GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive function

DOI:
10.1038/mp.2016.244

Document Version
Final published version

Link to publication record in Manchester Research Explorer

Citation for published version (APA):
Trampush, J. W., Yang, M. L. Z., Yu, J., Knowles, E., Davies, G., Liewald, D. C., Starr, J. M., Djurovic, S., Melle, I., Sundet, K., Christoforou, A., Reinvang, I., DeRosse, P., Lundervold, A. J., Steen, V. M., Espeseth, T., Räikkönen, K., Widen, E., Palotie, A., ... Lencz, T. (2017). GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive function: a report from the COGENT consortium. Molecular psychiatry, 22(3), 336-345. https://doi.org/10.1038/mp.2016.244

Published in:
Molecular psychiatry

Citing this paper
Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

General rights
Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Takedown policy
If you believe that this document breaches copyright please refer to the University of Manchester’s Takedown Procedures [http://man.ac.uk/04Y6Bo] or contact uml.scholarlycommunications@manchester.ac.uk providing relevant details, so we can investigate your claim.
GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive function: a report from the COGENT consortium

The complex nature of human cognition has resulted in cognitive genomics lagging behind many other fields in terms of gene discovery using genome-wide association study (GWAS) methods. In an attempt to overcome these barriers, the current study utilized GWAS meta-analysis to examine the association of common genetic variation (~8M single-nucleotide polymorphisms (SNP) with minor allele frequency ≥ 1%) to general cognitive function in a sample of 35,298 healthy individuals from European ancestry across 24 cohorts in the Cognitive Genomics Consortium (COGENT). In addition, we utilized individual SNP lookups and polygenic score analyses to identify genetic overlap with other relevant neurobehavioral phenotypes. Our primary GWAS meta-analysis identified two novel SNP loci (top SNPs: rs76114856 in the CENPO gene on chromosome 2 and rs6669072 near LOC10378853 on chromosome 1) associated with cognitive performance at the genome-wide significance level (P < 5 × 10^-8). Gene-based analysis identified an additional three Bonferroni-corrected significant loci at chromosomes 17q21.31, 17p13.1 and 1p13.3. Altogether, common variation across the genome resulted in a conservatively estimated SNP heritability of 21.5% (s.e. = 0.01%) for general cognitive function. Integration with prior GWAS of cognitive performance and educational attainment yielded several additional

1Division of Psychiatry Research, Zucker Hillside Hospital, Glen Oaks, NY, USA; 2Department of Psychology, Yale University School of Medicine, New Haven, CT, USA; 3Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, UK; 4Department of Psychology, University of Edinburgh, Edinburgh, UK; 5Department of Medical Genetics, Oslo University Hospital, University of Bergen, Bergen, Norway; 6NOMENT, K.G. Jebsen Centre for Psychosis Research, Bergen, Bergen, Norway; 7Division of Biological and Medical Psychology, University of Bergen, Bergen, Norway; 8Institute of Behavioural Sciences, University of Helsinki, Helsinki, Finland; 9Institute for Molecular Medicine Finland (FIMM), University of Helsinki, Helsinki, Finland; 10Welcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridge, UK; 11Department of Medical Genetics, University of Helsinki and University Central Hospital, Helsinki, Finland; 12National Institute for Health and Welfare, Helsinki, Finland; 13Department of General Practice and Primary Health Care, University of Helsinki, Helsinki, Finland; 14Helsinki University Central Hospital, Unit of General Practice, Helsinki, Finland; 15Folkhälsan Research Centre, Helsinki, Finland; 16Department of Psychiatry, Martin Luther University of Halle-Wittenberg, Halle, Germany; 17Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA; 18Mental Illness Research, Education, and Clinical Center (VISN 3), James J. Peters VA Medical Center, Bronx, NY, USA; 19Department of Psychology, University of Crete, Rethymno, Greece; 20Manchester Centre for Audiology and Deafness, Manchester Academic Health Science Centre, The University of Manchester, Manchester, UK; 21Division of Evolution and Genomic Sciences, School of Biological Sciences, The University of Manchester, Manchester, UK; 22Centre for Integrated Genomic Medical Research, Institute of Population Health, University of Manchester, Manchester, UK; 23Manchester Medical School, Institute of Brain, Behaviour and Mental Health, University of Manchester, Manchester, UK; 24Department of Neurology, Broyd Alzheimer’s Disease Research Center, and Center for Genomic and Computational Biology, Duke University Medical Center, Durham, NC, USA; 25Division of Brain Sciences, Department of Medicine, Imperial College, London, UK; 26Center for Applied Genomics and Precision Medicine, Duke University School of Medicine, Durham, NC, USA; 27Campbell Family Mental Health Institute, Centre for Addiction and Mental Health, University of Toronto, Toronto, ON, Canada; 28Department of Psychiatry, University of Athens School of Medicine, Greece; 29University Mental Health Research Institute, Athens, Greece; 30Neurobiology Research Institute, Theodor Theohari Cozika Foundation, Athens, Greece; 31Department of Psychiatry, Johns Hopkins University School of Medicine, Baltimore, MD, USA; 32Department of Psychiatry and McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA; 33UCLA Semel Institute for Neuroscience and Human Behavior, Los Angeles, CA, USA; 34Department of Psychology, Yale University, New Haven, CT, USA; 35Department of Psychology, Stanford University, Palo Alto, CA, USA; 36Robert and Beverly Lewis Center for Neuroimaging, University of Oregon, Eugene, OR, USA; 3723andMe, Inc., Mountain View, CA, USA; 38Department of Psychology & Neuroscience, Laboratory of Neurogenetics, Duke University, Durham, NC, USA; 39Clinical and Translational Neuroscience Branch, Intramural Research Program, National Institute of Mental Health, National Institute of Health, Bethesda, MD, USA; 40Lieber Institute for Brain Development, Johns Hopkins University Medical Campus, Baltimore, MD, USA; 41Department of Psychology, National University of Ireland, Galway, Ireland; 42Department of Psychiatry, Neuropsychiatric Genetics Research Group, Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin, Ireland; 43Department of Psychiatry and Behavioral Sciences, Faculty of Medicine, University of Crete, Heraklion, Greece; 44Helsinki Collegium for Advanced Studies, University of Helsinki, Helsinki, Finland; 45Institute for Behavioral Genetics, University of Colorado, Boulder, CO, USA; 46Institute of Clinical Medicine, University of Oslo, Oslo, Norway and 47Department of Psychiatry, Hofstra Northwell School of Medicine, Hempstead, NY, USA. Correspondence: Dr T Lencz, Division of Psychiatry Research, Zucker Hillside Hospital, 75-59 263rd Street, Glen Oaks, NY 11004, USA.

E-mail: tlencz@northwell.edu

These two authors contributed equally to this work.

Received 12 September 2016; revised 30 October 2016; accepted 3 November 2016; published online 17 January 2017
significant loci. Finally, we found robust polygenic correlations between cognitive performance and educational attainment, several psychiatric disorders, birth length/weight and smoking behavior, as well as a novel genetic association to the personality trait of openness. These data provide new insight into the genetics of neurocognitive function with relevance to understanding the pathophysiology of neuropsychiatric illness.

Molecular Psychiatry (2017) 22, 336–345; doi:10.1038/mp.2016.244; published online 17 January 2017

INTRODUCTION

Genome-wide association studies (GWAS) of complex quantitative phenotypes such as height\(^1\) and body mass index\(^2\) have successfully discovered and replicated hundreds of common variants meeting criteria for genome-wide significant association. By contrast, finding genetic loci associated with individual differences in cognitive ability using GWAS has proven challenging, despite considerable evidence from family and twin studies indicating that cognitive ability is highly heritable.\(^3\) For example, no genome-wide significant hits were detected in the earliest multi-cohort GWAS meta-analyses of general cognitive ability in \(\sim 3500\) adults,\(^4\) or in \(\sim 5000\) adults,\(^5\) or in \(\sim 18000\) youth.\(^6\) However, the first results attaining genome-wide significance, in three loci on chromosomes 6, 14 and 19, recently emerged in a GWAS meta-analysis of general cognitive function in 53 949 adults reported by the CHARGE Consortium.\(^7\) In addition, a recent study using data collected as part of the UK Biobank project reported three genomic regions significantly associated with performance on a test of verbal numerical reasoning (\(N = 36\) 035), and two independent loci were significantly associated with performance on a reaction time task (\(N = 111\) 483).\(^8\) However, in the same cohort, no genome-wide significant single-nucleotide polymorphisms (SNP)-based findings were detected for scores on a memory test, despite the large sample size (\(N = 112\) 067), perhaps due to the low reliability of the very brief assay available.\(^8\)

Recent GWAS meta-analyses of educational attainment, proposed as a proxy phenotype for cognition,\(^9,10\) have demonstrated that associations can be discovered with sufficient sample size, with the most recent analysis of 293 723 individuals yielding 74 independent SNPs that reached genome-wide significance.\(^9\) Nevertheless, this number of hits is an order of magnitude smaller than that reported for a similarly powered GWAS meta-analysis of height, which identified 697 variants that together explained \(\sim 16\%\) of the variance for adult height in a sample of 253 288 individuals.\(^1\) Thus, the complex nature of human cognition, exacerbated by challenges of precise and reliable measurement, has rendered it a more difficult phenotype with which to gain traction in the era of GWAS discovery.

The importance of uncovering the molecular genetic basis of cognitive functioning is underscored by the fact that neurocognitive deficits represent a critical component of many neuropsychiatric disorders and disease states that can affect health outcomes across the lifespan. As examples, most early appearing neurodevelopmental disorders such as autism spectrum disorder\(^13\) and attention-deficit/hyperactivity disorder\(^14\) are associated with moderate to relatively impairing comorbid deficits in neurocognitive functioning. Longitudinally, lower cognitive ability scores in childhood have been linked to decreased rates of smoking cessation in adulthood.\(^15\) The major neuropsychiatric disorders that typically emerge in early adulthood, such as schizophrenia,\(^16\) bipolar disorder,\(^17\) anxiety disorders\(^18\) and depression,\(^19\) are also associated with a range of deficits in neurocognitive function. Human personality traits such as openness to new experiences\(^20\) and negative affect\(^21\) are, respectively, associated with better and worse neurocognitive performance. Individuals with debilitating neurological illnesses such as Parkinson’s disease\(^22,23\) and (by definition) the dementia spectrum including Alzheimer’s disease\(^24\) also suffer from marked neurocognitive impairments. Further, early life cognitive performance can predict long-term development of illness,\(^25\) including mortality.\(^26\) From these findings, it has been suggested that general cognitive performance may index global bodily integrity, thereby permitting a potentially broad application of ‘cognitive epidemiology’.\(^27\)

Deciphering the genetic overlap between cognition and risk for neuropsychiatric illness and other health-relevant traits can provide useful etiological insights and help prioritize likely causal relationships among complex human traits.\(^28\) Methods to estimate the genome-wide genetic correlation (\(r_g\)) between two traits using summary GWAS statistics from published research studies utilizing linkage disequilibrium (LD) score regression procedures have recently become available.\(^26,27\) LD score regression has been used recently to show significant genetic correlations between cognition-related phenotypes and cardiovascular disease,\(^27\) physical health,\(^30\) and neuropsychiatric illness.\(^31\) However, the underlying causal variants and the genes through which they act have yet to be identified.\(^32\)

There were two major aims of the current study: (1) conduct a large-scale (\(n = 35\) 298) GWAS meta-analysis of general cognitive function in 24 independent cohorts, to identify SNP-based and gene-based loci associated with cognition; and (2) determine the extent of genetic correlation between general cognitive function and published neurobehavioral phenotypes of interest. These aims were executed within the context of the Cognitive Genomics Consortium (COGENT),\(^33,34\) an international collaborative effort designed to study the molecular genetics of cognitive function.

MATERIALS AND METHODS

Participants

To date, COGENT has acquired individual-level neuropsychological, demographic, clinical and SNP array data from 24 studies (comprised of 35 sub-studies) enrolling 35 298 individuals (51.4% females, mean age of 45.6 (s.d. \(\pm\) 8.6) years) of European ancestry (comprised of 35 sub-studies) enrolling 35 298 individuals (51.4% females, mean age of 45.6 (s.d. \(\pm\) 8.6) years) of European ancestry drawn from the general population in North America, the United Kingdom and the European continent. Table 1 provides details of the individual study cohorts. A few of the cohorts overlap with those previously reported by the CHARGE consortium,\(^7\) so comparisons to the CHARGE report\(^7\) utilize 30 sub-studies comprising 27 888 fully independent subjects. All subjects provided written, informed consent to protocols approved by their institutional ethics boards in accordance with the Helsinki Declaration.

General cognitive function phenotype

We examined general cognitive function (\(g'\)), a statistically derived broadband index of within-person performance on a neuropsychological test battery. The \(g\) phenotype estimates overall performance and is relatively invariant to the battery used and specific cognitive abilities assessed.\(^33,34\) As in our prior reports,\(^5,10\) for each cohort, \(g\) was determined using the first unrotated component extracted from a principal components
analysis of individual test scores. Details on phenotypic assessments are provided in the supplement. Briefly, each COGENT sub-study \( (n = 35) \) administered an average of 8 (s.d. ± 4) neuropsychological tests. To be included as a participant in COGENT, data from at least one neuropsychological measure across at least three domains of cognitive performance (for example, digit span for working memory; logical memory for verbal declarative memory; and digit symbol coding for processing speed), or the use of a validated g-sensitive measure was required. Digit symbol coding, digit span, verbal memory for words, visual memory, word reading, semantic fluency, verbal memory for stories, vocabulary, phonemic fluency and the trail-making test were the most common tests administered across cohorts. All individual test scores were adjusted (using multiple regression) for age\(^2\) and sex, common tests administered across cohorts. All individual test scores were adjusted (using multiple regression) for age\(^2\) and sex,

GWAS meta-analysis

In each COGENT sample, allelic association analysis of general cognitive function was conducted using imputed allele dosages and the first 10 principal components from the genotyped data to additionally adjust for population stratification. Cohorts of unrelated individuals (27 sub-cohorts) were analyzed using Plink 1.9.\(^{37}\) Samples including related individuals (8 sub-cohorts) were analyzed with BOLT-LMM\(^{38}\) using a mixed linear model association function that corrects for population stratification and relatedness.\(^{39}\) GWAS results were combined for meta-analysis of all 35 sub-studies using the inverse variance-weighted Z-score method in METAL.\(^{40}\) SNPs were filtered according to the following quality control thresholds: (1) minimum imputation quality INFO score of 0.60; (2) minor allele frequency at least 1%; and (3) minimum 10 000 samples successfully imputed. Application of these filters resulted in a total of 8 037 763 high-quality SNPs available for meta-analysis in up to 35 298 samples. The standard threshold for genome-wide significance (\( P < 5 \times 10^{-8} \)) was applied to SNP results of the GWAS meta-analysis.

Gene-based analysis

Individual SNP results from the meta-analysis were aggregated to conduct a gene-based analysis using MAGMA.\(^{41}\) SNPs were mapped to genes based on NCBI build 37.3 and defined by the start and stop site ± 5 kb, resulting in 18 164 autosomal genes. A genome-wide significance threshold for gene-based associations was calculated using the Bonferroni method (\( \alpha = 0.05/18\ 164; P = 2.75 \times 10^{-8} \)).

Genetic correlation analysis

LD score regression\(^{28,29}\) was used to derive genetic correlations among GWAS results for general cognitive function and publicly available GWAS results from multiple neurobehavioral phenotypes of potential relevance. LD score regression quantifies the extent to
which two phenotypes share genetic etiology (at least with respect to common variation captured by GWAS). GWAS summary statistics for 29 phenotypes were downloaded and processed similar to the pipeline of Bulik-Sullivan et al.\(^7\) The following phenotypes were included—cognition: childhood intelligence, educational attainment and obtaining a college degree; neurodevelopmental: autism and attention-deficit/hyperactivity disorder; neuropsychiatric: schizophrenia, bipolar disorder, anxiety and major depression; tobacco use: ever smoked cigarettes, number of cigarettes per day, age of onset of smoking and being a former smoker; personality: neuroticism, extraversion, openness, agreeableness and conscientiousness; brain volume: intracranial volume, nucleus accumbens, caudate nucleus, putamen, globus pallidus, hippocampus and thalamus; early childhood growth and development: infant head circumference, birth length and birth weight. URLs and references for data sources are provided in the Supplementary Information.

**RESULTS**

**GWAS of general cognitive function**

The QQ plot (Supplementary Figure 1) demonstrates \(\lambda = 1.12\). The LD score regression intercept of 1.04 indicates that polygenicity, rather than residual variance, coextensive with a known inversion polymorphism,\(^4\) results. In addition, a large 1.4 Mb region at chromosome 17q21.31, coextensive with a known inversion polymorphism,\(^4\) harbored 101 nearly significant SNPs (all \(P < 10^{-8}\), top SNP, rs916888, \(P = 8.18 \times 10^{-10}\)).

Gene-based analysis of general cognitive function

Seven genes in three chromosomal regions (WNT3, PLEKHM1 and ARHGAP27 at chromosome 17q21.31; TP53 and WRAP53 at chromosome 17p13.1; and ATXN7L2 and CYB561D1 at chromosome 1p13.3) were significantly associated with cognitive function after Bonferroni correction (Table 2). Several genes at these loci, including NSF, STH, KANSL1, CRHR1 and MAPT in the 17q21.31 inversion region, demonstrate association just below the Bonferroni-corrected threshold.

SNP lookups from published GWAS of cognition and educational attainment

We sought to expand the utility of our data by using it as a lookup table to confirm and extend associations previously reported in large-scale GWAS of cognition (from the CHARGE consortium\(^7\) and the UK Biobank\(^6\)) and educational attainment (from the SSGAC consortium\(^5\)). First, we looked up all cognitive SNPs nominally associated in the CHARGE study at \(P < 10^{-5}\). Importantly, because of partial sample overlap between COGENT and CHARGE, we ran our cognitive GWAS excluding five overlapping cohorts (CHS, FHS, HBCS, LBC1936 and NCNG). A meta-analytic \(P\)-value was then generated across the two studies for those loci with nominal \(P < 0.05\) in COGENT. As shown in Supplementary Table S3, we found support for the three genome-wide significant loci reported in CHARGE, as well as support for an additional, novel locus at chromosome 3p22.3 that attained a meta-analytic \(P\)-value surpassing the genome-wide significance threshold (rs1523041, \(P = 5.46 \times 10^{-10}\)). This SNP is intergenic; however, publicly available gene expression data (from GTEx\(^3\)) have shown that rs1523041 is an expression quantitative trait locus for the ARPP21 gene (Supplementary Figure 1).

Next, we examined the genome-wide significant SNPs reported from the UK Biobank study of verbal numerical reasoning and reaction time to determine if these were also associated with general cognitive ability. As shown in Supplementary Table S4, we found nominally significant support for the chromosome 22 locus associated with verbal numerical reasoning (top local SNP in COGENT, rs12170228, \(P = 0.0053\), same direction of effect in COGENT and UK Biobank). We also showed a similar trend for the chromosome 7 locus associated with verbal numerical reasoning (rs9771228, \(P = 0.074\) in COGENT, same direction of effect in COGENT and UK Biobank). The lone SNP on chromosome 14 that attained genome-wide significance in the UK Biobank GWAS of verbal numerical reasoning was a rare variant (MAF = 0.1%) that was not available in COGENT and has no known proxies. For the two loci reported to be associated with reaction time in UK Biobank, COGENT results demonstrated the same direction of allelic effects, but were not statistically significant (COGENT \(P > 0.6\)).

Finally, we looked up all SNPs that represented independent, genome-wide significant hits for educational attainment from the combined SSGAC+UK Biobank cohorts. Of a total of 164 SNPs meeting this criteria (as listed in Table 1.16 of that report\(^8\)), 143 SNPs were directly available in COGENT, and an additional 12 SNPs were tested by proxy (9 SNPs were unavailable in COGENT even by proxy). There were 31 educational attainment SNPs that were nominally significant at \(P < 0.05\) in COGENT, all in the same direction of effect as for educational attainment, representing a highly significant enrichment (\(P = 3.9 \times 10^{-11}\), binomial test) of overlap between years of education and cognitive function (Supplementary Table S5). Further, two SNPs (rs7593947 and rs2568955) were Bonferroni-corrected significant (for 155 tests; \(P < 3.23 \times 10^{-8}\), suggesting two specific loci for cognition were discovered using this ‘proxy-phenotype’ method.\(^11\) Notably, GTEx data reveal that rs2568955 is a strongly significant expression quantitative trait locus (Supplementary Figure 2) in brain tissue for RPL31P12, although this gene is annotated as a pseudogene; rs7593947 is an intronic variant in the BCL11A gene.

Genetic correlation of general cognitive function with related complex traits

SNP heritability (due to common variation) as calculated using LD score regression was 21.5% (s.e. = 0.01%). This value for \(h^2_{LD}\) is slightly lower than prior studies\(^7\) which utilized the GCTA
approach in single samples; this attenuation is expected based on the fact that LD score regression utilizes summary scores as opposed to full SNP data. The results of the LD score regression-based genetic correlations with other neurobehavioral phenotypes are detailed in Table 3. Note that our use of the LD score regression approach applied a stringent correction (unconstrained intercept) for potential sample overlap and population stratification, as well as a conservative Bonferroni correction for multiple phenotypes. Not surprisingly, strongly significant positive correlations were observed with the most closely related phenotypes: years of education ($P = 1.48 \times 10^{-63}$), obtaining a college degree ($P = 1.88 \times 10^{-23}$) and childhood intelligence ($P = 1.24 \times 10^{-13}$). A significant negative correlation was observed with schizophrenia ($P = 4.09 \times 10^{-4}$), such that genetic load for lower cognitive scores was associated with greater risk for schizophrenia, consistent with prior reports from COGENT$^5$ and others.$^{10,32}$ Similar trends were observed for attention-deficit/hyperactivity disorder and anxiety, with nominal (but not Bonferroni-corrected) levels of statistical significance. A Bonferroni-corrected significant positive correlation was observed for autism ($P = 6.00 \times 10^{-4}$), again consistent with prior reports.$^{30,44}$

A novel observation is a significant, positive genetic correlation between general cognitive ability and the personality trait of openness ($r_g = 0.48; P = 3.25 \times 10^{-4}$); no other correlations with personality traits were even nominally significant. (It should be noted that agreeableness was the only trait for which the
Table 3. Results of genetic correlation using LD score regression

| Group               | Phenotype                     | r_g  | s.e.   | z     | P-value      |
|---------------------|-------------------------------|------|--------|-------|--------------|
| Cognitive           | Childhood IQ                  | 0.89 | 0.12   | 7.41  | 1.24E–13     |
|                     | College degree                | 0.66 | 0.07   | 9.98  | 1.88E–23     |
|                     | Years of education            | 0.73 | 0.04   | 16.83 | 1.48E–63     |
|                     | ADHD                          | -0.35| 0.16   | -2.14 | 3.22E–02     |
|                     | Alzheimer's                   | -0.13| 0.11   | -1.17 | 2.41E–01     |
|                     | Anorexia                      | -0.02| 0.10   | -0.24 | 8.08E–01     |
|                     | Anxiety                       | -0.50| 0.19   | -2.57 | 1.03E–02     |
|                     | Autism                        | 0.28 | 0.08   | 3.43  | 6.00E–04     |
|                     | Bipolar                       | 0.00 | 0.08   | 0.06  | 9.52E–01     |
|                     | Major depression              | 0.10 | 0.10   | 0.96  | 3.35E–01     |
|                     | Schizophrenia                 | -0.17| 0.05   | -3.53 | 4.09E–04     |
| Personality         | Extraversion                  | -0.13| 0.10   | -1.36 | 1.74E–01     |
|                     | Agreeableness                 | 1.24 | 1.24   | 1.00  | 3.17E–01     |
|                     | Conscientousness              | 0.10 | 0.14   | 0.74  | 4.61E–01     |
|                     | Openness                      | 0.48 | 0.13   | 3.59  | 3.25E–04     |
|                     | Neuroticism                   | -0.18| 0.12   | -1.49 | 1.35E–01     |
|                     | Age of onset                  | 0.21 | 0.13   | 1.67  | 9.49E–02     |
|                     | Cigarettes per day            | 0.03 | 0.11   | 0.27  | 7.85E–01     |
|                     | Ever smoker                   | -0.24| 0.08   | -3.07 | 2.13E–03     |
|                     | Former smoker                 | 0.29 | 0.10   | 2.80  | 5.09E–03     |
| Smoking             | Accumbens                     | 0.26 | 0.15   | 1.74  | 8.18E–02     |
|                     | CAudate                       | 0.08 | 0.10   | 0.79  | 4.30E–01     |
|                     | Hippocampus                   | 0.24 | 0.13   | 1.88  | 6.06E–02     |
|                     | Intracranial volume           | 0.14 | 0.13   | 1.09  | 2.77E–01     |
|                     | Pallidum                      | 0.16 | 0.13   | 1.28  | 2.02E–01     |
|                     | Putamen                       | 0.13 | 0.09   | 1.44  | 1.50E–01     |
|                     | Thalamus                      | 0.13 | 0.12   | 1.07  | 2.83E–01     |
| Brain volume        | Birth length                  | 0.20 | 0.09   | 2.13  | 3.33E–02     |
|                     | Birth weight                  | 0.15 | 0.05   | 2.89  | 3.90E–03     |
|                     | Infant head Circumference     | 0.19 | 0.11   | 1.69  | 9.04E–02     |

Abbreviations: ADHD, attention-deficit/hyperactivity disorder; IQ, intelligence quotient; LD, linkage disequilibrium. Traits that are in bold font were nominally significant (P < 0.05); traits that are italicized were significant after Bonferroni correction.
SNP-based heritability did not significantly differ from zero \((h^2 = 0.016; \text{s.e.} = 0.029)\) and was, therefore, not included in correlative analyses. Higher cognitive ability was strongly genetically correlated with reduced rates of smoking \((P = 2.13 \times 10^{-6})\) and greater rates of quitting smoking \((P = 5.09 \times 10^{-3})\).

Nominally significant positive genetic correlations were observed with birth length and weight, with similar trends for infant head circumference. All genetic correlations with neuroanatomic measures trended in the positive direction (larger brain volumes associated with higher cognitive ability), consistent with a large literature revealing correlations at the phenotypic level, although none of the genetic correlations attained nominal significance under our conservative approach.

**DISCUSSION**

Our GWAS meta-analysis of general cognitive function in a sample of 35,298 individuals of European ancestry revealed two novel associated SNP loci, three novel gene-based loci, and provided added support for several previously reported associations. Strengths of our study included access to individual-level genetic and neuropsychological data, which allowed us to run each sample through uniform genotype and phenotype quality control pipelines. Specifically, the general cognitive function phenotype was well characterized as a composite score derived from relatively large batteries of both verbal and nonverbal neuropsychological tests. Genotype data were processed with the latest imputation platforms and analytic procedures.

Our top GWAS hit was rs76114856, of which the minor \(T\) allele was associated with reduced cognitive performance. This SNP is an intronic variant in the \textit{CENPO} gene, which encodes a component of the interphase centromere complex.\(^{46}\) This gene is highly expressed in the basal ganglia and thalamus of the human brain.\(^{65}\) \textit{CENPO} is located at chromosome 2p23.3 and has prior GWAS associations to height.\(^{1,47}\) The \textit{CENPO} gene also had a nominal association to cognition in our gene-based analysis at \(P < 0.05\), as did neighboring genes \textit{NCOA1}, \textit{PThrH1}, and \textit{ADCY3}.

The second strongest GWAS signal fell within a large intergenic non-coding RNA (lincRNA) of unknown function, \textit{RP4-665J23.1}. Neighboring protein-coding genes are poorly annotated and do not provide strong clues as to the potential biological mechanism underlying the association.

We also found evidence that the chromosome 17q21.31 inversion region is associated with cognitive function. The chromosome 17q21.31 inversion consists of two haplotypes (H1 and H2), and the absence of recombination across the \(~1.5\) Mb region between the inverted (H2) and the noninverted (H1) chromosomes has resulted in two families of chromosomes.\(^{48}\) H1 chromosomes comprise the common (~80% frequency in European samples) noninverted gene order, whereas the H2 haplotype comprises the inverted gene order (~20% in European samples).\(^{48}\) There are several sources of evidence that variation at this locus is associated with neurobehavioral phenotypes. For example, the 17q21.31 microdeletion syndrome is associated with the H2 haplotype, which carries additional low-copy repeats susceptible to non-allelic homologous recombination. The syndrome is characterized clinically by developmental delay/intellectual disability, neonatal/childhood hypotonia, friendly behavior and specific facial dysmorphisms.\(^{49}\) Notably, \textit{KANSL1} gene disruption is associated with the full clinical spectrum of 17q21.31 microdeletion syndrome.\(^{49}\) In addition, the region harbors the \textit{MAPT} gene, encoding microtubule-associated protein tau, a hallmark of multiple dementias.\(^{48,50–52}\) The H1 family of haplotypes has been associated with increased risk for late-life tauopathies, diseases marked by the accumulation of \textit{MAPT} neurofibrillary tangles in nerve cells, such as sporadic frontotemporal dementia,\(^{53}\) Alzheimer’s disease,\(^{54}\) Parkinson’s disease\(^{55}\) and progressive supranuclear palsy.\(^{55}\) By contrast, the H2 haplotype has been associated with developmental delay and learning difficulties,\(^{51,56–58}\) as well as reduced intracranial volume.\(^{52}\)

Consistent with these latter observations, our data suggest alleles corresponding to the H2 haplotype are associated with worse cognitive performance.

In addition to the loci attaining clear genome-wide significance through our primary SNP-based and gene-based analyses, our results confirmed and extended prior GWAS studies of cognitive and educational phenotypes. Although a prior COGENT report provided converging evidence for a role of a chromosome 6 locus (rs1906252),\(^{18}\) we now provide further support for the \textit{NPAS3/AKAP6} locus on chromosome 14 previously reported by the CHARGE consortium. \textit{NPAS3} is a promising candidate gene, as it has a role in neurodevelopment, and disruptions of this gene have been associated with psychiatric and intellectual disability phenotypes.\(^{59,60}\)

In the context of prior associations to cognitive and educational phenotypes, our data identified several loci with strong empirical support for a role in cognition. Of these, two are noteworthy for representing known expression quantitative trait locus, permitting inference of potential biological mechanisms underlying the statistical associations. Specifically, we found that the major \((C)\) allele at rs1523041 was strongly \((P = 5.46 \times 10^{-10})\) associated with better cognitive performance; this allele drives lower expression of the \textit{ARPP21} gene (Supplementary Figure S1). \textit{ARPP21} encodes a cAMP-regulated phosphoprotein, enriched in the basal ganglia and cerebellum, that has a central role in the integration of neurotransmitter inputs into striatal medium spiny neurons.\(^{61}\) Intriguingly, a deletion encompassing this gene segregated with syndromic intellectual disability in a multiply affected pedigree.\(^{62}\) Similarly, we found that the minor \((T)\) allele of rs2568955 was associated with poorer cognitive performance, and this allele is associated with greater expression of \textit{RPL31P12} (Supplementary Figure S2). It should be noted that the strongest expression quantitative trait locus associations for these SNPs were observed in non-brain tissue in the GTEx database, perhaps due to smaller sample sizes available for neuronal phenotypes; these results should be tested in larger studies of brain expression that will soon be forthcoming. \textit{BCL11A} is also a promising candidate gene for cognition. Haploinsufficiency of this gene has been associated with intellectual disability in a large clinical study, with the phenotype recapitulated in \textit{Bcl11a} knockout mice, which was shown to be mediated through downstream transcriptional dysregulation in the hippocampus and cortex.\(^{63}\)

Analysis of the genetic correlation between general cognitive function and various other phenotypes revealed that better cognitive performance was robustly genetically correlated with more years of schooling, decreased likelihood of smoking and decreased risk for several psychiatric disorders (as well as increased risk for autism). These results are generally consistent with recent genetic correlation studies of cognitive phenotypes\(^{10}\) and proxy phenotypes for cognition.\(^{9}\) The personality trait of openness, a core component of the ‘Big 5’ model of personality, was positively correlated with cognitive ability at the genetic level. This novel finding is consistent with a prior literature in which moderate phenotypic correlations (values for \(r\) ranging between 0.25 and 0.5) between openness and cognitive ability have been repeatedly observed,\(^{64,66}\) whereas cognition is generally uncorrelated with other personality dimensions. Moreover, phenotypic data from twin and family studies have suggested a specific genetic correlation between openness and general cognitive ability.\(^{65}\) Longitudinal studies have suggested a model in which openness may serve as a ‘buffer’ against cognitive decline, as has been proposed in the Openness-Fluid-Crystallized-Intelligence model applied to late adulthood.\(^{66}\) Positive genetic correlations with birth length and weight suggest a critical role for prenatal developmental factors in the subsequent manifestation of cognitive ability throughout the lifespan.
One limitation of the current study is the wide age range of subjects, both across cohorts and within cohorts. Although we sought to control for confounding effects of age using covariates, genetic influence on cognitive ability is somewhat reduced in early childhood and adolescence relative to adulthood.57 In addition, late-life effects of cognitive decline may be mediated through somewhat disparate molecular pathways; this may explain the relatively weak effect of variation at APOE compared with prior GWAS meta-analysis.5 Nevertheless, cognitive abilities are remarkably stable across the entire lifespan,68 so we chose to include all available samples in order to maximize sample size and power.

Similarly, we chose to include cohorts with widely disparate neuropsychiatric batteries, which undoubtedly contributed to noise surrounding the estimates of g. Moreover, as demonstrated in Supplementary Table 1, the degree to which the estimates of cognitive ability varied across cohorts, thereby reducing power in comparison with more easily measured quantitative traits such as height. Notably, it has been empirically demonstrated that such noise is more than compensated by increases in statistical power.70 Given the expense in conducting comprehensive cognitive assessments, we chose to include all available cohorts meeting our basic criteria.

Despite the statistical significance of the novel GWAS loci identified in the current report, it is important to emphasize that the effect sizes for individual SNPs are very small; each of our top two SNPs individually account for ~0.1% of the variance in cognitive performance. For context, these effect sizes are considerably smaller than those observed for the top individual loci associated with other quantitative anthropometric traits such as height and weight.1,2 This difference may reflect the complexity of the underlying genetic architecture of cognition, as ~80% of all genes are expressed in brain;21 this complexity has also slowed progress in identifying genetic loci for neuropsychiatric disorders, given the potentially large mutational target.70 This challenge is exacerbated by the fact that general cognitive ability is a latent trait that is only indirectly captured by the available phenotypic measures, which are also quite heterogeneous across cohorts. Moreover, the well-known winners’ curse phenomenon will likely result in even further reduction of the effect size estimates in future studies of independent cohorts. However, as described in the preceding paragraphs, results of the present study can provide important information about the molecular underpinnings of cognitive function as well as clues relevant to the etiology of neuropsychiatric disorders and other conditions relevant to human health. Although it is striking that general cognitive ability remains the quantitative trait most challenging to GWAS methodology, the recent success of very large-scale GWAS in educational attainment provides optimism that cognition is now at the beginning of the slope of increasing GWAS discovery that has been observed for all heritable complex traits.73

CONFLICT OF INTEREST
The authors declare no conflict of interest.

REFERENCES
1 Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S et al. Defining the role of common variation in the genome and biological architecture of adult human height. Nat Genet 2014; 46: 1173–1186.
2 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.
3 Deary IJ, Johnson W, Houlden LM. Genetic foundations of human intelligence. Hum Genet 2009; 126: 215–232.
4 Davies G, Tenesa A, Payton A, Yang J, Harris SE, Liewald D et al. Genome-wide association studies establish that human intelligence is highly heritable and polygenic. Mol Psychiatry 2011; 16: 996–1005.
5 Lencz T, Knowles E, Davies G, Guha S, Liewald DC, Starr JM et al. Molecular genetic evidence for overlap between general cognitive ability and schizophrenia: a report from the Cognitive Genomics consortiun (COGENT). Mol Psychiatry 2013; 19: 168–174.
6 Benyamin B, Bourcain B, Davies G, Hansell NK, Brion M-JA et al. Childhood intelligence is heritable, highly polygenic and associated with FBN1. Mol Psychiatry 2014; 19: 253–258.
7 Davies G, Armstrong N, Bis JC, Bressler J, Chouraki V, Giddaluru S et al. Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium (N=5 349). Mol Psychiatry 2015; 20: 183–192.
8 Davies G, Marioni RE, Liewald DC, Hill WD, Hagaers SP, Harris SE et al. Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=1 151). Mol Psychiatry 2016; 21: 758–767.
9 Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA et al. Genome-wide association study identifies 74 loci associated with educational attainment. Nature 2016; 533: 519–524.
10 Trampush JW, Lencz T, Knowles E, Davies G, Guha S, Pe’er I et al. Independent evidence for an association between general cognitive ability and a genetic locus for educational attainment. Am J Med Genet Part B Neuropsychiatr Genet 2015; 168B: 363–373.
GWAS of general cognitive function

JW Trampusch et al.
additional copy number variations in a large family. Am J Med Genet A 2013; 161A: 2890–2893.

63 Dias C, Estruch SB, Graham SA, McRae J, Sawiak SJ, Hurst JA et al. BCL11A haploinsufficiency causes an intellectual disability syndrome and dysregulates transcription. Am J Hum Genet 2016; 99: 253–274.

64 Zimprich D, Allemand M, Dellenbach M. Openness to experience, fluid intelligence, and crystallized intelligence in middle-aged and old adults. J Res Pers 2009; 43: 444–454.

65 Wainwright M, Wright MJ, Luciano M, Geffen GM, Martin NG. Genetic covariation among facets of openness to experience and general cognitive ability. Twin Res Hum Genet 2008; 11: 275–286.

66 Ziegler M, Cengia A, Mussel P, Gentorf D. Openness as a buffer against cognitive decline: the Openness-Fluid-Crystallized-Intelligence (OFCI) model applied to late adulthood. Psychol Aging 2015; 30: 573–588.

67 Haworth CM, Wright MJ, Luciano M, Martin NG, de Geus EJ, van Beijsterveldt CE et al. The heritability of general cognitive ability increases linearly from childhood to young adulthood. Mol Psychiatry 2010; 15: 1112–1120.

68 Deary IJ, Whalley LJ, Lemmon H, Crawford JR, Starr JM. The stability of individual differences in mental ability from childhood to old age: follow-up of the 1932 Scottish Mental Survey. Intelligence 2000; 28: 49–55.

69 Wechsler D. Wechsler Abbreviated Scale of Intelligence (WASI). Norwegian Manual Supplement. Pearson Assessment: Stockholm, 2007.

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

70 Chabris CF, Lee JJ, Benjamin DJ, Beauchamp JP, Glaeser EL, Borst G et al. Why it is hard to find genes associated with social science traits: theoretical and empirical considerations. Am J Public Health 2013; 103: S152–S166.

71 Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, Shen EH, Ng L, Miller JA et al. An anatomically comprehensive atlas of the adult human brain transcriptome. Nature 2012; 489: 391–399.

72 Zollner S, Pritchard JK. Overcoming the winner’s curse: estimating penetrance parameters from case-control data. Am J Hum Genet 2007; 80: 605–615.

73 Visscher PM, Brown MA, McCarthy MI, Yang J. Five years of GWAS discovery. Am J Hum Genet 2012; 90: 7–24.