npg

**ORIGINAL ARTICLE**

# GWAS for executive function and processing speed suggests involvement of the *CADM2* gene

CA Ibrahim-Verbaas[1,2,91], J Bressler[3,91], S Debette[4,5,6,91], M Schuur[1,2,91], AV Smith[7,8,91], JC Bis[9,91], G Davies[10,91], S Trompet[11,12], JA Smith[13], C Wolf[14], LB Chibnik[15], Y Liu[16], V Vitart[17], M Kirin[18], K Petrovic[19], O Polasek[20], L Zgaga[21], C Fawns-Ritchie[10], P Hoffmann[22,23,24], J Karjalainen[25], J Lahti[26,27], DJ Llewellyn[28], CO Schmidt[29], KA Mather[30], V Chouraki[31], Q Sun[32], SM Resnick[33], LM Rose[34], C Oldmeadow[35], M Stewart[18], BH Smith[36], V Gudnason[7,8], Q Yang[37,38], SS Mirza[39,40], JW Jukema[11], PL deJager[15], TB Harris[41], DC Liewald[10,42], N Amin[1], LH Coker[43], O Stegle[44], OL Lopez[45], R Schmidt[19], A Teumer[46], I Ford[47], N Karbalai[14], JT Becker[45,48,49], MK Jonsdottir[50], R Au[4,37], RSN Fehrmann[25], S Herms[23,24], M Nalls[51], W Zhao[13], ST Turner[52], K Yaffe[53], K Lohman[16], JC van Swieten[2], SLR Kardia[13], DS Knopman[54], WM Meeks[55], G Heiss[56], EG Holliday[35], PW Schofield[57], T Tanaka[58], DJ Stott[59], J Wang[38], P Ridker[34], AJ Gow[10,42], A Pattie[10], JM Starr[10,60], LJ Hocking[61], NJ Armstrong[30,62,63], S McLachlan[18], JM Shulman[64,65], LC Pilling[66], G Eiriksdottir[7], RJ Scott[35], NA Kochan[30,67], A Palotie[68,69,70], Y-C Hsieh[71], JG Eriksson[27,72,73,74,75], A Penman[76], RF Gottesman[77], BA Oostra[1], L Yu[78], AL DeStefano[4,37,38], A Beiser[4,37,38], M Garcia[41], JI Rotter[79,80,81], MM Nöthen[24,82], A Hofman[39,40], PE Slagboom[83], RGJ Westendorp[84], BM Buckley[85], PA Wolf[4,37], AG Uitterlinden[39,40,86], BM Psaty[9,87,88,89], HJ Grabe[90], S Bandinelli[91], DI Chasman[34], F Grodstein[32], K Räikkönen[26], J-C Lambert[31], DJ Porteous[92], on behalf of Generation Scotland, JF Price[18], PS Sachdev[30,67], L Ferrucci[58], JR Attia[35], I Rudan[18], C Hayward[17], AF Wright[17], JF Wilson[18], S Cichon[23,24,93], L Franke[25], H Schmidt[19], J Ding[94], AJM de Craen[12], M Fornage[95], DA Bennett[78,98], IJ Deary[10,42,98], MA Ikram[2,39,40,96,98], LJ Launer[41,98], AL Fitzpatrick[87,98], S Seshadri[4,37,98], CM van Duijn[1,40,98] and TH Mosley[97,98]

To identify common variants contributing to normal variation in two specific domains of cognitive functioning, we conducted a genome-wide association study (GWAS) of executive functioning and information processing speed in non-demented older adults from the CHARGE (Cohorts for Heart and Aging Research in Genomic Epidemiology) consortium. Neuropsychological testing was available for 5429–32 070 subjects of European ancestry aged 45 years or older, free of dementia and clinical stroke at the time of cognitive testing from 20 cohorts in the discovery phase. We analyzed performance on the Trail Making Test parts A and B, the Letter Digit Substitution Test (LDST), the Digit Symbol Substitution Task (DSST), semantic and phonemic fluency tests, and the Stroop Color and Word Test. Replication was sought in 1311-21860 subjects from 20 independent cohorts. A significant association was observed in the discovery cohorts for the single-nucleotide polymorphism (SNP) rs17518584 (discovery $P$-value $= 3.12 \times 10^{-8}$) and in the joint discovery and replication meta-analysis ($P$-value $= 3.28 \times 10^{-9}$ after adjustment for age, gender and education) in an intron of the gene cell adhesion molecule 2 (*CADM2*) for performance on the LDST/DSST. Rs17518584 is located about 170 kb upstream of the transcription start site of the major transcript for the *CADM2* gene, but is within an intron of a variant transcript that includes an alternative first exon. The variant is associated with expression of *CADM2* in the cingulate cortex ($P$-value $= 4 \times 10^{-4}$). The protein encoded by *CADM2* is involved in glutamate signaling ($P$-value $= 7.22 \times 10^{-15}$), gamma-aminobutyric acid (GABA) transport ($P$-value $= 1.36 \times 10^{-11}$) and neuron cell-cell adhesion ($P$-value $= 1.48 \times 10^{-13}$). Our findings suggest that genetic variation in the *CADM2* gene is associated with individual differences in information processing speed.

*Molecular Psychiatry* advance online publication, 14 April 2015; doi:10.1038/mp.2015.37

[1]Genetic Epidemiology Unit, Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands; [2]Department of Neurology, Erasmus University Medical Center, Rotterdam, The Netherlands; [3]Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA; [4]Department of Neurology, Boston University School of Medicine, Boston, MA, USA; [5]Institut National de la Santé et de la Recherche Médicale (INSERM), U897, Epidemiology and Biostatistics, University of Bordeaux, Bordeaux, France; [6]Department of Neurology, Bordeaux University Hospital, Bordeaux, France; [7]Icelandic Heart Association, Kopavogur, Iceland; [8]Faculty of Medicine, University of Iceland, Reykjavik, Iceland; [9]Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA; [10]Centre for Cognitive Ageing and Cognitive Epidemiology, The University of Edinburgh, Edinburgh, UK; [11]Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands; [12]Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands; [13]Department of Epidemiology, University of Michigan, Ann Arbor, MI, USA; [14]RG Statistical Genetics, Max Planck Institute of Psychiatry, Munich, Germany; [15]Program in Translational Neuropsychiatric Genomics, Department of Neurology, Brigham and Women's Hospital, Boston, MA, USA; [16]Department of Epidemiology, Wake Forest School of Medicine, Winston-Salem, NC, USA; [17]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK; [18]Centre for Population Health Sciences, University of Edinburgh, Edinburgh, UK; [19]Department of Neurology, Medical University and General Hospital of Graz, Graz, Austria; [20]Department of Public Health, University of Split, Split, Croatia; [21]Department of Public Health and Primary Care, Trinity College Dublin, Dublin, Ireland; [22]Institute of Neuroscience and Medicine (INM -1), Research Center Juelich, Juelich, Germany; [23]Division of Medical Genetics, Department of Biomedicine, University of Basel, Basel, Switzerland; [24]Department of Genomics, Life and Brain Research Center, Institute of Human Genetics, University of Bonn, Bonn, Germany; [25]Department of Genetics, University Medical Centre Groningen, University of Groningen, Groningen, The Netherlands; [26]Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland; [27]Folkhälsan Research Centre, Helsinki, Finland; [28]Institute of Biomedical and Clinical Sciences, University of Exeter Medical School, Exeter, UK; [29]Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany; [30]Centre for Healthy Brain Ageing, School of Psychiatry, UNSW Medicine, University of New South Wales, Sydney, Australia; [31]Inserm, U1167, Institut Pasteur de Lille, Université Lille-Nord de France, Lille, France; [32]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA; [33]Laboratory of Behavioral Neuroscience, National Institute on Aging, NIH, Baltimore, MD, USA; [34]Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA, USA; [35]Hunter Medical Research Institute and Faculty of Health, University of Newcastle, Newcastle, NSW, Australia; [36]Medical Research Institute, University of Dundee, Dundee, UK; [37]The National Heart Lung and Blood Institute's Framingham Heart Study, Framingham, MA, USA; [38]Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA; [39]Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands; [40]Netherlands Consortium for Healthy Ageing, Leiden, The Netherlands; [41]Laboratory of Epidemiology and Population Sciences, National Institute on Aging, Bethesda, MD, USA; [42]Department of Psychology, University of Edinburgh, Edinburgh, UK; [43]Division of Public Health Sciences and Neurology, Wake Forest School of Medicine, Winston-Salem, NC, USA; [44]Max Planck Institute for Developmental Biology, Max Planck

# INTRODUCTION

Cognitive function broadly refers to multiple dissociable, but inter-correlated cognitive domains, including memory, language, executive function, processing speed and visuospatial ability. Unimpaired cognitive abilities are an important determinant of quality of life. Impairment of cognitive abilities is seen with dementia, bipolar disorder, schizophrenia and attention deficit hyperactivity disorder.[1–3] Among the cognitive domains, processing speed is considered a fundamental process, reflecting the speed at which cognitive operations are performed.[4,5] Executive function reflects higher-order cognitive capabilities, presumably mediated by the frontal lobes, including response inhibition, attention, cognitive flexibility and planning.[6]

In addition to domain-specific variance, processing speed and executive function are also, in part, explained by an individual's general cognitive ability. The same holds true for the genetic variance of performance on individual cognitive tests.[7] The estimated heritability of general cognitive ability from twin studies of intelligence ranges from approximately 50 to 80%, and appears to increase with age.[8] The general intelligence construct *g* has a heritability of around 29%, for example, observed in a genome-wide association study (GWAS) based on the GCTA procedure.[9,10] Heritability of performance on tests within individual cognitive domains has been estimated from 12 to 68%[11–14] for processing speed and 16–63%[13,15,16] for executive function.

Of note, there is also considerable covariation between cognitive domains.[17,18] There is, for example, debate as to whether processing speed is merely one of the cognitive domains, or whether processing speed has a more unique role as a fundamental process underlying variation in more complex cognitive traits as well as in cognitive aging.[19,20]

Identifying variants that influence quantitative variation in processing speed and executive function may provide insight into the normal variation in these important cognitive functions, and may ultimately increase our understanding of diseases that disrupt these cognitive domains.

Although various genes have been identified as potential candidates affecting different dimensions of cognitive function, prior studies have yielded inconsistent results.[21] Candidate gene meta-analyses have shown associations of the apolipoprotein E (*APOE*) gene[22] and the *DTNBP1* (dystrobrevin binding protein 1) gene[23] to general cognitive ability, although these findings do not meet the current standard of genome-wide significance (*P-value* between 0.01 and 0.05 for *APOE*, *P-value* = 0.003 for *DTNBP1*). Linkage analyses of executive function tasks have identified regions on chromosomes 2q, 5q, 11q, 13q and 14q.[24–26] To our knowledge, there are currently five published GWAS on processing speed and executive traits in adults.[5,27–30] Processing speed was suggestively associated with several loci, of which the *TRIB3* (tribbles homolog 3) gene was the strongest and biologically most interesting.[5] For executive function, one study[30] identified a genome-wide significant association (*P-value* = $4.32 \times 10^{-8}$) of a single-nucleotide polymorphism (SNP) in the *WDR72* gene (chromosome 15) for a cognitive test similar to the Stroop interference test. The *WDR72* gene has also been associated with kidney function.[31] Prior studies of executive function and processing speed have been limited by small sample sizes (700 and 4000 subjects), resulting in limited power, and by application of lenient significance thresholds. Replication of prior findings has been lacking both across and within cognitive domains.

In this study, we performed a large-scale meta-analysis to identify genetic variants associated with executive function and information processing speed combining GWAS from multiple cohorts of non-demented middle-aged and older adults from the CHARGE (Cohorts for Heart and Aging Research in Genomic Epidemiology) consortium.[32]

# MATERIALS AND METHODS

## Study populations

The discovery phase included 20 cohorts contributing to one or more test (*N* per test = 5429–32 070) (Supplementary Table 1). The number of discovery cohorts and subjects varied based on the availability of each test.

Institute for Intelligent Systems, Tübingen, Germany; [45]Department of Neurology, University of Pittsburgh, Pittsburgh, PA, USA; [46]Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Greifswald, Germany; [47]Robertson Center for biostatistics, University of Glasgow, Glasgow, UK; [48]Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, USA; [49]Department of Psychology, University of Pittsburgh, Pittsburgh, PA, USA; [50]Landspitali Hospital, Reykjavik, Iceland; [51]Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD, USA; [52]Division of Nephrology and Hypertension, Department of Internal Medicine, Mayo Clinic, Rochester, MN, USA; [53]Departments of Psychiatry, Neurology and Epidemiology, University of California, San Francisco and San Francisco VA Medical Center, San Francisco, CA, USA; [54]Department of Neurology, Mayo Clinic, Rochester, MN, USA; [55]Department of Medicine, Division of Geriatrics, University of Mississippi Medical Center, Jackson, MS, USA; [56]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA; [57]School of Medicine and Public Health, Faculty of Health, University of Newcastle, Newcastle, SW, Australia; [58]Translational Gerontology Branch, National Institute on Aging, Baltimore, MD, USA; [59]Department of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK; [60]Alzheimer Scotland Research Centre, Edinburgh, UK; [61]Division of Applied Medicine, University of Aberdeen, Aberdeen, UK; [62]Cancer Research Program, Garvan Institute of Medical Research, Sydney, NSW, Australia; [63]School of Mathematics & Statistics and Prince of Wales Clinical School, University of New South Wales, Sydney, NSW, Australia; [64]Department of Neurology, Baylor College of Medicine, Houston, TX, USA; [65]Department of Molecular and Human Genetics, The Jan and Dan Duncan Neurological Research Institute, Texas Children's Hospital, Houston, TX, USA; [66]Epidemiology and Public Health Group, University of Exeter Medical School, Exeter, UK; [67]Neuropsychiatric Institute, The Prince of Wales Hospital, Sydney, NSW, Australia; [68]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridge, UK; [69]Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland; [70]Department of Medical Genetics, University of Helsinki and University Central Hospital, Helsinki, Finland; [71]School of Public Health, Taipei Medical University, Taipei, Taiwan; [72]Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland; [73]National Institute for Health and Welfare, Helsinki, Finland; [74]Helsinki University Central Hospital, Unit of General Practice, Helsinki, Finland; [75]Vasa Central Hospital, Vasa, Finland; [76]Center of Biostatistics and Bioinformatics, University of Mississippi Medical Center, Jackson, MS, USA; [77]Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, USA; [78]Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, IL, USA; [79]Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, CA, USA; [80]Institute for Translational Genomics and Population Sciences, Los Angeles BioMedical Research Institute at Harbor-UCLA Medical Center, Torrance, CA, USA; [81]Division of Genetic Outcomes, Department of Pediatrics, Harbor-UCLA Medical Center, Torrance, CA, USA; [82]German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany; [83]Department of Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands; [84]Leiden Academy of Vitality and Ageing, Leiden, The Netherlands; [85]Department of Pharmacology and Therapeutics, University College Cork, Cork, Ireland; [86]Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, The Netherlands; [87]Department of Epidemiology, University of Washington, Seattle, WA, USA; [88]Department of Health Services, University of Washington, Seattle, WA, USA; [89]Group Health Research Institute, Group Health, Seattle, WA, USA; [90]Department of Psychiatry and Psychotherapy, University Medicine Greifswald, HELIOS-Hospital Stralsund, Stralsund, Germany; [91]Geriatric Unit, Azienda Sanitaria Firenze (ASF), Florence, Italy; [92]Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK; [93]Institute of Neuroscience and Medicine (INM-1), Research Center Juelich, Juelich, Germany; [94]Department of Internal Medicine, Wake Forest University School of Medicine, Winston-Salem, NC, USA; [95]Institute for Molecular Medicine and Human Genetics Center, University of Texas Health Science Center at Houston, Houston, TX, USA; [96]Department of Radiology, Erasmus University Medical Center, Rotterdam, The Netherlands and [97]Department of Medicine and Neurology, University of Mississippi Medical Center, Jackson, MS, USA. Correspondence: Dr TH Mosley, Department of Medicine (Geriatrics), University of Mississippi Medical Center, 2500 North State Street, Jackson, 39216-4505 MS, USA.
E-mail: tmosley@umc.edu
[99]These authors contributed equally to this work.
Received 18 December 2013; revised 21 January 2015; accepted 11 February 2015

Each cohort had extensive phenotypic data on one or more traits, and genome-wide SNP data available. Details for each cohort are given in Supplementary Table 1. Subjects aged 45 and older who were free of stroke and dementia and of European ancestry were eligible for the study.

Twenty replication cohorts (total $N = 1311–21\,860$) (Supplementary Table 1) were selected based upon comparability of the study populations to the discovery cohorts, and availability of genotype data, and these cohorts were invited to share data from a GWAS run according to the same protocols. Applying the same inclusion criteria for age ($\geqslant 45$ years) and the absence of stroke or dementia, we also included cohorts of African-American ancestry ($N = 1004–3164$ depending on the trait) in the replication phase, partly to evaluate whether our findings could be extrapolated to persons of other ethnicities. Data from the discovery and replication studies were further meta-analyzed. Additional details are provided in the Supplementary Material. Each participant provided informed consent and all studies were approved by their local Institutional Review Boards.

## Executive function and processing speed tests

Each cohort included some tests of executive function and/or processing speed. Test batteries differed across cohorts. Tests of executive function included the Trail Making Test part B (Trail B), Stroop card 3 (color word card) and tests of phonemic (letters, see Supplementary Material) and semantic (animals) fluency. Tests of processing speed included Trail Making Test part A (Trail A), Stroop card 2 (color card), the Digit-Symbol Substitution Test (DSST), Letter-Digit Substitution Task (LDST), the Letter Digit Coding Test (LDCT) and the Symbol Digit Modalities Test (SDMT). All tests were administered in a standardized manner by an investigator unaware of any genetic information on the subjects. Details for the test administration and raw scores are given in Supplementary Table 1 and in the Supplementary Methods. Because the nature of the task and procedures for DSST, LDST, LDCT and SDMT were so similar, these measures were combined (that is, treated as the same test; referred to hereafter as DSST/LDST) in analysis. To assess the validity of combining these measures into a single meta-analysis, we conducted a substudy to determine the correlation between these measures in 102 volunteers not included in the meta-analysis cohorts. Details of the substudy are provided in the Supplementary Material, section 6 and Supplementary Table 5.

## Genotyping and imputation

Genotyping was independently performed by each cohort using commercially available arrays ranging from the Illumina 300 and 610 k to the Affymetrix GeneChip SNP Array 6.0 (Supplementary Table 2). Each cohort applied standard quality control filtering before genetic imputations. These filters included a SNP call rate of at least 90%, sample call rate of at least 92%, minor allele frequency of at least 0.01, and Hardy–Weinberg deviation *P-value* of at most $10^{-3}$ (Supplementary Table 3). Genetic data imputations were performed in each cohort using HapMap II CEPH (Utah residents with ancestry from northern and western Europe) as the reference panel, with the exceptions of ARIC and GENOA African-American cohorts that imputed their genetic data using both HapMap II CEPH and YRI (Yoruba in Ibadan) populations as a reference.[33] Details on genotyping and imputations are provided in Supplementary Tables 2–4.

## Genome-wide association analysis

Each cohort performed a linear regression model of test scores against the dosage of coded alleles for each SNP testing an additive effect of the genetic variants. Skewness and kurtosis were evaluated before the analyses. A detailed description of the screening for latent population substructure in each cohort is given in part 5 of the Supplementary Material and Supplementary Table 3. Where appropriate, principal components or family-specific methods such as mixed model score were applied to correct for familial relationships (Supplementary Material, part 5).

We used two models of association analysis; with and without education level as a covariate in addition to age, gender and other study specific confounders, for example, study site, familial relations or population substructure. This is because there is a dynamic, two-way relationship between cognitive function and level of education. The ability to score high on cognitive tests is influenced by background familiarity with, for example, the numbers and alphabet (Trails tests) or a large vocabulary (fluency tests), which are typically acquired during one's formal education. Estimates are that, even in the Netherlands (10%) and the USA (22%), functional illiteracy is common among adults aged 16–65.[34,35]

Conversely, there is a genetic correlation between education and cognitive ability.[36]

## Meta-analysis

The meta-analysis was performed in METAL.[37] For all tests except the Stroop test and LDST/DSST, meta-analysis was performed using the inverse variance method. For the Stroop test and LDST/DSST, a sample size weighted meta-analysis was performed because of the differences in the test methodology and measurement units that impeded the pooling of the beta coefficients. The z-statistic was weighted by the effective sample size (sample size × (observed dosage variance/expected dosage variance)) for each SNP. Genomic control was applied within each cohort before meta-analysis. The meta-analyses were restricted to autosomal SNPs common to all studies for each neurocognitive test.

## Additional analyses

*Expression quantitative trait locus analyses.* All variants with a discovery *P-value* $< 5 \times 10^{-6}$ were analyzed further to test whether they were associated with RNA expression. For this, we used the Genotype Tissue expression portal (GTex, Broad Institute, Boston, MA, USA; http://www.gtexportal.org/home/) to assess the variants for their influence on expression of their closest genes in brain tissue.[38] We also performed expression quantitative trait locus (eQTL) experiments in 138 human hippocampal cell lines obtained *in vivo* from patients undergoing surgery for treatment-resistant epilepsy. Details on the methods of this functional follow-up study are given in the Supplementary Material. Cognitive tests are never completely limited to a single cognitive domain. Even in executive function and processing speed tests, working memory (and other capabilities) may have a role. For example, one needs to remember the assignment and which words were already used in fluency tasks, and those who are better able to passively learn the key in the LDST or DSST should likewise produce a higher (better) score. Thus, some hippocampal involvement might be expected along with the predominant frontal processes of executive function and processing speed. Hence we used the hippocampal cell lines to identify associations between whole-genome SNP (Illumina Human 660 W array) and RNA expression data (Illumina HumanHT-12v3). The Bonferroni corrected level of eQTL significance for a given SNP can be assumed to be 0.05/n of genes for a given SNP and a given test, leading to a *P-value* of $0.05/20\,000 = 2.5 \times 10^{-6}$ for the hypothesis-free, genome-wide search for eQTL associations.

*Gene network and functional prediction analysis.* A gene network analysis was performed using the Gene Network database (http://www.genenetwork.nl/genenetwork, by Groningen University, The Netherlands) (Fehrmann et al., manuscript in preparation), which incorporates gene expression data from 77 840 human, mouse and rat Affymetrix microarrays from the Gene Expression Omnibus.[39,40] Using principal components analysis on the probeset correlation matrices, so-called transcriptional components were identified that describe major biological pathways. We combined these data into a multi-species gene network with 19 997 unique human genes. Predictions of gene function were made using biological databases such as the Kyoto Encyclopedia of Genes and Genomes (KEGG, www.genome.jp/kegg),[41] Gene Ontology (GO, www.geneontology.org) Database[42] and Reactome pathway database (www.reactome.org).[43] Additionally, we consulted the GTex website for tissue expression data on the genes of interest. More detailed information on the gene network approach is given in the Supplementary Material.

## RESULTS

### Phenotypes

Analyses in an independent group of 102 volunteers showed correlations between LDST, DSST and SDMT above 0.8 (*P-value* $< 0.01$). After partialling out age, correlations remained above 0.7 (*P-value* $< 0.01$). (Supplementary Table 6 and Supplementary Figure 1). We therefore deemed it appropriate to combine these three tests into one sample size-weighted meta-analysis.

### GWAS and eQTL analysis

Baseline characteristics of the study populations and mean test results for each trait are provided in Supplementary Table 1.

Supplementary Figure 2 shows the quantile-quantile plots for the discovery-phase GWAS meta-analysis of each trait. No inflation due to hidden substructure or cryptic relatedness was observed for any cohort.

In the meta-analysis for processing speed (LDST/DSST tests), only one genome-wide significant association was observed at an intronic variant (rs17518584, $P$-value $= 3.12 \times 10^{-8}$ in the model adjusted for age, sex and education and $P$-value $= 6.25 \times 10^{-7}$ in the model adjusted for age and sex) in the gene encoding cell adhesion molecule 2 (*CADM2*) on chromosome 3 (Supplementary Figure 3 and Table 1). This gene is also known as *SYNCAM2* (synaptic cell adhesion molecule 2). This variant explains 0.05% (age, sex and education-adjusted model) of the variance in scores on the LDST/DSST in the Rotterdam Study, one of the largest population-based cohorts.

Additional information was sought for all SNPs with a discovery $P$-value of below $5 \times 10^{-6}$ and a minor allele frequency of $> 0.05$. At this phase, 7 SNPs were analyzed for LDST/DSST, 15 SNPs for the Stroop test, 12 SNPs for Trail A, 17 SNPs for Trail B, 9 SNPs for letter fluency and 34 SNPs for semantic fluency. Of all the loci tested in the independent cohorts, including both subjects of European and African-American ancestries, a nominally significant association in the same direction as in the discovery analyses was observed only for the *CADM2* variant rs17518584 with LDST/DSST (nominal $P$-value $= 0.03$ for meta-analyses of results in the additional cohorts, adjusting for age and sex, and nominal $P$-value $= 0.05$ for meta-analysis of results after adjusting for age, sex and education). Meta-analysis of discovery and replication cohorts yielded genome-wide significant evidence for association of rs17518584 in *CADM2* with scores for the LDST/DSST processing speed tests, in the fully-adjusted model that included age, sex and education ($P$-value $= 3.28 \times 10^{-9}$) (Table 1). The findings of the individual studies (both discovery and replication) are shown in Table 2.

When examining association signals for loci previously identified in genetic analyses of cognitive function, rs7412 and rs429358 used to genotype the *APOE* ε4 allele associated with general cognition[22] were not present in the analyzed SNP sets. However, a single SNP proxy for *APOE* ε4 (rs4420638)[44] was associated with the LDST/DSST ($P$-value $= 2.11 \times 10^{-4}$) in the fully adjusted model. We found no effect of additional adjustments for rs4420638 on the association of rs17518584 with the LDST/DSST in the Rotterdam study cohorts. There was no evidence for association of SNPs in any of the other tested candidate genes that have previously been associated with executive and processing speed functions. (*DTNBP1:* lowest $P$-value $= 0.019$ for rs2619522 for the Stroop test, *TRIB3:* lowest $P$-value $= 0.155$ for phonemic fluency, *WDR72* lowest $P$-value $= 0.157$ for LDST/DSST in our GWAS meta-analysis for the fully adjusted model). Also, regions for which linkage was reported (2q, 5q, 11q, 13q and 14q) were not genome-wide significant. The strongest association was seen in the 11q.25 region in which we found a SNP rs2734839 to be suggestively associated with performance on the LDST/DSST ($P$-value $= 4.39 \times 10^{-7}$ for the age- and sex-adjusted model and $3.08 \times 10^{-6}$ for the model adjusted for age, sex and education). Rs2734839 is an intronic SNP in the *DRD2* gene encoding the D2 subtype of the dopamine receptor.

Bioinformatics analysis
The GTex tissue expression data (www.broadinstitute.org/gtex) show that *CADM2* is expressed more abundantly in different areas of the brain than in any other tissue, and most specifically in the frontal and anterior cingulate cortex (Supplementary Figure 4). In the GTex eQTL analysis, the top hit from the LDST/DSST GWAS, rs17518584, showed association with RNA expression levels of *CADM2* in the cingulate cortex ($P$-value $= 4 \times 10^{-4}$), general hemispheral cortex ($P$-value $= 0.004$), substantia nigra ($P$-value $= 0.007$), frontal cortex ($P$-value $= 0.02$) and cerebellum

**Table 1.** SNPs with $P$-value $< 10^{-6}$ in original meta-analysis

| Trait | SNP | Chr | Position | Gene | Feature | A1 | A2 | Average EAF | Discovery No. of subjects | Discovery Beta (SE) Age, sex adjusted | Discovery P-value Age, sex adjusted | Discovery Beta (SE) Age, sex, education adjusted | Discovery P-value Age, sex, education adjusted | Replication No. of subjects | Replication Beta (SE) Age, sex adjusted | Replication P-value Age, sex adjusted | Replication Beta (SE) Age, sex, education adjusted | Replication P-value Age, sex, education adjusted | Combined Beta (SE) Age, sex adjusted | Combined P-value Age, sex adjusted | Combined Beta (SE) Age, sex education adjusted | Combined P-value Age, sex education adjusted |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *Tests with summary values in beta (SE)* | | | | | | | | | | | | | | | | | | | | | | |
| Trail A | rs9514964 | 13 | 109679886 | MYO16 | Intron | A | G | 0.17 | 5429 | 0.048 (0.009) | 3.01E-07 | 0.044 (0.009) | 2.35E-06 | 9553 | -0.006 (0.007) | 0.37 | 0.007 (0.007) | 0.32 | 0.029 (0.007) | 1.55E-04 | 0.020 (0.006) | 3.03E-04 |
| | rs9559465 | 13 | 109683883 | MYO16 | Intron | A | G | 0.84 | | -0.048 (0.010) | 3.30E-07 | -0.043 (0.009) | 2.59E-06 | | 0.006 (0.008) | 0.43 | -0.007 (0.008) | 0.36 | -0.035 (0.008) | 1.58E-04 | -0.021 (0.006) | 2.35E-04 |
| | rs1230154 | 4 | 99988659 | METAP1, ADH5 | Intergenic | T | C | 0.71 | | -0.038 (0.007) | 7.69E-07 | -0.036 (0.007) | 1.64E-06 | | 0.003 (0.006) | 0.63 | 0.001 (0.006) | 0.93 | -0.030 (0.007) | 9.34E-04 | -0.013 (0.005) | 4.83E-03 |
| Trail B | rs11082233 | 18 | 37265119 | PIK3C3, KC6 | Intergenic | T | C | 0.11 | 6210 | 0.073 (0.013) | 6.95E-08 | 0.061 (0.013) | 2.28E-06 | 10817 | -0.012 (0.010) | 0.25 | -0.017 (0.010) | 0.09 | 0.020 (0.008) | 1.73E-02 | 0.0123 (0.0079) | 0.12 |
| | rs6583634 | 8 | 143546590 | BAI1 | Intron | T | C | 0.15 | 13454 | 1.046 (0.209) | 5.75E-07 | 0.914 (0.194) | 2.47E-06 | 12734 | -0.162 (0.230) | 0.48 | -0.009 (0.164) | 0.96 | 0.498 (0.155) | 1.00E-03 | 0.376 (0.125) | 0.0027 |
| Phonemic fluency | rs10481393 | 8 | 143556168 | BAI1 | Intron | T | C | 0.15 | | 1.050 (0.210) | 5.96E-07 | 0.914 (0.195) | 2.78E-06 | 21860 | 0.068 (0.479) | 0.88 | -0.015 (0.183) | 0.937 | 0.891 (0.193) | 3.69E-06 | 0.420 (0.134) | 0.0016 |
| Semantic fluency | rs6587905 | 1 | 61093597 | C1orf87, NFIA | Intergenic | T | C | 0.16 | 6383 | 0.664 (0.129) | 2.80E-07 | 0.656 (0.126) | 2.09E-07 | | 0.091 (0.099) | 0.36 | 0.138 (0.077) | 0.071 | 0.303 (0.079) | 1.16E-04 | 0.280 (0.065) | 1.90E-05 |
| | rs10115337 | 9 | 88711574 | LOC392358, GAS1 | Intergenic | T | C | 0.18 | | -0.494 (0.100) | 8.02E-07 | -0.478 (0.098) | 9.88E-07 | | -0.058 (0.079) | 0.46 | -0.083 (0.059) | 0.159 | -0.224 (0.062) | 2.88E-04 | -0.189 (0.050) | 1.65E-04 |
| | rs10780801 | 9 | 88710941 | LOC392358, GAS1 | Intergenic | T | C | 0.82 | | 0.493 (0.100) | 8.18E-07 | 0.477 (0.101) | 1.01E-06 | | 0.055 (0.079) | 0.49 | 0.081 (0.059) | 0.166 | 0.223 (0.062) | 3.19E-04 | 0.189 (0.050) | 1.79E-04 |
| *Tests with summary values in Z-scores* | | | | | | | | | | Z-score | P-value | Z-score | P-value | | Z-score | P-value | Z-score | P-value | Z-score | P-value | Z-score | P-value |
| Stroop | rs664154 | 6 | 19099538 | RP1-239K6.1 | Intergenic | T | C | 0.24 | 12866 | 5.22 | 1.84 E-07 | 4.79 | 1.67E-06 | 6381 | -0.96 | 0.34 | -1.06 | 0.29 | 5.03 | 5.01E-07 | 4.57 | 4.90E-06 |
| LDST/DSST | rs17518584 | 3 | 85687613 | CADM2 | Intron | T | C | 0.64 | 32070 | 4.98 | 6.25 E-07 | 5.54 | 3.12E-08 | 1311 | 2.23 | 0.03 | 1.95 | 0.05 | 5.43 | 5.56E-08 | 5.92 | 3.28E-09 |
| | rs2734839 | 11 | 112791700 | DRD2 | Intron | T | C | 0.61 | | 4.99 | 5.92 E-07 | 4.86 | 1.18E-06 | | 0.55 | 0.58 | 0.20 | 0.84 | 5.13 | 2.95E-07 | 4.91 | 9.18E-07 |
| | rs2118666 | 2 | 222808767 | PAX3 | Intron | T | G | 0.46 | | -4.95 | 7.33 E-07 | -4.25 | 2.17E-05 | | 0.10 | 0.92 | -0.07 | 0.94 | -4.95 | 7.45E-07 | -4.26 | 2.08E-05 |

Abbreviations: A1, effect allele; A2, reference allele; Beta, beta coefficient of effect allele; Chr, Chromosome; EAF, effect allele frequency; SE, standard error of beta coefficient; LDST/DSST, Letter-Digit or Digit-Symbol Substitution Task; SNP, single-nucleotide polymorphism.

**Table 2.** Effect size and direction of effects across cohorts for rs17518584

| Study | Test | Time (seconds) | Coded allele | Non-coded allele | Beta Coeff./SE (age, sex adjusted) | P-value | Beta Coeff./SE (age, sex, education adjusted) | P-value |
|---|---|---|---|---|---|---|---|---|
| Discovery | | | | | | | | |
| ARIC | DSST[53] | 90 | T | C | 0.235 (0.157) | $1.34 \times 10^{-1}$ | 0.275 (0.146) | $5.97 \times 10^{-2}$ |
| CHS | | 90 | T | C | 0.317 (0.387) | $4.12 \times 10^{-1}$ | 0.469 (0.358) | $1.90 \times 10^{-1}$ |
| GENOA | | 90 | T | C | 0.503 (0.550) | $3.61 \times 10^{-1}$ | 0.553 (0.534) | $3.02 \times 10^{-1}$ |
| ORCADES | | 90 | T | C | 1.516 (1.101) | $1.68 \times 10^{-1}$ | 1.541 (1.054) | $1.44 \times 10^{-1}$ |
| AGES | DSST[54] | 90 | T | C | 0.682 (0.282) | $1.56 \times 10^{-2}$ | 0.808 (0.259) | $1.84 \times 10^{-3}$ |
| Korcula | | 120 | T | C | − 0.350 (0.967) | $7.18 \times 10^{-1}$ | -0.161 (0.838) | $8.48 \times 10^{-1}$ |
| Split | | 60 | T | C | 0.495 (0.664) | $4.56 \times 10^{-1}$ | 0.549 (0.588) | $3.51 \times 10^{-1}$ |
| Vis | | 120 | T | C | 0.725 (1.155) | $5.31 \times 10^{-1}$ | 0.880 (1.089) | $4.19 \times 10^{-1}$ |
| Health ABC | DSST[55] | 90 | T | C | 0.448 (0.452) | $3.22 \times 10^{-1}$ | 0.458 (0.432) | $2.89 \times 10^{-1}$ |
| LBC 1921 | DSST[56] | 120 | T | C | 0.755 (1.178) | $5.22 \times 10^{-1}$ | 0.804 (1.140) | $4.80 \times 10^{-1}$ |
| LBC 1936 | | 120 | T | C | 1.074 (0.630) | $8.84 \times 10^{-2}$ | 0.776 (0.611) | $2.04 \times 10^{-1}$ |
| ASPS | LDST[57] | 60 | T | C | 0.850 (0.600) | $1.55 \times 10^{-1}$ | 0.545 (0.580) | $3.44 \times 10^{-1}$ |
| RS | | 60 | T | C | 0.176 (0.174) | $3.12 \times 10^{-1}$ | 0.224 (0.165) | $1.76 \times 10^{-1}$ |
| RS2 | | 60 | T | C | 0.370 (0.212) | $8.08 \times 10^{-2}$ | 0.329 (0.205) | $1.09 \times 10^{-1}$ |
| RS3 | | 60 | T | C | 0.648 (0.216) | $2.72 \times 10^{-3}$ | 0.599 (0.208) | $3.91 \times 10^{-3}$ |
| MAP | SDMT[58] | 90 | T | C | − 1.214 (0.530) | $2.15 \times 10^{-2}$ | -1.193 (0.520) | $2.14 \times 10^{-2}$ |
| ROS | | 90 | T | C | 1.154 (0.462) | $1.22 \times 10^{-2}$ | 1.218 (0.450) | $6.58 \times 10^{-3}$ |
| PROSPER | LDCT[59] | 60 | T | C | 0.336 (0.155) | $3.04 \times 10^{-2}$ | 0.303 (0.148) | $4.07 \times 10^{-2}$ |
| Replication | | | | | | | | |
| BLSA | DSST[53] | 90 | T | C | − 0.119 (0.663) | $8.58 \times 10^{-1}$ | − 0.155 (0.678) | $8.20 \times 10^{-1}$ |
| MAS | DSST[55] | 90 | T | C | 1.651 (0.586) | $4.81 \times 10^{-3}$ | 1.425 (0.576) | $1.33 \times 10^{-2}$ |

Abbreviations: Coeff, coefficient; DSST, Digit Symbol Substitution Test; LDCT, Letter Digit Coding Test; LDST, Letter-Digit Substitution Task; SDMT, Symbol Digit Modalities Test; SE, standard error of beta coefficient; Model 1: adjusted for age and gender; Model 2: adjusted for age, gender and education.

(P-value = 0.03). No significant *cis* or *trans* associations were identified in the *ex-vivo* hippocampal data. Gene network analysis shows that CADM2 is significantly expressed in human brain (P-value for expression in cerebral cortex = $2.47 \times 10^{-171}$, area under the curve (AUC) = 0.99; P-value for expression in the prefrontal cortex specifically = $1.29 \times 10^{-31}$, AUC = 1.0; P-value for expression in the hippocampus = $8.2 \times 10^{-36}$, AUC = 0.99). The gene network analyses suggest that CADM2 is likely involved in biological processes that include the glutamate signaling pathway (P-value = $7.22 \times 10^{-15}$), gamma-aminobutyric acid (GABA) transport (P-value = $1.36 \times 10^{-11}$) and neuron cell-cell adhesion (P-value = $1.48 \times 10^{-13}$). CADM2 shows strong positive co-expression with several genes involved in the GABA and other glutamate neurosignaling pathways, such as GABA A receptor alpha 1 and beta 2 (GABRA1, GABRB2) and glutamate receptor metabotropic 5 (GRM5). Further, there is strong positive co-expression with many members of the voltage-gated potassium channel group including KCNJ9, KCNJ10, KCNB2 and KCNC, and with the OPCML (opioid binding protein/cell adhesion molecule-like) gene and EPHA5 (EPH receptor A5) (Figure 1). Supplementary Table 7 and Supplementary Figure 4 provide the expression data in neuronal tissues. As expected, Gene Network and GTex show that DRD2 is predominantly expressed in the pituitary (GTex reads per kilobase per million (RPKM) = 50), putamen (gene network AUC = 1, P-value = $5 \times 10^{-12}$, GTex RPKM = 40), substantia nigra (gene network AUC = 0.98, P-value = $4 \times 10^{-15}$, GTex RPKM = 4) and nucleus accumbens (GTex RPKM = 28). However, the intronic variant rs2734839 is not associated with expression at nominal significance in any region reported (cingulate, frontal or general cortex, amygdala, caudate nucleus, nucleus accumbens, putamen, substantia nigra, cerebellum, hippocampus, hypothalamus, pituitary or spinal cord).

## DISCUSSION

The discovery phase of our study yielded genome-wide significance for an association of processing speed (LDST/DSST) with

rs17518584. *In silico* replication in additional independent cohorts yielded nominally significant support for association. In the combined discovery-replication analysis, the model adjusted for age, sex and education resulted in genome-wide significant association between rs17518584 and LDST/DSST performance. The variant is associated with CADM2 expression, at least at nominal significance.

Rs17518584 is located about 170 kb upstream of the transcription start site of the major transcript for the CADM2 gene, but is within an intron of a variant transcript that includes an alternative first exon. Variants in the CADM2 gene have been previously associated with body mass index,[37] but not with cognitive function. The gene has been studied as a candidate for autism spectrum disorders,[45] and showed a suggestive association with scores on the persistence items on a personality scale.[46] The protein is likely involved in long-term signal depression and potentiation and neuroactive ligand-receptor interaction (http://www.genome.jp/kegg/), and is a member of the immunoglobulin superfamily. The gene encodes a neuronal adhesion molecule that has been shown to be widely expressed in the developing and adult brain in laboratory mice,[47] as well as in human and rat brain tissue, with highest expression AUCs for the prefrontal cortex. In the GTex analyses, rs17518584 significantly influenced CADM2 expression levels in the brain, most specifically in the cingulate cortex. This finding is of interest in the background of diffusion tensor imaging experiments that have shown an association between fractional anisotropy in the cingulum and performance on executive and processing speed tasks.[48,49] The co-expression data from gene network analysis also revealed some interesting links to Alzheimer's disease (Figure 1). OPCML has been associated with a variety of cognitive domains in a follow-up of linkage regions for Alzheimer's disease.[50] EPHA5 is a member of the same family of ephrin receptors as the EPHA1 gene, which is associated with Alzheimer's disease.[51,52]

No significant *cis* or *trans* associations with gene expression were identified in the experiments using *ex-vivo* hippocampal tissue from epilepsy patients. However, both GTex and Gene
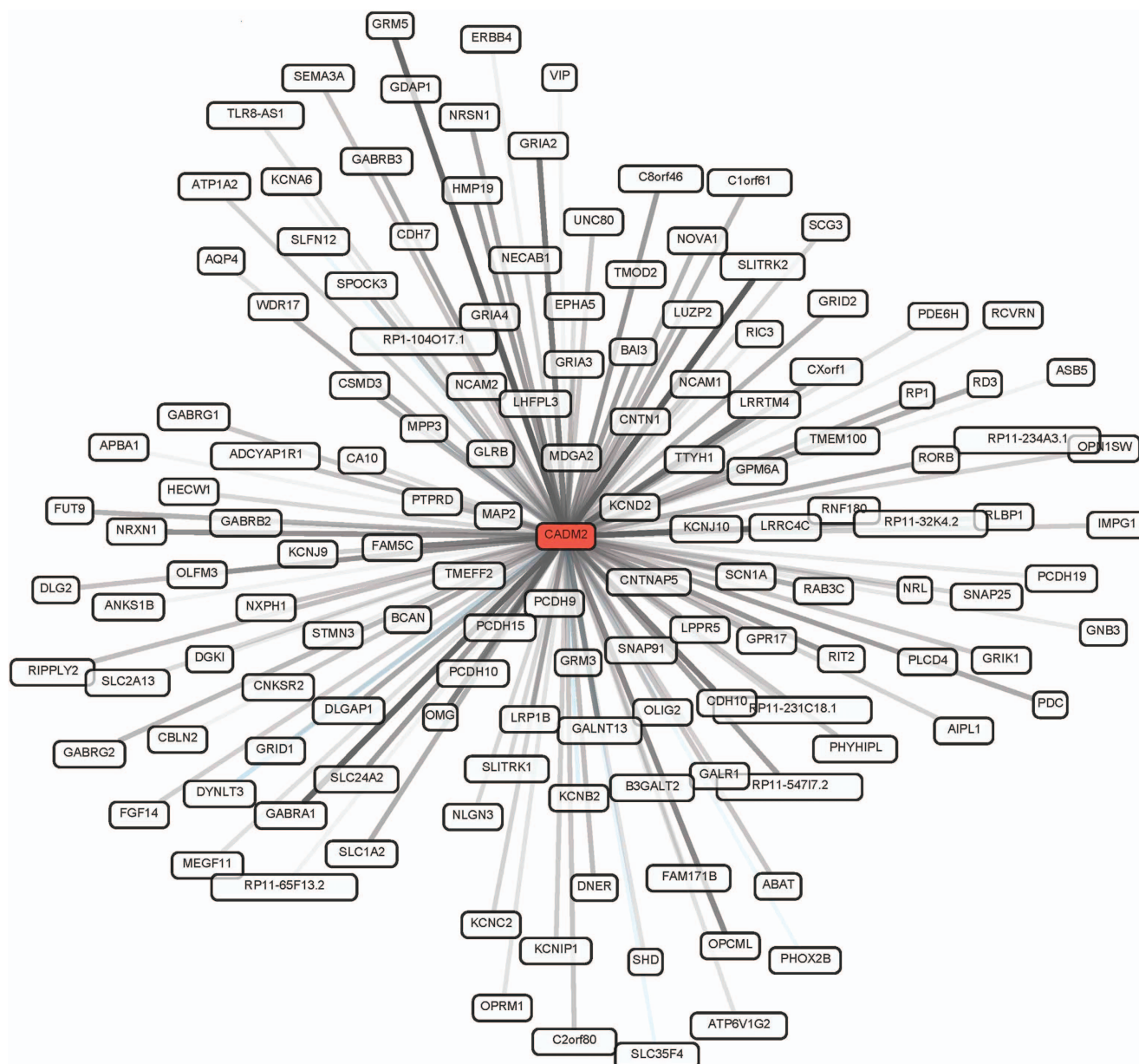
**Figure 1.** Gene network plot for *CADM2*, based on biological processes. Gray lines indicate positive co-expression and blue lines indicate negative co-expression, with the density of the line reflecting the strength of the co-expression.

Network bioinformatics sources show that *CADM2* is expressed in normal hippocampal tissue (obtained postmortem).

We identified only one genome-wide significant variant explaining a small part of the variance in performance on the LDST/DSST ($R^2 = 0.005$). These findings should be interpreted in light of the sample size studied. The number of persons analyzed varied considerably between outcomes, with 5555–32 900 subjects per individual test. This variability is a consequence of the nature of the CHARGE consortium, which uses the data previously collected in the individual cohorts, long before the GWAS era. As a consequence, there were differences in the test batteries administered in each cohort, and some tests were available in a larger number of cohorts than others. Limitations in sample size may have prevented us from finding variants (false negatives) associated with some outcomes. To minimize the detection of false positives, we sought replication in independent studies.

This may also explain, in part, why rs17518584 was identified only in LDST/DSST, which had the largest discovery sample. The strongest association for this SNP in other traits was in letter fluency, with a *P-value* of 0.076 in the discovery meta-analysis, with all studies showing a positive effect of the T allele (as in LDST/DSST). This may reflect the influence of processing speed on fluency performance previously described in the literature.[18] Our analyses lacked power to address genes with small effects affecting tests for which we had only a limited sample size. The lack of test-specific power should also be taken into account when interpreting the lack of association with candidate genes beyond *APOE* (*DTNBP1, TRIB3* and *WDR72*). There was only one other intronic variant (rs2734839) in the *DRD2* gene, encoding the D2 subtype of the dopamine receptor, that approached genome-wide significance. The bioinformatic analyses did not support a functional effect of the expression of the gene in the substantia

nigra. This variant therefore remains to be replicated. A further limitation is that the cognitive phenotypes were based on single assessments and thus we were unable to determine genetic effects on age-related cognitive decline. Despite some limitations, this is the first study to discover and replicate a genetic variant involved in processing speed.

The present report provides the most comprehensive meta-analysis of processing speed and executive function GWAS to date. We found a genome-wide significant association between widely-used tests of processing speed (LDST/DSST) and a SNP in the *CADM2* gene, which is involved in glutamatergic and GABA-ergic transmission, in middle-aged and older non-demented adults. This gene is a candidate for autism and personality, but based on the pathway and expression analyses it may also be relevant to a broad range of neuropsychiatric diseases including dementias.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## AUTHOR CONTRIBUTIONS

Study concept and design were performed by CAIV, JB, SD, MS, AVS, JCB, GD, ST, CW, BHS, VG, JWJ, TBH, OLL, IF, JTB, GH, STT, KY, JvS, LHC, SLRK, DSK, PWS, DJS, JMS, LJH, GE, RS, RJS, AP, RFG, BAO, JIR, MMN, AH, RGJW, PAW, AGU, BMP, HJG, SB, DIC, LF, JRA, IR, CH, AFW, JFW, MF, DAB, IJD, MAI, LJL, ALF, SS, CMvD and THM. Acquisition of data was carried out by CAIV, MS, ST, KP, OP, LZ, PH, JK, JL, SMR, FG, MS, BHS, VG, JWJ, PLDeJ, TBH, OLL, JTB, MKJ, RA, RSNF, SH, STT, JvS, LHC, SLRK, DSK, WMM, GH, PWS, DCL, PR, AJG, AP, LJH, NJA, SMcL, DJP, JMS, GE, RS, RJS, NAK, AP, JGE, RFG, BAO, AB, JIR, MMN, AH, PES, PAW, AGU, BMP, HJG, SB, DIC, FG, KR, JFP, PSS, LF, JRA, CH, SC, LF, JFW, MF, DAB, MAI, SS, CMvD and THM. Statistical analysis and interpretation of the data were performed by CAIV, JB, SD, MS, AVS, JCB, GD, ST, JS, CW, LBC, YL, VV, MK, LZ, JK, JL, DL, COS, KM, VC, QS, LMR, CO, QY, SSM, NA, OS,AT, NK, RSNF, WZ, KY, KL, SLRK, EGH, TT, JW, NJA, LCP, RJS, NK, AP, YCH, LY, ALdeS, AB, MG, MMN, JCL, PSS, JRA, CH, JD, AJMDeC, IJD, MAI and THM. The manuscript was drafted by CAIV, JB, SD, MS, JCB, NA, HS, SS, CMvD and THM. Critical revision of the manuscript was performed by CAIV, JB, SD, MS, AVS, JCB, GD, ST, JS, CV, LBC, YL, VV, MK, KP, OP, LZ, PH, JK, JL, DL, COS, KAM, VC, QS, SMR, LMR, CO, MS, BHS, VG, QY, SSM, JWJ, PLDeJ, TBH, DCL, NA, LHC, OS, OLL, RS, AT, IF, NK, JTB, MJK, RA, RSNF, SH, MN, WZ, STT, KY, KL, JCvS, SLRK, DSK, WMM, GH, EGH, PWS, TT, DJS, JW, PR, AJG, AP, JMS, LJH, NJA, SMcL, JMS, LCP, GE, RJS, NAK, AP, YCH, JGE, AP, RFG, BAO, LY, ALDeS, AB, MG, JIR, MN, AH, PES, RGJW, BMB, PAW, AGU, BMP, HJG, SB, DIC, FG, KR, JCL, DJP, JFP, PSS, LF, JRA, IR, CH, AFW, JFW, SC, LF, HS, JD, AJMDeC, MF, DAB, ID, MAI, LJL, AF, SS, CMvD and THM. Funding was obtained by ST, SMR, BHS, VG, JWJ, PLDeJ, TBH, OLL, IF, JTB, MN, STT, JCvS, SLRK, GH, PWS, JMS, AP, JGE, BAO, JIR, MMN, AH, PES, RGJW, BMB, PAW, AGU, BMP, HJG, SB, DIC, FG, KR, PSS, JRA, IR, AFW, JFW, AJMDeC, DAB, IJD, MAI, LJL, ALF, SS, CMvD and THM.

## REFERENCES

1 Green MF. Cognitive impairment and functional outcome in schizophrenia and bipolar disorder. *J Clin Psychiatry* 2006; **67**: 3–8, discussion 36-42.
2 Seidman LJ, Biederman J, Weber W, Hatch M, Faraone SV. Neuropsychological function in adults with attention-deficit hyperactivity disorder. *Biol Psychiatry* 1998; **44**: 260–268.
3 Huntley JD, Howard RJ. Working memory in early Alzheimer's disease: a neuropsychological review. *Int J Geriatr Psychiatry* 2009; **25**: 121–132.
4 Salthouse TA. The processing-speed theory of adult age differences in cognition. *Psychol Rev* 1996; **103**: 403–428.
5 Luciano M, Hansell NK, Lahti J, Davies G, Medland SE, Raikkonen K *et al*. Whole genome association scan for genetic polymorphisms influencing information processing speed. *Biol Psychol* 2011; **86**: 193–202.
6 Smith EE, Jonides J. Storage and executive processes in the frontal lobes. *Science (New York, NY)* 1999; **283**: 1657–1661.

7 Deary IJ, Penke L, Johnson W. The neuroscience of human intelligence differences. *Nat Rev Neurosci* 2010; **11**: 201–211.

8 Lee T, Henry JD, Trollor JN, Sachdev PS. Genetic influences on cognitive functions in the elderly: a selective review of twin studies. *Brain Res Rev* 2010; **64**: 1–13.

9 Marioni RE, Davies G, Hayward C, Liewald D, Kerr SM, Campbell A *et al*. Molecular genetic contributions to socioeconomic status and intelligence. *Intelligence* 2014; **44**: 26–32.

10 Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011; **88**: 76–82.

11 Wright M, De Geus E, Ando J, Luciano M, Posthuma D, Ono Y *et al*. Genetics of cognition: outline of a collaborative twin study. *Twin Res* 2001; **4**: 48–56.

12 Lee T, Mosing MA, Henry JD, Trollor JN, Lammel A, Ames D *et al*. Genetic influences on five measures of processing speed and their covariation with general cognitive ability in the elderly: the older Australian twins study. *Behav Genet* 2012; **42**: 96–106.

13 Cirulli ET, Kasperaviciute 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**: 815–820.

14 Tucker-Drob EM, Reynolds CA, Finkel D, Pedersen NL. Shared and unique genetic and environmental influences on aging-related changes in multiple cognitive abilities. *Dev Psychol* 2014; **50**: 152–166.

15 Lee T, Mosing MA, Henry JD, Trollor JN, Ames D, Martin NG *et al*. Genetic influences on four measures of executive functions and their covariation with general cognitive ability: the Older Australian Twins Study. *Behav Genet* 2012; **42**: 528–538.

16 Sleegers K, de Koning I, Aulchenko YS, van Rijn MJ, Houben MP, Croes EA *et al*. Cerebrovascular risk factors do not contribute to genetic variance of cognitive function: the ERF study. *Neurobiol Aging* 2007; **28**: 735–741.

17 Finkel D, Reynolds CA, McArdle JJ, Hamagami F, Pedersen NL. Genetic variance in processing speed drives variation in aging of spatial and memory abilities. *Dev Psychol* 2009; **45**: 820–834.

18 Elgamal SA, Roy EA, Sharratt MT. Age and verbal fluency: the mediating effect of speed of processing. *Can Geriatr J* 2011; **14**: 66–72.

19 Deary IJ, Spinath FM, Bates TC. Genetics of intelligence. *Eur J Hum Genet* 2006; **14**: 690–700.

20 Deary IJ, Ritchie SJ. 10 quick questions about processing speed. *Br Acad Rev* 2014; **24**: 4.

21 Payton A. The impact of genetic research on our understanding of normal cognitive ageing: 1995–2009. *Neuropsychol Rev* 2009; **19**: 451–477.

22 Wisdom NM, Callahan JL, Hawkins KA. The effects of apolipoprotein E on non-impaired cognitive functioning: A meta-analysis. *Neurobiol Aging* 2009; **32**: 63–74.

23 Zhang JP, Burdick KE, Lencz T, Malhotra AK. Meta-analysis of genetic variation in DTNBP1 and general cognitive ability. *Biol Psychiatry* 2010; **68**: 1126–1133.

24 Buyske S, Bates ME, Gharani N, Matise TC, Tischfield JA, Manowitz P. Cognitive traits link to human chromosomal regions. *Behav Genet* 2006; **36**: 65–76.

25 Rommelse NN, Arias-Vasquez A, Altink ME, Buschgens CJ, Fliers E, Asherson P *et al*. Neuropsychological endophenotype approach to genome-wide linkage analysis identifies susceptibility loci for ADHD on 2q21.1 and 13q12.11. *Am J Hum Genet* 2008; **83**: 99–105.

26 Almasy L, Gur RC, Haack K, Cole SA, Calkins ME, Peralta JM *et al*. A genome screen for quantitative trait loci influencing schizophrenia and neurocognitive phenotypes. *Am J Psychiatry* 2008; **165**: 1185–1192.

27 Seshadri S, DeStefano AL, Au R, Massaro JM, Beiser AS, Kelly-Hayes M *et al*. Genetic correlates of brain aging on MRI and cognitive test measures: a genome-wide association and linkage analysis in the Framingham Study. *BMC Med Genet* 2007; **8**: S15.

28 Cirulli ET, Kasperaviciute 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**: 815–820.

29 Need AC, Attix DK, McEvoy JM, Cirulli ET, Linney KL, Hunt P *et al*. A genome-wide study of common SNPs and CNVs in cognitive performance in the CANTAB. *Hum Mol Genet* 2009; **18**: 4650–4661.

30 LeBlanc M, Kulle B, Sundet K, Agartz I, Melle I, Djurovic S *et al*. Genome-wide study identifies PTPRO and WDR72 and FOXQ1-SUMO1P1 interaction associated with neurocognitive function. *J Psychiatr Res* 2012; **46**: 271–278.

31 Kottgen A, Pattaro C, Boger CA, Fuchsberger C, Olden M, Glazer NL *et al*. New loci associated with kidney function and chronic kidney disease. *Nat Genet* 2010; **42**: 376–384.

32 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.

33 International HapMap Consortium, Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL *et al*. A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007; **449**: 851–861.

34 Kirsch IS, Jungeblut A, Jenkins L, Kolstad A. *Adult literacy in America*. National Center for Education Statistics, U.S. Department of Education Office of Educational Research and Improvement: Washington, DC, USA, 2002.

35 Fourage D, Houtkoop W, Van der Velden R. *Laaggeletterdheid in Nederland Utrecht/'s-Hertogenbosch*. Expertisecentrum Beroepsonderwijs: The Netherlands, 2011 September. Report No.: Contract No.: ISBN/EAN 978-94-6052-039-6.

36 Calvin CM, Deary IJ, Webbink D, Smith P, Fernandes C, Lee SH *et al*. Multivariate genetic analyses of cognition and academic achievement from two population samples of 174,000 and 166,000 school children. *Behav Genet* 2012; **42**: 699–710.

37 Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU *et al*. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet* 2010; **42**: 937–948.

38 GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013; **45**: 580–585.

39 Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M *et al*. NCBI GEO: archive for functional genomics data sets--update. *Nucleic Acids Res* 2013; **41**: D991–D995.

40 Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 2002; **30**: 207–210.

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

42 Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM *et al*. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000; **25**: 25–29.

43 Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G *et al*. The Reactome pathway knowledgebase. *Nucleic Acids Res* 2014; **42**: D472–D477.

44 Coon KD, Myers AJ, Craig DW, Webster JA, Pearson JV, Lince DH *et al*. A high-density whole-genome association study reveals that APOE is the major susceptibility gene for sporadic late-onset Alzheimer's disease. *J Clin Psychiatry* 2007; **68**: 613–618.

45 Casey JP, Magalhaes T, Conroy JM, Regan R, Shah N, Anney R *et al*. A novel approach of homozygous haplotype sharing identifies candidate genes in autism spectrum disorder. *Hum Genet* 2012; **131**: 565–579.

46 Service SK, Verweij KJ, Lahti J, Congdon E, Ekelund J, Hintsanen M *et al*. A genome-wide meta-analysis of association studies of Cloninger's Temperament Scales. *Transl Psychiatry* 2012; **2**: e116.

47 Thomas LA, Akins MR, Biederer T. Expression and adhesion profiles of SynCAM molecules indicate distinct neuronal functions. *J Comp Neurol* 2008; **510**: 47–67.

48 Peters BD, Ikuta T, Derosse P, John M, Burdick KE, Gruner P *et al*. Age-related differences in white matter tract microstructure are associated with cognitive performance from childhood to adulthood. *Biol Psychiatry* 2013; **75**: 248–256.

49 Sasson E, Doniger GM, Pasternak O, Tarrasch R, Assaf Y. White matter correlates of cognitive domains in normal aging with diffusion tensor imaging. *Front Neurosci* 2013; **7**: 32.

50 Liu F, Arias-Vasquez A, Sleegers K, Aulchenko YS, Kayser M, Sanchez-Juan P *et al*. A genomewide screen for late-onset Alzheimer disease in a genetically isolated Dutch population. *Am J Hum Genet* 2007; **81**: 17–31.

51 Naj AC, Jun G, Beecham GW, Wang LS, Vardarajan BN, Buros J *et al*. Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nat Genet* 2011; **43**: 436–441.

52 Hollingworth P, Harold D, Sims R, Gerrish A, Lambert JC, Carrasquillo MM *et al*. Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nat Genet* 2011; **43**: 429–435.

53 Wechsler D. *WAIS-R Manual*. Psychological Corporation: Cleveland, OH, USA, 1981.

54 Wechsler D. *The Measurement and Appraisal of Adult Intelligence Manual for the Wechsler Adult Intelligence Scale*. Psychological Corporation: New York, NY, USA, 1955.

55 Wechsler D. *Wechsler Adult Intelligence Scale – Third Edition (WAIS III)*. Psychological Corporation: San Antonio, TX, USA, 1997.

56 Wechsler D. *WAIS-IIIUK Administration and Scoring Manual London*. Psychological Corporation: UK, 1998.

57 Jolles J, Houx PJ, Van Boxtel MPJ, Ponds RWHM. *The Maastricht Aging Study: Determinants of Cognitive Aging*. Neuropsych Publishers: Maastricht, Netherlands, 1995.

58 Smith A. *Symbol Digit Modalities Test Manual - Revised*. Western Psychological Services: Los Angeles, CA, USA, 1992.

59 Houx PJ, Shepherd J, Blauw GJ, Murphy MB, Ford I, Bollen EL *et al*. Testing cognitive function in elderly populations: the PROSPER study. PROspective Study of Pravastatin in the Elderly at Risk. *J Neurol Neurosurg Psychiatry* 2002; **73**: 385–389.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp)