ARTICLE   OPEN

Check for updates

Association of low-frequency and rare coding variants with information processing speed

Jan Bressler [1,48✉], Gail Davies [2,3,48], Albert V. Smith[4], Yasaman Saba[5], Joshua C. Bis[6], Xueqiu Jian[7], Caroline Hayward [8,9], Lisa Yanek[10], Jennifer A. Smith [11,12], Saira S. Mirza[13,14,15], Ruiqi Wang[16,17], Hieab H. H. Adams [13,18], Diane Becker[10], Eric Boerwinkle[1,19], Archie Campbell[20,21,22], Simon R. Cox [2,3], Gudny Eiriksdottir[4], Chloe Fawns-Ritchie[2,3], Rebecca F. Gottesman [23,24], Megan L. Grove [1], Xiuqing Guo[25], Edith Hofer[26,27], Sharon L. R. Kardia[11], Maria J. Knol [13], Marisa Koini[26], Oscar L. Lopez [28,29], Riccardo E. Marioni[20], Paul Nyquist[23], Alison Pattie[2,3], Ozren Polasek[30,31], David J. Porteous[20,32], Igor Rudan[33], Claudia L. Satizabal [17,34,35], Helena Schmidt[5], Reinhold Schmidt[26], Stephen Sidney[36], Jeannette Simino[37], Blair H. Smith [38], Stephen T. Turner[39], Sven J. van der Lee [13], Erin B. Ware[12], Rachel A. Whitmer[36], Kristine Yaffe[40], Qiong Yang[16,17], Wei Zhao [11], Vilmundur Gudnason[4,41], Lenore J. Launer [42], Annette L. Fitzpatrick[43,44], Bruce M. Psaty[6,44,45], Myriam Fornage [7], M. Arfan Ikram [13], Cornelia M. van Duijn[13,46], Sudha Seshadri[17,34,35], Thomas H. Mosley[47] and Ian J. Deary [2,3]

Measures of information processing speed vary between individuals and decline with age. Studies of aging twins suggest heritability may be as high as 67%. The Illumina HumanExome Bead Chip genotyping array was used to examine the association of rare coding variants with performance on the Digit-Symbol Substitution Test (DSST) in community-dwelling adults participating in the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium. DSST scores were available for 30,576 individuals of European ancestry from nine cohorts and for 5758 individuals of African ancestry from four cohorts who were older than 45 years and free of dementia and clinical stroke. Linear regression models adjusted for age and gender were used for analysis of single genetic variants, and the T5, T1, and T01 burden tests that aggregate the number of rare alleles by gene were also applied. Secondary analyses included further adjustment for education. Meta-analyses to combine cohort-specific results were carried out separately for each ancestry group. Variants in *RNF19A* reached the threshold for statistical significance ($p = 2.01 \times 10^{-6}$) using the T01 test in individuals of European descent. *RNF19A* belongs to the class of E3 ubiquitin ligases that confer substrate specificity when proteins are ubiquitinated and targeted for degradation through the 26S proteasome. Variants in *SLC22A7* and *OR51A7* were suggestively associated with DSST scores after adjustment for education for African-American participants and in the European cohorts, respectively. Further functional characterization of its substrates will be required to confirm the role of *RNF19A* in cognitive function.

INTRODUCTION

Cognitive function can be classified into a number of domains such as reasoning, memory, verbal ability and information processing speed [1]. Measures of processing speed vary between individuals and decline on average with age [2, 3]. Low scores on the Digit-Symbol Substitution Test (DSST), the psychometric test examined in the current analysis, have been associated with both incident mild cognitive impairment and dementia [4, 5]. In addition to being considered as a possible endophenotype for age-related neurological disorders [6, 7] and other psychiatric conditions such as schizophrenia and attention deficit hyperactivity disorder [8, 9], processing speed has sometimes been seen as a relatively basic cognitive function that explains some of the differences in other cognitive abilities [10]. Since heritability evaluated in twin studies has been estimated to be as high as 67% for inter-individual variation in performance on tests of processing speed [11–14], three genome-wide association studies (GWAS)

have been performed to identify common genetic variants that may contribute to this cognitive phenotype [15–17]. To date, there is evidence for genome-wide association with the rs17518584 variant located within an intron of *CADM2* and information processing speed in a sample of 32,070 older adults of European ancestry [17], whereas no significant associations were found in two smaller studies in which there were 4038 participants from four cohorts [15] or 1086 young adults [16]. More recently, associations of intronic single nucleotide polymorphisms (SNPs) in *SH2B3* (rs10849947) and *SPATS2* (rs10931898) and reaction time measured using a computerized game were detected in a GWAS that included 111,483 individuals in UK Biobank. Twenty three genes were also significantly associated with reaction time in the same study in gene-based analyses [18].

An exome genotyping array containing over 200,000 coding variants discovered through exome sequencing in ~12,000 individuals has become available to comprehensively evaluate

rare coding variants. Variants that affect protein structure were selected if they were found in two or more individuals in more than two sequencing projects, and thus collectively, the array represents nearly all non-synonymous coding and splice variation with a > 1:1000 allele frequency in the European population [19]. The goal of this study was to test the hypothesis that rare coding variants in addition to common genetic polymorphisms contribute to scores on a test of processing speed in non-demented community-dwelling adults by combining results across studies participating in the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium [20].

MATERIALS AND METHODS
Study populations
Nine population-based epidemiological cohort studies contributed to the discovery phase of the analysis: Age, Gene/Environment Susceptibility–Reykjavik Study (AGES-Reykjavik); Atherosclerosis Risk in Communities (ARIC) Study; Cardiovascular Health Study (CHS); Coronary Artery Risk Development in Young Adults (CARDIA); CROATIA-Korčula study (Korčula); Generation Scotland: Scottish Family Health Study (GS: SFHS); Genetic Epidemiology Network of Arteriopathy (GENOA); Lothian Birth Cohort 1921 (LBC1921); and Lothian Birth Cohort 1936 (LBC1936); Details for each cohort are described in the Supplementary Material. Written informed consent was obtained from all participants, and the investigators in each of the cohort studies obtained approval from their institutional review board or equivalent committee. All individuals in the study were 45 years or older and determined to be free of stroke and dementia using criteria established within each individual cohort. The total sample size was 36,334 and included 30,576 individuals of European ancestry and 5758 African-Americans. Replication was sought in 1697 individuals of European descent from two independent cohorts: Austrian Stroke Prevention Study (ASPS) and Rotterdam Study (RS).

Genotyping
Most of the study participants were genotyped using the HumanExome Bead Chip v1.0 (Illumina, Inc. San Diego, CA), and variant calling was performed jointly for AGES, ARIC, CARDIA, CHS, GENOA, and RS at the University of Texas Health Science Center at Houston [19]. LBC1921, LBC1936, and CROATIA-Korčula were called in Genome Studio (Illumina Inc.) based on the CHARGE Consortium joint calling cluster file. Quality control procedures included checking concordance with previously collected GWAS genotyping data; exclusion of individuals missing more than 5% of genotypes; population clustering outliers; those with high inbreeding coefficients, rates of heterozygosity, or unexpectedly high identity-by-descent; and participants with gender mismatches. All genetic variants were coded additively with respect to the minor allele in the jointly called dataset. GS:SFHS was genotyped using the HumanExome Bead Chip v1_A and variant calling was performed using GenCall. ASPS genotyping was performed at the Helmholtz Zentrum München using the Illumina HumanExome v1.1 chip and Genome Studio Version V2011.1 software. Samples were excluded if there was contamination with other DNA, sex mismatch, cryptic relatedness, excess heterozygosity, duplicates on the chip or low call rate (<95%). SNPs were excluded based on low call rate (<95%) and Hardy–Weinberg equilibrium p value $< 10^{-6}$.

Cognitive tests
The DSST of the Wechsler Adult Intelligence Scale-Revised (WAIS-R) [21] requires the participant to translate numbers (1–9) to symbols using a key provided at the top of the test page. The WAIS-R was used by ARIC, CHS, and GENOA and was scored as the number of correct translations completed within 90 s. The duration of the version of the test [22] administered by AGES and CROATIA-Korčula was 90 s for AGES, and was 120 s for CROATIA-Korčula. GS, LBC1921, and LBC1936 used the test from the WAIS-III UK [23] with a test duration of 120 s. CARDIA administered the test from the WAIS III [24] with the score calculated as the number correct within 90 s. For all of the DSST tests, a higher score indicates a higher measure of cognition.

Statistical analysis
Single variant and gene-based inverse variance meta-analyses were conducted using the seqMeta package (https://cran.r-project.org/web/packages/seqMeta/seqMeta.pdf). For the single variant analyses,

non-synonymous and splice variants with a minor allele frequency (MAF) > 0.1% were evaluated for both African-American study participants and individuals of European ancestry. In the gene-based analyses using the T5, T1, and T01 tests that aggregate the total number of rare alleles at each locus [25], results were filtered using a cumulative MAF $\geq 0.05\%$, and were limited to genes in which there were at least 2 variants contributing to the test. These three tests incorporated non-synonymous and splice site variants with a MAF of <5%, <1%, and <0.01%, respectively. Two different statistical models were applied in all cohorts. In the first model, linear regression models were adjusted for age, gender, study center if appropriate, family relationship if appropriate, and principal components to correct for population structure with cognitive test scores examined as quantitative traits. The second model included all of the covariates specified in the first model and was further adjusted for educational attainment. Effective sample size (N-weighted) meta-analyses were carried out for individuals of European ancestry due to the different DSST protocols used among the various cohorts and were performed using METAL [26] after initially conducting two separate inverse variance meta-analyses for groups categorized on the basis of test duration (90 s and 120 s). For African-Americans, all cohorts included in the discovery set used a common test duration so inverse variance weighted meta-analysis was performed. A two-sided p value < 0.05/number of single variants (European ancestry: $N = 51{,}043$ variants, $p < 9.8 \times 10^{-7}$; African ancestry: $N = 78{,}701$ variants, $p < 6.4 \times 10^{-7}$) or <0.05/20,000 genes ($p < 2.5 \times 10^{-6}$) was considered statistically significant after Bonferroni correction for multiple comparisons.

Candidate genes
The meta-analysis results for single variants were examined for association with low-frequency polymorphisms in genes previously associated with processing speed [17], or recently identified in studies of Alzheimer's disease using either an exome-wide genotyping array or whole exome sequencing [27–33].

Gene expression, gene-set enrichment, and molecular network analyses
Gene expression in human tissues was assessed using the Genotype Tissue expression portal (GTEx, Broad Institute of MIT and Harvard, Cambridge, MA; (http://www.gtexportal.org/home) [34]. Differential gene expression in multiple brain regions over the human lifespan was explored using data from the Human Brain Transcriptome Project (http://hbatlas.org/pages/hbtd) [35]. Summary statistics from the meta-analyses of the individuals of European ancestry and African ancestry were analyzed separately using FUnctional Mapping and Annotation of genetic associations (FUMA) to explore enrichment in biological pathways [36]. Evidence for over-representation of prioritized genes in gene sets represented in the MsigDB [37] and WikiPathways [38] databases was obtained using the hypergeometric test implemented in the GENE2FUNC function. Genes reaching a p value $\leq 1 \times 10^{-4}$ in either the single variant or gene-based meta-analyses adjusted for age and gender (model 1), or age, gender, and education (model 2) were included in the analyses. The same genes were also considered the focus or input genes in the gene network analyses conducted using the core analysis function in Ingenuity Pathway Analysis (IPA) software (QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis) to generate a set of networks based on the relationships between these and other molecules cataloged in the Ingenuity Knowledge Base [39]. The resulting networks are scored using a right tailed Fisher's exact test and the $-\log_{10}$ (p value) to test the null hypothesis that the association of the focus genes and a set of genes selected from the database and added to the network is due to chance. A score > 1.3 ($p < 0.05$) was chosen as the a priori level of statistical significance. The IPA core analysis function was also used to identify biological functions associated with the focus genes.

RESULTS
The demographic characteristics, mean DSST score, and the DSST test duration for each cohort contributing results to either the discovery or replication meta-analyses are shown in Table 1 stratified by ancestry. Educational attainment was classified into 4–5 categories to include the lowest and highest levels of education.as appropriate in each study (Supplementary Table 1). Quantile-quantile plots revealed no inflation of test statistics for any of the individual cohorts or for the meta-analyses (Supplementary Figs. 1, 2).

Table 1. Discovery and replication cohorts.

Cohort	N[a]	N[b]	Age (years) (mean, SD)	Female (%)	DSST Test duration (seconds)	DSST (digit-symbol pairs) (mean, SD)
Discovery European ancestry						
AGES	2583	2583	75.90 ± 5.49	57.99	90	29.85 ± 14.10
ARIC	10 201	10 201	57.21 ± 5.68	53.39	90	49.09 ± 11.46
CARDIA	1694	1692	50.73 ± 3.37	53.36	90	74.68 ± 14.73
CHS	3754	3754	72.53 ± 5.42	57.03	90	38.89 ± 12.52
GENOA	940	940	59.94 ± 9.48	59.15	90	50.92 ± 12.65
GS	9559	8744	51.86 ± 13.45	59.23	120	70.41 ± 17.02
Korčula	665	629	55.37 ± 10.08	62.86	120	50.75 ± 20.26
LBC1921	245	245	83.35 ± 0.54	58.59	120	42.75 ± 12.98
LBC1936	935	935	69.54 ± 0.84	49.42	120	56.94 ± 12.83
Total	30 576	29 723				
Discovery African ancestry						
ARIC	3103	3103	56.11 ± 5.72	64.45	90	31.33 ± 13.23
CARDIA	1319	1315	49.49 ± 3.84	59.67	90	64.34 ± 15.72
CHS	639	639	72.00 ± 5.03	62.91	90	30.10 ± 12.55
GENOA	697	697	61.82 ± 9.38	72.17	90	33.65 ± 14.28
Total	5758	5754				
Replication European ancestry					LDST Test duration (seconds)	
RS	1561	1546	72.64 ± 6.86	48.94	60	26.33 ± 7.11
ASPS	136	136	70.92 ± 6.86	62.50	60	25.40 ± 6.62
Total	1697	1682				

SD standard deviation, *DSST* Digit-Symbol Substitution Test, *LDST* Letter Digit Substitution Test.
N[a] number of participants contributing to meta-analysis adjusted for age and gender.
N[b] number of participants contributing to meta-analysis adjusted for age, gender, and educational attainment.

Table 2. Gene-based analysis of association of low-frequency variants and DSST scores.

Ancestry	Test	Gene	Chr	CMAF (%)	Beta[a] (SE)	Beta[b] (SE)	Variants (*N*)	Z-Score	*p* value[1]	Z-Score	*p* value[2]
Genome-Wide											
EA	T01	*RNF19A*	8q22.2	0.18	–	–	–	−4.75	2.01×10^{-6}	−4.51	6.31×10^{-6}
Suggestive											
EA	T5/T1	*OR51A7*	11p15.4	0.25	–	–	–	−4.42	9.76×10^{-6}	−4.66	3.13×10^{-6}
Suggestive											
AA	T01	*SLC22A7*	6p21.1	0.078	17.74 (3.97)	15.79 (3.37)	2	–	8.12×10^{-6}	–	2.78×10^{-6}

DSST Digit-Symbol Substitution Test, *EA* European ancestry, *AA* African ancestry, *Chr* chromosome, *CMAF* cumulative minor allele frequency of variants identified in burden test, *SE* standard error, *N* number of variants contributing to burden test, *Z-score* summary value for sample size weighted meta-analysis.
[1]*p* value adjusted for age and gender; [2]*p* value adjusted for age, gender, and educational attainment.
[a]Beta beta coefficient for model adjusted for age and gender.
[b]Beta beta coefficient for model adjusted for age, gender, and educational attainment.

When low-frequency variants were tested individually for association with performance on the DSST, no polymorphic loci that met the a priori significance thresholds were identified for either ancestry group under an additive genetic model.

In the gene-based analyses adjusting for age and gender, there was one genome-wide significant result ($p < 2.5 \times 10^{-6}$) for the ring finger protein 19A, RBR E3 protein ubiquitin ligase (*RNF19A*) gene on chromosome 8q22 that was associated with DSST scores ($p = 2.01 \times 10^{-6}$) in individuals of European ancestry using the T01 test (Table 2, Supplementary Table 2), although this association was attenuated after further adjustment for education ($p = 6.31 \times 10^{-6}$) (Table 2, Supplementary Table 3). *RNF19A* has previously been implicated in Parkinson's disease in which there is slowed information processing [40, 41] and in amyotrophic lateral sclerosis [42, 43]. Examination of the GTEx tissue expression data (Supplementary Fig. 3) revealed that *RNF19A* was most highly expressed in the endocervix, testis, uterus and bladder, whereas it appeared to be transcribed at a relatively low level in all brain regions analyzed. When *RNF19A* was evaluated in the human brain across the lifespan (Supplementary Fig. 4), there was evidence of differential expression in some brain regions. Expression of *RNF19A* was highest in the first year of life in the neocortex and medial nucleus of the thalamus, whereas in the cerebellar cortex the level of *RNF19A* was lowest during the same time period and reached its maximum in early adulthood.

In addition, olfactory receptor family 51 subfamily A member 7 (*OR51A7*; chromosome 11p15.4) was suggestively associated with scores on the DSST using both the T5 test ($p = 3.13 \times 10^{-6}$) and T1 test ($p = 3.13 \times 10^{-6}$) after adjustment for education in the European cohort participants (Table 2, Supplementary Table 3). Two variants in solute carrier family 22 member 7 (*SLC22A7*) contributed to a suggestive association with performance on the

Table 3. Replication analysis of association of genes and DSST scores —RS and ASPS ($N = 1682$).

Gene	Test	Beta (SE)	p^a	CMAF	Variants (N)
RNF19A	T01	−44 (1.92)	0.82	0.0033	4

DSST Digit-Symbol Substitution Test, *Beta* beta coefficient, *SE* standard error, *p* p value, *CMAF* cumulative minor allele frequency of variants identified in burden test, *N* number of variants contributing to burden test.
[a]Adjusted for age and gender.

DSST in African-American individuals using the T01 test only when educational attainment was added as a covariate to the regression models ($\beta = 15.79$ (SE $= 3.37$) number–symbol pairs; $p = 2.78 \times 10^{-6}$) (Table 2, Supplementary Tables 4, 5). The protein encoded by *SLC22A7* is involved in facilitative transport of cGMP and other guanine nucleotides in multiple tissues [44].

Replication was attempted for genes meeting the threshold for genome-wide significance in the gene-based tests after meta-analyzing the results for 1561 participants in the RS and 136 participants in the ASPS that used the Letter Digit Substitution Test (LDST) [45, 46], a related test of processing speed previously shown to be correlated with the DSST ($r = 0.87$, $p < 0.01$) in a group of 102 volunteers [17]. There was no evidence of association with performance on the LDST in these cohorts (Table 3).

In study participants of both ethnicities, the meta-analysis results were examined for rare single nucleotide variants in the gene for cell adhesion molecule 2 (*CADM2*), a locus previously identified in a large GWAS of processing speed [17], and in several candidate genes previously reported to be associated with Alzheimer's disease [27–33]. Among the variants that were included on the exome array, there was one nominally significant association with A-kinase anchoring protein 9 (*AKAP9*) rs149979685 ($\beta = -3.117$ (SE $= 1.538$) number–symbol pairs; $p = 0.043$) in African-Americans after adjusting for age and gender that did not survive adjustment for multiple comparisons (Supplementary Tables 6, 7). The *AKAP9* genetic variant was initially identified by whole exome sequencing in an African-American discovery cohort of Alzheimer's disease cases and controls [27]. In addition, SNPs with a p value $\leq 10^{-4}$ in the meta-analyses of the results for the discovery cohorts in the earlier CHARGE consortium GWAS of processing speed [17] and that were also present on the exome array were evaluated for association with performance on the DSST in participants of both ancestries (Supplementary Table 8). Two exonic variants in ankyrin repeat and kinase domain containing 1 (*ANKK1*) met these criteria in both the meta-analysis adjusted for age and gender and the meta-analysis adjusted for age, gender, and educational attainment, but were not significantly associated with DSST scores in the current study using the exome array genotyping data (all $p \geq 2.98 \times 10^{-3}$).

Gene-set enrichment analysis of genes prioritized using FUMA, and IPA network analysis were also performed separately by ancestry to assess shared biological functions. Enrichment in immune system-related pathways was found for both African-Americans and individuals of European ancestry, consistent with the discovery of a central role for the immune response in genetic studies of the risk of Alzheimer's disease [31, 47]. In the cohorts of European ancestry, *RNF19A* was one of five genes that overlapped with the tested gene set for the Reactome adaptive immune pathway [48]. Enrichment in a KEGG calcium-signaling pathway [49] was seen only in African-Americans (Supplementary Table 9). IPA network analysis in African-Americans revealed that the most highly scored network ($p = 1 \times 10^{-23}$) was associated with cardiovascular disease development and function, organismal development, and tissue morphology and included an indirect interaction between *SLC22A7* and the nonessential amino acid L-glutamic acid. In individuals of European ancestry, the most highly scored network ($p = 1 \times 10^{-25}$) included *RNF19A* and was associated with embryonic development, organismal development, and tissue development. Supplementary Figs. 5, 6 show the network diagrams generated separately by IPA for each ancestry group. In addition, Supplementary Table 10 shows the most significantly associated biological functions related to the focus genes used to produce the molecular networks and that were detected in the IPA core analysis for each ancestry group. The top biological function in the physiological systems development category in African-Americans was nervous system development with 5 associated genes (p value range $= 4.77 \times 10^{-2}$–8.51×10^{-4}) and neurological disease was the second most highly associated biological function in the diseases and disorders category for individuals of European descent with 14 associated genes (p value range $= 2.82 \times 10^{-2}$–4.80×10^{-5}).

DISCUSSION

When the exome array was used to evaluate 30,576 individuals of European descent and 5758 African-Americans, only *RNF19A* was found to exceed the a priori threshold for genome-wide significant association with DSST scores in European ancestry cohorts after conducting the T01 test and adjusting for age and gender. Ubiquitination of proteins targeted for degradation through the 26S proteasome requires the successive activity of an E1 ubiquitination-activation enzyme, an E2 ubiquitin-conjugating enzyme, and an E3 ubiquitin ligase [50, 51]. Mutations in E3 ubiquitin ligases have previously been reported to be associated with both common and rare neurological disorders including autism spectrum disorder and Angelman syndrome [52]. *RNF19A* is a RING finger-type E3 ubiquitin ligase [53] that has been shown to localize to Lewy bodies, a characteristic neuronal inclusion in the brain of patients with Parkinson's disease, and to ubiquitylate synphilin-1. Synphilin-1 was demonstrated to interact in a yeast two-hybrid screen with α-synuclein, another component of Lewy bodies known to cause neuronal degeneration when overexpressed in transgenic flies and mice [42]. *RNF19A* also appears to play a role in familial amyotrophic lateral sclerosis (ALS) by ubiquitylating mutant superoxide dismutase (SOD-1) proteins and promoting their degradation, thereby contributing to the protection of surviving motor neurons [43]. In addition, *Rnf19a*-deficient mice have been found to have reduced adult neurogenesis and enhanced long-term potentiation in the dentate gyrus [54]. There were no genome-wide significant results identified for African-Americans using any of the gene-based tests.

Though it is possible to speculate that efficient quality control of cellular proteins mediated by *RNF19A* is implicated in processing speed in cognitively normal individuals, the identity of its substrate targets and the stage of development during which it may influence cognitive function are currently unknown. Whereas many of the previous reports described above indicate that *RNF19A* is expressed in neurons in humans, their primary focus was the role of *RNF19A* in neuronal inclusions containing insoluble protein aggregates that are not found in the absence of a neurodegenerative disease [42, 43, 53, 55–58]. The results of the network analysis suggest that *RNF19A* may play a role in embryonic development, and interacts with several genes that have been identified in GWAS of either Alzheimer's disease or vascular risk factors associated with cognitive decline in late life. A direct interaction between *RNF19A* and nuclear receptor coactivator 3 (*NCOA3*) was observed. *NCOA3*, a member of the p160 steroid coactivator (SRC) family that modulates transcriptional activation by nuclear receptors in response to hormones, has been implicated in retinoic acid signaling in mouse fetal cortical neurons and is expressed in the murine and human adult brain [59–63]. An intronic variant in *NCOA3* (rs13042367) was recently reported to be associated with HDL-cholesterol levels in a GWAS of circulating lipoproteins [64, 65]. Variants in other interaction partners of *NCOA3* including insulin growth factor 1 (*IGF1* rs5742643), HNF1 homeobox A (*HNF1A* rs1800574 and

rs56348580), and IQ motif containing K (*IQCK* rs7185636) were associated with systolic blood pressure [66, 67], type 2 diabetes [68–72], or Alzheimer's disease [47], respectively. Links to genes implicated in Alzheimer's disease were also found in the results of the gene expression network analysis for *CADM2* identified in the previous CHARGE GWAS of processing speed [17].

The strengths of the study include the well-phenotyped study populations, the representation of individuals of both European and African ancestry, and joint calling of the variants present on the exome array across the participating cohorts. In addition, this is to our knowledge the largest sample size reported for an analysis of rare genetic variants and a single test of processing speed. There are also limitations. The detection of a single gene associated with performance on the DSST suggests that an even larger study may be required to identify additional genome-wide significant findings as has been previously observed for common variants in GWAS of other complex traits, such as height and body mass index [73, 74]. Because the coding and splice site variants present on the exome array are only a subset of the total number of variants in the human genome, and since rare variants found in only one individual were not included by design, it is possible that analysis of whole exome or whole genome sequencing data will be required to fully characterize the role of low-frequency genetic variation in information processing speed.

DATA AVAILABILITY
Summary statistics for the meta-analyses will be available via dbGaP study accession phs000930.v9.p1 (CHARGE (Consortium for Heart and Aging Research in Genomic Epidemiology) Consortium Summary Results from Genomic Studies).

REFERENCES
1. Deary IJ, Penke L, Johnson W. The neuroscience of human intelligence differences. Nat Rev Neurosci. 2010;11:201–11.
2. Salthouse TA. The processing-speed theory of adult age differences in cognition. Psychol Rev. 1996;103:403–28.
3. Der G, Deary IJ. Age and sex differences in reaction time in adulthood: results from the United Kingdom Health and Lifestyle Survey. Psychol Aging. 2006;21:62–73.
4. Lopez OL, Jagust WJ, Dulberg C, Becker JT, DeKosky ST, Fitzpatrick A, et al. Risk factors for mild cognitive impairment in the Cardiovascular Health Study Cognition Study: part 2. Arch Neurol. 2003;60:1394–9.
5. Twamley EW, Ropacki SA, Bondi MW. Neuropsychological and neuroimaging changes in preclinical Alzheimer's disease. J Int Neuropsychol Soc. 2006;12:707–35.
6. Reitz C, Mayeux R. Endophenotypes in normal brain morphology and Alzheimer's disease: a review. Neuroscience. 2009;164:174–90.
7. Wilson RS, Bienias JL, Berry-Kravis E, Evans DA, Bennett DA. The apolipoprotein E epsilon 2 allele and decline in episodic memory. J Neurol Neurosurg Psychiatry. 2002;73:672–7.
8. Dickinson D, Ramsey ME, Gold JM. Overlooking the obvious: a meta-analytic comparison of digit symbol coding tasks and other cognitive measures in schizophrenia. Arch Gen Psychiatry. 2007;64:532–42.
9. Bidwell LC, Willcutt EG, Defries JC, Pennington BF. Testing for neuropsychological endophenotypes in siblings discordant for attention-deficit/hyperactivity disorder. Biol Psychiatry. 2007;62:991–8.
10. Deary IJ. Looking down on human intelligence: from psychometrics to the brain. New York, NY: Oxford University Press; 2000.
11. Swan GE, Carmelli D, Reed T, Harshfield GA, Fabsitz RR, Eslinger PJ. Heritability of cognitive performance in aging twins. The National Heart, Lung, and Blood Institute Twin Study. Arch Neurol. 1990;47:259–62.
12. McClearn GE, Johansson B, Berg S, Pedersen NL, Ahern F, Petrill SA, et al. Substantial genetic influence on cognitive abilities in twins 80 or more years old. Science. 1997;276:1560–3.
13. Finkel D, Pedersen NL. Contribution of age, genes, and environment to the relationship between perceptual speed and cognitive ability. Psychol Aging. 2000;15:56–64.
14. Swan GE, Carmelli D. Evidence for genetic mediation of executive control: a study of aging male twins. J Gerontol B Psychol Sci Soc Sci. 2002;57:P133–43.
15. Luciano M, Hansell NK, Lahti J, Davies G, Medland SE, Räikkönen K, et al. Whole genome association scan for genetic polymorphisms influencing information processing speed. Biol Psychol. 2011;86:193–202.
16. Cirulli ET, Kaspareviciūtė D, Attix DK, Need AC, Ge D, Gibson G, et al. Common genetic variation and performance on standardized cognitive tests. Eur J Hum Genet. 2010;18:815–20.
17. Ibrahim-Verbaas CA, Bressler J, Debette S, Schuur M, Smith AV, Bis JC, et al. GWAS for executive function and processing speed suggests involvement of the CADM2 gene. Mol Psychiatry. 2016;21:189–97.
18. Davies G, Marioni RE, Liewald DC, Hill WD, Hagenaars SP, Harris SE, et al. Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N = 112 151). Mol Psychiatry. 2016;21:758–67.
19. Grove ML, Yu B, Cochran BJ, Haritunians T, Bis JC, Taylor KD, et al. Best practices and joint calling of the HumanExome BeadChip: the CHARGE Consortium. PLoS ONE. 2013;8:e68095.
20. Psaty BM, O'Donnell CJ, Gudnason V, Lunetta KL, Folsom AR, Rotter JI, et al. Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium: Design of prospective meta-analyses of genome-wide association studies from 5 cohorts. Circ Cardiovasc Genet. 2009;2:73–80.
21. Wechsler D. Wechsler adult intelligence scale-revised. New York: Psychological Corporation; 1981.
22. Wechsler D. Manual for the Wechsler Adult Intelligence Scale. New York: Psychological Corporation; 1955.
23. Wechsler D. WAIS-III UK administration and scoring manual. London: Psychological Corporation; 1998.
24. Wechsler D. Wechsler Adult Intelligence Scale-III (WAIS-III). New York: Psychological Corporation; 1997.
25. Li B, Leal SM. Methods for detecting associations with rare variants for common diseases: application to analysis of sequence data. Am J Hum Genet. 2008;83:311–21.
26. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genome-wide association scans. Bioinformatics. 2010;26:2190–1.
27. Logue MW, Schu M, Vardarajan BN, Farrell J, Bennett DA, Buxbaum JD, et al. Two rare AKAP9 variants are associated with Alzheimer's disease in African Americans. Alzheimers Dement. 2014;10:609–18.e11.
28. N'Songo A, Carrasquillo MM, Wang X, Burgess JD, Nguyen T, Asmann YW, et al. African American exome sequencing identifies potential risk variants at Alzheimer disease loci. Neurol Genet. 2017;3:e141.
29. Sims R, van der Lee SJ, Naj AC, Bellenguez C, Badarinarayan N, Jakobsdottir J, et al. Rare coding variants in PLCG2, ABI3, and TREM2 implicate microglial-mediated innate immunity in Alzheimer's disease. Nat Genet. 2017;49:1373–84.
30. Vardarajan BN, Ghani M, Kahn A, Sheikh S, Sato C, Barral S, et al. Rare coding mutations identified by sequencing of Alzheimer disease genome-wide association studies loci. Ann Neurol. 2015;78:487–98.
31. Bis JC, Jian X, Kunkle BW, Chen Y, Hamilton-Nelson KL, Bush WS, et al. Whole exome sequencing study identifies novel rare and common Alzheimer's-Associated variants involved in immune response and transcriptional regulation. Mol Psychiatry.2020;25:1859–75.
32. Raghavan NS, Brickman AM, Andrews H, Manly JJ, Schupf N, Lantigua R, et al. Whole-exome sequencing in 20,197 persons for rare variants in Alzheimer's disease. Ann Clin Transl Neurol. 2018;5:832–42.
33. Patel D, Mez J, Vardarajan BN, Staley L, Chung J, Zhang X, et al. Association of rare coding mutations with Alzheimer disease and other dementias among adults of European ancestry. JAMA Netw Open. 2019;2:e191350.
34. GTEx Consortium. The genotype-tissue expression (GTEx) project. Nat Genet. 2013;45:580–5.
35. Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M, et al. Spatio-temporal transcriptome of the human brain. Nature. 2011;478:483–9.
36. Watanabe K, Taskesen E, van Bochoven A, Posthuma D. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8:1826.
37. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics. 2011;27:1739–40.
38. Kutmon M, Riutta A, Nunes N, Hanspers K, Willighagen EL, Bohler A, et al. WikiPathways: capturing the full diversity of pathway knowledge. Nucleic Acids Res. 2016;44:D488–94.
39. Kramer A, Green J, Pollard J Jr, Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics. 2014;30:523–30.
40. Pate DS, Margolin DI. Cognitive slowing in Parkinson's and Alzheimer's patients: distinguishing bradyphrenia from dementia. Neurology. 1994;44:669–74.
41. Peavy GM, Salmon D, Bear PI, Paulsen JS, Cahn DA, Hofstetter CR, et al. Detection of mild cognitive deficits in Parkinson's disease patients with the WAIS-R NI. J Int Neuropsychol Soc. 2001;7:535–43.
42. Ito T, Niwa J, Hishikawa N, Ishigaki S, Doyu M, Sobue G. Dorfin localizes to Lewy bodies and ubiquitylates synphilin-1. J Biol Chem. 2003;278:29106–14.
43. Niwa J, Ishigaki S, Hishikawa N, Yamamoto M, Doyu M, Murata S, et al. Dorfin ubiquitylates mutant SOD1 and prevents mutant SOD1-mediated neurotoxicity. J Biol Chem. 2002;277:36793–8.

J. Bressler et al.

44. Cropp CD, Komori T, Shima JE, Urban TJ, Yee SW, More SS, et al. Organic anion transporter 2 (SLC22A7) is a facilitative transporter of cGMP. Mol Pharmacol. 2008;73:1151–8.

45. Jolles J, Houx P, van Boxtel M, Ponds R. The Maastricht aging study: determinants of cognitive aging. Maastricht: Neuropsych Publishers; 1995.

46. Lezak M. Neuropsychological assessment. New York: Oxford University Press; 1995.

47. Kunkle BW, Grenier-Boley B, Sims R, Bis JC, Damotte V, Naj AC, et al. Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. Nat Genet. 2019;51:414–30.

48. Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, et al. The Reactome pathway knowledgebase. Nucleic Acids Res. 2014;42:D472–7.

49. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;28:27–30.

50. Ciechanover A, Heller H, Katz-Etzion R, Hershko A. Activation of the heat-stable polypeptide of the ATP-dependent proteolytic system. Proc Natl Acad Sci USA. 1981;78:761–5.

51. Hershko A, Heller H, Elias S, Ciechanover A. Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. J Biol Chem. 1983;258:8206–14.

52. George AJ, Hoffiz YC, Charles AJ, Zhu Y, Mabb AM. A comprehensive atlas of E3 ubiquitin ligase mutations in neurological disorders. Front Genet. 2018;9:29.

53. Niwa J, Ishigaki S, Doyu M, Suzuki T, Tanaka K, Sobue G. A novel centrosomal ring-finger protein, dorfin, mediates ubiquitin ligase activity. Biochem Biophys Res Commun. 2001;281:706–13.

54. Park H, Yang J, Kim R, Li Y, Lee Y, Lee C, et al. Mice lacking the PSD-95-interacting E3 ligase, Dorfin/Rnf19a, display reduced adult neurogenesis, enhanced long-term potentiation, and impaired contextual fear conditioning. Sci Rep. 2015;5:16410.

55. Hishikawa N, Niwa J, Doyu M, Ito T, Ishigaki S, Hashizume Y, et al. Dorfin localizes to the ubiquitylated inclusions in Parkinson's disease, dementia with Lewy bodies, multiple system atrophy, and amyotrophic lateral sclerosis. Am J Pathol. 2003;163:609–19.

56. Takeuchi H, Niwa J, Hishikawa N, Ishigaki S, Tanaka F, Doyu M, et al. Dorfin prevents cell death by reducing mitochondrial localizing mutant superoxide dismutase 1 in a neuronal cell model of familial amyotrophic lateral sclerosis. J Neurochem. 2004;89:64–72.

57. Ishigaki S, Hishikawa N, Niwa J, Iemura S, Natsume T, Hori S, et al. Physical and functional interaction between Dorfin and Valosin-containing protein that are colocalized in ubiquitylated inclusions in neurodegenerative disorders. J Biol Chem. 2004;279:51376–85.

58. Niwa J, Yamada S, Ishigaki S, Sone J, Takahashi M, Katsuno M, et al. Disulfide bond mediates aggregation, toxicity, and ubiquitylation of familial amyotrophic lateral sclerosis-linked mutant SOD1. J Biol Chem. 2007;282:28087–95.

59. Dasgupta S, Lonard DM, O'Malley BW. Nuclear receptor coactivators: master regulators of human health and disease. Annu Rev Med. 2014;65:279–92.

60. Chai Z, Yang L, Yu B, He Q, Li WI, Zhou R, et al. p38 mitogen-activated protein kinase-dependent regulation of SRC-3 and involvement in retinoic acid receptor alpha signaling in embryonic cortical neurons. IUBMB Life. 2009;61:670–8.

61. Tetel MJ, Acharya KD. Nuclear receptor coactivators: regulators of steroid action in brain and behaviour. J Neuroendocrinol. 2013;25:1209–18.

62. Sun Z, Xu Y. Nuclear receptor coactivators (NCOAs) and corepressors (NCORs) in the brain. Endocrinology. 2020;161:1–12.

63. Fagerberg L, Hallström BM, Oksvold P, Kampf C, Djureinovic D, Odeberg J, et al. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol Cell Proteom. 2014;13:397–406.

64. Reitz C, Tang MX, Schupf N, Manly JJ, Mayeux R, Luchsinger JA. Association of higher levels of high-density lipoprotein cholesterol in elderly individuals and lower risk of late-onset Alzheimer disease. Arch Neurol. 2010;67:1491–7.

65. Richardson TG, Sanderson E, Palmer TM, Ala-Korpela M, Ference BA, Davey Smith G, et al. Evaluating the relationship between circulating lipoprotein lipids and apolipoproteins with risk of coronary heart disease: a multivariable Mendelian randomisation analysis. PLoS Med. 2020;17:e1003062.

66. Arvanitakis Z, Capuano AW, Lamar M, Shah RC, Barnes LL, Bennett DA, et al. Late-life blood pressure association with cerebrovascular and Alzheimer disease pathology. Neurology. 2018;91:e517–25.

67. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. Nat Genet. 2018;50:1412–25.

68. Ott A, Stolk RP, van Harskamp F, Pols HA, Hofman A, Breteler MM. Diabetes mellitus and the risk of dementia: the Rotterdam Study. Neurology. 1999;53:1937–42.

69. Knopman D, Boland LL, Mosley T, Howard G, Liao D, Szklo M, et al. Cardiovascular risk factors and cognitive decline in middle-aged adults. Neurology. 2001;56:42–8.

70. Peila R, Rodriguez BL, Launer LJ. Honolulu-Asia Aging Study. Type 2 diabetes, APOE gene, and the risk for dementia and related pathologies: The Honolulu-Asia Aging Study. Diabetes. 2002;51:1256–62.

71. Mahajan A, Wessel J, Willems SM, Zhao W, Robertson NR, Chu AY, et al. Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes. Nat Genet. 2018;50:559–71.

72. Vujkovic M, Keaton JM, Lynch JA, Miller DR, Zhou J, Tcheandjieu C, et al. Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. Nat Genet. 2020;52:680–91.

73. Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. Nat Genet. 2014;46:1173–86.

74. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature. 2015;518:197–206.

AUTHOR CONTRIBUTIONS
JB, GD, and IJD drafted the paper, and IJD, GD, JB, and THM conceived and designed the study. Phenotype data acquisition and quality control were carried out by AVS, GE, RFG, THM, ALF, OLL, KY, S.Sidney, RAW, OP, AC, STT, SLRK, CFR, AP, SRC, and IJD. Genotype data acquisition and quality control were conducted by AVS, EB, MG, HS, JCB, XG, MF, CH, JAS, WZ, SLRK, GD, and REM. Statistical methods and analyses were implemented by AV, JB, EB, THM, YS, JCB, XJ, MF, RW, QY, CH, JAS, WZ, EBW, GD, SS, MJK, SJvdL, and HHHA, LJL, VG, AVS, GE, JB, EB, JS, THM, MF, QY, S.Seshadri, GD, IJD, and CMvD contributed to interpreting the results. All authors critically reviewed and approved the paper.

COMPETING INTERESTS
RFG reports serving as an associate editor for *Neurology*. BMP serves on the Steering Committee of the Yale Open Data Access Project funded by Johnson & Johnson. All other authors declare no competing interests.

ADDITIONAL INFORMATION
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41398-021-01736-6.

Correspondence and requests for materials should be addressed to Jan Bressler.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

[1]Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA. [2]Lothian Birth Cohorts, University of Edinburgh, Edinburgh, UK. [3]Department of Psychology, University of Edinburgh, Edinburgh, UK. [4]Icelandic Heart Association, Kopavogur, Iceland. [5]Gottfried Schatz Research Center for Cell Signaling, Metabolism and Aging, Medical University of Graz, Graz, Austria. [6]Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA. [7]Brown Foundation Institute of Molecular Medicine, University of Texas Health Science Center at Houston, Houston, TX, USA. [8]Medical Research Council Human Genetics Unit, Institute of Genetics and Cancer, University of Edinburgh, Edinburgh, UK. [9]Generation Scotland, Centre for Genomic and Experimental Medicine, Institute of Genetics and Cancer, University of Edinburgh, Edinburgh, UK. [10]Department of General Internal Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA. [11]Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, USA. [12]Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, USA. [13]Department of Epidemiology, Erasmus MC University Medical Center, Rotterdam, The Netherlands. [14]Department of Neurology, Sunnybrook Health Sciences Centre, Toronto, ON, Canada. [15]Department of Medicine, University of Toronto, Toronto, ON, Canada. [16]Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA. [17]Framingham Heart Study, Framingham, MA, USA. [18]Department of Neurology and Alzheimer Center, Amsterdam Neuroscience, Vrieje Universiteit Medical Center, Amsterdam, The Netherlands. [19]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA. [20]Centre for Genomic and Experimental Medicine, Institute of Genetics and Cancer, University of Edinburgh, Edinburgh, UK. [21]Usher Institute, University of Edinburgh, Edinburgh, UK. [22]Health Data Research UK, University of Edinburgh, Edinburgh, UK. [23]Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, USA. [24]Department of Epidemiology, Johns Hopkins University Bloomberg School of Public Health, Baltimore, MD, USA. [25]The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, The Lundquist Institute for Biomedical Innovation at Harbor-UCLA Medical Center, Torrance, CA, USA. [26]Clinical Division of Neurogeriatrics, Department of Neurology, Medical University of Graz, Graz, Austria. [27]Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Graz, Austria. [28]Department of Neurology, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. [29]Department of Psychiatry, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. [30]Faculty of Medicine, University of Split, Split, Croatia. [31]Gen-Info Ltd, Zagreb, Croatia. [32]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK. [33]Centre for Global Health Research, Usher Institute for Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, UK. [34]Glenn Biggs Institute for Alzheimer's and Neurodegenerative Diseases, UT Health San Antonio, San Antonio, TX, USA. [35]Department of Neurology, Boston University School of Medicine, Boston, MA, USA. [36]Kaiser Permanente Northern California Division of Research, Oakland, CA, USA. [37]Department of Data Science, University of Mississippi Medical Center, Jackson, MS, USA. [38]Division of Population Health and Genomics, Ninewells Hospital and Medical School, University of Dundee, Dundee, UK. [39]Division of Nephrology and Hypertension, Mayo Clinic, Rochester, MN, USA. [40]Department of Psychiatry, University of California San Francisco, San Francisco, CA, USA. [41]Faculty of Medicine, University of Iceland, Reykjavik, Iceland. [42]Laboratory of Epidemiology and Population Science, National Institute on Aging, Intramural Research Program, National Institutes of Health, Bethesda, MD, USA. [43]Department of Family Medicine, University of Washington, Seattle, WA, USA. [44]Department of Epidemiology, University of Washington, Seattle, WA, USA. [45]Department of Health Services, University of Washington, Seattle, WA, USA. [46]Department of Public Health, Oxford University, Oxford, UK. [47]Division of Geriatrics, Department of Medicine, University of Mississippi Medical Center, Jackson, MS, USA. [48]These authors contributed equally: Jan Bressler, Gail Davies. ✉email: jan.bressler@uth.tmc.edu