

Genetic markers of ADHD-related variations in intracranial volume

Marieke Klein¹, Raymond K. Walters^{2,3,4}, Ditte Demontis^{5,6,7}, Jason L. Stein⁸, Derrek P. Hibar⁹, Hieab H. Adams^{10,11}, Janita Bralten¹, Nina Roth Mota¹, Russell Schachar¹², Edmund Sonuga-Barke¹³, Manuel Mattheisen^{5,6,7,14,15}, Benjamin M. Neale^{16,17}, Paul M. Thompson^{9,18}, Sarah E. Medland¹⁹, Anders D. Børglum^{5,6,7}, Stephen V. Faraone^{20,21}, Alejandro Arias-Vasquez^{1,22,23#}, Barbara Franke^{1,22#*}

¹ Radboud university medical center, Donders Institute for Brain, Cognition and Behaviour, Department of Human Genetics, Nijmegen, The Netherlands

² Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA

³ Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA

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

⁵ Department of Biomedicine, Aarhus University, Aarhus, Denmark

⁶ Centre for Integrative Sequencing, iSEQ, Aarhus University, Aarhus, Denmark

⁷ The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Denmark

⁸ Department of Genetics & Neuroscience Center, University of North Carolina, Chapel Hill, North Carolina, USA

⁹ Imaging Genetics Center, USC Mark and Mary Stevens Neuroimaging & Informatics Institute, Keck School of Medicine of University of Southern California, Los Angeles, USA

¹⁰ Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands

¹¹ Department of Radiology and Nuclear Medicine, Erasmus Medical Center, Rotterdam, The Netherlands

¹² Department of Psychiatry and Research Institute, The Hospital for Sick Children, University of Toronto, Toronto, Canada

¹³ Department of Child and Adolescent Psychiatry, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

¹⁴ Department of Clinical Neuroscience, Centre for Psychiatric Research, Karolinska Institutet, Stockholm, Sweden

¹⁵ Stockholm Health Care Services, Stockholm County Council, Stockholm, Sweden

¹⁶ Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, USA

¹⁷ Analytic and Translational Genetics Unit, Department of Medicine, Massachusetts General Hospital, Boston, USA

¹⁸ Department of Neurology, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

¹⁹ Quantitative Genetics Laboratory, QIMR Berghofer Medical Research Institute, Brisbane, Australia

²⁰ Departments of Psychiatry and of Neuroscience and Physiology, SUNY Upstate Medical University, Syracuse, NY, USA

²¹ K.G. Jebsen Centre for Neuropsychiatric Disorders, University of Bergen, Bergen, Norway

²² Radboud university medical center, Donders Institute for Brain, Cognition and Behaviour, Department of Psychiatry, Nijmegen, The Netherlands

²³ Radboud university medical center, Donders Institute for Brain, Cognition and Behaviour, Department of Cognitive Neuroscience, Nijmegen, The Netherlands

shared final responsibility

* Corresponding author

Corresponding author:

Barbara Franke, PhD

Department of Human Genetics (route 855)

Radboud university medical center

Postbus 9101, 6500 HB Nijmegen, The Netherlands

Telephone: +31243610181

Email: barbara.franke@radboudumc.nl

Running title: Genetic overlap between ADHD and brain volumes

Keywords: ADHD, ICV, MRI brain imaging, genetics, neurite outgrowth

Word count main text: 3,002 /4000

Word count online methods: 2,777/3,000

References: 47 (without methods)/50

ABSTRACT

Attention-Deficit/Hyperactivity Disorder (ADHD) is a common and highly heritable neurodevelopmental disorder with a complex pathophysiology, where genetic risk is hypothesized to be mediated by alterations in structure and function of diverse brain networks. We tested one aspect of this hypothesis by investigating the genetic overlap between ADHD (n=55,374) and (mainly subcortical) brain volumes (n=11,221-24,704), using the largest publicly available studies. At the level of common variant genetic architecture, we discovered a significant negative genetic correlation between ADHD and intracranial volume (ICV). Meta-analysis of individual variants found significant loci associated with both ADHD risk and ICV; additional loci were identified for ADHD and amygdala, caudate nucleus, and putamen volumes. Gene-set analysis in the ADHD-ICV meta-analytic data showed significant association with variation in neurite outgrowth-related genes. In summary, our results suggest new hypotheses about biological mechanisms involved in ADHD etiology and highlight the need to study additional brain parameters.

Abstract length: 147 words

INTRODUCTION

Attention-Deficit/Hyperactivity Disorder (ADHD) is a common and highly heritable neurodevelopmental disorder with a complex and presumably heterogeneous pathophysiology. Pathways towards disease are hypothesized to be mediated by alterations in diverse brain networks¹. A recent neuroimaging mega-analysis reported subtle but consistent differences in the volumes of subcortical brain regions and intracranial volume (ICV) in ADHD, across diverse cohorts worldwide: compared to healthy controls, patients with ADHD showed decreased volumes of the nucleus accumbens, amygdala, caudate nucleus, hippocampus, putamen, and ICV². How such alterations contribute to the disease phenotype, is still poorly understood. However, brain volume alterations are also present, on average, in unaffected relatives of patients with ADHD^{3,4}, and both ADHD and brain volumes have high heritability (60-70%⁵⁻⁷ and 70-90%⁸, respectively). This suggests that genetic variants underlying ADHD pathophysiology may also influence brain volume variation. Recently, the first genome-wide significant loci for ADHD were identified and a SNP-based heritability of 20.16% has been reported⁹. The current study aimed to determine, if common genetic variants are shared between ADHD risk and subcortical brain volumes or ICV.

Genome-wide association studies (GWAS) have identified ten genome-wide significant loci associated with hippocampal volume^{8,10-13}, eight loci with ICV^{8,10,14,15}, four loci with putamen volume⁸, and one locus with caudate volume⁸. These variants explain only a small fraction of the heritability of the brain structure volumes^{8,10-12,14}. Recently, Franke and colleagues reported a battery of statistical tools to comprehensively examine genetic overlap between brain volumes and risk for brain disease at the genome-wide level and at the level of individual risk variants, using schizophrenia as an example¹⁶. Here, we applied a similar set of methods to identify and dissect genetic sharing between ADHD and brain volumes implicated in ADHD based on the latest mega-analysis². We used the largest available data sets, i.e., the GWAS meta-analysis (GWAS-MA) of ADHD from the Psychiatric Genomics Consortium (PGC) and The Lundbeck Foundation Initiative for Integrative Psychiatric Research

(iPSYCH)⁹ and the GWAS-MAs of brain volumes from the ENIGMA Consortium (Enhancing Neuroimaging Genetics through Meta-Analysis)^{8,13,17} and the Neurology Working Group of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium^{18,19}.

RESULTS

We analyzed GWAS-MA data for ADHD (20,183 cases and 35,191 controls) and six MRI-based measures of brain volumes (11,221–24,704 individuals; subcortical brain volumes were corrected for ICV). Detailed information on the numbers of individuals and genetic variants of these different cohorts are given in **Supplementary Table 1**. These data were used in a comprehensive set of analyses of common variant genetic sharing between ADHD and brain volumetric measures evaluating potential overlap at global and single variant levels.

Comparison of common variant genetic architectures

Linkage disequilibrium score regression. The SNP-based heritability estimates for the MRI measures were consistent with prior reports^{13,16,19} and ranged from 13.32% (nucleus accumbens) to 28.15% (putamen; **Table 1**). The amygdala mean χ^2 was too small to allow a valid analysis. In the analysis of the genetic correlations, we observed a significant negative genetic correlation between ADHD and ICV ($r_g = -0.227$, $P = 0.00015$). All other correlations were non-significant (**Table 1**; **Supplementary Table 2** shows the results when using constrained intercepts).

SNP effect concordance analysis. We used SNP Effect Concordance Analysis (SECA)²⁰ to determine the extent and directionality of genetic overlap between ADHD and each brain volume. We found significant evidence of global pleiotropy for variants affecting ADHD risk for volumes of four subcortical brain regions and ICV (ADHD pleiotropy only reached nominal significance for the nucleus accumbens, **Table 2** and **Supplementary Fig. 1**). The discordant SNP effects for ADHD and ICV were

significant, i.e. variants increasing the risk for ADHD were associated with decreased ICV ($P_{ICV} < 0.001$; **Table 2** and **Supplementary Fig. 2**). Evidence for concordant SNP-effects reached significance for ADHD and nucleus accumbens ($P_{accumbens} = 0.002$) and caudate nucleus ($P_{caudate} = 0.004$, **Table 2** and **Supplementary Fig. 2**); analyses for other structures did not reach study-wide significance (**Table 2**).

Sign tests. Based on the phenotypic observation that patients with ADHD have, on average, smaller brain volumes compared to healthy controls², we had expected discordant rather than concordant SNP effects. As both discordant and concordant effects were observed in the SECA among the brain volumes, we subsequently used sign tests to specifically determine directionality of genetic overlap between ADHD and each brain volume for the top-associations per trait. In this, we zoomed in further on the most strongly associated and LD-independent SNPs and compared the signs of the regression coefficients of those top-associations per trait across ADHD and the MRI volumetric measures. None of the sign tests showed a consistent direction of discordance, after correction for multiple testing (**Supplementary Table 3**).

Analyses at the single genetic variant level

Weighted SNP meta-analyses. Based on the findings of both concordant and discordant links between ADHD and the brain volume SNPs, we performed a genome-wide search for specific genetic loci associated with both ADHD and each of the brain traits. For this, we used a weighted SNP meta-analysis design allowing the combination of findings from GWAS of binary and quantitative variables⁹. This enabled us to specifically look for concordant effects at the level of single genetic variants; there is currently no suitable method to study discordant effects. The weighted GWAS-MA for ADHD and ICV, comprising 79,398 individuals from independent cohorts, identified two loci of interest, one on chromosome 15 (*SEMA6D*) and one on chromosome 16 (intergenic). Those loci were genome-wide significant in the cross-phenotype meta-analysis and/or had a cross-phenotype p-value, which was improved by at least one order of magnitude, and a cross-phenotype z-score, which

(at least) equaled the ones observed in the GWAS-MAs for the individual phenotypes (**Table 3** and **Figures 1** and **2**). Four additional LD-independent genetic loci passed the study-wide threshold for genome-wide significance of $P=8.33 \times 10^{-9}$, but those were related to a single phenotype and did not meet the criteria for cross-disorder relevance (**Figure 1**).

We also performed weighted GWAS-MAs for ADHD and four subcortical brain structures. For amygdala volume, a naïve sample size-weighted meta-analysis was performed, as no genetic correlation with ADHD had been estimated. Manhattan plots for these analyses are in **Supplementary Figures 3-7**; the six novel and/or improved LD-independent genome-wide significant loci observed in these analyses are summarized in **Table 3**. Among those, the *SEMA6D* locus was significantly associated with ADHD and putamen volume ($P=3.62 \times 10^{-9}$; **Table 3**, **Figure 2**, and **Supplementary Fig. 7**).

Reciprocal lookup of genome-wide significant associations. No significant associations were observed between the 14 previously identified genome-wide significant ADHD SNPs⁹ and brain volumes, (**Supplementary Table 4**). Conversely, among the 21 SNPs previously associated with the brain volumes^{8,13,19}, two associations with ADHD survived correction for multiple testing (**Supplementary Table 5**; $P<0.00238$). A proxy variant (rs8756; $r^2=1$; chr12:66359752, located in exon 5 of *HMGA2*) of the ICV-associated variant rs138074335 (chr12:66374247, intergenic and upstream of *HMGA2*; increasing ICV) was associated with increased risk for ADHD (OR=1.042, $P=0.00186$). A second variant, rs2195243 (chr12:102922986, intergenic and upstream of *IGF1*), was associated with decreased ICV and increased risk for ADHD (OR=1.0565, $P=0.00127$).

Expression quantitative loci and brain gene expression

The recently published ADHD GWAS-MA paper described biological annotations implementing the co-localized genes for some of the genome-wide significant loci⁹. Among those, it was shown that

many SNPs in the *SEMA6D* locus on chromosome 15, which we found overlapping between ADHD risk and ICV, were strongly associated with the expression of *SEMA6D* in fibroblasts²¹. Indeed, repeating this analysis for the most strongly associated variants identified by the (weighted) cross-phenotype SNP meta-analyses using the GTEx data²¹, we found that rs281320 on chromosome 15, the SNP with the strongest association in the ADHD+ICV GWAS-MA ($P=9.95 \times 10^{-11}$), to be a significant expression quantitative trait locus (eQTL) for *SEMA6D* in transformed fibroblast tissue ($P=1.2 \times 10^{-20}$; **Supplementary Table 6**). Likewise, rs281323 was also strongly associated with *SEMA6D* expression ($P=1.2 \times 10^{-21}$). The alternative alleles of both rs281320 and rs281323, which are associated with increased risk for ADHD and larger ICV, also increased *SEMA6D* expression (**Supplementary Fig. 9, a and b**). Additionally, rs12653396 was a significant eQTL for the *CTC-498M16.4* and *MEF2C* genes, in brain ($P=2 \times 10^{-7}$)²¹ and blood tissue ($P=6.53 \times 10^{-7}$)²², respectively, with the disease-associated A-allele also being associated with increased *MEF2C* expression (**Supplementary Table 6**). All other top-SNPs were not present in either of the two eQTL databases.

We determined mRNA expression for *SEMA6D* and *MEF2C*, the only protein-coding genes identified in the cross-phenotype GWAS-MAs, in six brain regions of the developing and adult human brain using data from the Human Brain Transcriptome Project²³. Both genes were found to be globally expressed in the developing and adult human brain, with highest mRNA expression in prenatal periods (**Supplementary Fig. 8**).

Neurite outgrowth gene-set analysis

Based on our finding that *SEMA6D* – a gene with a prominent role in neuronal migration and axonal path finding - is a key locus contributing to both ADHD risk and ICV, we investigated, whether genes involved in directed neurite outgrowth, more generally, have a role in ADHD–ICV genetic overlap. We found a significant association between a pre-defined gene-set of 45 neurite outgrowth genes²⁴ and the meta-analytic data for ADHD+ICV using MAGMA ($P_{\text{competitive}}=0.00338$). There were no significant

associations with either ADHD or ICV, separately (**Supplementary Table 7**). As expected, we found a significant association of the neurite outgrowth gene-set with ADHD for the self-contained test ($P_{\text{self-contained}}=5.53 \times 10^{-9}$). Using the ADHD+ICV GWAS-MA data, the most strongly associated individual gene among the set of neurite outgrowth genes was *CREB5* ($P=0.000553$), and a total of nine additional neurite outgrowth genes showed nominally significant associations (**Supplementary Table 8**).

DISCUSSION

We found significant - though modest - genetic covariation between ADHD risk and brain volumes, both on global and single variant levels. On the global level, significant negative genetic correlation between ADHD and ICV was demonstrated. The direction of effect was supported by the SNP effect concordance analysis. For most of the subcortical brain volumes affected in ADHD, pleiotropic effects were also found. On the single variant level, meta-analyses found significant loci associated with both ADHD risk and brain volumes. We identified *SEMA6D* as a potential key-locus contributing to both ADHD risk and ICV, and gene-set analysis revealed a significant association of ADHD-ICV overlap with variation in previously ADHD-related neurite outgrowth genes.

In contrast to prior studies across different mental disorders, such as schizophrenia^{15,18}, major depressive disorder²⁵, and autism¹⁹, which provided no evidence for a sharing of genetic architecture and brain volumes, we found a modest but significant negative global genetic correlation between ADHD and ICV. This finding was highly consistent across approaches (i.e. LDSR and SECA; and reciprocal look-up for rs2195243) and in the expected direction, given the previous observation that patients with ADHD have smaller ICV relative to controls². Recent studies also support a genetic correlation between ICV and adult height, Parkinson's disease, and cognitive function in children and adults¹⁹. With ADHD being negatively genetically correlated with general cognitive function²⁶ and human intelligence²⁷, it remains to be tested, whether the same genes are involved.

We observed significant pleiotropy also between ADHD and amygdala, caudate nucleus, hippocampus, and putamen volumes. This global genetic covariation was substantiated by the local effects, which we observed in the weighted cross-phenotype meta-analyses. In addition to ICV, caudate nucleus, and putamen volumes variation also showed significant genetic concordance with ADHD. However, whereas the results for ICV were in line with our expectation (variants increasing risk for ADHD decrease brain volume; ADHD patients have overall smaller brain volume²), the concordant effects for ADHD and nucleus accumbens and caudate nucleus volume are counterintuitive (ADHD patients have smaller volumes for these structures²), suggesting a reverse or more complex pattern of causation.

We investigated genetic overlap on global and single variant levels. Combining both approaches is highly informative, as they answer different questions. Whereas on the global level, we are interested in the general genetic architecture underlying two phenotypes, the single variant level aims to identify specific loci that can inform us about shared biological mechanisms. In line with this, we performed a meta-analysis of ADHD and brain volumes even for those volumes that showed no detectable genetic correlation.

On the single variant level, we only had tools available to perform a meta-analysis by looking at concordant effects⁹, we therefore had to ignore locus-specific discordant effects. Still, the strongest association of single genetic markers was observed for ADHD and ICV, additional associations were identified for ADHD and the subcortical volumes. The weighted meta-analysis of ADHD and ICV found two potentially pleiotropic loci. One of those was *SEMA6D*, coding for the semaphorin 6D, a transmembrane molecule which plays a role in maintenance and remodeling of neuronal connections²⁸. Animal studies have shown that it acts as a ligand for PlexinA1, which is involved in critical steps of neuronal development in the spinal cord²⁹ and cardiac development^{30,31}. This finding is consistent with prior research implicating neurite outgrowth dysregulation in ADHD^{24,32}, and highlighting it as potential neural mediator of genetic risk for ADHD. This should be tested in future

studies by applying imaging genetics mediation models³³ or Mendelian randomization approaches³⁴. We also found the *SEMA6D* locus in the cross-phenotype meta-analysis for ADHD and putamen volume, even though this volume had been corrected for ICV, suggesting that genetic variation in *SEMA6D* may influence specific brain regions to a different extent.

The current results raise a number of questions concerning the way that alterations in the brain mediate etiological risk pathways in ADHD. First, given the cross-sectional nature of the available data, we must be unbiased as to direction of causation. As long as we have not fully clarified the role of brain structure alterations in psychiatric disorders, i.e. whether those are cause or consequence of the disorders, we should keep an open mind regarding the direction of effects. Observed effects might not always be direct, but may involve reciprocal or transactional processes. Second, the degree of sharing was statistically modest. At first sight, this seems to be inconsistent with the general hypothesis that ADHD is a genetic-based brain disorder. However, there are a number of possible explanations for this modest amount of shared effects. First, we examined brain structure at a gross anatomical measure: atlas-based brain segmentations might be too coarse to identify subtler volumetric differences. In line with this, a surface-based analysis localized allelic effects of a SNP in *KTN1* to regions along the superior and lateral putamen⁸, suggesting that genetic variation does not affect structure across different regions of the brain in the same way. Additionally, cellular changes (e.g., altered spine density) may not manifest at such gross levels at all. Moreover, there is a limited link between structure and function at the level of the MRI-derived volumes, so other neuroimaging phenotypes (e.g., structural and functional connectivity measures) could be more informative. As pointed out previously¹⁶, the limited SNP-heritability of the (subcortical) brain volumes further challenges the identification of genetic overlap. Non-significance of some findings might also be related to sample size. However, based on the comparison of the LDSR results for ENIGMA-only and for ENIGMA+CHARGE, we can conclude that sample size affected the standard errors and therefore the significance of this analysis, but not the point estimates. On the other hand, the cross-phenotype meta-analysis results were mainly driven by the association signal of the larger GWAS-MA, namely

the ADHD data set, although several loci improved after combining the two phenotypes. However, even larger GWAS-MA data sets are becoming available for the (sub)cortical regional brain volumes³⁵, so that the disorder and brain GWAS-MAs start to become equally well-powered. Finally, it is also possible that some of ADHD's association with reduced brain volumes is driven by environmental effects, either independently or in interaction with genetic factors. Adverse prenatal and postnatal environmental exposures have been linked to both ADHD and general brain growth and development¹. In line with this idea, monozygotic twins discordant for ADHD were found to have significant epigenetic differences among genes highly expressed in the developing cerebellum, striatum, and thalamus, in addition to genes functionally related to neurodevelopment and neurotransmitter regulation³⁶.

The findings described here need to be interpreted in light of several strengths and limitations. Previous brain imaging genetics studies in ADHD mainly focused on single genetic variant investigations, as recently reviewed by Klein et al.³⁷, and were hampered by limited sample sizes. The current study combined the largest data sets available to investigate the genetic overlap between ADHD and brain volumes. For this, a complementary battery of statistical methods and analyses was used, allowing the study of genetic covariation on both global and single variant levels and across the entire genome. Nevertheless, some limitations apply. Firstly, this study focused on a limited set of mainly subcortical MRI measures, and future work should be extended to cortical regions and white matter integrity^{3,38,39}. Secondly, for the cross-phenotype GWAS-MA, we used a recently described weighted meta-analysis method⁹. However, we observed that with low and moderate genetically correlated phenotypes, the association signals did generally not improve over a naïve meta-analysis, performed without adding additional weights (data not shown). In addition, we could only investigate concordant SNP effects with this method. Novel methods for joint analysis of summary statistics from GWASs of different phenotypes⁴⁰ and co-localization approaches⁴¹ are currently being developed and might help in the future to overcome these limitations and increase power. Thirdly, it is possible that this study underestimates genetic correlations, as we did not take into account the

known role of rare variants in the genetic architecture of ADHD⁴²⁻⁴⁴. In line with the omnigenic model for complex traits⁴⁵, and as the heritability of complex traits and diseases is spread broadly across the genome^{46,47}, future studies investigating heritability and genetic correlation in low-minor allele frequency and low-LD regions may identify stronger relationships between ADHD and brain volumes.

This is the first study to show significant global and single variant level genetic correlations derived from polygenic overlap between ADHD and brain volumes. The subtle but significant genetic overlap between ADHD and variation in brain volumes is consistent with models implicating alterations in brain structure in ADHD-related genetic risk pathways and stimulate new hypotheses about neurobiological mechanisms involved in ADHD. However, the rather modest overlap reinforces the need to consider other structural and functional brain metrics and explore the role of environmental factors.

ONLINE METHODS

Ethics statement

The current study used summary statistics of GWAS meta-analyses (GWAS-MA) that had been approved by the local ethics committees and had the required informed consents, as described in the earlier publications^{8,9,13,19}.

Participant samples

ADHD Working Group of the PGC and the ADHD iPSYCH-SSI-Broad collaboration

ADHD GWAS-MA summary statistics data were acquired from the ADHD Working Group of the PGC and the ADHD iPSYCH-SSI-Broad collaboration (n=55,374⁹, <https://www.med.unc.edu/pgc/results-and-downloads>). Detailed quality control and imputation parameters are described in the original publication⁹. Briefly, genotype imputation was done using the bioinformatic pipeline “ricopili” and with the pre-phasing/imputation stepwise approach implemented in IMPUTE2/SHAPEIT using the

haplotypes from the 1000 Genomes Project, phase 3, version 5 (1KGP3v5)⁴⁸ data. Association analyses using the imputed marker dosages were performed separately for the 11 PGC samples and the 23 waves in iPSYCH by an additive logistic regression model using PLINK v1.9⁴⁹, with the derived principal components included as covariates as described in the original publication⁹. Subsequently, meta-analysis, including summary statistics from GWASs of the 23 waves in iPSYCH and 11 PGC samples, was conducted using an inverse-weighted fixed effects model. In total, 20,183 cases and 35,191 controls were used for the original analysis (**Supplementary Table 1**). Only SNPs with imputation quality (INFO score) >0.8 and MAF >0.01 were included in the meta-analysis. PGC+iPSYCH ADHD GWAS-MA summary statistics data only included markers which were supported by an effective sample size greater than 70% (8,047,420 markers)⁹.

ENIGMA

In the GWAS-MAs on subcortical volumes those volumes had been adjusted for ICV to identify specific genetic contributions to individual volumes. Five subcortical volumes and ICV were selected for this study since significant volume reductions in patients with ADHD compared to controls were reported². Access to the summary statistics of ENIGMA can be requested via their website (<http://enigma.ini.usc.edu/download-enigma-gwas-results/>). For the initial GWAS-MA analysis, MRI brain scans and genome-wide genotype data were available for 11,840 subjects from 22 cohorts. Genomic data were imputed to a reference panel (1000 Genomes, phase1, v3 (1KGP1v3)⁵⁰) comprising only European samples and with monomorphic SNPs removed. Imputation was performed at each site using MaCH for phasing and minimac for imputation⁵¹. Only SNPs with an imputation score of RSQ >0.5 and minor allele counts >10 within each site were included. Tests of association were conducted separately for eight MRI volumetric phenotypes (nucleus accumbens, amygdala, caudate nucleus, hippocampus, pallidum, putamen, thalamus and ICV) with the following covariates in a multiple linear regression framework: age, age2, sex, four MDS components (to account for population structure) and ICV (for subcortical brain phenotypes). GWA statistics from

each of the 22 sites were combined using a fixed-effect inverse variance-weighted meta-analysis as implemented in METAL⁵². Prior to all analyses, a cohort including ADHD cases (NeuroIMAGE cohort, n=154) was removed from the ENIGMA data.

CHARGE

We obtained genome-wide GWAS-MA summary statistics data on ICV and hippocampal volume from the CHARGE Consortium (n=12,803 and n=13,039, respectively^{13,19}) and CHARGE summary statistics data had been requested by the principal investigator of the study described by Adams et al.¹⁹. GWAS of ICV and hippocampal volumes were performed for each site separately, controlling for age, sex, and, when applicable age², population stratification variables, study site, and diagnosis (when applicable). Summary statistics, including effect estimates of the genetic variant with ICV or hippocampal volume under an additive model, were exchanged to perform a fixed-effects meta-analysis weighting for sample size in METAL⁵². After the final meta-analysis, variants were excluded if they were only available for fewer than 5,000 individuals.

Removal of duplicated individuals

Subject overlap between all PGC ADHD and ENIGMA cohorts was evaluated using a checksum algorithm to ensure the robustness of our results, given that some analyses could be sensitive to the presence of duplicate individuals. For each individual, ten checksum numbers were created based on ten batches of 50 SNP genotypes and compared between individuals from both consortia. Based on these comparisons no subjects needed to be removed from the data sets. As no Danish cohort was included in the ENIGMA or CHARGE study, we assumed that there is no sample overlap between cohorts studying brain volume and iPSYCH.

GWAS meta-analysis of ENIGMA and CHARGE data sets

To increase the sample size for the hippocampal volume and ICV data, summary statistics of GWAS-MA results from ENIGMA⁸ (after removal of ADHD cases) and CHARGE^{13,19} were combined using a fixed-effects sample size-weighted meta-analysis framework as implemented in METAL⁵². After the final meta-analysis, variants were excluded if they were only available for fewer than 5,000 individuals or a MAF ≤ 0.005 . After filtering, the meta-analyses results included more than 9,145,464 markers.

Linkage disequilibrium score regression (LDSR)

For LDSR, each GWAS-MA data set underwent additional filtering. Only markers overlapping with HapMap Project Phase 3 SNPs and passing the INFO score ≥ 0.9 and MAF ≥ 0.01 filters were included (where available). SNPs with missing values, duplicate rs-numbers, too low a sample size (where available SNPs with an effective sample size less than 0.67 times the 90th percentile of sample size were removed), or that were strand-ambiguous - as well as indels - were removed. As described in the original ADHD GWAS-MA paper⁹, for LDSR analysis the European only subset was used ($n_{\text{cases}}=19,099$ and $n_{\text{controls}}=34,194$), since LDSR requires linkage disequilibrium [LD] data from a sample of comparable ethnic background). For the ENIGMA amygdala results, the mean χ^2 was too low (1.0) to reliably estimate SNP heritability using LDSR.

The analyses used a two-step procedure with the LD-scoring analysis package⁵³. An unconstrained regression estimated regression intercepts for each pair of phenotypes. Since we adopted protocols to exclude sample overlap, we also performed the analyses with regression intercept for the genetic correlation analysis defined as zero (**Supplementary Table 2**). To compute p-values, standard errors were estimated using a block jackknife procedure.

SNP effect concordance analysis (SECA)

Post-processing of genetic data

To statistically compare the ADHD and six brain volume GWAS-MAs, we used SNPs passing quality control and filtering rules (for ADHD GWAS-MA INFO ≥ 0.9 and MAF ≥ 0.01 and for ENIGMA and CHARGE GWAS-MA RSQ ≥ 0.5 and MAF ≥ 0.005) in all data sets. With these data, we performed a clumping procedure in PLINK⁵⁴ to identify an independent SNP from every LD block across the genome. The clumping procedure was performed separately for each of the brain volume GWAS-MAs using a 500 kb window, with SNPs in LD ($r^2 > 0.2$) in the European reference samples from the 1KGP1v3⁵⁰. The SNP with the lowest p-value within each LD block was selected as the index SNP representing that LD block and all other SNPs in the LD block were dropped from the analysis. The result after applying the clumping procedure was sets of independent SNPs representing the total variation explained across the genome conditioned on the significance in each brain volume GWAS-MA. For each of these sets of SNPs, we then determined the corresponding ADHD GWAS-MA test statistic for each independent, index SNP and used these data sets for the subsequent analyses.

Tests of pleiotropy and concordance

We used SNP Effect Concordance Analysis (SECA)²⁰ to determine the extent and directionality of genetic overlap between ADHD and each brain volume. Within SECA we performed a global test of pleiotropy using a binomial test at 12 p-value levels: $P \leq (0.01, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9)$. For a given brain volume and ADHD paired set, we separately ordered SNPs based on their p-value for association with each trait. For each of the 12 p-value levels, we determined the total number of SNPs overlapping between the two traits at each p-value threshold and compared that number to the expected random overlap under the null hypothesis of no pleiotropy using a binomial test. In total, 144 comparisons were performed. We tallied the number of comparisons with evidence of overlap at a nominally significant level of $P \leq 0.05$. To evaluate the global level of pleiotropy, we generated 1,000 permuted data sets for a given brain volume to ADHD comparison and determined, if the number of significance thresholds with genetic overlap was significantly greater than chance.

Similarly, we estimated concordance, the agreement in SNP effect directions across two traits. We determined whether or not there was a significant ($P \leq 0.05$) positive or negative trend in the effect of the overlapping SNPs at each of the 12 p-value thresholds. This was done using a two-sided Fisher's exact test. The direction of effect for each SNP was determined by the sign of the SNP regression coefficient (OR or beta value) from each meta-analysis. In the ADHD GWAS-MA, an odds ratio >1 for a SNP indicates that the A1 reference allele was associated with an increased risk of developing ADHD (an odds ratio <1 indicates a protective allele). A positive Beta value for a SNP in a brain volume GWAS-MA indicates that the A1 reference allele of that SNP is associated with an increase in brain volume (a negative Beta value indicates that the A1 reference allele of that SNP is associated with a reduction in brain volume). We estimated the global level of concordance between a given brain trait and ADHD by generating 1,000 permuted data sets, repeating the Fisher's exact test procedure, and determined if the number of significant overlapping thresholds was significantly greater than chance (see Nyholt et al., 2014²⁰ for details of the SECA analysis).

In total, we tested for pleiotropy and concordance between ADHD and six brain volumes. In accordance with the number of tests performed, we set a Bonferroni-corrected significance level at $P=0.05/(2*6)=4.17 \times 10^{-3}$.

Independent genome-wide significant markers and loci

LD-independent markers associated at $P < 1 \times 10^{-5}$ were defined using the clump flag in PLINK v1.9⁴⁹. Clumping was used to group additional associated markers within a 0.5 Mb window surrounding the index SNP. Markers were grouped to the index SNPs if they were also associated ($P < 0.001$) and were in LD with the index SNP ($r^2 > 0.1$). A genome-wide significant locus was defined as the physical region containing the identified LD independent index SNPs and their correlated variants ($r^2 > 0.8$) with $P < 0.001$. Associated loci within 250 kb of each other were merged. All LD statistics were calculated using the 1KGP3v5⁴⁸ reference haplotypes.

SNP sign test in the top GWAS-MA findings

To investigate a potential accumulation of same- or opposite-direction effects of SNPs between ADHD and brain volumes, we counted the number of *opposite* direction effects (as expected from the imaging results in Hoogman et al.²) for the top findings from the ADHD data set in the different brain structure data sets. The ADHD GWAS-MA data were clumped to define independent loci for all variants with $P < 1 \times 10^{-5}$ in the ADHD GWAS-MA using 1KGP3v5⁴⁸ data on European ancestry populations as reference.

The proportion of variants with discordant direction of effect in the individual brain GWAS-MA was evaluated using a binomial test against a null hypothesis of 0.5 (i.e. the signs are discordant between the two analyses by chance). This test was done for loci passing p-value thresholds of $P < 5 \times 10^{-8}$ (14 LD-independent genome-wide significant SNPs), $P < 1 \times 10^{-6}$ (44 LD-independent SNPs), and $P < 1 \times 10^{-5}$ (132 LD-independent SNPs) in the ADHD GWAS-MA.

Weighted meta-analysis of ADHD and brain volume data sets

Independent of the results of the global overlap analyses, we also performed meta-analyses combining the results from the ADHD GWAS-MA with results from GWAS-MAs of brain volumes (amygdala, nucleus accumbens, caudate nucleus, hippocampus, putamen, and ICV). This was done using a modified sample size-based weighting method, integrating the binary ADHD trait (ADHD risk) with the continuous trait (brain volume traits), as described in Demontis et al.⁹. For the meta-analyses, modified sample size-based weights were derived to account for the respective heritability, genetic correlation, and measurement scale of the GWASs. The adjusted samples sizes reflect differences in power between the studies due to measurement scale and relative heritability that is not captured by sample size. Thereby, the contribution of the continuous phenotype's GWAS to the meta-analysis is reduced based on imperfect correlation with the dichotomous phenotype of interest (in this case ADHD risk). The adjustments are computed based on the sample and population prevalence of the dichotomous phenotype, the estimated SNP heritability of the two phenotypes

(liability scale for dichotomous phenotype), and the genetic correlation between the two phenotypes, as well as the average SNP LD score, and the number of SNPs. Heritability and genetic correlation values to compute these weights are computed using LD score regression⁵³ as described before. For a comprehensive description of the method for meta-analysis of continuous and dichotomous phenotype and notes on the implementation please see the supplementary information of the original ADHD GWAS-MA publication⁹. For all brain volumes, we also performed naive meta-analyses given the low genetic correlations with ADHD risk observed. Correcting for meta-analyzing six brain phenotypes with ADHD, we set the threshold for genome-wide significance at $P=5 \times 10^{-8}/6=8.33 \times 10^{-9}$.

Reciprocal lookup of significant GWAS-MA loci

Evidence for an effect of the independent ADHD-associated SNPs on brain volume was studied through a look-up of results in the ENIGMA(+CHARGE) GWAS-MAs. LD-independent loci with corresponding index SNPs were obtained by clumping summary statistics of the ADHD GWAS-MA⁹. Similarly, effects of the 21 independent SNPs significantly associated with brain volumes as reported in the original publications of the brain volume GWAS-MAs^{8,13,19} on ADHD risk were looked-up in the ADHD GWAS-MA data. If the index variant was not present in the other data set, a proxy variant was selected by using LDlink (<https://analysistools.nci.nih.gov/LDlink/>). Given the number of tests performed, we set a Bonferroni-corrected significance levels at $P=0.05/(14 \times 8)=0.000446$ for the look-up of ADHD SNPs in brain volume GWAS-MAs and at $P=0.05/21=0.002381$ for brain volume SNPs in ADHD GWAS-MA data.

Expression quantitative trait loci and brain gene expression

To assess potential functionality in (brain) tissues, we tested the identified risk variants (**Table 3**) for association with gene expression. Expression quantitative trait loci (eQTL) were examined using data

from the GTEx portal (<https://www.gtexportal.org/home/>)²¹. In addition, blood eQTL data were queried using the Blood eQTL Browser (<http://genenetwork.nl/bloodeqtlbrowser/>)²².

We also investigated the spatio-temporal expression pattern in brain tissue for genes with significantly associated variants in the approaches described above (**Table 3**) using data from the Human Brain Transcriptome Project (<http://hbatlas.org>). We assessed messenger RNA (mRNA) expression trajectories in six regions of the developing and adult human brain. Spanning periods from embryonic development to late adulthood, this data set provides genome-wide exon-level transcriptome data generated using the Affymetrix GeneChip Human Exon 1.0 SS Arrays from over 1,340 tissue samples sampled from both hemispheres of *postmortem* human brains (n=57)²³. Gene expression over the lifespan from the spatio-temporal atlas was graphed using custom R scripts²³.

Gene-based and gene-set analyses for ADHD+ICV GWAS-MA data

Based on our finding that *SEMA6D* is a key locus contributing to both ADHD risk and ICV – a loci involved in neuronal migration and axonal path finding – we investigated, whether neurite outgrowth-related genes in general have a role in ADHD–ICV genetic overlap. Genome-wide summary statistics of the (i) weighted meta-analysis for ADHD and ENIGMA+CHARGE ICV GWAS-MA, (ii) the ADHD GWAS-MA, and (iii) the ENIGMA+CHARGE GWAS-MA on ICV were used as input for gene-based and gene-set analyses. For the ADHD+ICV GWAS-MA, only SNPs shared between the ADHD and ICV data set were included. Statistical analyses used the Multi-marker Analysis of GenoMic Annotation (MAGMA) software package (version 1.05)⁵⁵. Genome-wide SNP data from a reference panel 1KGP1v3⁵⁰ was annotated to NCBI Build 37.3 gene locations using a symmetric 100 kb flanking window. The gene annotation file was used to map genome-wide SNP data from the different studies, to assign SNPs to genes followed by the calculation of gene-based p-values. For the gene-based analyses, single SNP p-values within a gene were transformed into a gene-statistic by taking the mean of the χ^2 -statistic among the SNPs in each gene. To account for LD, the 1KGP1v3⁵⁰ was used as a reference to estimate the LD between SNPs within (the vicinity of) the genes

(http://ctglab.nl/software/MAGMA/ref_data/g1000_ceu.zip). Subsequent to the genome-wide gene-based analysis, we also tested, whether genes in the neurite-outgrowth gene-set (defined previously, $n_{\text{genes}}=45$; ref. ²⁴) were jointly associated with results of the weighted meta-analytic data of ADHD+ICV using self-contained and competitive testing⁵⁶. Since self-contained tests do not take into account the overall level of association across the genome, gene size, and gene density, we were particularly interested in the competitive test for the current analysis. Subsequently, we tested whether the gene-set was associated with the two individual data sets as well. In this, the same procedure was followed for analysis of the ADHD GWAS-MA and ENIGMA+CHARGE ICV GWAS-MA summary statistics individually. *Post-hoc*, the individual genes in the set were investigated, by reviewing gene test-statistics of the weighted ADHD+ICV GWAS-MA results. Genes were considered gene-wide significant, if they reached the adjusted Bonferroni correction threshold ($P=0.05/45=0.00111$). Subsequently, we reviewed gene-based associations in the ADHD GWAS-MA and ENIGMA+CHARGE ICV GWAS-MA results separately.

URLs

<http://enigma.ini.usc.edu/download-enigma-gwas-results/>

<https://www.med.unc.edu/pgc/results-and-downloads>

<https://github.com/bulik/ldsc>

<https://neurogenetics.qimrberghofer.edu.au/SECA/>

<https://analysistools.nci.nih.gov/LDlink/>

<https://www.gtexportal.org/home/>

<http://hbatlas.org>

http://ctglab.nl/software/MAGMA/ref_data/g1000_ceu.zip

<http://locuszoom.sph.umich.edu/>

DATA AVAILABILITY

The genome-wide summary statistics that support the findings of this study are available at the consortia websites.

PGC ADHD working group and the ADHD iPSYCH-SSI-Broad collaboration: <https://www.med.unc.edu/pgc/results-and-downloads>

ENIGMA and ENIGMA+CHARGE for ICV and hippocampus: <http://enigma.ini.usc.edu/download-enigma-gwas-results/>

ACKNOWLEDGMENTS

Funding: This work was supported by the Netherlands Organization for Scientific Research (NWO), i.e. the NWO Brain & Cognition Excellence Program (grant 433-09-229) and the Vici Innovation Program (grant 016-130-669 to BF). Additional support was received from the European Community's Seventh Framework Programme (FP7/2007 – 2013) under grant agreements n° 602805 (Aggressotype), n° 602450 (IMAGEMEND), and n° 278948 (TACTICS) as well as from the European Community's Horizon 2020 Programme (H2020/2014 – 2020) under grant agreements n° 643051 (MiND) and n° 667302 (CoCA). The work was also supported by the Lundbeck Foundation (grant R102-A9118 and R155-2014-1724 to iPSYCH/A.D.B.), by a grant to the ENIGMA Consortium (grant number U54 EB020403) from the BD2K Initiative, a cross-NIH partnership, and by the ECNP Network 'ADHD across the Lifespan'. SEM is supported by an Australian National Health and Medical Research Council fellowship (SRFB-1103623). This work was partly carried out on the Dutch national e-infrastructure with the support of SURF Foundation.

The Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA) Consortium provided summary statistics of the consortium findings to this project. The original publication of those findings as well as the list of contributing samples and authors may be found on the ENIGMA website: <http://enigma.ini.usc.edu>. The Neurology Working Group of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium also contributed with an independent set

of summary statistics of consortium findings. The contributing cohorts and authors may be found on previous publications of the CHARGE consortium as listed on the website: <http://www.chargeconsortium.com/>. The ADHD working group of the Psychiatric Genomics Consortium (PGC) and the iPSYCH-SSI-Broad collaboration ADHD Working Group contributed with an independent set of summary of consortium findings. The data and a complete list of contributing samples and people can be obtained from the PGC download website (<https://www.med.unc.edu/pgc/results-and-downloads>).

AUTHOR CONTRIBUTIONS

Study conception and supervision: M.K., A.A.-V., B.F. Obtained funding: B.F., A.D.B. Provided samples and/or data: D.D, D.P.H., H.H.A., M.M., S.E.M. Conducted analyses: M.K. Contributed to analyses and data interpretation: R.K.W., J.L.S. Writing group: M.K., S.V.F., A.A.-V., B.F., A.D.B., J.B., N.R.M., R.S., E.S.-B., P.M.T. All authors reviewed and approved the final version of this manuscript.

CONFLICT OF INTEREST

Barbara Franke discloses having received educational speaking fees from Merz and Shire. Derrek P Hibar is now an employee of Janssen Research and Development, LLC. None of the other authors report conflicts of interest.

REFERENCES

- 1 Faraone SV, Asherson P, Banaschewski T, Biederman J, Buitelaar JK, Ramos-Quiroga JA et al. Attention-deficit/hyperactivity disorder. *Nature Reviews Disease Primers* 2015: 15020.
- 2 Hoogman M, Bralten J, Hibar DP, Mennes M, Zwiers MP, Schweren LS et al. Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis. *Lancet Psychiatry* 2017.
- 3 Bralten J, Greven CU, Franke B, Mennes M, Zwiers MP, Rommelse NN et al. Voxel-based morphometry analysis reveals frontal brain differences in participants with ADHD and their unaffected siblings. *J Psychiatry Neurosci* 2016; 41: 272-279.
- 4 Greven CU, Bralten J, Mennes M, O'Dwyer L, van Hulzen KJ, Rommelse N et al. Developmentally stable whole-brain volume reductions and developmentally sensitive caudate and putamen volume alterations in those with attention-deficit/hyperactivity disorder and their unaffected siblings. *JAMA Psychiatry* 2015; 72: 490-499.
- 5 Chang Z, Lichtenstein P, Asherson PJ & Larsson H. Developmental twin study of attention problems: high heritabilities throughout development. *JAMA Psychiatry* 2013; 70: 311-318.
- 6 Faraone SV, Perlis RH, Doyle AE, Smoller JW, Goralnick JJ, Holmgren MA et al. Molecular genetics of attention-deficit/hyperactivity disorder. *Biol Psychiatry* 2005; 57: 1313-1323.
- 7 Larsson H, Chang Z, D'Onofrio BM & Lichtenstein P. The heritability of clinically diagnosed attention deficit hyperactivity disorder across the lifespan. *Psychol Med* 2014; 44: 2223-2229.
- 8 Hibar DP, Stein JL, Renteria ME, Arias-Vasquez A, Desrivieres S, Jahanshad N et al. Common genetic variants influence human subcortical brain structures. *Nature* 2015; 520: 224-229.
- 9 Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E et al. Discovery Of The First Genome-Wide Significant Risk Loci For ADHD. *bioRxiv* 2017.
- 10 Stein JL, Medland SE, Vasquez AA, Hibar DP, Senstad RE, Winkler AM et al. Identification of common variants associated with human hippocampal and intracranial volumes. *Nat Genet* 2012; 44: 552-561.

- 11 Bis JC, DeCarli C, Smith AV, van der Lijn F, Crivello F, Fornage M et al. Common variants at 12q14 and 12q24 are associated with hippocampal volume. *Nat Genet* 2012; 44: 545-551.
- 12 Melville SA, Buros J, Parrado AR, Vardarajan B, Logue MW, Shen L et al. Multiple loci influencing hippocampal degeneration identified by genome scan. *Ann Neurol* 2012; 72: 65-75.
- 13 Hibar DP, Adams HH, Jahanshad N, Chauhan G, Stein JL, Hofer E et al. Novel genetic loci associated with hippocampal volume. *Nat Commun* 2017; 8: 13624.
- 14 Ikram MA, Fornage M, Smith AV, Seshadri S, Schmidt R, Debette S et al. Common variants at 6q22 and 17q21 are associated with intracranial volume. *Nat Genet* 2012; 44: 539-544.
- 15 Adams HH, Hibar DP, Chouraki V, Stein JL, Nyquist PA, Renteria ME et al. Novel genetic loci underlying human intracranial volume identified through genome-wide association. *Nat Neurosci* 2016.
- 16 Franke B, Stein JL, Ripke S, Anttila V, Hibar DP, van Hulzen KJ et al. Genetic influences on schizophrenia and subcortical brain volumes: large-scale proof of concept. *Nat Neurosci* 2016; 19: 420-431.
- 17 Thompson PM, Stein JL, Medland SE, Hibar DP, Vasquez AA, Renteria ME et al. The ENIGMA Consortium: large-scale collaborative analyses of neuroimaging and genetic data. *Brain Imaging Behav* 2014; 8: 153-182.
- 18 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.
- 19 Adams HH, Hibar DP, Chouraki V, Stein JL, Nyquist PA, Renteria ME et al. Novel genetic loci underlying human intracranial volume identified through genome-wide association. *Nat Neurosci* 2016; 19: 1569-1582.

- 20 Nyholt DR. SECA: SNP effect concordance analysis using genome-wide association summary results. *Bioinformatics* 2014; 30: 2086-2088.
- 21 GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013; 45: 580-585.
- 22 Westra HJ, Peters MJ, Esko T, Yaghootkar H, Schurmann C, Kettunen J et al. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* 2013; 45: 1238-1243.
- 23 Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M et al. Spatio-temporal transcriptome of the human brain. *Nature* 2011; 478: 483-489.
- 24 Poelmans G, Pauls DL, Buitelaar JK & Franke B. Integrated genome-wide association study findings: identification of a neurodevelopmental network for attention deficit hyperactivity disorder. *Am J Psychiatry* 2011; 168: 365-377.
- 25 Wigmore EM, Clarke T-K, Adams MJ, Fernandez-Pujals AM, Gibson J, Davies G et al. Do Regional Brain Volumes and Major Depressive Disorder Share Genetic Architecture: a study in Generation Scotland (n=19,762), UK Biobank (n=24,048) and the English Longitudinal Study of Ageing (n=5,766). *bioRxiv* 2016.
- 26 Trampush JW, Yang ML, Yu J, Knowles E, Davies G, Liewald DC et al. GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive function: a report from the COGENT consortium. *Mol Psychiatry* 2017; 22: 336-345.
- 27 Sniekers S, Stringer S, Watanabe K, Jansen PR, Coleman JRI, Krapohl E et al. Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. *Nat Genet* 2017; 49: 1107-1112.
- 28 Qu X, Wei H, Zhai Y, Que H, Chen Q, Tang F et al. Identification, characterization, and functional study of the two novel human members of the semaphorin gene family. *J Biol Chem* 2002; 277: 35574-35585.

- 29 Yoshida Y, Han B, Mendelsohn M & Jessell TM. PlexinA1 signaling directs the segregation of proprioceptive sensory axons in the developing spinal cord. *Neuron* 2006; 52: 775-788.
- 30 Toyofuku T, Zhang H, Kumanogoh A, Takegahara N, Suto F, Kamei J et al. Dual roles of Sema6D in cardiac morphogenesis through region-specific association of its receptor, Plexin-A1, with off-track and vascular endothelial growth factor receptor type 2. *Genes Dev* 2004; 18: 435-447.
- 31 Toyofuku T, Zhang H, Kumanogoh A, Takegahara N, Yabuki M, Harada K et al. Guidance of myocardial patterning in cardiac development by Sema6D reverse signalling. *Nat Cell Biol* 2004; 6: 1204-1211.
- 32 Mooney MA, McWeeney SK, Faraone SV, Hinney A, Hebebrand J, Nigg JT et al. Pathway analysis in attention deficit hyperactivity disorder: An ensemble approach. *Am J Med Genet B Neuropsychiatr Genet* 2016; 171: 815-826.
- 33 van der Meer D, Hoekstra PJ, Zwiers M, Mennes M, Schweren LJ, Franke B et al. Brain Correlates of the Interaction Between 5-HTTLPR and Psychosocial Stress Mediating Attention Deficit Hyperactivity Disorder Severity. *Am J Psychiatry* 2015; 172: 768-775.
- 34 Pickrell JK, Berisa T, Liu JZ, Segurel L, Tung JY & Hinds DA. Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet* 2016; 48: 709-717.
- 35 Satizabal CL, Adams HHH, Hibar DP, White CC, Stein JL, Scholz M et al. Genetic Architecture of Subcortical Brain Structures in Over 40,000 Individuals Worldwide. *bioRxiv* 2017.
- 36 Chen YC, Sudre G, Sharp W, Donovan F, Chandrasekharappa SC, Hansen N et al. Neuroanatomic, epigenetic and genetic differences in monozygotic twins discordant for attention deficit hyperactivity disorder. *Mol Psychiatry* 2017.
- 37 Klein M, Onnink M, van Donkelaar M, Wolfers T, Harich B, Shi Y et al. Brain imaging genetics in ADHD and beyond - Mapping pathways from gene to disorder at different levels of complexity. *Neurosci Biobehav Rev* 2017; 80: 115-155.

- 38 Onnink AM, Zwiers MP, Hoogman M, Mostert JC, Dammers J, Kan CC et al. Deviant white matter structure in adults with attention-deficit/hyperactivity disorder points to aberrant myelination and affects neuropsychological performance. *Prog Neuropsychopharmacol Biol Psychiatry* 2015; 63: 14-22.
- 39 Wu ZM, Bralten J, Cao QJ, Hoogman M, Zwiers MP, An L et al. White Matter Microstructural Alterations in Children with ADHD: Categorical and Dimensional Perspectives. *Neuropsychopharmacology* 2017; 42: 572-580.
- 40 Turley P, Walters RK, Maghzian O, Okbay A, Lee JJ, Fontana MA et al. MTAG: Multi-Trait Analysis of GWAS. *bioRxiv* 2017.
- 41 Giambartolomei C, Vukcevic D, Schadt EE, Franke L, Hingorani AD, Wallace C et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet* 2014; 10: e1004383.
- 42 Demontis D, Lescai F, Borglum A, Glerup S, Ostergaard SD, Mors O et al. Whole-Exome Sequencing Reveals Increased Burden of Rare Functional and Disruptive Variants in Candidate Risk Genes in Individuals With Persistent Attention-Deficit/Hyperactivity Disorder. *J Am Acad Child Adolesc Psychiatry* 2016; 55: 521-523.
- 43 Williams NM, Zaharieva I, Martin A, Langley K, Mantripragada K, Fossdal R et al. Rare chromosomal deletions and duplications in attention-deficit hyperactivity disorder: a genome-wide analysis. *Lancet* 2010; 376: 1401-1408.
- 44 Zayats T, Jacobsen KK, Kleppe R, Jacob CP, Kittel-Schneider S, Ribases M et al. Exome chip analyses in adult attention deficit hyperactivity disorder. *Transl Psychiatry* 2016; 6: e923.
- 45 Boyle EA, Li YI & Pritchard JK. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* 2017; 169: 1177-1186.
- 46 Loh PR, Bhatia G, Gusev A, Finucane HK, Bulik-Sullivan BK, Pollack SJ et al. Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance-components analysis. *Nat Genet* 2015; 47: 1385-1392.

- 47 Shi H, Kichaev G & Pasaniuc B. Contrasting the Genetic Architecture of 30 Complex Traits from Summary Association Data. *Am J Hum Genet* 2016; 99: 139-153.
- 48 The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* 2015; 526: 68-74.
- 49 Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM & Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 2015; 4: 7.
- 50 Genomes Project Consortium, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM et al. A map of human genome variation from population-scale sequencing. *Nature* 2010; 467: 1061-1073.
- 51 Fuchsberger C, Abecasis GR & Hinds DA. minimac2: faster genotype imputation. *Bioinformatics* 2015; 31: 782-784.
- 52 Willer CJ, Li Y & Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* 2010; 26: 2190-2191.
- 53 Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics C et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* 2015; 47: 291-295.
- 54 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; 81: 559-575.
- 55 de Leeuw CA, Mooij JM, Heskes T & Posthuma D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* 2015; 11: e1004219.
- 56 de Leeuw CA, Neale BM, Heskes T & Posthuma D. The statistical properties of gene-set analysis. *Nat Rev Genet* 2016; 17: 353-364.
- 57 Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010; 26: 2336-2337.

DISPLAY ITEMS (2 figures, 3 tables)

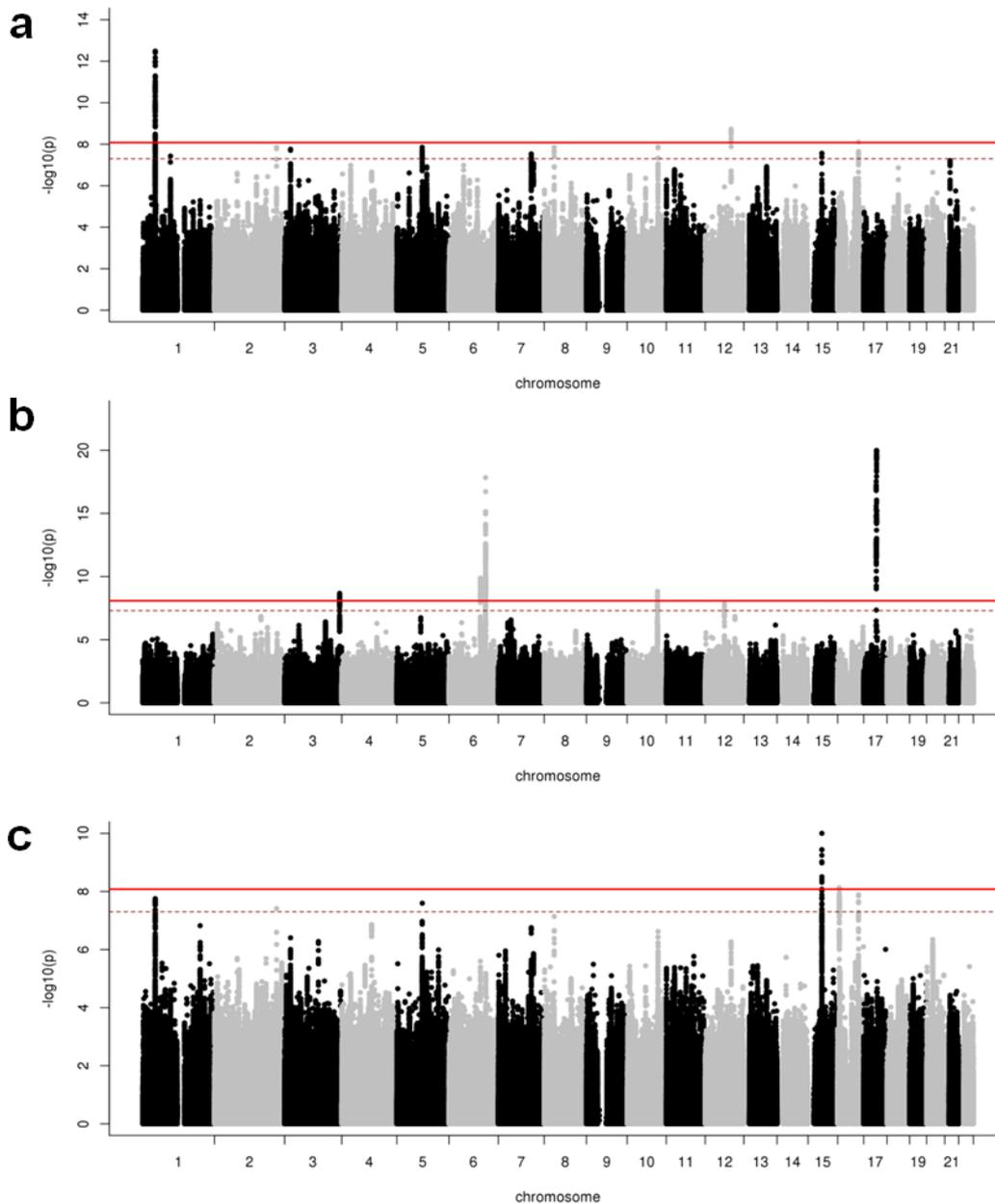


Figure 1: Common genetic variants associated with ADHD, ICV and ADHD+ICV. Shown here are Manhattan plots, in which every point represents a single genetic variant plotted according to its genomic position (x-axis) and its $-\log_{10}(P)$ for association with the respective trait (y-axis). The solid bright red line represents the study-wide genome-wide significance of $P < 8.33 \times 10^{-9}$, and the dashed dark red line represents the genome-wide significance of $P < 5 \times 10^{-8}$. **(a)** PGC+iPSYCH ADHD GWAS-MA. **(b)** ENIGMA+CHARGE ICV GWAS-MA. **(c)** ADHD+ICV weighted GWAS-MA.

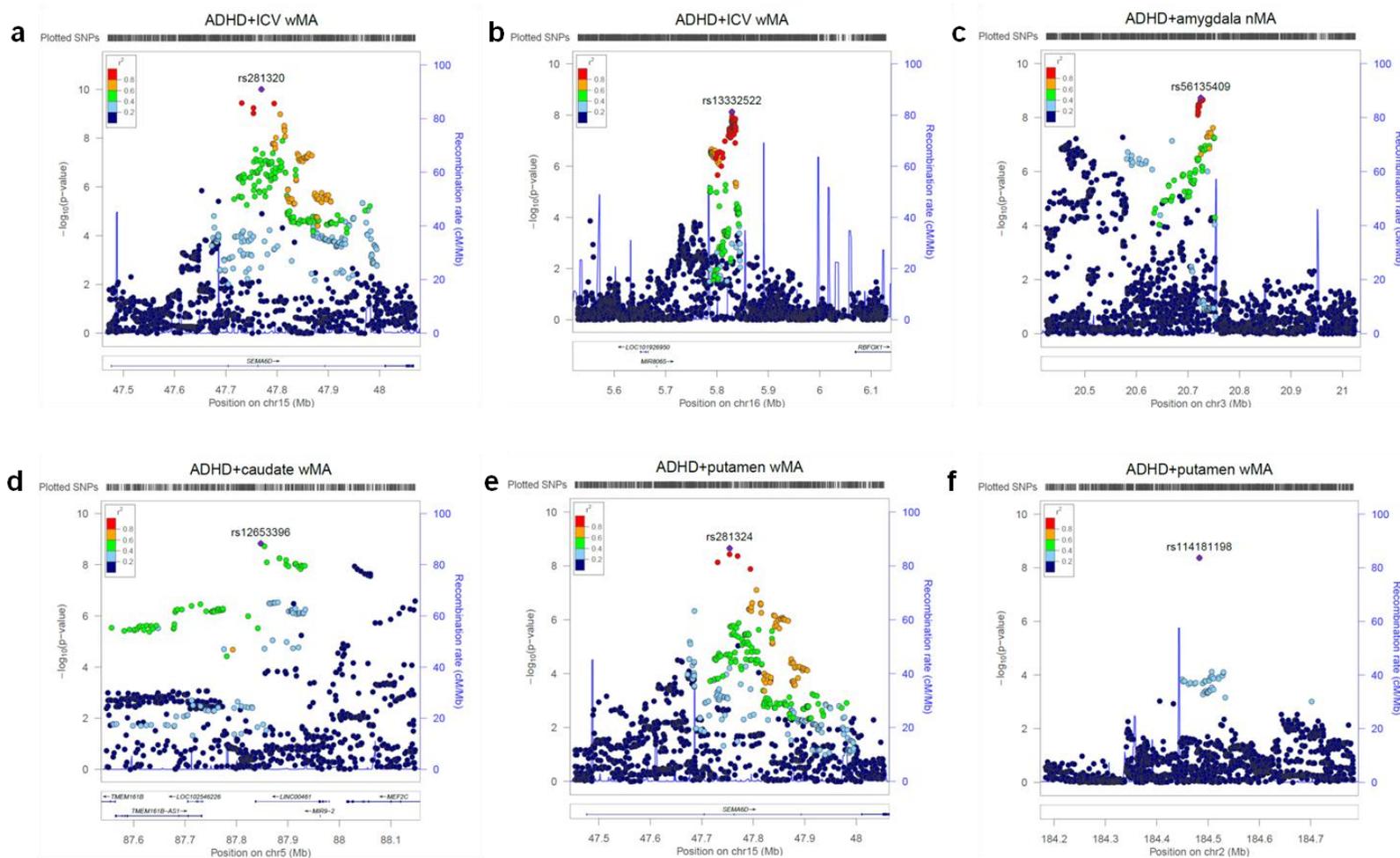


Figure 2: Regional association of genome-wide significant loci of ADHD and brain volume GWAS-MAs. For each panel (a-f), zoomed-in association plots (\pm 300 kb from the top SNP, indexed by purple dot) are shown. Plots are zoomed to highlight the genomic region that likely harbors the causal variant(s). Plots were made using the Locuszoom software⁵⁷.

Table 1. SNP heritability analyses for MRI brain volumes and genetic correlation with ADHD*.

| Brain region | N | Heritability | SE | Genetic correlation with ADHD | SE | Z | P |
|----------------------------------|--------|--------------|--------|-------------------------------|---------|---------|------------------|
| Nucleus accumbens | 11,709 | 0.1332 | 0.0518 | 0.005558 | 0.09487 | 0.05858 | 0.9533 |
| Caudate nucleus | 11,772 | 0.2456 | 0.0455 | -0.06426 | 0.07321 | -0.8778 | 0.3801 |
| Hippocampus [#] | 24,704 | 0.1418 | 0.0286 | -0.02354 | 0.06902 | -0.3411 | 0.733 |
| Intracranial volume [#] | 24,024 | 0.2318 | 0.0325 | -0.2266 | 0.05989 | -3.784 | 0.0001543 |
| Putamen | 11,646 | 0.2815 | 0.056 | 0.006433 | 0.07077 | 0.09089 | 0.9276 |
| Hippocampus ENIGMA only | 11,665 | 0.1363 | 0.0488 | -0.04202 | 0.09965 | -0.4217 | 0.6733 |
| Intracranial volume ENIGMA only | 11,221 | 0.1745 | 0.0461 | -0.2348 | 0.09452 | -2.485 | 0.01296 |

*Amygdala mean χ^2 was too small to allow a valid analysis (N=11,757). [#]Using GWAS-MA summary statistics from the meta-analysis of ENIGMA and CHARGE cohorts. Heritability and genetic correlation were estimated by using free intercepts. In accordance with the number of tests performed, we set a Bonferroni-corrected significance level at $P=0.05/6=0.0083$. P-values in bold are significant after Bonferroni correction.

Table 2. Results of pleiotropy and concordance test of SNP Effect Concordance Analysis. Brain volume GWAS-MA was conditioned on ADHD GWAS-MA.

| Brain volume | P _{pleiotropy} | CI _{pleiotropy} | P _{concordance} | CI _{concordance} | Direction of SNP effects |
|----------------------------------|-------------------------|--------------------------------|--------------------------|--------------------------------|--------------------------|
| Nucleus accumbens | 0.034 | 0.0244-0.0471 | 0.002 | 0.000548-0.00726 | concordant |
| Amygdala | <0.001 | 5.12x10 ⁻⁵ -0.00564 | 0.006 | 0.00275-0.013 | discordant |
| Caudate nucleus | <0.001 | 5.12x10 ⁻⁵ -0.00564 | 0.004 | 0.00156-0.0102 | concordant |
| Hippocampus [#] | 0.002 | 0.000548-0.00726 | 1 | 0.996-1 | / |
| Intracranial volume [#] | <0.001 | 5.12x10 ⁻⁵ -0.00564 | <0.001 | 5.12x10 ⁻⁵ -0.00564 | discordant |
| Putamen | <0.001 | 5.12x10 ⁻⁵ -0.00564 | 0.01 | 0.00544-0.0183 | concordant |
| Hippocampus ENIGMA only | 0.005 | 0.00214-0.0116 | 1 | 0.996-1 | / |
| Intracranial volume ENIGMA only | <0.001 | 5.12x10 ⁻⁵ -0.00564 | <0.001 | 5.12x10 ⁻⁵ -0.00564 | discordant |

P-values and confidence intervals (CI) were obtained based on 1,000 permutations. [#]Using GWAS-MA summary statistics from the meta-analysis of ENIGMA and CHARGE cohorts. In accordance with the number of tests performed, we set a Bonferroni-corrected significance level at $P=0.05/(2*6)=0.00416$. P-values in bold are significant after Bonferroni correction.

Table 3: Results of (weighted) meta-analyses of ADHD and brain volumes showing novel or improved independent genome-wide significant loci.

| Brain trait | Chr | Pos | Gene | SNP | A1 | Zscore | P-value | Direction ADHD / brain | Zscore _{ADHD} | Zscore _{brain} |
|---------------------|-----|-----------|------------------|--------------------------|----|---------|------------------------|------------------------|------------------------|-------------------------|
| Intracranial volume | 15 | 47769424 | <i>SEMA6D</i> | rs281320 ^a | T | -6.4677 | 9.95x10 ⁻¹¹ | -/- | -5.5487 | -3.33 |
| Intracranial volume | 16 | 5829204 | intergenic | rs13332522 ^b | C | 5.7810 | 7.43x10 ⁻⁹ | +/+ | 4.6162 | 3.508 |
| Amygdala | 3 | 20725016 | intergenic | rs56135409 ^a | A | -6.0105 | 1.85x10 ⁻⁹ | -/- | -5.3525 | -2.7482 |
| Caudate nucleus | 5 | 87847273 | <i>LINC00461</i> | rs12653396 ^b | A | 5.3826 | 1.46 x10 ⁻⁹ | +/+ | 5.3777 | 2.9499 |
| Putamen | 15 | 47754027 | <i>SEMA6D</i> | rs281323 ^a | A | 5.9005 | 3.62 x10 ⁻⁹ | +/+ | 5.4775 | 2.2114 |
| Putamen | 2 | 184483159 | intergenic | rs114181198 ^b | A | 5.1308 | 4.28 x10 ⁻⁹ | +/+ | 5.0859 | 3.2596 |

Results for six genome-wide significant index SNPs identified in the GWAS meta-analyses of ADHD+brain volumes. The allele (A1) for the effect of the Z-score is given. Direction gives information about the direction of the association in the two samples, “+” indicates that A1 increases risk/volume and “-” indicates that A1 decreases risk/volume. The threshold for genome-wide significance was set at $P < 8.33 \times 10^{-9}$ ($5 \times 10^{-8} / 6$ traits). ^aThis locus improved by one order of magnitude after the meta-analysis. ^bThis locus is a new genome-wide significant locus in the ADHD+brain volume meta-analysis.