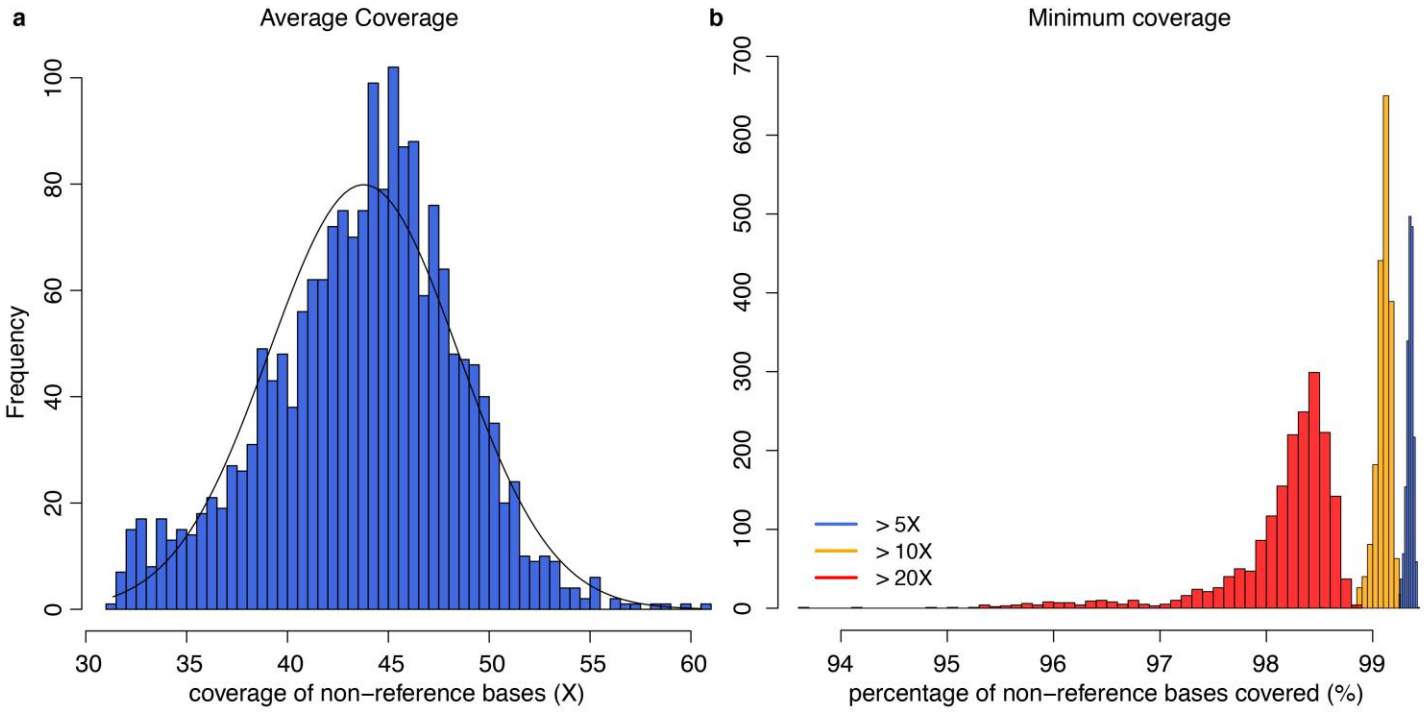
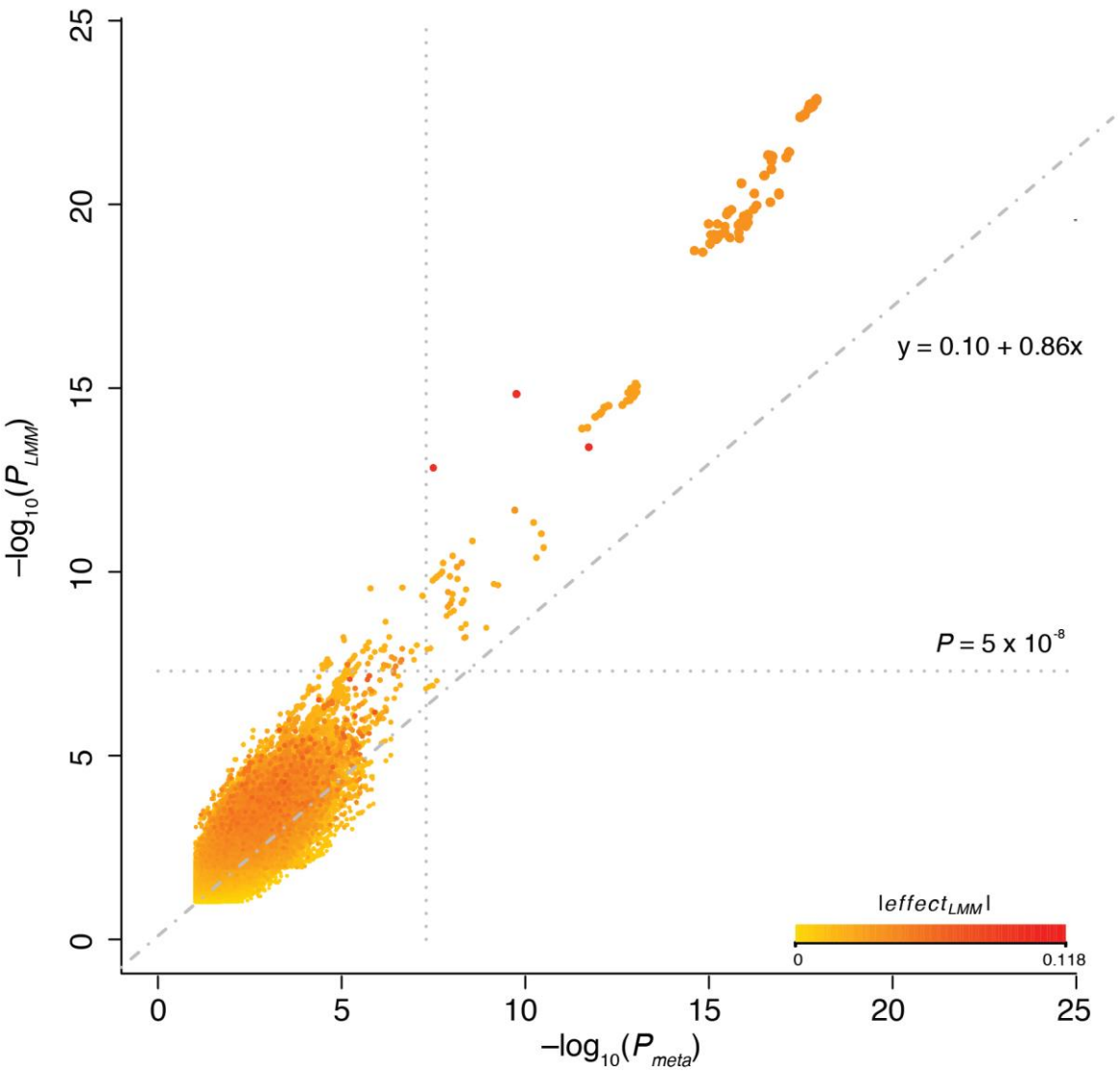


Supplementary Figure 1

Coverage distribution for reference panel.

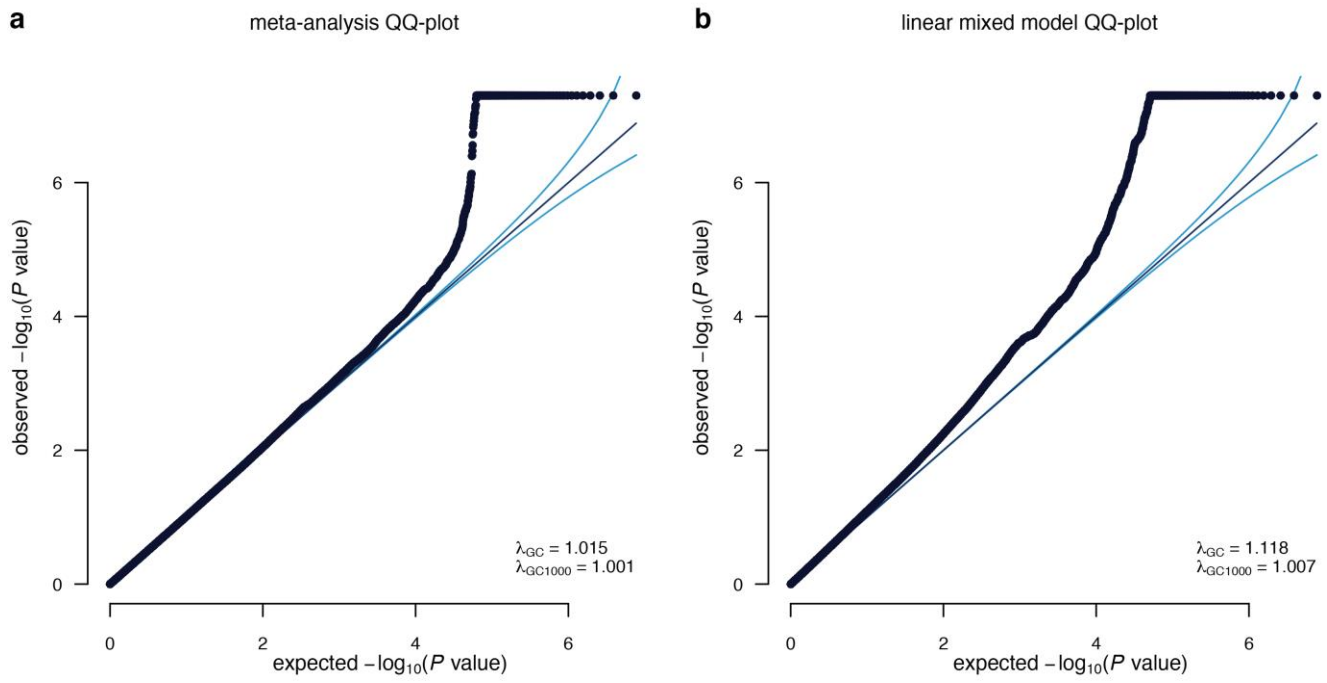
(**a**) Average coverage for all non-reference bases for all samples in the custom reference panel. An average coverage of 43.7 reads per base was achieved. (**b**) Percentage of non-reference bases covered by at least 5 (blue), 10 (yellow) or 20 (red) reads. Histograms include all individuals, before quality control.



Supplementary Figure 2

Comparison between meta-analysis and linear mixed-model results.

Regressing the $-\log_{10}$ (*P* values) derived from the linear mixed model on the meta-analysis *P* values yielded a slope below the diagonal (β = 0.86), implying no overall inflation of the test statistic. Strongly associated SNPs, however, deviated from the regression line representing the increased power of a linear mixed model in comparison to a meta-analysis for association testing.

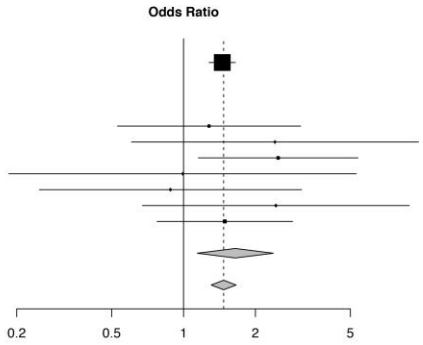


Supplementary Figure 3

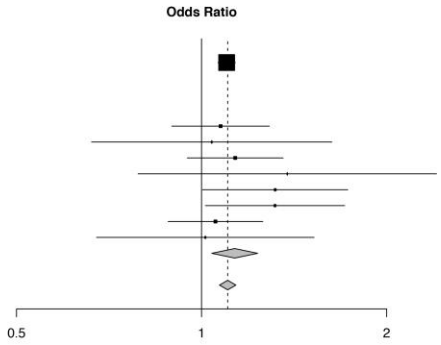
Quantile–quantile plots.

(**a**,**b**) Meta-analysis (**a**) and linear mixed model (**b**). For presentation purposes, *P* values <5 × 10^{-8} are plotted at 5 × 10^{-8}.

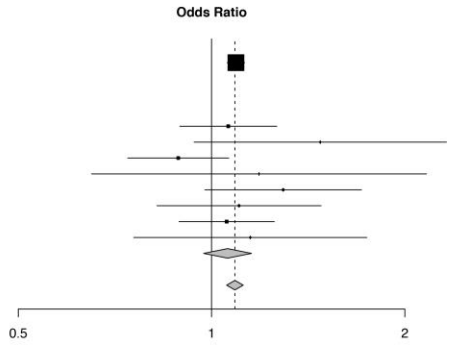
rs75087725 (*C21orf2*)	Cases	Contols	MAF(cases)	MAF(controls)	OR	95%–CI	P
Discovery							
Linear Mixed Model	12577	23475	0.019	0.014	1.45	[1.28; 1.65]	**2.65e-09**
Replication							
Australia	519	661	0.010	0.008	1.28	[0.53; 3.10]	
Belgium	93	92	0.038	0.016	2.41	[0.60; 9.64]	
France	606	530	0.021	0.008	2.49	[1.15; 5.39]	
Germany	116	46	0.022	0.022	0.99	[0.19; 5.30]	
Ireland	209	359	0.007	0.008	0.88	[0.25; 3.12]	
Italy	348	227	0.016	0.007	2.44	[0.67; 8.83]	
Netherlands	547	778	0.016	0.011	1.49	[0.77; 2.86]	
Turkey	141	74	0.000	0.014	NA		
Replication	**2579**	**2767**			**1.65**	**[1.14; 2.38]**	**3.89e-03**
Discovery + replication	**15156**	**26242**			**1.47**	**[1.31; 1.66]**	**3.08e-10**



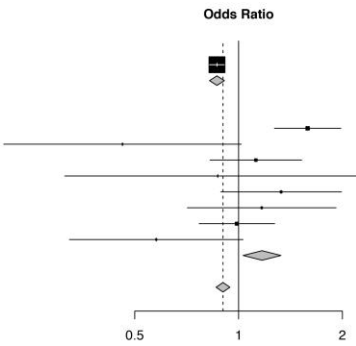
rs616147 (*MOBP*)	Cases	Contols	MAF(cases)	MAF(controls)	OR	95%–CI	P
Discovery							
Linear Mixed Model	12577	23475	0.303	0.276	1.10	[1.06; 1.14]	**1.43e-08**
Replication							
Australia	519	661	0.279	0.265	1.07	[0.89; 1.29]	
Belgium	93	92	0.280	0.272	1.04	[0.66; 1.63]	
France	606	530	0.310	0.283	1.13	[0.95; 1.36]	
Germany	116	46	0.291	0.228	1.38	[0.79; 2.41]	
Ireland	209	359	0.301	0.246	1.32	[1.00; 1.73]	
Italy	348	227	0.362	0.303	1.32	[1.01; 1.71]	
Netherlands	547	778	0.299	0.289	1.05	[0.88; 1.26]	
Turkey	141	74	0.375	0.372	1.01	[0.67; 1.52]	
Replication	**2579**	**2767**			**1.13**	**[1.04; 1.23]**	**2.35e-03**
Discovery + replication	**15156**	**26242**			**1.10**	**[1.07; 1.14]**	**4.19e-10**



rs10139154 *(SCFD1)*	Cases	Contols	MAF(cases)	MAF(controls)	OR	95%–CI	P
Discovery							
Linear Mixed Model	12577	23475	0.337	0.312	1.09	[1.06; 1.13]	**4.95e-08**
Replication							
Australia	519	661	0.311	0.298	1.06	[0.89; 1.26]	
Belgium	93	92	0.367	0.285	1.48	[0.94; 2.32]	
France	606	530	0.307	0.333	0.89	[0.74; 1.06]	
Germany	116	46	0.306	0.275	1.19	[0.65; 2.16]	
Ireland	209	359	0.293	0.246	1.29	[0.98; 1.71]	
Italy	348	227	0.408	0.384	1.10	[0.82; 1.48]	
Netherlands	547	778	0.331	0.320	1.06	[0.89; 1.25]	
Turkey	141	74	0.383	0.351	1.15	[0.76; 1.75]	
Replication	**2579**	**2767**			**1.06**	**[0.97; 1.15]**	**9.55e-02**
Discovery + replication	**15156**	**26242**			**1.09**	**[1.06; 1.12]**	**3.45e-08**



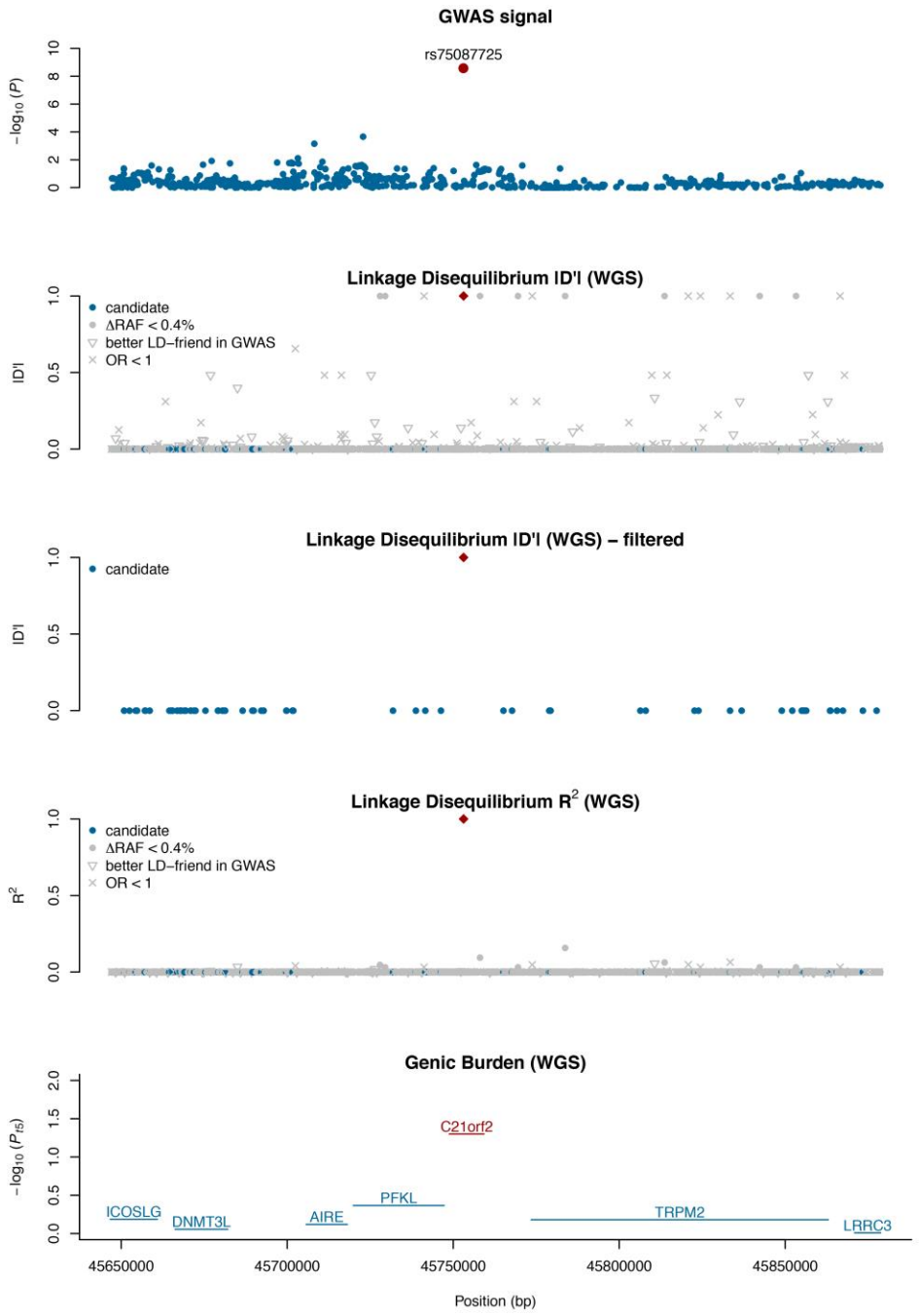
rs7813314 (*lncRNA*)	Cases	Contols	MAF(cases)	MAF(controls)	OR	95%–CI	P
Discovery							
Linear Mixed Model	12577	23475	0.093	0.102	0.87	[0.82; 0.91]	**3.14e-08**
Replication							
Australia	519	661	0.161	0.101	1.59	[1.27; 1.98]	
Belgium	93	92	0.063	0.12	0.46	[0.21; 1.02]	
France	606	530	0.098	0.089	1.12	[0.83; 1.53]	
Germany	116	46	0.072	0.081	0.87	[0.31; 2.42]	
Ireland	209	359	0.116	0.091	1.33	[0.89; 1.99]	
Italy	348	227	0.109	0.096	1.17	[0.71; 1.92]	
Netherlands	547	778	0.110	0.111	0.99	[0.77; 1.27]	
Turkey	141	74	0.101	0.164	0.58	[0.32; 1.03]	
Replication	**2579**	**2767**			**1.17**	**[1.03; 1.33]**	**7.75e-03**
Discovery + replication	**15156**	**26242**			**0.90**	**[0.86; 0.94]**	**1.05e-05**



Supplementary Figure 4

Replication results.

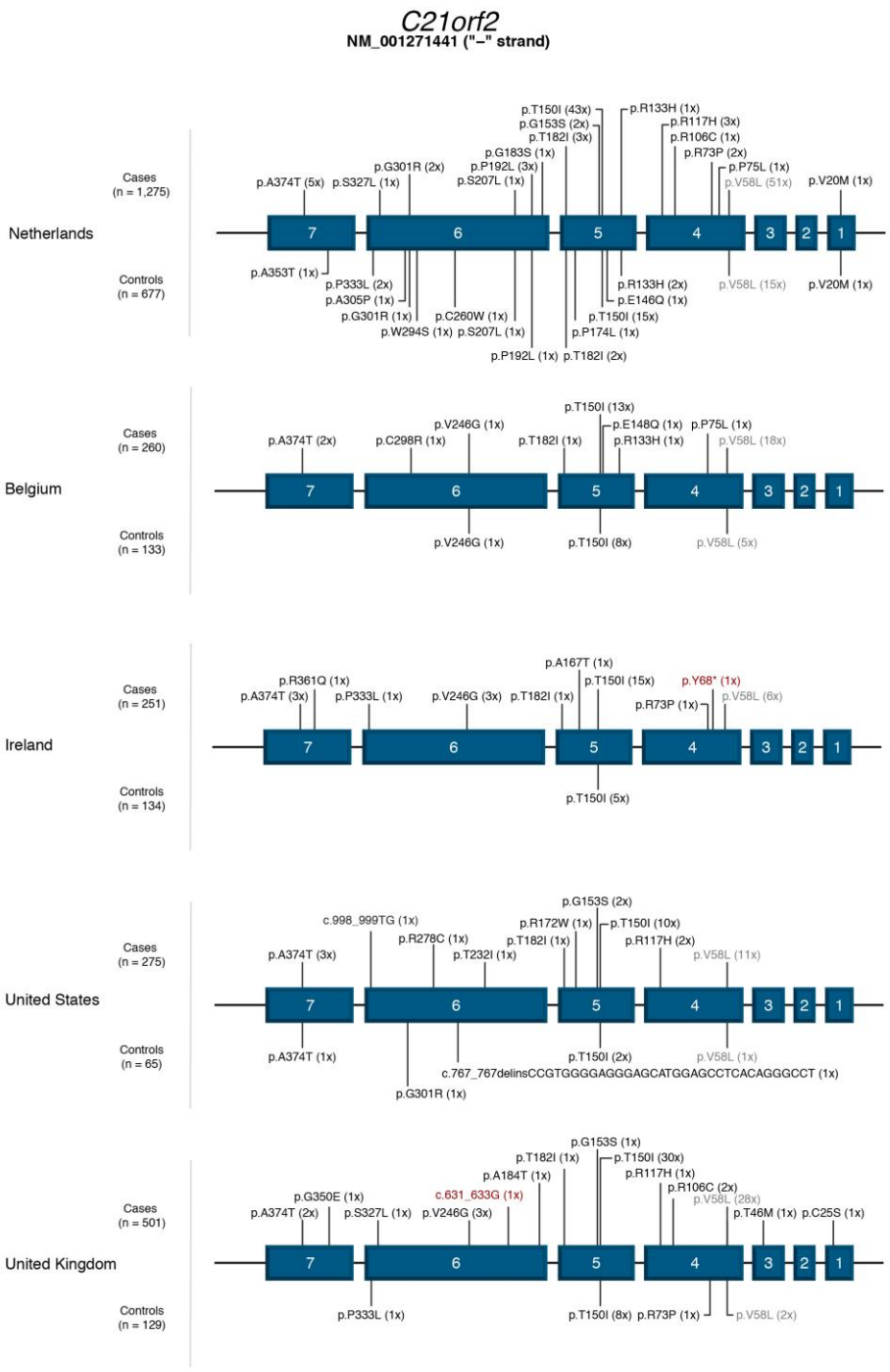
Forest plot for the inverse-variance-weighted, fixed-effect meta-analysis of the discovery phase and replication cohorts. OR, odds ratio; CI, confidence interval.



Supplementary Figure 5

Fine-mapping of the *C21orf2* locus.

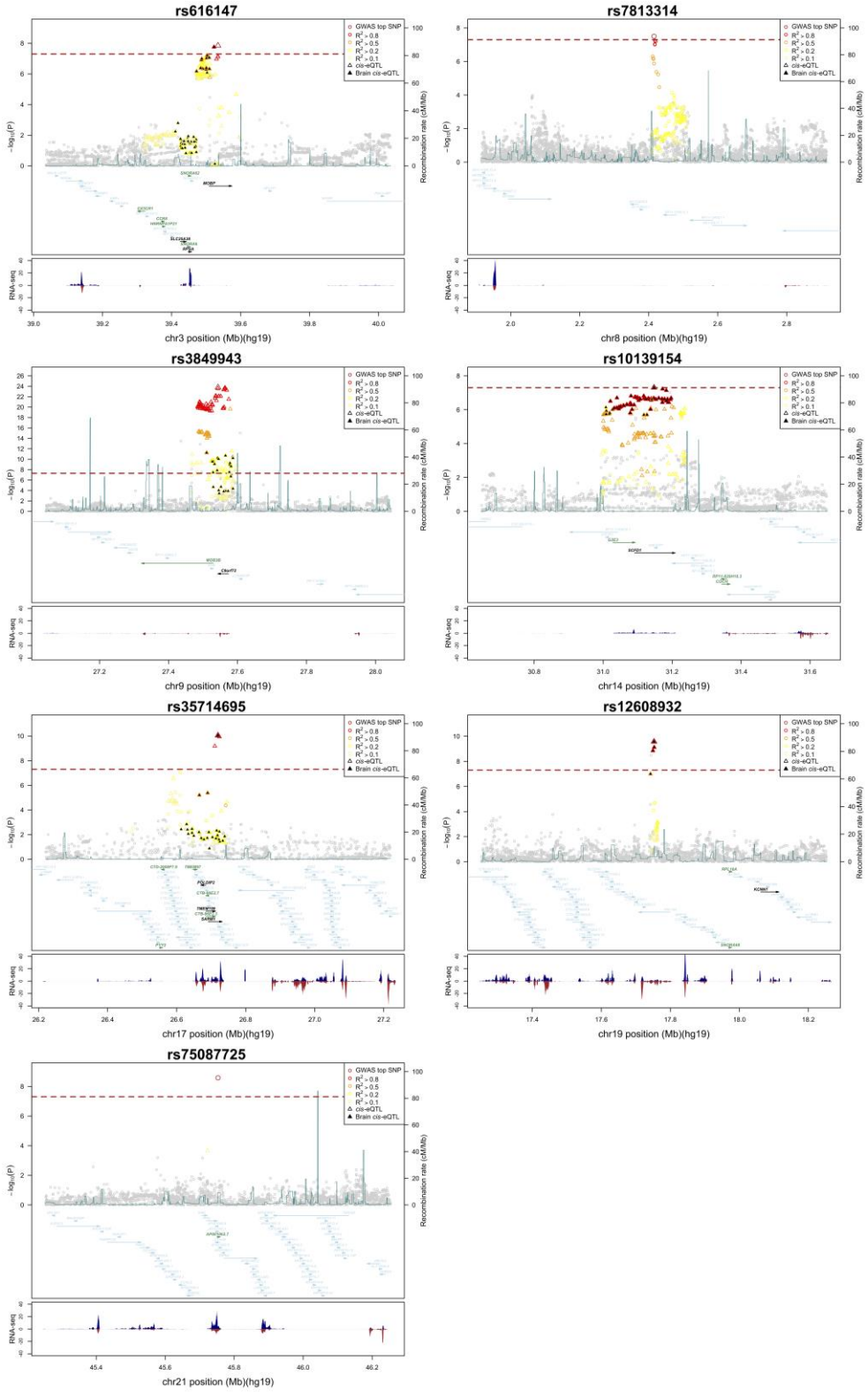
(**a**) Associations for all SNPs analyzed in the GWAS in the *C21orf2* locus where only rs75087725 reaches genome-wide significance. (**b**) LD as defined by |*D'*| (absolute *D*-prime value) for all variants obtained from whole-genome sequencing data of the custom reference panel. Possible causal variants for driving factors of the GWAS association were considered when they met the following criteria: (i) risk allele frequency difference between cases and controls exceeding 0.4% (risk allele frequency difference for rs75087725 = 0.8%), (ii) rs75087725 as the best genotyped GWAS tag for the variant and (iii) minor allele with a risk-increasing effect on ALS. (**c**) No such variant was in LD with rs75087725. (**d**) LD defined by R^2 is very sparse. (**e**) *C21orf2* is the only gene in this locus with an increased burden of rare nonsynonymous variants in the sequencing data of our custom reference panel, indicating that *C21orf2* is indeed the ALS risk gene.



Supplementary Figure 6

***C21orf2* rare variant burden.**

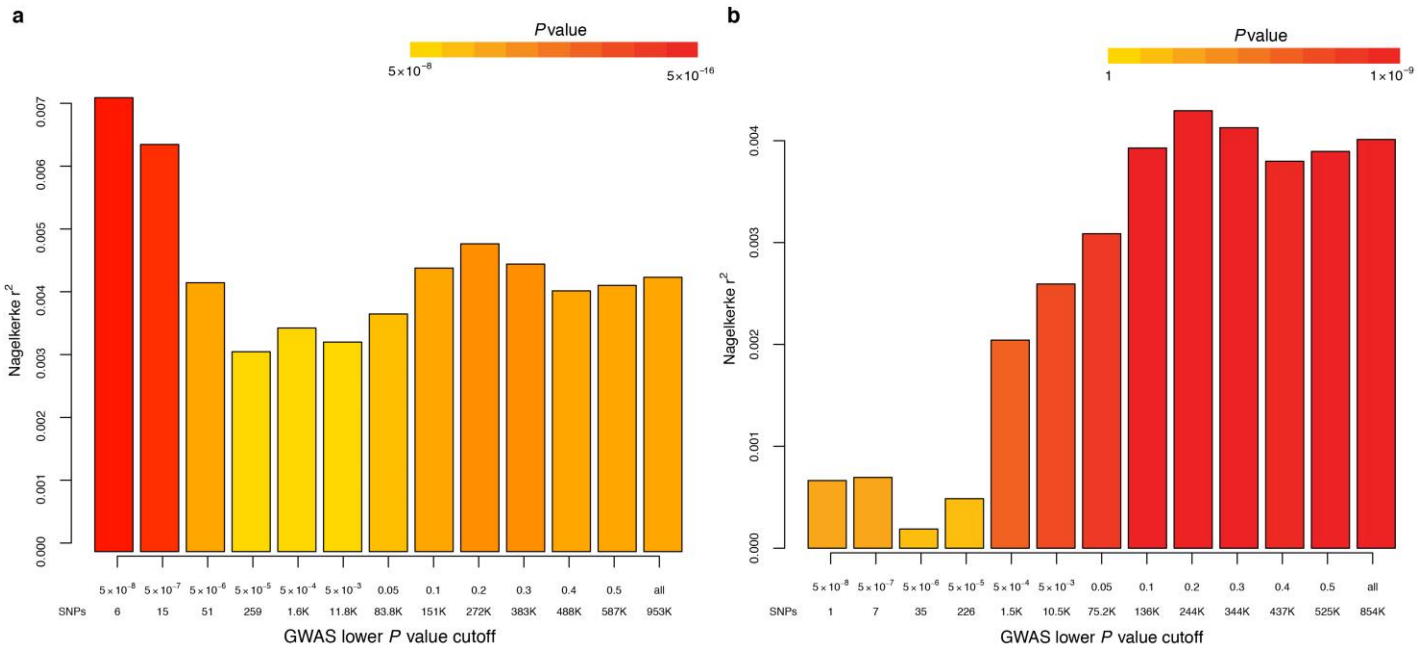
Summary of the rare (MAF < 0.05) nonsynonymous and loss-of-function mutations in the canonical transcript of *C21orf2*. Conditioning on the SNP found to be associated in the GWAS (rs75087725, p.V58L; gray), there was an increased burden of nonsynonymous and loss-of-function mutations among ALS cases ($P_{T5} = 9.2 \times 10^{-5}$, $P_{T1} = 0.01$). Odds ratios (calculated by counting alleles in cases and controls per stratum, unadjusted for principal components, combined in a Cochran–Mantel–Haenszel test) are 1.63 and 1.48 for T5 and T1 burden, respectively. The two loss-of-function mutations observed in cases are colored red.



Supplementary Figure 7

***cis*-eQTL regional plots for six genome-wide-significant loci.**

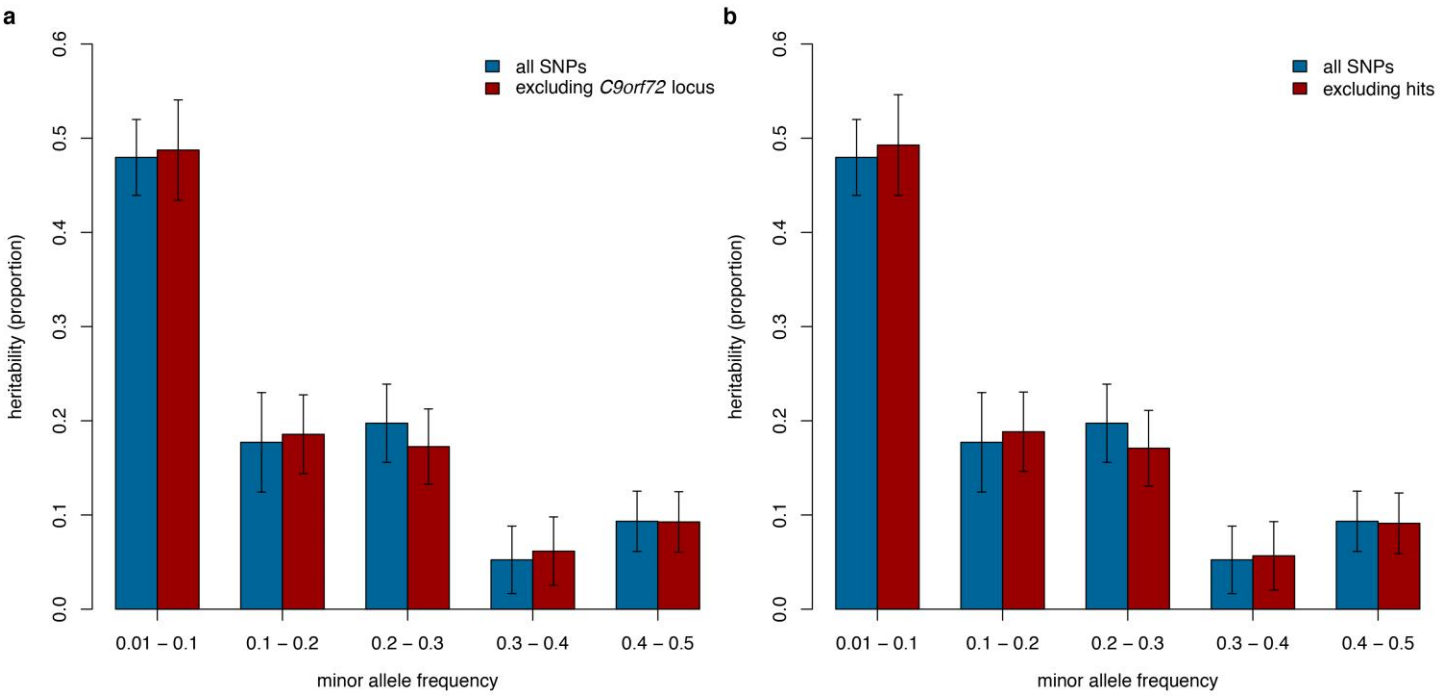
Highlighted are *cis*-eQTLs acquired from several resources (genes in green) and brain *cis*-eQTLs (genes in black). Stranded RNA-seq data for one fetal brain (3,000-bp sliding window) are shown on a separate track. For the *MOBP* region, there was an eQTL effect for *MOBP* ($P = 7.12 \times 10^{-18}$), but the effect SNP, rs1707953, did not show any association with ALS in our GWAS ($P = 0.74$). Further details of highlighted brain *cis*-eQTLs and non-brain *cis*-eQTLs are given in **Supplementary Table 10** and the **Supplementary Data Set**. eQTL annotation and LD data are shown only for SNPs present in the 1000 Genomes Project p1v3 CEU population.



Supplementary Figure 8

Polygenic risk scores.

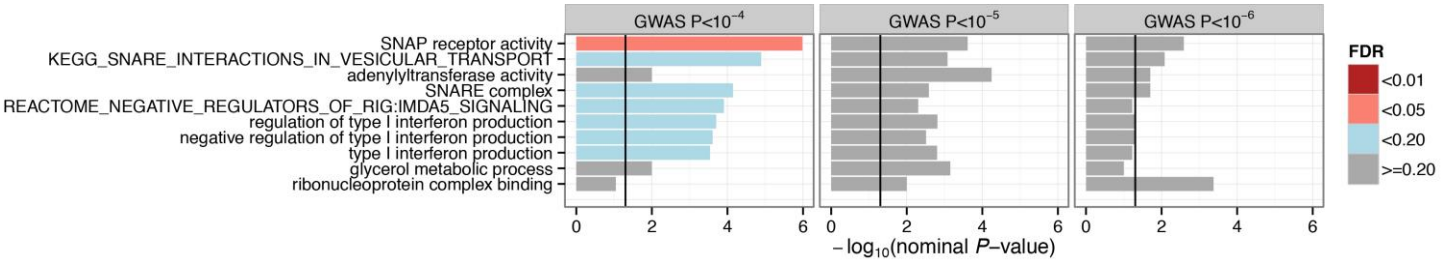
(**a**) Polygenic risk score analyses where nine cohorts were used as targets. Best predictions were made when including the six genome-wide-significant SNPs from the *C9orf72* and *UNC13A* loci only. (**b**) Polygenic risk score analyses excluding all variants on chromosome 9. Increased polygenic risk score predictions were made when including more variants by lowering the *P*-value threshold. Note that the overall prediction accuracy is lower than when SNPs on chromosome 9 were included.



Supplementary Figure 9

Partitioned heritability excluding candidate loci.

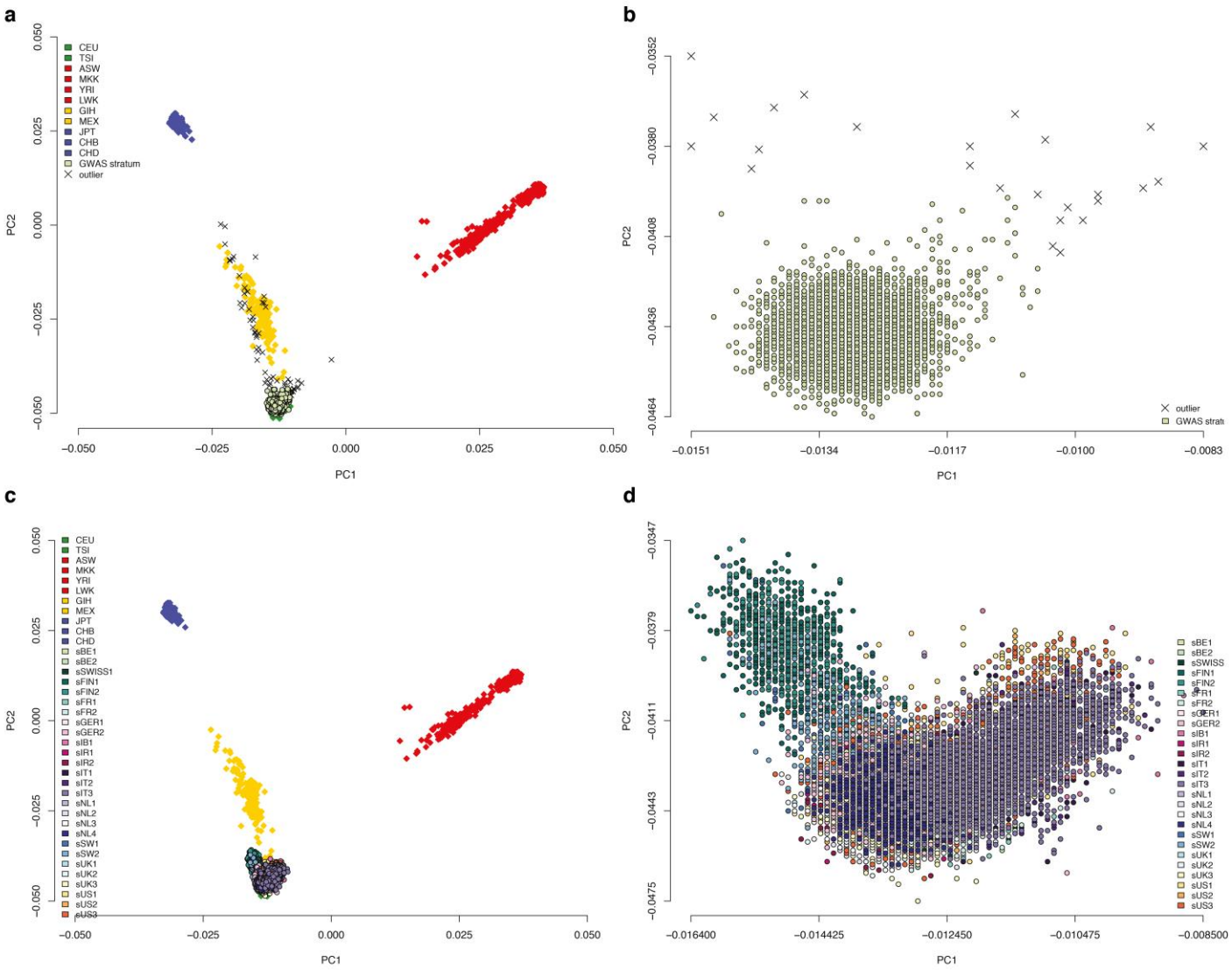
(**a**) SNPs in the *C9orf72* locus (within 1 Mb of rs3849943 and r^2 >0.2) were excluded from heritability estimates. (**b**) SNPs within 1 Mb of the top associated SNP and r^2 >0.2 for all loci exceeding genome-wide significance were excluded for heritability estimates. In both instances, most heritability was explained by low-frequency variants (MAF < 0.1).



Supplementary Figure 10

DEPICT biological pathway analysis.

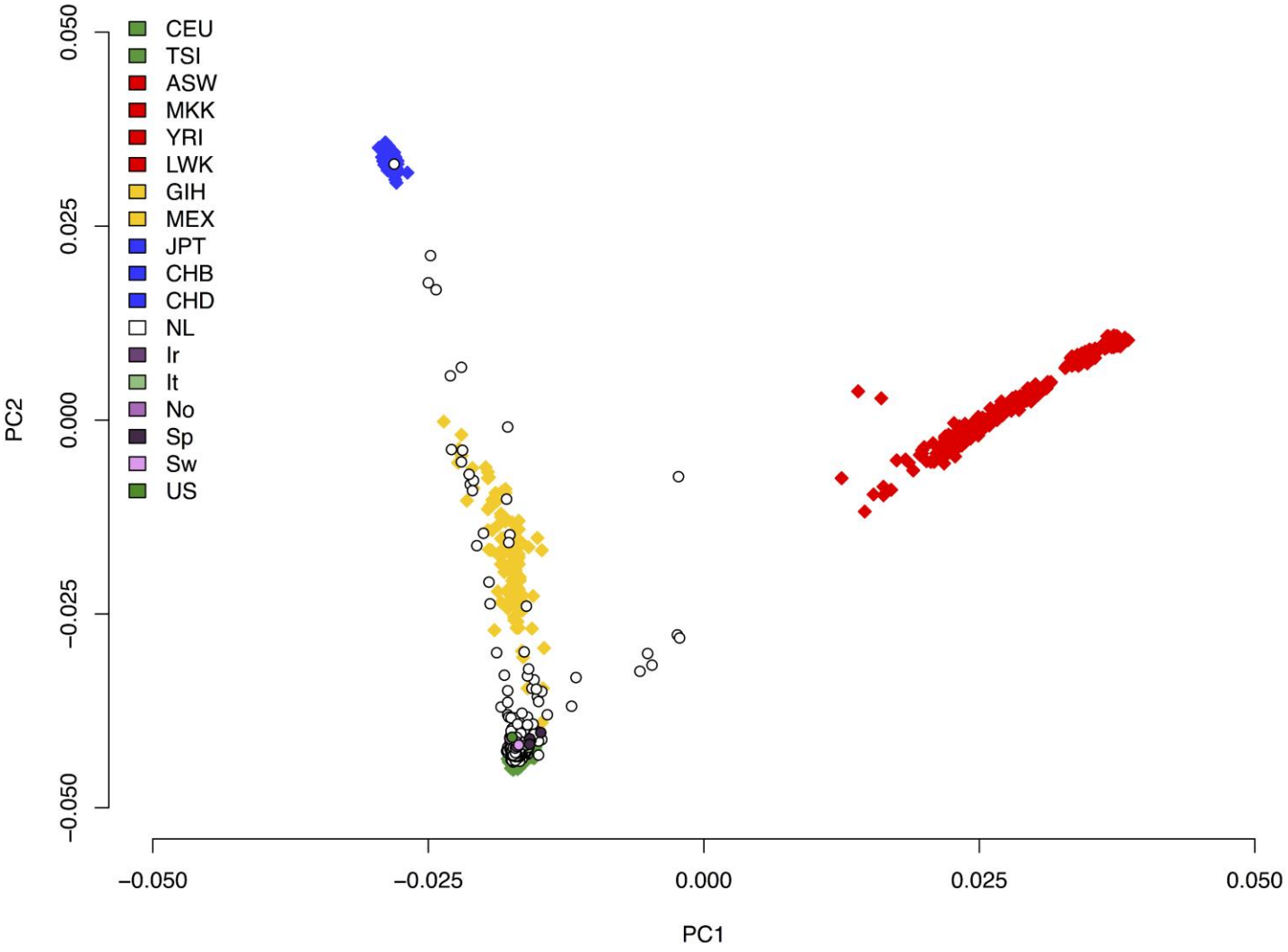
The top ten terms from Gene Ontology, KEGG and Reactome pathways that are most enriched for genes tagged by SNPs in the GWAS are displayed. Different thresholds of significance were used to select SNPs for the DEPICT analyses. The lengths of the bars correspond to the nominal significance levels, the black line indicates the *P*-value threshold of 0.05 and color corresponds to FDR, determined by 200 permutations. When all SNPs with a *P* value $<1 \times 10^{-4}$ in the linear mixed-model analysis were included, it identifies the Gene Ontology category SNAP receptor (SNARE) activity as the only significantly enriched term after correction for multiple testing.



Supplementary Figure 11

Population outlier removal.

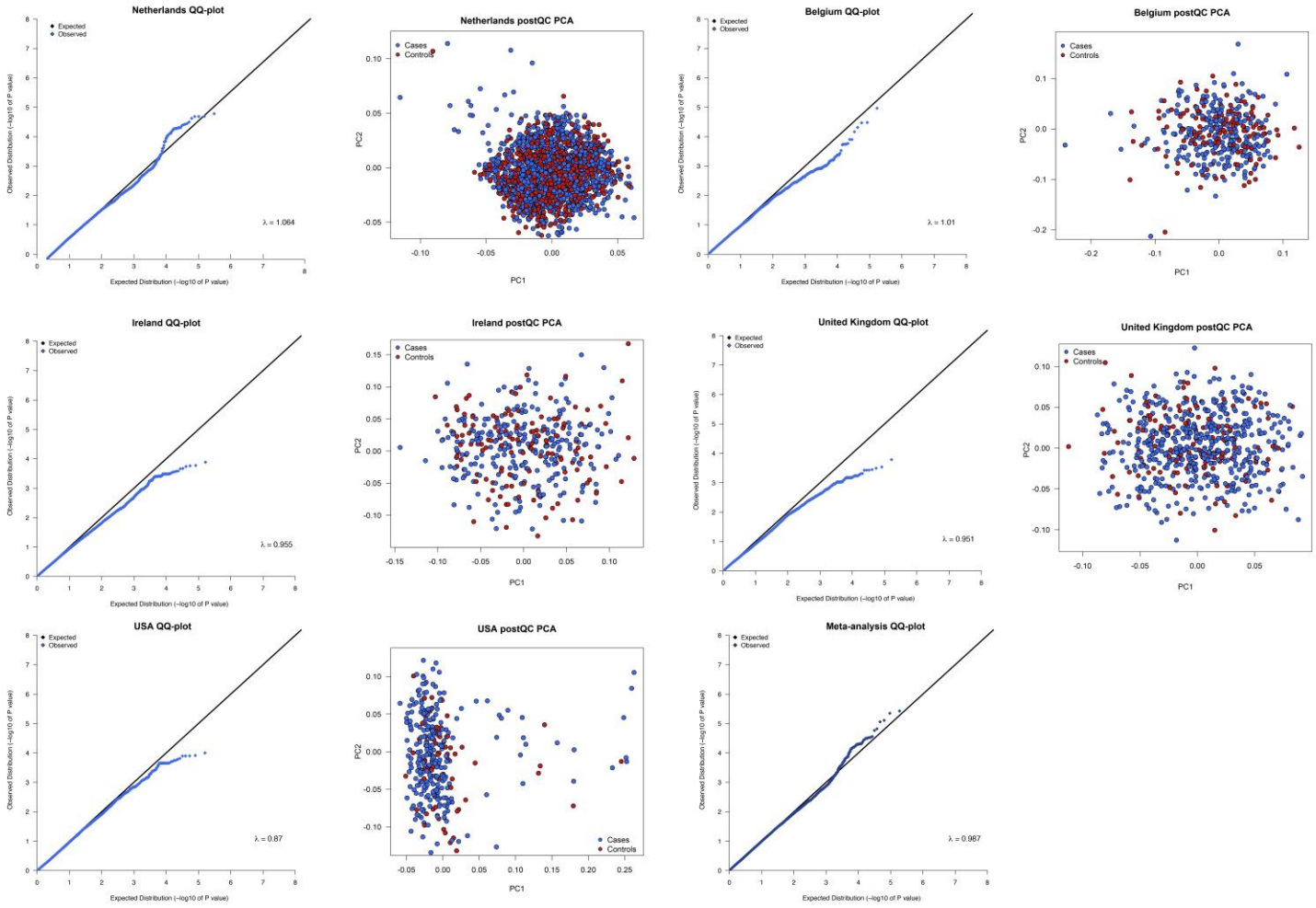
(**a**) Example of stratum sNL2 projected onto the first two principal components calculated on HapMap 3 individuals. Individuals of non-European ancestry (±10 s.d. from the HapMap CEU mean on PC1–PC4) were removed. (**b**) Subsequent removal of samples deviating by more than 4 s.d. from the stratum mean eigenvalues on PC1 and PC2 (HapMap 3). (**c**,**d**) Individuals remaining after removing population outliers resulting in a homogenous European sample where the Scandinavian individuals (mainly Finnish), as expected, depart from the other strata.



Supplementary Figure 12

Population structure of the custom reference panel.

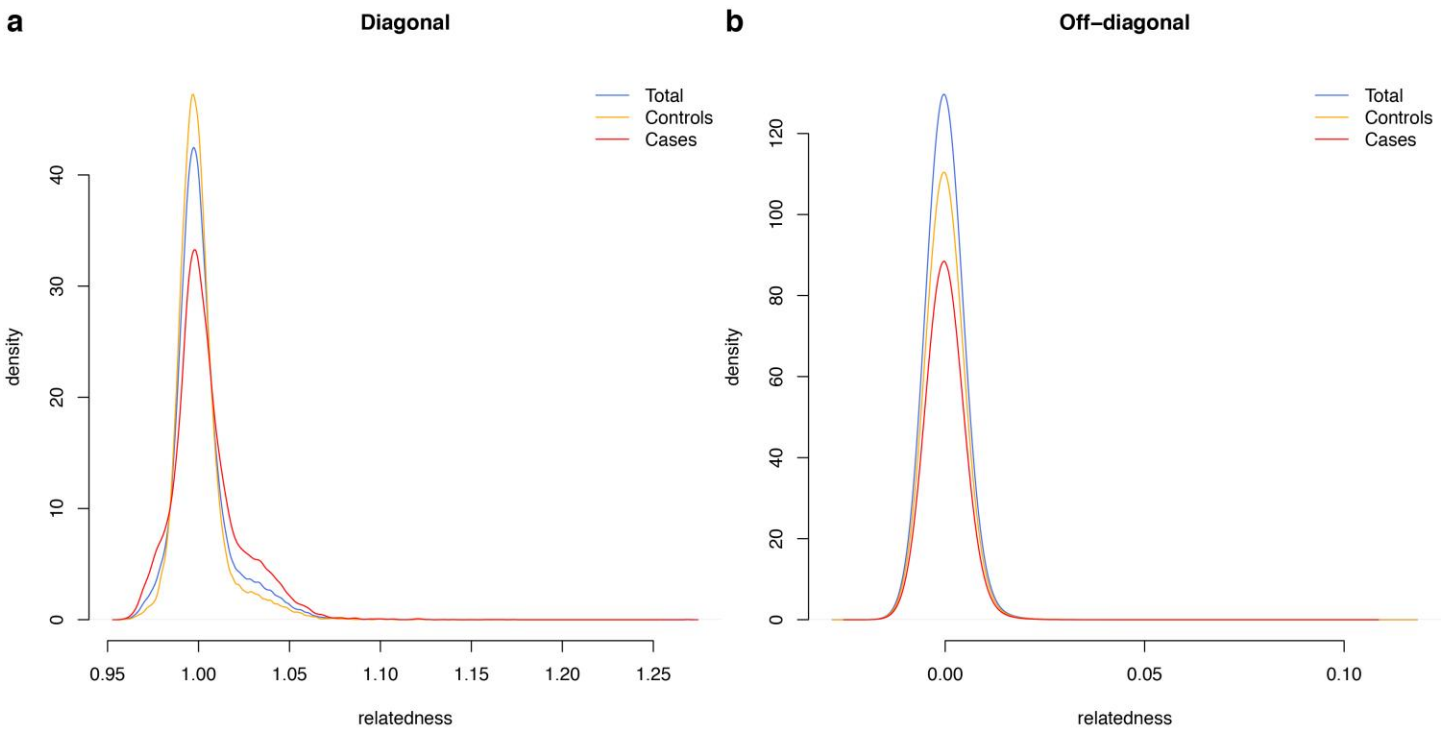
Individuals projected onto the first two principal components calculated on HapMap 3 individuals. NL, Netherlands; Ir, Ireland; It, Italy; No, Norway; Sp, Spain; Sw, Sweden; US, United States.



Supplementary Figure 13

***C21orf2* rare variant analysis quality control.**

Quantile–quantile plots for the single-SNV regression of chromosome 21 including the first ten principal components as covariates are displayed on the left. Principal-components plots for each stratum are shown on the right. Population stratification was successfully corrected for, resulting in well-behaved burden tests without evidence for overall inflation of the test statistic.



Supplementary Figure 14

Genetic relationship matrix distribution.

(**a**) Diagonal values of the SNP-based GRM. (**b**) Distribution of the off-diagonal values representing the relatedness between samples. Pairs of individuals whose relatedness exceeded 0.05 were excluded from the heritability estimations. GRM, genetic relationship matrix.

Supplementary Note

Custom reference panel sequencing and quality control

Whole genome sequencing

ALS cases from The Netherlands and population-based controls matched for sex, age and geographical region were selected for whole genome sequencing as well as the affected individuals of 11 identical discordant twin pairs. All DNA samples were sequenced with Illumina's FastTrack services (San Diego, USA) using PCR free library preparation and paired-end (100bp) sequencing on the HiSeq 2500 platform (Illumina®, San Diego, Illumina) to yield 35X coverage at minimum. Reads were aligned to the hg19 human genome build using BWA alignment software and the Isaac variant caller was used to call single nucleotide variants. Quality scores for reference calls were retrieved from genome VCFs (gVCF) and variants not passing Isaac's quality filter were set to missing.

SNV quality control

In total 75,818,355 SNVs were called. Quality control was performed using PLINK v1.9. First multi-allelic and non-autosomal variants were excluded (n = 12,430,295). Subsequently SNVs with a call-rate 0.98 (for reference and variant alleles) and those deviating from Hardy-Weinberg equilibrium ($p \leq 1.0 \times 10^{-4}$) were removed (n = 23,480,030). Finally singletons (n = 21,166,529), that are notoriously hard to phase, were excluded. Ultimately, 18,741,501 SNVs were included in our ALS-enriched reference panel.

Sample quality control

Samples discordant for reported and genetically determined gender were excluded (n = 5). Comparing genotypes derived from whole genome sequencing and the Illumina2.5M SNP array revealed high concordance between both methods (mean 99.3%, SD 2.1%) and very little discordant calls (mean 0.03%, SD 0.2%) meaning that most differences were caused by missing calls in the sequencing data or SNP-array (mean 0.6%, SD 2.0%). Five individuals were removed due to higher sequencing versus SNP-array discordance (> 1.0%). For the remaining 1,861 individuals over 98% of the 18,741,510 SNVs were successfully called (mean 99.8%). To assess population substructures all individuals were projected along the first two principal components calculated on the HapMap3 individuals (**Supplementary figure 11**). Since population diversity will improve imputation accuracy no population outliers were removed. Finally, 1,246 ALS cases and 615 healthy controls passed quality control and were included in the reference panel.

GWAS discovery phase sample ascertainment

The final number of included cases and controls for this GWAS may differ from the original papers due to overlapping cohorts. All cases and controls gave written informed consent and the relevant institutional review boards approved this study.

1. NL1, Population based ALS registry, The Netherlands.

Cases were diagnosed with probable or definite ALS according to the 1994 El-Escorial Criteria by neurologists specialized in motor neuron diseases[1]. Tertiary referral centers for ALS were University Medical Center Utrecht, Academic Medical Centre Amsterdam and Radboud University Medical Center Nijmegen. Cases with a family history for ALS were excluded. Control individuals were free of any neuromuscular disease and matched for age, gender and ethnicity. Healthy controls were either spouses of ALS patients or ascertained through a population-based study on ALS in the Netherlands[2]. All participants were from Dutch descent where all four grandparents were born in the Netherlands. More details were described previously[3].

2. BE1, University Hospital Gasthuisberg Leuven, Belgium.

Cases were diagnosed with probable or definite ALS according to the 1994 El-Escorial Criteria by neurologists specialized in motor neuron diseases[1]. Cases with a family history for ALS were excluded. Control individuals were free of any neuromuscular disease and matched for age, gender and ethnicity. Healthy control individuals were married into families that were participating in other genetic studies of neurological diseases. All participants reported Flemish descent for at least three generations. More details were described previously[4].

3. NL2, Population based ALS registry, The Netherlands.

Cases were diagnosed with probable or definite ALS according to the 1994 El-Escorial Criteria by neurologists specialized in motor neuron diseases[1]. Tertiary referral centers for ALS were University Medical Center Utrecht, Academic Medical Centre Amsterdam and Radboud University Medical Center Nijmegen. Cases with a family history for ALS were excluded. Control individuals were free of any neuromuscular disease and matched for age, gender and ethnicity ascertained through a population based study on ALS in the Netherlands[2]. Additional controls were included from a genome-wide association study on schizophrenia; these were healthy individuals without a psychiatric history. All participants were from Dutch descent where at least three out of four grandparents were born in the Netherlands. More details were described previously[4].

4. SW1, Umeå University ALS clinic, Sweden.

Cases were diagnosed with probable or definite ALS according to the 1994 El-Escorial Criteria by neurologists specialized in motor neuron diseases[1]. Cases with a family history for ALS were excluded. Control individuals were free of any neuromuscular disease and matched for age, gender and ethnicity. Healthy controls were spouses of ALS patients or patients with other neurological diseases. All participants were from Swedish descent all reporting Northern Swedish citizenship for at least three generations. More details were described previously[4].

5. NL3, Rotterdam Study, Erasmus University Medical Centre Rotterdam, The Netherlands.

Healthy controls were participants in the Rotterdam Elderly study. This is a prospective population based study among elderly living in Ommoord, Rotterdam. All healthy controls are 55 years or older. More details were described previously[5].

6. FR1, Hôpitaux de Paris, Hôpital de la Salpêtrière Paris, France.

Cases were diagnosed with probable or definite ALS according to the 1994 El-Escorial Criteria[1]. Cases with a positive family history for ALS were excluded. Controls were healthy individuals ascertained in France. More details were described previously[6].

7. UK1, Institute of Psychiatry, King's College London, United Kingdom.

Cases were diagnosed with probable or definite ALS according to the 1994 El-Escorial Criteria by neurologists specialized in motor neuron diseases[1]. Cases with a positive family history for ALS were excluded. Control samples were collected from neurologically normal, unrelated individuals, either spouses of ALS patients, carers or blood donors from the same geographical region. More details were described previously[6].

8. US1, Massachusetts General Hospital Boston and Emory University Hospital Atlanta, United States of America.

Cases were diagnosed with probable or definite ALS according to the 1994 El-Escorial Criteria by neurologists specialized in motor neuron diseases[1]. Cases with a positive family history for ALS were excluded. Control samples were healthy volunteers from Boston or spouses of ALS patients. Additional healthy control samples were purchased from the Coriell Institute for Medical Research. More details were described previously[6].

9. IR1, Population based ALS registry, Ireland.

Cases were diagnosed with probable or definite ALS according to the 1994 El-Escorial Criteria by neurologists specialized in motor neuron diseases at Beaumont Hospital in

Dublin[1]. Patients were referred from all regions in Ireland and were part of an ongoing population-based prospective ALS registry. Original ascertainment criteria excluded patients with a self-reported family history of ALS; post-hoc detailed family history review revealed that 4.07% of patients in fact had a family history of ALS, defined as a first, second or third degree relative with ALS or FTD. Control samples were healthy individuals matched for gender and age. They were either spouses or those accompanying patients to the ALS clinic. All individuals reported Irish ancestry for at least three generations. More details were described previously[7].

10. IR2, Population based ALS registry, Ireland.

Cases were diagnosed with probable or definite ALS according to the 1994 El-Escorial Criteria by neurologists specialized in motor neuron diseases Beaumont Hospital in Dublin[1]. Patients were referred from all regions in Ireland and were part of an ongoing population-based prospective ALS registry. 19.44% of ALS patients in the cohort had a family history of ALS, defined as a first, second or third degree relative with ALS or FTD. Control samples were healthy individuals matched for gender and age. They were either spouses or those accompanying patients to the ALS clinic. All individuals reported Irish ancestry for at least three generations. More details were described previously[8].

11. UK2, UK Biobank for Motor Neuron Disease Research, United Kingdom.

Cases were diagnosed with ALS in one of 20 UK hospitals by neurologists specialized in motor neuron diseases. Patients had no family history for ALS. The all participated in the UK National Biobank for Motor Neuron Disease Research. Patients had no family history for ALS and were of self-reported European descent. More details were described previously[9].

12. US2, Coriell Institute for Medical Research, United States of America.

Controls were healthy individuals obtained from the NINDS Neurogenetics repository at the Coriell Institute for Medical Research. More details were described previously[10]. Data were accessed via the dbGaP (phs000101.v3.p1, phg000018v2).

13. US3, Coriell Institute for Medical Research, United States of America.

All cases were diagnosed with probable, probable lab supported or definite ALS according to the revised El-Escorial Criteria[11]. Patients were diagnosed at different sites throughout the United States of America. Patients with a family history of ALS were excluded. More details were described previously[12]. Data were accessed via dbGaP (phs000101.v3.p1, phg000073v1).

14. IT1, Population based registry (PARALS), Italy.

All cases were diagnosed with probable, probable lab supported or definite ALS according to the revised El-Escorial Criteria[11]. Both patients with and without a positive family history for ALS were included. Patients were ascertained through a population-based registry in the northwestern Italy (PARALS). More details were described previously[13]. Data were accessed via dbGaP (phs000101.v3.p1, phg0001161v1). The number of cases differ from that reported due to sample overlap with other cohorts accessed via dbGaP.

15. IT2, Coriell Institute for Medical Research, United States of America.

Cases were obtained from the NINDS Neurogenetics repository at the Coriell Institute for Medical Research. They were diagnosed with probable, probable lab-supported or definite ALS according to the revised El-Escorial Criteria[11]. The patients had no family history for ALS and were white non-Hispanic individuals. More details were described previously[10]. Data were accessed via dbGaP (phs000101.v3.p1, phg000073v1).

16. UK3, Wellcome Trust Case Control Consortium UK National Blood Service, United Kingdom.

Healthy controls were obtained from the WTCCC National Blood Service control group. Controls were blood donors from England and North Wales. More details are provided at www.wtccc.org.uk.

17. UK4, Wellcome Trust Case Control Consortium 1958 Birth Cohort, United Kingdom.

Healthy controls were obtained from the WTCCC 1958 Birth Cohort control group. Control individuals are part of a population-based cohort that included individuals born in a single week in 1958 in England, Scotland and Wales. More details are provided at http://www2.le.ac.uk/projects/birthcohort/1958bc and www.wtccc.org.uk.

18. US4, Coriell Institute for Medical Research, United States of America.

Controls were elderly obtained from the NINDS Neurogenetics repository at the Coriell Institute for Medical Research. Individuals were non-Hispanic whites from the United States. They were age matched to a previous case cohort in Parkinson's disease. More details were described previously[14]. Data were accessed via dbGaP (phs000126.v1.p1, phg000022.v1.p1)

19. US5, The NeuroGenetics Research Consortium (NGRC), United States of America.

Controls were volunteers or spouses of individuals with Parkinson's disease. All individuals were healthy and free of neurodegenerative diseases. They were self-reported of European

descent. More details were described previously[15]. Data were accessed via dbGaP (phs000196.v1.p1, phg000066.v1.p1)

20. FIN1, Population-based cohort, Finland.

Controls were obtained from a population-based study of the elderly in Finland. All individuals were Finnish citizens and not demented at the age of 85 years. More details were described previously[16,17]. Data were accessed via dbGaP (phs000344.v1.p1, phg000129.v1.p1)

21. FIN2, Helsinki University Central Hospital, Finland.

Cases were diagnosed with ALS according to the 1994 El-Escorial criteria by a neurologist specialized in motor neuron diseases[1]. Individuals were referred to the specialized ALS clinic from throughout Finland. Both patients with and without a family history for ALS were included. More details were described previously[16]. Data were accessed via dbGaP (phs000344.v1.p1, phg000129.v1.p1)

22. FIN3, Population-based cohort, Finland.

Controls were obtained from a population-based study of elderly in Finland. All individuals were Finnish citizens and not demented at the age of 85 years. More details were described previously[16,17]. Data were accessed via dbGaP (phs000344.v1.p1, phg000129.v1.p1)

23. IT3, Italian Consortium for the Genetics of ALS (SLAGEN), Italy.

Cases and controls were collected by the Italian Consortium for the Genetics of ALS (SLAGEN) including six different neurological hospitals in Italy (IRCCS Istituto Auxologico Italiano Milano, IRCCS Istituto Neurologico Besta Milano, IRCCS Mondino, Pavia, Università degli Studi del Piemonte Orientale "Amedeo Avogadro" Novara, IRCCS Ospedale Maggiore Policlinico, Milano, Università degli Studi di Padova). All cases were diagnosed with probable or definite ALS according to the revised El-Escorial Critera[11]. Cases had no family history for ALS. Control samples were healthy blood donors collected via the Istituto Auxologico Italiano. More details were described previously[18].

24. FR2, 3C study, Population-based French controls, France

Controls were part of the Three-City (3C) Study in France. This is a population-based study in the cities Bordeaux, Dijon and Montpellier on the relation between vascular diseases and dementia in persons aged 65 years and older. Controls were free of any neurological disease. More details were described previously[19].

25. GER1, PopGen, Population-based German controls, Germany

Controls were selected form the PopGen biobank. These were population-based controls ascertained in Northern Germany and were free of any neurological disease. More details were described previously[20].

26. NL4, Population based ALS registry, The Netherlands.

Cases were diagnosed with probable or definite ALS according to the revised El-Escorial Criteria by neurologists specialized in motor neuron diseases[11]. Both cases with and without a family history for ALS were included. Tertiary referral centers for ALS were University Medical Center Utrecht, Academic Medical Centre Amsterdam and Radboud University Medical Center Nijmegen. Both patients with and without a family history for ALS were included. Control individuals were free of any neuromuscular disease and matched for age, gender and ethnicity ascertained through a population based study on ALS in the Netherlands[2].

27. GER2, Hannover Medical School and University of Wuerzburg, Germany.

Hannover. Patients were diagnosed with possible, probable or definite ALS according to the revised El-Escorial criteria[11]. They were seen by neurologists specialized in motor neuron diseases at the ALS outpatient clinic at Hannover Medical School.

Wuerzburg. Patients were diagnosed with probable or definite ALS according to the revised El-Escorial criteria[11]. Both patients with and without a family history for ALS were included. Healthy controls were obtained at the Department of Transfusion Medicine and Immunohematology at the University of Wuerzburg.

28. IT4, Population based ALS registries (SLALOM and SLAP), Italy.

Patients were diagnosed with possible, probable or definite ALS according to the revised El-Escorial criteria[11]. They are population-based cases from the Piemonte and Valle d'Aosta register(PARALS). Both patients with and without a family history for ALS (one first degree relative or two second/third degree relatives) were included. Control subjects were healthy individuals free of neuromuscular diseases, matched for age, gender and geographical region (Piemonte registry).

29. PU1, University of Lisbon, Portugal.

Patients were diagnosed with possible, probable or definite ALS according to the revised El-Escorial criteria[11] by neurologists specialized in motor neuron diseases. Both cases with and without a family history (third degree relatives) were included. Control subjects were spouses or those accompanying patients to the clinic.

30. SP1, University of Madrid, University of Barcelona, Spain.

Patients were diagnosed with possible, probable or definite ALS according to the revised the El Escorial criteria[11]. They were seen by neurologists specialized in motor neuron diseases at the ALS Clinic of the Hospital Carlos III in Madrid and Motor Neuron Diseases Clinic at the Hospital de la Santa Creu i Sant Pau in Barcelona. Patients did not have a family history for ALS. Control samples were healthy individuals, free of neuromuscular disease, matched for gender and age. They were spouses and, when not possible, political relatives or those accompanying patients to the ALS clinic.

31. SWISS1, Kantonsspital St. Gallen, Switzerland.

Cases were diagnosed with probable or definite ALS according to the revised El-Escorial Criteria by neurologists specialized in motor neuron diseases[11]. All patients were diagnosed at tertiary referral centers. Patients had no family history for ALS. Control samples were aged-matched healthy blood donors and primary care givers of ALS patients.

32. BE2, University Hospital Gasthuisberg Leuven, Belgium.

Cases were diagnosed with probable or definite ALS according to the revised El-Escorial Criteria by neurologists specialized in motor neuron diseases[11]. Both patients with and without a family history for ALS were included. Control individuals were spouses accompanying patients to the University Hospital Gasthuisberg in Leuven Belgium. More details were described previously[21].

33. FIN4, ALS clinics, Finland (via Umeå University, Sweden)

Patients were diagnosed with probable or definite ALS according to the revised El-Escorial Criteria[11] by neurologists that were members of clinical ALS teams throughout Finland. Both patients with and without a family history for ALS were included. Ascertainment was coordinated at Umeå University, Sweden. All participants were of Finnish descent.

34. IR3, Population based ALS registry, Ireland.

Cases were diagnosed with possible, probable or definite ALS according to the revised El-Escorial Criteria[11] by neurologists specialized in motor neuron diseases at Beaumont Hospital in Dublin. Patients were referred from all regions in Ireland and were part of an ongoing population-based prospective ALS registry. 12.92% of ALS patients in the cohort had a family history of ALS, defined as a first, second or third degree relative with ALS or FTD. Control samples were healthy individuals matched for gender and age. They were either spouses, those accompanying patients to the ALS clinic or neurologically healthy individuals sampled from across Ireland.

35. SW2, ALS clinics, via Umeå University, Sweden.

Patients were diagnosed with probable or definite ALS according to the revised El-Escorial Criteria[11] by neurologists that were members of clinical ALS teams throughout Sweden. Ascertainment was coordinated at Umeå University, Sweden. Both cases with and without a family history for ALS were included. Control individuals were free of any neuromuscular disease and matched for age, gender and ethnicity. Healthy controls were spouses of ALS patients or patients with other neurological diseases (migraine, epilepsy etc.). All participants were of Swedish descent for at least three generations.

36. US6, University of California Los Angeles, United States of America.

Cases were identified at the ALS clinical centers of University of California Los Angeles and University of California San Francisco. Patients fulfilled the El-Escorial Criteria[11] for definite or probable ALS. Patients had no family history for ALS or FTD. Control participants were population-based individuals.

37. GER3, University of Ulm, Germany.

Cases were identified at the ALS clinical center at the Ulm University. All patients were diagnosed according to the EFNS Consensus criteria[22]. Patients with and without a family history of ALS were included.

38. FR3, Paris and Tours, France.

Cases were diagnosed with probable or definite ALS according to the revised El-Escorial Criteria by neurologists specialized in motor neuron diseases at the ALS National Reference Center of Pitié-Salpêtrière Hospital (Paris). Additional cases were diagnosed with probable or definite ALS according to the 1994 El-Escorial Criteria[1] by neurologists specialized in motor neuron diseases in the ALS Center of the Regional Hospital of Tours. Both patients with and without a family history for ALS were included.

39. US7, Massachusetts General Hospital, Boston, Emory University, Atlanta and the Penn ALS Center, University of Pennsylvania, United States of America.

Cases were diagnosed with probable or definite ALS according to the revised El-Escorial Criteria by neurologists specialized in motor neuron disease[11]. Cases were of self-reported European-American descent. Ascertainment was coordinated from. Researchers from Boston (Massachusetts General Hospital) and Atlanta (Emory University) coordinated the ascertainment. Additional cases were ascertained through an integrated clinical database at the Penn ALS Center of the University of Pennsylvania. These patients were diagnosed with ALS

according to the El-Escorial Criteria[11] by neurologists specialized in motor neuron diseases. Both patients with and without a family history for ALS were included. More details were described previously[23].

40. UK5, UK Biobank for Motor Neuron Disease Research, United Kingdom.

Cases were diagnosed with ALS in one of 20 UK hospitals by neurologists specialized in motor neuron diseases. They all participated in the UK National Biobank for Motor Neuron Disease Research. Patients had no family history for ALS and were of self-reported European descent.

41. NL5, Population based ALS registry, The Netherlands.

Cases were diagnosed with probable or definite ALS according to the revised El-Escorial Criteria by neurologists specialized in motor neuron diseases[11]. Tertiary referral centers for ALS were University Medical Center Utrecht, Academic Medical Centre Amsterdam and Radboud University Medical Center Nijmegen. Both cases with and without a family history for ALS were included. Control individuals were free of any neuromuscular disease and matched for age, gender and ethnicity ascertained through a population based study on ALS in the Netherlands[2].

GWAS replication phase sample ascertainment

1. Sydney, Australia

Patients and controls were ascertained from Macquarie Neurology at Macquarie University, Sydney, neurogenetic clinics at Concord Hospital, Sydney, as well as the Australian MND DNA bank. Patients were diagnosed with definite or probable ALS according to the revised El-Escorial Criteria[11]. Patients with a family history for ALS were excluded. Control subjects were healthy individuals free of neuromuscular diseases.

2. *University Hospital Gasthuisberg Leuven, Belgium.*

Cases were diagnosed with probable or definite ALS according to the revised El-Escorial Criteria by neurologists specialized in motor neuron diseases[11]. Both patients with and without a family history for ALS were included. Control individuals were spouses accompanying patients to the University Hospital Gasthuisberg in Leuven Belgium.

3. University Hospital Tours, France.

Cases were diagnosed with probable or definite ALS according to the 1994 El-Escorial Criteria by neurologists specialized in motor neuron diseases in the ALS Center of the

Regional University Hospital of Tours[1]. Both patients with and without a family history for ALS were included. Control individuals were free of any neuromuscular disease.

4. *Jena, Germany*

Patients were diagnosed with possible, probable or definite ALS according to the revised El-Escorial Criteria by neurologists specialized in motor neuron diseases at Hans-Berger-Department of Neurology, University Hospital Jena[11]. Both cases with and without a family history of ALS were included. Control samples were healthy individuals.

5. Population based ALS registry, Ireland.

Cases were diagnosed with possible, probable or definite ALS according to the revised El-Escorial Criteria by neurologists specialized in motor neuron diseases at Beaumont Hospital in Dublin.[11] Patients were referred from all regions in Ireland and were part of an ongoing population-based prospective ALS registry. Both cases with an without a family history for ALS were included. Control samples were healthy individuals matched for gender and age. They were either spouses or those accompanying patients to the ALS clinic.

6. Population based ALS registries, Italy.

Patients were diagnosed with possible, probable or definite ALS according to the revised El-Escorial criteria[11]. They were seen by neurologists specialized in motor neuron diseases at tertiary referral centers throughout Italy. Both cases with and without a family history for ALS were included. Control subjects were healthy individuals free of neuromuscular diseases, matched for age, gender and geographical region (Piemonte and Puglia only) within Italy.

7. Population based ALS registry, The Netherlands.

Cases were diagnosed with probable or definite ALS according to the revised El-Escorial Criteria by neurologists specialized in motor neuron diseases[11]. Tertiary referral centers for ALS were University Medical Center Utrecht, Academic Medical Centre Amsterdam and Radboud University Medical Center Nijmegen. Both cases with and without a family history for ALS were included. Control individuals were free of any neuromuscular disease and matched for age, gender and ethnicity ascertained through a population based study on ALS in the Netherlands[2].

8. Population based ALS registry, Turkey

Patients were diagnosed with ALS possible, probable or definite ALS according to the revised El-Escorial Criteria by neurologists specialized in motor neuron diseases at the Boğaziçi

University in Istanbul, Turkey[11]. Cases with a positive family history for ALS were excluded. Control samples were healthy individuals without any known history of neurological disorders collected from the Microbiology Department of Haydarpaşa State Hospital in Istanbul.

GWAS quality control

Quality control (QC) was first performed per cohort to remove low quality SNPs and individuals using PLINK 1.9. SNPs were first annotated according to dbSNP137 and mapped to the hg19 reference genome. Subsequently, multi-allelic and AT/CG SNPs were removed as well as SNPs with a call-rate < 98%, < 10 minor allele observations per cohort, biased missingness as determined by haplotype and non-autosomal SNPs. Individuals with gender mismatches or an excessive number of heterozygous SNPs (F < -0.2) were removed. To check strand inconsistencies or annotation errors, allele frequencies between each cohort were compared to those observed in the European population represented in 1000GP. The number of SNPs and individuals failing each QC step within each cohort are described in **Supplementary Table 2**. Considering the low number of overlapping SNPs between all different platforms (n = 48,229), cohorts and the presence of cohorts with cases only or controls only, cohorts were combined based on reported nationality and genotyping platform. Quality control per stratum included removal of SNPs that deviated from Hardy-Weinberg equilibrium ($p < 1 \times 10^{-5}$ and $p < 1 \times 10^{-9}$ in controls and cases respectively) and those with biased missingness between cases and controls ($p < 1 \times 10^{-3}$). Subsequently, related and duplicate individuals across all strata (pi-hat > 0.1) were removed. Individuals were projected along the first four principal components (PC) calculated on HapMap3 individuals using EIGENSTRAT. Population outliers, defined as deviation > 10SD from the HapMap CEU population mean on PCs 1-4 or > 4SD from its stratum mean on PC1-2, were removed. After removing population outliers, PCs were recalculated on a LD-pruned set of SNPs for each stratum and again outliers (> 5SD on PC1-4) were removed. The procedure for removing outliers, using PCs and its results are summarized in **Supplementary Figure 12**. After removing all outliers, PCs were again recalculated. Based on scree plots per stratum the eigenvectors for the first 1-4 PCs were included in the logistic regression. To calculate genomic inflation factors per stratum, the test statistic's empirical quantiles were obtained applying logistic regression in an additive model. The number of SNPs and individuals failing each QC step across all strata are displayed in **Supplementary Table 3**.

Genotyping experiments in replication phase

TaqMan assay (rs75087725 and rs616147)

The reaction mix included 1.0 µl of genomic DNA (10 ng/µl), 0.25 µl TaqMan genotyping assay 20X (Life Technologies), 2.5 µl MasterMix 2X (Life Technologies) and 1.25 ul MilliQ. The thermocycler program included 30 sec at 60˚C, 10 min at 95˚C followed by 40 cycles of 15 sec at 95˚C and 1 min at 60˚C and a final step of 30 sec at 60˚C. Fluorescent signals were analyzed on a QuantStudio 6 Flex Real-Time PCR System and genotypes were determined by allelic discrimination using QuantStudio 6 Flex Real-Time PCR System Software (Life Technologies).

Sanger sequencing (rs10139154 and rs7813314)

The oligonucleotide primers were used for amplification (**Supplementary Table 7**). PCR reactions consisted of 1.0 µL genomic DNA (50 ng/µl), 1.0 µL 10xNH4 reaction buffer, 0.2 µL dinucleotide triphosphate (dNTP; 10mM each), 0.3 µL MgCl2 (50mM), 0.1 µL Biotaq (5U/µL), 7.0 µL milli-Q, 0.2 µL of each primer (10 µM) in a total volume of 10 µL. The thermocycler program included initialization of 3 min at 96˚C followed by 35 cycles of 30 sec at 96˚C, 45 sec at 57˚C and 1 min at 72˚C and a final step of 5 min at 72˚C. The PCR products were electrophoresed in 1.2-1.5% agarose gel containing 0.02‰ Ethidium Bromide and visualized using a Proxima AQ 4.2 Imager. Based on initial sequencing results, the reverse primer was used for analysis of both loci. The thermocycler program for sequencing consisted of 1 min at 96˚C followed by 25 cycles of 10 sec at 96˚C, 5 sec at 50˚C and 4 min at 60˚C. Sequencing was performed on an Applied Biosystems 3730 DNA-Analyzer using BigDye Terminator 3.1 sequencing kit (Applied Biosystems).

Cases and controls were randomly assigned on plates and experimenters were blinded for case-control status when calling genotypes.

Rare variant association analysis for *C21orf2*

Data merging

For burden testing in the *C21orf2* locus, cases and controls were available from the following collection sites: the Netherlands (N = 2,089), Belgium (N = 421), Ireland (N = 386), the United Kingdom (N = 687), and the United States (N = 393). Samples were received as individual-level genome VCF (gVCF) files containing single nucleotide variants (SNVs) and insertions/deletions (indels). The gVCF files were merged together by cohort using the Illumina aggregation ("agg") tool available on GitHub (https://github.com/Illumina/agg). The agg tool ensures that all samples in a given cohort were genotyped at the union of all sites

observed across all samples. The resulting file is a (per-chromosome) VCF file that contains only those sites where at least one non-reference allele is observed in the set of cases and controls.

Sample quality control

Sample and site QC were performed using PLINK1.9 and the most recent version of VCFtools. After the samples within each cohort had been merged into per-chromosome VCF files, we removed all sites with a quality (QUAL) score < 30. Genotypes with a genotype quality < 10 were set to missing on an individual-level basis. We then checked all samples for missingness across the autosomal chromosomes; all samples had missingness < 10% and thus no samples were excluded at this point in QC. Principal components were calculated on a set of LD-pruned ($R^2 > 0.2$) high-quality sites (MAF > 0.1, call-rate > 99.9%, not A/T or C/G, biallelic, outside the MHC, LCT locus and inversions on chromosome 8 and 17). All cases and controls were projected on to the reference HapMap3 populations using EIGENSTRAT. Individuals not of European ancestry (using the European populations in HapMap 3 as a reference) were removed from the analysis. The cleaned principal component plots are displayed in **Supplementary figure 13**. Using this same set of high-quality SNPs, we also removed samples with excessively high or low inbreeding coefficients (further than 3 standard deviations from the mean of the inbreeding coefficient distribution in the cohort). A single individual, preferably cases or those with the lowest overall missingness, from each related pair (pi-hat > 0.125, indicating a cousin relationship or closer) was removed. Finally, we checked the average depth of coverage across each of the samples; all samples looked appropriately covered (within 6 SD of the mean of the average depth distribution) and no sample was removed at this step.

Site quality control

We then performed site QC on sites on chromosome 21 only (as this is the chromosome containing the *C21orf2* locus). All sites with missingness < 5% or with an average read depth < 5 or > 50 were removed from the analysis. Sites out of Hardy-Weinberg in controls ($p < 10^{-6}$) or with excessive differential missingness between cases and controls ($p < 10^{-6}$) were also removed. Once sample and site QC was complete, the following samples were available for burden testing:

(a) The Netherlands: 1275 cases, 677 controls

(b) Ireland: 251 cases, 134 controls

(c) Belgium: 260 cases, 133 controls

(d) United States: 275 cases, 65 controls

(e) United Kingdom: 501 cases, 129 controls

Functional annotation

Sites passing QC were functionally annotated using ANNOVAR. All variants with a MAF < 0.05 that were either non-synonymous or loss of function (stop gain or frameshift indel) were kept for burden testing.

Association testing

Principal components were recalculated per cohort on cases and controls that passed quality control. To check that the calculated PCs appropriately corrected for population stratification, we ran single-variant analysis (per cohort) on all of chromosome 21 using logistic regression and including the top ten principal components as covariates. We plotted quantile-quantile plots of the results and checked the genomic inflation factor for signs of population stratification. The data looked appropriately distributed (highest lambda = 1.06, in the Netherlands cohort) and thus we proceeded to burden testing in *C21orf2* (**Supplementary Figure 13**).

Per cohort, we extracted all non-synonymous and loss of function variants in *C21orf2* with MAF < 5% for burden testing. Burden tests were corrected with the top ten principal components as covariates. Burden testing was performed using ScoreSeq, which performs five burden tests: T1, T5, Frequency-weighted (also called the Madsen-Browning test), SKAT, and EREC. The first three tests assume the same direction of effect of all variants across a given locus. The last two tests allow for multiple directions of effect in the same locus. We set the software to perform 1 million permutations to evaluate significance of the burden test. Subsequently, these results were meta-analyzed using MASS, the companion software to ScoreSeq.

eQTL analysis

Linkage disequilibrium data and *cis*-eQTLs for different tissues were acquired from SNiPA webtool (http://snipa.helmholtz-muenchen.de/snipa/, accessed 28.03.2016)[24], using 1000G p1v3 for European population and ENSEMBL v80 genome annotations. Brain *cis*-eQTLs were manually curated from several brain eQTL studies (**Supplementary Table 7**)[25-36]. Additionally, *cis*-eQTLs effects observed in non-brain tissues were described for all SNPs within the genome-wide significant loci[36-40]. For BRAINEAC database, *cis*-eQTLs were

queried for SNP-gene combinations locating in the suggestive GWAS loci (LMM, $p < 10^{-4}$), as defined by DEPICT (v139, $R^2 > 0.5$).

Stranded RNA-seq tracks were acquired from Epigenome Roadmap Project data portal (http://egg2.wustl.edu/roadmap/web_portal/) for one fetal brain[41].

Supplementary Tables

Supplementary Table 1: Description of cohorts

Cohort	*Name*	*Country*	*Cases*	*Controls*	*Platform*
1	NL1	The Netherlands	461	450	Illumina317K
2	BE1	Belgium	311	371	Illumina370K
3	NL2	The Netherlands	582	629	Illumina370K
4	SW1	Sweden	493	500	Illumina370K
5	NL3	The Netherlands	0	5,974	Illumina550K
6	FR1	France	251	724	Illumina317K
7	UK1	United Kingdom	245	221	Illumina317K
8	US1	United States	753	811	Illumina317K
9	IR1	Ireland	221	211	Illumina550K
10	IR2	Ireland	103	127	Illumina610K
11	UK2	United Kingdom	661	0	Illumina550K
12	US2	United States	0	527	Illumina550K
13	US3	United States	276	0	Illumina550K
14	IT1	Italy	141	0	Illumina610K
15	IT2	Italy	261	246	Illumina550K
16	UK3	United Kingdom	0	2,501	Illumina1M
17	UK4	United Kingdom	0	2,699	Illumina1M
18	US4	United States	0	867	Illumina370K
19	US5	United States	0	1,986	Illumina1M
20	FIN1	Finland	0	201	Illumina370K
21	FIN2	Finland	401	191	Illumina370K
22	FIN3	Finland	0	103	Illumina1M
23	IT3	Italy	1,792	1,107	Illumina660W
24	FR2	France	0	1,100	Illumina550K
25	GER1	Germany	0	677	Illumina550K
26	NL4	The Netherlands	1,226	2,262	IlluminaOmniExpress
27	GER2	Germany	580	286	IlluminaOmniExpress
28	IT4	Italy	311	100	IlluminaOmniExpress
29	PU1	Portugal	40	54	IlluminaOmniExpress
30	SP1	Spain	105	63	IlluminaOmniExpress
31	SWISS1	Switzerland	228	236	IlluminaOmniExpress
32	BE2	Belgium	225	250	IlluminaOmniExpress
33	FIN4	Finland	144	0	IlluminaOmniExpress
34	IR3	Ireland	268	478	IlluminaOmniExpress
35	SW2	Sweden	281	271	IlluminaOmniExpress
36	US6	United States	65	42	IlluminaOmniExpress
37	GER3	Germany	1,519	0	IlluminaCoreExome
38	FR3	France	363	0	IlluminaOmniExpress
39	US7	United States	573	0	IlluminaOmniExpress
40	UK5	United Kingdom	1,214	14	IlluminaOmniExpress
41	NL5	The Netherlands	697	619	Illumina2.5M
Total			**14,791**	**26,898**	
New			**7,839**	**4,675**	

Supplementary Table 2: GWAS cohort SNP quality control

Cohort	*SNPs pre-QC*	*SNPs unannotated (dbSNP 137)*	*SNPs genotyping rate < 0.98*	*MAF threshold*	*SNPs MAF < threshold*	*SNPs biassed missingness*	*SNPs non-autosomal*	*SNPs post-cohortQC*	*SNP callrate*
1	317,503	2,208	5,454	0.005	181	434	8,797	300,429	0.9993
2	370,404	30,592	9,648	0.007	6,652	368	9,164	313,980	0.9985
3	370,404	30,592	2,265	0.004	6,450	685	9,726	320,686	0.9990
4	370,404	30,592	3,406	0.005	6,544	425	9,610	319,827	0.9990
5	561,466	4,556	22,125	0.001	14,479	26,506	12,281	481,519	0.9988
6	307,790	1,952	18,533	0.005	132	584	6,420	280,169	0.9984
7	307,790	1,952	33,724	0.011	240	230	7,609	264,035	0.9985
8	307,790	1,952	23,788	0.003	93	2,706	7,708	271,543	0.9985
9	561,466	3,848	12,332	0.012	24,913	228	12,428	507,717	0.9990
10	620,901	36,687	4,629	0.022	40,887	0	12,990	525,708	0.9997
11	584,414	7,613	152	0.008	26,843	139	13,157	536,510	0.9997
12	561,466	3,832	6,758	0.009	22,256	270	12,647	515,703	0.9992
13	555,351	4,768	14,891	0.018	27,330	97	12,307	495,958	0.9989
14	620,901	37,132	8,080	0.035	48,213	0	12,579	514,897	0.9996
15	555,352	3,870	8,585	0.010	21,179	262	12,587	508,869	0.9991
16	934,848	22,706	631	0.002	10,891	3,323	26,932	870,365	0.9992
17	934,010	22,904	566	0.002	10,975	3,503	26,918	869,144	0.9991
18	344,301	4,973	6,891	0.006	5,811	564	9,501	316,561	0.9988
19	1,012,895	268,736	3,118	0.003	71,887	1,831	15,562	651,761	0.9996
20	346,831	6,517	23,484	0.025	9,202	0	9,008	298,620	0.9983
21	345,111	4,948	2,849	0.008	7,703	159	9,419	320,033	0.9995
22	1,199,187	86,910	26,923	0.049	25,8479	0	23,359	803,516	0.9994
23	545,914	3,085	6,472	0.002	10,565	4,110	0	521,682	0.9995
24	582,892	7,360	23,900	0.005	24,825	1,319	11,801	513,687	0.9987
25	561,466	3,756	13,268	0.007	22,195	830	12,505	508,912	0.9989
26	719,665	8,463	19,197	0.001	55,899	9,179	14,261	612,666	0.9993
27	719,665	8,333	13,478	0.006	68,044	1,236	14,281	614,293	0.9996
28	719,665	8,333	6,673	0.012	74,986	252	14,477	614,944	0.9997
29	719,665	7,252	15,662	0.053	12,1072	0	4,905	570,774	0.9998
30	719,665	7,798	11,024	0.030	91,290	0	13,810	595,743	0.9997
31	719,665	7,252	23,449	0.011	74,919	598	13,376	600,071	0.9996
32	719,665	7,429	2,760	0.011	76,744	152	14,679	617,901	0.9996
33	719,665	7,429	6,104	0.035	107,408	0	13,402	585,322	0.9997
34	719,665	7,429	2,327	0.007	73,839	338	14,828	620,904	0.9998
35	719,665	7,429	8,392	0.009	75,575	415	14,465	613,389	0.9997
36	964,193	252,092	2,402	0.047	107,033	0	13,826	588,840	0.9996
37	529,948	273,756	2,743	0.003	9,946	680	6,652	236,171	0.9995
38	730,525	9,899	23,081	0.014	76,288	263	14,172	606,822	0.9994
39	730,525	9,899	21,804	0.009	71,357	662	14,418	612,385	0.9995
40	730,525	9,899	19,986	0.004	64,667	2,533	14,813	618,627	0.9993
41	2,391,739	109,675	62,083	0.004	702,535	5,961	30,024	1,481,461	0.9987

Supplementary Table 3: GWAS cohort individual quality control

Cohort	*Inidividuals pre-QC*	*Individuals genotyping rate < 0.98*	*Individuals |F| > 0.2*	*Individuals Gender mismatch*	*Cases post-cohortQC*	*Controls post-cohortQC*	*Observed MAF vs 1KG (r^2)*
1	911	26	0	3	444	438	0.982
2	682	38	0	8	305	331	0.984
3	1,211	9	0	5	582	615	0.983
4	993	1	0	20	483	489	0.974
5	5,974	20	0	3	0	5,951	0.987
6	975	76	0	4	232	663	0.984
7	466	122	0	5	174	165	0.981
8	1,564	169	0	9	660	726	0.987
9	432	4	0	0	217	211	0.977
10	230	0	0	1	103	126	0.970
11	661	0	0	14	647	0	0.986
12	527	6	0	1	0	520	0.986
13	276	1	0	0	275	0	0.982
14	141	1	0	0	140	0	0.955
15	507	2	0	1	259	245	0.968
16	2,501	1	0	0	0	2,500	0.987
17	2,699	11	0	0	0	2,688	0.987
18	867	4	0	1	0	862	0.987
19	1,986	0	0	1	0	1985	0.990
20	201	5	0	1	0	195	0.944
21	592	11	0	3	393	185	0.947
22	103	5	0	1	0	97	0.928
23	2,899	0	0	0	1,792	1,107	0.971
24	1,100	68	0	2	0	1,030	0.985
25	677	21	0	1	0	655	0.984
26	3,488	121	0	23	1,207	2,137	0.988
27	866	13	0	22	558	273	0.988
28	411	0	0	6	309	96	0.970
29	94	0	0	10	39	45	0.940
30	168	1	0	1	103	63	0.963
31	464	0	0	14	219	231	0.984
32	475	6	0	9	215	245	0.985
33	144	0	0	1	143	0	0.937
34	746	2	0	6	268	470	0.979
35	552	0	0	4	279	269	0.977
36	107	1	0	1	63	42	0.956
37	1,519	12	0	21	1,486	0	0.987
38	363	3	0	0	360	0	0.982
39	573	16	0	7	550	0	0.986
40	1,228	20	0	1	1,194	13	0.988
41	1,316	15	0	7	679	615	0.989

Supplementary Table 4: GWAS stratum SNP quality control

Stratum	*Cohort(s)*	*Name*	*SNPs pre-stratumQC*	*SNPs HWE (controls) $p < 10^{-5}$*	*SNPs HWE (cases) $p < 10^{-9}$*	*SNPs diff. missing $p < 10^{-3}$*	*SNPs post-stratumQC*
1	1	sNL1	300,429	61	6	9	300,353
2	2	sBE1	313,980	68	1	3	313,908
3	3+5	sNL2	278,421	527	1	1,778	276,115
4	4	sSW1	319,827	108	5	9	319,705
5	6	sFR1	280,169	106	0	642	279,421
6	7	sUK1	264,035	13	0	0	264,022
7	8+18	sUS1	269,380	195	0	2,393	266,792
8	9+10	sIR1	490,128	108	2	1	490,017
9	11+16	sUK2	505,806	478	6	763	504,559
10	12+13	sUS2	488,366	170	0	0	488,196
11	14+15	sIT1	474,737	48	8	119	474,562
12	20+21	sFIN1	291,763	117	6	16	291,624
13	26	sNL3	612,666	648	4	1,479	610,535
14	27	sGER1	614,293	56	25	5	614,207
15	28	sIT2	614,944	14	8	12	614,910
16	29+30	sIB1	564,171	9	3	0	564,159
17	31	sSWISS1	600,071	51	1	0	600,019
18	32	sBE2	617,901	82	0	4	617,815
19	35	sSW2	613,389	89	0	0	613,300
20	22+33	sFIN2	437,421	18	2	0	437,401
21	34	sIR2	620,904	211	1	2	620,690
22	19+36+39	sUS3	549,833	371	1	3,802	545,659
23	24+38	sFR2	297,657	153	0	332	297,172
24	17+40	sUK3	460,024	284	6	7,648	452,086
25	25+37	sGER2	118,407	13	8	2,961	115,425
26	23	sIT3	521,682	356	52	1,946	519,328
27	41	sNL4	1,481,461	14,583	1542	2,775	1,462,561

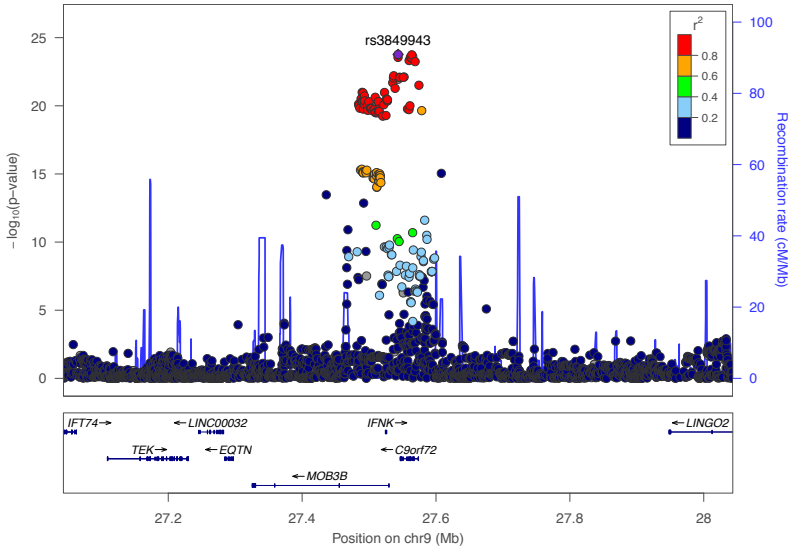
* > 10 SD from CEU (PC1-4), > 4 SD from stratum (PC 1-2)

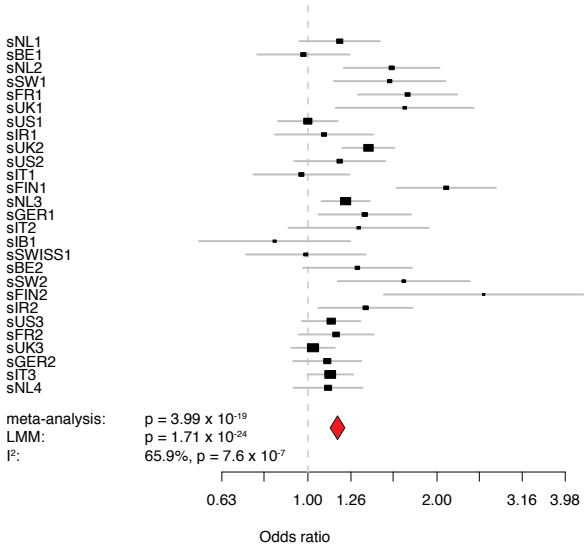
Supplementary Table 5: GWAS stratum individual quality control

Stratum	*Population outliers**	*Individuals Duplicated pi-hat > 0.9*	*Individuals Related pi-hat > 0.1*	*Individuals PC-outliers**	*Cases post-stratumQC*	*Controls post-stratumQC*	$\lambda_{(GC)}$ *Unimputed*	$\lambda_{(GC)}$ *Imputed*
1	14	10	11	4	423	420	1.001	1.008
2	4	5	9	2	299	317	1.004	1.013
3	79	964	1,040	38	145	4,882	1.011	1.025
4	21	373	22	0	288	268	1.006	1.020
5	13	70	0	3	155	654	1.017	1.024
6	8	0	1	3	168	159	1.015	1.015
7	21	277	5	8	598	1,339	0.994	1.003
8	3	10	4	1	308	331	1.001	1.010
9	25	8	1	0	614	2,687	1.008	1.013
10	6	4	3	3	266	513	1.007	1.012
11	0	14	2	2	382	244	1.008	1.020
12	1	14	2	0	378	378	1.020	1.022
13	89	392	80	2	952	1,829	1.010	1.012
14	11	23	13	8	518	258	1.013	1.010
15	4	11	1	6	290	93	1.006	1.022
16	14	3	4	4	126	99	1.018	1.044
17	11	0	11	4	203	221	1.012	1.022
18	3	1	8	1	205	242	1.003	1.008
19	7	48	22	4	232	235	1.010	1.017
20	0	6	2	0	135	97	1.026	1.072
21	4	20	5	2	264	443	1.001	1.007
22	45	21	7	5	559	2,003	1.007	1.005
23	29	10	10	9	327	1,005	1.003	1.000
24	59	95	13	6	1,032	2,502	1.024	1.015
25	56	23	11	4	1,399	648	1.000	1.009
26	8	41	9	51	1,715	1,075	1.015	1.013
27	56	71	32	6	596	533	1.001	0.998

* > 5 SD from stratum on PC 1–4

Supplementary Table 6: rs3849943 details

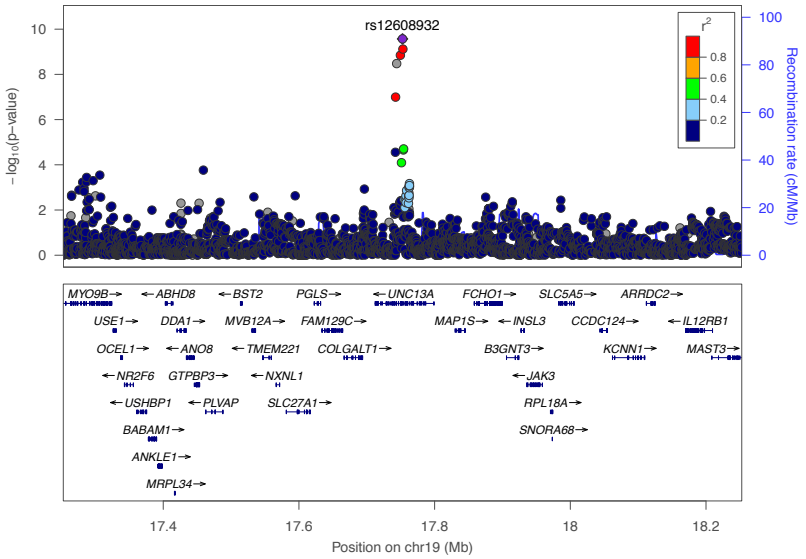


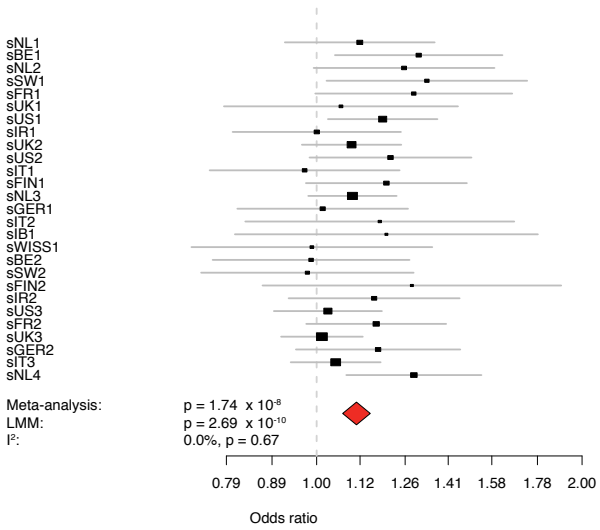


Stratum information rs3849943

NAME	IMPUTED	MAF_{CASES}	$MAF_{CONTROLS}$	HWE	INFO	$OR_{STRATUM}$	$P_{STRATUM}$
sNL1	1	0.2636	0.2298	0.1281	1.0000	1.1864	0.119
sBE1	1	0.2475	0.2524	0.1362	1.0000	0.9761	0.849
sNL2	1	0.3207	0.2292	0.2738	0.9998	1.5676	8.06×10^{-4}
sSW1	1	0.2813	0.1977	0.8460	0.9999	1.5511	3.64×10^{-3}
sFR1	1	0.3458	0.2342	0.5863	0.9978	1.7078	9.60×10^{-5}
sUK1	1	0.3036	0.2107	0.6294	0.9999	1.6815	0.005
sUS1	1	0.2416	0.2401	0.5499	0.9999	0.9995	0.996
sIR1	1	0.2371	0.2221	0.4261	0.9995	1.0905	0.520
sUK2	1	0.3038	0.2407	0.1130	0.9996	1.3843	6.03×10^{-6}
sUS2	1	0.2669	0.2359	0.6243	0.9999	1.1859	0.173
sIT1	1	0.2745	0.2848	0.6382	0.9991	0.9656	0.790
sFIN1	1	0.2963	0.1667	0.7117	1.0000	2.1011	2.27×10^{-8}
sNL3	1	0.2589	0.2233	0.2512	1.0000	1.2245	0.002
sGER1	1	0.2799	0.2229	0.5928	1.0000	1.3566	0.015
sIT2	1	0.2862	0.2313	0.7714	0.9996	1.3134	0.149
sIB1	1	0.2778	0.3172	0.1598	0.9981	0.8364	0.389
sSWISS1	1	0.2537	0.2489	0.3690	0.9997	0.9884	0.944
sBE2	1	0.2854	0.2314	0.2788	1.0000	1.3042	0.075
sSW2	1	0.2565	0.1702	0.4949	1.0000	1.6732	0.004
sFIN2	1	0.2889	0.1701	0.7313	1.0000	2.5689	2.97×10^{-4}
sIR2	1	0.2728	0.2168	0.5754	0.9996	1.3633	0.016
sUS3	1	0.2612	0.2319	1.0000	0.9999	1.1331	0.123
sFR2	1	0.2599	0.2254	0.0458	1.0000	1.1632	0.142
sUK3	1	0.2650	0.2598	0.6400	0.9998	1.0279	0.647
sGER2	1	0.2605	0.2277	0.2164	0.9646	1.1095	0.265
sIT3	1	0.2946	0.2721	0.8776	0.9997	1.1273	0.053
sNL4	1	0.2500	0.2298	0.0099	1.0000	1.1144	0.252

Supplementary Table 7: rs12608932 details

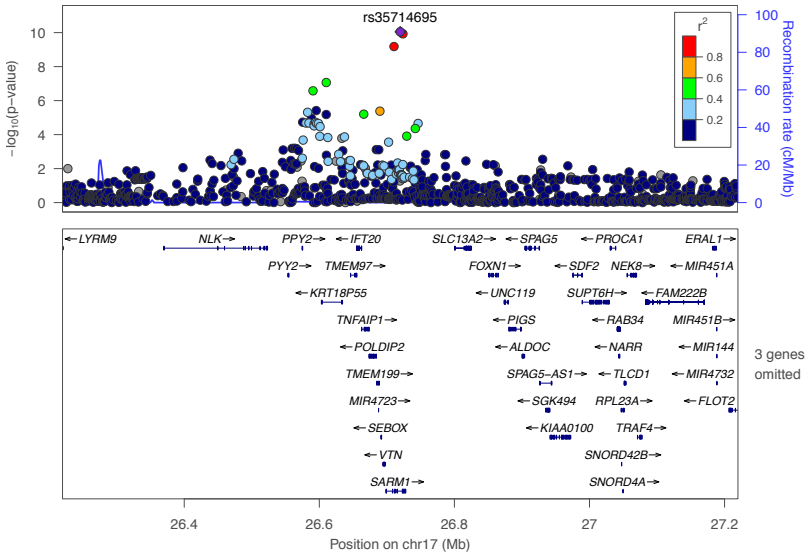


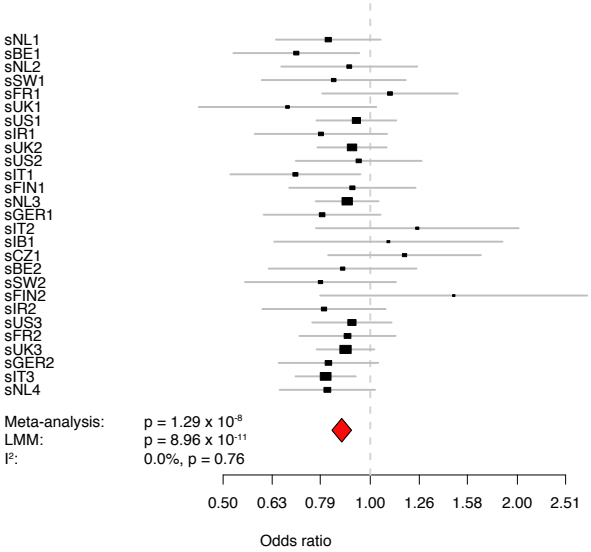


Stratum information rs12608932

NAME	IMPUTED	MAF_{CASES}	$MAF_{CONTROLS}$	HWE	INFO	$OR_{STRATUM}$	$P_{STRATUM}$
sNL1	0	0.403	0.376	0.406	1.000	1.119	0.259
sBE1	0	0.433	0.363	0.003	1.000	1.305	0.016
sNL2	0	0.407	0.347	0.343	1.000	1.256	0.060
sSW1	0	0.380	0.317	0.574	1.000	1.334	0.030
sFR1	0	0.377	0.320	1.000	1.000	1.289	0.054
sUK1	0	0.384	0.368	0.236	1.000	1.066	0.684
sUS1	0	0.370	0.330	0.537	1.000	1.188	0.018
sIR1	0	0.349	0.349	0.116	1.000	1.000	0.998
sUK2	0	0.371	0.352	0.704	1.000	1.095	0.169
sUS2	0	0.361	0.314	0.759	1.000	1.213	0.074
sIT1	0	0.325	0.324	0.143	1.000	0.969	0.804
sFIN1	0	0.495	0.429	0.675	1.000	1.200	0.089
sNL3	0	0.378	0.357	0.222	1.000	1.098	0.113
sGER1	0	0.337	0.333	0.211	1.000	1.016	0.890
sIT2	0	0.328	0.290	1.000	1.000	1.179	0.354
sIB1	0	0.389	0.338	0.823	1.000	1.200	0.366
sSWISS1	0	0.350	0.342	0.101	1.000	0.988	0.938
sBE2	0	0.380	0.384	0.057	1.000	0.986	0.912
sSW2	0	0.407	0.402	0.892	1.000	0.976	0.864
sFIN2	0	0.478	0.402	1.000	1.000	1.282	0.209
sIR2	0	0.377	0.342	0.526	1.000	1.162	0.188
sUS3	0	0.352	0.357	0.243	1.000	1.030	0.683
sFR2	0	0.365	0.327	0.474	1.000	1.168	0.096
sUK3	0	0.359	0.356	0.632	1.000	1.014	0.800
sGER2	1	0.367	0.358	0.494	0.537	1.174	0.142
sIT3	0	0.303	0.295	0.609	1.000	1.051	0.406
sNL4	0	0.381	0.326	0.075	1.000	1.289	0.005

Supplementary Table 8: rs35714695 details

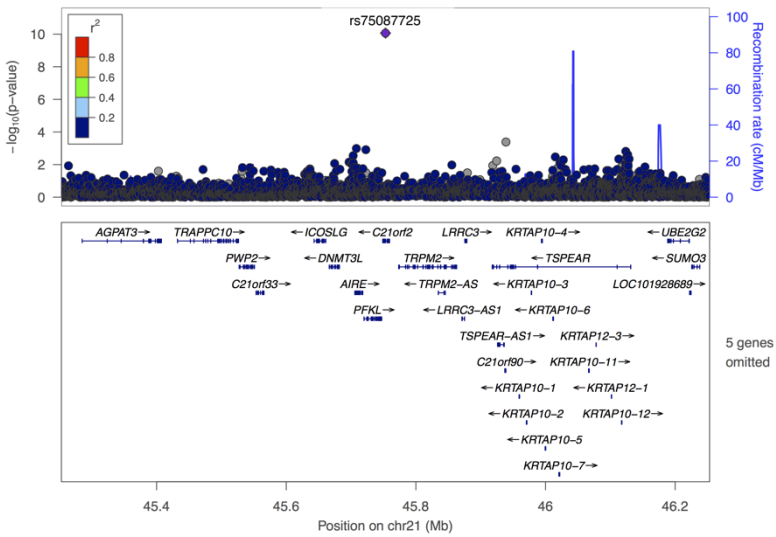


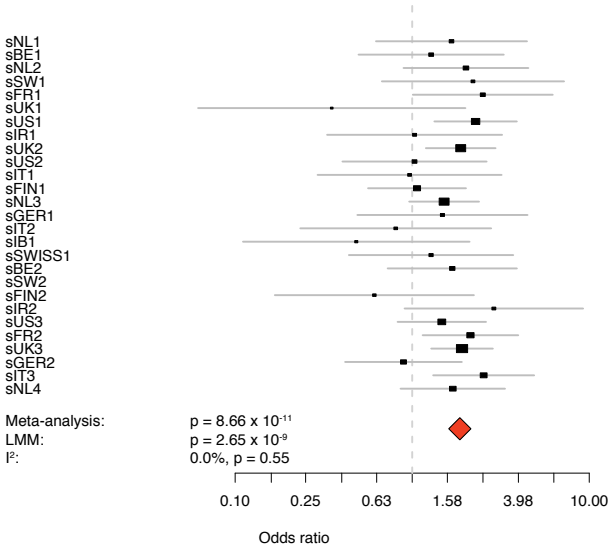


Stratum information

NAME	IMPUTED	MAF_{CASES}	$MAF_{CONTROLS}$	HWE	INFO	$OR_{STRATUM}$	$P_{STRATUM}$
sNL1	1	0.162	0.191	0.428	0.992	0.820	0.115
sBE1	1	0.160	0.210	0.612	0.993	0.706	0.020
sNL2	1	0.166	0.182	0.336	0.995	0.906	0.541
sSW1	1	0.152	0.173	0.522	0.995	0.842	0.319
sFR1	1	0.201	0.187	0.248	0.992	1.098	0.570
sUK1	1	0.146	0.200	0.609	0.987	0.677	0.065
sUS1	1	0.165	0.174	0.568	0.978	0.936	0.494
sIR1	1	0.131	0.161	0.307	0.995	0.792	0.142
sUK2	1	0.175	0.187	0.799	0.996	0.917	0.300
sUS2	1	0.158	0.165	0.423	0.989	0.947	0.718
sIT1	1	0.161	0.199	0.419	0.992	0.703	0.024
sFIN1	1	0.134	0.150	1.000	0.996	0.919	0.579
sNL3	1	0.168	0.184	0.349	0.996	0.897	0.148
sGER1	1	0.165	0.198	1.000	0.993	0.797	0.108
sIT2	1	0.171	0.143	0.058	0.982	1.248	0.357
sIB1	1	0.155	0.136	1.000	0.999	1.089	0.755
sSWISS1	1	0.197	0.181	0.820	0.998	1.175	0.379
sBE2	1	0.165	0.184	1.000	0.993	0.878	0.463
sSW2	1	0.155	0.198	0.676	0.995	0.791	0.195
sFIN2	1	0.137	0.113	0.342	1.000	1.481	0.215
sIR2	1	0.148	0.179	0.143	0.996	0.805	0.138
sUS3	1	0.162	0.183	0.297	0.995	0.917	0.363
sFR2	1	0.179	0.192	0.032	0.991	0.898	0.351
sUK3	1	0.171	0.188	1.000	0.995	0.890	0.090
sGER2	1	0.156	0.177	0.787	0.802	0.821	0.100
sIT3	1	0.162	0.193	0.555	0.987	0.810	0.004
sNL4	1	0.163	0.193	0.211	0.998	0.817	0.077

Supplementary Table 9: rs75087725 details





Stratum information rs75087725

NAME	IMPUTED	MAF_{CASES}	$MAF_{CONTROLS}$	HWE	INFO	$OR_{STRATUM}$	$P_{STRATUM}$
sNL1	1	0.014	0.009	1.000	0.882	1.667	0.302
sBE1	1	0.018	0.015	1.000	0.875	1.280	0.608
sNL2	1	0.025	0.014	1.000	0.883	2.012	0.122
sSW1	1	0.017	0.008	1.000	0.925	2.204	0.169
sFR1	1	0.028	0.013	1.000	0.834	2.504	0.059
sUK1	1	0.010	0.019	1.000	0.660	0.351	0.219
sUS1	1	0.027	0.014	1.000	0.813	2.281	0.003
sIR1	1	0.012	0.012	1.000	0.795	1.031	0.957
sUK2	1	0.025	0.014	1.000	0.875	1.879	0.009
sUS2	1	0.015	0.015	1.000	0.868	1.031	0.950
sIT1	1	0.012	0.013	1.000	0.755	0.967	0.956
sFIN1	1	0.032	0.027	0.228	0.959	1.065	0.847
sNL3	1	0.019	0.013	1.000	0.925	1.514	0.076
sGER1	1	0.014	0.010	1.000	0.827	1.482	0.475
sIT2	1	0.019	0.023	1.000	0.842	0.807	0.739
sIB1	1	0.016	0.027	1.000	0.850	0.483	0.328
sSWISS1	1	0.025	0.017	1.000	0.882	1.278	0.652
sBE2	1	0.035	0.022	1.000	0.936	1.684	0.221
sSW2	1	0.010	0.002	0.891	1.000	-	-
sFIN2	1	0.019	0.034	1.000	0.950	0.611	0.454
sIR2	1	0.017	0.007	1.000	0.796	2.892	0.068
sUS3	1	0.018	0.014	1.000	0.864	1.470	0.201
sFR2	1	0.030	0.016	0.210	0.866	2.133	0.020
sUK3	1	0.023	0.013	1.000	0.877	1.912	0.002
sGER2	1	0.016	0.013	1.000	0.758	0.893	0.770
sIT3	1	0.015	0.008	1.000	0.762	2.533	0.003
sNL4	1	0.021	0.013	1.000	0.984	1.695	0.120

* SNP failed imputation QC because of MAF < 0.01

Supplementary Table 10: functional details for rs75087725

General info		Functional prediction		ExAC allele frequency	
Chromosome	21	PolyPhen-2	0.015 (Benign)	Finnish	0.026
Basepair	45753117			European	0.011
Gene	*C21orf2*	**Conservation scores**		Latino	0.003
SNP	rs75087725	PhyloP	-0.118	African	0.002
Mutation	p.V58L	PhastCons	0.026	South Asian	0.001
				East Asian	0.000

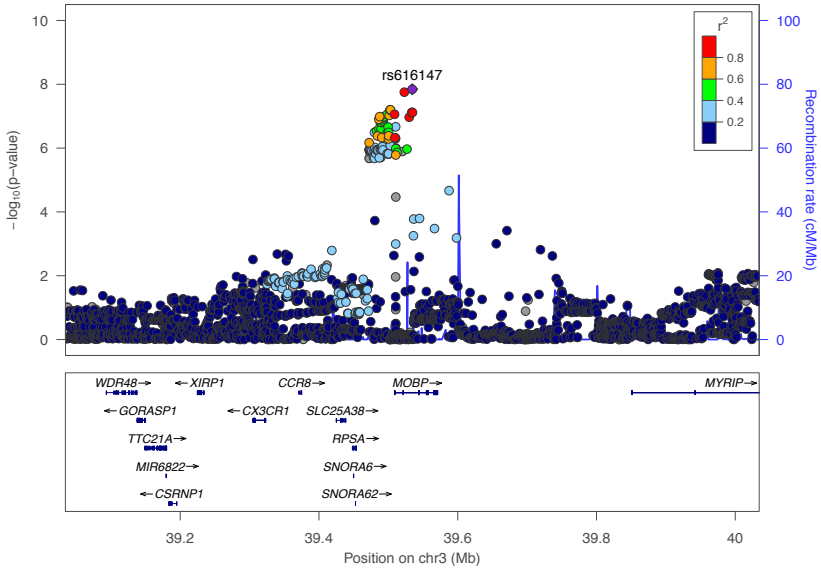
Amino-acid conservation

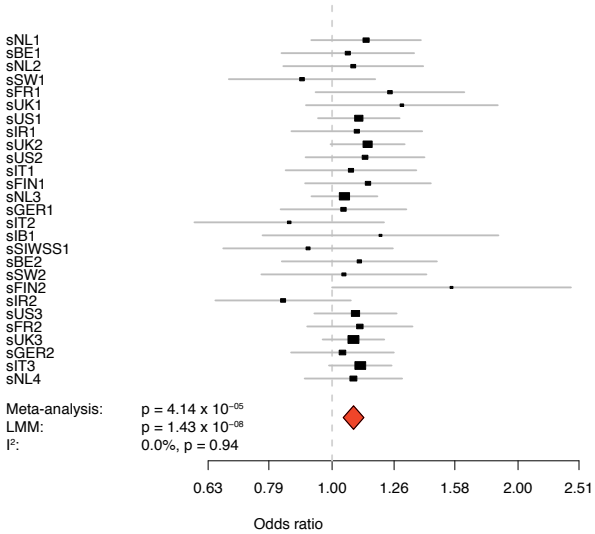
Human	C	R	S	**V**	P	E	L
Rhesus	C	R	S	**V**	P	E	L
Mouse	C	S	R	**V**	P	E	L
Dog	C	Q	S	**V**	P	E	L
Zebrafish	C	S	S	**L**	H	E	L

Supplementary Table 11: rs75087725 imputation comparison

Stratum	*1000GP imputation*			*Custom panel imputation*		
	MAF	*INFO*	*Passed QC*	*MAF*	*INFO*	*Passed QC*
sNL1	0.0077	0.77	0	0.0116	0.88	1
sBE1	0.0099	0.83	0	0.0167	0.87	1
sNL2	0.0097	0.76	0	0.0140	0.88	1
sSW1	0.0049	0.74	0	0.0126	0.92	1
sFR1	0.0096	0.77	0	0.0160	0.83	1
sUK1	0.0098	0.66	0	0.0142	0.66	1
sUS1	0.0130	0.77	1	0.0179	0.81	1
sIR1	0.0085	0.73	0	0.0120	0.79	1
sUK2	0.0114	0.81	1	0.0162	0.88	1
sUS2	0.0127	0.85	1	0.0147	0.87	1
sIT1	0.0071	0.66	0	0.0126	0.76	1
sFIN1	0.0270	0.94	1	0.0291	0.96	1
sNL3	0.0102	0.77	1	0.0149	0.93	1
sGER1	0.0100	0.76	1	0.0131	0.83	1
sIT2	0.0147	0.64	1	0.0201	0.84	1
sIB1	0.0117	0.75	1	0.0204	0.85	1
sSWISS1	0.0147	0.80	1	0.0204	0.88	1
sBE2	0.0157	0.73	1	0.0283	0.94	1
sSW2	0.0042	0.74	0	0.0049	0.89	0
sFIN2	0.0259	0.94	1	0.0254	0.95	1
sIR2	0.0058	0.80	0	0.0110	0.80	1
sUS3	0.0098	0.75	0	0.0148	0.86	1
sFR2	0.0133	0.77	1	0.0193	0.87	1
sUK3	0.0098	0.74	0	0.0161	0.88	1
sGER2	0.0093	0.60	0	0.0154	0.76	1
sIT3	0.0099	0.72	0	0.0125	0.76	1
sNL4	0.0109	0.80	1	0.0171	0.98	1
$P_{\text{(meta-analysis)}}$	3.24×10^{-5}			8.65×10^{-11}		

Supplementary Table 12: rs616147 details

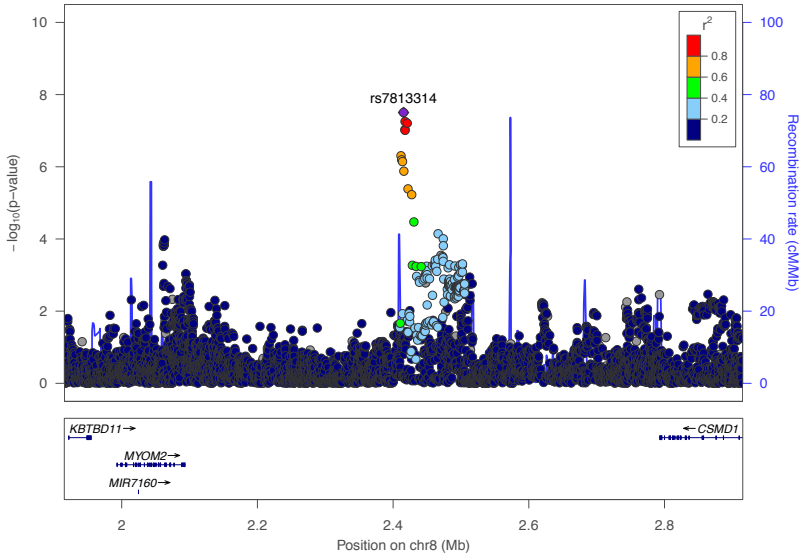


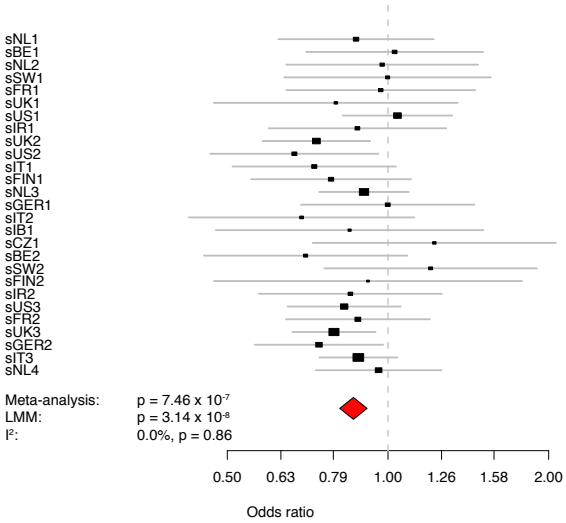


Stratum information rs616147

NAME	IMPUTED	MAF_{CASES}	$MAF_{CONTROLS}$	HWE	INFO	$OR_{STRATUM}$	$P_{STRATUM}$
sNL1	1	0.302	0.275	0.086	0.983	1.135	0.223
sBE1	1	0.320	0.308	0.597	0.975	1.061	0.643
sNL2	1	0.288	0.268	0.610	0.983	1.082	0.556
sSW1	1	0.286	0.313	1.000	0.988	0.894	0.420
sFR1	1	0.316	0.273	0.430	0.982	1.241	0.129
sUK1	1	0.300	0.252	0.668	0.975	1.297	0.152
sUS1	1	0.294	0.273	0.679	0.982	1.105	0.200
sIR1	1	0.291	0.272	0.782	1.000	1.096	0.459
sUK2	1	0.287	0.261	0.689	1.000	1.141	0.062
sUS2	1	0.316	0.289	0.160	1.000	1.131	0.278
sIT1	1	0.341	0.328	0.558	0.999	1.073	0.570
sFIN1	1	0.323	0.274	0.242	0.987	1.143	0.263
sNL3	1	0.289	0.281	0.817	1.000	1.047	0.464
sGER1	1	0.279	0.270	1.000	0.999	1.043	0.722
sIT2	1	0.342	0.376	0.178	0.998	0.852	0.376
sIB1	1	0.278	0.242	0.780	0.999	1.197	0.420
sSWISS1	1	0.293	0.310	1.000	1.000	0.914	0.579
sBE2	1	0.300	0.279	0.874	0.999	1.107	0.491
sSW2	1	0.306	0.296	0.754	1.000	1.045	0.779
sFIN2	1	0.337	0.278	0.456	1.000	1.560	0.047
sIR2	1	0.258	0.291	0.298	0.999	0.833	0.154
sUS3	1	0.296	0.274	0.465	0.999	1.091	0.266
sFR2	1	0.307	0.291	0.939	0.996	1.108	0.303
sUK3	1	0.282	0.266	0.474	0.999	1.083	0.175
sGER2	1	0.291	0.259	0.355	0.876	1.040	0.692
sIT3	1	0.346	0.320	0.944	0.998	1.111	0.077
sNL4	1	0.292	0.277	0.665	0.999	1.082	0.391

Supplementary Table 13: rs7813314 details

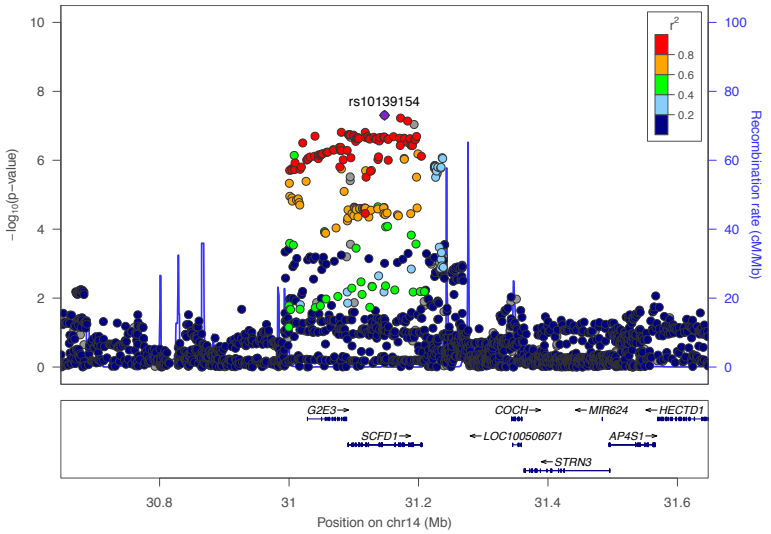


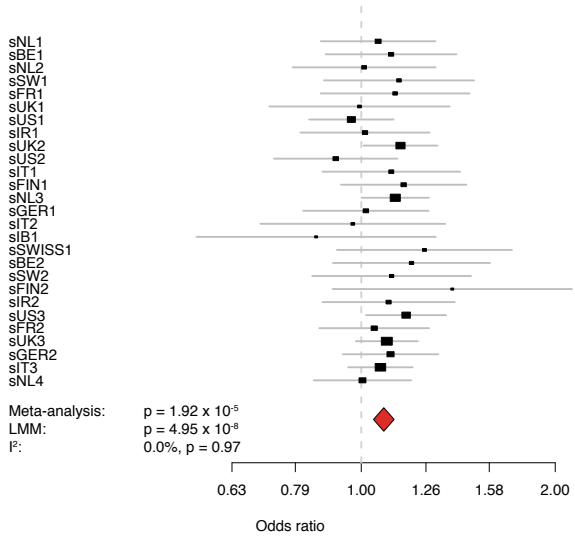


Stratum information rs7813314

NAME	IMPUTED	MAF_{CASES}	$MAF_{CONTROLS}$	HWE	INFO	$OR_{STRATUM}$	$P_{STRATUM}$
sNL1	1	0.080	0.092	0.065	0.995	0.871	0.420
sBE1	1	0.100	0.097	0.752	0.963	1.029	0.883
sNL2	1	0.096	0.100	0.425	0.941	0.976	0.905
sSW1	1	0.090	0.095	0.709	0.982	0.998	0.995
sFR1	1	0.100	0.103	0.830	0.990	0.969	0.880
sUK1	1	0.086	0.105	0.676	0.989	0.799	0.401
sUS1	1	0.101	0.098	0.276	0.911	1.042	0.735
sIR1	1	0.087	0.098	1.000	0.988	0.876	0.499
sUK2	1	0.075	0.101	0.915	0.994	0.734	0.007
sUS2	1	0.081	0.117	0.386	0.989	0.668	0.025
sIT1	1	0.109	0.138	0.588	0.994	0.728	0.078
sFIN1	1	0.098	0.116	0.802	0.985	0.782	0.162
sNL3	1	0.087	0.096	0.136	0.998	0.902	0.292
sGER1	1	0.083	0.083	1.000	0.996	0.999	0.995
sIT2	1	0.116	0.157	0.682	0.996	0.689	0.139
sIB1	1	0.107	0.122	0.630	0.997	0.847	0.574
sSWISS1	1	0.098	0.084	0.375	0.996	1.221	0.454
sBE2	1	0.095	0.128	0.777	0.998	0.701	0.109
sSW2	1	0.111	0.089	0.678	0.995	1.202	0.431
sFIN2	1	0.078	0.093	0.540	0.999	0.917	0.799
sIR2	1	0.081	0.094	0.156	0.998	0.850	0.417
sUS3	1	0.089	0.099	0.032	0.994	0.828	0.125
sFR2	1	0.089	0.104	0.866	0.992	0.878	0.408
sUK3	1	0.085	0.105	0.335	0.988	0.792	0.010
sGER2	1	0.082	0.095	1.000	0.981	0.742	0.036
sIT3	1	0.110	0.124	0.399	0.990	0.880	0.138
sNL4	1	0.096	0.100	0.012	1.000	0.961	0.770

Supplementary Table 14: rs10139154 details





Stratum information rs10139154

NAME	IMPUTED	MAF_{CASES}	$MAF_{CONTROLS}$	HWE	INFO	$OR_{STRATUM}$	$P_{STRATUM}$
sNL1	1	0.331	0.318	0.573	1.000	1.062	0.565
sBE1	1	0.337	0.313	0.512	0.993	1.112	0.372
sNL2	1	0.314	0.310	0.593	0.994	1.010	0.938
sSW1	1	0.347	0.312	0.668	1.000	1.144	0.324
sFR1	1	0.342	0.316	0.588	0.993	1.129	0.374
sUK1	1	0.304	0.305	0.850	1.000	0.994	0.969
sUS1	1	0.310	0.318	0.032	0.993	0.965	0.645
sIR1	1	0.312	0.309	0.093	0.999	1.013	0.910
sUK2	1	0.328	0.296	0.517	1.000	1.150	0.039
sUS2	1	0.336	0.357	0.847	0.998	0.913	0.416
sIT1	1	0.361	0.348	0.776	0.999	1.113	0.392
sFIN1	1	0.343	0.308	0.469	0.996	1.163	0.185
sNL3	1	0.339	0.314	0.233	1.000	1.129	0.048
sGER1	1	0.323	0.320	0.086	1.000	1.016	0.886
sIT2	1	0.376	0.383	0.829	0.999	0.970	0.858
sIB1	1	0.294	0.328	0.359	0.999	0.851	0.459
sSWISS1	1	0.350	0.296	0.630	0.999	1.253	0.156
sBE2	1	0.312	0.273	0.627	0.999	1.197	0.210
sSW2	1	0.332	0.309	0.760	0.999	1.115	0.454
sFIN2	1	0.381	0.294	0.626	0.999	1.384	0.133
sIR2	1	0.301	0.281	0.813	1.000	1.103	0.420
sUS3	1	0.345	0.306	0.155	1.000	1.174	0.029
sFR2	1	0.333	0.320	0.277	0.980	1.047	0.645
sUK3	1	0.317	0.298	0.534	1.000	1.096	0.107
sGER2	1	0.341	0.316	0.275	0.964	1.110	0.229
sIT3	1	0.365	0.348	0.382	0.999	1.071	0.243
sNL4	1	0.337	0.334	0.846	1.000	1.004	0.962

Supplementary Table 15: previously associated loci

		Discovery GWAS			*Current GWAS*[*]		
Locus	*SNP*	MAF_{cases}	$MAF_{controls}$	*p-value*	*OR*	*p-value*	*Power*
FGGY[56]	rs6700125	0.32	0.41	1.8×10^{-5}	1.06	0.07	1.00
ITPR2[3]	rs2306677	0.11	0.07	3.3×10^{-6}	1.03	0.54	1.00
SUN3[10]	rs2708909	0.45	0.50	7.0×10^{-7}	0.97	0.28	1.00
C7orf57[10]	rs2708851	0.45	0.50	1.2×10^{-6}	0.97	0.24	1.00
DPP6[4]	rs10260404	0.42	0.35	5.4×10^{-8}	0.97	0.20	1.00
CAMK1G[57]	rs6703183	0.41	0.34	2.9×10^{-8}	0.96	0.13	1.00
SUSD2[57]	rs8141797	0.15	0.10	2.4×10^{-9}	1.02	0.70	1.00
18q11.2[18]	rs1788776	0.41	0.38	8.4×10^{-6}	1.02	0.43	1.00
CYP27A1[58]	rs4674345	-	-	1.8×10^{-4}	0.99	0.57	-
CENPV[59]	rs7477	~ 0.5	~ 0.5	2.9×10^{-7}	1.01	0.70	-
8q24.13[59]	rs12546767	~ 0.1	~ 0.1	2.7×10^{-6}	0.93	0.09	-

[*] Newly genotyped individuals only, excluding possible sample overlap with all discovery GWAS.

Supplementary Table 16: description of eQTL datasets

Study	*Tissue*	*N*	*Neuropathology*	*N*	*Nr of tissue samples in analysis*
Myers et al.[23]	Pooled cortex	193	Normal	193	193
Heinzen et al.[24]	Frontal cortex	93	Normal	93	93
Webster et al.[25]	Pooled cortex	364	Normal	188	188
			Alzheimer disease	176	176
Gibbs et al.[26]	Cerebellum	143	Normal	572	143
	Caudal pons	142			142
	Frontal cortex	143			143
	Temporal cortex	144			144
Coluantoni et al.[27]	Prefrontal cortex	269	Normal	269	269
Liu et al.[28]	Prefrontal cortex	127	Normal	40	
			Bipolar	39	
			Schizophrenia	37	
			Depression	11	
					127
Kim et al.[31]	Frontal cortex	NA	Normal	15	
	Temporal cortex	NA	Schizophrenia	15	
	Thalamus	NA	Bipolar	15	
	Cerebellum	NA	Major Depression	15	
					60
	Frontal cortex	NA	Normal	35	
	Hippocampus	NA	Schizophrenia	35	
			Bipolar	35	
					105
Zou et al.[32]	Cerebellum	197	Alzheimer disease	177	
	Temporal cortex	202	Other brain pathologies	197	
					~400
Ramasamy et al.[33]	Cerebellar cortex	134			
	Frontal cortex	134			
	Hippocampus	134			
	Medulla	134			
	Putamen	134			
	Substantia nigra	134			
	Thalamus	134			
	Temporal cortex	134			
	Intralobular white matter	134			
					1,231
Deelen et al.[34]	NA	42	NA	42	42
Kim et al.[35]	Frontal cortex	56			
	Frontal cortex	124			
	Frontal cortex	24			
	Pooled cortex	189			
	Frontal cortex	31			
					373
GTex[36]	Anterior cingulate	72			72
	Caudate basal ganglia	100			100
	Cerebellar hemisphere	89			89
	Cerebellum	103			103
	Brain Cortex	96			96
	Brain Frontal Cortex	92			92
	Hippocampus	81			81
	Hypothalamus	81			81
	Nucleus accumbens	93			93
	Putamen basal ganglia	82			82

Supplementary Table 17: details for top brain *cis*-eQTLs

GWAS top SNP	*Brain cis-eQTL SNP*	R^2	*GWAS* $p_{(LMM)}$	*Brain cis-eQTL gene*	*Brain tissue*	*eQTL p**	*Source*
rs616147	rs1768208	0.99	1.8×10^{-8}	*RPSA*	cerebellum	7.7×10^{-4}	GTEx[36]
	rs2965067	0.65	6.3×10^{-8}	*RPSA*	cerebellar hemisphere	9.8×10^{-4}	GTEx[36]
	rs1472508	0.28	0.026	*RPSA*	cortex	0.023	GTEx[36]
	rs1472508	0.28	0.026	*RPSA*	caudate basal ganglia	0.019	GTEx[36]
	rs12638676	0.31	8.3×10^{-7}	*SLC25A38*	meta-analysis	5.9×10^{-3}	Kim et al.[35]
	rs2039845	0.15	0.027	*SLC25A38*	mixed brain sample	1.0×10^{-3}	Webster et. al[27]
	rs1707953	0.10	0.74	*MOBP*	cerebellum / temporal cortex	7.1×10^{-18}	Zou et al.[32]
rs3849943	rs10812605	0.56	5.8×10^{-12}	*C9orf72*	cerebellum	7.7×10^{-4}	GTEx[36]
	rs2492816	0.45	2.0×10^{-11}	*C9orf72*	cerebellar hemisphere	9.8×10^{-4}	GTEx[36]
	rs2492816	0.45	2.0×10^{-11}	*C9orf72*	nucleus accumbens	0.041	GTEx[36]
	rs4879541	0.33	8.5×10^{-10}	*C9orf72*	frontal cortex	0.013	GTEx[36]
	rs2244606	0.24	2.5×10^{-10}	*C9orf72*	meta-analysis	5.3×10^{-7}	Kim et al.[35]
rs10139154	rs10139154	1.00	5.0×10^{-8}	*SCFD1*	cerebellum	7.7×10^{-4}	GTEx[36]
	rs10139154	1.00	5.0×10^{-8}	*SCFD1*	cerebellar hemisphere	9.8×10^{-4}	GTEx[36]
rs35714695	rs35714695	1.00	9.0×10^{-11}	*POLDIP2*	cortex	2.3×10^{-3}	GTEx[36]
	rs739438	0.54	4.1×10^{-6}	*POLDIP2*	substantia nigra	0.038	BRAINEAC[33]
	rs9913833	0.22	0.012	*POLDIP2*	nucleus accumbens	0.012	GTEx[36]
	rs4795434	0.22	0.013	*POLDIP2*	putamen basal ganglia	0.048	GTEx[36]
	rs7212510	0.21	0.016	*TMEM199*	anterior cingulate cortex	0.011	GTEx[36]
	rs12947270	0.33	0.014	*SARM1*	cortex	0.043	GTEx[36]
rs12608932	rs12608932	1.00	2.7×10^{-10}	*KCNN1*	frontal cortex	7.0×10^{-3}	BRAINEAC[33]

For each *cis*-eQTL in each brain region, details are provided only for the SNP in strongest LD with the top SNP in the GWAS. A full description of all brain and non-brain *cis*-eQTLs is provided in the supplementary Excel file.

* Corrected *p*-values as reported in original article; for BRAINEAC Benjamini-Hockhberg correction was performed for all eQTL SNPs overlapping with suggestive DEPICT GWAS loci (GWAS LMM $p < 10 \times 10^{-4}$; $r^2 > 0.5$); for GTEx, q-value for eGene is reported.

Supplementary Table 18: software packages

Software	Version	Web source	Reference
ANNOVAR	2014-07-14	http://annovar.openbioinformatics.org	42
DEPICT	1.139	http://www.broadinstitute.org/mpg/depict	43
EIGENSTRAT	3.0	https://github.com/DReichLab/EIG	44
GCTA	1.24	http://www.complextraitgenomics.com/software/gcta/	45
IMPUTE2	2.3.1	https://mathgen.stats.ox.ac.uk/impute/impute_v2.html	46
MASS	7.0	http://dlin.web.unc.edu/software/mass/	47
METAL	2011-03-25	http://genome.sph.umich.edu/wiki/METAL_Program	48
LDSC	1.0.0	https://github.com/bulik/ldsc/	49
PCGC	2014-12	https://sites.google.com/site/davidgolanshomepage/software/pcgc	50
PLINK 1.9	beta3p	https://www.cog-genomics.org/plink2/	51
R	3.0.2	http://cran.r-project.org	-
ScoreSeq	7.0	http://dlin.web.unc.edu/software/score-seq/	52
SHAPEIT2	2.727	https://mathgen.stats.ox.ac.uk/genetics_software/shapeit/shapeit.html	53
SNPTEST	2.5	https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html	54
VCFtools	Dec 2015	https://vcftools.github.io/index.html	55

Supplementary Table 19: IMPUTE2 imputation concordance scores

Stratum	Type 0 concordance			Type 1 concordance		
	cases	*controls*	*p**	*cases*	*controls*	*p**
1	0.982	0.982	0.956	0.975	0.975	0.944
2	0.977	0.977	0.997	0.967	0.967	0.996
3	0.983	0.982	0.915	0.977	0.975	0.895
4	0.981	0.982	0.965	0.973	0.974	0.960
5	0.968	0.968	0.984	0.955	0.954	0.983
6	0.971	0.971	0.996	0.960	0.960	0.999
7	0.971	0.971	0.987	0.960	0.960	0.986
8	0.985	0.985	0.995	0.979	0.979	0.986
9	0.987	0.987	0.981	0.981	0.981	0.980
10	0.985	0.985	0.972	0.978	0.978	0.961
11	0.977	0.974	0.901	0.967	0.962	0.878
12	0.985	0.984	0.936	0.979	0.978	0.919
13	0.999	0.992	0.275	0.998	0.989	0.187
14	0.987	0.987	0.990	0.981	0.981	0.995
15	0.982	0.981	0.971	0.974	0.972	0.964
16	0.981	0.981	0.986	0.973	0.972	0.984
17	0.986	0.986	0.995	0.979	0.979	0.994
18	0.988	0.988	0.999	0.983	0.983	0.998
19	0.990	0.990	0.991	0.986	0.986	0.988
20	0.989	0.989	0.975	0.985	0.984	0.961
21	0.988	0.988	0.993	0.983	0.983	0.992
22	0.986	0.987	0.761	0.980	0.981	0.698
23	0.969	0.968	0.968	0.955	0.955	0.961
24	0.984	0.984	0.783	0.978	0.977	0.360
25	0.922	0.930	0.548	0.886	0.899	0.435
26	0.978	0.979	0.912	0.969	0.970	0.887
27	0.999	0.999	0.989	0.999	0.999	0.966
Total**	0.983	0.984	0.839	0.999	0.999	0.648

* logistic regression of concordance scores between cases and controls.

** logistic regression of concordance scores with stratum as covariate.

Supplementary Table 20: oligonucleotide primers for Sanger sequencing

SNP	Forward primers	Reverse primer
rs10139154	*5′-GCTAGAAGATTGTGTCATGTAGAGG-3′*	*5′-ACAACCAGCAAGTGGGAAAG-3′*
rs7813314	5′-TGCTGATACGGTGACCAGAG-3′	5′-TTGCTATTGAATGCTCCTTTGC-3′

Consortium members

FALS Sequencing Consortium members

Bradley N. Smith, Nicola Ticozzi, Claudia Fallini, Athina Soragia Gkazi, Simon Topp, Jason Kost, Emma L. Scotter, Kevin P. Kenna, Pamela Keagle, Jack W. Miller, Cinzia Tiloca, Caroline Vance, Eric W. Danielson, Claire Troakes, Claudia Colombrita, Safa Al-Sarraj, Elizabeth A. Lewis, Andrew King, Daniela Calini, Viviana Pensato, Barbara Castellotti, Jacqueline de Belleroche, Frank Baas, Anneloor L.M.A. ten Asbroek, Peter C. Sapp, Diane McKenna-Yasek, Russell L. McLaughlin, Meraida Polak, Seneshaw Asress, Jesús Esteban-Pérez, José Luis Muñoz-Blanco, Sandra D'Alfonso, Letizia Mazzini, Giacomo P. Comi, Roberto Del Bo, Mauro Ceroni, Stella Gagliardi, Giorgia Querin, Cinzia Bertolin, Wouter van Rheenen, Frank P. Diekstra, Rosa Rademakers, Marka van Blitterswijk, Kevin B. Boylan, Giuseppe Lauria, Stefano Duga, Stefania Corti, Cristina Cereda, Lucia Corrado, Gianni Sorarù, Kelly L. Williams, Garth A. Nicholson, Ian P. Blair, Claire Leblond-Manry, Guy A. Rouleau, Orla Hardiman, Karen E. Morrison, Jan H. Veldink, Leonard H. van den Berg, Ammar Al-Chalabi, Hardev Pall, Pamela J. Shaw, Martin R. Turner, Kevin Talbot, Franco Taroni, Alberto García-Redondo, Zheyang Wu, Jonathan D. Glass, Cinzia Gellera, Antonia Ratti, Robert H. Brown, Jr., Vincenzo Silani, Christopher E. Shaw and John E. Landers

Italian Consortium for the Genetics of ALS (SLAGEN) Consortium members

Daniela Calini, Isabella Fogh, Antonia Ratti, Vincenzo Silani, Nicola Ticozzi, Cinzia Tiloca, Barbara Castellotti, Cinzia Gellera, Viviana Pensato, Franco Taroni, Cristina Cereda, Mauro Ceroni, Stella Gagliardi, Giacomo Comi, Stefania Corti, Roberto Del Bo, Lucia Corrado, Sandra D'Alfonso, Letizia Mazzini, Elena Pegoraro, Giorgia Querin and Gianni Sorarù

Registro Lombardo Sclerosi Laterale Amyotrofica (SLALOM) group members

Francesca Gerardi, Fabrizio Rinaldi, Maria Sofia Cotelli, Luca Chiveri, Maria Cristina Guaita and Patrizia Perrone

Piemonte and Valle d'Aosta Registry for Amyotrophic Lateral Sclerosis (PARALS) group members

Stefania Cammarosano, Antonio Canosa, Dario Cocito, Leonardo Lopiano, Luca Durelli, Bruno Ferrero, Antonio Bertolotto, Alessandro Mauro, Luca Pradotto, Roberto Cantello, Enrica Bersano, Dario Giobbe, Maurizio Gionco, Daniela Leotta, Lucia Appendino, Roberto Cavallo, Enrico Odddenino, Claudio Geda, Fabio Poglio, Paola Santimaria, Umberto

Massazza, Antonio Villani, Roberto Conti, Fabrizio Pisano, Mario Palermo, Franco Vergnano, Paolo Provera, Maria Teresa Penza, Marco Aguggia, Nicoletta Di Vito, Piero Meineri, Ilaria Pastore, Paolo Ghiglione, Danilo Seliak, Nicola Launaro, Giovanni Astegiano and Bottacchi Edo

Sclerosi Laterale Amyotrofica-Puglia (SLAP) registry members
Isabella Laura Simone, Stefano Zoccolella, Michele Zarrelli and Franco Apollo

Neuroprotection and Natural History in Parkinson Plus Syndromes (NNIPPS) study group members
William Camu, Jean Sebastien Hulot, Francois Viallet, Philippe Couratier, David Maltete, Christine Tranchant, Marie Vidailhet.

Nature Genetics: doi:10.1038/ng.3622

Acknowledgments

This study was supported by the ALS Foundation Netherlands, the Belgian ALS Liga and Agency for Innovation by Science and Technology (IWT), and the MND association (UK) (Project MinE, www.projectmine.com).

Research leading to these results has received funding from the European Community's Health Seventh Framework Programme (FP7/2007-2013).

This study was supported by ZonMW under the frame of E-Rare-2, the ERA Net for Research on Rare Diseases (PYRAMID).

This is an EU Joint Programme–Neurodegenerative Disease Research (JPND) project (STRENGTH, SOPHIA). The project is supported through the following funding organizations under the aegis of JPND: UK, Medical Research Council and Economic and Social Research Council; Ireland, Health Research Board; Netherlands, ZonMw; Italy, Ministry of Health and Ministry of Education, University and Research; France, L'Agence nationale pour la recherche.

Samples used in this research were in part obtained from the UK National DNA Bank for MND Research, funded by the MND Association and the Wellcome Trust.

We acknowledge sample management undertaken by Biobanking Solutions funded by the Medical Research Council at the Centre for Integrated Genomic Medical Research, University of Manchester.

The study was supported by the Slovenian Research Agency (P3-0338, P4-0127, J3-6789, J3-5502, 1000-15-0510).

PopGen 2.0 network is supported by a grant from the German Ministry for Education and Research (01EY1103).

This work was supported by a grant from IWT (Agency for Innovation by Science and Technology), the University of Leuven [GOA 11/014] and the Interuniversity Attraction Poles program P7/16 of the Belgian Federal Science Policy Office.

This study was supported by Telethon Genetic BioBank (GTB12001D) and the Eurobiobank network.

This work was supported by grants from NIH (AG032953 and AG017586) and the ALS Association (USA).

This work was supported by a grant-in-aid from the MND Research Institute of Australia and a Leadership grant from MND Australia.

This project is supported by the Netherlands Organisation for Health Research and Development (Vici scheme to L.H. van den Berg and veni scheme to M.A. van Es).

M.A. van Es is supported by the Thierry Latran Foundation, the Dutch ALS foundation and the Rudolf Magnus Brain Center Talent Fellowship.
Christopher E. Shaw and Ammar Al-Chalabi receive salary support from the National Institute for Health Research (NIHR) Dementia Biomedical Research Unit and Biomedical Research Centre in Mental Health at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.
Russell. L. McLaughlin is supported by the Thierry Latran Foundation (ALSIBD) and the ALS Association (2284).
Isabella Fogh was supported by the MND association (UK) (Grant Ref: Al-Chalabi/Feb14/905-793).
An Goris is supported by the Research Fund KU Leuven.
Nicola Ticozzi, Isabella Fogh, Cinzia Gellera, Cinzia Tiloca, Antonia Ratti and Vincenzo Silani are supported by Agenzia Italiana per la Ricerca sulla SLA-AriSLA (grant NOVALS 2012 cofinanced with the contribution of 5 x 1000, Healthcare Research support of the Ministry of Health), the Italian Ministry of Health (GR-2011-02347820, IRiS-ALS) and Associazione Amici "Centro Dino Ferrari".
Philip Van Damme holds a senior clinical investigatorship from FWO-Vlaanderen and is supported by the ALS liga België.
The University of Queensland group is supported by the Australian National Health and Medical Research Council (NHMRC grants 1083187, 1078901, 1078037, 1048853 and 1050218), with Jian Yang supported by a Charles and Sylvia Viertel Senior Medical Research Fellowship.
Stéphanie Millecamps is financed by the Association pour la Recherche sur la Sclérose latérale amyotrophique et autres maladies du motoneurone (ARSla, France) and the Association Française contre les Myopathies (AFM, France).
Andrea Calvo was supported by Fondazione Vialli e Mauro
Adriano Chiò was supported by Ministero della Salute (Ricerca Sanitaria Finalizzata, 2010, grant RF-2010-2309849).
Pietro Fratta is funded by an MRC/MNDA LEW Clinical Fellowship and NIHR UCLH BRC
Wim Robberecht is supported through the E. von Behring Chair for Neuromuscular and Neurodegenerative Disorders, the Laevers Fund for ALS Research, the ALS Liga België, the fund 'Een Hart voor ALS' and the fund 'Opening the Future'.
This work was supported by the UK Medical Research Council (MRC) through the MRC Sudden Death Brain Bank, a Project Grant (G0901254 to John Hardy).

This study represents independent research part funded by the National Institute for Health Research (NIHR) Biomedical Research Centre and Dementia Unit at South London and Maudsley NHS Foundation Trust and IoPPN Genomics & Biomarker Core Facility King's College London The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.
Genotyping the Turkish individuals in the replication phase was performed by Doruk Savas, Cemre Coskun and Irmak Sahbaz.
The French control work was supported by the National Foundation for Alzheimer's disease and related disorders (Fondation Plan Alzheimer), Inserm, the Institut Pasteur de Lille, the Centre National de Génotypage and the LABEX (excellence laboratory, program investment for the future) DISTALZ (Development of Innovative Strategies for a Transdisciplinary approach to Alzheimer's disease).
Michael Sendtner and Carsten Drepper were supported by the Deutsche Forschungsgemeinschaft, SE697/4-1, BMBF EnergI and the Deutsche Gesellschaft für Muskelkranke, Project He 2/2.
We would like to thanks people with ALS/MND, their families and control individuals for their participation in this project.

References

1. Brooks, B. R. El Escorial World Federation of Neurology criteria for the diagnosis of amyotrophic lateral sclerosis. Subcommittee on Motor Neuron Diseases/Amyotrophic Lateral Sclerosis of the World Federation of Neurology Research Group on Neuromuscular Diseases and the El Escorial 'Clinical limits of amyotrophic lateral sclerosis' workshop contributors. *J. Neurol. Sci.* **124 Suppl,** 96–107 (1994).
2. Huisman, M. H. B. *et al.* Population based epidemiology of amyotrophic lateral sclerosis using capture-recapture methodology. *J. Neurol. Neurosurg. Psychiatry* **82,** 1165–1170 (2011).
3. van Es, M. A. *et al.* ITPR2 as a susceptibility gene in sporadic amyotrophic lateral sclerosis: a genome-wide association study. *Lancet Neurol.* **6,** 869–877 (2007).
4. van Es, M. A. *et al.* Genetic variation in DPP6 is associated with susceptibility to amyotrophic lateral sclerosis. *Nat. Genet.* **40,** 29–31 (2008).
5. Hofman, A. *et al.* The Rotterdam Study: objectives and design update. *Eur. J. Epidemiol.* **22,** 819–829 (2007).
6. Landers, J. E. *et al.* Reduced expression of the Kinesin-Associated Protein 3 (KIFAP3) gene increases survival in sporadic amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci.* **106,** 9004–9009 (2009).
7. Cronin, S. *et al.* A genome-wide association study of sporadic ALS in a homogenous Irish population. *Hum. Mol. Genet.* **17,** 768–774 (2008).
8. van Es, M. A. *et al.* Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. *Nat. Genet.* **41,** 1083–1087 (2009).
9. Shatunov, A. *et al.* Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study. *Lancet Neurol.* **9,** 986–994 (2010).
10. Chiò, A. *et al.* A two-stage genome-wide association study of sporadic amyotrophic lateral sclerosis. *Hum. Mol. Genet.* **18,** 1524–1532 (2009).
11. Brooks, B. R., Miller, R. G., Swash, M., Munsat, T. L. & World Federation of Neurology Research Group on Motor Neuron Diseases. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler. Other Motor Neuron Disord.* **1,** 293–299 (2000).
12. Schymick, J. C. *et al.* Genome-wide genotyping in amyotrophic lateral sclerosis and neurologically normal controls: first stage analysis and public release of data. *Lancet*

Neurol. **6,** 322–328 (2007).

13. Traynor, B. J. *et al*. Kinesin-associated protein 3 (KIFAP3) has no effect on survival in a population-based cohort of ALS patients. *Proc. Natl. Acad. Sci*. **107,** 12335–12338 (2010).
14. Pankratz, N. *et al*. Genomewide association study for susceptibility genes contributing to familial Parkinson disease. *Hum. Genet*. **124,** 593–605 (2009).
15. Hamza, T. H. *et al*. Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson's disease. *Nat. Genet*. **42,** 781–785 (2010).
16. Laaksovirta, H. *et al*. Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study. *Lancet Neurol*. **9,** 978–985 (2010).
17. Polvikoski, T. *et al*. Prevalence of Alzheimer's disease in very elderly people: a prospective neuropathological study. *Neurology* **56,** 1690–1696 (2001).
18. Fogh, I. *et al*. A genome-wide association meta-analysis identifies a novel locus at 17q11.2 associated with sporadic amyotrophic lateral sclerosis. *Hum. Mol. Genet*. **23,** 2220–2231 (2014).
19. 3C Study Group. Vascular factors and risk of dementia: design of the Three-City Study and baseline characteristics of the study population. *Neuroepidemiology* **22,** 316–325 (2003).
20. Krawczak, M. *et al*. PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. *Community Genet*. **9,** 55–61 (2006).
21. Goris, A. *et al*. No evidence for shared genetic basis of common variants in multiple sclerosis and amyotrophic lateral sclerosis. *Hum. Mol. Genet*. **23,** 1916–1922 (2014).
22. Andersen, P. M. *et al*. EFNS guidelines on the clinical management of amyotrophic lateral sclerosis (MALS)--revised report of an EFNS task force. *Eur. J. Neurol*. **19,** 360–375 (2012).
23. McCluskey, L. *et al*. ALS-Plus syndrome: non-pyramidal features in a large ALS cohort. *J. Neurol. Sci*. **345,** 118–124 (2014).
24. Arnold, M., Raffler, J., Pfeufer, A., Suhre, K. & Kastenmüller, G. SNiPA: an interactive, genetic variant-centered annotation browser. *Bioinformatics* **31**, 1334–1336 (2015).
25. Myers, A. J. *et al*. A survey of genetic human cortical gene expression. *Nat. Genet*. **39,** 1494–1499 (2007).
26. Heinzen, E. L. *et al*. Tissue-specific genetic control of splicing: implications for the study of complex traits. *PLoS Biol*. **6,** e1 (2008)

27. Webster, J. A. *et al*. Genetic control of human brain transcript expression in Alzheimer disease. *Am. J. Hum. Genet*. **84,** 445–458 (2009).
28. Gibbs, J. R. *et al*. Abundant quantitative trait loci exist for DNA methylation and gene expression in human brain. *PLoS Genet* **6,** e1000952 (2010).
29. Colantuoni, C. *et al*. Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature* **478,** 519–523 (2011).
30. Liu, C. Brain expression quantitative trait locus mapping informs genetic studies of psychiatric diseases. *Neurosc. Bull*. **27,** 123–133 (2011).
31. Kim, S., Cho, H., Lee, D. & Webster, M. J. Association between SNPs and gene expression in multiple regions of the human brain. *Transl. Psychiatry* **2,** e113 (2012).
32. Zou, F. *et al*. Brain expression genome-wide association study (eGWAS) identifies human disease-associated variants. *PLoS Genet*. **8,** e1002707 (2012).
33. Ramasamy, A. *et al*. Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat. Neurosc*. **17,** 1418–1428 (2014).
34. Deelen, P. *et al*. Calling genotypes from public RNA-sequencing data enables identification of genetic variants that affect gene-expression levels. *Genome Med*. **7,** 30 (2015).
35. Kim, Y. *et al*. A meta-analysis of gene expression quantitative trait loci in brain. *Transl. Psychiatry* **4,** e459 (2014).
36. GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348,** 648–660 (2015).
37. Westra, H.-J. *et al*. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat. Genet*. **45,** 1238–1243 (2013).
38. Grundberg, E. *et al*. Mapping cis- and trans-regulatory effects across multiple tissues in twins. *Nat. Genet*. **44,** 1084–1089 (2012).
39. Fairfax, B. P. *et al*. Genetics of gene expression in primary immune cells identifies cell type-specific master regulators and roles of HLA alleles. *Nat. Genet*. **44,** 502–510 (2012).
40. Zeller, T. *et al*. Genetics and beyond--the transcriptome of human monocytes and disease susceptibility. *PLoS ONE* **5,** e10693 (2010)
41. Roadmap Epigenomics Consortium *et al*. Integrative analysis of 111 reference human epigenomes. Nature **518**, 317–330 (2015).
42. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. **38,** e164 (2010).
43. Pers, T. H. *et al*. Biological interpretation of genome-wide association studies using

predicted gene functions. *Nat. Commun.* **6,** 5890–20 (2015).

44. Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38,** 904–909 (2006).
45. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88,** 76–82 (2011).
46. Howie, B., Marchini, J. & Stephens, M. Genotype imputation with thousands of genomes. *G3* **1,** 457-70 (2011).
47. Tang, Z.-Z. & Lin, D.-Y. MASS: meta-analysis of score statistics for sequencing studies. *Bioinformatics* **29,** 1803–1805 (2013).
48. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26,** 2190–2191 (2010).
49. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47,** 291–295 (2015).
50. Golan, D., Lander, E. S. & Rosset, S. Measuring missing heritability: inferring the contribution of common variants. *Proc. Natl. Acad. Sci.* **111,** E5272–81 (2014).
51. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81,** 559–575 (2007).
52. Lin, D.-Y. & Tang, Z.-Z. A general framework for detecting disease associations with rare variants in sequencing studies. *Am. J. Hum. Genet.* **89,** 354–367 (2011).
53. Delaneau, O., Marchini, J. & 1000 Genomes Project Consortium. Integrating sequence and array data to create an improved 1000 Genomes Project haplotype reference panel. *Nat. Commun.* **5,** 3934 (2014).
54. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* **39,** 906–913 (2007).
55. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27,** 2156–2158 (2011).
56. Dunckley, T. *et al.* Whole-genome analysis of sporadic amyotrophic lateral sclerosis. *N. Engl. J. Med.* **357,** 775–788 (2007).
57. Deng, M. *et al.* Genome-wide association analyses in Han Chinese identify two new susceptibility loci for amyotrophic lateral sclerosis. *Nat. Genet.* **45,** 697–700 (2013).
58. Diekstra, F. P. *et al.* Mapping of Gene Expression Reveals CYP27A1 as a Susceptibility Gene for Sporadic ALS. *PLoS ONE* **7,** e35333 (2012).
59. Diekstra, F. P. *et al.* C9orf72 and UNC13A are shared risk loci for amyotrophic lateral sclerosis and frontotemporal dementia: A genome-wide meta-analysis. *Ann. Neurol.*

76, 120–133 (2014).