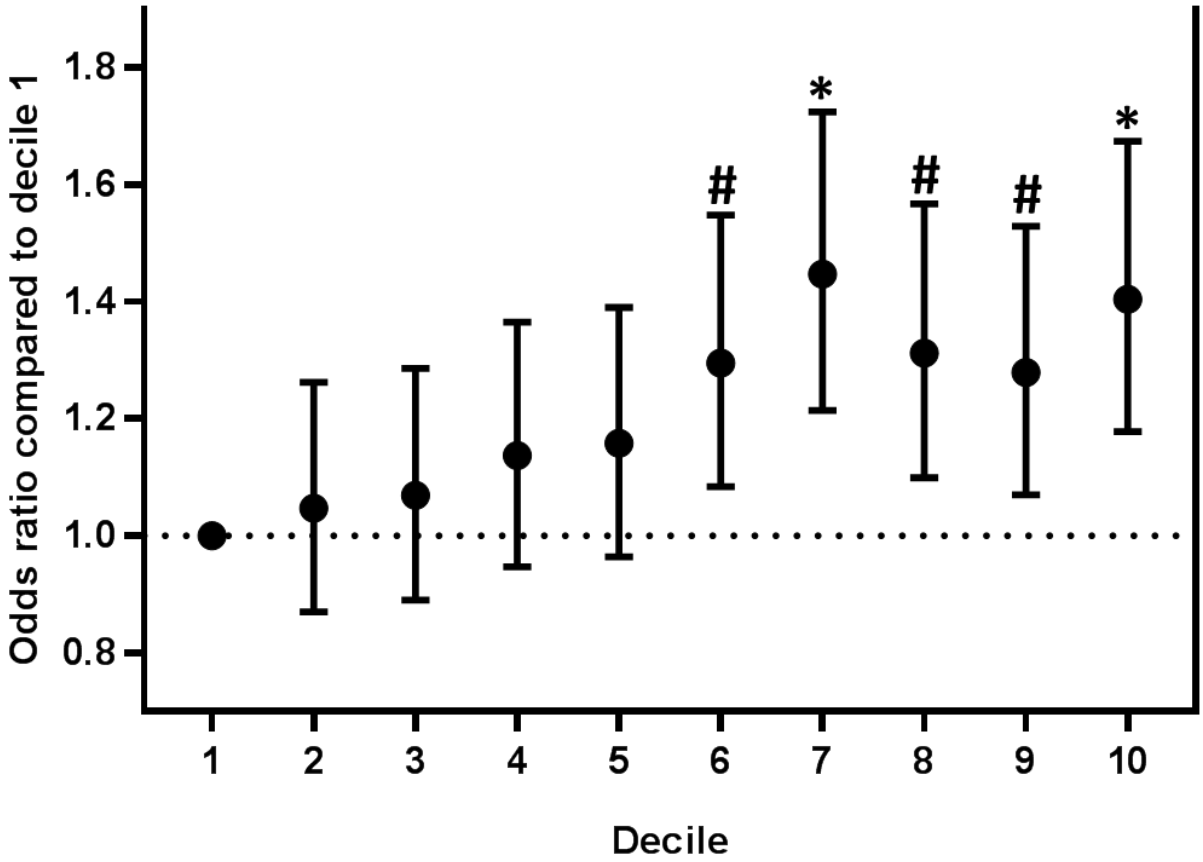
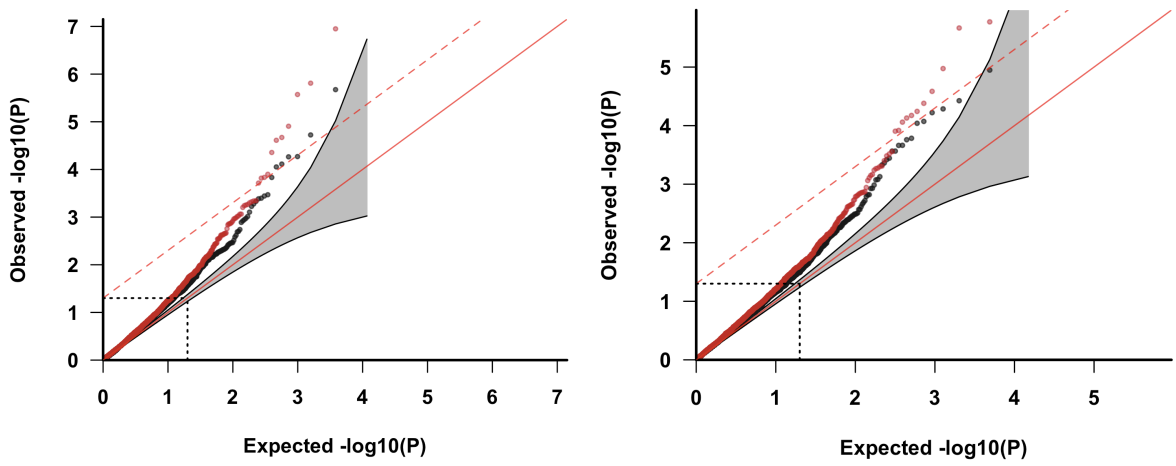


*Figure 2C: Odds ratio for GGE by PRS decile for beta-power oscillations. PRS values with the best-fit P-value threshold (P_t=0.06245) were calculated for 3,122 GGE cases and 22,078 controls. The subjects were distributed into evenly sized quantiles, where decile 1 contains the 10% subjects with the lowest PRS scores and decile 10 contains the 10% subjects with the highest scores. Logistic regression was used to assess whether the odds of having GGE is higher in decile 2 to 10 compared with decile 1 (reference). Sex and 10 PCs were included as covariates. Odds ratios ± 95% confidence intervals are displayed. #: P<0.05; *: P<0.001.*

Meta-analysis, gene-based testing and gene expression analysis

The meta-analysis of the discovery epilepsy GGE subset and EEG GWASs yielded no novel significant loci (Supplemental figures 2 and 3). However, brain tissue enrichment increased relative to the published epilepsy GWAS (International League Against Epilepsy Consortium on Complex, 2018). Caudate and basal ganglia constituted the primary implicated tissues in the meta-analysis (up from the fourth most enriched region in the epilepsy-only GWAS); and 3 out of the 5 most enriched tissues were now brain segments (up from 1 out of 5 in the epilepsy-only GWAS; supplemental Figure 4).

Moreover, using S-Predixcan we found significant associations between genetically imputed expression levels of brain-expressed genes and the GGE + beta oscillations meta-analyzed phenotype for all brain tissues. **Figure 3** shows the Q-Q plots for these associations (hippocampus and frontal cortex are given as examples; Q-Q plots for the other tissues are provided in supplemental figure 5). Finally, relative to the GGE-only GWAS and beta oscillations GWAS, the meta-analysis showed an increased number of significantly associated genes across the brain (Supplemental table 3**).**



A *B*

Figures 3: S-PrediXcan gene-expression effects increased in strength after meta-analyzing the original GGE GWAS (black) with beta power (GGE + beta, red). This resulted in many more significant discoveries than in the original GGE GWAS (FDR-corrected, dashed line; Supplemental table 3). (A) Q-Q plot for hippocampal expression; (B) for frontal cortical expression.

Mendelian Randomization analyses

Eight GGE-associated SNPs were selected as instrumental variables at a strict LD prune threshold (r^2<0.01, LD window =10Mb). These were used in fixed effects IVW, weighted median, MR Egger, and GSMR (r^2<0.01) models. After loosening the LD threshold, 11 (r^2<0.1), 12 (r^2<0.15) and 14 (r^2<0.2) SNPs were selected as instrumental variables for GSMR models. Causal effects of GGE loci on beta oscillations were found at the LD r^2< 0.15 and <0.2 thresholds (OR=1.79, CI=1.189-2.707, P-value= $5.2x10^{-3}$; and OR=1.723, CI=1.180-2.516, P-value=$4.8x10^{-3}$, respectively, **Figure 4**, Supplemental Table 4, Supplemental Figure 5). Significant heterogeneity was detected in the fixed-effect IVW model (Q-statistics=18.188, df=7, p= 0.01) and MR Egger model (Q-statistics=14.594, df=6, p= 0.02). No SNPs altered the pooled β coefficient in the leave-one-out sensitivity analysis (β =0.374, p= 0.314) in the fixed-effect IVW model. We found no evidence of horizontal pleiotropic effects. Similarly, the HEIDI test detected no SNPs as pleiotropic outliers.

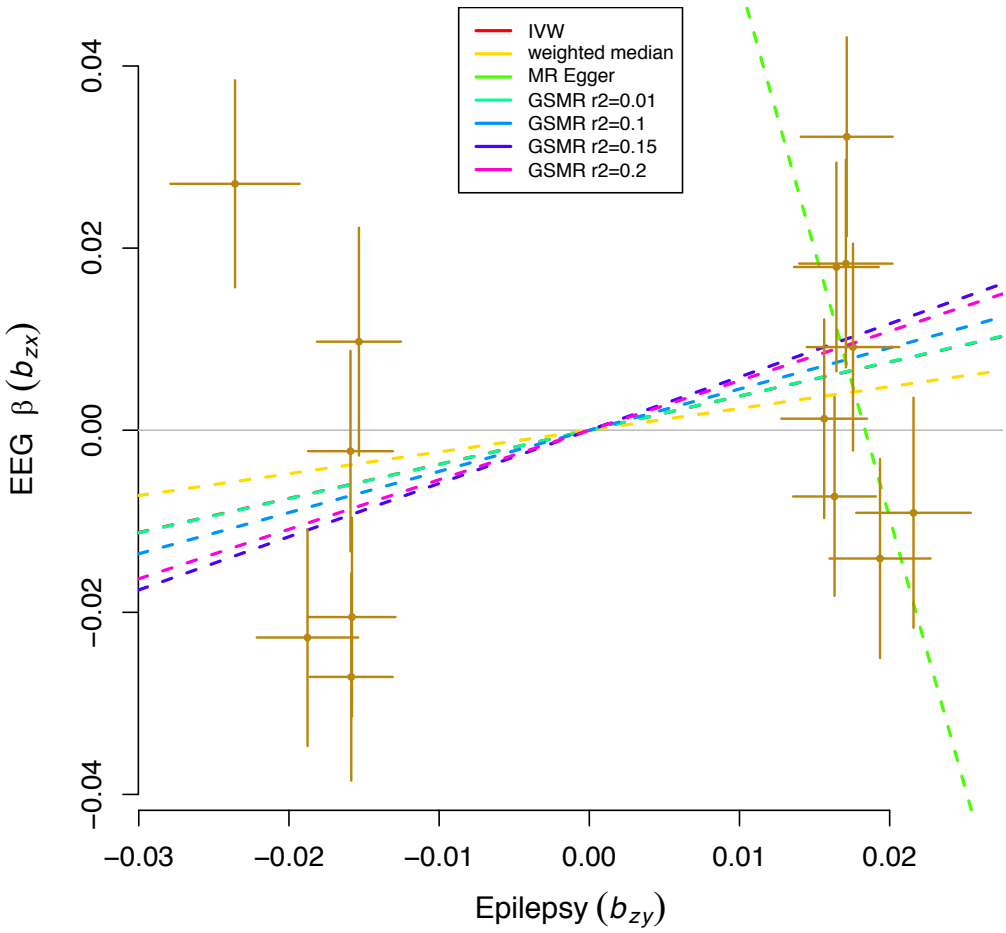


Figure 4. Scatter plot of instrumental SNPs associated with generalized epilepsy tested for EEG beta oscillations. The instrumental variables passed Heterogeneity in Dependent Instruments (HEIDI) testing for instrumental outliers from GSMR detection with multiple LD prune thresholds. Lines in red, yellow, light green and dark green represent β for fixed effect IVW, weighted median, MR Egger, and GSMR models using 8 instruments (r^2<0.01). The lines in blue purple and pink represent βs from GSMR model with 11 (r^2<0.1), 12 (r^2<0.15) and 14 (r^2<0.2) instruments in GSMR.

Discussion

Here we leveraged the largest currently available GWASs of epilepsy and of background EEG oscillations to estimate the genetic correlation between these two traits and conduct follow-up analyses. We demonstrate a strong genetic correlation between beta and theta power with GGE. We replicated this genetic correlation in an independent epilepsy cohort that was statistically significant for beta power. Polygenic risk scoring confirmed our findings and quantified the increase in risk of suffering from genetic generalized epilepsy from the lowest to highest beta oscillations decile. Furthermore, in a meta-analysis of GGE and beta power GWASs we show that gene expression enrichment shifts towards more CNS tissues relative to the discovery epilepsy GWAS. Finally, Mendelian Randomisation analyses pointed to causal effects of genetic liability to GGE on beta power.

Previous studies comparing EEG background oscillations between epilepsy patients and controls are inconsistent: some show increased power in all frequency bands (alpha, beta, delta, theta), whereas others show only increases in specific frequency bands or even decreases in power (Clemens *et al.*, 2000; Clemens, 2004; Clemens *et al.*, 2007; Clemens, 2008; Tikka *et al.*, 2013; Clemens *et al.*, 2014). This heterogeneity likely reflects multiple variables that are difficult to control for in clinical studies, such as AED usage, sleep deprivation, fluctuations in (inter-)ictal epileptic brain activity, EEG processing and electrode placement. We overcame such caveats by determining the genetic underpinnings of EEG frequency bands in people without epilepsy who are AED-naïve, and with a consistent electrode placement and signal processing. We applied several statistical models to assess this overlap and found that people with generalized, but not focal, epilepsy carry a relative abundance of genetic variation associated with higher theta and beta oscillations. The results are consistent with the notion that epilepsy exists on a continuum rather than as a binary state, implying that some individuals are prone to develop epileptic seizures only under certain circumstances (e.g. stroboscopic light), while others relentlessly suffer from seizures despite the use of multiple AEDs. However, prospective studies must confirm whether altered oscillatory activity could be a prodromal state of GGE. Alternatively, aberrant loading of beta oscillations on an EEG may constitute a feature of epilepsy.

We did not find genetic correlations between background EEG and focal epilepsy. Although power was limited for this analysis, this finding is consistent with the low SNP-based heritability of focal epilepsy and the lack of genetic overlap between focal and

generalized epilepsy (International League Against Epilepsy Consortium on Complex Epilepsies, 2018). A likely explanation of the lack of genetic correlation between focal epilepsy and background EEG oscillations may lie in the physiological differences between focal epilepsy and GGE. In contrast to focal epilepsy, the EEG discharges that characterize generalized epilepsy are dependent on the thalomocortical system (Gloor, 1978; Steriade and Contreras, 1998). Background oscillations have also been functionally attributed to the thalamocortical system (Steriade *et al.*, 1996a; Steriade *et al.*, 1996b) and are related to functional states of the brain. Interestingly, fast oscillations have also been implicated in cognitive processes (Gray and Singer, 1989; MacKay and Mendonca, 1995), and are also part of the sleep EEG (Steriade *et al.*, 1996b).

Our results should be interpreted in light of several limitations. First, we are aware of the possible advantages of using Genome Complex Trait Analysis (GCTA) relative to LDSC but since no subject-level genotype data are available for the EEG GWAS, we restricted our genetic correlation estimates to the standard LDSC method, which is based on summary statistics. Nonentheless, LDSC has proven to be a reliable method for genetic correlation estimates and results between LDSC and GCTA have proven consistent. Second, we found that alleles are shared between GGE and background EEG, but preferably prospective studies are needed to establish whether GGE patients have significant differences on their EEGs compared to controls and those with focal epilepsy. Third, only one-way Mendelian Randomisation analyses were performed due to lack of genome-wide significant loci in the EEG GWAS. We therefore cannot exclude the possibility of bidirectional causality between EEG and GGE. Fourth, we had insufficient data available to carry out subgroup analyses on subjects with non-lesional focal epilepsy patients.

Increasingly large EEG GWASs may increase power to detect pleiotropic loci and also empower pathway analyses, thus possibly providing mechanistic insight into the overlap between GGE and background EEG. Future studies may also integrate transcranial magnetic stimulation – EEG (TMS-EEG) and/or event-related potentials to examine whether beta and theta powers correlate with altered brain excitability in subjects with high epilepsy liability. Furthermore, machine-learning studies integrating background EEG with other sources of clinical and demographic data may one day increase diagnostic accuracy of epilepsy. Finally, our findings may be followed up in clinical trials examining whether clinical decisions with a high degree of uncertainty, e.g. about if or when to start an AED in medication-naïve subjects, may be partly informed by background EEG analyses.

Acknowledgements

We are grateful to the patients and volunteers who participated in this research. We thank the following clinicians and research scientists for their contribution through sample collection (cases and controls), data analysis, and project support: Geka Ackerhans, Muna Alwaidh, R E Appleton, Willem Frans Arts, Guiliano Avanzini, Paul Boon, Sarah Borror, Kees Braun, Oebele Brouwer, Hans Carpay, Karen Carter, Peter Cleland, Oliver C Cockerell, Paul Cooper, Celia Cramp, Emily de los Reyes, Chris French, Catharine Freyer, William Gallentine, Michel Georges, Peter Goulding, Micheline Gravel, Rhian Gwilliam, Lori Hamiwka, Steven J Howell, Adrian Hughes, Aatif Husain, Monica Islam, Floor Jansen, Mary Karn, Mark Kellett, Ditte B Kjelgaard, Karl Martin Klein, Donna Kring, Annie WC Kung, Mark Lawden, Jo Ellen Lee, Benjamin Legros, Leanne Lehwald, Edouard Louis, Colin HT Lui, Zelko Matkovic, Jennifer McKinney, Brendan McLean, Mohamad Mikati, Bethanie Morgan-Followell, Wim Van Paesschen, Anup Patel, Manuela Pendziwiat, Marcus Reuber, Richard Roberts, Guy Rouleau, Cathy Schumer, B Sharack, Kevin Shianna, NC Sin, Saurabh Sinha, Laurel Slaughter, Sally Steward, Deborah Terry, Chang-Yong Tsao, TH Tsoi, Patrick Tugendhaft, Jaime-Dawn Twanow, Jorge Vidaurre, Sarah Weckhuysen, Pedro Weisleder, Kathleen White, Virginia Wong, Raju Yerra, Jacqueline Yinger and all contributing clinicians from the Department of Clinical and Experimental Epilepsy at the National Hospital for Neurology and Neurosurgery and UCL Institute of Neurology.

Funding

We would like to thank the Ming Fund for providing funding for R.S. This work was in part supported by an award by a Translational Research Scholars award from the Health Research Board of Ireland (C.D.W.), 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. Further funding sources include: Wellcome Trust (grant 084730); Epilepsy Society, UK, NIHR (08-08-SCC); GIHE: NIH R01-NS-49306-01 (R.J.B.); NIH R01-NS-053998 (D.H.L); GSCFE: NIH R01-NS-064154-01 (R.J.B. and Ha.Ha.); NIH: UL1TR001070, Development Fund from The Children's Hospital of Philadelphia (Ha.Ha.); NHMRC Program Grant ID: 1091593 (S.F.B., I.E.S., K.L.O., and K.E.B.); The Royal Melbourne Hospital Foundation Lottery Grant (S.P.); The RMH Neuroscience Foundation (T.J.O'B.); European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 279062 (EpiPGX) and 602102, Department of Health's NIHR Biomedical Research Centers funding scheme, European Community (EC:

FP6 project EPICURE: LSHM-CT2006-037315); German Research Foundation (DFG: SA434/4-1/4-26-1 (Th.Sa.), WE4896/3-1); EuroEPINOMICS Consortium (European Science Foundation/DFG: SA434/5-1, NU50/8-1, LE1030/11-1, HE5415/3-1 (Th.Sa., P.N., H.L., I.H.), RO 3396/2- 1); the German Federal Ministry of Education and Research, National Genome Research Network (NGFNplus/EMINet: 01GS08120, and 01GS08123 (Th.Sa., H.L.); IntenC, TUR 09/I10 (Th.Sa.)); The Netherlands National Epilepsy Fund (grant 04-08); EC (FP7 project EpiPGX 279062). Research Grants Council of the Hong Kong Special Administrative Region, China project numbers HKU7623/08 M (S.S.C, P.K., L.W.B., P.C.S), HKU7747/ 07 M (S.S.C., P.C.S.) and CUHK4466/06 M (P.K., L.B). Collection of Belgian cases was supported by the Fonds National de la Recherche Scientifique, Fondation Erasme, Université Libre de Bruxelles. GlaxoSmithKline funded the recruitment and data collection for the GenEpA Consortium samples. We acknowledge the support of Nationwide Children's hospital in Columbus, Ohio, USA. The Wellcome Trust (WT066056) and The NIHR Biomedical Research Centres Scheme (P31753) supported UK contributions. Further support was received through the Intramural Research Program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development (Contract: N01HD33348). The project was also supported by the popgen 2.0 network through a grant from the German Ministry for Education and Research (01EY1103). Parts of the analysis of this work were performed on resources of the High Performance Center of the University of Luxembourg and Elixir-Luxembourg. The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ. The International League Against Epilepsy (ILAE) facilitated the Consortium through the Commission on Genetics and by financial support; however, the opinions expressed in the manuscript do not necessarily represent the policy or position of the ILAE.

Competing interests

The authors declare no competing interests.

References

Barbeira AN, Dickinson SP, Bonazzola R, Zheng J, Wheeler HE, Torres JM*, et al.* Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. Nat Commun 2018; 9(1): 1825.

Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C*, et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet 2015; 47(3): 291-5.

Burgess S, Scott RA, Timpson NJ, Smith GD, Thompson SG, Consortium E-I. Using published data in Mendelian randomization: a blueprint for efficient identification of causal risk factors. European Journal of Epidemiology 2015; 30(7): 543-52.

Clemens B. Pathological theta oscillations in idiopathic generalised epilepsy. Clinical Neurophysiology 2004; 115(6): 1436-41.

Clemens B. Valproate decreases EEG synchronization In a use-dependent manner in idiopathic generalized epilepsy. Seizure-European Journal of Epilepsy 2008; 17(3): 224-33.

Clemens B, Bessenyei M, Piros P, Toth M, Seress L, Kondakor I. Characteristic distribution of interictal brain electrical activity in idiopathic generalized epilepsy. Epilepsia 2007; 48(5): 941-9.

Clemens B, Puskas S, Besenyei M, Kovacs NZ, Spisak T, Kis SA*, et al.* Valproate treatment normalizes EEG functional connectivity in successfully treated idiopathic generalized epilepsy patients. Epilepsy Research 2014; 108(10): 1896-903.

Clemens B, Szigeti G, Barta Z. EEG frequency profiles of idiopathic generalised epilepsy syndromes. Epilepsy Research 2000; 42(2-3): 105-15.

Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E*, et al.* Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. Nat Genet 2019; 51(1): 63-75.

Devinsky O, Vezzani A, O'Brien TJ, Jette N, Scheffer IE, de Curtis M*, et al.* Epilepsy. Nature Reviews Disease Primers 2018; 4.

Euesden J, Lewis CM, O'Reilly PF. PRSice: Polygenic Risk Score software. Bioinformatics 2015; 31(9): 1466-8.

Gamazon ER, Wheeler HE, Shah KP, Mozaffari SV, Aquino-Michaels K, Carroll RJ*, et al.* A gene-based association method for mapping traits using reference transcriptome data. Nat Genet 2015; 47(9): 1091-8.

Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM*, et al.* A global reference for human genetic variation. Nature 2015; 526(7571): 68-74.

Gloor P. Generalized epilepsy with bilateral synchronous spike and wave discharge. New findings concerning its physiological mechanisms. Electroencephalogr Clin Neurophysiol Suppl 1978(34): 245-9.

Gloor P. Generalized epilepsy with spike-and-wave discharge: a reinterpretation of its electrographic and clinical manifestations. The 1977 William G. Lennox Lecture, American Epilepsy Society. Epilepsia 1979; 20(5): 571-88.

Gloor P, Fariello RG. Generalized epilepsy: some of its cellular mechanisms differ from those of focal epilepsy. Trends Neurosci 1988; 11(2): 63-8.

Gray CM, Singer W. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. Proc Natl Acad Sci U S A 1989; 86(5): 1698-702.

Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D*, et al.* The MR-Base platform supports systematic causal inference across the human phenome. Elife 2018; 7.

International League Against Epilepsy Consortium on Complex E. Genome-wide mega-analysis identifies 16 loci and highlights diverse biological mechanisms in the common epilepsies. Nat Commun 2018; 9(1): 5269.
MacKay WA, Mendonca AJ. Field potential oscillatory bursts in parietal cortex before and during reach. Brain Res 1995; 704(2): 167-74.
Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. Bioinformatics 2010; 26(22): 2867-73.
Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D*, et al.* PLINK: A tool set for whole-genome association and population-based linkage analyses. American Journal of Human Genetics 2007; 81(3): 559-75.
Smit DJA, Posthuma D, Boomsma DI, De Geus EJC. Heritability of background EEG across the power spectrum. Psychophysiology 2005; 42(6): 691-7.
Smit DJA, Wright MJ, Meyers JL, Martin NG, Ho YYW, Malone SM*, et al.* Genome-wide association analysis links multiple psychiatric liability genes to oscillatory brain activity. Human Brain Mapping 2018; 39(11): 4183-95.
Steriade M. Corticothalamic resonance, states of vigilance and mentation. Neuroscience 2000; 101(2): 243-76.
Steriade M. Sleep, epilepsy and thalamic reticular inhibitory neurons. Trends Neurosci 2005; 28(6): 317-24.
Steriade M, Amzica F, Contreras D. Synchronization of fast (30-40 Hz) spontaneous cortical rhythms during brain activation. J Neurosci 1996a; 16(1): 392-417.
Steriade M, Contreras D. Spike-wave complexes and fast components of cortically generated seizures. I. Role of neocortex and thalamus. J Neurophysiol 1998; 80(3): 1439-55.
Steriade M, Contreras D, Amzica F, Timofeev I. Synchronization of fast (30-40 Hz) spontaneous oscillations in intrathalamic and thalamocortical networks. J Neurosci 1996b; 16(8): 2788-808.
Thomas RH, Berkovic SF. The hidden genetics of epilepsy-a clinically important new paradigm. Nature Reviews Neurology 2014; 10(5): 283-92.
Tikka SK, Goyal N, Umesh S, Nizamie SH. Juvenile myoclonic epilepsy: Clinical characteristics, standard and quantitative electroencephalography analyses. J Pediatr Neurosci 2013; 8(2): 97-103.
Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 2011; 88(1): 76-82.
Zhu ZH, Zheng ZL, Zhang FT, Wu Y, Trzaskowski M, Maier R*, et al.* Causal associations between risk factors and common diseases inferred from GWAS summary data. Nature Communications 2018; 9.