SUPPLEMENTARY FIGURES

Figure S1. Manhattan plot for SNP effects for delta, theta, alpha, and beta EEG power at the vertex (Cz), occipital alpha power, and occipital alpha peak frequency. Dashed line is the standard threshold for genome-wide significance (5·10⁻⁸). Two genome-wide SNP effect were observed for Cz alpha power. On chr 4 rs984924 (p=4.9·10⁻⁸) is a SNP intronic to the protein kinase cGMP-dependent type II (PRKG2) gene. On chr 7 rs10231372 (P=2.9·10⁻⁸) is an intron variant in the long non-coding RNA gene LINC00996.

Figure S2. Regional association/LD plots. Chromosome 3 showed a range of low p-values and significant genes for Cz (and occipital) alpha power from about 52.2 to 52.8 Mbp (hg19/b37). (top) Cz alpha regional association plot with color-coded LD shows that the cluster of large p-values may be partially due to LD in an area spanning from ALAS1 to NEK4. SNP rs7614727 showed top association. Evidence for multiple causal variants may be found in the relatively high –log10(p) values for SNPs with moderate (green) and low (blue) LD with the top SNP in the 52.5 to 52.8 Mb region. (bottom) the same region was implicated for bipolar disorder and schizophrenia (48, 49). The regional association plot for schizophrenia showed LD and p-value patterns similar to Cz alpha power, and also evidence for multiple independent genetic effects (e.g., low p-values in the 52.5 to 52.8 region in low/moderate LD with rs7614727, and low p-values above the 52.8 Mb position with very low LD).

Figure S3. Cz alpha Q-Q plots revealed strong inflation for most GTEx brain tissue eQTLs. FDR-corrected significant SNP effects are circled. eQTLs reached FDR significant discovery for frontal and anterior cingulate cortex, and hypothalamus. The genes associated with these eQTLs were ITIH4, GLN3, and MTERF4 for the frontal cortex, and MTERF4 for the hypothalamus and anterior cingulate. Red dashed line is FDR q=0.05.

Figure S4. Occipital alpha quantile-quantile plots for GTEx brain tissue eQTLs (as in figure S3). FDR discoveries were eQTLs for immune genes IL18R1 (Hippocampus, Caudate), IL1RL1 (Hippocampus, Caudate, Nucleus Accumbens), as well as CLHC1 (Caudate, Nucleus Accumbens), ITIH4 (Hippocampus, Caudate) and GLYCTK (Hippocampus). Red dashed line is FDR q=0.05.

Figure S5. LD score regression based SNP heritability was significant for EEG power phenotypes, not for occipital peak frequency. A relatively low proportion of total twin/family
based heritability estimates are captured by common genetic variants. Error bars represent 95% CI around the null hypothesis.

Figure S6. Bivariate LD score regression estimates of the genetic correlation $r_G$ heatmaps (left) and associated p-values (right; white is not significant) for the EEG phenotypes. Strong genetic overlap existed between low frequency power (theta and delta), and between occipital alpha, vertex alpha, and beta power. All power phenotypes showed highly significant $r_G$ except between theta and beta power. Peak frequency showed significant negative genetic correlations with delta and theta power.

Figure S7. Genetic correlations comparing results with other GWAS results by means of LD hub. No significant results were obtained after multiple-tests correction. Nominal significance was obtained for reduced theta power and autism/chronotype, Cz alpha power and ever smoked/heart rate/chronotype, occipital alpha power and neuroticism, beta power and generalized epilepsy/college completed/pallidum volume/chronotype, occipital frequency and femoral neck BMD/ever smoked. The nominally significant $r_G$ between beta power and epilepsy is consistent with GABA influencing both (41, 42, 66, 94). The nominally significant $r_G$ between beta power and globus pallidus volume is consistent with both phenotypes being associated with the regulation of (voluntary) movement, the involvement of dopamine pathways, and the links of both to symptoms in Parkinson’s disease (8, 58–60, 121–123). BMD=Bone Mineral Density.
Figure S3: Cz alpha eQTL Q-Q plot
Figure S4: Occipital alpha eQTL Q-Q plot
Figure S5: LD score regression SNP heritability
Figure S6: Genetic correlations between EEG parameters
Figure S7: Genetic correlations with published GWAS results
Supplementary Methods

PART I: Sample description and EEG phenotyping

la NTR

Subjects
The EEG data were collected as part of multiple longitudinal studies into the genetics and development of brain function in twin family cohorts (1–3). The studies were approved by the medical ethical review board. Informed consent was obtained from all subjects (in the case of adults), or their parents/guardians. All experiments were conducted in accordance with the Declaration of Helsinki.

A total number of 839 subjects (495 and 398 females and males respectively) had resting EEG and genomics data available out of 1675 twins and siblings who accepted an invitation for extensive EEG measurement and behavioral testing. Subjects were invited in mixed cross-sectional/longitudinal waves (mean ages ~5, ~7, ~16, ~18, ~25, and ~50 years; age range 5.2 to 71.0 years). For the childhood and adolescent age groups only twins participated, while for the adults, their singleton siblings were additionally invited. Between 5–7 and between 16–18 years consisted of longitudinal measurements.

EEG acquisition and preprocessing
Due to the long time-span of the project, three different EEG systems were used. The childhood (5 and 7) and adolescent (16 and 18) EEG were recorded with tin electrodes in an ElectroCap connected to a Nihon Koden PV-441A polygraph with time constant 5 s and lowpass of 35 Hz, digitized at 250 Hz using an in-house built 12-bit A/D converter board. Leads were Fp1, Fp2, F7, F3, F4, F8, C3, C4, T5, P3, P4, T6, O1, O2, and bipolar horizontal and vertical EOG derivations. Electrode impedances were kept below 5 kOhm. Following the recommendation by Pivik et al. (4), tin earlobe electrodes (A1, A2) were fed to separate high-impedance pre-amplifiers, after which the electrically linked output signals served as reference to the EEG signals. Sine waves of 100 uV were used for calibration of the amplification/AD conversion before measurement of each subject.

Young adult and middle-aged cohort EEG was recorded with Ag/AgCl electrodes mounted in an ElectroCap. Signal registration was conducted using an AD amplifier developed by Twente Medical Systems (TMS; Enschede, The Netherlands) for 657 subjects and NeuroScan SynAmps 5083 amplifier for 103 subjects. Standard 10-20 positions were F7, F3, F1, Fz, F2, F4, F8, T7, C3, Cz, C4, T8, T5, P3, Pz, P4, T6, O1 and O2. For subjects measured with NeuroScan also Fp1, Fp2, and Oz were recorded. The vertical electro-oculogram (EOG) was recorded bipolarly between two Ag/AgCl electrodes, affixed one cm below the right eye and one cm above the eyebrow of the right eye. The horizontal EOG was recorded bipolarly between two Ag/AgCl electrodes affixed one cm left from the left eye and one cm right from the right eye. An Ag/AgCl electrode placed on the forehead was used as a ground electrode. Impedances of all EEG electrodes were kept below 5 kOhm, and impedances of the EOG electrodes were kept below 10 kOhm.

EEG data consisted of 3–4 minutes of eyes-closed rest recordings. All EEG was sampled at 250 Hz. All EEG was individually inspected. Bad channels or channels with excessive artifacts were removed. Episodes with artifacts and suspected drowsiness (and even sleep) were removed. EEG signals were broadband filtered 1–37 Hz followed by an
ENIGMA-EEG Genome-wide association

extended ICA decomposition using the EEGLAB toolbox in MATLAB (5–7). Components reflecting blinks and eye movement were identified based on visual inspection of the IC trace, high loading onto frontal electrodes, and high correlation with EOG signal, and removed.

Analysis proceeded with power analysis for channels Cz, O1, and O2; and peak frequency analysis for O1 and O2. Conform (8) O1 and O2 were rereferenced against T5 and T6 respectively. The average of C3 and C4 replaced Cz power in the childhood/adolescent subjects, correlating .95 with Cz power in the other subject groups. After cleaning, 1349 EEG recordings were available in 835 subjects (Cz, 463 female) and 834 subjects (O1/O2, 463 female). A minimum of 120 seconds artifact free was required for power estimations. Of the 1625 subject, 30 subjects (all adults) recordings were not completed. EEG cleaning of the remaining 1595 removed a further 95 subjects’ data (6.0%).

Figure. Age histogram for the NTR sample (N=1349 recordings including longitudinal data).

I.b COGA

Subjects
COGA recruited DSM-III-R and DSM-IV AD probands from inpatient and outpatient treatment facilities through six participating sites: State University of New York Downstate Medical Center, University of Connecticut Health Science Center, Indiana University School of Medicine, University of Iowa School of Medicine, University of California School of Medicine, and Washington University School of Medicine. Recruitment and assessment procedures, including a clinical interview, neurophysiological assessments and DNA collection have been described previously (Begleiter et al., 1995; Foroud et al., 2000). All participants (proband, family members, and community controls) were administered the Semi-Structured Assessment for the Genetics of Alcoholism (SSAGA), a poly-diagnostic interview (Bucholz et al. 1994; Hesselbrock et al. 1999). Individuals below the age of 18 were administered an adolescent version of the SSAGA. The laboratory and data-collection procedures were identical at each of the sites (Begleiter et al., 1998). Institutional review boards at all sites approved the study.
The analytic sample for the current study consisted of families of European Ancestry (EA fGWAS) with both resting state EEG and GWAS data available: 1,492 individuals (783 females and 709 males; average age: 31.12) from 117 multi-generational families affected with AD. Family sizes ranged from 4 to 39 individuals with an average of 13.4 individuals (with EEG data) per family. In addition, an unrelated case-control sample of alcohol dependent cases and controls (ccGWAS) were also collected, and have been described previously (Edenberg et al., 2010). Both resting state EEG and GWAS data were available on 1,128 unrelated individuals. Subjects included in the ccGWAS who were also part of the EA fGWAS were excluded from the ccGWAS sample, leaving an independent EA ccGWAS sample of 660 unrelated individuals (319 females and 341 males; average age: 42.24) for analysis. Identical phenotypic definitions, covariates, SNP QC standards, MAF thresholds and imputation protocols were used to process the samples.

**EEG acquisition and preprocessing**
Prior to neurophysiological assessments, participants were required to have abstained from alcohol for a minimum of 3 weeks. Individual were excluded from neurophysiological assessment if they had any of the following: (1) recent substance or alcohol use (i.e., positive breath-analyzer test); (2) hepatic encephalopathy/cirrhosis of the liver; (3) significant history of head injury, seizures or neurosurgery; (4) uncorrected sensory deficits; (5) taking medication known to influence brain functioning; and (6) other acute/chronic medical illnesses that affect brain function.

Participants were seated comfortably in a dimly lit sound-attenuated temperature-regulated booth (Industrial Acoustics, Bronx, NY). They were instructed to keep their eyes closed and remain relaxed, but to not fall asleep. EEG data were collected in the awake, eyes-closed condition for 4.25 minutes. Each participant wore a fitted electrode cap (Electro-Cap International, Eaton, OH) using either the 19-channel montage as specified according to the 10-20 international system (Supplementary Figure 1) or montages with a greater number of electrodes (32 or 64). The nose was used as a reference, and a forehead electrode served as the ground electrode. Electrode impedances were maintained below 5 kΩ. Electrical activity was amplified 10,000 times by Sensorium (Charlotte, VT) EPA-2 electrophysiology amplifiers with either a bandpass between 0.02 and 50 Hz and digitized on a Concurrent (Atlanta, GA) 5550 computer at a sampling rate of 256 Hz or a band pass between 0.02 Hz and 100.0 Hz on a Neuroscan system (Version 4.1 to 4.5) (Neurosoft, Inc., El Paso, TX) at sampling rates of 500 Hz or 512 Hz. All six collection sites used identical experimental procedures and EEG acquisition hardware and software programs.

A continuous interval comprising 256 seconds of eyes-closed resting EEG data was analyzed from the CZ, O1, and O2 electrodes. Data was detrended and filtered using a bandpass FIRLS filter spanning .2 Hz to 40 Hz. One second intervals were checked for artifact and absence of signal before being analyzed according to the ENIGMA-EEG protocol which called for power in the delta, theta, alpha, and beta bands and peak frequency in the alpha band obtained from overlapping two second intervals by Fourier transform methods. A minimum of 128 seconds of artifact-free data was required for the inclusion of the subject in the study.
About 5% of the candidate subjects were excluded as a result of excessive artifact or problems in the recording of the data.

Figure. Age histogram for the COGA (a) European Ancestry family sample N=1,492, and (b) the Case-Control (CC, unrelated) sample N=1,128.

I.c MTFS

Subjects
EEG data were collected in two independent age cohorts of twins and their parents being studied in large, ongoing longitudinal studies of the Minnesota Twin Family Study (MTFS) (9, 10). EEG recordings for twins was from their age-17 assessment, representing the intake assessment of one cohort and the second follow-up of the other cohort. EEG data were also collected from the parents, who visited when the twins were minors. Informed consent was obtained from the parents regarding their children’s participation as well as their own. Assent was obtained from minor children, consent from those at least 18 years old. The sample consisted of 4,026 subjects with usable EEG data, 2,056 males and 1,970 females. There were 2,383 adolescents (1,230 females; mean age, 17.7; sd, 0.5) and 1,643 parents (740 females; mean age, 45.1; sd, 5.7). The sample for molecular genetic studies consisted of
ENIGMA-EEG Genome-wide association

white subjects, who make up the vast majority of the larger sample, to avoid allelic stratification.

Figure. Age distribution of the MTFS EEG recordings in two age cohorts (N = 4,026)

EEG acquisition and preprocessing

Data were collected over the course of more than 20 years using two different recording systems. For participants in the MTFS older and younger cohorts (71% of the sample), EEG was recorded using Ag/AgCl electrodes in an electrode cap via Grass Neurodata 12 systems (128-Hz sampling rate, pass-band from 1 to 30 Hz with a roll-off of 6 dB). For all subjects, EEG data were collected in the awake, eyes-closed condition for 5 minutes. EEG was recorded from bipolar derivations O1–P7 and O2–P8 and from Cz referenced to linked earlobes. The recording montage came to include F3 and F4, but these leads were not available for all subjects (N = 958 without). A transverse electrode arrangement, with one electrode superior to the eye and one on the outer canthus, served to record eye blinks and other eye movements. EEG was recorded from participants in the ES sample from a BioSemi ActiveTwo system, which uses an active electrode technology, with 61 scalp electrodes arrayed following the 10/10 convention (Fp1, Fpz, Fp2, AF8, AF4, AFz, AF3, AF7, F7, F5, F3, F1, Fz, F2, F4, F6, F8, FT8, FC6, FC4, FC2, FCz, FC1, FC3, FC5, FT7, T7, C5, C3, C1, Cz, C2, C4, C6, T8, CP6, CP4, CP2, CPz, CP1, CP3, CP5, TP7, P7, P5, P3, P1, Pz, P2, P4, P6, P8, PO8, PO4, POZ, PO3, PO7, O1, Oz and O2) and four EOG electrodes (superior to and inferior to the right eye, lateral to each eye), was used to collect continuously recorded EEG data with a sampling rate of 1024 Hz. ActiveTwo amplifiers are DC-coupled. All signals are monopolar and were low-pass-filtered using a digital 5th-order Bessel antialiasing sinc filter with a cutoff frequency (3-dB attenuation) of 205 Hz.

The electrodes available for the original Grass system were extracted from Biosemi recordings and were transformed to be similar to the Grass data. The preprocessing procedure is described in detail in Malone et al. (8). All data were subsequently processed using the identical approach. EEG segments containing transient artifacts and excessively small or large voltage deflections were tagged for exclusion by computer algorithm. Multivariate outliers with respect to power in all bands and all leads were identified using
ENIGMA-EEG Genome-wide association

the robustbase package (Rousseeuw et al., 2011; robustbase: Basic Robust Statistics (Version R package version 0.7-3) in the R statistical programming environment (R Development Core Team, 2010). If a subject’s robust Mahalanobis distance fell in the upper 10th percentile of the cumulative chi-square distribution, data from this subject were flagged for review and excluded if visual inspection of the raw data indicated that the data were contaminated by artifact. Individual recording sites were excluded from analyses if fewer than 45 artifact-free sweeps were available.

Based on notes recorded when the data were originally collected, 110 subjects were excluded and problematic channels dropped. An additional 68 subjects were excluded for use of alcohol, marijuana or other illicit drug the day of the assessment; a history of serious head injury; neurological disorders; medication likely to affect psychophysiological responses; not refraining from taking medication for ADHD, such as methylphenidate, the day of their assessment, as was requested of all adolescent participants. Finally, 252 subjects who reported sleeping during the session or who were noted to have fallen asleep, were also excluded (cf. Malone et al., 2014; 10.1111/psyp.12344). Out of a total of 4,450 subjects eligible for this investigation, 427 subjects were excluded due to recording- or subject-related issues. One hundred sixty-five subjects were previously eliminated as a result of QA screens of DNA samples (genotype quality assurance steps and procedures are described thoroughly elsewhere; Miller et al., 2012).

I.d QIMR
Participants
EEG data was collected from adolescent twins and their singleton siblings as a component of the Cognition Study – one of a series of studies conducted under the umbrella of the Brisbane Adolescent Twin Study (11, 12). Exclusion criteria were parental report of head injury, neurological or psychiatric illness, substance abuse or dependence, and current use of medication with known effects on the central nervous system. All participants were right-handed. Written, informed consent was obtained from all participants, including a parent or guardian for those aged less than 18 years. The study was approved by the Human Research Ethics Committee at the QIMR Berghofer Medical Research Institute. Resting EEG was available for 1209 subjects, of whom genomics data were available for 971 participants (498 females and 473 males). The distribution of age in the cohort is shown in the figure below.
**EEG acquisition and preprocessing**

Resting EEG comprised two 4 minute recordings (eyes closed, following by eyes open) in a semi-darkened, electrically shielded, and sound-attenuated cubicle. EEG was recorded from 15 scalp locations (Fp1, Fp2, Fz, F3, F4, F7, F8, Cz, C3, C4, Pz, P3, P4, O1, O2) using an electrode cap with tin electrodes arranged according to the International electrode (10-20) placement system, and referenced to physically linked ears, with the ear impedances matched at the beginning of the recording session. The ground lead was located just anterior to the Fz electrode. Ocular potentials (electro-ocularogram or EOG) were recorded from single tine electrodes and were located on the outer canthus and the centre of the supraorbital ridge above the left eye. Impedance readings were all below 5 k\(\Omega\). EOG, Fp1 and Fp2 were amplified with a factor 5K and all other channels with a factor 20K by Grass preamplifiers (model P511K), and recordings were filtered with a band pass filter of 0.01 to 30 Hz (6dB per octave) and a 50 Hz notch filter.

Software controlling the recording determined that the maximum length of continuously recorded EEG was 12s with a discontinuity of 2s between successive 12s blocks. Twenty 12s blocks were recorded for both eyes closed and eyes open conditions. All EEG was resampled at 250 Hz, and individually visually inspected. Bad channels or channels with excessive artifacts were removed. Episodes with artifacts and suspected drowsiness (and even sleep) were removed. EEG signals were broadband filtered 1–37 Hz followed by an extended ICA decomposition using the EEGLAB toolbox in MATLAB (5–7). Components reflecting blinks and eye movement were identified based on visual inspection of the IC trace, high loading onto frontal electrodes, and high correlation with EOG signal, and removed.

Analysis proceeded with power analysis for channels Cz, O1, and O2; and peak frequency analysis for O1 and O2. Conform (8) O1 and O2 were rereferenced against T5 and T6 respectively. A minimum of 120 s was required for power calculations. After cleaning 1117 out of 1209 subjects remained (7.6% drop out).

**PART II: Genome-wide SNP arrays and imputation**

**II.a NTR**

**Genotyping**

DNA was isolated with the GENTRA Puregene kit (13). Genotyping was done on multiple arrays with a number of overlapping samples following manufacturer protocols. The following arrays were used: Affymetrix Perlegen 5.0 (N=1,718), Illumina 370 (N=424), Illumina 660 (N=1,103), Illumina Omni Express 1 M (N=346) and Affymetrix 6.0 (N=3602). Genotype calls were made with the platform specific software Birdsute, APT-Genotyper and Beadstudio. Sample and SNP QC was done first within, and then between platforms using the PLINK software (14). With the LiftOver tool ("http://genome.sph.umich.edu/wiki/LiftOver") the individual SNP markers were lifted over to build 37 (HG19) of the Human reference genome for each platform. All data were then strand aligned with the 1000 Genomes GIANT phase1 release v3 20,101,123 SNPs INDELS SVS ALL reference panel. SNPs were removed if they had ambiguous locations, mismatching alleles with the reference set or the allele frequencies differed more than 0.20 compared to
the reference. Per platform SNPs were excluded if any of the following criteria were met: a Minor Allele Frequency (MAF) <1%, Hardy–Weinberg Equilibrium (HWE) with p < 0.00001, and call rate <95%. Samples were excluded when their expected sex did not match their genotyped sex, when the genotype missing rate was >10% or the F inbreeding value was >0.10 or <−0.10. After these steps, the data of the individual arrays were merged into a single dataset. Within this set, identity by state (IBS) sharing was calculated between all possible pairs of participants and compared to the known NTR family structures. Samples were removed if the data did not match their expected IBS sharing. The concordance rate of DNA samples on multiple platforms for overlapping SNPs generally exceeded 99.0% after data cleaning. The HWE, MAF- and the reference allele frequency difference <0.20 filters were subsequently re-applied in the combined data. SNPs with C/G and A/T allele combinations were removed when the MAF was between 0.35 and 0.50 to avoid incorrect strand alignment. Phasing of all samples and imputing cross-missing platform SNPs was done with MACH1(15). The phased data were then imputed with MINIMAC(16) in batches of around 500 individuals for the autosomal genome using the 1000G reference panel for 561 chromosome chunks, which were obtained by the CHUNKCHROMOSOME program (17). To avoid issues having SNPs from different platforms partly imputed and partly genotyped we took the re-imputed calls for all genotyped SNPs. After imputation, we generally find a high concordance between re-imputed SNPs and the original genotypes of around 0.9868. The mean imputation quality R² metric is 0.38, based on all 30,051,533 imputed autosomal SNPs. After imputation, SNPs were filtered based on the Mendelian error rate in families, which was calculated from the best guess genotypes in families (trios or sib-pairs with parents) using first GTOOL to calculate best guess genotypes and then PLINK to analyze the data. SNPs were removed if the Mendelian error rate >0.02.

GWA Statistics

Association was performed using R on 7,335,750 SNPs further filtered for, call rate .90, imputation R²>0.3. Generalized Estimating Equations (GEE) with the independence working correlation matrix was used to correct for dependency within subject and between family members including MZ and DZ twins (18). Three geographically-related PCs (19) plus 10 PCs reflecting stratification plus 9 PCs reflecting SNP batch effects were used to correct for spurious effects. Further filtering is described in the main text.

II.b COGA

Genotyping, Imputation and Quality Review

Genotyping was performed at the Center for Inherited Disease Research (CIDR) using the Illumina (Illumina, San Diego, CA, USA) OmniExpress array for the EA family sample and the Illumina 1M array for the case-control sample. COGA’s quality control (QC) approach has been previously reported (20). Briefly, individuals with a genotype rate <98% were excluded from analysis, and SNPs with a genotyping rate <98% were excluded from analysis. The 795 genotyped founders were used to remove SNPs which violated Hardy-Weinberg equilibrium (HWE; p<10⁻⁶). SNPs with minor allele frequency (MAF) less than 3% in the founders were also removed from further analysis. The reported pedigree structure was assessed using a pruned set of 1,519,440 SNPs. Pairwise identity by descent estimates were computed in
ENIGMA-EEG Genome-wide association

PLINK (14) to detect pairs of individuals whose allele sharing was not consistent with the reported family relationship. Family structures were altered as needed, and then SNP genotypes were tested for Mendelian inconsistencies (Pedcheck) (21) with the revised family structure. The cleaned genotype data were imputed to 1000 genomes (EUR and AFR, Phase 3, b37, October 2014) with build hg19 using SHAPEIT (https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html) and IMPUTE2 (https://mathgen.stats.ox.ac.uk/impute/impute_v2.html). To avoid ambiguities in strand designation, SNPs with A/T or C/G alleles were removed. After imputation, genotype probabilities ≥ 0.90 were changed to genotypes. Mendelian errors in the imputed SNPs were reviewed and resolved as described in Wetherill et al., 2015 (20). All SNPs with imputation genotyping rate < 98% and MAF < 0.03 were excluded from association analyses. Further filtering is reported in the main text.

GWA Statistics
Association was performed on 12,972,748 SNPs pre-filtered for HWE 1E-6, call rate .90, imputation R²>0.3. Generalized Estimating Equations (GEE) with the independence working correlation matrix was used to correct for dependency within subject and between family members (18). Sex, log-transformed age at the time of EEG recording, and 10 PCs reflecting stratification were used as covariates. Given COGA’s study design and the previous associations observed among EEG and alcohol dependence, analyses also included DSM-IV alcohol dependence as a covariate.

II.c MTFS
Genotyping
Genotyping was conducted using the Illumina 660W Quad array. The pipeline for extracting and processing DNA and for quality control has been described in detail (22). Problematic markers identified by Illumina were dropped. In addition, the following filters were used: call rate less than 99%; mismatches in duplicate samples; minor allele frequency (MAF) than 1%; a significant deviation from Hardy-Weinberg genotype frequencies or Mendelian inconsistencies across families; a significant association with participant gender or processing batch; excessive heterozygous calls for markers on the X chromosome in males or in mitochondrial DNA in the sample. The 527,829 markers surviving quality control filters were imputed using minimac (16) after genotypes had been phased in Beagle (23), which uses known familial structure to improve phasing accuracy, and 9,331,500 SNPs imputed using 1000 Genomes reference haplotypes. The MAF exceeded .001 and the imputation r² exceeded .30 for 9,160,194 SNPs.

GWA Statistics
Genome-wide association was conducted on these SNPs using rapid feasible generalized least squares (RFGLS) implemented in an R package developed for molecular genetic analysis of family data (24). RFGLS is a computationally efficient form of generalized least squares, which accommodates correlated data (25). The matrix of within-family correlations was estimated separately for MZ and DZ twin families. To minimize the number of parameters to be estimated, parent-offspring correlations were constrained across sex of
ENIGMA-EEG Genome-wide association

parent and were equal for both twins. Variances were constrained equal for both twins and were identical for MZ and DZ twins. Step-parents were considered independent observations. Chronological age, age cohort (adolescent or parent), sex, lab (Biosemi or Grass), and 10 PCs derived using EIGENSTRAT to control for residual ethnic stratification in the otherwise white sample served as covariates. Accounting for the distinct age groups (adolescents and their middle-aged parents) obviated the need to account for anything other than a linear effect of age.

II.d QIMR

Genotyping
Participants were genotyped on the Illumina Human610-Quad SNP chips. These samples were genotyped in the context of a larger genome-wide association project that resulted in the genotyping of 15,945 individuals using the Illumina 317, 370, 610, 660 SNP chips which included data from twins, their siblings and their parents. Genotype data were screened for genotyping quality (GenCall < 0.7), SNP and individual call rates (< 0.95), HWE failure (P < 10^-6) and MAF (< 0.01). As these samples were genotyped in the context of a larger project, the data were integrated with the larger QIMR genotype project and the data were checked for pedigree, sex and Mendelian errors and for non-European ancestry. Individuals were imputed to the 1000 Genomes Project phase 3 version 5 references using a set of SNPs common to the first generation genotyping platforms (N ~ 278,000). Imputation was performed using MACH.

GWA Statistics

GWAS analyses were conducted using merlin-offline which uses known pedigree and zygosity information to explicitly correct for relatedness (http://www.sph.umich.edu/csg/abecasis/Merlin/). Analyses included 10 Ancestry based principal components and corrections for Sex, Age and Age^2.

PART III: Covariate effects

NTR
ENIGMA-EEG Genome-wide association

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>peakOcc</th>
<th>alphaOcc</th>
<th>deltaCz</th>
<th>thetaCz</th>
<th>alphaCz</th>
<th>betaCz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>0.027</td>
<td>0.082</td>
<td>0.008</td>
<td>7.61E-08</td>
<td>2.62E-06</td>
<td>0.043</td>
</tr>
<tr>
<td>Age*</td>
<td>0.804</td>
<td>0.653</td>
<td>0.775</td>
<td>0.590</td>
<td>0.625</td>
<td>0.834</td>
</tr>
<tr>
<td>AgeSq*</td>
<td>0.880</td>
<td>0.447</td>
<td>0.539</td>
<td>0.635</td>
<td>0.576</td>
<td>0.770</td>
</tr>
<tr>
<td>PC1_NL</td>
<td>0.058</td>
<td>0.359</td>
<td>0.428</td>
<td>0.057</td>
<td>0.086</td>
<td>0.399</td>
</tr>
<tr>
<td>PC2_NL</td>
<td>0.382</td>
<td>0.276</td>
<td>0.611</td>
<td>0.211</td>
<td>0.330</td>
<td>0.660</td>
</tr>
<tr>
<td>PC3_NL</td>
<td>0.134</td>
<td>0.472</td>
<td>0.408</td>
<td>0.912</td>
<td>0.691</td>
<td>0.464</td>
</tr>
<tr>
<td>PC4</td>
<td>0.599</td>
<td>0.137</td>
<td>0.853</td>
<td>0.219</td>
<td>0.301</td>
<td>0.983</td>
</tr>
<tr>
<td>PC5</td>
<td>0.192</td>
<td>0.065</td>
<td>0.556</td>
<td>0.056</td>
<td>0.073</td>
<td>0.768</td>
</tr>
<tr>
<td>PC6</td>
<td>0.750</td>
<td>0.094</td>
<td>0.607</td>
<td>0.182</td>
<td>0.085</td>
<td>0.520</td>
</tr>
<tr>
<td>PC7</td>
<td>0.722</td>
<td>0.892</td>
<td>0.926</td>
<td>0.203</td>
<td>0.481</td>
<td>0.271</td>
</tr>
<tr>
<td>PC8</td>
<td>0.906</td>
<td>0.156</td>
<td>0.956</td>
<td>0.736</td>
<td>0.988</td>
<td>0.936</td>
</tr>
<tr>
<td>PC9</td>
<td>0.373</td>
<td>0.370</td>
<td>0.575</td>
<td>0.019</td>
<td>0.052</td>
<td>0.416</td>
</tr>
<tr>
<td>PC10</td>
<td>0.280</td>
<td>0.566</td>
<td>0.560</td>
<td>0.778</td>
<td>0.888</td>
<td>0.598</td>
</tr>
<tr>
<td>PC11</td>
<td>0.603</td>
<td>0.821</td>
<td>0.926</td>
<td>0.355</td>
<td>0.430</td>
<td>0.315</td>
</tr>
<tr>
<td>PC12</td>
<td>0.482</td>
<td>0.079</td>
<td>0.107</td>
<td>0.067</td>
<td>0.056</td>
<td>0.030</td>
</tr>
<tr>
<td>PC13</td>
<td>0.595</td>
<td>0.745</td>
<td>0.709</td>
<td>0.357</td>
<td>0.454</td>
<td>0.369</td>
</tr>
<tr>
<td>PC14</td>
<td>0.476</td>
<td>0.711</td>
<td>0.708</td>
<td>0.787</td>
<td>0.944</td>
<td>0.663</td>
</tr>
<tr>
<td>PC15</td>
<td>0.192</td>
<td>0.114</td>
<td>0.451</td>
<td>0.633</td>
<td>0.375</td>
<td>0.505</td>
</tr>
<tr>
<td>PC16</td>
<td>0.104</td>
<td>0.026</td>
<td>0.069</td>
<td>0.139</td>
<td>0.042</td>
<td>0.024</td>
</tr>
<tr>
<td>PC17</td>
<td>0.857</td>
<td>0.741</td>
<td>0.987</td>
<td>0.836</td>
<td>0.935</td>
<td>0.802</td>
</tr>
<tr>
<td>PC18</td>
<td>0.902</td>
<td>0.136</td>
<td>0.283</td>
<td>0.404</td>
<td>0.238</td>
<td>0.592</td>
</tr>
<tr>
<td>PC19</td>
<td>0.564</td>
<td>0.685</td>
<td>0.620</td>
<td>0.786</td>
<td>0.820</td>
<td>0.737</td>
</tr>
<tr>
<td>PC20</td>
<td>0.561</td>
<td>0.763</td>
<td>0.198</td>
<td>0.360</td>
<td>0.356</td>
<td>0.247</td>
</tr>
<tr>
<td>PC21</td>
<td>0.427</td>
<td>0.150</td>
<td>0.168</td>
<td>0.021</td>
<td>0.028</td>
<td>0.420</td>
</tr>
<tr>
<td>PC22</td>
<td>0.383</td>
<td>0.358</td>
<td>0.432</td>
<td>0.275</td>
<td>0.250</td>
<td>0.242</td>
</tr>
</tbody>
</table>

Note: Effects were estimated in R with Generalized Estimating Equations (gee package) with family identifier as clustering variable. PC_NL are geographical PCs reflecting a north-south gradient, east-west gradient, and central region, PCs 1-9 reflect SNP genotyping platform effects (different platforms were used for different subsamples), PCs 11-20 reflect remaining population stratification. Effects of the PCs after correcting for Sex, age group. P-values for Age and AgeSq are after the removal of age group means (groups 5, 7, 16, 18, 25 and 50) show that the age group dummy variable was highly effective in removing age effects. PC5 is missing due to lack of variation.
<table>
<thead>
<tr>
<th></th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>peak Occ</td>
</tr>
<tr>
<td>Sex</td>
<td>0.917</td>
</tr>
<tr>
<td>Age (log transformed)</td>
<td>0.703</td>
</tr>
<tr>
<td>Age (log transformed)</td>
<td>0.983</td>
</tr>
<tr>
<td>PC1</td>
<td>0.108</td>
</tr>
<tr>
<td>PC2</td>
<td>0.161</td>
</tr>
<tr>
<td>PC3</td>
<td><strong>0.000</strong></td>
</tr>
<tr>
<td>PC4</td>
<td><strong>0.000</strong></td>
</tr>
<tr>
<td>PC5</td>
<td>0.258</td>
</tr>
<tr>
<td>PC6</td>
<td>0.368</td>
</tr>
<tr>
<td>PC7</td>
<td>0.484</td>
</tr>
<tr>
<td>PC8</td>
<td>0.716</td>
</tr>
<tr>
<td>PC9</td>
<td><strong>0.022</strong></td>
</tr>
<tr>
<td>PC10</td>
<td>0.924</td>
</tr>
</tbody>
</table>

Note: Subjects included families of European ancestry from COGA. Age is chronological age. The 10 PCs were EIGENSTRAT to control for residual ancestral variation. p-values in bold face are significant at a level of p < 0.05, those in red were significant at a Bonferroni-corrected level of 0.0005. Covariates accounted for between 1.8% and of the variance in each phenotype.
## ENIGMA-EEG Genome-wide association

### MTFS

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>peak Occ</th>
<th>alpha Occ</th>
<th>delta Cz</th>
<th>theta Cz</th>
<th>alpha Cz</th>
<th>beta Cz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>.000</td>
<td>.000</td>
<td>.048</td>
<td>.000</td>
<td>.011</td>
<td>.330</td>
</tr>
<tr>
<td>Sex</td>
<td>.023</td>
<td>.171</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Age Cohort</td>
<td>.000</td>
<td>.567</td>
<td>.000</td>
<td>.044</td>
<td>.173</td>
<td>.127</td>
</tr>
<tr>
<td>Lab</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.204</td>
<td>.804</td>
<td>.000</td>
</tr>
<tr>
<td>PC1</td>
<td>.380</td>
<td>.213</td>
<td>.583</td>
<td>.278</td>
<td>.254</td>
<td>.249</td>
</tr>
<tr>
<td>PC2</td>
<td>.030</td>
<td>.191</td>
<td>.369</td>
<td>.187</td>
<td>.132</td>
<td>.736</td>
</tr>
<tr>
<td>PC3</td>
<td>.066</td>
<td>.179</td>
<td>.026</td>
<td>.112</td>
<td>.830</td>
<td>.383</td>
</tr>
<tr>
<td>PC4</td>
<td>.006</td>
<td>.013</td>
<td>.657</td>
<td>.284</td>
<td>.020</td>
<td>.797</td>
</tr>
<tr>
<td>PC5</td>
<td>.883</td>
<td>.118</td>
<td>.592</td>
<td>.428</td>
<td>.928</td>
<td>.566</td>
</tr>
<tr>
<td>PC6</td>
<td>.705</td>
<td>.652</td>
<td>.846</td>
<td>.989</td>
<td>.395</td>
<td>.434</td>
</tr>
<tr>
<td>PC7</td>
<td>.253</td>
<td>.934</td>
<td>.461</td>
<td>.719</td>
<td>.728</td>
<td>.660</td>
</tr>
<tr>
<td>PC8</td>
<td>.356</td>
<td>.001</td>
<td>.158</td>
<td>.086</td>
<td>.007</td>
<td>.039</td>
</tr>
<tr>
<td>PC9</td>
<td>.255</td>
<td>.569</td>
<td>.776</td>
<td>.572</td>
<td>.659</td>
<td>.763</td>
</tr>
<tr>
<td>PC10</td>
<td>.120</td>
<td>.631</td>
<td>.885</td>
<td>.237</td>
<td>.747</td>
<td>.464</td>
</tr>
</tbody>
</table>

Note: p-values are based on Rapid Feasible Generalized Least Squares (RFGLS). RFGLS is a computationally efficient implementation of Generalized Least Squares (GLS), which can accommodate correlated data and therefore the fact that data were from 4-member families. Subjects were adolescent twins or their parents. The Age Cohort dummy variable served to estimate mean differences between these two age groups. Lab refers to the recording system used to collect the data (Biosemi or Grass). Age is chronological age. The 10 PCs were derived from EIGENSTRAT to control for residual ethnic variation. p-values in bold face are significant at a level of p < .05, whereas those in red were significant at a Bonferroni-correlated level of .0006. Covariates accounted for between 1.7% and 23.5% of the variance in each phenotype using the method of Fite (1973) for estimating R^2 in GLS.

## QIMMR

### QIMMR Table. Effects of covariates on the EEG phenotypes using Linear Mixed Models in GCTA.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>peak Occ</th>
<th>alpha Occ</th>
<th>delta Cz</th>
<th>theta Cz</th>
<th>alpha Cz</th>
<th>beta Cz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>0.022</td>
<td>0.588</td>
<td>0.818</td>
<td>8.48E-06</td>
<td>5.65E-05</td>
<td>5.69E-14</td>
</tr>
<tr>
<td>Age</td>
<td>0.033</td>
<td>0.474</td>
<td>0.479</td>
<td>0.753</td>
<td>0.340</td>
<td>0.214</td>
</tr>
<tr>
<td>AgeSq</td>
<td>0.033</td>
<td>0.449</td>
<td>0.511</td>
<td>0.787</td>
<td>0.321</td>
<td>0.223</td>
</tr>
<tr>
<td>PC1</td>
<td>0.530</td>
<td>0.890</td>
<td>0.323</td>
<td>0.446</td>
<td>0.486</td>
<td>0.172</td>
</tr>
<tr>
<td>PC2</td>
<td>0.588</td>
<td>0.791</td>
<td>0.162</td>
<td>0.211</td>
<td>0.982</td>
<td>0.857</td>
</tr>
<tr>
<td>PC3</td>
<td>0.650</td>
<td>0.446</td>
<td>0.251</td>
<td>0.851</td>
<td>0.232</td>
<td>0.195</td>
</tr>
<tr>
<td>PC4</td>
<td>0.275</td>
<td>0.925</td>
<td>0.557</td>
<td>0.053</td>
<td>0.473</td>
<td>0.979</td>
</tr>
<tr>
<td>PC5</td>
<td>0.543</td>
<td>0.888</td>
<td>0.410</td>
<td>0.352</td>
<td>0.437</td>
<td>0.213</td>
</tr>
<tr>
<td>PC6</td>
<td>0.989</td>
<td>0.239</td>
<td>0.779</td>
<td>0.784</td>
<td>0.868</td>
<td>0.092</td>
</tr>
<tr>
<td>PC7</td>
<td>0.591</td>
<td>0.436</td>
<td>0.249</td>
<td>0.093</td>
<td>0.549</td>
<td>0.862</td>
</tr>
<tr>
<td>PC8</td>
<td>0.340</td>
<td>0.044</td>
<td>0.721</td>
<td>0.335</td>
<td>0.098</td>
<td>0.888</td>
</tr>
<tr>
<td>PC9</td>
<td>0.053</td>
<td>0.558</td>
<td>0.284</td>
<td>0.991</td>
<td>0.843</td>
<td>0.856</td>
</tr>
<tr>
<td>PC10</td>
<td>0.254</td>
<td>0.065</td>
<td>0.232</td>
<td>0.037</td>
<td>0.309</td>
<td>0.442</td>
</tr>
</tbody>
</table>
Supplementary Methods References


---

Supplementary table pre-meta-analysis QC steps

Table. SNP removal at pre-meta-analysis QC steps using EasyQC

<table>
<thead>
<tr>
<th>QC step</th>
<th>NTR</th>
<th>COGA_CC</th>
<th>COGA_EA</th>
<th>MTFS</th>
<th>QIMR</th>
</tr>
</thead>
<tbody>
<tr>
<td>ref vs sample allele freq. outlier</td>
<td>54</td>
<td>90</td>
<td></td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>duplicates</td>
<td>13</td>
<td>90</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>invalid effect allele</td>
<td>6511</td>
<td>166</td>
<td>3151</td>
<td>74821</td>
<td></td>
</tr>
<tr>
<td>sample effect allele freq.</td>
<td>1161577</td>
<td>3046801</td>
<td>4073771</td>
<td>2567458</td>
<td>2057384</td>
</tr>
<tr>
<td>reference set effect allele freq.</td>
<td>8952</td>
<td>17854</td>
<td>9919</td>
<td>3045</td>
<td></td>
</tr>
<tr>
<td>low call rate/low SNP N</td>
<td>123097</td>
<td>3199971</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Missing P</td>
<td></td>
<td>568</td>
<td>120</td>
<td>519</td>
<td>22941</td>
</tr>
<tr>
<td>HWE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Imputation quality (R2/INFO)</td>
<td>31854</td>
<td>25735848</td>
<td>79781</td>
<td>325094</td>
<td>99582</td>
</tr>
</tbody>
</table>
ENIGMA-EEG Genome-wide association


Supplementary acknowledgements and references for main text

LD Hub acknowledges the following consortia:

“We gratefully acknowledge all the studies and databases that made GWAS summary data available: ADIPOGen (Adiponectin genetics consortium), C4D (Coronary Artery Disease Genetics Consortium), CARDIoGRAM (Coronary ARtery Disease Genome wide Replication and Meta-analysis), CKDGen (Chronic Kidney Disease Genetics consortium), dbGAP (database of Genotypes and Phenotypes), DIAGRAM (Diabetes Genetics Replication And Meta-analysis), ENIGMA (Enhancing Neuro Imaging Genetics through Meta Analysis), EAGLE (Early Genetics & Lifecourse Epidemiology Eczema Consortium, excluding 23andMe), EGG (Early Growth Genetics Consortium), GABRIEL (A Multidisciplinary Study to Identify the Genetic and Environmental Causes of Asthma in the European Community), GCAN (Genetic Consortium for Anorexia Nervosa), GEFOS (Genetic Factors for Osteoporosis Consortium), GIANT (Genetic Investigation of ANthropometric Traits), GIS (Genetics of Iron Status consortium), GLGC (Global Lipids Genetics Consortium), GPC (Genetics of Personality Consortium), GUGC (Global Urate and Gout consortium), HaemGen (haematological and platelet traits genetics consortium), HRgene (Heart Rate consortium), IBDGC (International Inflammatory Bowel Disease Genetics Consortium), ILCCO (International Lung Cancer Consortium), IMMSGC (International Multiple Sclerosis Genetic Consortium), MAGIC (Meta-Analyses of Glucose and Insulin-related traits Consortium), MESA (Multi-Ethnic Study of Atherosclerosis), PGC (Psychiatric Genomics Consortium), Project MinE consortium, ReproGen (Reproductive Genetics Consortium), SSGAC (Social Science Genetics Association Consortium) and TAG (Tobacco and Genetics Consortium), TRICL (Transdisciplinary Research in Cancer of the Lung consortium), UK Biobank.

“We gratefully acknowledge the contributions of Alkes Price (the systemic lupus erythematosus GWAS and primary biliary cirrhosis GWAS) and Johannes Kettunen (lipids metabolites GWAS).”

GWAS References for LD Hub obtained from
https://academic.oup.com/bioinformatics/article/33/2/272/2525718
Anderson, C. et al. (2011) Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. Nature Genetics, 43, 246-252.


Benyamin, B. et al. (2013) Childhood intelligence is heritable, highly polygenic and associated with FNBP1L. Molecular Psychiatry, 19, 253-258.


ENIGMA-EEG Genome-wide association

ENIGMA-EEG Genome-wide association


Full acknowledgement

University of Minnesota: Funding by National Institutes of Health (NIH) DA 05147, DA 38216, DA 024417, AA 09367, DA 13240

COGA: The Collaborative Study on the Genetics of Alcoholism (COGA), Principal Investigators B. Porjesz, V. Hesselbrock, H. Edenberg, L. Bierut, includes eleven different centers: University of Connecticut (V. Hesselbrock); Indiana University (H.J. Edenberg, J. Nurnberger Jr., T. Foroud); University of Iowa (S. Kuperman, J. Kramer); SUNY Downstate (B. Porjesz); Washington University in St. Louis (L. Bierut, J. Rice, K. Bucholz, A. Agrawal); University of California at San Diego (M. Schuckit); Rutgers University (J. Tischfield, A. Brooks); Department of Biomedical and Health Informatics, The Children's Hospital of Philadelphia; Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia PA (L. Almasy), Virginia Commonwealth University (D. Dick), Icahn School of Medicine at Mount Sinai (A. Goate), and Howard University (R. Taylor). Other COGA collaborators include: L. Bauer (University of Connecticut); J. McClintick, L. Wetherill, X. Xuei, Y. Liu, D. Lai, S. O’Connor, M. Plawecki, S. Lourens (Indiana University); G. Chan (University of Iowa; University of Connecticut); J. Meyers, D. Chorlian, C. Kamarajan, A. Pandey, J. Zhang (SUNY Downstate); J.-C. Wang, M. Kapoor, S. Bertelsen (Icahn School of Medicine at Mount Sinai); A. Anokhin, V. McCutcheon, S. Saccone (Washington University); J. Salvatore, F. Aliev, B. Cho (Virginia Commonwealth University); and Mark Kos (University of Texas Rio Grande Valley). A. Parsian and M. Reilly are the NIAAA Staff Collaborators.

We continue to be inspired by our memories of Henri Begleiter and Theodore Reich, founding PI and Co-PI of COGA, and also owe a debt of gratitude to other past organizers of COGA, including Ting-Kai Li, P. Michael Conneally, Raymond Crowe, and Wendy Reich, for their critical contributions. This national collaborative study is supported by NIH Grant U10AA008401 from the National Institute on Alcohol Abuse and Alcoholism (NIAAA) and the National Institute on Drug Abuse (NIDA). Funding for GWAS genotyping, which was performed at the Johns Hopkins University Center for Inherited Disease Research, was provided by NIAAA, the NIH GEI (U01HG004438) and the NIH contract ‘High through-put genotyping for studying the genetic contributions to human disease’ (HHSN268200782096C).

Netherlands Twin Registry/Vrije Universiteit: Funding by Netherlands Organization for Scientific Research (NWO) and The Netherlands Organisation for Health Research and Development (ZonMW) grants 904-61-090, 985-10-002, 912-10-020, 904-61-193, 480-04-004, 463-06-001, 451-04-034, 400-05-717, Addiction-31160008, Middelgroot-911-09-032, Spinozapremie 56-464-14192, Biobanking and Biomolecular Resources Research Infrastructure (BBMRI –NL, 184.021.007, NWO-Groot 480-15-001/674). VU Institute for Health and Care Research (EMGO+ ); the European Community’s Seventh Framework Program (FP7/2007-2013), ENGAGE (HEALTH-F4-2007-201413); the European Research Council (ERC Advanced, 230374, ERC Starting grant 284167), Rutgers University Cell and DNA Repository (NIHM U24 MH068457-06), the Avera Institute, Sioux Falls, South Dakota (USA) and the National Institutes of Health (NIH, R01DK0042157-01A, MH081802; R01 DK092127-04, Grand Opportunity grants 1RC2 MH089951 and 1RC2 MH089995). Part of the genotyping and analyses were funded by the Genetic Association Information Network (GAIN) of the Foundation for the National Institutes of Health. Computing was supported by BiG Grid, the Dutch e-Science Grid, which is financially supported by NWO.

Personal funding: Netherlands Organisation for Scientific Research NWO 480-04- 004, NWO/SPI 56-464-14192, and NWO 911-09-032 to D.B., NWO/MagW VENI-451-08-026 to D.S.; European Research Council ERC-230374 to D.B.; BBR Foundation (NARSAD) Young
Investigator grant 21668 to D.S.; VU University VU-USF 96/22 to D.B.; Human Frontiers of Science Program RG0154/1998-B to D.B., and E.d.G.; KNAW Academy Professor Award (PAH/6635) to D.B.; Australian Research Council (A79600334, A79906588, A79801419, DP0212016) to N.M. and M.W.; National Health and Medical Research Council (Medical Bioinformatics Genomics Proteomics Program 389891) to N.M.; Fellowship APP1103623 to S.M.; NIH K01DA037914 to J.L.M.