HGG Advances

ARTICLE

Exome-wide analysis reveals role of *LRP1* and additional novel loci in cognition

Shreya Chakraborty[1,2] and Bratati Kahali[1,3,*]

Summary

Cognitive functioning is heritable, with metabolic risk factors known to accelerate age-associated cognitive decline. Identifying genetic underpinnings of cognition is thus crucial. Here, we undertake single-variant and gene-based association analyses upon 6 neurocognitive phenotypes across 6 cognition domains in whole-exome sequencing data from 157,160 individuals of the UK Biobank cohort to expound the genetic architecture of human cognition. We report 20 independent loci associated with 5 cognitive domains while controlling for *APOE* isoform-carrier status and metabolic risk factors; 18 of which were not previously reported, and implicated genes relating to oxidative stress, synaptic plasticity and connectivity, and neuroinflammation. A subset of significant hits for cognition indicates mediating effects via metabolic traits. Some of these variants also exhibit pleiotropic effects on metabolic traits. We further identify previously unknown interactions of *APOE* variants with *LRP1* (rs34949484 and others, suggestively significant), *AMIGO1* (rs146766120; pAla25Thr, significant), and *ITPR3* (rs111522866, significant), controlling for lipid and glycemic risks. Our gene-based analysis also suggests that *APOC1* and *LRP1* have plausible roles along shared pathways of amyloid beta (Aβ) and lipid and/or glucose metabolism in affecting complex processing speed and visual attention. In addition, we report pairwise suggestive interactions of variants harbored in these genes with *APOE* affecting visual attention. Our report based on this large-scale exome-wide study highlights the effects of neuronal genes, such as *LRP1*, *AMIGO1*, and other genomic loci, thus providing further evidence of the genetic underpinnings for cognition during aging.

Introduction

Cognition refers to a plethora of mental processes that guide acquisition, transformation, storage, recovery, and implementation of information, and is key to good health. Understanding genetic predispositions for inter-individual differences in age-related cognitive decline is of paramount importance in healthy aging. Genome-wide studies on cognition have shown that intelligence in humans is heritable and that individual differences can be explained by genetic variations.[1–5] Previous GWAS on cognitive functioning have yielded significant positions in the genome affecting various domains of cognition, such as episodic memory, processing speed, reaction time, fluid intelligence, and general intelligence.[5–10] Non-invasive neuropsychological cognitive assessments serve as dependable endophenotypes to assess brain functioning in healthy aging and dementia.[11,12]

The *APOE* (MIM: 107741) locus confers the highest genetic risk for Alzheimer dementia and is also known to be associated with nonpathological cognitive aging.[13] ApoE is the major apolipoprotein that plays a central role in maintaining homeostasis in the brain via transport and clearance of lipids and amyloid beta (Aβ). Several other age-associated metabolic disorders, namely, obesity, type 2 diabetes, dyslipidemia, and cardiovascular disease, can act as modifiable risk factors for cognitive impairment.[14] Interplay among ApoE, lipid homeostasis, brain glucose, and Aβ trafficking in animal models of Alzheimer disease has been reported.[15]

In this study, we decipher the genetic underpinnings of cognitive functioning while considering the effects of putative interrelations with metabolic risk factors in the UK Biobank. We also identify variants in crucial genes that work in conjunction and interact with *APOE* in influencing cognitive functioning at a granularity of specific cognitive domains in the presence of lipid and glycemic metabolic risk factors.

Material and methods

Samples and participants

We present our analysis based on whole exomes of 200,643 individuals enrolled in the UK Biobank (approved project ID 55652).[16]

Phenotypes

We consider six cognitive domains of simple processing speed, episodic memory, fluid intelligence, working memory, visual attention, and complex processing speed corresponding to which "Reaction time," "Pairs," "Reasoning," "Digit recall," "Trail making," and "Digit-symbol substitution" cognitive tests were administered on the UK Biobank participants

[1]Centre for Brain Research, Indian Institute of Science, Bangalore, Karnataka 560012, India; [2]Interdisciplinary Mathematical Sciences, Indian Institute of Science, Bangalore, Karnataka 560012, India
[3]Lead contact
*Correspondence: bratati@iisc.ac.in
https://doi.org/10.1016/j.xhgg.2023.100208.

(web resources: id=8481) (further details in supplemental information). All the cognition phenotypes that we have selected have been measured on a continuous scale, and these fields had maximum available data points across all individuals with reported sample size >100,000 for each test, and also recommended by UK Biobank as the primary item of interest for each cognitive test. Table S1 (rows tagged as "Available") shows the mean and standard deviation of the test scores pertaining to each cognitive test for the number of individuals for whom information on at least the cognitive test score was available.

As part of covariates, we also consider age in years (quantitative), genetic sex (binary–male or female), educational qualification (categorical: based on (1) college or university degree, (2) A levels/AS levels/equivalent, (3) either of O levels/general or secondary certificate/higher national diploma or certificate/national vocational qualification/equivalent) (further details in supplemental information). We also use serum lipid (HDL, LDL, TG, TC) and serum glucose levels quantified using clinical chemistry analyser Beckman Coulter AU5800 (web resources: serum_biochemistry.pdf) in the UK Biobank assessments as covariates. We also use serum HbA1c levels obtained using five Bio-Rad Variant II Turbo analyzer assay as glycemic covariates (web resources: serum_HbA1c.pdf).

Genetic data and quality control

We download the UK Biobank population-level exome OQFE files for ∼200k exomes in pVCF format (Field ID: 23156) using the "gfetch" utility. After extensive quality checks (details in supplemental information), we retain 157,160 individuals with 211,012 variants (Table S2). There are 71,566 males and 85,461 females among these individuals with ages ranging from 38 to 72 years. These individuals are predominantly White British (91%) (Table S3) and their educational attainment levels with the categories considered are provided in Table S4.

Heritability

Before proceeding to association analyses, we assess the heritability of the six cognition phenotypes based on unrelated individuals using the LDAK[17] model (supplemental information). Our heritability estimates (Table S5) show good concordance with evidence from previous family-based studies and GWAS ATLAS resource.[18]

APOE-carrier status determination

Out of 157,160 samples, 93 have missing genotype information for *APOE* at either rs7412 or rs429358 or both. We determine *APOE*-carrier status, by flagging samples with at least one copy of $\epsilon 4$ allele as risk, with at least one copy of $\epsilon 2$ as protective/beneficial, $\epsilon 1/\epsilon 3$ and $\epsilon 3/\epsilon 3$ carriers as neutral, to include as a categorical covariate in association models (Table S6).

Statistical analyses

Single-variant association

With genetic data on the resultant 157,067 samples and 211,012 variants, we perform single-variant Wald test using rvtests.[19] Our baseline model controls for age, gender, educational qualification, top 10 principal components, and *APOE*-carrier status (functional form of the model and other details are provided in supplemental information). Furthermore, to control for age-related metabolic conditions that can adversely affect cognition, we add lipid levels (serum total cholesterol, HDL and LDL direct cholesterol, triglycerides), glucose, and HbA1c levels separately as covariates to the baseline model in models 2 and 3 (supplemental information). We obtain the residuals and test the inverse-normalized residuals against genotypes of each variant (supplemental information). We obtain Manhattan plots (Figures S1–S5) and QQ plots (Figures S5 and S6) to visualize our results and significant hits. We use the LDtrait module of LDlink[20] to check if variants, which are in high LD ($r^2 > 0.8$) and fall within ±500 kbp with our significant variant, were previously associated with any trait listed in the EBI-GWAS catalog.[21]

Coding region-specific analysis

The exome sequencing data dispensed to us by the UK Biobank does not have the variants functionally annotated. The sequencing design was targeted toward 39 Mbp of the human genome and included variants in target regions and 100 bp flanking regions upstream and downstream of each capture target (web resources: UK Biobank exome release). Since it is well established that non-coding regions in the genome harbor variants associated with traits, and also regulate gene expression, control mRNA transport and assembly, and have several other functional roles,[22,23] in this paper we highlight the results that we obtain from our analyses including all variants (211,012) in the exome sequencing data in the main text of this article, even though they are in introns and UTRs. Thus, we obtain the annotations (Table S7) and filter coding region-specific variants (filtering criteria, methodological details pertaining to coding region-specific analysis in supplemental information) and perform separate single-variant associations, and subsequent analysis pertaining to 83,673 variants in the coding region only (Tables S8–S13). We do not repeat the gene-based tests for the coding variants separately because rvtests uses annotation information curated from several bioinformatic databases to specify a gene unit that contains only coding regions.

Gene-based association

We performed gene-based association tests in 157,067 individuals with kernel-based (SKAT)[24] and unified kernel- and burden-based methods (SKAT-O[25]) to detect the cumulative burden in genes that work concomitantly with *APOE* (functional form of the models are provided in supplemental information). Figure 1 represents the possible pathways in which *APOE* affects neuronal dysfunction and hence cognition. Consideration of genes along these pathways ensures capturing the genetic basis of cognition in association with lipids homeostasis, glucose metabolism, and Aβ pathogenesis that plausibly play vital roles in modulating cognition with age.

As an ancillary analysis, we also carry out whole-exome-wide gene-based analysis. We test 16,915 genes using the same gene-based association setup (supplemental information). We also carry out enrichment analysis of the whole-exome-wide significant genes using the KEGG pathway database[26] integrated in the ShinyGo tool,[27] and filter pathways with a fold enrichment value >10, false discovery rate <0.05, and having at least three genes in the gene set. We compare our exome-wide gene-based hits with the EBI-GWAS catalog "MAPPED_GENE" column to find which of our hits had been previously mapped with cognition traits. This column contains gene(s) mapped to the strongest SNP identified from GWAS. If the SNP is located within a gene (or overlapping genes), that gene(s) is listed. If the SNP is intergenic, the upstream and downstream genes are listed, separated by a hyphen. We search the "DISEASE/TRAIT" column of the catalog for terms containing "cognition," "cognitive performance," "intelligence," "Alzheimer disease," "dementia," and investigate if the mapped gene from these published single-variant associations contain

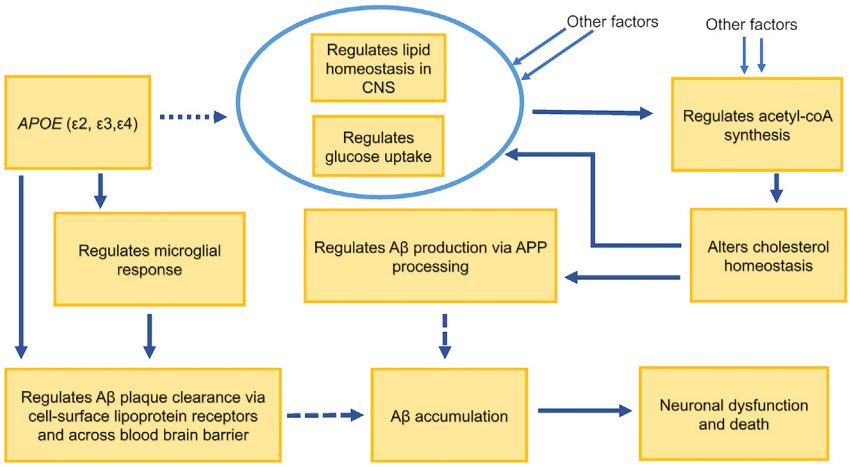

Figure 1. Putative pathways in which APOE isoforms regulate neuronal dysfunction

any of our exome-wide significant genes, and thereby assess the novelty of our exome-wide gene-based hits. However, we do not consider the genes for the traits which reported interaction or pleiotropy. We note that this comparison is based on a mapping of the strongest associated SNP in the region and is not truly a gene-gene comparison.

Pairwise epistasis

We further uncover the interactions of the significant loci discovered by our single-variant association tests through pairwise epistasis analysis (using plink-1.9.0 software)[28] with each of the two *APOE* isoform-defining SNPs (rs429358 and rs7412) (functional form of the model and other details are provided in supplemental information). We control for all covariates used in models 2 and 3 except *APOE*-carrier status and obtain the inverse-normalized residuals before testing for interaction effects.

Next, we conduct a pairwise epistasis test for the variants in significant gene hits from either the SKAT or SKAT-O models ($p < 0.0025$), with *APOE* isoform-defining rs429358 and rs7412, as well as all *APOE* SNPs in two different models (supplemental information).

Bivariate association tests

For each of the significant variants from our single-variant association tests, we conduct bivariate association tests for cognitive measures and lipid levels/glycemic traits, respectively, with the summary statistics obtained from the single-variant Wald test results using metaMANOVA and metaUSAT[29] (functional forms of the models are provided in supplemental information).

Annotation and tissue expression analysis

We annotate our exome-wide significant hits by mapping them to nearest genes (NCBI, dbSNP; UCSC) and calculating their deleteriousness using CADD[30] scores, where higher scores indicate more deleteriousness. We also assess if our identified variants are rare (minor allele frequency [MAF] < 1%), low frequency (MAF between 1% and 5% both inclusive), or common (MAF > 5%). We perform functional annotation of the variants uncovered using SnpEff[31] and calculate LofTool[32] scores. The lower the LoFtool score, the more intolerant is the mapped gene to functional changes. Also, we investigate, using GTEx,[33] if our variants are eQTL loci or lie near eQTL loci or are in linkage disequilibrium (pairwise LD obtained using 1000G EUR) with eQTL loci that significantly regulate expression of the respectively annotated genes in brain regions. We also use RegulomeDB[34] to identify the putative regulatory potential of our association and interaction signals, which provide supporting evidence about how likely our variants are to be functional.

Results

Identification of exome-wide significant variants for cognitive domains

From our exome-wide single-variant association analysis on 211,012 variants in 157,067 individuals, we identify 20 independent loci associated with 5 different domains of cognition (Table 1) with and without controlling for metabolic risks.

Fluid intelligence

We identify a rare variant rs115865641 in *PCDHB16* (MIM: 606345) (3′ UTR) associated with fluid intelligence (Table 1). Controlling for HDL and glucose separately, we obtain two independent significant hits—rs17876162 in *PON2* (MIM: 602447) (intronic) and rs3824734 (synonymous) in *CPEB3* (MIM: 610606) (Table 1). Our synonymous rs3824734 (*CPEB3*) with a CADD score of 11.74 (Table 1), implying that it is predicted to be among 7% of the most deleterious substitutions to the genome, can be crucial for pinpointing the role of this gene in cognition. This *CPEB3* variant is also obtained as a significant signal from the same models from our coding region-specific analysis (Table S8).

Simple processing speed

We detect rs3813363 (5′ UTR of *SAMD3*), and rs17662853 (missense variant in *KANSL1* [MIM: 612452]) (CADD score of 23.9) (Table 1) to be associated with simple processing speed in the baseline as well as all models controlling for lipid and glycemic traits. rs3813363 is within 500 kbp and in high LD ($r^2 > 0.8$) of both rs11154580 and rs6937866, known to be associated with reaction time[5] (Figure 2). Highly deleterious rs17662853 is in high LD ($r^2 = 0.856$) with intronic rs10775404 (CADD score of 1.782) previously associated with reaction time (Figure 2),[5] highlighting that our missense variant can be more impactful, and more likely to be causal. rs17662853 is also obtained as a significant hit from our coding region-specific analysis from all models (Table S8). We also identify two variants—rs73922480 and rs77285514 (synonymous and intronic *GPR108* [MIM: 618491], respectively, 80 bp apart) to be associated with mean reaction time in the baseline model as well as controlling for HbA1c (Table 1). rs73922480 in *GPR108* is also obtained as a significant hit from the baseline model as well as models controlling from HDL and glucose, respectively, from our coding region-specific analysis (Table S8). Controlling for HbA1c, we additionally identify rs201404149 (synonymous

Table 1. Single-variant association analysis

Phenotype/domain:	Fluid intelligence score/reasoning, fluid intelligence cognitive domain								
Models	Chr:Position; rsid – REF/ALT (ALT = effect allele); effect allele frequency	Informative samples	Effect size: β (effect allele)	Standard error (SE)	p value	Variance explained (%)	Previously known for cognition	Genomic annotation; mapped gene	CADD score (Phred scaled); LofTool score
Baseline[a]	chr5:141185287; rs115865641 – **G**/A; 0.009	36,547	0.036	0.006	7.33E−08	0.002	no	3′ UTR variant; *PCDHB16*	3.05; 0.71
Baseline + HDL	chr7:95406923; rs17876162 – **A**/G; 0.001	31,885	0.035	0.007	6.04E−08	3.16E−04	no	intron variant; *PON2*	5.37; 0.82
	chr10:92240093; rs3824734 – A/**G**; 0.594	31,885	−0.042	0.008	1.48E−07	0.086	no	synonymous variant; *CPEB3*	11.74; 0.06
Baseline + Glucose	chr7:95406923; rs17876162 – **A**/G; 0.001	31,852	0.036	0.007	4.75E−08	3.20E−04	no	intron variant; *PON2*	5.37; 0.82
	chr10:92240093; rs3824734 – A/**G**; 0.594	31,852	−0.043	0.008	1.13E−07	0.087	no	synonymous variant; *CPEB3*	11.74; 0.06
Phenotype/Domain:	Mean time to identify matches/simple processing speed cognitive domain								
Models	Chr:Position; rsid – REF/ALT (ALT = effect allele); effect allele frequency	Informative samples	Effect size: β (effect allele)	Standard error (SE)	p value	Variance explained (%)	Previously known for cognition	Genomic annotation; mapped gene	CADD score (Phred scaled); LofTool score
Baseline	chr6:130365270; rs3813363 - **C**/T; 0.324	121,127	−0.024	0.004	5.91E−08	0.025	yes	5′ UTR variant; *SAMD3*	7.30; 0.94
	chr17:46171482; rs17662853 - **G**/A; 0.149	121,127	−0.025	0.005	1.04E−07	0.016	yes	missense variant (p.Thr221Ile); *KANSL1*	23.90; –
	chr19:6732114; rs73922480 - **C**/T; 0.001	121,127	−0.584	0.106	3.66E−08	0.097	no	synonymous variant; *GPR108*	4.16; 0.76
	chr19:6732194; rs77285514 - **C**/T; 0.002	121,127	−0.448	0.085	1.45E−07	0.066	no	intron variant; *GPR108*	0.59; 0.76

(*Continued on next page*)

Table 1. ***Continued***

Phenotype/Domain:	Mean time to identify matches/simple processing speed cognitive domain								
Models	Chr:Position; rsid – REF/ALT (ALT = effect allele); effect allele frequency	Informative samples	Effect size: β (effect allele)	Standard error (SE)	p value	Variance explained (%)	Previously known for cognition	Genomic annotation; mapped gene	CADD score (Phred scaled); LofTool score
Baseline + HDL	chr6:130365270; rs3813363 - **C**/T; 0.324	105,977	−0.025	0.005	1.59E−07	0.026	yes	5′ UTR variant; *SAMD3*	7.30; 0.94
	chr12:440788; rs11062991 - **G**/T; 0.011	105,977	−0.109	0.021	1.78E−07	0.026	no	intron variant; *CCDC77*	4.22; 1.00
	chr15:70891939; rs2959174 - **G**/T; 0.407	105,977	−0.022	0.004	1.96E−07	0.024	no	intron/synonymous variant; *LRRC49/THAP10*	2.75; 0.48
	chr17:46171482; rs17662853 - **G**/A; 0.149	105,977	−0.027	0.005	1.22E−07	0.018	yes	missense variant (p.Thr221Ile); *KANSL1*	23.90; –
Baseline + LDL	chr6:130365270; rs3813363 - **C**/T; 0.324	115,316	−0.025	0.004	3.89E−08	0.027	yes	5′ UTR variant; *SAMD3*	7.30; 0.94
	chr17:46171482; rs17662853 - **G**/A; 0.149	115,316	−0.026	0.005	9.70E−08	0.017	yes	missense variant (p.Thr221Ile); *KANSL1*	23.90; –
Baseline + TC	chr6:130365270; rs3813363 - **C**/T; 0.324	115,520	−0.024	0.004	4.79E−08	0.026	yes	5′ UTR variant; *SAMD3*	7.30; 0.94
	chr15:70891939; rs2959174. - **G**/T; 0.407	115,520	−0.022	0.004	8.61E−08	0.024	no	intron/synonymous variant; *LRRC49/THAP10*	2.75; 0.48
	chr17:46171482; rs17662853 - **G**/A; 0.149	115,520	−0.028	0.005	1.17E−08	0.019	yes	missense variant (p.Thr221Ile); *KANSL1*	23.90; –
Baseline + TG	chr6:130365270; rs3813363 - **C**/T; 0.324	115,425	−0.024	0.004	6.03E−08	0.026	yes	5′ UTR variant; *SAMD3*	7.30; 0.94
	chr17:46171482; rs17662853 - **G**/A; 0.149	115,425	−0.027	0.005	2.83E−08	0.018	yes	missense variant (p.Thr221Ile); *KANSL1*	23.90; –
Baseline + Glucose	chr6:130365270; rs3813363 - **C**/T; 0.324	105,898	−0.025	0.005	8.10E−08	0.028	yes	5′ UTR variant; *SAMD3*	7.30; 0.94
	chr12:440788; rs11062991 - **G**/T; 0.011	105,898	−0.110	0.021	1.46E−07	0.027	no	intron variant; *CCDC77*	4.22; 1.00
	chr17:46171482; rs17662853 - **G**/A; 0.149	105,898	−0.028	0.005	4.27E−08	0.019	yes	missense variant (p.Thr221Ile); *KANSL1*	23.90; –

(*Continued on next page*)

Table 1. ***Continued***

Phenotype/Domain:	Mean time to identify matches/simple processing speed cognitive domain								
Models	Chr:Position; rsid – REF/ALT (ALT = effect allele); effect allele frequency	Informative samples	Effect size: β (effect allele)	Standard error (SE)	p value	Variance explained (%)	Previously known for cognition	Genomic annotation; mapped gene	CADD score (Phred scaled); LofTool score
Baseline + HbA1c	chr1:25826774; rs201404149 - **C**/T; 0.001	115,563	−0.311	0.060	1.93E−07	0.027	no	synonymous variant; *MTFR1L*	9.64; –
	chr6:130365270; rs3813363 - **C**/T; 0.324	115,563	−0.025	0.004	4.50E−08	0.026	yes	5′ UTR variant; *SAMD3*	7.30; 0.94
	chr15:70832754; rs3825970 - G/**A**; 0.586	115,563	0.022	0.004	1.64E−07	0.024	no	synonymous variant; *LARP6*	0.895; 0.18
	chr15:70832865; rs1549317 - A/**G**; 0.589	115,563	0.023	0.004	9.71E−08	0.025	no	synonymous variant; *LARP6*	6.676; 0.18
	chr17:46171482; rs17662853 - **G**/A; 0.149	115,563	−0.025	0.005	1.64E−07	0.016	yes	missense variant (p.Thr221Ile); *KANSL1*	23.90; –
	chr19:6732114; rs73922480 - **C**/T; 0.001	115,563	−0.595	0.112	1.03E−07	0.101	no	synonymous variant; *GPR108*	4.16; 0.76
	chr19:6732194; rs77285514 - **C**/T; 0.002	115,563	−0.476	0.089	1.07E−07	0.074	no	intron variant; *GPR108*	0.59; 0.76
Phenotype/Domain:	Maximum correct symbol-digit substitutions/complex processing speed cognitive domain								
Models	Chr:Position; rsid – REF/ALT (ALT = effect allele); effect allele frequency	Informative samples	Effect size: β (effect allele)	Standard error (SE)	p value	Variance explained (%)	Previously known for cognition	Genomic annotation; mapped gene	CADD score (Phred scaled); LofTool score
Baseline	chr16:27462539; rs12932325 – G/**A**; 0.142	35,174	−0.040	0.007	8.56E−08	0.039	no	intron variant; *GTF3C1*	0.88; 0.28
Baseline + LDL	chr12:112482065; rs12301915-**C**/A; 0.013	33,433	0.039	0.007	8.69E−08	0.004	no	intron variant; *PTPN11*	14.34; 0.05
Baseline + TC	chr12:112482065; rs12301915-**C**/A; 0.013	33,495	0.039	0.007	1.12E−07	0.004	no	intron variant; *PTPN11*	14.34; 0.05
Baseline + HbA1c	chr11:70324575; rs71467481- **G**/A; 0.021	33,503	0.030	0.006	1.15E−07	0.004	no	intron variant; *PPFIA1*	0.09; 0.23
Phenotype/Domain:	Proportion of incorrect pair matches/episodic memory domain								
Models	Chr:Position; rsid – REF/ALT (ALT = effect allele); effect allele frequency	Informative samples	Effect size: β (effect allele)	Standard error (SE)	p value	Variance explained (%)	Previously known for cognition	Genomic annotation; mapped gene	CADD score (Phred scaled); LofTool score
Baseline[a]	chr1:109508840; rs146766120-**C**/T; 0.001	34,120	−1.080	0.180	3.93E−08	0.319	no	missense variant (p.Ala25Thr); *AMIGO1*	15.97; –
	chr2:130356125; rs77807661 - **C**/T; 0.002	34,120	−2.157	0.378	1.82E−07	1.711	no	synonymous variant; *PTPN18*	6.70; 0.41
	chr6:33688801; rs111522866-**C**/T; 0.002	34,120	−0.882	0.147	4.62E−08	0.243	no	intron variant; *ITPR3*	0.06; 0.05

(*Continued on next page*)

Table 1. ***Continued***

Phenotype/Domain:	Proportion of incorrect pair matches/episodic memory domain								
Models	Chr:Position; rsid – REF/ALT (ALT = effect allele); effect allele frequency	Informative samples	Effect size: β (effect allele)	Standard error (SE)	p value	Variance explained (%)	Previously known for cognition	Genomic annotation; mapped gene	CADD score (Phred scaled); LofTool score
Baseline + LDL[a]	chr5:90502412; rs7725495 – **G**/A; 0.002	32,423	−0.034	0.006	1.86E−07	3.74E−04	no	intron variant; *POLR3G*	0.59; 0.53
Baseline + HbA1c[a]	chr1:109508840; rs146766120-**C**/T; 0.001	32,494	−1.091	0.186	7.80E−08	0.326	no	missense variant (p.Ala25Thr); *AMIGO1*	15.97; –
	chr2:130356125; rs77807661-**C**/T; 0.002	32,494	−2.598	0.408	5.81E−09	2.481	no	synonymous variant; *PTPN18*	6.70; 0.41
	chr2:236419198; rs3754644 – **T**/C; 0.001	32,494	−0.971	0.158	1.92E−08	0.235	no	missense variant (p.Gln369Arg); *IQCA1*	7.61; –
	chr6:33688801; rs111522866 – **C**/T; 0.002	32,494	−0.927	0.149	1.30E−08	0.269	no	intron variant; *ITPR3*	0.06; 0.05
	chr16:28904132; rs73529530-**T**/C; 0.003	32,494	−0.044	0.008	1.31E−07	0.001	no	intron variant; *ATP2A1*	0.08; 0.08
Phenotype/Domain:	Duration of alphanumeric trail/visual attention domain								
Models	Chr:Position; rsid – REF/ALT (ALT = effect allele); effect allele frequency	Informative samples	Effect size: β (effect allele)	Standard error (SE)	p value	Variance explained (%)	Previously known for cognition	Genomic annotation; mapped gene	CADD score (Phred scaled); LofTool score
Baseline + HDL[a]	chr1:46001049; rs11589562-**T**/C; 0.419	27,214	−0.050	0.009	7.64E−08	0.122	no	intron variant; *MAST2*	7.84; 0.84

Chr: chromosome; β: effect size of the association; SE: standard error of β.
Combined annotation-dependent depletion (CADD) score evaluates the deleteriousness of the variants. Higher the CADD score, more deleterious is the variant.
Loftool score: a gene intolerance score based on loss of function variants. A lower score indicates more intolerance of the mapped gene to functional variation.
We note that the effect allele given in our single-variant association allele is not always the risk allele for cognitive decline. The direction of association has different implication on cognition for different phenotypes. For example, when the direction of association of the effect allele is negative for average trail duration, mean reaction time and proportion of incorrect matches, the effect allele is actually the protective allele for cognitive decline, whereas a positive association implies that the effect allele is the risk increasing allele for cognitive decline. Similarly, negative association of effect allele with fluid intelligence score as well as with maximum symbol-digit substitutions imply that the effect allele is in essence the risk allele affecting cognition in their corresponding domains and vice versa.
The alleles denoted in bold are the alleles that are affecting cognition adversely, and hence are termed "risk alleles" in the main text. Since all variants are biallelic, risk allele frequency = 1 – effect allele frequency, in cases where risk allele is the non-effect allele.
[a]GC corrected models.

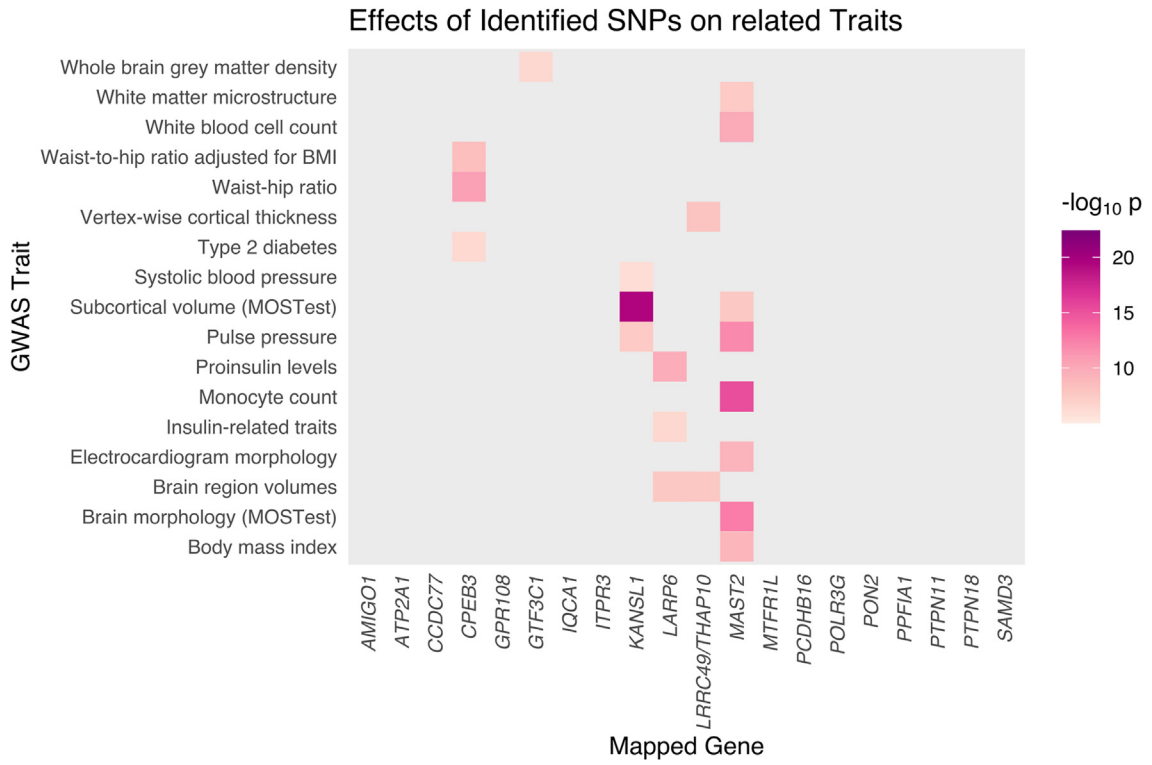

Figure 2. Effects of cognition-associated variants with related metabolic and brain structure traits
Previously known statistically significant effects of our exome-wide significant cognition-associated loci (mapped to nearest genes) on related metabolic and brain structure (obtained from the EBI-GWAS catalog) is highlighted in a pink-purple gradient. Darker color signifies more significant association (lower p value). Gray signifies no significant association.

MTFR1L), rs3825970, and rs1549317 (synonymous *LARP6* [MIM: 611300]) associated with mean reaction time (Table 1). We identify hits rs11062991 (intronic *CCDC77*) and rs2959174 (synonymous *THAP10* [MIM: 612538]) while controlling for serum HDL (Table 1). rs11062991and rs2959174 are also associated with mean reaction time when we control for glucose and total cholesterol, respectively. rs2959174 in *THAP10* is also obtained as a significant hit from the HDL-adjusted model from our coding region-specific analysis (Table S8).

Complex processing speed

We find rs12932325 (intronic *GTF3C1* [MIM: 603246]) associated with complex processing speed from the baseline model. Controlling for LDL and total cholesterol levels independently, we detect rs12301915 (intronic *PTPN11* [MIM: 176876]) (Table 1). Controlling for HbA1c, we identify rs71467481 (intronic *PPFIA1* [MIM: 611054]) as another significant hit (Table 1).

Episodic memory

We identify three variants, rs146766120 (missense *AMIGO1* [MIM: 615689]), rs77807661 (synonymous *PTPN18* [MIM: 606587]), and rs111522866 (intronic *ITPR3* [MIM: 147267]), to be associated with episodic memory in the baseline as well as controlling for serum HbA1c levels (Table 1). The missense *AMIGO1* variant is also yielded as a significant hit from all models except the one controlling for HDL in our coding region-specific analysis (Table S8). rs146766120 (CADD score of 15.97) is among the top ~3% deleterious variants. Similarly, the *PTPN18* variant is also a significant hit from all models in our coding region-specific analysis (Table S8). From the HbA1c-controlled models, we additionally identify rs3754644 (missense *IQCA1*) and rs73529530 (intronic *ATP2A1* [MIM: 108730]) to be associated with episodic memory (Table 1). Missense rs3754644 (*IQCA1*) is also significant from all except the models controlling for HDL and glucose in our coding region-specific analysis (Table S8). Upon controlling for LDL, we find rs7725495 (intronic *POLR3G* [MIM: 617456]) to be associated with episodic memory (Table 1).

Visual attention

We detect rs11589562 (intronic *MAST2* [MIM: 612257]) as associated with visual attention, measured by alphanumeric trail duration, when controlling for HDL levels (Table 1). rs11589562 can significantly control expression of several nearby genes, including *MAST2* in different brain regions.

Genes implicated in cognition from kernel and burden tests

We identify *APOC1* (MIM: 107710) to be significantly associated with complex processing speed and visual attention in baseline model and in models controlling for LDL, total cholesterol, triglycerides, and HbA1c (Table S14). *APOC1* (~5 kb downstream of *APOE*) encodes the smallest of all lipoproteins participating in lipid transport and metabolism and is known to be pleiotropically associated with serum HDL, LDL, triglyceride and cholesterol, and HbA1c

levels.[35] Animal model studies have indicated the role of *APOC1*, expressed in astrocytes and endothelial cells of hippocampus, in cognitive processes in both an *APOE*-dependent and -independent manner.[36] rs4420638 (*APOC1*) has been implicated in general intelligence[6] and CSF biomarker levels.[37] However, we report evidence of *APOC1* influencing two specific cognitive domains through the collective burden of all variants in the gene in a human population that was previously undocumented. We could uncover this effect of *APOC1* after filtering out the more plausible effects of *APOC1* in lipid and glycemic pathways, thereby highlighting the independent role of *APOC1* in cognition and the importance of considering appropriate co-occurring metabolic risks in genetic epidemiological studies.

Controlling for HDL and the baseline covariates, we also identify *LRP1* (MIM: 107770) as a significant gene influencing visual attention (Table S14). A few targeted studies indicate that *LRP1* SNPs and haplotypes influence cognitive performance in Chinese patients with risk of Alzheimer disease.[38,39] This gene encodes the low-density lipoprotein receptor-related protein1, an endocytotic receptor with over 40 ligands including ApoE and Aβ, regulating Aβ uptake and clearance across the blood-brain barrier along with its signaling role in Alzheimer disease pathology.[40] Our results provide evidence from large-scale human whole-exome-based analysis on the role of the elusive *LRP1* in visual attention, which was not known before.

Our ancillary analyses in exome-wide gene-based tests reveal 514 genes associated across six cognition domains (Tables S15–S21). It is worth noting that some of our single-variant hits replicate in this analysis, with respect to mapped gene (further details in supplemental information under section gene-based association test). We find significant inter-domain overlap between the genes uncovered, especially between episodic memory, visual attention, complex processing speed, and fluid intelligence (Table S22), highlighting the integrative role of these domains in general cognition.[41–43] Pathway enrichment analysis carried out with exome-wide significant genes associated with each of the six cognitive domains reveals five pathways (Table S23): (1) long-term depression (*PPP2R1A* [MIM: 605983], *GRIA2* [MIM: 138247], *PRKG2* [MIM: 601591], *PLA2G4F*, *PLA2G4E*), (2) mannose type O-glycan biosynthesis (*POMT2* [MIM: 607439], *POMGNT1* [MIM: 606822], *POMT1* [MIM: 607423]), (3) ras signaling pathway (*GNB1* [MIM: 139380], *MAPK10* [MIM: 602897], *RASGRP1* [MIM: 603962], *PTPN11*), (4) spinocerebellar ataxia (*MAPK10*, *OMA1* [MIM: 617081], *DAB1* [MIM: 603448]), and (5) peptidyl-tyrosine dephosphorylation involved in activation of protein kinase activity (*PTPRH* [MIM: 602510], *PTPRB* [MIM: 176882], *PTPRO* [MIM: 600579]).

In total, we identify 514 genes across 6 cognitive domains and observe that 139 genes for fluid intelligence, 6 for simple processing speed, 8 for working memory, 33 for complex processing speed, 47 for visual attention, and 144 for episodic memory have never been reported to be associated with cognition, intelligence, or Alzheimer disease and related dementia.

Pleiotropy and mediation

Out of the 20 independent loci (Figure 3A), 15 independent loci (*PCDHB16*, *PON2*, *MTFR1L*, *SAMD3*, *LARP6*, *KANSL1*, *GPR108*, *PPFIA1*, *PTPN11*, *AMIGO1*, *PTPN18*, *IQCA1*, *POLR3G*, *ITPR3*, and *ATP2A1*) exhibit pleiotropic effects on lipid and/or glycemic phenotypes (Tables S24–S28). Interestingly, we identify suggestive mediating effects of 4 of these 20 loci on their respective cognitive domains. rs115865641 (*PCDHB16*), associated with fluid intelligence in our baseline model, is also found to be associated with HDL and glucose, but shows effect sizes reduced in magnitude when we control for HDL and glucose levels separately, and is also pleiotropically associated with lipid and glycemic traits (Table S24). This suggests that serum HDL and glucose levels could partially mediate the effect of rs115865641 on fluid intelligence along with its pleiotropic effect. Similar effects are observed for rs201404149 (*MTFR1L*) associated with simple processing speed controlling for HbA1c. rs201404149 is significant from the baseline model, pleiotropically associated with serum glucose levels and this variant shows a reduced effect size on mean reaction time when controlling for serum glucose levels (Table S25), indicating that the effect of this variant on reaction time could be partially mediated through its effect on serum glucose levels, providing further evidence of metabolic risk affecting cognition. Similarly, the *PPFIA1* variant rs71467481 is significant in the baseline model, and is associated with serum glucose levels but shows a reduced effect size compared with baseline when controlling for glucose, implying that the effect of this variant may be mediated through glucose homeostasis in influencing complex processing speed. This variant also shows pleiotropic association with HDL, LDL, and glucose levels (Table S27). rs73529530 in *ATP2A1* shows association with HDL and glucose levels in addition to pleiotropic association with cognition phenotype and all lipid levels and serum glucose levels. rs73529530 may also affect episodic memory by mediation through serum HDL and glucose levels as reflected by the reduction in magnitude of effect size compared with baseline when the phenotype is controlled for HDL and glucose levels (Table S28).

From our analysis restricted to coding region, we observe rs17662853 (*KANSL1*), rs73922480 (*GPR108*), rs146766120 (*AMIGO1*), rs77807661 (*PTPN18*), and rs3754644 (*IQCA1*) to exhibit pleiotropic effects with lipid and glycemic phenotypes as well (Tables S11 and S12). In addition, this analysis shows evidence of rs3824734 (*CPEB3*) exhibiting pleiotropic effect on cognition and both lipid and glycemic phenotypes, respectively (Table S10).

Expression profile analysis

eQTL analysis of significant loci associated with fluid intelligence

We identify rs115865641 (3′ UTR of *PCDHB16*) associated with fluid intelligence scores from the baseline

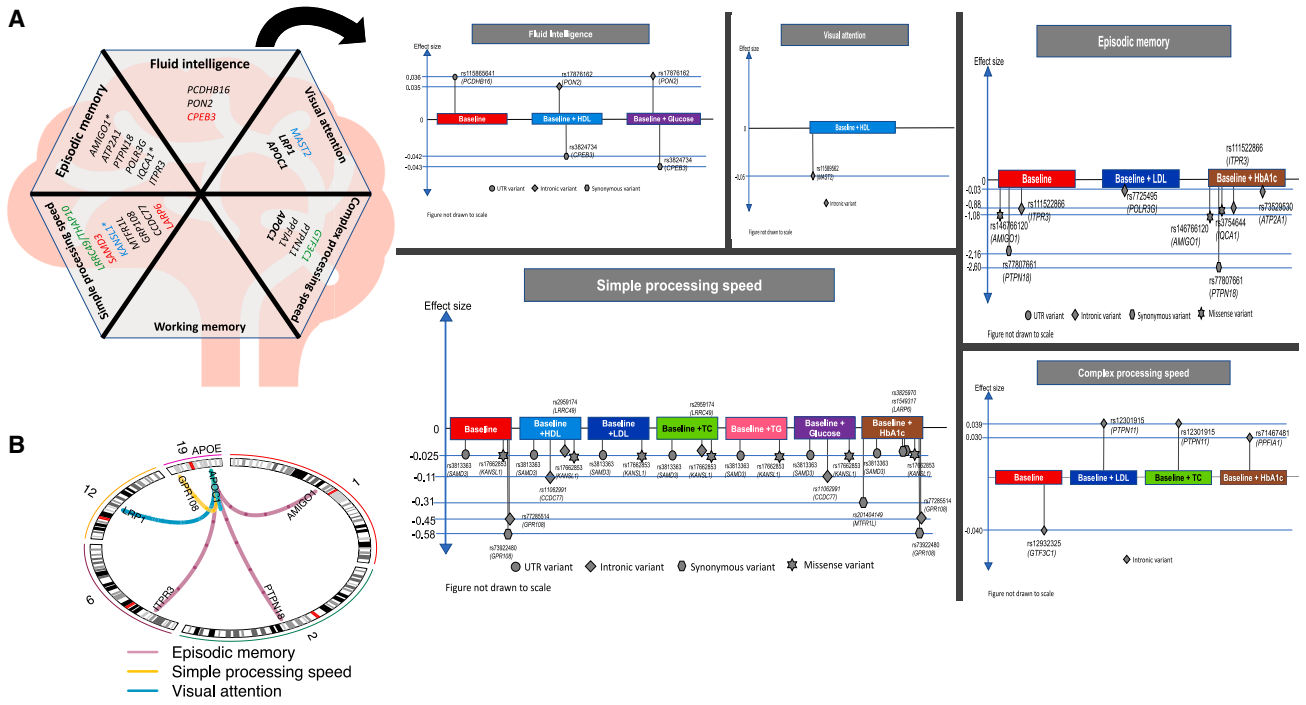

Figure 3. Summary figure showing association hits mapped to nearby genes corresponding to diverse cognition domains and their interactions with *APOE*

(A) Variants and genes we have uncovered associated with diverse cognition domains. The genes to which our single-variant hits have been annotated and the genes identified from gene-based tests (given in bold) are represented here. The variants that are eQTLs for the genes they have been mapped to are represented in red; those that are eQTLs for nearby genes are in green; and those that are eQTLs for both their annotated and nearby genes are represented in light blue. The genes corresponding to variants that are suggestive eQTLs (because of their proximity to eQTL variants) are shown in black. The missense variants are represented with asterisk sign (*) beside their corresponding genes. The figures in the subpanels (indicated by the black arrows) show the effect size of each significant variant uncovered from the different statistical models used for each of the cognition domains. In each subplot, the color-coded rectangles denote the model from which the variants have been shown to be significant. The color coding for different models is consistent with the color coding used for obtaining Manhattan and QQ plots as well. These variants have been plotted according to their position in the genome, assuming that the breadth of each of the rectangles represent the genome in order (not drawn to scale). Information on the genomic annotation of the significant hits are represented by symbols given in the legend of each subplot. UTR variants are represented by circles, intronic variants by diamonds, synonymous variants by hexagon, and missense variants are marked with a star symbol. Effect size of a variant should be interpreted as the amount by which the phenotype is increased or decreased (depending on the sign: positive or negative) when one copy of the effect allele increases keeping all other factors constant.

(B) Circos plot showing pairwise interactions (significant and suggestive) of loci with *APOE* influencing diverse domains of cognition. The numbers on the periphery of the circle represent the chromosome. The purple lines represent interactions influencing episodic memory, the yellow lines represent interactions influencing simple processing speed, and the turquoise lines represent interactions affecting visual attention. Tables 2 and 3 contain related details.

model. Controlling for HDL and glucose levels separately, we obtain two independent significant hits—rs17876162 in *PON2* and rs3824734 in *CPEB3*. rs3824734 is an eQTL controlling significant expression of *CPEB3* in cerebellar hemispheres (NES = 0.22, p = 3.8×10^{-5}) (Table S29), which is known for its role in influencing intelligence.[44] Even though rs115865641 (rare variant) is not a significant eQTL controlling expression of *PCDHB16* as per GTEx data, we find that all eQTL variants lying within ±500 kb of our variant significantly control expression of *PCDHB16* in the cerebrum, which contains the prefrontal cerebral cortex—the postulated seat of fluid intelligence[44,45]—and also in cerebellar hemispheres, hippocampus, and basal ganglia (Table S30). Tissue-specific expression data reveal that *PON2* is highly expressed in frontal cortex, anterior cingulate cortex, and basal ganglia (Figure S7), which are areas in the brain correlated with fluid intelligence.[45–47]

eQTL analysis of significant loci associated with simple processing speed

We find that rs3813363 in *SAMD3*, the association hit for mean reaction time from all models, as a significant eQTL controlling expression of *SAMD3* in the cortical regions of the brain (NES = −0.4, p = 1.1×10^{-5}) (Table S29). Several studies have established that cortical regions of the brain are well correlated with reaction time phenotypes assessing the domain of simple processing speed.[48] Similarly, rs17662853—the missense hit in *KANSL1* for reaction time, is an eQTL with significant expression for *KANSL1* in the cerebellum (NES = −0.4; p = 2.3×10^{-5}) and anterior cingulate cortex (NES = −0.49; p = 2.5×10^{-5}) (Table S29), regions responsible for perception and motor response whose co-ordination

is necessary for completion of a reaction time task.[48,49] rs17662853 is also an eQTL controlling expression of *NSFP1*, *LRRC37A* (MIM: 616555), *ARL17A*, *ARL17B*, *RP11*-798G7.8, *NSF* (MIM: 601633), *NSFP1*, and *FAM215B* in several brain regions including cortex and cerebellum (Table S29), highlighting the importance of significantly associated variants obtained from exome-wide analysis that could regulate expression of nearby genes relevant to the biology of the trait. eQTL variants in *GPR108* significantly control expression of *GPR108* in various brain tissues with the lead eQTLs within 500 kbp of our lead SNP controlling *GPR108* expression significantly in the cortex (Table S30). Our eQTL analysis reveals that loci around 500 kbp of rs11062991 (intronic *CCDC77*) most significantly regulates expression of *CCDC77* in hypothalamus and cerebellar hemispheres (Table S30). rs2959174 (synonymous *THAP10*/intronic *LRRC49*) is a significant eQTL regulating the high expression of *LARP6* in cerebellum, cerebellar hemispheres, and putamen of basal ganglia, hippocampus, and cortex (Table S29). Thus, our eQTL analysis reveals another relevant gene *LARP6* for understanding the biology of cognition, even though the identified variant itself annotates to *LRRC49* and *THAP10* with less relevance to cognition.[50] The *LARP6* loci identified from our analysis (rs3825970 and rs1549317) is also a significant eQTL controlling *LARP6* expression in cerebellum, cerebellar hemispheres, and putamen of basal ganglia (Table S29). eQTLs within 500 kbp of rs201404149 (synonymous *MTFR1L*) are significant for expression of *MTFR1L* in cerebellum, cortex, frontal cortex, cerebellar hemispheres, and caudate nucleus of basal ganglia (Table S30).

eQTL analysis of significant loci associated with complex processing speed

The baseline model for this domain yields one significant hit—rs12932325 in the intronic region of *GTF3C1*. rs12932325 is an eQTL for *IL21R* (Table S29) that impacts Alzheimer disease pathology by enhancing brain and peripheral immune and inflammatory responses and leads to increased deposition of Aβ plaques.[51] Both the models controlling for LDL and total cholesterol levels independently yield an intronic variant in *PTPN11* as a significant hit for complex processing speed. The lead eQTL variant near ±500 kb of this variant significantly controls expression of *PTPN11* in the substantia nigra of the brain (Table S30). Research has shown that Parkinson disease causes loss of dopamine-producing neurons in the substantia nigra and dopaminergic processes are shown to be involved in cognitive functions such as processing speed.[52] Controlling for HbA1c, we identify rs71467481 in the intronic region of *PPFIA1* as another significant hit for complex processing speed. Our eQTL analysis shows that variants around 500 kbp of this SNP can significantly regulate expression of *PPFIA1* in many brain regions (Figure S8; Table S30). We find that variants in *APOC1* act as significant eQTLs for regulating its expression in basal ganglia and amygdala (Table S31), where basal ganglia is known to be involved in playing an important role in complex processing speed.[53]

eQTL analysis of significant loci associated with episodic memory

We identify a significant missense variant rs146766120 in *AMIGO1* to be associated with episodic memory with and without controlling for serum HbA1c levels. *AMIGO1* is expressed in the astrocytes, hippocampus, and cortical neurons, and it is postulated to influence neuron survival.[54] In our eQTL analysis too, we find that variants within 500 kb of rs146766120 significantly influences expression of *AMIGO1* in the brain, especially in the cortex (NES = 0.2, $p = 8.2 \times 10^{-12}$) and hippocampus (NES = 0.17, $p = 1.4 \times 10^{-11}$) (Figure S9; Table S30), areas in the brain that interact among each other to encode and retrieve episodic memory,[55,56] thus highlighting the importance of our identified hit in influencing episodic memory. Similarly we identify another synonymous variant rs77807661 in *PTPN18* both with and without controlling for serum HbA1c levels. Significant eQTL variants around 500 kbp of rs77807661 can regulate expression of *PTPN18* in cortex, prefrontal cortex, cerebellum, cerebellar hemispheres, caudate basal ganglia, and nucleus accumbens (Figure S10; Table S30), thus pinpointing the established crucial role of the cerebellum in episodic memory via cortical-cerebellar brain networks.[57] Studies also suggest that memory formation in the hippocampus is guided by motivational significance of events whose effect on memory is thought to depend on interactions between the hippocampus, ventral tegmental area, and nucleus accumbens.[58] The baseline model as well as the model controlling for HbA1c also yield rs111522866 in the intronic region of *ITPR3* as another significant variant for episodic memory. eQTL variants around 500 kb of this variant significantly influence expression of *ITPR3* in cerebellar hemispheres as well as in caudate basal ganglia (Table S30). Upon controlling for LDL, we find another variant, rs7725495 in the intronic region of *POLR3G*, to be associated with episodic memory. eQTL variants within 500 kb of rs7725495 significantly influences expression of *POLR3G* in the cerebellum, cortex, anterior cingulate cortex, hypothalamus, nucleus accumbens, and putamen (Table S30). We identify two additional hits—missense rs3754644 (*IQCA1*) and rs73529530 (intronic *ATP2A1*)—to be associated with episodic memory when controlled for HbA1c levels. eQTL variants around 500 kb of rs3754644 also significantly control *IQCA1* expression in the amygdala, cerebellar hemispheres, cerebellum, cortex, frontal cortex, anterior cingulate cortex, and nucleus accumbens (Table S30). eQTL variants around 500 kb of rs73529530 significantly regulate expression of *ATP2A1* in the hypothalamus (Table S30).

eQTL analysis of significant loci associated with visual attention

We get an association signal of rs11589562 for visual attention when we adjust for HDL level. This variant is located in the intronic region of the *MAST2* gene. eQTL analyses

Table 2. Interaction analysis with variants identified from single-variant association tests

Domain:	Episodic memory								
Model	Interacting SNP1 (chr:Pos_REF_ALT; rsid mapped gene)	Main effect size (SNP1); p value	ALT_AF (SNP1)	Interacting SNP2 (chr:Pos_REF_ALT; rsid mapped gene)	Main effect size (SNP2); p value	ALT_AF (SNP2)	Interaction effect size	Test statistic: CHISQ	p value
Baseline	chr19:44908684_T_C; rs429358 (*APOE*)	0.036; 0.001	0.1531	chr1:109508840_C_T; rs146766120 (*AMIGO1*)	−0.885; 4.68E−11	0.0014	−1.151	7.869	0.0050[a]
	chr19:44908822_C_T; rs7412 (*APOE*)	−0.001; 0.924	0.0799	chr6:33688801_C_T; rs111522866 (*ITPR3*)	−0.853; 6.46E−11	0.0016	−1.271	9.14	0.0025[a]
Baseline + HbA1c	chr19:44908684_T_C; rs429358 (*APOE*)	0.035; 0.002	0.1531	chr1:109508840_C_T; rs146766120 (*AMIGO1*)	−0.769; 4.68E−09	0.0014	−1.154	7.723	0.0055[b]
	chr19:44908684_T_C; rs429358 (*APOE*)	0.034; 0.002	0.1531	chr2:130356125_C_T; rs77807661 (*PTPN18*)	−1.993; 2.42E−10	0.0018	−1.774	4.203	0.0404
	chr19:44908822_C_T; rs7412 (*APOE*)	0.004; 0.790	0.0799	chr6:33688801_C_T; rs111522866 (*ITPR3*)	−0.840; 5.26E−10	0.0016	−0.968	5.296	0.0214
Domain:	Simple Processing Speed								
Model	Interacting SNP1 (chr:Pos_REF_ALT; rsid (mapped gene)	Main effect size (SNP1); p value	ALT_AF (SNP1)	Interacting SNP2 (chr:Pos_REF_ALT; rsid (mapped gene)	Main effect size (SNP2); p value	ALT_AF (SNP2)	Interaction effect size	Test statistic: CHISQ	p value
Baseline + HbA1c	chr19:44908822_C_T; rs7412 (*APOE*)	−0.009; 0.213	0.0799	chr19:6732114_C_T; rs73922480 (*GPR108*)	−0.496; 9.55E−08	0.0014	−0.733	5.096	0.0240
	chr19:44908822_C_T; rs7412 (*APOE*)	−0.009; 0.213	0.0799	chr19:6732194_C_T; rs77285514 (*GPR108*)	−0.418; 9.63E−08	0.0016	−0.542	3.865	0.0493

[a]This interaction remains significant after Bonferroni's correction with five multiple comparisons ($p = 0.05/5 = 0.01$).
[b]This interaction remains significant after Bonferroni's correction with nine multiple comparisons ($p = 0.05/9 = 0.056$).

show that this variant controls expression for *MAST2* in the cerebral cortex (NES = 0.21, $p = 7.70 \times 10^{6}$) and cerebellum (NES = 0.26, $p = 4.6 \times 10^{6}$) (Table S29). It is known that the posterior parietal lobe of the cortex assesses the visual scene and interacts with the frontal lobes in choosing object of interest to plan visually guided movement.[59] This variant is also an eQTL significantly influencing expression of *CCDC163*, *TESK2* (MIM: 604746), and *PIK3R3* (MIM: 606076) in several brain regions (Table S29). *MAST2* is highly expressed in the hypothalamus and substantia nigra (Figure S11). Several studies have found oxytocin, synthesized in several nuclei of the hypothalamus, to regulate visual attention and eye movements to external sensory/social stimuli.[60] In addition, studies have shown dopamine-producing neurons in the ventral tegmental area and substantia nigra to be related to multiple aspects of visual attention.[60] We find that variants in *APOC1* and *LRP1* also act as significant eQTLs for regulating their expression in basal ganglia and cerebellar hemispheres (Table S31), crucial to visual attention.[60]

Interaction analyses

Our epistasis analysis conducted with significant variants from the single-variant analysis reveals four pairs of significant epistatic interactions (at nominal level of significance) with the *APOE* isoform-defining variants (rs7412 and rs429358) for episodic memory and simple processing speed (Table 2; Figure 3B). Each of the variants that interact with either of the two *APOE* variants exerts a significant effect on the phenotype in addition to its interaction effect. These variants are rare, with large effect sizes conforming to the general consensus that rarer variants have larger effect sizes. Out of these interactions, we find the interaction between rs429358 (*APOE*) and rs14676612 (*AMIGO1*) and between rs429358 and rs111522866 (*ITPR3*) of particular interest. The interaction between rs429358 (*APOE*) and rs14676612 (*AMIGO1*) is significant even after Bonferroni's correction (five valid tests in baseline model, cutoff = 0.05/5 = 0.01 and nine valid tests in the HbA1c-controlled model, cutoff = 0.05/9 = 0.056) (indicated with footnotes a and b, respectively, in Table 2). Interactions between rs429358 (*APOE*) and rs14676612 (*AMIGO1*) from all models (except HDL adjusted) remain significant at nominal level when we restrict our analysis to the coding region (Table S9). The *APOE-AMIGO1* interaction obtained from the model controlling for glucose in our coding region-specific analysis, is also significant after multiple hypothesis correction (Table S9). Similarly, the interaction between rs7412 and rs111522866 (*ITPR3*) in the baseline model is also significant after Bonferroni's multiple testing adjustment ($p < 0.01$) (indicated by footnote a in Table 2). We also see another rs429358 (*APOE*) - rs77807661 (*PTPN18*) interaction at nominal level of significance when we tease out the effect of serum HbA1c on episodic memory (Table 2; Figure 3B). Significant *GPR108* variants also exhibit suggestive (nominally significant) interaction effect with *APOE* for simple processing speed (Table 2; Figure 3B). In addition, we also find evidence of interaction between *APOE* variants (rs429358, rs7412) and rs3754644 (*IQCA1*) from models controlling for LDL and triglyceride in coding region-specific analysis (Table S9).

Table 3. Interaction analysis of variants in genes identified from gene-based association tests

Domain:	Visual attention				
Model	Interacting SNP1 (chr:Pos_REF_ALT; rsid mapped gene)	Interacting SNP2 (chr:Pos_REF_ALT; rsid mapped gene)	Interaction effect size	Test statistic: CHISQ	p value
Baseline	chr19:44908822_C_T; rs7412 (*APOE*)	chr19:44919304_T_G; rs1064725 (*APOC1*)	0.196	4.543	0.0331
Baseline + HDL	chr19:44905910_C_G; rs440446 (*APOE*)	chr12:57143889_C_T; rs185694830 (*LRP1*)	0.463	5.829	0.0158
	chr19:44905910_C_G; rs440446 (*APOE*)	chr12:57193273_G_A; rs138348495 (*LRP1*)	0.177	6.501	0.0108[a]
	chr19:44906731_C_T; rs143063029 (*APOE*)	chr12:57183854_C_T; rs138993371 (*LRP1*)	−2.357	5.449	0.0196
	chr19:44906745_G_A; rs769449 (*APOE*)	chr12:57201197_C_T; rs34949484 (*LRP1*)	−0.238	7.135	0.0076
	chr19:44908684_T_C; rs429358 (*APOE*)	chr12:57183939_G_A; rs34423990 (*LRP1*)	−1.523	4.066	0.0438
	chr19:44908684_T_C; rs429358 (*APOE*)	chr12:57201197_C_T; rs34949484 (*LRP1*)	−0.230	9.178	0.0024
	chr19:44908822_C_T; rs7412 (*APOE*)	chr12:57190715_C_T; rs1800183 (*LRP1*)	0.269	4.873	0.0273
	chr19:44908822_C_T; rs7412 (*APOE*)	chr12:57194415_C_T; rs138980324 (*LRP1*)	0.464	5.211	0.0224[a]
	chr19:44908822_C_T; rs7412 (*APOE*)	chr12:57195284_C_T; rs1800142 (*LRP1*)	0.288	6.502	0.0108[a]
Baseline + HbA1c	chr19:44908822_C_T; rs7412 (*APOE*)	chr19:44919304_T_G; rs1064725 (*APOC1*)	0.186	3.898	0.0483

[a]These interactions are also significant in the coding region-specific interaction analysis, at nominal level of significance.

The epistasis analysis conducted on the basis of gene-based tests reveals nine significant epistatic interactions between five *APOE* (rs440446, rs143063029, rs769449 rs429358, and rs7412) and eight *LRP1 SNPs* (Table 3) to be associated with visual attention, adjusting for HDL levels at nominal level of significance. Out of these, three interactions (footnote a in Table 3)—namely between rs440446 and rs138348495 (*LRP1*, MAF = 0.0085), rs7412 and rs138980324 (*LRP1*, MAF = 0.0028), rs7412 and rs1800142 (*LRP1*, MAF = 0.0083)—remain significant in the separate analysis restricted to coding region variants at the nominal level. We detect another nominally significant interaction between rs7412 (*APOE*) and rs1064725 (*APOC1*) associated with visual attention while controlling for baseline covariates and for HbA1c independently (Table 3).

Discussion

Our study is a comprehensive analysis to understand the genetic architecture of human cognition via single-variant-based, gene-based association, pairwise interaction, mediation, and pleiotropy analyses (Figure 4).

Our single-variant and gene-based association identifies independent loci in *PCDHB16*, *PON2*, *CPEB3*, *LRRC49/THAP10*, *CCDC77*, *LARP6*, *MTFR1L*, *GPR108*, *GTF3C1*, *PTPN11*, *PPFIA1*, *AMIGO1*, *ITPR3*, *PTPN18*, *IQCA1*, *ATP2A1*, *POLR3G*, *MAST2*, *APOC1*, and *LRP1*, which have not been previously reported, and previously known *KANSL1* and *SAMD3* as associated with diverse cognition domains (Figure 3A) in baseline as well as while adjusting for serum lipids and glycemic levels, which are postulated to be modifiable metabolic risk factors for cognition. These genes are known to function in pathways relevant to cognition. For instance, *PCDHB16* localizes mainly in the post-synaptic compartment and serves as a candidate gene for specification of synaptic connectivity and neuronal networks,[61] a key element for cognition. *PON2* (Paraoxonase-2), a mitochondrial enzyme, has higher expression in dopaminergic regions such as striatum, striatal astrocytes, and cortical microglia,[62] which highlights its role in protecting cells from oxidative damage and neuroinflammation.[62] *CPEB3* is involved in synaptic protein regulation, acting as a negative regulator of AMPA receptor subunits GluA1 and GluA2 to maintain long-term synaptic plasticity.[63] Koolen-de Vries syndrome/17q21.31 microdeletion syndrome characterized by intellectual disability has been attributed to mutations in *KANSL1*.[64] One study showed that autophagosome accumulation at excitatory synapses in *KANSL1*-deficient neurons led to reduced synaptic density, reduced transmission via GRIA/AMPA receptors, along with impairment of neuronal network activity.[65] A recent study showed that *MTFR1L* expression changed in

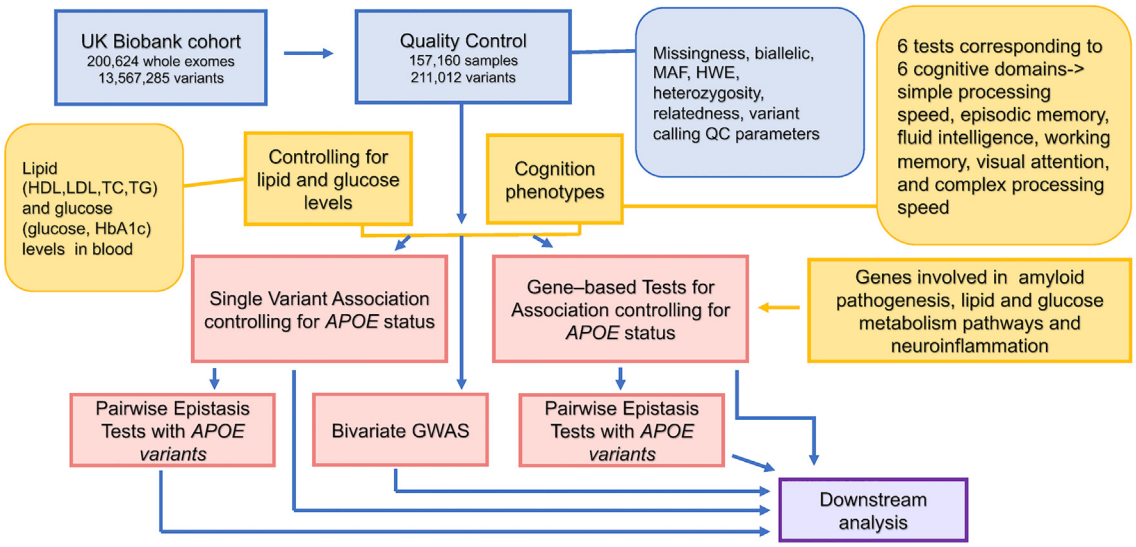

Figure 4. Workflow of this study

Blue boxes represent the information about the data and quality checks performed; the yellow boxes are indicative of the phenotypes and genes considered for gene-level analysis. Red boxes highlight the statistical tests performed; and the purple box indicates downstream analysis performed, such as annotations, gene expression analysis, and mediation analysis.

the hippocampus and cerebral cortex in memantine-treated transgenic Alzheimer-diseased mice.[66] Memantine is an FDA-approved prescription drug administered to improve learning and memory for moderate-to-severe Alzheimer cases, suggesting the importance of our identified loci in *MTFR1L* for understanding cognition in the pathophysiology of Alzheimer disease. A *GTF3C1* locus has been found to be significantly associated with entorhinal cortical thickness, an Alzheimer disease-related neuroimaging biomarker (Figure 2),[67,68] which is in linkage disequilibrium with our *GTF3C1* variant, signifying that our variants could potentially influence such related traits, and thus allude to the shared genetic mechanisms of cognition, Alzheimer disease, and gray matter density. *PTPN11* is a tyrosine phosphatase that activates MAPK pathway, plays a critical role in synaptic plasticity and memory formation,[69] and interacts with tau in Alzheimer patients.[70] Mutations in *PTPN11* have been associated with numerous syndromes; especially known is the Noonan syndrome, which is known to affect human cognition,[67] along with cardiovascular abnormalities and congenital heart defects.[71] *PPFIA1* encodes the neuronal scaffold protein liprin-α1 functioning in active synaptic zones and post-synaptic sites,[72] and has been proposed as a candidate gene involved in late-onset Alzheimer disease etiology.[73] *AMIGO1*, required for myelination of developing neurons and playing important role in neural plasticity in adult nervous system, is a commonly altered marker gene in Alzheimer patients.[74] *PTPN18* is a non-receptor tyrosine phosphatase expressed in neural tissues, likely influencing Alzheimer disease progression.[75] *ITPR3* encodes inositol 1,4,5-trisphosphate receptor, type 3, which mediates release of intracellular calcium and facilitates crucial intra-organellar Ca^{2+} signal transmission from the endoplasmic reticulum (ER) to the mitochondria[76] to maintain proper cognition. *IQCA1*, found to be upregulated in the hippocampus of Alzheimer-like monkeys compared with normal aged monkeys, is postulated to be associated with brain AMPKα2 activity, playing a pivotal role in *de novo* protein synthesis, an indispensable phenomenon for long-term synaptic plasticity and memory formation.[77] *ATP2A1* encodes proteins associated with mitochondria-associated-ER membrane (MAM),[78] and disruption at this locus could perturb MAM functioning, which is posited to play a role in Alzheimer disease pathogenesis. We thus note that these implicated variants and genes, through their functioning in synaptic plasticity and connectivity, oxidative stress, and neuroinflammation, could play vital roles in influencing cognition, and thus in some instances possibly impact the pathology of Alzheimer disease and related dementias. Interestingly, all-risk alleles of single-variant hits (highlighted in bold in Table 1, detailed interpretation in footnotes and legend) affecting cognition adversely are common in the population with allele frequency >5%, thus highlighting the significance of our work for studying the genetic context of cognitive abilities of individuals in the general population to understand the risk factors for cognitive decline. We also obtain significant hits harbored in the coding region that are in LD with genotyped variants identified by Davies et al.[5] associated with reaction time, thus highlighting the importance of exome-based analysis in uncovering likely causal associations.

MAF is one of the widely used metrics to study genetic variation at population level and the proportion of rare variants (MAF < 1%) contributing to complex traits is an important metric for understanding their heritability and genetic architecture. Functional annotation of our significant hits shows that, among the variants mapped to genes

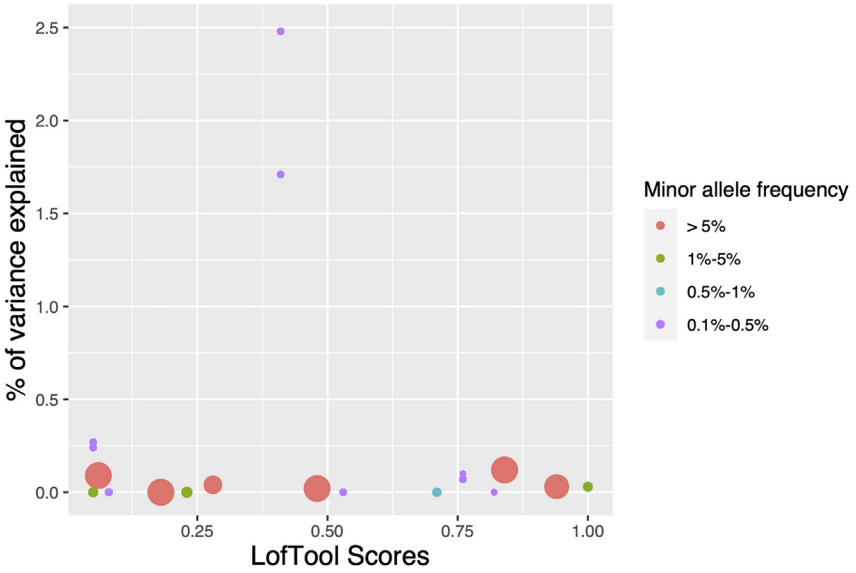

Figure 5. Phenotypic variance explained with respect to evolutionary constraint acting on the variants identified from single-variant association tests

This figure represents the proportion of phenotypic variance explained by the exome-wide significantly associated variants for diverse cognitive domains vs. their intolerance to genic functional changes as indicated by their LofTool scores. The lower the LoFtool score, the more intolerant are the mapped genes to functional changes. Coral and green solid circles represent common, low-frequency variants, respectively, while turquoise and violet solid circles represent rare variants. The sizes of the circles are proportional to the minor allele frequency of the corresponding variants represented by the circles.

with lower tolerance to loss of function effects (LoFTool score < 0.25), the rare variants (MAF $< 1\%$) explain higher proportion of variance compared with low frequency (MAF 1%–5%) and common variants (MAF $> 5\%$), a phenomenon conforming to the general consensus that rare variants exhibit higher effects (in terms of magnitude of effect size) (Figure 5). Nonetheless, even among the variants mapped to genes exhibiting moderate intolerance to loss-of function (LoFTool score between 0.25 and 0.5), we see that the common variants explain a lower proportion of phenotypic variation than the rare variants (Figure 5). Thus, our results show that common variants that explain comparatively less proportion of phenotypic variation can also map to genes that are moderately intolerant to loss of function. Most of our association signals from the single-variant association analysis are most likely to affect transcription factor (TF) binding and affect expression of a gene target (RegulomeDB rank: 1d, 1f, 2a) (Table S32). The other hits (rank: 4–6) have binding evidence due to being located in a functional region implicated in TF binding and DNase peak demonstrating protein binding sites (Table S32). Similarly, for the suggestively interacting variants in *APOE*, *APOC1*, and *LRP1*, we find evidence that these variants are likely to be highly functional (rank: 1b, 1f, 2b, 4) via their regulatory effects in the genome (Table S32).

Out of the 20 independent loci (Table 1; Figure 3A), rs3824734 (*CPEB3*) implicated in fluid intelligence, rs3813363 (*SAMD3*) and rs3825970 (*LARP6*) implicated in simple processing speed are eQTL loci, significantly controlling expression of their respective genes in cerebellum, cortex, and basal ganglia. rs2959174 (*LRRC49/THAP10*) and rs12932325 (*GTF3C1*), associated with simple and complex processing speed, respectively, are significant eQTL, controlling expression of nearby genes such as *LARP6* and *IL21R*, respectively. rs11589562 (*MAST2*) associated with visual attention and rs17662853 (*KANSL1*) associated with simple processing speed are significant eQTLs, controlling expression of the respective mapped genes as well as nearby genes. For the remaining loci, namely rs115865641 (*PCDHB16*), rs73922480 (*GPR108*), rs11062991 (*CCDC77*), rs201404149 (*MTFR1L*), rs12301915 (*PTPN11*), rs71467481 (*PPFIA1*), rs146766120 (*AMIGO1*), rs77807661 (*PTPN18*), rs111522866 (*ITPR3*), rs7725495 (*POLR3G*), rs3754644 (*IQCA1*), and rs73529530 (*ATP2A1*), we observe eQTLs in the vicinity (± 500 kbp) controlling the expression of their respectively annotated genes in different brain regions pertinent to cognition (Table S30). The genes mapped to these 12 variants are highly expressed in brain regions deemed responsible for completion of neuropsychological tasks corresponding to respective cognitive domains. All of these 12 variants are either rare or low frequency (MAF ranging from 0.1% to 2%) and that could be one possible reason for them to have lower LD R^2 values with the identified eQTLs in the vicinity. But, interestingly, all these variants are likely to affect either TF binding or are located in the DNase peak, thus enhancing chromatin accessibility (Table S32, RegulomeDB ranks: 1f, 2a, 4, 5, 6) and could possibly regulate transcriptional activity. On the other hand, the eQTLs in their vicinity are also likely to be highly functional by virtue of them being eQTLs. Thus, these variants could have some regulatory role in expression of genes responsible for cognition.

Our study reveals a previously unknown evidence of *LRP1* association with cognition. Furthermore, we find that six out of the eight *LRP1* SNPs that interact with *APOE* (at nominal level of significance) are rare and the remaining two are of low frequency. Targeted experiments have shown roles for *APOC1* and *LRP1*[79] in cognitive decline or neurodegeneration; however, our interaction analysis (Figure 3B; Table 3) now identifies suggestive SNPs in *APOC1* and *LRP1* acting in conjunction with well-recognized *APOE* in governing cognitive abilities, thus providing evidence for the role of *LRP1* on cognition. In total, we identify 2 Bonferroni-level significant and 12 nominally significant pairwise

interactions relevant to episodic memory, simple processing speed, visual attention between *APOE* and our exome-wide associated hits, many of them are interestingly rare (MAF 0.12%–4%). Our study shows evidence of interactions especially between *APOE* and *AMIGO1*, *APOE*, and *ITPR3*, and suggestive interactions of *APOE* with *PTPN18* and *GPR108* in influencing cognition or neurodegeneration, which was hitherto unreported.

We uncover in total 514 loci associated with different domains of cognition by performing exome-wide burden- and kernel-based association testing at gene level, the high number of hits being anticipated because whole-exome-based approaches lead to large-scale discoveries. For instance, Davies et al.[5] identified 709 genes for general cognition where their gene sets and gene property analysis points to neurodevelopmental pathways implicated in general cognition with expression levels significant in cerebellum and cortex. In our exome-wide gene-based analysis we find five pathways related to long-term depression, ras signaling, mannose type O-glycan biosynthesis, spinocerebellar ataxia, and peptidyl-tyrosine dephosphorylation involved in inactivation of protein kinase activity. Even though the pathways uncovered are non-specific to cognition, the genes co-occurring in the pathways have been linked to cognition and Alzheimer disease. For example, the *PTPR* family genes are known to play vital roles in cell adhesion, neurite growth, and cell differentiation.[80] *PPP2R1A* has been found to be associated with Alzheimer disease and visuospatial domain.[81] There exists evidence of *de novo* mutations in *GRIA2* potentially causing neurodevelopmental disorders with language impairment and behavioral abnormalities.[82] *PLA2G4E* has been identified as a candidate gene for resilience in Alzheimer disease.[83] mRNAs for O-glycan biosynthesis genes like *POMT2* have been found to be altered in Alzheimer disease in multiple regions of the brain.[84] In general, the relationships between glycosylation modulation and neuroinflammation have been explored, which are key for cognitive functioning.[85] *OMA1* is a gene known to function in mitochondrial dynamics, which gets affected in pathological conditions of Alzheimer disease.[86] *DAB1* plays a regulatory role in Reelin signaling in the adult brain through synaptic functioning and is required for normal functioning.[87] Genome-wide analysis in *APOE*-ϵ4 homozygotes as well functional enrichment studies have identified *DAB1* as a protective candidate gene in Alzheimer disease etiology.[88]

Nevertheless, with exome-wide analysis yielding a large number of significant results, it becomes difficult to pinpoint genes and provide holistic interpretation. On the contrary, incorporation of *a priori* knowledge of biological pathways into the analysis is often found to yield relevant biological insights that cannot be detected by focusing on genes that provide strongest attestation of differential expression.[89,90] Here, our knowledge-driven approach (based on 20 genes) leverages the widely established information of biological pathways in which *APOE* functions and is relevant to Aβ deposition degradation and clearance with relevant metabolic risk factors; the major processes, when disrupted, could lead to memory impairment and cognitive decline.[91] Our detailed eQTL analysis of the implicated genes and variants in these pathways also reveal their significant regulatory roles in expression in cerebellum, cortex, and basal ganglia. Thus, our context specific findings provide potential basis for experimental interrogation.

Upon analyzing the subset of variants belonging to only exon-coding regions, we find significant variants in *NOSTRIN* (MIM: 607496), *OMA1* (MIM: 216081), *HIGD1B*, *DDX27* (MIM: 616621), *AMTN* (MIM: 610912), *NKX2-5* (MIM: 600584), *BBS9* (MIM: 607968), *PTF1A* (MIM: 607194), *E2F7* (MIM: 612046), *FAM210A* (MIM: 617975), *PGAM1* (MIM:171900), *KMT2D* (MIM: 602113), and *HPR* (MIM:140210) (Table S8) for four domains of cognition. The relevance of many of these hits to biological basis of cognition is not clear. Notably, though, the *KMT2D* variant—rs3782357— is in LD ($R^2 > 0.97$) with variants previously associated with intelligence, verbal numerical reasoning, and reaction time.[5,6,92] Also, *NOSTRIN* binds an enzyme responsible for nitric oxide production, which is a mediator in neurotransmission, inflammation, and vascular homeostasis.[93] The hits in *AMIGO1*, *CPEB3*, *GPR108*, *IQCA1*, *KANSL1*, *PTPN18*, and *THAP10* are significant from the single-variant association analysis comprising variants from capture region plus flanking regions as well as in this coding region-specific analysis, and these genes have established biological roles in neuronal processes including cognition. Overall, our results highlight that coding missense and disruptive variants as well as variants flanking exonic transcripts are crucial to understanding the genetic architecture of complex traits, such as cognition, because potential functional regulatory roles of variants are important and ought to be considered for interpretation of their effects on traits.

Despite several strengths of this study, we acknowledge the fact that the results reported herein must be considered in the light of some limitations. Firstly, even though the initial sample size is quite large (∼157,000), effective sample sizes varies for each test (∼27,000–121,000), and is less because we have ensured that each participant has non-missing data on all variables of interest (phenotype and covariates) for all models (Table S1). Secondly, cognition can be affected by a multitude of factors apart from age-related metabolic conditions of dyslipidemia and glycemic risk, for instance, obesity, cardiovascular disease, hypertension, physical activity, mood, personality, social activity, smoking status, alcohol consumption, income, etc.,[94,95] which we have not considered as covariates in our model. So there also lies the possibility of residual confounding in our results. Thirdly, our analyses has been based on individuals from primarily European ancestry. So caution must be exercised while generalizing the results for diverse ancestries.

Data and code availability

All phenotype and genotype data used in this study for analysis are available at UK Biobank. The custom codes used in our analysis are given in GitHub.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.xhgg.2023.100208.

Acknowledgments

We are extremely thankful to Dr. Balaji Jayaprakash for his support and enthusiastic encouragement toward this work. We would also like to thank Mr. Sheldon D'Silva for his help in carrying out quality control processes for the genetic data. We are also grateful to our system administrators Mr. Naveenan Srinivasan, Mr. Karthik Sundaram, and Mr. Anand Kumar E for their support in managing technical challenges encountered while carrying out this computational work. We also thank the funding agencies: Science & Engineering Research Board (SERB), Government of India (GOI) (ECR/2018/001429), Department of Biotechnology, Government of India (BT/RLF/29/2016), and National Supercomputing Mission, Government of India (DST/NSM/R&D_HPC_Applications/2021/03.12) for funding equipment and data acquisition for this study. We conducted this research using the UK Biobank Resource under application no. 55652. We thus extend our sincere gratitude to all UK Biobank participants, researchers, clinicians, technicians, administrative staff, and funding authorities who enabled curation of this enriched biomedical resource. We acknowledge the Prime Minister's Research Fellowship (GOI) to Shreya Chakraborty. Statement of ethics: the UK Biobank Study received approval from the North West Multicentre Research Ethics Committee (MREC) as a Research Tissue Bank approval. This approval means that researchers do not require separate ethical clearance and can operate under the RTB approval.

Author contributions

B.K. conceived and designed the study. S.C. and B.K. performed the data analysis. S.C. and B.K. wrote the manuscript. S.C. and B.K. prepared the figures. Both authors have read and approved the final manuscript.

Declaration of interests

The authors declare no competing interests.

Received: December 9, 2022
Accepted: May 16, 2023

Web resources

UK Biobank, https://www.ukbiobank.ac.uk

UK Biobank, https://www.ukbiobank.ac.uk/media/cfulxh52/uk-biobank-exome-release-faq_v9-december-2020.pdf

UK Biobank, https://biobank.ctsu.ox.ac.uk/crystal/refer.cgi?id=8481

UK Biobank, https://biobank.ndph.ox.ac.uk/showcase/ukb/docs/serum_biochemistry.pdf

UK Biobank, https://biobank.ndph.ox.ac.uk/showcase/ukb/docs/serum_hb1ac.pdf

1000 Genomes Project, https://www.internationalgenome.org/

LDtrait, https://ldlink.nci.nih.gov/?tab=ldtrait

UniProt, https://www.uniprot.org/

GWAS ATLAS, https://atlas.ctglab.nl/

VEP (LofTool), https://asia.ensembl.org/Homo_sapiens/Tools/VEP/

Harmonizome, https://maayanlab.cloud/Harmonizome/

Gtex, https://gtexportal.org/home/

NCBI, https://www.ncbi.nlm.nih.gov/

dbSNP;, https://www.ncbi.nlm.nih.gov/snp/

UCSC, https://genome.ucsc.edu/

OMIM, https://www.omim.org/

RegulomeDb, https://beta.regulomedb.org/regulome-search/

GitHub, https://github.com/BratatiKahaliLab/Exome-based-analysis-on-human-cognition

References

1. Davies, G., Tenesa, A., Payton, A., Yang, J., Harris, S.E., Liewald, D., Ke, X., le Hellard, S., Christoforou, A., Luciano, M., et al. (2011). Genome-wide association studies establish that human intelligence is highly heritable and polygenic. Mol. Psychiatry *16*, 996–1005. https://doi.org/10.1038/mp.2011.85.
2. Davies, G., Armstrong, N., Bis, J.C., Bressler, J., Chouraki, V., Giddaluru, S., Hofer, E., Ibrahim-Verbaas, C.A., Kirin, M., Lahti, J., et al. (2015). Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium (N=53 949). Mol. Psychiatry *20*, 183–192. https://doi.org/10.1038/mp.2014.188.
3. Trampush, J.W., Yang, M.L.Z., Yu, J., Knowles, E., Davies, G., Liewald, D.C., Starr, J.M., Djurovic, S., Melle, I., Sundet, K., et al. (2017). GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive function: a report from the COGENT consortium. Mol. Psychiatry *22*, 336–345. https://doi.org/10.1038/mp.2016.244.
4. Sniekers, S., Stringer, S., Watanabe, K., Jansen, P.R., Coleman, J.R.I., Krapohl, E., Taskesen, E., Hammerschlag, A.R., Okbay, A., Zabaneh, D., et al. (2017). Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. Nat. Genet. *49*, 1107–1112. https://doi.org/10.1038/ng.3869.
5. Davies, G., Lam, M., Harris, S.E., Trampush, J.W., Luciano, M., Hill, W.D., Hagenaars, S.P., Ritchie, S.J., Marioni, R.E., Fawns-Ritchie, C., et al. (2018). Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. Nat. Commun. *9*, 2098. https://doi.org/10.1038/s41467-018-04362-x.
6. de la Fuente, J., Davies, G., Grotzinger, A.D., Tucker-Drob, E.M., and Deary, I.J. (2021). A general dimension of genetic sharing across diverse cognitive traits inferred from molecular data. Nat. Hum. Behav. *5*, 49–58. https://doi.org/10.1038/s41562-020-00936-2.
7. Hatoum, A.S., Morrison, C.L., Mitchell, E.C., Lam, M., Benca-Bachman, C.E., Reineberg, A.E., Palmer, R.H.C., Evans, L.M.,

Keller, M.C., and Friedman, N.P. (2023). Genome-wide association study shows that executive functioning is influenced by GABAergic processes and is a neurocognitive genetic correlate of psychiatric disorders. Biol. Psychiatry *93*, 59–70. https://doi.org/10.1016/j.biopsych.2022.06.034.

8. Ramanan, V.K., Nho, K., Shen, L., Risacher, S.L., Kim, S., McDonald, B.C., Farlow, M.R., Foroud, T.M., Gao, S., Soininen, H., et al. (2015). FASTKD2 is associated with memory and hippocampal structure in older adults. Mol. Psychiatry *20*, 1197–1204. https://doi.org/10.1038/mp.2014.142.
9. Huentelman, M.J., Papassotiropoulos, A., Craig, D.W., Hoerndli, F.J., Pearson, J.v., Huynh, K.-D., Corneveaux, J., Hänggi, J., Mondadori, C.R.A., Buchmann, A., et al. (2007). Calmodulin-binding transcription activator 1 (CAMTA1) alleles predispose human episodic memory performance. Hum. Mol. Genet. *16*, 1469–1477. https://doi.org/10.1093/hmg/ddm097.
10. Davies, G., Marioni, R.E., Liewald, D.C., Hill, W.D., Hagenaars, S.P., Harris, S.E., Ritchie, S.J., Luciano, M., Fawns-Ritchie, C., Lyall, D., et al. (2016). Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=112 151). Mol. Psychiatry *21*, 758–767. https://doi.org/10.1038/mp.2016.45.
11. Gupta, A., and Kahali, B. (2020). Machine learning-based cognitive impairment classification with optimal combination of neuropsychological tests. Alzheimers Dement. *6*, e12049. https://doi.org/10.1002/trc2.12049.
12. Homann, J., Osburg, T., Ohlei, O., Dobricic, V., Deecke, L., Bos, I., Vandenberghe, R., Gabel, S., Scheltens, P., Teunissen, C.E., et al. (2022). Genome-wide association study of Alzheimer's disease brain imaging biomarkers and neuropsychological phenotypes in the European medical information framework for Alzheimer's disease multimodal biomarker discovery dataset. Front. Aging Neurosci. *14*, 840651. https://doi.org/10.3389/fnagi.2022.840651.
13. Davies, G., Harris, S.E., Reynolds, C.A., Payton, A., Knight, H.M., Liewald, D.C., Lopez, L.M., Luciano, M., Gow, A.J., Corley, J., et al. (2014). A genome-wide association study implicates the APOE locus in nonpathological cognitive ageing. Mol. Psychiatry *19*, 76–87. https://doi.org/10.1038/mp.2012.159.
14. Raisi-Estabragh, Z., M'Charrak, A., McCracken, C., Biasiolli, L., Ardissino, M., Curtis, E.M., Aung, N., Suemoto, C.K., Mackay, C., Suri, S., et al. (2022). Associations of cognitive performance with cardiovascular magnetic resonance phenotypes in the UK Biobank. Eur. Heart J. Cardiovasc. Imaging *23*, 663–672. https://doi.org/10.1093/ehjci/jeab075.
15. Lane, R.M., and Farlow, M.R. (2005). Lipid homeostasis and apolipoprotein E in the development and progression of Alzheimer's disease. J. Lipid Res. *46*, 949–968. https://doi.org/10.1194/jlr.M400486-JLR200.
16. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. Nature *562*, 203–209. https://doi.org/10.1038/s41586-018-0579-z.
17. Speed, D., Hemani, G., Johnson, M.R., and Balding, D.J. (2012). Improved heritability estimation from genome-wide SNPs. Am. J. Hum. Genet. *91*, 1011–1021. https://doi.org/10.1016/j.ajhg.2012.10.010.
18. Watanabe, K., Stringer, S., Frei, O., Umićević Mirkov, M., de Leeuw, C., Polderman, T.J.C., van der Sluis, S., Andreassen, O.A., Neale, B.M., and Posthuma, D. (2019). A global overview of pleiotropy and genetic architecture in complex traits. Nat. Genet. *51*, 1339–1348. https://doi.org/10.1038/s41588-019-0481-0.
19. Zhan, X., Hu, Y., Li, B., Abecasis, G.R., and Liu, D.J. (2016). RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. Bioinformatics *32*, 1423–1426. https://doi.org/10.1093/bioinformatics/btw079.
20. Machiela, M.J., and Chanock, S.J. (2015). LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics *31*, 3555–3557. https://doi.org/10.1093/bioinformatics/btv402.
21. Buniello, A., MacArthur, J.A.L., Cerezo, M., Harris, L.W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E., et al. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res. *47*, D1005–D1012. https://doi.org/10.1093/nar/gky1120.
22. Jo, B.-S., and Choi, S.S. (2015). Introns: the functional benefits of introns in genomes. Genomics Inform. *13*, 112–118. https://doi.org/10.5808/GI.2015.13.4.112.
23. Ye, Y., and Zeng, Y. (2019). Whole exome sequencing identifies a novel intron heterozygous mutation in TSC2 responsible for tuberous sclerosis complex. Sci. Rep. *9*, 4456. https://doi.org/10.1038/s41598-019-38898-9.
24. Wu, M.C., Lee, S., Cai, T., Li, Y., Boehnke, M., and Lin, X. (2011). Rare-variant association testing for sequencing data with the sequence kernel association test. Am. J. Hum. Genet. *89*, 82–93. https://doi.org/10.1016/j.ajhg.2011.05.029.
25. Lee, S., Emond, M.J., Bamshad, M.J., Barnes, K.C., Rieder, M.J., Nickerson, D.A., NHLBI GO Exome Sequencing Project—ESP Lung Project Team, Christiani, D.C., Wurfel, M.M., and Lin, X. (2012). Optimal unified approach for rare-variant association testing with application to small-sample case-control whole-exome sequencing studies. Am. J. Hum. Genet. *91*, 224–237. https://doi.org/10.1016/j.ajhg.2012.06.007.
26. Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M., and Ishiguro-Watanabe, M. (2023). KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Res. *51*, D587–D592. https://doi.org/10.1093/nar/gkac963.
27. Ge, S.X., Jung, D., and Yao, R. (2020). ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics *36*, 2628–2629. https://doi.org/10.1093/bioinformatics/btz931.
28. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575. https://doi.org/10.1086/519795.
29. Ray, D., and Boehnke, M. (2018). Methods for meta-analysis of multiple traits using GWAS summary statistics. Genet. Epidemiol. *42*, 134–145. https://doi.org/10.1002/gepi.22105.
30. Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. Nat. Genet. *46*, 310–315. https://doi.org/10.1038/ng.2892.
31. Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) *6*, 80–92. https://doi.org/10.4161/fly.19695.

32. Fadista, J., Oskolkov, N., Hansson, O., and Groop, L. (2017). LoFtool: a gene intolerance score based on loss-of-function variants in 60 706 individuals. Bioinformatics *33*, 471–474. https://doi.org/10.1093/bioinformatics/btv602.
33. GTEx Consortium, Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., et al. (2013). The genotype-tissue expression (GTEx) project. Nat. Genet. *45*, 580–585. https://doi.org/10.1038/ng.2653.
34. Boyle, A.P., Hong, E.L., Hariharan, M., Cheng, Y., Schaub, M.A., Kasowski, M., Karczewski, K.J., Park, J., Hitz, B.C., Weng, S., et al. (2012). Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. *22*, 1790–1797. https://doi.org/10.1101/gr.137323.112.
35. Ligthart, S., de Vries, P.S., Uitterlinden, A.G., Hofman, A., CHARGE Inflammation working group, Franco, O.H., Chasman, D.I., and Dehghan, A. (2015). Pleiotropy among common genetic loci identified for cardiometabolic disorders and C-reactive protein. PLoS One *10*, e0118859. https://doi.org/10.1371/journal.pone.0118859.
36. Fuior, E.v., and Gafencu, A.v. (2019). Apolipoprotein C1: its pleiotropic effects in lipid metabolism and beyond. Int. J. Mol. Sci. *20*, 5939. https://doi.org/10.3390/ijms20235939.
37. Liu, C., Chyr, J., Zhao, W., Xu, Y., Ji, Z., Tan, H., Soto, C., Zhou, X.; and Alzheimer's Disease Neuroimaging Initiative (2018). Genome-wide association and mechanistic studies indicate that immune response contributes to Alzheimer's disease development. Front. Genet. *9*, 410. https://doi.org/10.3389/fgene.2018.00410.
38. Cao, W., Tian, S., Zhang, H., Zhu, W., An, K., Shi, J., Yuan, Y., and Wang, S. (2020). Association of low-density lipoprotein receptor-related protein 1 and its rs1799986 polymorphism with mild cognitive impairment in Chinese patients with type 2 diabetes. Front. Neurosci. *14*, 743. https://doi.org/10.3389/fnins.2020.00743.
39. Shi, Y.M., Zhou, H., Zhang, Z.J., Yu, H., Bai, F., Yuan, Y.G., Deng, L.L., and Jia, J.P. (2009). Association of the LRP1 gene and cognitive performance with amnestic mild cognitive impairment in elderly Chinese. Int. Psychogeriatr. *21*, 1072–1080. https://doi.org/10.1017/S104161020999072X.
40. Shinohara, M., Tachibana, M., Kanekiyo, T., and Bu, G. (2017). Role of LRP1 in the pathogenesis of Alzheimer's disease: evidence from clinical and preclinical studies. J. Lipid Res. *58*, 1267–1281. https://doi.org/10.1194/jlr.R075796.
41. Aizpurua, A., and Koutstaal, W. (2010). Aging and flexible remembering: contributions of conceptual span, fluid intelligence, and frontal functioning. Psychol. Aging *25*, 193–207. https://doi.org/10.1037/a0018198.
42. Cochrane, A., Simmering, V., and Green, C.S. (2019). Fluid intelligence is related to capacity in memory as well as attention: evidence from middle childhood and adulthood. PLoS One *14*, e0221353. https://doi.org/10.1371/journal.pone.0221353.
43. Jaeger, J. (2018). Digit symbol substitution test. J. Clin. Psychopharmacol. *38*, 513–519. https://doi.org/10.1097/JCP.0000000000000941.
44. Yoon, Y.B., Shin, W.-G., Lee, T.Y., Hur, J.-W., Cho, K.I.K., Sohn, W.S., Kim, S.-G., Lee, K.-H., and Kwon, J.S. (2017). Brain structural networks associated with intelligence and visuomotor ability. Sci. Rep. *7*, 2177. https://doi.org/10.1038/s41598-017-02304-z.
45. Chen, P.-Y., Chen, C.-L., Hsu, Y.-C., Cam-CAN, and Tseng, W.Y.I. (2020). Fluid intelligence is associated with cortical volume and white matter tract integrity within multiple-demand system across adult lifespan. Neuroimage *212*, 116576. https://doi.org/10.1016/j.neuroimage.2020.116576.
46. Burgaleta, M., MacDonald, P.A., Martínez, K., Román, F.J., Álvarez-Linera, J., Ramos González, A., Karama, S., and Colom, R. (2014). Subcortical regional morphology correlates with fluid and spatial intelligence. Hum. Brain Mapp. *35*, 1957–1968. https://doi.org/10.1002/hbm.22305.
47. Rhein, C., Mühle, C., Richter-Schmidinger, T., Alexopoulos, P., Doerfler, A., and Kornhuber, J. (2014). Neuroanatomical correlates of intelligence in healthy young adults: the role of basal ganglia volume. PLoS One *9*, e93623. https://doi.org/10.1371/journal.pone.0093623.
48. Paraskevopoulou, S.E., Coon, W.G., Brunner, P., Miller, K.J., and Schalk, G. (2021). Within-subject reaction time variability: role of cortical networks and underlying neurophysiological mechanisms. Neuroimage *237*, 118127. https://doi.org/10.1016/j.neuroimage.2021.118127.
49. Coon, W.G., Gunduz, A., Brunner, P., Ritaccio, A.L., Pesaran, B., and Schalk, G. (2016). Oscillatory phase modulates the timing of neuronal activations and resulting behavior. Neuroimage *133*, 294–301. https://doi.org/10.1016/j.neuroimage.2016.02.080.
50. Rouillard, A.D., Gundersen, G.W., Fernandez, N.F., Wang, Z., Monteiro, C.D., McDermott, M.G., and Ma'ayan, A. (2016). The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. Database *2016*, baw100. https://doi.org/10.1093/database/baw100.
51. Agrawal, S., Baulch, J.E., Madan, S., Salah, S., Cheeks, S.N., Krattli, R.P., Subramanian, V.S., Acharya, M.M., and Agrawal, A. (2022). Impact of IL-21-associated peripheral and brain crosstalk on the Alzheimer's disease neuropathology. Cell. Mol. Life Sci. *79*, 331. https://doi.org/10.1007/s00018-022-04347-6.
52. Nguyen, H., Hall, B., Higginson, C.I., Sigvardt, K.A., Zweig, R., and Disbrow, E.A. (2017). Theory of cognitive aging in Parkinson disease. J. Alzheimers Dis. Parkinsonism *7*. https://doi.org/10.4172/2161-0460.1000369.
53. Botzung, A., Philippi, N., Noblet, V., Loureiro de Sousa, P., and Blanc, F. (2019). Pay attention to the basal ganglia: a volumetric study in early dementia with Lewy bodies. Alzheimer's Res. Ther. *11*, 108. https://doi.org/10.1186/s13195-019-0568-y.
54. Chen, Y., Hor, H.H., and Tang, B.L. (2012). AMIGO is expressed in multiple brain cell types and may regulate dendritic growth and neuronal survival. J. Cell. Physiol. *227*, 2217–2229. https://doi.org/10.1002/jcp.22958.
55. Dickerson, B.C., and Eichenbaum, H. (2010). The episodic memory system: neurocircuitry and disorders. Neuropsychopharmacology *35*, 86–104. https://doi.org/10.1038/npp.2009.126.
56. Meenakshi, P., and Balaji, J. (2017). Neural circuits of memory consolidation and generalisation. J. Indian Inst. Sci. *97*, 487–495. https://doi.org/10.1007/s41745-017-0042-4.
57. Andreasen, N.C., O'Leary, D.S., Paradiso, S., Cizadlo, T., Arndt, S., Watkins, G.L., Boles Ponto, L.L., and Hichwa, R.D. (1999). The cerebellum plays a role in conscious episodic memory retrieval. Hum. Brain Mapp. *8*, 226–234. https://doi.org/10.1002/(SICI)1097-0193(1999)8:4<226::AID-HBM6>3.0.CO;2-4.
58. Kahn, I., and Shohamy, D. (2013). Intrinsic connectivity between the hippocampus, nucleus accumbens, and ventral tegmental area in humans. Hippocampus *23*, 187–192. https://doi.org/10.1002/hipo.22077.
59. Das, M., Bennett, D.M., and Dutton, G.N. (2007). Visual attention as an important visual function: an outline of

manifestations, diagnosis and management of impaired visual attention. Br. J. Ophthalmol. *91*, 1556–1560. https://doi.org/10.1136/bjo.2006.104844.

60. Lockhofen, D.E.L., and Mulert, C. (2021). Neurochemistry of visual attention. Front. Neurosci. *15*, 643597. https://doi.org/10.3389/fnins.2021.643597.
61. Junghans, D., Heidenreich, M., Hack, I., Taylor, V., Frotscher, M., and Kemler, R. (2008). Postsynaptic and differential localization to neuronal subtypes of protocadherin β16 in the mammalian central nervous system. Eur. J. Neurosci. *27*, 559–571. https://doi.org/10.1111/j.1460-9568.2008.06052.x.
62. Giordano, G., Cole, T.B., Furlong, C.E., and Costa, L.G. (2011). Paraoxonase 2 (PON2) in the mouse central nervous system: a neuroprotective role? Toxicol. Appl. Pharmacol. *256*, 369–378. https://doi.org/10.1016/j.taap.2011.02.014.
63. Qu, W.R., Sun, Q.H., Liu, Q.Q., Jin, H.J., Cui, R.J., Yang, W., Song, D.B., and Li, B.J. (2020). Role of CPEB3 protein in learning and memory: new insights from synaptic plasticity. Aging *12*, 15169–15182. https://doi.org/10.18632/aging.103404.
64. Moreno-Igoa, M., Hernández-Charro, B., Bengoa-Alonso, A., Pérez-Juana-del-Casal, A., Romero-Ibarra, C., Nieva-Echebarria, B., and Ramos-Arroyo, M.A. (2015). KANSL1 gene disruption associated with the full clinical spectrum of 17q21.31 microdeletion syndrome. BMC Med. Genet. *16*, 68. https://doi.org/10.1186/s12881-015-0211-0.
65. Linda, K., Lewerissa, E.I., Verboven, A.H.A., Gabriele, M., Frega, M., Klein Gunnewiek, T.M., Devilee, L., Ulferts, E., Hommersom, M., Oudakker, A., et al. (2022). Imbalanced autophagy causes synaptic deficits in a human model for neurodevelopmental disorders. Autophagy *18*, 423–442. https://doi.org/10.1080/15548627.2021.1936777.
66. Zhou, X., Wang, L., Xiao, W., Su, Z., Zheng, C., Zhang, Z., Wang, Y., Xu, B., Yang, X., and Hoi, M.P.M. (2019). Memantine improves cognitive function and alters hippocampal and cortical proteome in triple transgenic mouse model of Alzheimer's disease. Exp. Neurobiol. *28*, 390–403. https://doi.org/10.5607/en.2019.28.3.390.
67. Miller, J.E., Shivakumar, M.K., Risacher, S.L., Saykin, A.J., Lee, S., Nho, K., and Kim, D. (2018). Codon bias among synonymous rare variants is associated with Alzheimer's disease imaging biomarker. Pac. Symp. Biocomput. *23*, 365–376.
68. Velayudhan, L., Proitsi, P., Westman, E., Muehlboeck, J.-S., Mecocci, P., Vellas, B., Tsolaki, M., Kłoszewska, I., Soininen, H., Spenger, C., et al. (2013). Entorhinal cortex thickness predicts cognitive decline in Alzheimer's disease. J. Alzheimers Dis. *33*, 755–766. https://doi.org/10.3233/JAD-2012-121408.
69. Kusakari, S., Saitow, F., Ago, Y., Shibasaki, K., Sato-Hashimoto, M., Matsuzaki, Y., Kotani, T., Murata, Y., Hirai, H., Matsuda, T., et al. (2015). Shp2 in forebrain neurons regulates synaptic plasticity, locomotion, and memory formation in mice. Mol. Cell Biol. *35*, 1557–1572. https://doi.org/10.1128/MCB.01339-14.
70. Kim, Y., Liu, G., Leugers, C.J., Mueller, J.D., Francis, M.B., Hefti, M.M., Schneider, J.A., and Lee, G. (2019). Tau interacts with SHP2 in neuronal systems and in Alzheimer's disease. J. Cell Sci. *132*, jcs229054. https://doi.org/10.1242/jcs.229054.
71. Linglart, L., and Gelb, B.D. (2020). Congenital heart defects in Noonan syndrome: diagnosis, management, and treatment. Am. J. Med. Genet. C Semin. Med. Genet. *184*, 73–80. https://doi.org/10.1002/ajmg.c.31765.
72. Xie, X., Liang, M., Yu, C., and Wei, Z. (2021). Liprin-α-Mediated assemblies and their roles in synapse formation. Front. Cell Dev. Biol. *9*, 653381. https://doi.org/10.3389/fcell.2021.653381.
73. Scholz, C.-J., Weber, H., Jungwirth, S., Danielczyk, W., Reif, A., Tragl, K.-H., Fischer, P., Riederer, P., Deckert, J., and Grünblatt, E. (2018). Explorative results from multistep screening for potential genetic risk loci of Alzheimer's disease in the longitudinal VITA study cohort. J. Neural. Transm. *125*, 77–87. https://doi.org/10.1007/s00702-017-1796-6.
74. Bayraktar, A., Lam, S., Altay, O., Li, X., Yuan, M., Zhang, C., Arif, M., Turkez, H., Uhlén, M., Shoaie, S., and Mardinoglu, A. (2021). Revealing the molecular mechanisms of Alzheimer's disease based on network analysis. Int. J. Mol. Sci. *22*, 11556. https://doi.org/10.3390/ijms222111556.
75. Stewart, A.F.R., and Chen, H.-H. (2020). Activation of tyrosine phosphatases in the progression of Alzheimer's disease. Neural Regen. Res. *15*, 2245–2246. https://doi.org/10.4103/1673-5374.284986.
76. Bartok, A., Weaver, D., Golenár, T., Nichtova, Z., Katona, M., Bánsághi, S., Alzayady, K.J., Thomas, V.K., Ando, H., Mikoshiba, K., et al. (2019). IP3 receptor isoforms differently regulate ER-mitochondrial contacts and local calcium transfer. Nat. Commun. *10*, 3726. https://doi.org/10.1038/s41467-019-11646-3.
77. Wang, X., Zhou, X., Uberseder, B., Lee, J., Latimer, C.S., Furdui, C.M., Keene, C.D., Montine, T.J., Register, T.C., Craft, S., et al. (2021). Isoform-specific dysregulation of AMP-activated protein kinase signaling in a non-human primate model of Alzheimer's disease. Neurobiol. Dis. *158*, 105463. https://doi.org/10.1016/j.nbd.2021.105463.
78. Schon, E.A., and Area-Gomez, E. (2013). Mitochondria-associated ER membranes in Alzheimer disease. Mol. Cell. Neurosci. *55*, 26–36. https://doi.org/10.1016/j.mcn.2012.07.011.
79. Zang, F., Zhu, Y., Zhang, Q., Tan, C., Wang, Q., Xie, C.; and Alzheimer's Disease Neuroimaging Initiative* (2021). APOE genotype moderates the relationship between LRP1 polymorphism and cognition across the Alzheimer's disease spectrum via disturbing default mode network. CNS Neurosci. Ther. *27*, 1385–1395. https://doi.org/10.1111/cns.13716.
80. Stoker, A.W. (2015). RPTPs in axons, synapses and neurology. Semin. Cell Dev. Biol. *37*, 90–97. https://doi.org/10.1016/j.semcdb.2014.09.006.
81. Miron, J., Picard, C., Labonté, A., Auld, D., Breitner, J., Poirier, J.; United Kingdom Brain Expression Consortium; and PREVENT-AD research group (2019). Association of PPP2R1A with Alzheimer's disease and specific cognitive domains. Neurobiol. Aging *81*, 234–243. https://doi.org/10.1016/j.neurobiolaging.2019.06.008.
82. Salpietro, V., Dixon, C.L., Guo, H., Bello, O.D., Vandrovcova, J., Efthymiou, S., Maroofian, R., Heimer, G., Burglen, L., Valence, S., et al. (2019). AMPA receptor GluA2 subunit defects are a cause of neurodevelopmental disorders. Nat. Commun. *10*, 3094. https://doi.org/10.1038/s41467-019-10910-w.
83. Pérez-González, M., Mendioroz, M., Badesso, S., Sucunza, D., Roldan, M., Espelosín, M., Ursua, S., Luján, R., Cuadrado-Tejedor, M., and Garcia-Osta, A. (2020). PLA2G4E, a candidate gene for resilience in Alzheimer's disease and a new target for dementia treatment. Prog. Neurobiol. *191*, 101818. https://doi.org/10.1016/j.pneurobio.2020.101818.
84. Tang, X., Yu, L., Yang, J., Guo, W., Liu, Y., Xu, Y., and Wang, X. (2021). Brain-region-specific, glycosylation-related transcriptomic alterations in Alzheimer's disease. J. Clin. Sleep Med. *17*, 779–789. https://doi.org/10.1002/alz.051117.

85. Rebelo, A.L., Chevalier, M.T., Russo, L., and Pandit, A. (2022). Role and therapeutic implications of protein glycosylation in neuroinflammation. Trends Mol. Med. *28*, 270–289. https://doi.org/10.1016/j.molmed.2022.01.004.
86. Flannery, P.J., and Trushina, E. (2019). Mitochondrial dynamics and transport in Alzheimer's disease. Mol. Cell. Neurosci. *98*, 109–120. https://doi.org/10.1016/j.mcn.2019.06.009.
87. Trotter, J., Lee, G.H., Kazdoba, T.M., Crowell, B., Domogauer, J., Mahoney, H.M., Franco, S.J., Müller, U., Weeber, E.J., and D'Arcangelo, G. (2013). Dab1 is required for synaptic plasticity and associative learning. J. Neurosci. *33*, 15652–15668. https://doi.org/10.1523/JNEUROSCI.2010-13.2013.
88. Bracher-Smith, M., Leonenko, G., Baker, E., Crawford, K., Graham, A.C., Salih, D.A., Howell, B.W., Hardy, J., and Escott-Price, V. (2022). Whole genome analysis in APOE4 homozygotes identifies the DAB1-RELN pathway in Alzheimer's disease pathogenesis. Neurobiol. Aging *119*, 67–76. https://doi.org/10.1016/j.neurobiolaging.2022.07.009.
89. Wang, K., Li, M., and Bucan, M. (2007). Pathway-based approaches for analysis of genomewide association studies. Am. J. Hum. Genet. *81*, 1278–1283. https://doi.org/10.1086/522374.
90. Zhao, J., Gupta, S., Seielstad, M., Liu, J., and Thalamuthu, A. (2011). Pathway-based analysis using reduced gene subsets in genome-wide association studies. BMC Bioinf. *12*, 17. https://doi.org/10.1186/1471-2105-12-17.
91. Long, J.M., and Holtzman, D.M. (2019). Alzheimer disease: an update on pathobiology and treatment strategies. Cell *179*, 312–339. https://doi.org/10.1016/j.cell.2019.09.001.
92. Savage, J.E., Jansen, P.R., Stringer, S., Watanabe, K., Bryois, J., de Leeuw, C.A., Nagel, M., Awasthi, S., Barr, P.B., Coleman, J.R.I., et al. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. Nat. Genet. *50*, 912–919. https://doi.org/10.1038/s41588-018-0152-6.
93. Zimmermann, K., Opitz, N., Dedio, J., Renné, C., Müller-Esterl, W., and Oess, S. (2002). NOSTRIN: a protein modulating nitric oxide release and subcellular distribution of endothelial nitric oxide synthase. Proc. Natl. Acad. Sci. USA *99*, 17167–17172. https://doi.org/10.1073/pnas.252345399.
94. Ritchie, S.J., Tucker-Drob, E.M., Cox, S.R., Corley, J., Dykiert, D., Redmond, P., Pattie, A., Taylor, A.M., Sibbett, R., Starr, J.M., and Deary, I.J. (2016). Predictors of ageing-related decline across multiple cognitive functions. Intelligence *59*, 115–126. https://doi.org/10.1016/j.intell.2016.08.007.
95. Salthouse, T.A. (2014). Correlates of cognitive change. J. Exp. Psychol. Gen. *143*, 1026–1048. https://doi.org/10.1037/a0034847.