

The impact of rare protein coding genetic variation on adult cognitive function

Chia-Yen Chen^{1*}, Ruoyu Tian^{1§}, Tian Ge²⁻⁴, Max Lam⁴⁻⁶, Gabriela Sanchez-Andrade⁷, Tarjinder Singh^{4,8}, Lea Urpa¹⁰, Jimmy Z. Liu^{1^}, Mark Sanderson⁷, Christine Rowley⁷, Holly Ironfield⁷, Terry Fang¹, Biogen Biobank Team, the SUPER-Finland study, the Northern Finland Intellectual Disability study, Mark Daly^{4,8,10,11}, Aarno Palotie^{4,8,10,11}, Ellen A. Tsai¹, Hailiang Huang^{4,8,9}, Matthew E. Hurles⁷, Sebastian S. Gerety⁷, Todd Lencz^{5,6,12}, Heiko Runz^{1*}

¹Research and Development, Biogen Inc.; Cambridge, MA, USA

²Psychiatric and Neurodevelopmental Genetics Unit, Center for Genomic Medicine, Massachusetts General Hospital; Boston, MA, USA

³Department of Psychiatry, Massachusetts General Hospital, Harvard Medical School; Boston, MA, USA

⁴Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard; Cambridge, MA, USA

⁵Division of Psychiatry Research, The Zucker Hillside Hospital, Northwell Health; Glen Oaks, NY, USA

⁶Institute of Behavioral Science, Feinstein Institutes for Medical Research; Manhasset, NY, USA

⁷Wellcome Sanger Institute; Hinxton, Cambridge

⁸Analytic and Translational Genetics Unit, Massachusetts General Hospital; Boston, MA, USA

⁹Department of Medicine, Harvard Medical School; Boston, MA, USA

¹⁰Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki; Helsinki, Finland

¹¹Program in Medical and Population Genetics, Broad Institute of MIT and Harvard; Cambridge, Massachusetts, USA

¹²Departments of Psychiatry and Molecular Medicine, Zucker School of Medicine at Hofstra/Northwell; Hempstead, NY, USA

[§]Current address: Dewpoint Therapeutics; Boston, MA, USA

[^]Current address: GlaxoSmithKline; Upper Providence, Philadelphia, PA, USA

*Corresponding authors. Email: chiayenc@gmail.com (C.-Y.C.); heiko.runz@gmail.com (H.R.)

Supplementary Materials

Table of Contents

Supplementary note	3
Brief summary of cognitive function-associated genes identified	3
Biogen Biobank team	6
The SUPER-Finland study researchers	6
Supplementary Figures	9
Reference	33

Supplementary note

Brief summary of cognitive function-associated genes identified

ADGRB2. *ADGRB2* (adhesion G protein-coupled receptor B2; also known as *BAI2*) encodes an adhesion G protein coupled receptor (GPCR) that is one of the main mediators of signal transduction in the central nervous system. *ADGRB2* is considered as an orphan GPCR (oGPCR), for which endogenous ligands have not yet been identified ¹. *ADGRB2* is primarily expressed in the brain (neurons and astrocytes in hippocampus, amygdala and cerebral cortex) ^{2,3}. Variants near *ADGRB2* have been associated with educational attainment in a genome-wide association study ⁴, and also found associated with other traits such as body mass index ⁵, smoking, Intraocular pressure ⁶, or parental longevity ⁷.

KDM5B. *KDM5B* (lysine demethylase 5B; also known as *JARID1B* or *PLUI*) encodes a lysine-specific histone demethylase in the jumonji/ARID domain-containing family of histone demethylases. The encoded protein can demethylate tri-, di- and monomethylated lysine 4 of histone H3 (H3K4me1/2/3) ⁸⁻¹⁰, which is broadly associated with enhancers and promoters of actively transcribed genomic loci. Mutations in *KDM5B* are the cause for an autosomal-recessive intellectual disability syndrome ¹¹ (OMIM # 618109) and have further been found associated with schizophrenia ¹² and autism spectrum disorder ^{13,14} in sequencing studies, where disrupted neuronal differentiation was suggested as a potential mechanism. A search on GWAS Catalog (<https://www.ebi.ac.uk/gwas/>; accessed on Feb. 6, 2022) did not identify significant associations of *KDM5B* variants in previous GWAS. However, we note that the association of *KDM5B* with RT may be influenced by its association with reduced handgrip strength we observed in UKB, which might contribute to the epidemiological observation in UKB that hand grip strength and cognitive function share common mechanisms ¹⁵.

GIGYF1. *GIGYF1* (GRB10 interacting GYF protein 1) encodes an adaptor protein (a member of the gyf family) that binds growth factor receptor-bound 10 (GRB10), which in turn binds activated insulin receptors and insulin-like growth factor-1 (IGF-1) receptors ^{16,17}. By influencing the insulin and IGF-1 signaling pathway, *GIGYF1* plays a role in metabolic diseases and related anthropometric traits. For instance, significant associations were identified in previous GWAS for hemoglobin ¹⁸, total cholesterol, low density lipoprotein cholesterol, glucose and apolipoprotein B levels ¹⁹. *GIGYF1* was also associated with mosaic loss of chromosome Y (LOY) ²⁰ and metabolic diseases including glucose and HbA1c levels and type 2 diabetes ²¹ in previous exome sequencing studies.

ANKRD12. *ANKRD12* (ankyrin repeat domain 12; also known as *ANCO-2*) encodes a member of the ankyrin repeats-containing cofactor (ANCO) family. ANCOs are transcriptional co-regulators that interact with both co-activators and co-repressors²². *ANKRD12* interacts with the p160 co-activators (by recruiting HDACs [histone deacetylases]) and the co-activator ADA3 (alteration/deficiency in activation 3)^{22,23}. *ANKRD12* was found to be associated with corpuscular measures in GWAS^{5,24}.

SLC8A1. *SLC8A1* (solute carrier family 8 member A1; also known as *NCX1*) encodes a bidirectional calcium transporter, the cardiac sarcolemmal Na(+)-Ca(2+) exchanger, which is the primary mechanism for cardiac myocyte returning to its resting state following excitation (through extrusion of calcium) and plays a critical role in cardiac contractility²⁵. *SLC8A1* expression is enriched in human heart tissue. *SLC8A1* has been shown to be associated with bone mineral density²⁶, blood pressure²⁷, blood biomarkers (for example IGF-1¹⁹), electrocardiographic traits (PR interval²⁸, QT interval²⁹, etc.) and hand grip strength³⁰ among others.

RC3H2. *RC3H2* (ring finger and CCCH-type domains 2) encodes roquin-2 that belongs to a family of highly conserved RNA-binding proteins (roquins) that regulate their target genes on the post-transcriptional level. Roquins contain a RING (Really Interesting New Gene)-type E3 ubiquitin ligase domain, followed by a ROQ domain and a CCCH-type ZnF domain³¹⁻³³. Roquins play key roles in maintaining peripheral immunological tolerance and autoimmune diseases³⁴. It has been shown that *RC3H2* (and *RC3H1*) restricts T-cell activation and costimulation via *ICOS* and *OX40* to prevent inappropriate Tfh cell differentiation³⁵. Roquin-2 is widely expressed in all human tissues. *RC3H2* showed genome-wide significant association with insomnia³⁶ and HbA1c⁵ in GWAS.

CACNA1A. *CACNA1A* (calcium voltage-gated channel subunit alpha1 A) encodes the alpha-1A subunit of the voltage-dependent calcium channels. It is primarily expressed in neuronal tissue. Mutations in *CACNA1A* are a cause for type 2 episodic ataxia (OMIM #108500), spinocerebellar ataxia 6 (OMIM #183086), developmental and epileptic encephalopathy 42 (OMIM #617106) and familial hemiplegic migraine (OMIM #141500). *CACNA1A* was implicated in a previous educational attainment GWAS⁴, but the top associated SNP and LD peak do not fall into the *CACNA1A* gene region, but rather located in the intergenic region between *CACNA1A* and *RPL12P42*. Other GWAS associations for *CACNA1A* include depressive symptoms³⁷, age at first birth³⁸ and brain region volume³⁹.

BCAS3. *BCAS3* (BCAS3 microtubule associated cell migration factor) encodes a large, highly conserved cytoskeletal protein involved in human embryogenesis and tumor angiogenesis^{40,41}. It has recently been shown that *BCAS3* loss-of-function variants can cause Hengel-Marooofian-Schols syndrome (HEMARS; OMIM # 619641), which is an autosomal recessive

neurodevelopmental disorder characterized by severe global developmental delay starting from infancy or early childhood with facial dysmorphism and brain abnormalities ⁴¹. *BCAS3* has also been associated with glomerular filtration rate ⁴², bone mineral density ²⁶, serum creatinine level ⁵, hemoglobin concentration ²⁴, serum urate level ⁴³, red blood cell count ⁵, ophthalmologic measures (e.g. macular thickness ⁴⁴), coronary artery disease ⁴⁵ and additional traits in GWAS.

Biogen Biobank team

Steering team

Ellen Tsai, Christopher D. Whelan, Paola Bronson, David Sexton, Sally John, Heiko Runz

Data management team

Eric Marshall, Mehool Patel, Saranya Duraisamy, Timothy Swan

Extended Scientific team

Dennis Baird, Chia-Yen Chen, Susan Eaton, Jake Gagnon, Feng Gao, Cynthia Gubbels, Yunfeng Huang, Varant Kupelian, Kejie Li, Dawei Liu, Stephanie Loomis, Helen McLaughlin, Adele Mitchell, Nilanjana Sadhu, Benjamin Sun, Ruoyu Tian

The SUPER-Finland study researchers

	AFFILIATION
Aarno Palotie	Institute for Molecular Medicine, Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland; Broad Institute of MIT and Harvard, Cambridge, MA, USA; Massachusetts General Hospital Massachusetts General Hospital, Boston, MA, USA
Aija Kyttälä	Finnish Institute for Health and Welfare (THL), Helsinki, Finland
Amanda Elliott	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland; Broad Institute, Cambridge, MA, USA and Massachusetts General Hospital, Boston, MA, USA
Andre Sourander	Department of Child Psychiatry, University of Turku, Turku, Finland
Annamari Tuulio-Henriksson	Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland
Anssi Solismaa	Tampere University and Tampere University Hospital, Tampere, Finland
Antti Tanskanen	Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden and Impact Assessment Unit, Finnish Institute for Health and Welfare (THL), Helsinki, Finland
Ari Ahola-Olli	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Helsinki, Finland
Arto Mustonen	University of Turku, Turku, Finland
Arttu Honkasalo	University of Helsinki, Helsinki, Finland
Asko Wegelius	Department of Psychiatry, University of Helsinki and Helsinki University Hospital, Finland
Atiqul Mazumder	Unit of Clinical Neuroscience, Faculty of Medicine, University of Oulu, Oulu, Finland
Auli Toivola	Finnish Institute for Health and Welfare (THL), Helsinki, Finland
Benjamin Neale	Broad Institute of MIT and Harvard, Cambridge, MA, USA
Elina Hietala	Tampere University and Tampere University Hospital
Elmo Saarentaus	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Helsinki, Finland
Erik Cederlöf	Finnish Institute for Health and Welfare (THL), Helsinki, Finland
Erkki Isometsä	Department of Psychiatry, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

Heidi Taipale	Kuopio Research Center of Geriatric Care, University of Eastern Finland, Kuopio, Finland and School of Pharmacy, University of Eastern Finland, Kuopio, Finland, Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden
Imre Västrik	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Helsinki, Finland
Jaana Suvisaari	Mental Health Unit, Finnish Institute for Health and Welfare, Helsinki, Finland
Jari Tiihonen	Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden and Department of Forensic Psychiatry, Niuvanniemi Hospital, University of Eastern Finland, Kuopio, Finland
Jarmo Hietala	Department of Psychiatry, Turku University Hospital, Turku, Finland
Johan Ahti	Department of Psychiatry, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
Jonne Lintunen	Department of Forensic Psychiatry, Niuvanniemi Hospital, University of Eastern Finland, Kuopio, Finland
Jouko Lönnqvist	Finnish Institute for Health and Welfare (THL), Helsinki, Finland and University of Helsinki, Helsinki, Finland
Juha Veijola	Department of Psychiatry, Research Unit of Clinical Neuroscience, University of Oulu, Oulu, Finland and Department of psychiatry, University Hospital of Oulu, Oulu, Finland
Julia Moghadampour	Tampere University and Tampere University Hospital, Helsinki, Finland
Jussi Niemi-Pynttäre	Department of Psychiatry, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
Kaisla Lahdensuo	Mehiläinen, Helsinki, Finland
Katja Häkkinen	Department of Forensic Psychiatry, Niuvanniemi Hospital, University of Eastern Finland, Kuopio, Finland
Katriina Hakakari	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
Kimmo Suokas	Tampere University Hospital, Tampere, Finland and Department of Psychiatry, Pirkanmaa Hospital District, Tampere, Finland
Lea Urpa	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Helsinki, Finland
Marjo Taivalantti	Research unit of clinical neuroscience, Faculty of medicine, University of Oulu, Oulu, Finland
Mark Daly	Institute for Molecular Medicine, Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland; Broad Institute of MIT and Harvard, Cambridge, MA, USA; Massachusetts General Hospital Massachusetts General Hospital, Boston, MA, USA
Markku Lähteenvuo	Department of Forensic Psychiatry, Niuvanniemi Hospital, University of Eastern Finland, Kuopio, Finland
Martta Kerkelä	Research Unit of Clinical Neuroscience, University of Oulu, Oulu, Finland
Minna Holm	Mental Health Unit, Finnish Institute for Health and Welfare, Helsinki, Finland
Nina Lindberg	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
Noora Ristiluoma	Finnish Institute for Health and Welfare (THL), Helsinki, Finland
Olli Kampman	Tampere University and Tampere University Hospital, Tampere, Finland
Olli Pietiläinen	Neuroscience Center, HiLIFE, University of Helsinki, Helsinki, Finland

Risto Kajanne	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Helsinki, Finland
Sari Lång-Tonteri	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
Solja Niemelä	Department of Psychiatry, University of Turku, Turku, Finland
Steven E. Hyman	Broad Institute of MIT and Harvard, Cambridge, MA, USA
Susanna Rask	Tampere University and Tampere University Hospital
Tarjinder Singh	Broad Institute of MIT and Harvard, Cambridge, MA, USA
Teemu Männynsalo	Department of Psychiatry, University of Helsinki and Helsinki University Hospital, Helsinki, Finland
Tiina Paunio	Mental Health Unit, Finnish Institute for Health and Welfare, Helsinki, Finland
Tuomas Jukuri	Department of Psychiatry, Oulu University Hospital, Oulu, Finland
Tuomo Kiiskinen	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Helsinki, Finland
Tuula Kiesepä	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
Ville Mäkipelto	University of Helsinki, Helsinki, Finland
Willehard Haaki	Department of Psychiatry, University of Turku, Turku, Finland and Department of Psychiatry, Turku University Hospital, Turku, Finland
Zuzanna Misiewicz	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Helsinki, Finland

Supplementary Figures

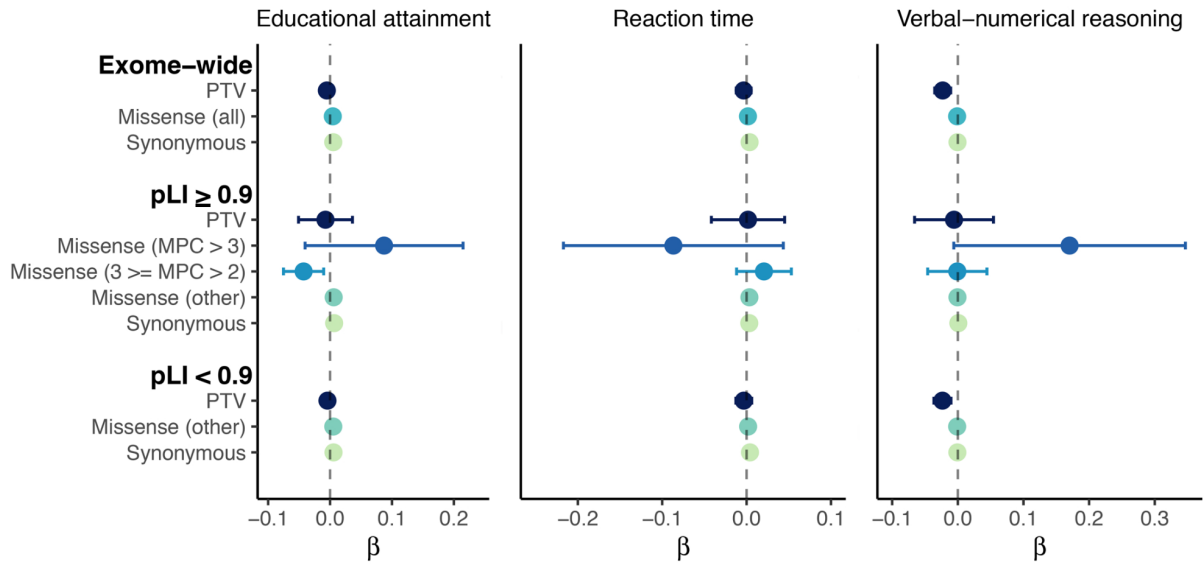


Fig. S1. Impact of exome-wide burden of rare protein coding variants on educational attainment (EDU), reaction time (RT) and verbal-numerical reasoning (VNR) in South Asian (SAS) samples in the UK Biobank.

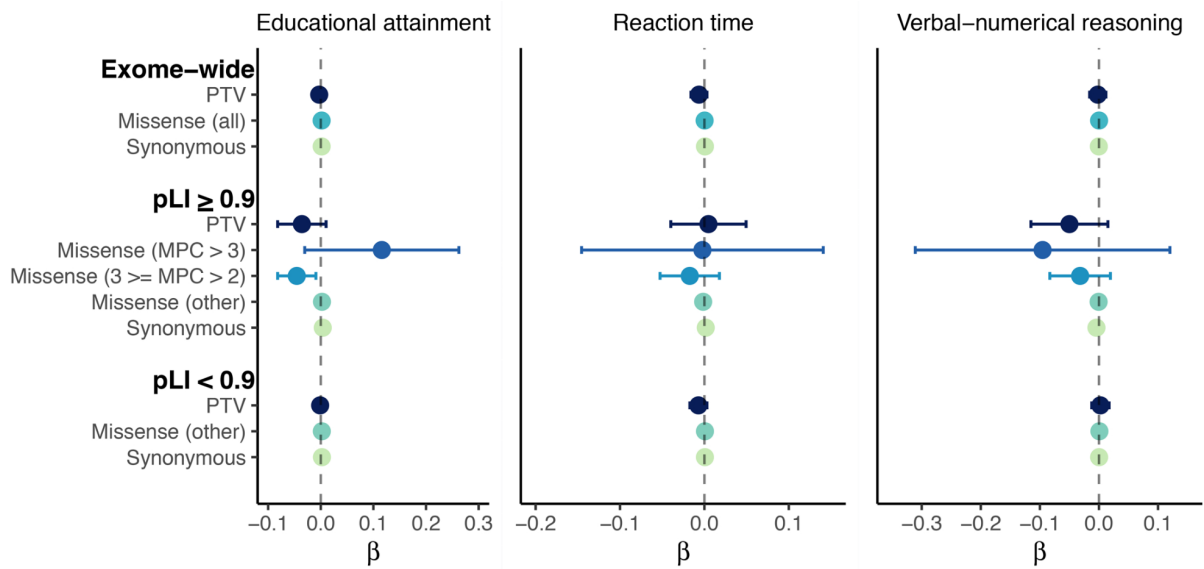


Fig. S2. Impact of exome-wide burden of rare protein coding variants on educational attainment (EDU), reaction time (RT) and verbal-numerical reasoning (VNR) in African samples (AFR) in the UK Biobank.

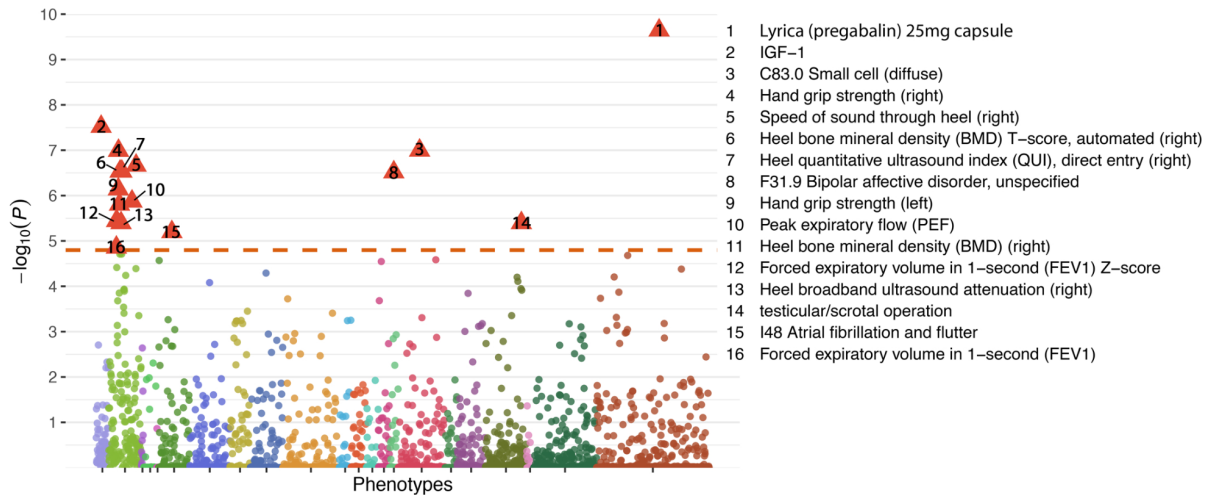
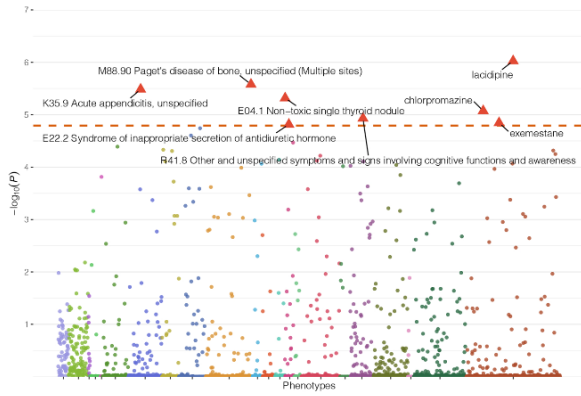
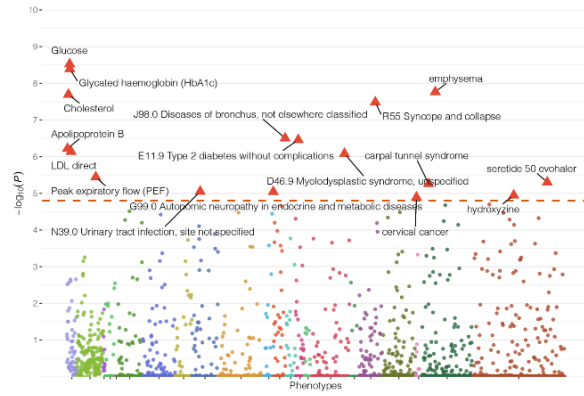


Fig. S3. PTV burden-based phenome-wide association analysis (3,150 phenotypes) for KDM5B in the UK Biobank European samples. FEV1 Z-score is Inverted GLI 2012 z-score for FEV1. Phenotypes were grouped and color-coded from left to right in the following categories: biomarker; composite phenotypes; family history; ICD-10 cause of death, ICD-10 congenital malformations; deformations and chromosomal abnormalities; ICD-10 diseases of the circulatory system; ICD-10 diseases of the digestive system; ICD-10 diseases of the eye and adnexa; ICD-10 diseases of the genitourinary system; ICD-10 diseases of the musculoskeletal system and connective tissue; ICD-10 diseases of the nervous system; ICD-10 diseases of the respiratory system; ICD-10 diseases of the skin and subcutaneous tissue; ICD-10 endocrine, nutritional and metabolic diseases; ICD-10 mental, behavioral and neurodevelopmental disorders; ICD-10 neoplasms; ICD-10 pregnancy, childbirth and the puerperium; ICD-10 symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified; operation code; self-reported illness: cancer; self-reported illness: non-cancer; self-reported medication.

ADGRB2



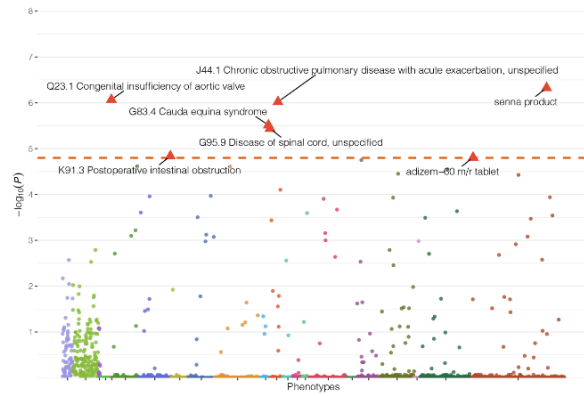
GIGYF1



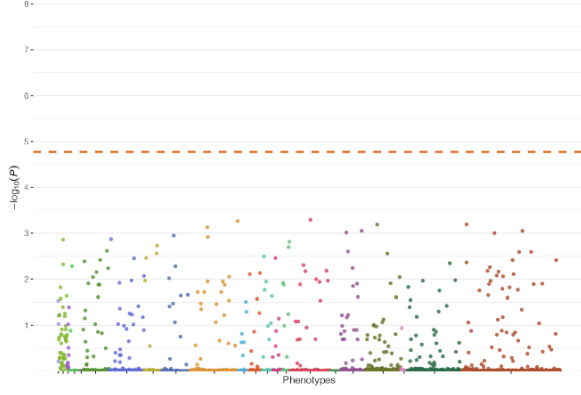
ANKRD12



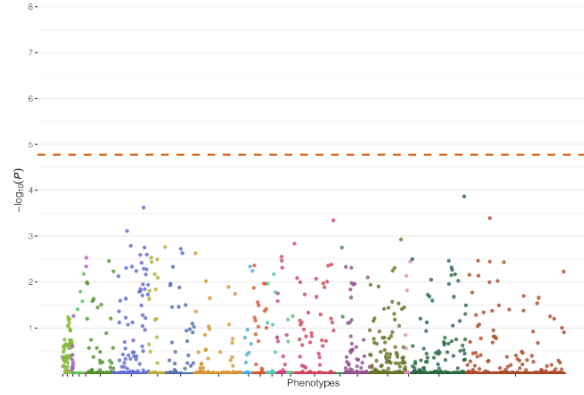
SLC8A1



RC3H2



CACNA1A



BCAS3

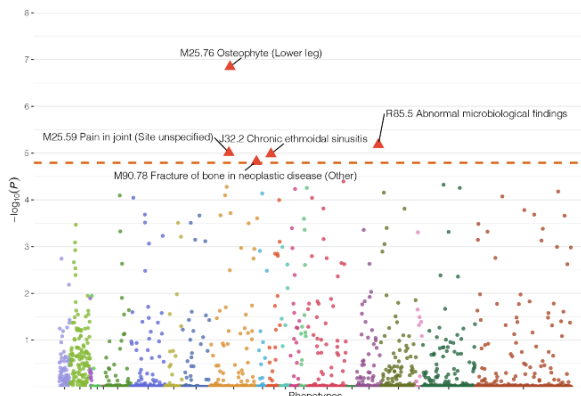


Fig. S4. Phenome-wide association analysis (3,150 phenotypes) for *ADGRB2*, *GIGYF1*, *ANKRD12*, *SLC8A1*, *RC3H2*, *CACNA1A* and *BCAS3* in unrelated European samples in the UK Biobank. Phenotypes were grouped and color-coded from left to right in the following categories: biobmarker; composite phenotypes; family history; ICD-10 cause of death, ICD-10 congenital malformations; deformations and chromosomal abnormalities; ICD-10 diseases of the circulatory system; ICD-10 diseases of the digestive system; ICD-10 diseases of the eye and adnexa; ICD-10 diseases of the genitourinary system; ICD-10 diseases of the musculoskeletal system and connective tissue; ICD-10 diseases of the nervous system; ICD-10 diseases of the respiratory system; ICD-10 diseases of the skin and subcutaneous tissue; ICD-10 endocrine, nutritional and metabolic diseases; ICD-10 mental, behavioral and neurodevelopmental disorders; ICD-10 neoplasms; ICD-10 pregnancy, childbirth and the puerperium; ICD-10 symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified; operation code; self-reported illness: cancer; self-reported illness: non-cancer; self-reported medication.

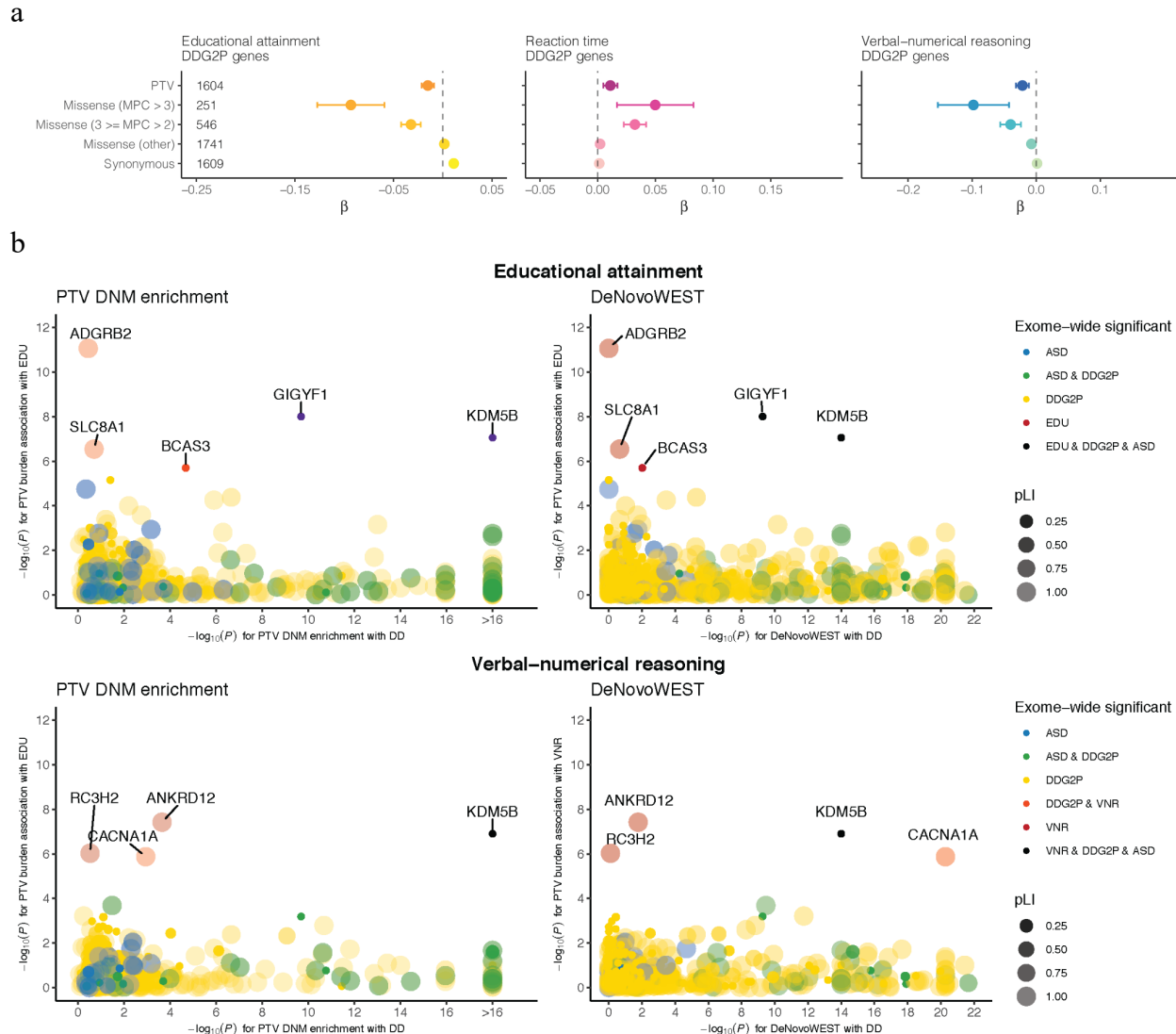


Fig. S5. Impact of rare coding variants in genes identified in the Developmental Disorder Genotype - Phenotype Database (DDG2P) on cognitive function.

a. The effects of protein-truncating, missense (stratified by MPC) and synonymous variant burden in exome sequencing study identified DDG2P on EDU, RT and VNR. DDG2P database (<https://www.deciphergenomics.org/ddd/ddgenes>) was accessed on December 23, 2020.

Missense variants were classified by deleteriousness (MPC) into 3 tiers: tier 1 with MPC > 3; tier 2 with $3 \geq \text{MPC} > 2$; tier 3 includes all missense variants not in tier 1 or 2. We note that the effect of damaging missense variants out scaled that of PTV burden for DDG2P genes. This is most likely explained by UKB participants being depleted for highly penetrant PTVs in this gene set that cause disease onset in childhood⁴⁶.

b. Comparison between gene-based associations for genes from DDG2P database, EDU and VNR (PTV DNM enrichment and DeNovoWEST for DD; rare PTV burden for EDU and VNR). Each dot represents a gene that is identified for DD in Kaplanis et al. 2020 and for EDU or VNR in the current exome analysis. The dots are color-coded according to the phenotypes (DD, ASD, or EDU) that the gene is exome-wide significantly associated with. The size and shade of the dots are representing the pLI for the gene. EDU and VNR genes are labeled with gene names.

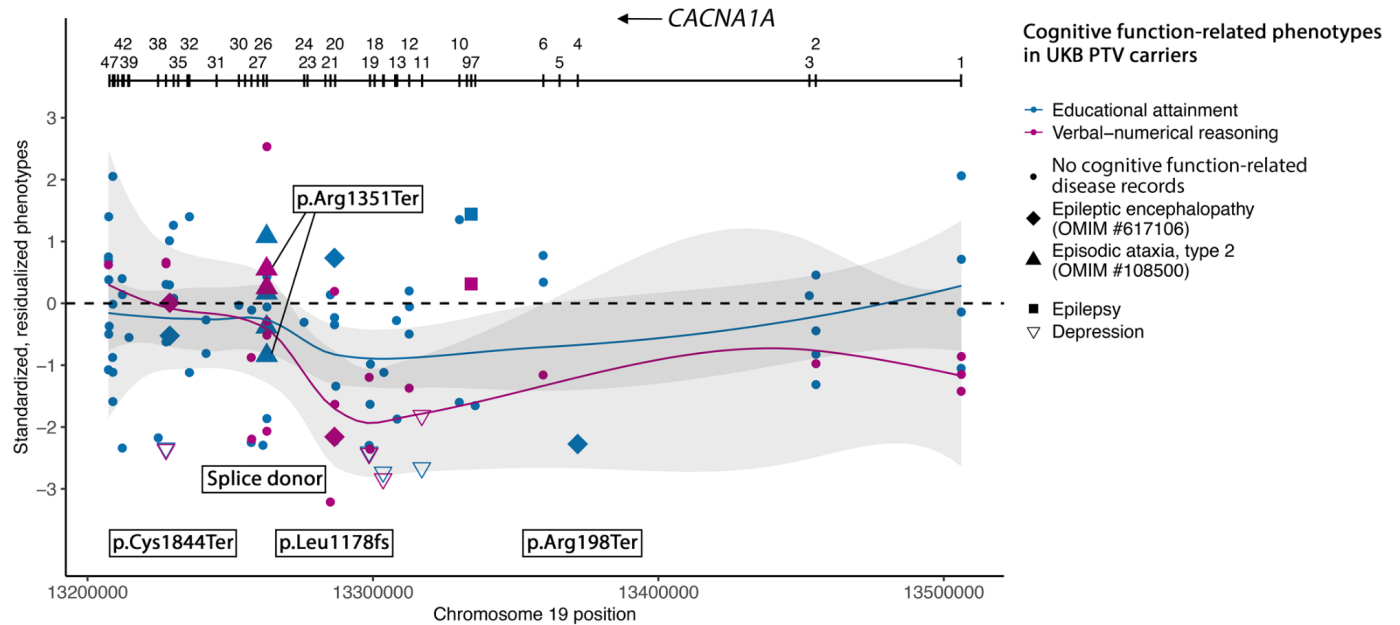


Fig. S6. Distribution of cognitive phenotypes (educational attainment and verbal-numerical reasoning) for *CACNA1A* PTV carriers. ClinVar pathogenic/likely pathogenic variants for epileptic encephalopathy (OMIM #617106) and/or type 2 episodic ataxia (OMIM #108500) was annotated. Samples with inpatient ICD-10 (International Classification of Diseases version-10) records of psychiatric (schizophrenia, bipolar disorder, depression, substance use disorder and/or anxiety and stress disorders), neurodegenerative and neurodevelopmental disorders were annotated. Phenotypes were residualized by sex, age, age², sex by age interaction, sex by age² interaction, top 20 PCs, and recruitment center and inverse rank-based normal transformed.

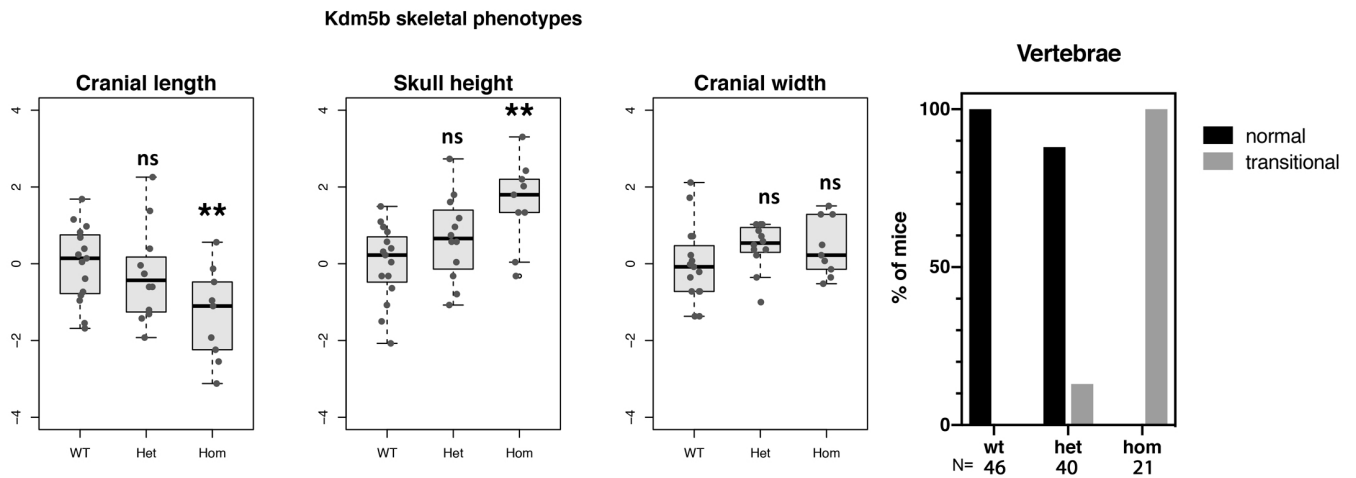


Fig. S7. *Kdm5b* loss-of-function impacts craniofacial and skeletal features in mice in a dose-dependent manner. An intermediate effect on cranial length (additive genotype effect $P=0.0268$) and height (additive genotype effect $P=0.0056$) is detected in *Kdm5b*^{+/-} mice, but not in cranial width (additive genotype effect $P=0.3090$). A fully penetrant transitional vertebrae phenotype seen in *Kdm5b*^{-/-} mice (N=21, Fisher's exact test vs *Kdm5b*^{+/+} [N=46] $P<0.001$) is observed at a lower frequency in *Kdm5b*^{+/-} mice (N=40, Fisher's exact test vs *Kdm5b*^{+/+} $P=0.0189$).

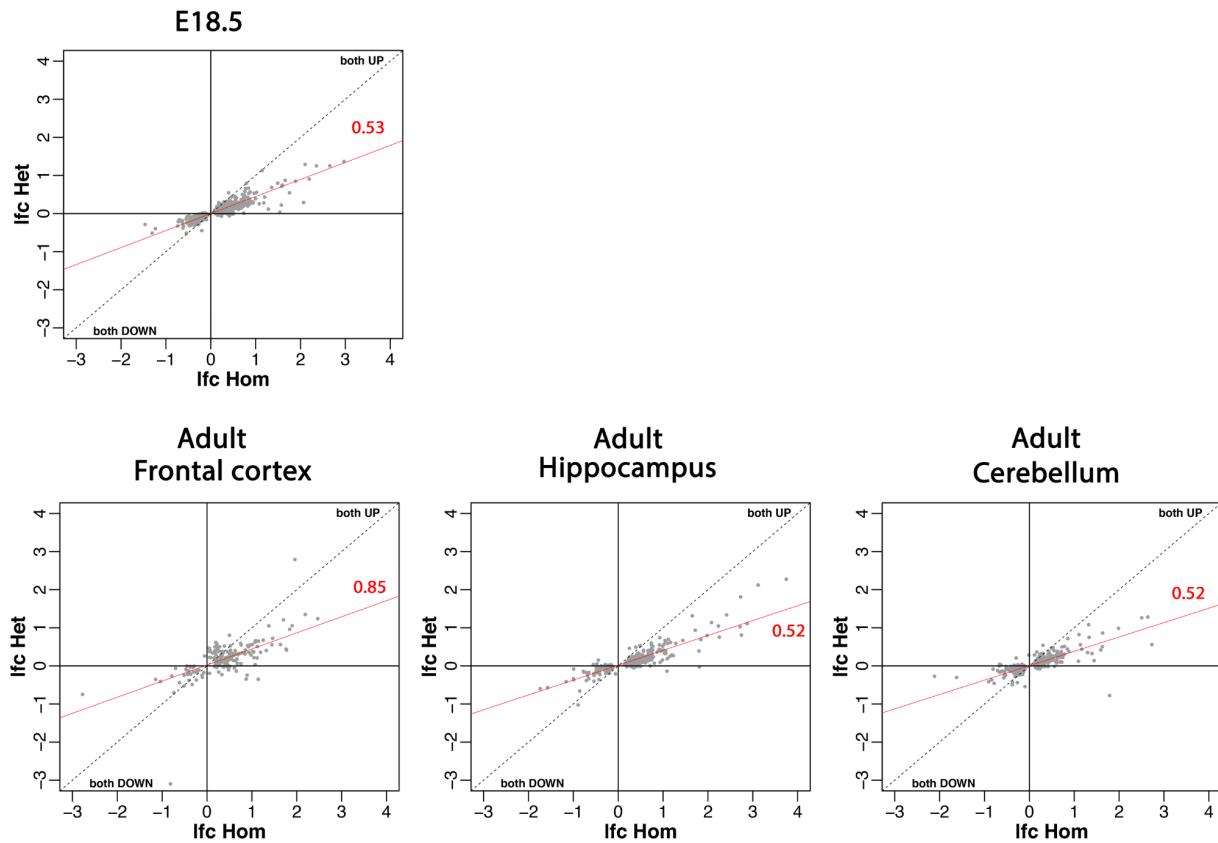


Fig. S8. Correlation in differential gene expression between heterozygous and homozygous *Kdm5b* mutant mice. LOG2-fold change of differentially expressed genes plotted for *Kdm5b*^{+/-} (y-axis) and *Kdm5b*^{-/-} (x-axis) mice across embryonic and adult brain tissues as indicated. There is a strong correlation between direction of change in expression in both mutant genotypes (robust linear regression line and slope shown in red).

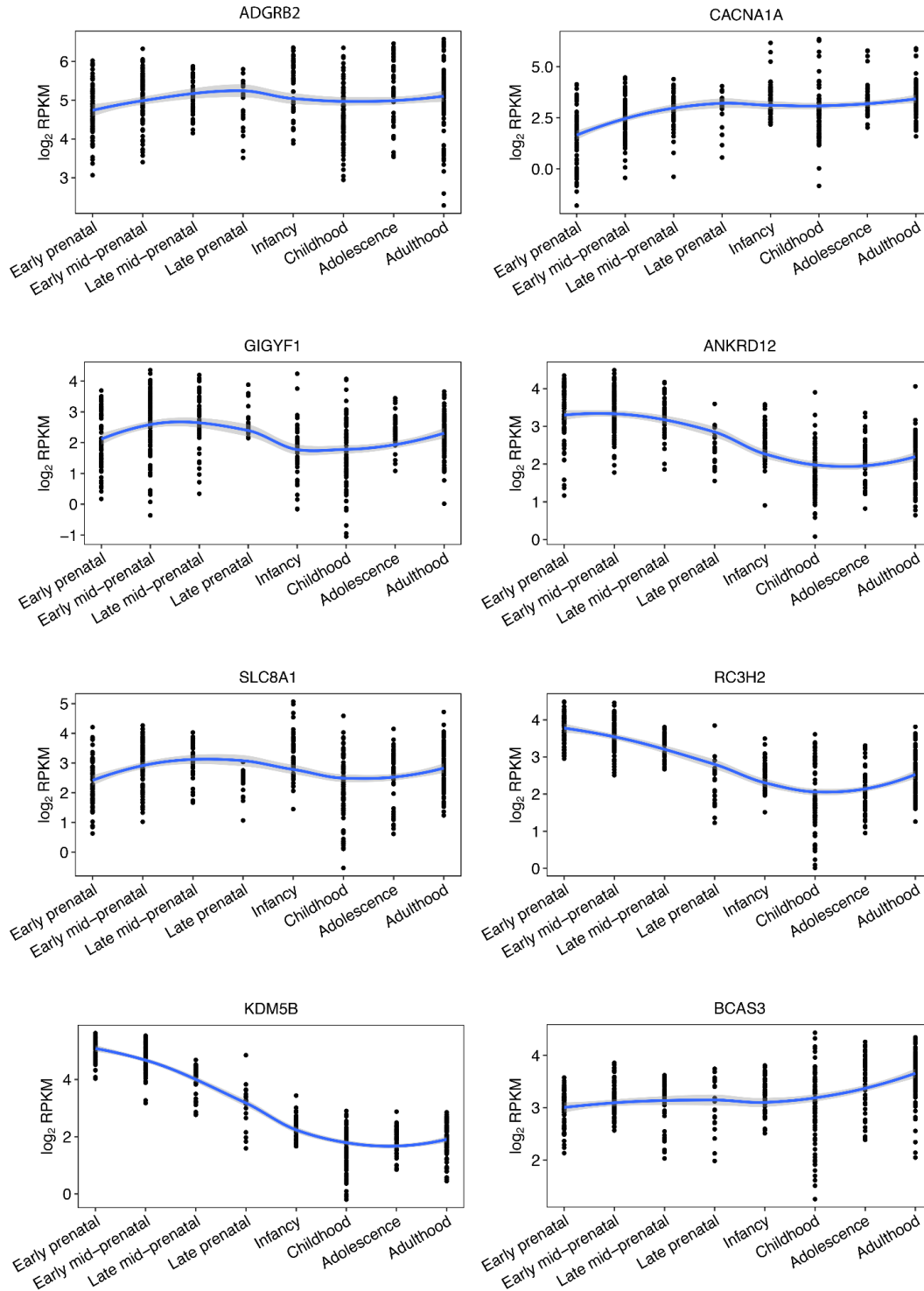
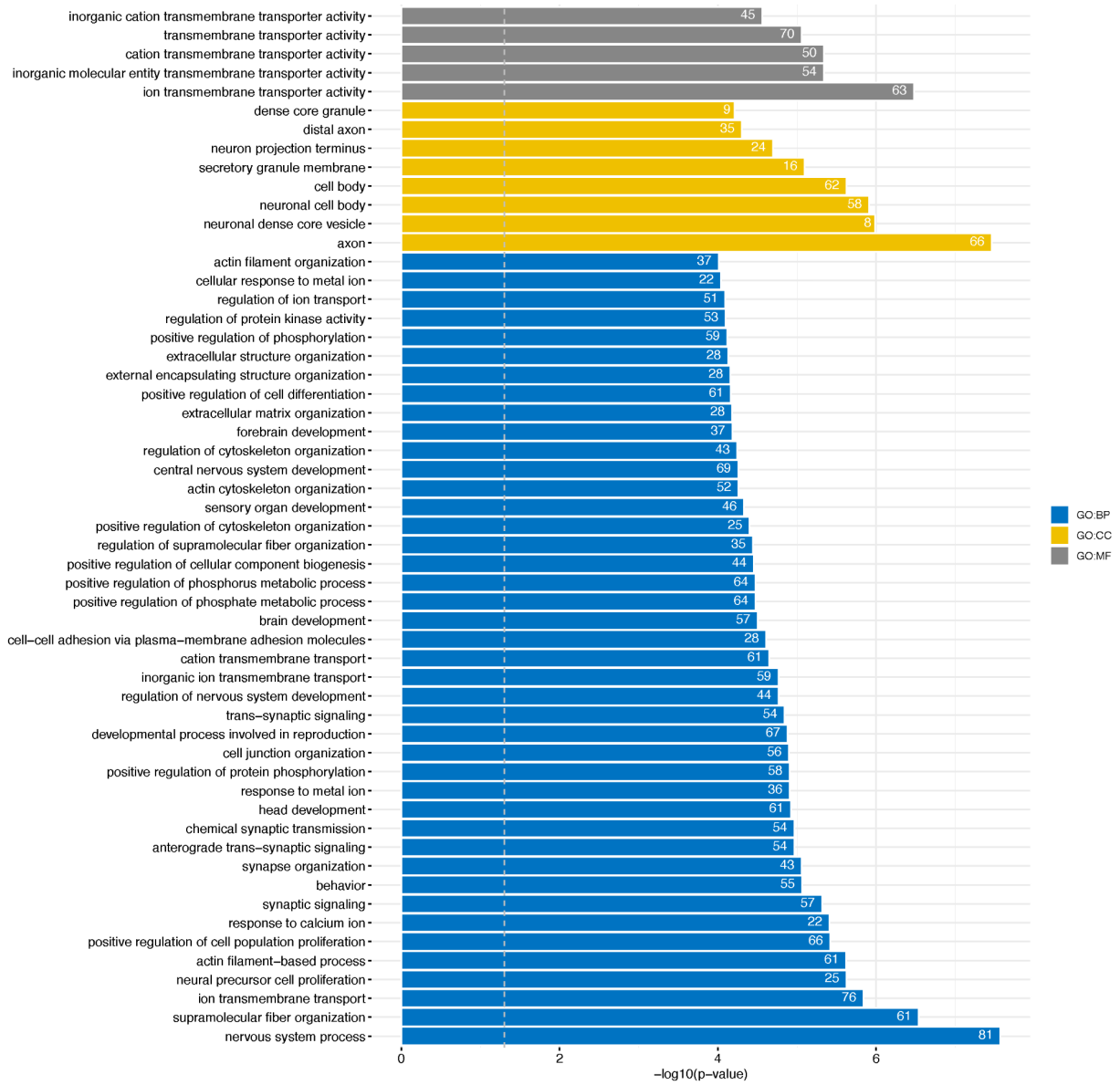
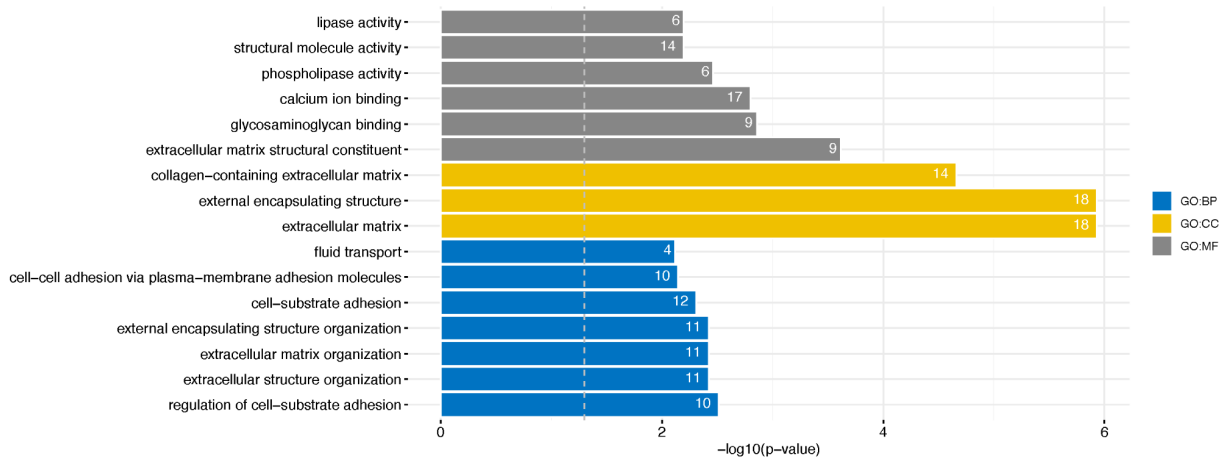


Fig. S9. Cognitive function gene expression in brain tissue at different developmental stages. RNA-seq data obtained from BrainSpan⁴⁷. Blue line represents fitted loess regression on *KDM5B* expression cross development stages.

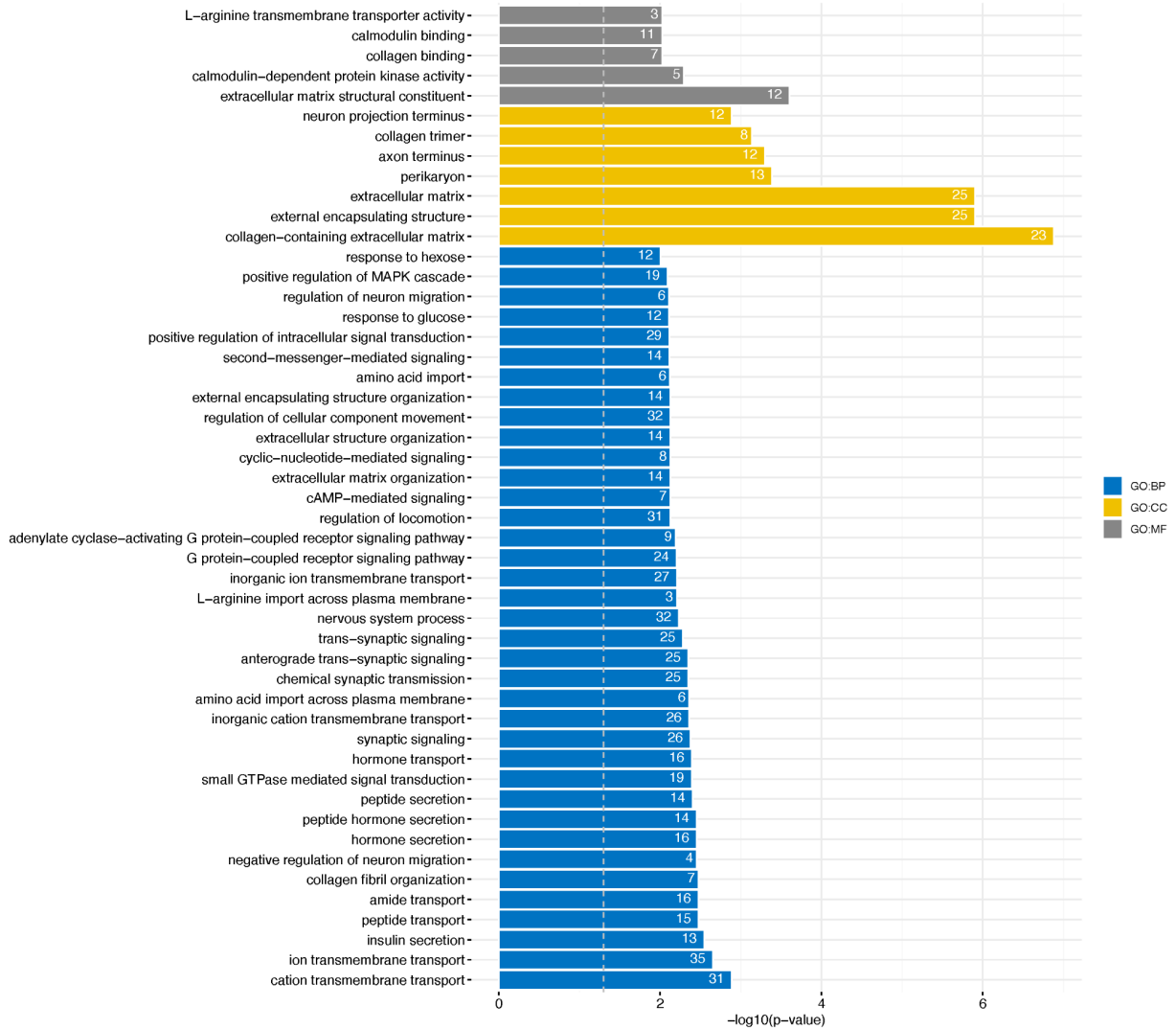
E18.5



Adult Frontal Cortex



Adult Hippocampus



Adult Cerebellum

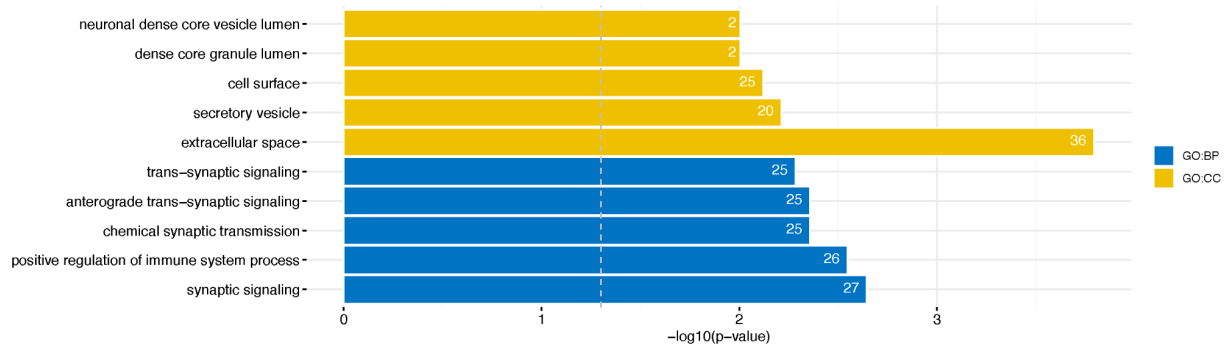


Fig. S10. GO term enrichment for differentially expressed genes in *Kdm5b* mutant mice. Differentially expressed genes (DEGs) from E18.5 and adult brain tissues of *Kdm5b*^{+/-} and *Kdm5b*^{-/-} mice were subject to Gene ontology (GO) pathway enrichment analysis using the gprofiler R package, with a threshold of 5% FDR and an enrichment significance threshold of $P < 0.05$ (hypergeometric test with FDR correction for multiple testing). For the E18.5 sample, we only showed results with enrichment p-value < 0.0001 (for display purposes). Full results are provided in Table S14. The European Nucleotide Archive accession numbers for the RNA-seq sequences reported are provided in Table S15. Background comprised only expressed genes in each tissue of interest.

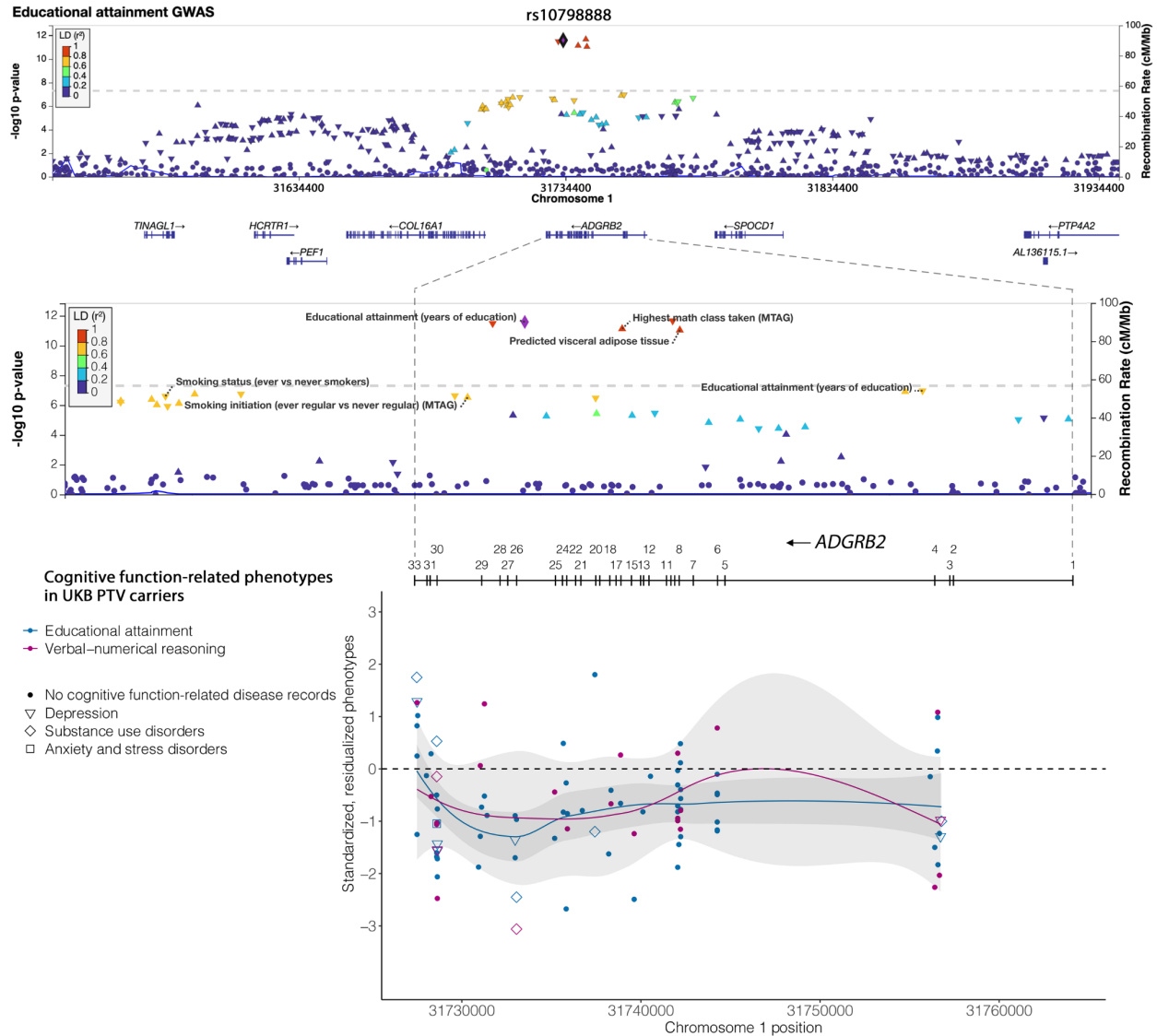


Fig. S11. Overlap between educational attainment GWAS (Lee et al. 2018⁴) locus on chromosome 1 and *ADGRB2* identified in PTV burden analysis in UKB. Regional plot of educational attainment GWAS association test results were generated around top independent SNP rs10798888. Additional associations from GWAS catalog were annotated with the associated phenotypes in the regional plot. EDU and VNR score for *ADGRB2* PTV carriers in UKB were plotted (both phenotypes were residualized by sex, age, age², sex by age², top 20 PCs and recruitment centers and were inverse rank-based normal transformed). Samples with inpatient ICD-10 (International Classification of Diseases version-10) records of psychiatric, neurodegenerative, and neurodevelopmental disorders were annotated.

Cognitive function GWAS

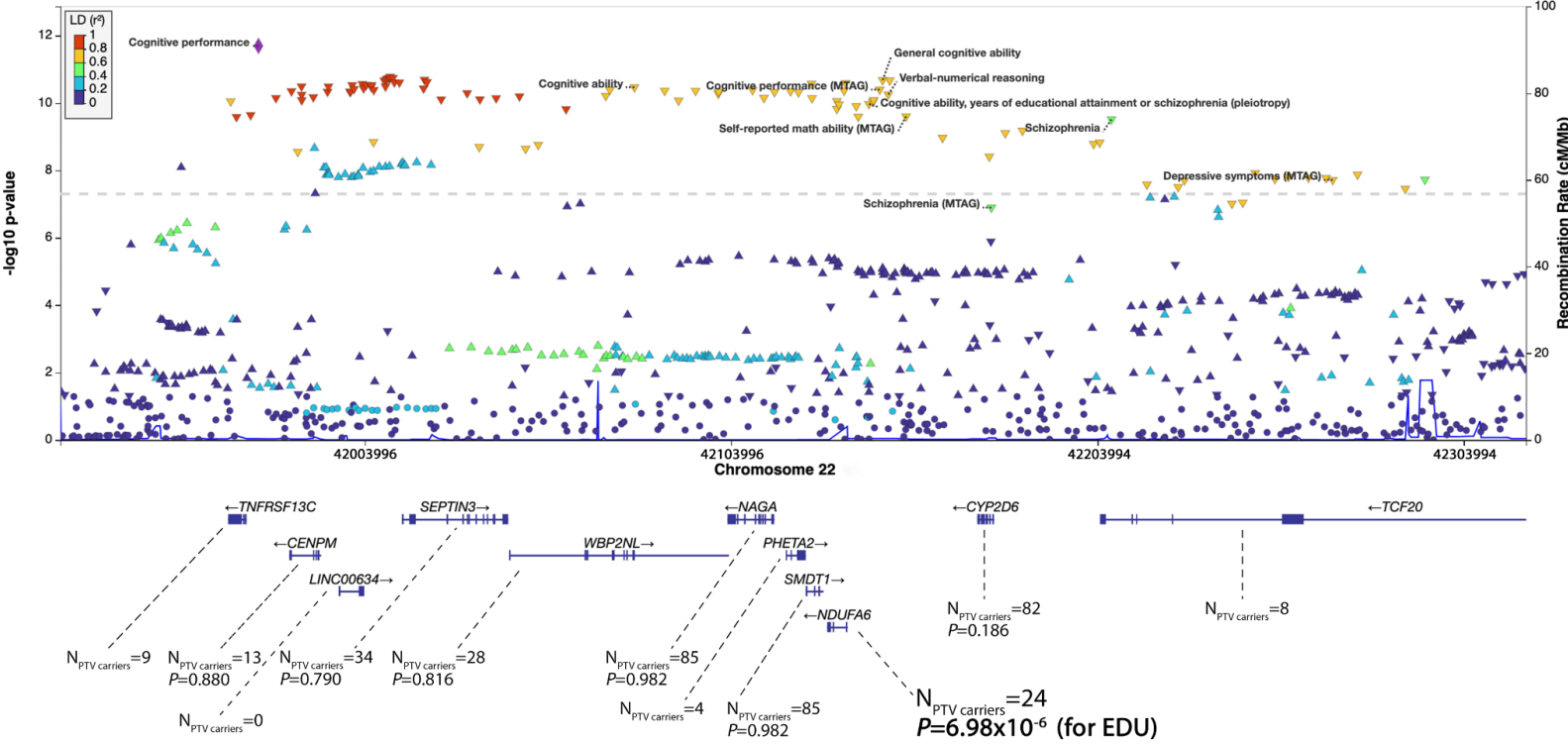


Fig. S12. Overlap between cognitive function GWAS (Lam et al. 2021⁴⁸) locus on chromosome 22 and *NDUFA6* identified in PTV burden analysis in UKB (FDR significant for EDU). Regional plot of cognitive function GWAS association test results were generated for top independent SNP rs5751191 and the extended LD region. Additional associations from GWAS catalog were annotated with the associated phenotypes in the regional plot. Number of PTV carriers and gene-based PTV burden association p-value were extracted for genes in the region.

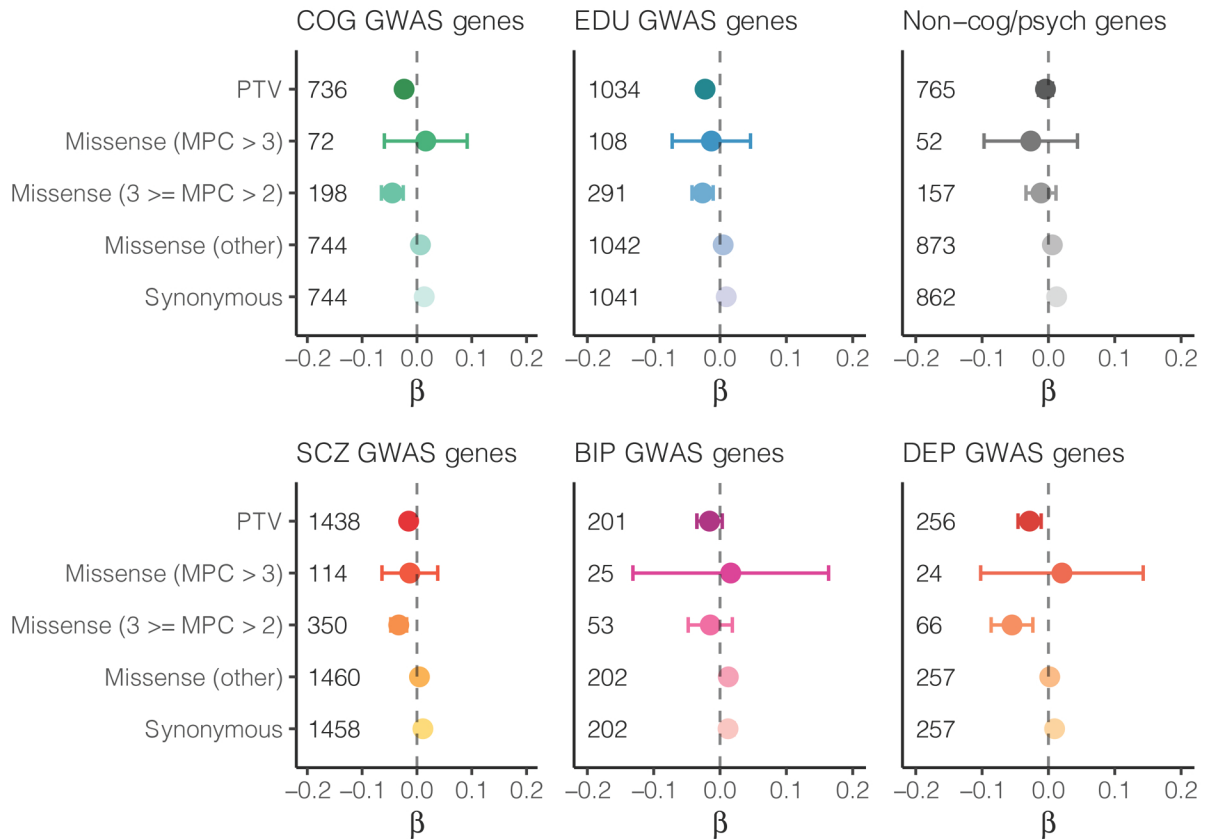


Fig. S13. Rare coding variant burden in genes identified in GWAS for cognitive function, educational attainment, schizophrenia, bipolar disorder and depression and non-cognitive function related genes on educational attainment (EDU). The impact of rare coding variant burden in genes identified through common variant association in GWAS for cognitive function (COG), educational attainment (EDU), schizophrenia (SCZ), bipolar disorder (BIP) and depression (DEP) and in non-cognitive function/non-psychiatric disorder-related (non-cog/psych) genes on EDU. Missense variants were classified by deleteriousness (MPC) into 3 tiers: MPC>3; 3 \geq MPC>2; and all missense variants not in the previous two tiers. The number of genes included in each burden was labeled in each panel.

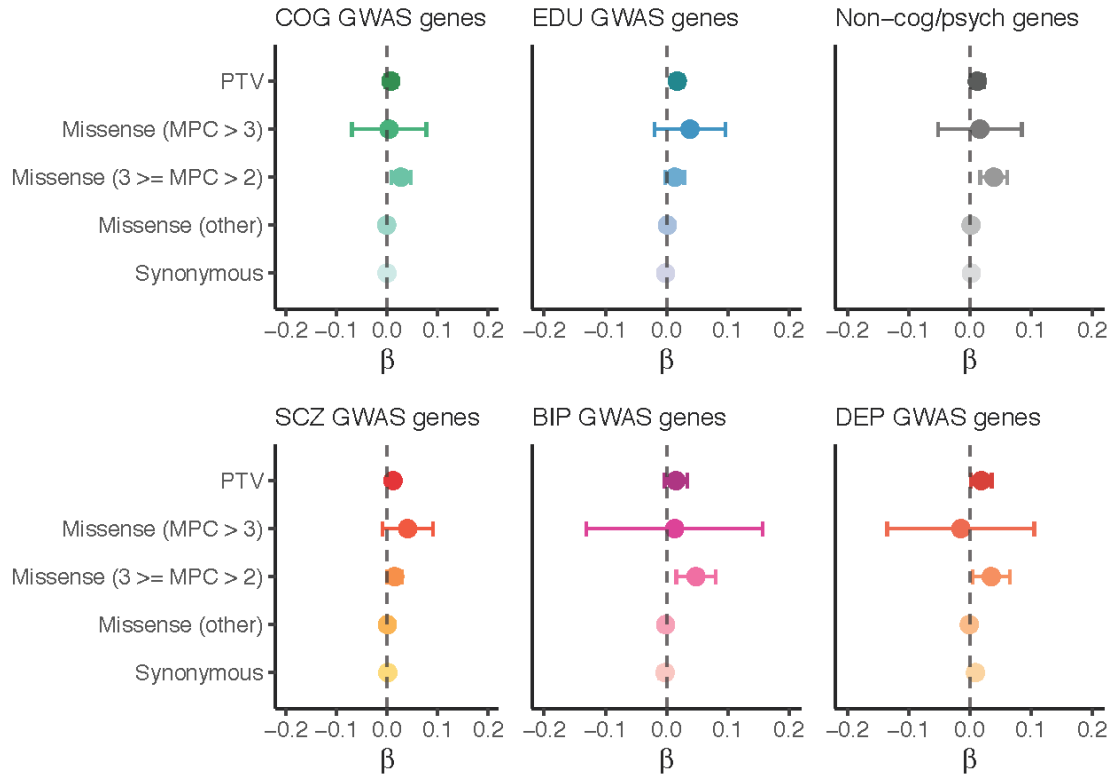


Fig. S14. Rare coding variant burden in genes identified in GWAS for cognitive function, educational attainment, schizophrenia, bipolar disorder and depression and non-cognitive function related genes on reaction time (RT).

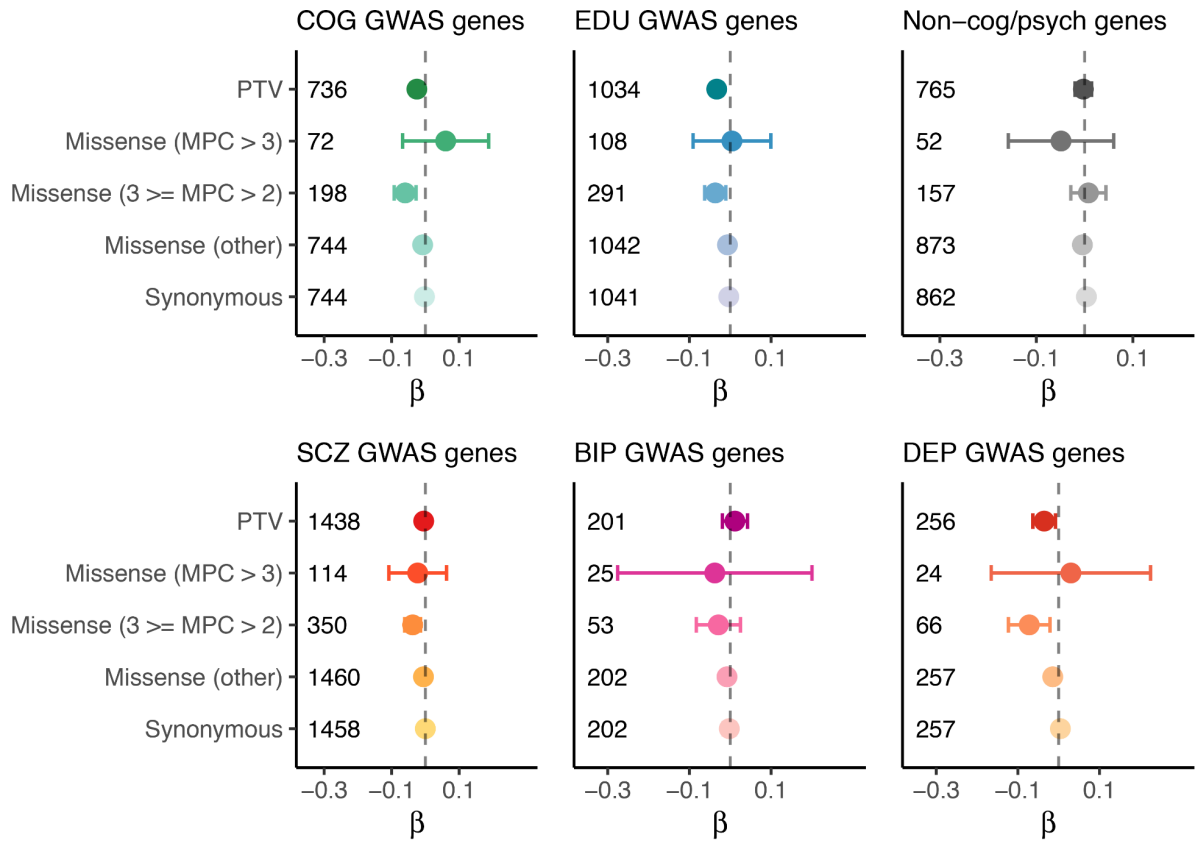
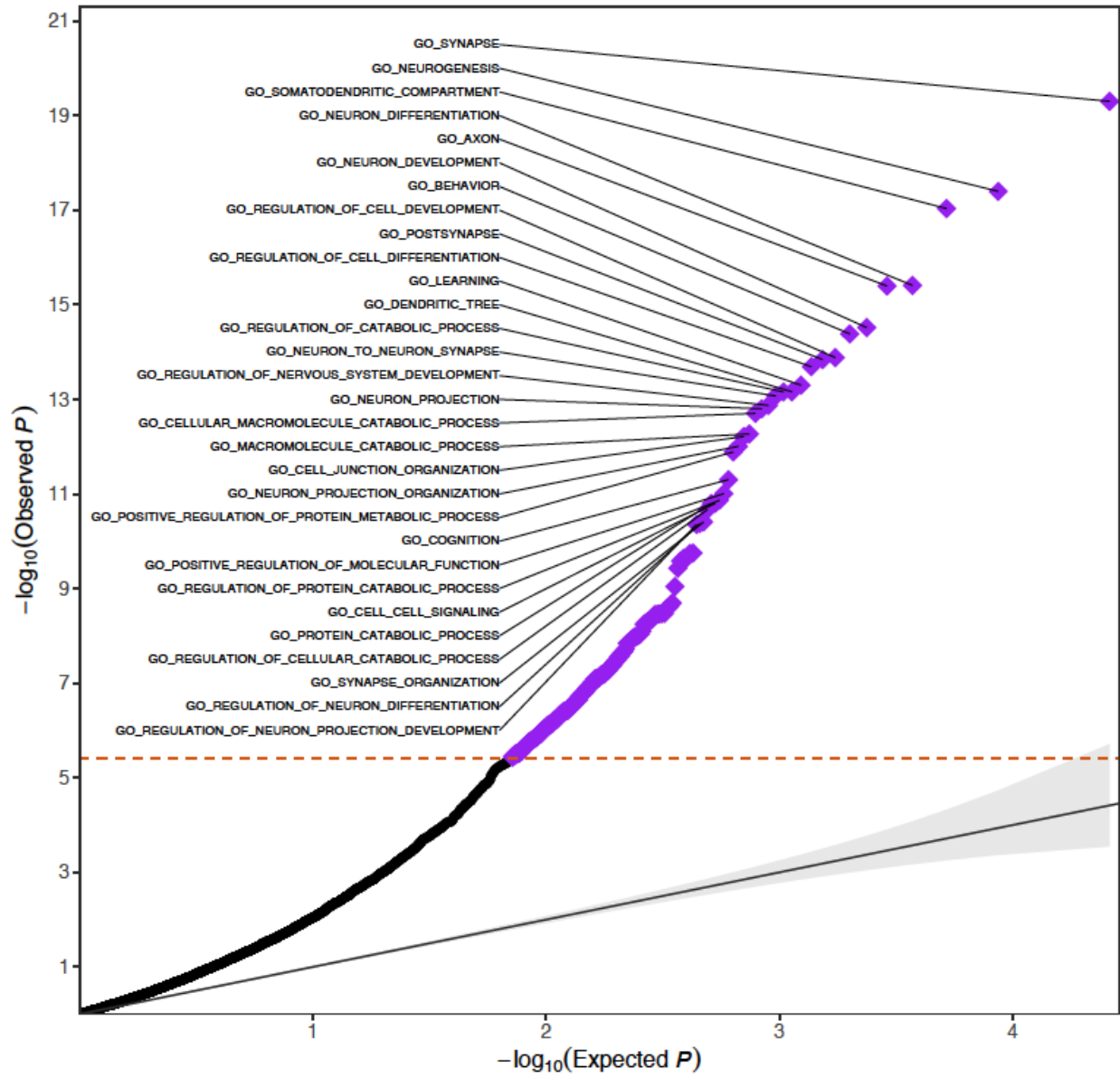
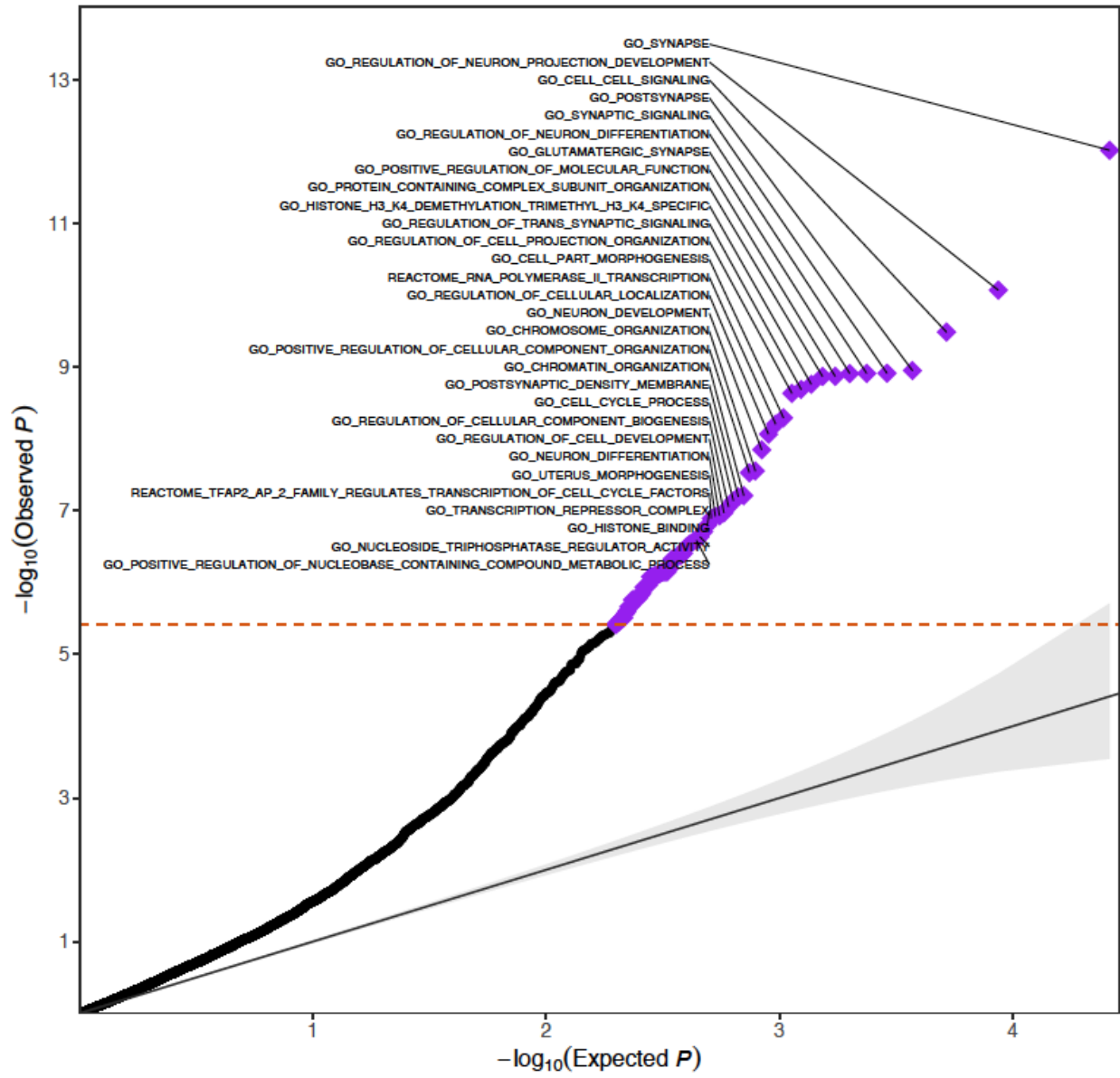


Fig. 15. The impact of rare coding variant burdens in genes identified in GWAS for cognitive function, educational attainment, schizophrenia, bipolar disorder and depression and non-cognitive function related genes on verbal-numerical reasoning (VNR).

Educational attainment



Reaction time



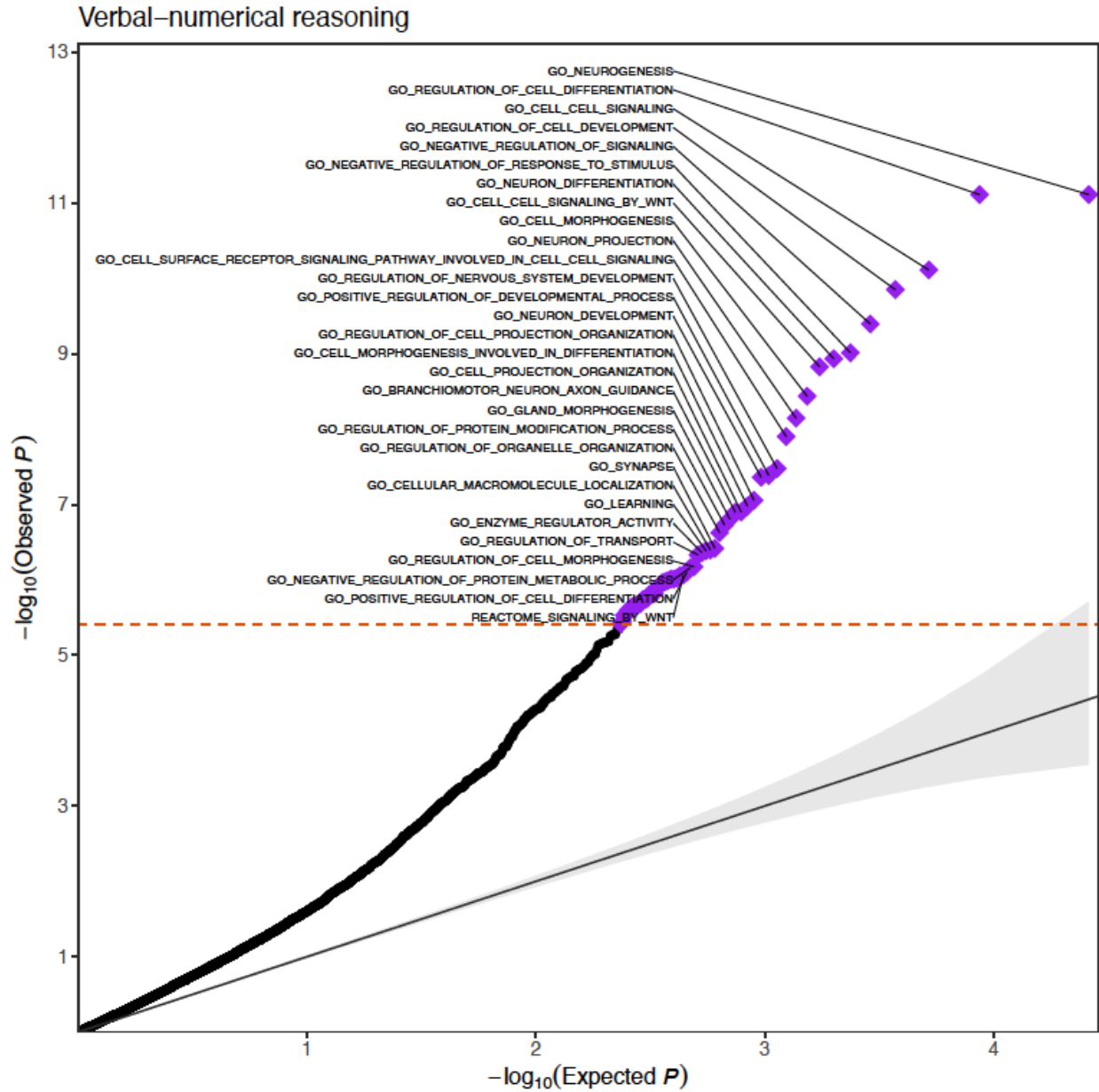


Fig. S16. Gene set-based PTV burden analysis in European samples in the UK Biobank for educational attainment, Reaction time and verbal-numerical reasoning. Top 30 gene sets were labeled in the figure. A total of 13,011 gene set from MSigDB v7.2 were identified, including C2 canonical pathways (N=2,808) and C5 Gene Ontology biological process (N = 7,531), cellular component (N = 996), and molecular function (N = 1,676).

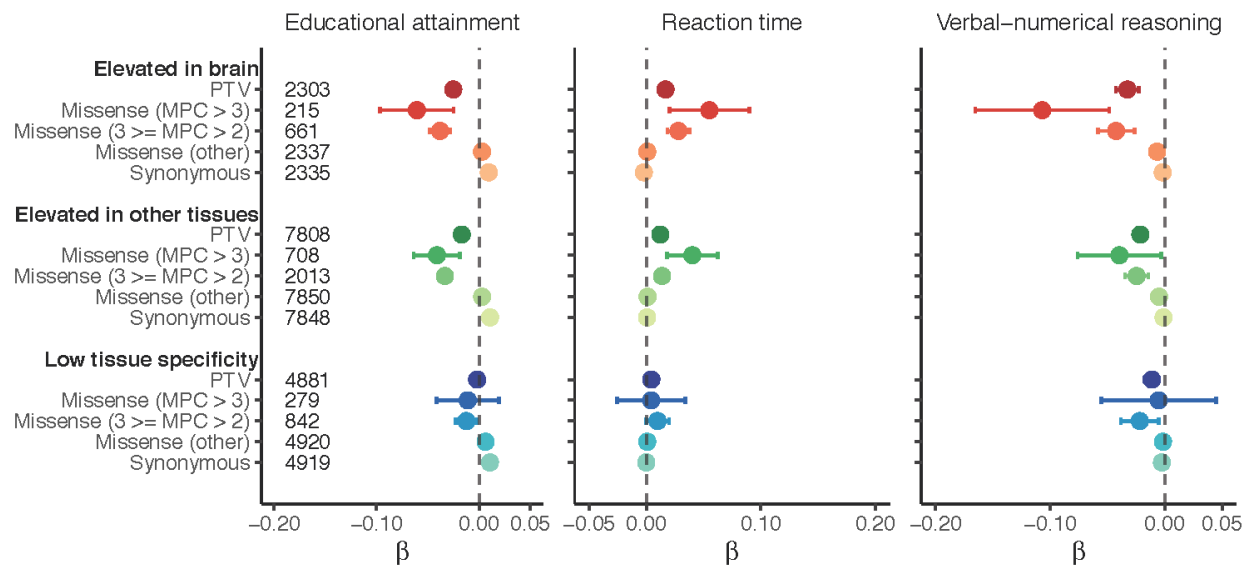


Fig. S17. The effects of protein-truncating, missense (stratified by MPC) and synonymous variant burden in genes stratified by brain-specific expression. Genes were stratified by elevated expression in brain tissue (2,587 genes), elevated expression in other tissues but also expressed in brain (5,298 genes) and no tissue specific expression (8,342 genes). Number of genes included in the burden is annotated for each set.

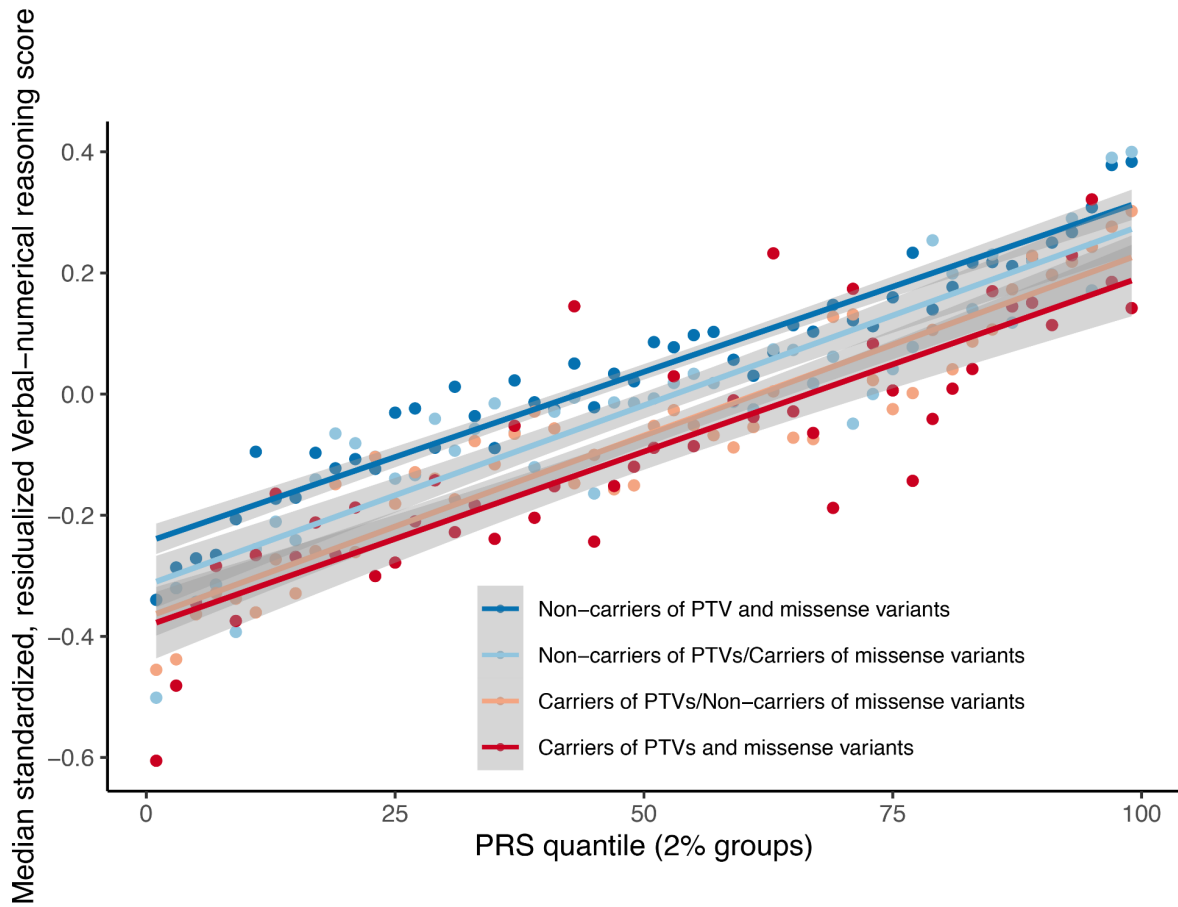


Fig. S18. The impact of cognitive function polygenic score and carrier status of PTV and/or damaging missense variants ($MPC > 2$) in LoF intolerant genes ($pLI > 0.9$) on verbal-numerical reasoning.

Reference

1. Watkins, L. R. & Orlandi, C. Orphan G Protein Coupled Receptors in Affective Disorders. *Genes* **11**, (2020).
2. Kee, H. J. *et al.* Expression of brain-specific angiogenesis inhibitor 2 (BAI2) in normal and ischemic brain: involvement of BAI2 in the ischemia-induced brain angiogenesis. *J. Cereb. Blood Flow Metab.* **22**, 1054–1067 (2002).
3. Fagerberg, L. *et al.* Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell. Proteomics* **13**, 397–406 (2014).
4. Lee, J. J. *et al.* Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nat. Genet.* **50**, 1112–1121 (2018).
5. Sakaue, S. *et al.* A cross-population atlas of genetic associations for 220 human phenotypes. *Nat. Genet.* **53**, 1415–1424 (2021).
6. Khawaja, A. P. *et al.* Genome-wide analyses identify 68 new loci associated with intraocular pressure and improve risk prediction for primary open-angle glaucoma. *Nat. Genet.* **50**, 778–782 (2018).
7. Wright, K. M. *et al.* A Prospective Analysis of Genetic Variants Associated with Human Lifespan. *G3* **9**, 2863–2878 (2019).
8. Vallianatos, C. N. & Iwase, S. Disrupted intricacy of histone H3K4 methylation in neurodevelopmental disorders. *Epigenomics* **7**, 503–519 (2015).
9. Han, M., Xu, W., Cheng, P., Jin, H. & Wang, X. Histone demethylase lysine demethylase 5B in development and cancer. *Oncotarget* **8**, 8980–8991 (2017).
10. Xhabija, B. & Kidder, B. L. KDM5B is a master regulator of the H3K4-methylome in stem cells, development and cancer. *Semin. Cancer Biol.* **57**, 79–85 (2019).
11. Najmabadi, H. *et al.* Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature* **478**, 57–63 (2011).

12. Takata, A. *et al.* Loss-of-function variants in schizophrenia risk and SETD1A as a candidate susceptibility gene. *Neuron* **82**, 773–780 (2014).
13. De Rubeis, S. *et al.* Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* **515**, 209–215 (2014).
14. Iossifov, I. *et al.* The contribution of de novo coding mutations to autism spectrum disorder. *Nature* **515**, 216–221 (2014).
15. Firth, J. *et al.* Grip Strength Is Associated With Cognitive Performance in Schizophrenia and the General Population: A UK Biobank Study of 476559 Participants. *Schizophr. Bull.* **44**, 728–736 (2018).
16. Giovannone, B. *et al.* Two novel proteins that are linked to insulin-like growth factor (IGF-I) receptors by the Grb10 adapter and modulate IGF-I signaling. *J. Biol. Chem.* **278**, 31564–31573 (2003).
17. Dufresne, A. M. & Smith, R. J. The adapter protein GRB10 is an endogenous negative regulator of insulin-like growth factor signaling. *Endocrinology* **146**, 4399–4409 (2005).
18. Astle, W. J. *et al.* The Allelic Landscape of Human Blood Cell Trait Variation and Links to Common Complex Disease. *Cell* **167**, 1415–1429.e19 (2016).
19. Sinnott-Armstrong, N. *et al.* Genetics of 35 blood and urine biomarkers in the UK Biobank. *Nat. Genet.* **53**, 185–194 (2021).
20. Zhao, Y. *et al.* GIGYF1 loss of function is associated with clonal mosaicism and adverse metabolic health. *Nat. Commun.* **12**, 4178 (2021).
21. Deaton, A. M. *et al.* Gene-level analysis of rare variants in 379,066 whole exome sequences identifies an association of GIGYF1 loss of function with type 2 diabetes. *Sci. Rep.* **11**, 21565 (2021).
22. Li, C.-W., Dinh, G. K., Zhang, A. & Chen, J. D. Ankyrin repeats-containing cofactors interact with ADA3 and modulate its co-activator function. *Biochem. J* **413**, 349–357 (2008).
23. Zhang, A. *et al.* Identification of a novel family of ankyrin repeats containing cofactors for p160

- nuclear receptor coactivators. *J. Biol. Chem.* **279**, 33799–33805 (2004).
24. Chen, M.-H. *et al.* Trans-ethnic and Ancestry-Specific Blood-Cell Genetics in 746,667 Individuals from 5 Global Populations. *Cell* **182**, 1198–1213.e14 (2020).
 25. Shieh, B. H. *et al.* Mapping of the gene for the cardiac sarcolemmal Na⁽⁺⁾-Ca²⁺ exchanger to human chromosome 2p21-p23. *Genomics* **12**, 616–617 (1992).
 26. Morris, J. A. *et al.* An atlas of genetic influences on osteoporosis in humans and mice. *Nat. Genet.* **51**, 258–266 (2019).
 27. Giri, A. *et al.* Trans-ethnic association study of blood pressure determinants in over 750,000 individuals. *Nat. Genet.* **51**, 51–62 (2019).
 28. Ntalla, I. *et al.* Multi-ancestry GWAS of the electrocardiographic PR interval identifies 202 loci underlying cardiac conduction. *Nat. Commun.* **11**, 2542 (2020).
 29. Arking, D. E. *et al.* Genetic association study of QT interval highlights role for calcium signaling pathways in myocardial repolarization. *Nat. Genet.* **46**, 826–836 (2014).
 30. Willems, S. M. *et al.* Large-scale GWAS identifies multiple loci for hand grip strength providing biological insights into muscular fitness. *Nat. Commun.* **8**, 16015 (2017).
 31. Siess, D. C. *et al.* A human gene coding for a membrane-associated nucleic acid-binding protein. *J. Biol. Chem.* **275**, 33655–33662 (2000).
 32. Vinuesa, C. G. *et al.* A RING-type ubiquitin ligase family member required to repress follicular helper T cells and autoimmunity. *Nature* **435**, 452–458 (2005).
 33. Zhang, Q. *et al.* New Insights into the RNA-Binding and E3 Ubiquitin Ligase Activities of Roquins. *Sci. Rep.* **5**, 15660 (2015).
 34. Heissmeyer, V. & Vogel, K. U. Molecular control of Tfh-cell differentiation by Roquin family proteins. *Immunol. Rev.* **253**, 273–289 (2013).
 35. Vogel, K. U. *et al.* Roquin paralogs 1 and 2 redundantly repress the Icos and O_x40 costimulator mRNAs and control follicular helper T cell differentiation. *Immunity* **38**, 655–668 (2013).
 36. Jansen, P. R. *et al.* Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk

- loci and functional pathways. *Nat. Genet.* **51**, 394–403 (2019).
37. Baselmans, B. M. L. *et al.* Multivariate genome-wide analyses of the well-being spectrum. *Nat. Genet.* **51**, 445–451 (2019).
 38. Mills, M. C. *et al.* Identification of 371 genetic variants for age at first sex and birth linked to externalising behaviour. *Nat Hum Behav* **5**, 1717–1730 (2021).
 39. Zhao, B. *et al.* Genome-wide association analysis of 19,629 individuals identifies variants influencing regional brain volumes and refines their genetic co-architecture with cognitive and mental health traits. *Nat. Genet.* **51**, 1637–1644 (2019).
 40. Siva, K., Venu, P., Mahadevan, A., S K, S. & Inamdar, M. S. Human BCAS3 expression in embryonic stem cells and vascular precursors suggests a role in human embryogenesis and tumor angiogenesis. *PLoS One* **2**, e1202 (2007).
 41. Hengel, H. *et al.* Bi-allelic loss-of-function variants in BCAS3 cause a syndromic neurodevelopmental disorder. *Am. J. Hum. Genet.* **108**, 1069–1082 (2021).
 42. Wuttke, M. *et al.* A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nat. Genet.* **51**, 957–972 (2019).
 43. Tin, A. *et al.* Target genes, variants, tissues and transcriptional pathways influencing human serum urate levels. *Nat. Genet.* **51**, 1459–1474 (2019).
 44. Gao, X. R., Huang, H. & Kim, H. Genome-wide association analyses identify 139 loci associated with macular thickness in the UK Biobank cohort. *Hum. Mol. Genet.* **28**, 1162–1172 (2019).
 45. Koyama, S. *et al.* Population-specific and trans-ancestry genome-wide analyses identify distinct and shared genetic risk loci for coronary artery disease. *Nat. Genet.* **52**, 1169–1177 (2020).
 46. Kaplanis, J. *et al.* Evidence for 28 genetic disorders discovered by combining healthcare and research data. *Nature* **586**, 757–762 (2020).
 47. Kang, H. J. *et al.* Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483–489 (2011).
 48. Lam, M. *et al.* Identifying nootropic drug targets via large-scale cognitive GWAS and transcriptomics. *Neuropsychopharmacology* **46**, 1788–1801 (2021).

