University of Groningen

Genomic patterns linked to gray matter alterations underlying working memory deficits in adults and adolescents with attention-deficit/hyperactivity disorder

Duan, Kuaikuai; Chen, Jiayu; Calhoun, Vince D; Jiang, Wenhao; Rootes-Murdy, Kelly; Schoenmacker, Gido; Silva, Rogers F; Franke, Barbara; Buitelaar, Jan K; Hoogman, Martine

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Link to publication in University of Groningen/UMCG research database

Check for updates

ARTICLE   OPEN

Genomic patterns linked to gray matter alterations underlying working memory deficits in adults and adolescents with attention-deficit/hyperactivity disorder

Kuaikuai Duan [1,2], Jiayu Chen [2,3], Vince D. Calhoun [1,2,3,4], Wenhao Jiang[4], Kelly Rootes-Murdy[4], Gido Schoenmacker [5], Rogers F. Silva [2], Barbara Franke [5,6], Jan K. Buitelaar [7], Martine Hoogman[5,6], Jaap Oosterlaan[8], Pieter J. Hoekstra[9], Dirk Heslenfeld[10], Catharina A. Hartman [11], Emma Sprooten[7], Alejandro Arias-Vasquez[5,6], Jessica A. Turner [2,4] and Jingyu Liu [2,3] ✉

Attention-deficit/hyperactivity disorder (ADHD) is a highly heritable neurodevelopmental disorder, with onset in childhood and a considerable likelihood to persist into adulthood. Our previous work has identified that across adults and adolescents with ADHD, gray matter volume (GMV) alteration in the frontal cortex was consistently associated with working memory underperformance, and GMV alteration in the cerebellum was associated with inattention. Recent knowledge regarding ADHD genetic risk loci makes it feasible to investigate genomic factors underlying these persistent GMV alterations, potentially illuminating the pathology of ADHD persistence. Based on this, we applied a sparsity-constrained multivariate data fusion approach, sparse parallel independent component analysis, to GMV variations in the frontal and cerebellum regions and candidate risk single nucleotide polymorphisms (SNPs) data from 341 unrelated adult participants, including 167 individuals with ADHD, 47 unaffected siblings, and 127 healthy controls. We identified one SNP component significantly associated with one GMV component in superior/middle frontal regions and replicated this association in 317 adolescents from ADHD families. The association was stronger in individuals with ADHD than in controls, and stronger in adults and older adolescents than in younger ones. The SNP component highlights 93 SNPs in long non-coding RNAs mainly in chromosome 5 and 21 protein-coding genes that are significantly enriched in human neuron cells. Eighteen identified SNPs have regulation effects on gene expression, transcript expression, isoform percentage, or methylation level in frontal regions. Identified genes highlight MEF2C, CADM2, and CADPS2, which are relevant for modulating neuronal substrates underlying high-level cognition in ADHD, and their causality effects on ADHD persistence await further investigations. Overall, through a multivariate analysis, we have revealed a genomic pattern underpinning the frontal gray matter variation related to working memory deficit in ADHD.

INTRODUCTION

Attention-deficit/hyperactivity disorder (ADHD) is a neurodevelopmental disorder characterized by pervasive symptoms of inattention and/or hyperactivity/impulsivity that affect normal functioning and development [1]. About 9.8% of children aged 3–17 years are estimated to have ADHD in the United States [2]. Based on long-term follow-up studies, 78% of boys [3] and 77.1% of girls [4] with ADHD showed persistence of ADHD into adulthood. Beyond symptoms, cognitive impairments are common in adults with ADHD, as indicated by studies of IQ [5], working memory [6, 7], response inhibition [8], executive functions [9], and educational outcomes [3, 4, 10]. The persistence of ADHD

symptoms and associated cognitive impairments significantly elevates the risk of adverse outcomes for individuals with ADHD and increases the burden for the involved families and the whole society. Neuronal and genetic determinants of ADHD in adults remain largely unknown. Characterizing the underlying biological mechanisms would help delineate the pathology of ADHD and aid early interventions.

Converging evidence from anatomical studies suggest that children/adolescents with ADHD have global gray matter volume (GMV) reductions [11–15] and regional GMV reductions in the basal ganglia [12, 16–18], the frontal lobe [13, 18–20], the anterior cingulate cortex [21], and the cerebellum [11, 14, 18, 22]. Adults

[1]Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA. [2]Tri-Institutional Center for Translational Research in Neuroimaging and Data Science (TReNDS), Atlanta, GA, USA. [3]Department of Computer Science, Georgia State University, Atlanta, GA, USA. [4]Department of Psychology, Georgia State University, Atlanta, GA, USA. [5]Department of Human Genetics, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center, Nijmegen, The Netherlands. [6]Department of Psychiatry, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center, Nijmegen, The Netherlands. [7]Department of Cognitive Neuroscience, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center, Nijmegen, The Netherlands. [8]Department of Clinical Neuropsychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands. [9]Department of Child and Adolescent Psychiatry, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands. [10]Faculty of Behavioral and Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands. [11]Department of Psychiatry, University of Groningen, Groningen, The Netherlands. ✉email: jliu75@gsu.edu

with ADHD also showed reduced total (cortical) GMV [23–25], reduced GMV in the frontal lobe [23, 25–28], caudate nucleus [29, 30], and cingulate cortices [16, 25, 27, 31], but the reduction is less conclusive as the recent ENIGMA-ADHD studies revealed no significant ADHD versus control differences for adults [32, 33]. Meanwhile, studies on ADHD-related persistent brain alterations are scarce, and the reported findings are inconsistent[30, 34].

Despite the persistent symptom and cognitive impairments in ADHD patients, the associated brain substrates remain understudied and inconsistent. Smaller GMV in the cerebellum was related to severe ADHD symptom (especially inattention) in children, adolescents [35, 36] and adults (18–63 years old) [28, 36] with ADHD. While in elderly (≥65 years old) patients with ADHD, greater GMV in the left cerebellum was associated with severer symptom [23]. Although the fronto-parietal network is well documented underlying working memory functions [37–45], its involvement in ADHD-related working memory deficit is understudied and inconclusive. Smaller GMV in the frontal region was related to poorer working memory performance in adults with ADHD [28]. But this relationship was absent in children with ADHD. Instead, smaller GMV [46] and thinner cortical thickness [47] in the temporal region were associated with poorer working memory performance in children with ADHD. The inconsistent results may be due to small sample sizes, methodological differences, and heterogeneity of the disorder [48, 49]. Recently, we leveraged structural magnetic resonance imaging (sMRI) data from reasonable large adult and adolescent samples to study brain alterations underlying persistent ADHD symptoms and cognitive impairments using consistent analytical approaches. We identified that GMV of the superior/middle/inferior frontal regions was consistently associated with working memory performance, and cerebellum GMV was consistently related to inattention symptoms in both adults and adolescents with ADHD [22, 28, 36]. Our findings emphasize that not only the symptom and impaired cognitive functions but also the underlying neurological alterations persist from childhood to adulthood in individuals with ADHD.

Since ADHD is highly heritable (estimated heritability: 77–88% [50]) and 12 independent single nucleotide polymorphism (SNP) loci showed significant differences between individuals with ADHD and controls in a recent ADHD genome-wide association study (GWAS) [51], ADHD susceptibility at the genomic level might underlie persistent brain alterations. A comprehensive ADHD imaging-genetic review paper summarized that thirteen candidate genes might underlie brain alterations of ADHD, while most existing studies only focused on a single gene or a few candidate genes in a specific pathway (e.g., dopamine-related pathway) [52]. Candidate gene-based analysis works fine if the prior knowledge is solid and well supported, but it cannot provide an overall picture of genes collectively affecting brain structure/function in ADHD. In contrast, polygenic risk score (PRS)-based analyses allow evaluations of the relationship between the overall genetic risk of disease and altered brain patterns. Recently, we examined the associations between PRS of ADHD and brain alterations in frontal and cerebellum regions that related to working memory performance and inattention in adults with ADHD [53]. We did not find any linear relationships between PRS and these brain alterations, although the ADHD-PRS was associated with hyperactivity symptoms, indicating that the lumpsum risk from the whole genome may not have the specificity to link to ADHD-related brain alterations.

Multivariate data-driven imaging-genetic fusion approaches may pose a promise for delineating the genetic factors underpinning GMV alterations. These approaches study a comprehensive set of SNPs and brain voxels, aiming to strike a balance between a few candidate variables and all-in-all lumpsum scores. They extract factors in a data-driven manner and assess associations at the multivariate factor level, which reduces the

number of tests and boosts statistical power. Sparse parallel independent component analysis (spICA) [54], as one of the sparsity-regularized multivariate data fusion approaches, shows great promise for imaging genetics: spICA identified stable and replicable imaging-genetic pairs from the whole-brain sMRI and whole-genome SNP data of 35,692 adult participants in the UK biobank [54]. Here, spICA was performed on GMV and SNP array data of 341 unrelated adults aggregated from the NeuroIMAGE [55] and IMpACT-NL [7, 56] projects (Fig. 1a) and replicated on data of 461 adolescents from the NeuroIMAGE project [55]. Specifically, we investigated three gray mater networks in superior/middle/inferior frontal and cerebellum regions [28, 36], previously reported to be associated with working memory impairments (superior/middle/inferior frontal region) and inattention (cerebellum) in both adults and adolescents with ADHD, together with a set of SNPs that moderately discriminate individuals with ADHD from controls as implicated in the recent ADHD GWAS study [51].

Using ADHD adult samples for discovery exploration enables us to probe the genetic underpinning of gray matter variations in the frontal and cerebellar regions that are associated with persist ADHD symptom and cognitive impairment. Replicating the results in ADHD adolescent samples will potentially facilitate the identification of the timeline when the genetic factors exert the effects on altering gray matter, which is relevant to ADHD persistence.

MATERIALS AND METHODS
Discovery dataset
The discovery cohort was composed of 341 non-related European Caucasian adults (age: 18–63 years) aggregated from two projects: 198 samples from the NeuroIMAGE project [55] and 143 participants from the IMpACT-NL project [7, 56], including 127 healthy controls (HCs), 167 participants with ADHD and 47 unaffected siblings (unaffected siblings and ADHD cases were from different families). The IMpACT-NL project was approved by the regional ethics committee (Centrale Commissie Mensgebonden Onderzoek: CMO Regio Arnhem–Nijmegen; Protocol number III.04.0403). The NeuroIMAGE study was approved by the same regional ethics committee (2008/163; ABR: NL23894.091.08) and the medical ethical committee of the VU University Medical Center. Written informed consent was collected from all included participants.

All participants had an IQ >70, and did not have a psychotic disorder, addictions in the past 6 months, a current major depression, a diagnosis of autism spectrum disorder, epilepsy, neurological disorders, sensorimotor disabilities, or any medical or genetic disorders which might be confounded with ADHD [55–57]. The inclusion criteria for adult ADHD were slightly adapted (see details later) [55, 58] from the DSM-IV (NeuroIMAGE project) or DSM-IV-TR (IMpACT-NL). In addition, the presence of ADHD in childhood was required. Specifically, the 18 DSM-IV symptoms were assessed in all participants to evaluate their inattention and hyperactivity-impulsivity. Each symptom outcome ranged from 0 to 9, and the larger the value, the more severe the disorder. In a nutshell, adult participants with ADHD had five or more scores in the inattention and/or hyperactivity-impulsivity domain [55, 58]. Unaffected siblings were selected to have a score less than five in both the inattention and hyperactivity-impulsivity domains. HCs were screened to have a score less than 2 on total symptom counts [28]. See more details on the discovery sample enrollment and working memory performance in supplement S1. Table 1 lists the demographics and characteristics of 341 adult participants.

Replication dataset
The replication dataset included 461 European Caucasian adolescents (144 HCs, 129 unaffected siblings, and 188 ADHD) from 309 families (age: 7–17 years) recruited in the NeuroIMAGE project [55]. All participants provided written informed consent. Out of the 461 subjects, 452 were over 10 years old, and 403 were over 12 years old. Thus, we named this group adolescents throughout the whole paper. The exclusion criteria, evaluations of inattention and hyperactivity/impulsivity symptoms, assessments of working memory performance, as well as the grouping criteria for
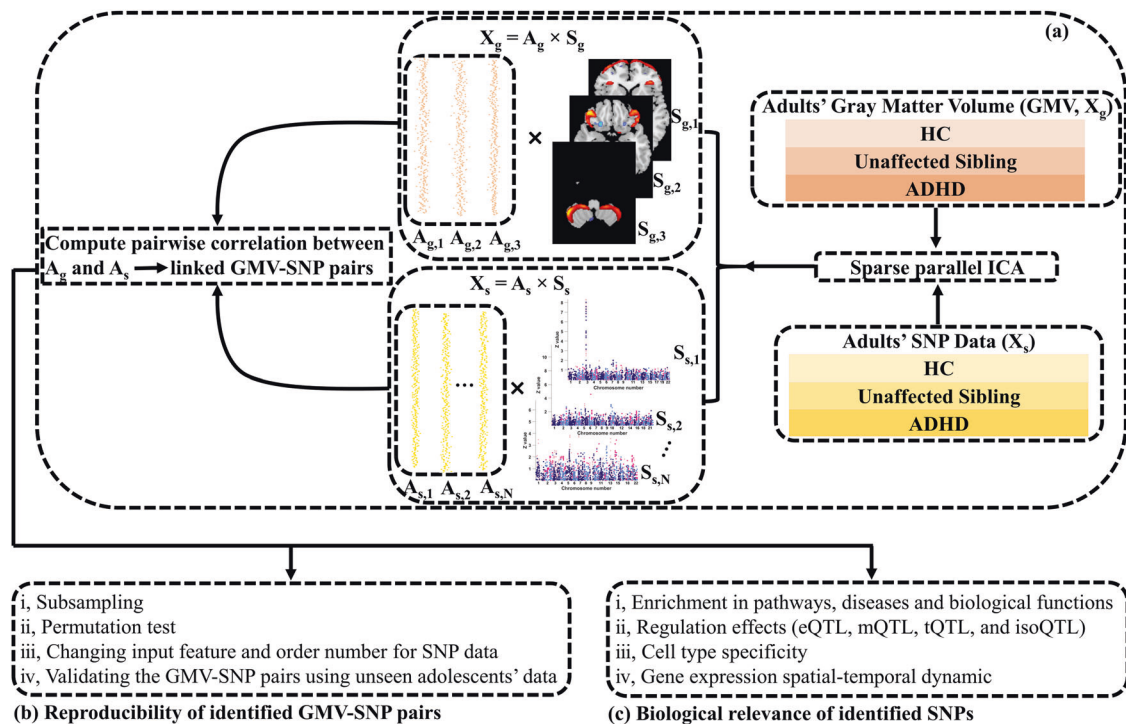
Fig. 1 Analysis pipeline of the study. a Applying spICA to GMV and SNP data of 341 unrelated adult participants to identify linked GMV-SNP pairs. **b** Examining the stability and replicability of the identified GMV-SNP pairs by various robustness analyses and external validation. **c** Performing a series of downstream analyses to investigate the biological relevance of the identified SNPs.

Table 1. Demographics and characteristics of the adult sample.

Variable	Diagnosis group (#)		
	ADHD (167)	**Unaffected siblings (47)**	**Controls (127)**
Age in years[a]	25.44 ± 8.65	20.91 ± 2.08	29.68 ± 12.35
Sex (male)[b]	95 (57%)	24 (51%)	36 (28%)
Estimated IQ[a]	103.81 ± 16.48	103.60 ± 12.71	108.61 ± 15.36
Inattention[a]	7.19 ± 1.74	1.72 ± 2.10	0.52 ± 1.18
Hyperactivity-impulsivity[a]	5.86 ± 2.37	1.51 ± 1.59	0.70 ± 1.00
Forward digit span score[a]	8.89 ± 1.98	9.00 ± 1.74	9.59 ± 1.93
Backward digit span score[a]	6.42 ± 2.31	5.96 ± 2.10	7.45 ± 2.04
History of stimulants[b]	78 (47%)	0	1 (0.8%)
Scan site (Nijmegen)[b]	115 (69%)	21 (45%)	108 (85%)
Scan site (Amsterdam)[b]	52 (31%)	26 (55%)	19 (15%)

Note, [a]Data represented as mean ± standard deviation.
[b]Data is represented as a number (percentage).

controls were the same as for adults in the aforementioned discovery cohort (see more details in supplement S2). The only difference was that adolescent participants with ADHD were required to have a score ≥6 in inattention and/or hyperactivity/impulsivity, and unaffected siblings had scores <6 in both symptom domains. Supplemental Table S1 lists the demographics and subject characteristics of 461 adolescents.

sMRI data preprocessing
T1-weighted MRI images of both the discovery and replication samples were collected with 1.5T scanners with comparable settings across projects. Extensive quality controls have been performed on all T1 images using the method described in [28, 55, 56] and metrics derived from MRIQC. The included images were segmented into six types of tissues using SPM 12 (https://www.fil.ion.ucl.ac.uk/spm/software/spm12/) with the default tissue probability map (for adults) or a customized tissue probability map (for adolescents) generated by TOM 8 [36, 59]. Subsequently, gray matter images were normalized into Montreal Neurological Institute space, followed by modulation and smoothing with a $6 \times 6 \times 6 \, \text{mm}^3$ Gaussian kernel [36]. Further quality control, gray matter refinement (see below), and voxel-wise confounding effect (i.e., age, sex, and site, if applicable) adjustment were performed separately on adults and adolescents. After quality control, only those subjects with a correlation >0.8 with the mean gray matter map were kept. Gray matter refinement (i.e., masking) selected voxels with a mean gray matter volume >0.2 for further analyses, yielding 456,921 voxels for adults and 479,770 voxels for adolescents. 441,258 common voxels between adults and adolescents were used.

Identifying the genomic factors underlying GMV variations in the superior/middle/inferior frontal and cerebellum was our ultimate goal. Thus, we reconstructed GMV data of adults and adolescents to only include variations from these three regions of interest (ROIs, Fig. S1, ICs 2–4)

[28, 36]. See more details about scanners, quality control, and the reconstruction of GMV data in supplement S3.

Genetic data preprocessing

Both the NeuroIMAGE and the IMpACT-NL project used the Illumina Psych Array to genotype DNA extracted from blood. Quality control and imputation were then performed. All samples included in this study were European Caucasians. We further controlled subgroup differences by using five genomic ancestry components in later analyses. See supplement S4 for details on genetic data preprocessing.

Guided by a recent large GWAS study on ADHD [51] and the simulation results of spICA [54, 60], we selected 2108 SNPs by focusing on SNPs showing promising ADHD vs. HC difference ($p < 1 \times 10^{-3}$) in the GWAS summary statistics [51], and applying a light linkage disequilibrium (LD) pruning ($r^2 < 0.9$ for p-value informed clumping). Light pruning does not inflate our statistical test, as correlated SNPs end up together in the same component after ICA factorization/decomposition [61].

Identification of linked GMV-SNP pairs and examination of their replicability

We first applied spICA to the reconstructed GMV and risk SNP data from 341 unrelated adults to identify linked GMV-SNP pairs and then comprehensively examined their stability and replicability.

Identification of linked GMV-SNP pairs in discovery dataset using spICA. We performed spICA to decompose 341 adults' reconstructed GMV and SNP data into three independent GMV components and 37 independent SNP components (Fig. 1a). Multivariate spICA decomposes GMV ($\mathbf{X_g}$) and SNP ($\mathbf{X_s}$) data into the product of a loading matrix ($\mathbf{A_g}$, $\mathbf{A_s}$ for GMV and SNP, respectively) and a component matrix ($\mathbf{S_g}$, $\mathbf{S_s}$ for GMV and SNP, respectively, $\mathbf{X_g} = \mathbf{A_g} \times \mathbf{S_g}$, $\mathbf{X_s} = \mathbf{A_s} \times \mathbf{S_s}$). Each row of the component matrix ($\mathbf{S_g}$ or $\mathbf{S_s}$) is one independent brain/genomic component and each brain/genomic component is statistically independent of other brain/genomic components, thus enabling component/network-based analyses. Values in one brain/genomic component reflect the contributions of individual variables (voxels/SNPs) to the brain/genomic component. Each column of the loading matrix ($\mathbf{A_g}$ or $\mathbf{A_s}$) is the loading vector, and values of the loading vector represent weights of the corresponding brain/genomic component across participants. The sparsity control in spICA regulates the number of major contributing variables for each brain and genomic component. The spICA code will be released in the Fusion ICA Toolbox (FIT, https://trendscenter.org/software/fit).

The GMV component number was set to three due to an expectation that the decomposed GMV components would resemble the three predefined ROIs (ICs 2–4 in Fig. S1). The sparsity regularizer was initialized as one. No sparsity constraint was imposed on GMV data since signals can be easily separated from backgrounds for these three ROI priors. The Hoyer constraint threshold of SNP data was set as 0.4 based on the estimation strategy described in our previous paper [54]. The SNP component number was estimated based on Chen's consistency measure [62]. Ten runs of spICA were performed, and ICASSO [63] was employed to select the most stable run to calculate GMV-SNP associations. Bonferroni correction was applied at $p < 0.05$ for comparison of 3*37 GMV-SNP associations.

Replicability of identified GMV-SNP pairs. Replicability of identified GMV-SNP pairs (Fig. 1b) was examined by (i) subsampling and permutation tests, (ii) changing input feature (applying different pruning and preselection p-value thresholds) and order (i.e., component) number for SNP data, (iii) validating identified GMV-SNP pairs using a replication dataset consisting of sMRI and SNP data of 461 adolescents recruited by the NeuroIMAGE project [55], and (iv) performing univariate genetic association analysis of the identified GMV component.

Subsampling and permutation tests: To test the stability of the identified GMV-SNP pairs, spICA was performed on 100 subsampled sets (Fig. 1b), where stratified sampling was employed to randomly select 90% subsamples from each of the ADHD, HC, and unaffected sibling groups. To test the likelihood of overfitting for the identified GMV-SNP associations, 1000 random permutations of the subjects were performed separately in the GMV and SNP data (Fig. 1b). We then applied spICA to the permuted data, and a null distribution was obtained based on the resulting pairwise

associations of loadings. Then a p-value was computed as the percentage of pairs yielding significant GMV-SNP associations (Bonferroni correction was applied at $p < 0.05$ for 1000*37*3 pairs), which reflected the probability of overfitting of the spICA model.

Varying input features and order numbers for SNP data: We further checked the stability of the identified SNP component under two scenarios (Fig. 1b): (1) the number of SNP component (i.e. order number of SNP data) varied from 5 to 60 (SNP data were fixed with $p < 1 \times 10^{-3}$, $r^2 < 0.9$ preselection); (2) SNP preselection *p*-value varied from 10^{-4} to 10^{-2} ($r^2 < 0.9$) and the number of SNP component was estimated for each resulting SNP data according to Chen's consistency measure [62]. The stability of the identified GMV-SNP associations (Fig. 1b) was also examined by performing spICA on heavily pruned SNP data ($r^2 = 0.2$, $p < 1 \times 10^{-3}$ preselection).

Validation of linked GMV-SNP pairs in the replication dataset: To investigate the identified discovery GMV-SNP pairs in the replication dataset (Fig. 1b), we projected the identified GMV and SNP components into adolescent GMV and SNP data to obtain the corresponding loadings [64] (see supplement S5 for details on the projection method), and examined their associations using a mixed-effect linear regression model (model 5 in section 2.6). A false discovery rate (FDR) at $p < 0.05$ was applied for multiple comparisons of identified GMV-SNP pairs.

Validation of linked GMV-SNP pairs in different age groups: To investigate whether participants from different age groups presented similar GMV-SNP associations within the 317 adolescents from ADHD families (i.e., 188 cases and 129 unaffected siblings) in the replication dataset, we evaluated GMV-SNP associations in five subgroups partitioned by age distribution (as listed in Table S2), separately.

Univariate genetic association of the identified GMV component: We performed univariate association analyses between individual SNP loci and loadings of the identified GMV components, controlling for five genomic ancestry components, and reported the univariate association p-values in the Manhattan plot, which was compared to the identified SNP component via visual inspection.

Statistical association analyses

For the discovery data (unrelated adults' data), the associations among GMV components, SNP components, working memory, and symptom scores were examined using the following linear regression models:

1. a GMV component loading = a SNP component loading + five genomic ancestry components.
2. working memory or symptom variable = age + sex + a GMV component loading.
3. working memory or symptom variable = age + sex + a SNP component loading + five genomic ancestry components.

    In model 1, each GMV/SNP component was tested separately. Age and sex effects were not considered for GMV components in model 1 because they have been regressed out voxel-wisely in the preprocessing step (the same for models 5 and 6 below). Working memory performance was measured by maximum forward and backward digit span count. The symptom variable included inattention and hyperactivity/impulsivity symptoms. The ADHD cases vs. HCs difference of each GMV component was evaluated using a two-sample t-test, and the ADHD cases vs. HCs difference of each SNP component was assessed using the regression model 4:

4. a SNP component loading = diagnosis (ADHD case/HC) + five genomic ancestry components.

    For the replication data, family structure and medication history (binary values) were considered in the association analyses, as medication did affect the three brain ROIs studied here in adolescents included in this study [36]. Thus, the association between a GMV component and an SNP component and the ADHD cases vs. HCs difference of a GMV/SNP component were tested using the following linear mixed-effect models:

5. a GMV component loading = a SNP component loading + five genomic ancestry components + family ID + medication status.
6. a GMV component loading = diagnosis (ADHD case/HC) + medication status + family ID.

7. a SNP component loading = diagnosis (ADHD case/HC) + family ID + five genomic ancestry components.

In models 5–7, the family structure was modeled as a random effect, and other predictors, i.e., an SNP component loading, diagnosis, medication, and five genomic ancestry components, were treated as fixed effects. Unaffected siblings were not used for ADHD cases vs. HCs comparison.

Confounding effects, including medication, history of major depressive disorder (MDD), IQ, and diagnosis group (for GMV-SNP association), were examined for both discovery and replication data by adding them as a covariate, one at a time, to the models above. A previous study [56] reported that adult ADHD patients with previous MDD showed smaller hippocampus volume compared to adult ADHD patients without previous MDD. Part of the samples in [56] was included in this study (see details in supplemental material S1 and S2). So, we controlled the history of MDD to rule out any potential confounding effects from previous MDD.

The model to test the interaction between the family label and SNP loading for predicting GMV IC 1 loading by controlling five genomic ancestry components is listed below:

8. GMV IC 1 loading = SNP loading + five genomic ancestry components + family label (ADHD/HC family) + family label*SNP loading

In model 8, the family label was coded as a binary variable.

Interpretation of identified GMV and SNP components and their biological relevance
Each of the identified GMV and SNP components was normalized to have a zero mean and unit standard deviation. GMV component regions were selected with $|z| > 2.5$ and mapped into the Talairach atlas [65].

SNPs with absolute weights larger than two were selected for further examinations of their biological relevance (Fig. 1c), including (i) gene annotation and enrichment in pathways, diseases, and biological functions by using Ingenuity Pathway Analysis (IPA, QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis) and Gene Ontology (http://geneontology.org/), (ii) regulation effects in prefrontal regions via expression quantitative trait loci (eQTL), transcript expression via transcript QTL (tQTL), isoform percentage via isoform QTL (isoQTL) based on the summary statistics available on the PsychENCODE website (http://resource.psychencode.org/) and methylation via methylation QTL (mQTL) based on Jaffe's study [66], (iii) enrichment of annotated genes in five human brains cell-type-specific gene sets based on Zhang's study [67], (iv) gene expression temporal dynamic in the dorsolateral prefrontal cortex based on BrainSpain RNA-seq data [68–70].

RESULTS
Discovery results
For the discovery data, spICA identified that two GMV components (IC 1 and IC 2 in Fig. 2a, b) showed significant and positive associations with one SNP component in Fig. 2c (IC 1-SNP_$p = 6.41 \times 10^{-13}$, IC 1-SNP_η_p^2 (partial eta squared) = 0.35; IC 2-SNP_$p = 2.46 \times 10^{-4}$, IC 2-SNP_$\eta_p^2 = 0.12$) after controlling for population structure and applying Bonferroni correction (p-value threshold was 0.05/(3*37)). Among 100 stratified subsampling tests, 89 subsamples stably identified GMV IC 1-SNP and GMV IC 2-SNP pairs with similar correlation strengths as in the full sample. 1000 permutation tests generated tail probabilities of $p = 4.10 \times 10^{-2}$ for GMV IC 2-SNP and $p = 1.20 \times 10^{-2}$ for GMV IC 1-SNP. Subsampling and permutation results supported that the identified GMV-SNP associations were stable in adults, and the corresponding correlation coefficients were likely not overfitted. Positive associations between GMV ICs 1–2 and the identified SNP component indicated that more counts of minor/reference/effect alleles in SNPs with positive weights (hot-colored loci in Fig. 2c) were related to higher GMV in brain regions with positive weights (hot-colored regions in Fig. 2a, b), and vice-versa. Figure 2d, e plot the associations between loadings of the SNP component and loadings of GMV IC 1 and IC 2, respectively. We observed that the GMV IC 1-SNP component association was much stronger in

unaffected siblings and ADHD patients compared to controls (sibling_$\eta_p^2 = 0.60$, ADHD_$\eta_p^2 = 0.43$, control_$\eta_p^2 = 0.17$). GMV IC 3-SNP component association (supplemental Table S3) was not significant in discovery samples ($p = 0.56$).

Further, we treated subjects from ADHD families (i.e., cases and unaffected siblings) as one group and participants from control families as another group. We observed that this categorical family label significantly interacted with SNP loading for GMV IC 1 prediction ($p = 8.50 \times 10^{-3}$) using model 8 in section 2.6. Diagnosis group, medication, IQ, and comorbidity did not confound the identified GMV-SNP associations.

We confirmed that the identified SNP component was robust to a range of order numbers for SNP data (SNP component number range: [24, 40]) and SNP preselection p-values (p: [0.0005, 0.005]). Applying spICA to the heavily pruned SNP data ($r^2 = 0.2$, $p < 1 \times 10^{-3}$), we identified a similar pair as reported in the main finding (the GMV IC 1-SNP pair), indicating that the discovered GMV IC 1-SNP association was unlikely biased by LD structure. See supplement S6 for details.

Replication results
In 461 adolescents, GMV IC 1-SNP component association was nominally significant after controlling for medication and five genomic ancestry components (uncorrected $p = 4.20 \times 10^{-2}$, $\eta_p^2 = 0.17$). GMV IC 2-SNP component association was not replicated. Similarly, we observed a stronger association for GMV IC 1-SNP component in unaffected siblings compared to ADHD patients and controls; the GMV IC1-SNP component associations in three groups were sibling_$\eta_p^2 = 0.73$, ADHD_$\eta_p^2 = 0.29$, control_$\eta_p^2 = 0.31$. Moreover, the categorical family label (ADHD/HC) significantly interacted with the SNP loading for GMV IC 1 prediction ($p = 4.83 \times 10^{-3}$) with controlling for medication effect. Focusing on adolescents from ADHD families, the association between GMV IC 1 and SNP component became significant (Fig. 2f, corrected $p = 5.74 \times 10^{-3}$, $\eta_p^2 = 0.37$, corrected for two GMV-SNP pairs identified in the discovery samples). The association between GMV IC 2 and the SNP component remained not significant. Moreover, GMV IC 1-SNP component association was significant in 165 older (15–17 years) adolescents (corrected $p = 4.60 \times 10^{-2}$, $\eta_p^2 = 0.42$) but not in 152 younger ones (age: 7–15 years, $p = 0.20$). See more results of sub-age adolescent groups in supplement S7. The diagnosis group, IQ, and major depression did not confound the association between GMV IC 1 and the SNP component in adolescents from ADHD families. GMV IC 3-SNP component association (supplemental Table S3) was not significant in 461 adolescents ($p = 0.80$), also not significant in 317 adolescents from ADHD families ($p = 0.17$).

Genetic association analysis of the GMV loading
None of the 2108 SNPs showed a significant association with GMV IC 1 after FDR correction. However, the p-value map of the association highlighted SNP loci in chromosomes 5, 1, and 9, which were largely consistent with the identified top SNPs. See supplement S8 for details.

Brain region identification
The identified GMV IC 1 (z-scored) is illustrated in Fig. 2a, highlighting superior and middle frontal gyri ($|z| > 2.5$). GMV IC 1 did not show significant ADHD versus control difference in adults but showed significant GMV reduction in adolescents with ADHD after controlling for the medication effect ($p = 8.72 \times 10^{-3}$, $t = 2.64$, DF = 329). Loadings of GMV IC 1 were significantly and positively associated with backward ($p = 2.09 \times 10^{-2}$, $\eta_p^2 = 0.03$) and forward ($p = 2.41 \times 10^{-2}$, $\eta_p^2 = 0.10$) digit span performance in adults and adolescents, respectively.

Highlighted SNPs identification, their annotations, and biological relevance
Figure 2c shows the absolute values of the z-scored SNP component. The identified SNP component showed no significant
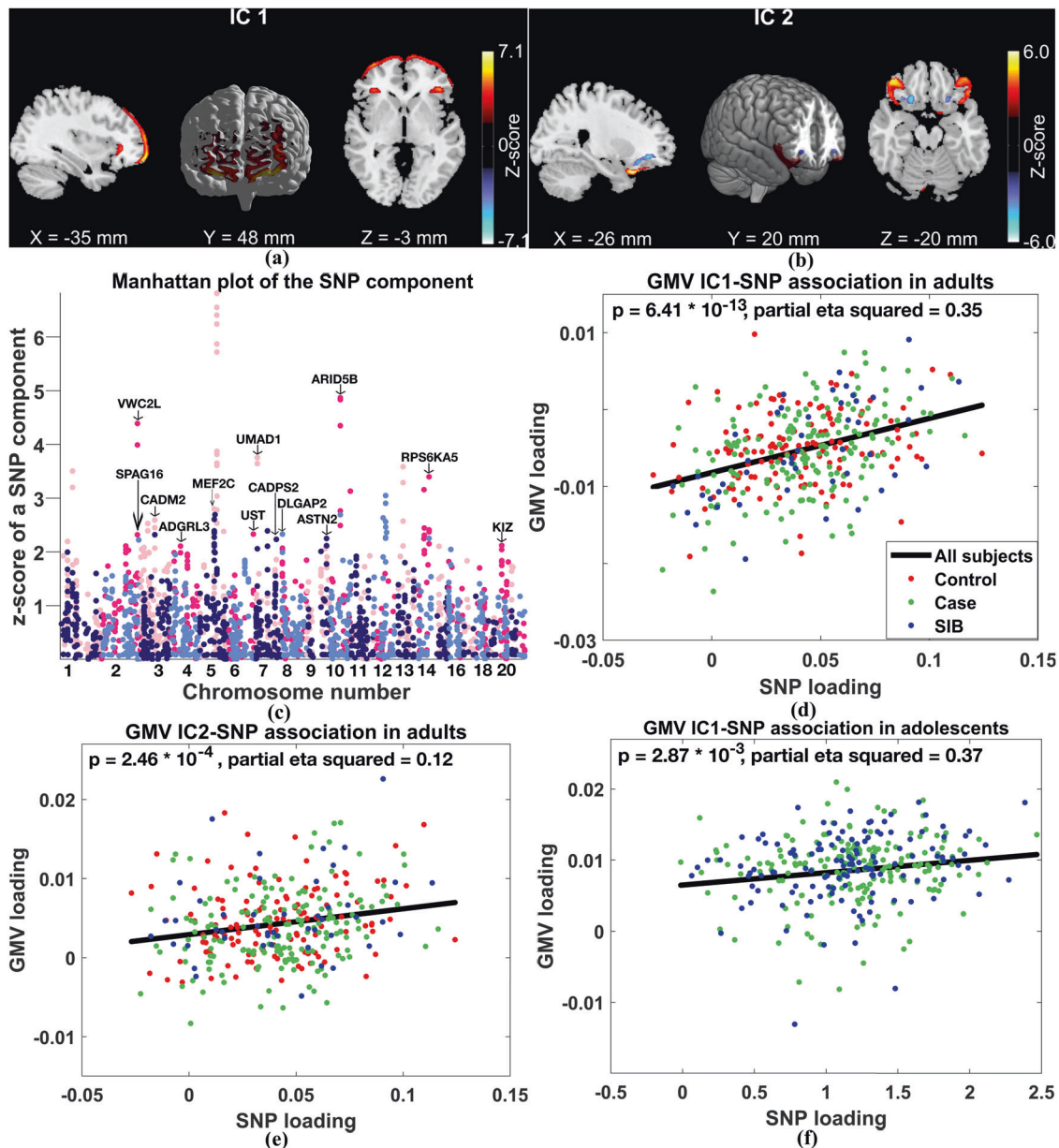
Fig. 2 Results of spICA in discovery and replication sets. The identified (**a**) GMV IC 1, (**b**) GMV IC 2, and (**c**) SNP component. The association between GMV IC 1 and SNP component was (**d**) significant in adults (the discovery set) and (**f**) replicated in adolescents (the replication set). **e** The association between GMV IC 2 and SNP component was significant in adults. In subplots **a**–**c**, hot colors represent positive weights, and cold colors denote negative weights. In subplots **d**–**f**, black, red, green, and blue colors represent all subjects, controls, cases, and unaffected siblings (SIB).

ADHD cases vs. HCs difference and was not significantly associated with forward/backward digit span performances or symptom scores in both discovery and replication sets.

The identified SNP component highlighted 93 top SNPs ($|z| > 2$, see supplemental spreadsheet). Using the human hg 19 build, 93 top SNPs were mapped to long non-coding RNAs (lncRNAs) mainly in chromosome 5, and 21 protein-coding genes, which were not significantly enriched in any particular pathways of Gene Ontology. Performing IPA on the 21 protein-coding genes, five genes, including MEF2C, DLGAP2, CADPS2, CADM2, and ADGRL3, were identified as being involved in cell-to-cell signaling and interaction (p-value range: [2.24×10^{-4}, 4.18×10^{-2}]). Six genes, i.e., MEF2C, DLGAP2, CADPS2, CADM2, ADGRL3, and UST, played a role in the nervous system development function (p-value range: [2.24×10^{-4}, 4.18×10^{-2}]). In addition, MEF2C, DLGAP2, and CADPS2 exerted effects on excitatory postsynaptic potential ($p = 2.24 \times 10^{-4}$).

Out of the 93 top SNPs, 4 acted as cis-eQTLs of three protein-coding genes and one lncRNA (Table S4), 6 were cis-isoQTLs of five unique transcripts (Table S5), and 3 were cis-tQTLs of three unique transcripts (Table S6). Moreover, 10 SNPs significantly regulated methylation levels of 9 unique CPG sites (Table S7). See supplement S9 for details of regulation effects. Furthermore, the annotated 21 genes were significantly enriched in human brain neuron cells ($p = 0.02$). Take MEF2C and CADPS2 as examples; their expression level in DLPFC from early fetal to middle adulthood is presented in supplement S10.

DISCUSSION

In this study, we applied spICA to GMV and SNP data from an ADHD adult cohort, and we identified one SNP component significantly and positively associated with two GMV components (GMV ICs 1–2 in Fig. 2a, b). The association between SNP and GMV IC 1 was further replicated in adolescents from ADHD families, but with overall weaker strength compared to adults. Meanwhile, within adolescent samples, the GMV IC 1-SNP association was stronger in older participants than in younger ones, indicating that the identified genetic component consistently exerted effects on perturbating GMV in the superior and middle frontal regions across development, but may exert their effects fully later in the development. Overall, participants from ADHD families carrying a higher load of this SNP component had larger gray matter volume in the superior and middle frontal gyri, and this relationship strengthened with age. We did not observe significant associations among control participants. Since individuals with ADHD showed no significant difference in GMV of superior and middle frontal gyri compared to controls in adults but demonstrated a significant reduction in adolescents, we speculate that a higher load of the identified SNP component likely compensates for GMV reduction in the frontal cortex in individuals with ADHD, and this compensation effect may be exerted fully in adulthood.

The identified superior/middle frontal regions are well documented in ADHD neuroimaging studies. Hoogman, et al. [33] reported that widespread cortical surfaces, including superior frontal, presented significantly reduced area in children with ADHD, and the reduction was attenuated in adolescents and adults in a large-scale ENIGMA-ADHD study. Zhao and colleagues [71] characterized GMV reduction in the left superior frontal and right middle frontal gyri in ADHD adolescents. Our results are in line with these findings showing that GMV reduction in superior and middle frontal gyri occurs in children/adolescents with ADHD but diminishes in adults with ADHD, suggesting a delayed maturation in the frontal cortex [72–75]. Better working memory performance has been related to higher GMV in the prefrontal region [76, 77] and larger surface area in superior and medial-orbital frontal gyri [45], which lends support to our results that higher GMV in superior and middle frontal gyri relates to better working memory performance in both adults and adolescents.

The top SNPs identified via spICA highlighted SNPs in lncRNAs in chromosome 5 and SNPs in 21 protein-coding genes, including MEF2C, CADPS2, and CADM2. The lncRNA transcript, RP11_6N13.1, has been associated with educational attainment [78] and broad depression [79], even though its biological function is unclear. MEF2C, highly expressed in DLPFC from early fetal to middle adulthood based on BrainSpan RNA-seq data [68–70], has a clear role in neuronal development and function [80–82]. For instance, mice were hyperactive and showed impaired motor coordination [83] after conditionally knocking out Mef2c. MEF2C mutations have been associated with ADHD [84], intellectual disability [85, 86], and other mental disorders[87–91]. Other important genes are CADM2 and CADPS2. They are both highly expressed in the frontal cortex [68–70, 92] and were associated with ADHD [93] and other mental disorders [94–99].

The findings presented here should be considered in context with their strengths and limitations. This study leveraged the recent ADHD GWAS study [51] to provide us with good candidates to investigate SNPs underlying brain alterations related to ADHD. However, samples from the NeuroIMAGE project included in this study were also utilized in the ADHD GWAS [51]. It is worth noting that the ADHD GWAS summaries were only used to select candidate SNPs as input for spICA and were not involved in the spICA analysis to identify SNP components related to brain patterns. This mitigates inflation due to overlapping samples. Another limitation is that the sample size of this study is relatively small compared to other genomic studies (e.g., GWAS), future work with a larger discovery and replication sample size is needed,

particularly for replicating GMV IC2-SNP association. Another potential limitation is that our ADHD and control groups were not matched on age and sex. Thus, there might be residual confounding effects on case vs. control differences, even after we regressed them beforehand.

This study showed the first application of our novel spICA to a clinical population. This study, together with another study in the general population (i.e., UK Biobank) [54], indicated that spICA has great potential to reveal stable and replicable genomic features underlying brain alterations related to diseases or brain development. In the future, we expect to apply spICA to other clinical populations (e.g., schizophrenia, bipolar disorder, etc.) to uncover new genomic features relevant to the neurobiology of these diseases.

In summary, spICA identified that one SNP component related to GMV in superior and middle frontal gyri underlying working memory performance in adults and adolescents with ADHD. This association was more significant in ADHD families and in older participants than in controls and younger participants. The top contributing SNPs resided in lncRNAs in chromosome 5 and a set of genes that were enriched in human brain neuron cells. Altogether, we present a new application of our novel multivariate method, which allows new ways to link genome-neuroanatomical variation to uncover new features relevant to ADHD neurobiology.

REFERENCES

1. American Psychiatric Association. Diagnostic and statistical manual of mental disorders 5th edn. Washington: AP; 2013.
2. Bitsko RH, Claussen AH, Lichstein J, Black LI, Jones SE, Danielson ML, et al. Mental health surveillance among children—United States, 2013–2019. MMWR Suppl. 2022;71:1.
3. Biederman J, Petty CR, Evans M, Small J, Faraone SV. How persistent is ADHD? A controlled 10-year follow-up study of boys with ADHD. Psychiatry Res. 2010;177:299–304.
4. Biederman J, Petty CR, O'Connor KB, Hyder LL, Faraone SV. Predictors of persistence in girls with attention deficit hyperactivity disorder: results from an 11-year controlled follow-up study. Acta Psychiatr Scand. 2012;125:147–56.
5. Agnew-Blais JC, Polanczyk GV, Danese A, Wertz J, Moffitt TE, Arseneault L. Evaluation of the persistence, remission, and emergence of attention-deficit/hyperactivity disorder in young adulthood. JAMA Psychiatry. 2016;73:713–20.
6. Alderson RM, Kasper LJ, Hudec KL, Patros CHG. Attention-Deficit/Hyperactivity Disorder (ADHD) and working memory in adults: a meta-analytic review. Neuropsychology 2013;27:287–302.
7. Mostert JC, Onnink AMH, Klein M, Dammers J, Harneit A, Schulten T, et al. Cognitive heterogeneity in adult attention deficit/hyperactivity disorder: a systematic analysis of neuropsychological measurements. Eur Neuropsychopharmacol. 2015;25:2062–74.
8. McAuley T, Crosbie J, Charach A, Schachar R. The persistence of cognitive deficits in remitted and unremitted ADHD: a case for the state-independence of response inhibition. J Child Psychol Psychiatry. 2014;55:292–300.
9. Miller M, Ho J, Hinshaw SP. Executive functions in girls with ADHD followed prospectively into young adulthood. Neuropsychology 2012;26:278–87.
10. Biederman J, Petty CR, Clarke A, Lomedico A, Faraone SV. Predictors of persistent ADHD: An 11-year follow-up study. J Psychiatr Res. 2011;45:150–5.
11. Carmona S, Vilarroya O, Bielsa A, Tremols V, Soliva JC, Rovira M, et al. Global and regional gray matter reductions in ADHD: a voxel-based morphometric study. Neurosci Lett. 2005;389:88–93.
12. Nakao T, Radua J, Rubia K, Mataix-Cols D. Gray matter volume abnormalities in ADHD: voxel-based meta-analysis exploring the effects of age and stimulant medication. Am J Psychiatry. 2011;168:1154–63.
13. Batty MJ, Liddle EB, Pitiot A, Toro R, Groom MJ, Scerif G, et al. Cortical gray matter in attention-deficit/hyperactivity disorder: a structural magnetic resonance imaging study. J Am Acad Child Adolesc Psychiatry. 2010;49:229–38.
14. Castellanos FX, Lee PP, Sharp W, Jeffries NO, Greenstein DK, Clasen LS, et al. Developmental trajectories of brain volume abnormalities in children and adolescents with attention-deficit/hyperactivity disorder. JAMA 2002;288:1740–8.
15. Greven CU, Bralten J, Mennes M, O'Dwyer L, van Hulzen KJE, 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. JAMA Psychiatry 2015;72:490–9.
16. Frodl T, Skokauskas N. Meta-analysis of structural MRI studies in children and adults with attention deficit hyperactivity disorder indicates treatment effects. Acta Psychiatr Scand. 2012;125:114–26.

K. Duan et al.

17. Ellison-Wright I, Ellison-Wright Z, Bullmore E. Structural brain change in attention deficit hyperactivity disorder identified by meta-analysis. Bmc Psychiatry. 2008;8:51.

18. Seidman LJ, Valera EM, Makris N. Structural brain imaging of attention-deficit/hyperactivity disorder. Biol Psychiatry. 2005;57:1263–72.

19. Mostofsky SH, Cooper KL, Kates WR, Denckla MB, Kaufmann WE. Smaller prefrontal and premotor volumes in boys with attention-deficit/hyperactivity disorder. Biol Psychiatry. 2002;52:785–94.

20. Durston S, Pol HEH, Schnack HG, Buitelaar JK, Steenhuis MP, Minderaa RB, et al. Magnetic resonance imaging of boys with attention-deficit/hyperactivity disorder and their unaffected siblings. J Am Acad Child Adolesc Psychiatry. 2004;43:332–40.

21. Bonath B, Tegelbeckers J, Wilke M, Flechtner HH, Krauel K. Regional gray matter volume differences between adolescents with ADHD and typically developing controls: further evidence for anterior cingulate involvement. J Atten Disord. 2018;22:627–38.

22. Jiang W, Duan K, Rootes-Murdy K, Hoekstra PJ, Hartman C, Oosterlaan J, et al. Structural brain alterations and their association with cognitive function and symptoms in attention-deficit/hyperactivity disorder families. Neuroimage Clin. 2020;27:102273.

23. Klein M, Souza-Duran FL, Menezes AKPM, Alves TM, Busatto G, Louzã MR. Gray matter volume in elderly adults with ADHD: associations of symptoms and comorbidities with brain structures. J Atten Disord. 2019;25:829–38.

24. Maier S, Perlov E, Graf E, Dieter E, Sobanski E, Rump M, et al. Discrete global but no focal gray matter volume reductions in unmedicated adult patients with attention-deficit/hyperactivity disorder. Biol Psychiat. 2016;80:905–15.

25. Seidman LJ, Valera EM, Makris N, Monteaux MC, Boriel DL, Kelkar K, et al. Dorsolateral prefrontal and anterior cingulate cortex volumetric abnormalities in adults with attention-deficit/hyperactivity disorder identified by magnetic resonance imaging. Biol Psychiatry. 2006;60:1071–80.

26. Depue BE, Burgess GC, Bidwell LC, Willcutt EG, Banich MT. Behavioral performance predicts grey matter reductions in the right inferior frontal gyrus in young adults with combined type ADHD. Psychiatry Res. 2010;182:231–7.

27. Bralten J, Greven CU, Franke B, Mennes M, Zwiers MP, Rommelse NNJ, et al. Voxel-based morphometry analysis reveals frontal brain differences in participants with ADHD and their unaffected siblings. J Psychiatr Neurosci. 2016;41:272–9.

28. Duan K, Chen J, Calhoun VD, Lin D, Jiang W, Franke B, et al. Neural correlates of cognitive function and symptoms in attention-deficit/hyperactivity disorder in adults. Neuroimage Clin. 2018;19:374–83.

29. Seidman LJ, Biederman J, Liang L, Valera EM, Monteaux MC, Brown A, et al. Gray matter alterations in adults with attention-deficit/hyperactivity disorder identified by voxel based morphometry. Biol Psychiatry. 2011;69:857–66.

30. Proal E, Reiss PT, Klein RG, Mannuzza S, Gotimer K, Ramos-Olazagasti MA, et al. Brain gray matter deficits at 33-year follow-up in adults with attention-deficit/hyperactivity disorder established in childhood. Arch Gen Psychiat. 2011;68:1122–34.

31. Amico F, Stauber J, Koutsouleris N, Frodl T. Anterior cingulate cortex gray matter abnormalities in adults with attention deficit hyperactivity disorder: A voxel-based morphometry study. Psychiat Res-Neuroim. 2011;191:31–5.

32. Hoogman M, Bralten J, Hibar DP, Mennes M, Zwiers MP, Schweren LSJ, 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;4:310–9.

33. Hoogman M, Muetzel R, Guimaraes JP, Shumskaya E, Mennes M, Zwiers MP, et al. Brain imaging of the cortex in ADHD: a coordinated analysis of large-scale clinical and population-based samples. Am J Psychiatry. 2019;176:531–42.

34. Luo YY, Halperin JM, Li XB. Anatomical substrates of symptom remission and persistence in young adults with childhood attention deficit/hyperactivity disorder. Eur Neuropsychopharm. 2020;33:117–25.

35. Castellanos FX, Lee PP, Sharp W, Jeffries NO, Greenstein DK, Clasen LS, et al. Developmental trajectories of brain volume abnormalities in children and adolescents with attention-deficit/hyperactivity disorder. JAMA-J Am Med Assoc. 2002;288:1740–8.

36. Duan K, Jiang W, Rootes-Murdy K, Schoenmacker GH, Arias-Vasquez A, Buitelaar JK, et al. Gray matter networks associated with attention and working memory deficit in ADHD across adolescence and adulthood. Transl Psychiatry. 2021;11:184.

37. Osaka M, Osaka N, Kondo H, Morishita M, Fukuyama H, Aso T, et al. The neural basis of individual differences in working memory capacity: an fMRI study. Neuroimage 2003;18:789–97.

38. Owen AM, McMillan KM, Laird AR, Bullmore E. N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. Hum Brain Mapp. 2005;25:46–59.

39. Chein JM, Moore AB, Conway AR. Domain-general mechanisms of complex working memory span. Neuroimage 2011;54:550–9.

40. Kim C, Kroger JK, Calhoun VD, Clark VP. The role of the frontopolar cortex in manipulation of integrated information in working memory. Neurosci Lett. 2015;595:25–9.

41. Chai WJ, Abd Hamid AI, Abdullah JM. Working memory from the psychological and neurosciences perspectives: a review. Front Psychol. 2018;9:401.

42. Rypma B, D'Esposito M. The roles of prefrontal brain regions in components of working memory: effects of memory load and individual differences. Proc Natl Acad Sci USA. 1999;96:6558–63.

43. Kharitonova M, Martin RE, Gabrieli JDE, Sheridan MA. Cortical gray-matter thinning is associated with age-related improvements on executive function tasks. Dev Cogn Neuros-Neth. 2013;6:61–71.

44. Rapado-Castro M, Villar-Arenzana M, Janssen J, Fraguas D, Bombin I, Castro-Fornieles J, et al. Fronto-parietal gray matter volume loss is associated with decreased working memory performance in adolescents with a first episode of psychosis. J Clin Med. 2021;10:3929.

45. Nissim NR, O'Shea AM, Bryant V, Porges EC, Cohen R, Woods AJ. Frontal structural neural correlates of working memory performance in older adults. Front Aging Neurosci. 2016;8:328.

46. Si FF, Liu L, Li HM, Sun L, Cao QJ, Lu H, et al. Cortical morphometric abnormality and its association with working memory in children with attention-deficit/hyperactivity disorder. Psychiat Invest 2021;18:679-+.

47. Pastura G, Kubo TT, Regalla MA, Mesquita CM, Coutinho G, Gasparetto EL, et al. Working memory and left medial temporal cortical thickness. Arq Neuropsiquiatr. 2016;74:785–90.

48. Faber J, Fonseca LM. How sample size influences research outcomes. Dent Press J Orthod. 2014;19:27–9.

49. Melsen WG, Bootsma MC, Rovers MM, Bonten MJ. The effects of clinical and statistical heterogeneity on the predictive values of results from meta-analyses. Clin Microbiol Infect. 2014;20:123–9.

50. Faraone SV, Larsson H. Genetics of attention deficit hyperactivity disorder. Mol Psychiatry. 2019;24:562–75.

51. Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E, et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. Nat Genet. 2019;51:63–75.

52. 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 R 2017;80:115–55.

53. Jiang W, Rootes-Murdy K, Duan K, Schoenmacker G, Hoekstra PJ, Hartman CA, et al. Discrepancies of polygenic effects on symptom dimensions between adolescents and adults with ADHD. Psychiatry Res Neuroimaging. 2021;311:111282.

54. Duan K, Chen J, Fu Z, Silva RF, Calhoun VD, Dell'Orco M, et al. Sparse parallel independent component analysis and its application to identify stable and replicable imaging-genomic association patterns in UK biobank. medRxiv. 2020. https://doi.org/10.1101/2022.06.27.22276981.

55. von Rhein D, Mennes M, van Ewijk H, Groenman AP, Zwiers MP.Oosterlaan J, et al. The NeuroIMAGE study: a prospective phenotypic, cognitive, genetic and MRI study in children with attention-deficit/hyperactivity disorder. Design and descriptives . Eur Child Adoles Psychiatry. 2015;24:265–81.

56. Onnink AMH, Zwiers MP, Hoogman M, Mostert JC, Kan CC, Buitelaar J, et al. Brain alterations in adult ADHD: effects of gender, treatment and comorbid depression. Eur Neuropsychopharm. 2014;24:397–409.

57. Arias-Vasquez A, Groffen AJ, Spijker S, Ouwens KG, Klein M, Vojinovic D, et al. A potential role for the STXBP5-AS1 gene in adult ADHD symptoms. Behav Genet. 2019;49:270–85.

58. Polanczyk G, Caspi A, Houts R, Kollins SH, Rohde LA, Moffitt TE. Implications of extending the ADHD age-of-onset criterion to age 12: results from a prospectively studied birth cohort. J Am Acad Child Adolesc Psychiatry. 2010;49:210–6.

59. Wilke M, Holland SK, Altaye M, Gaser C. Template-O-Matic: a toolbox for creating customized pediatric templates. Neuroimage 2008;41:903–13.

60. Liu JY, Bixler JN, Calhoun VD. A multimodality ICA study—integrating genomic single nucleotide polymorphisms with functional neuroimaging data. IEEE 2008;1:151.

61. Chen JY, Calhoun VD, Lin DD, Perrone-Bizzozero NI, Bustillo JR, Pearlson GD, et al. Shared genetic risk of Schizophrenia and gray matter reduction in 6p22.1. Schizophrenia Bull. 2019;45:222–32.

62. Chen J, Calhoun VD, Liu J. ICA order selection based on consistency: application to genotype data. Conf Proc IEEE Eng Med Biol Soc. 2012;2012:360–3.

63. Himberg J, Hyvarinen A. ICASSO: Software for investigating the reliability of ICA estimates by clustering and visualization. IEEE 2003;1085:259–68.

64. Chen J, Calhoun VD, Arias-Vasquez A, Zwiers MP, van Hulzen K, Fernandez G, et al. G-protein genomic association with normal variation in gray matter density. Hum Brain Mapp. 2015;36:4272–86.

65. Lancaster JL, Woldorff MG, Parsons LM, Liotti M, Freitas CS, Rainey L, et al. Automated Talairach atlas labels for functional brain mapping. Hum Brain Mapp. 2000;10:120–31.

66. Jaffe AE, Gao Y, Deep-Soboslay A, Tao R, Hyde TM, Weinberger DR, et al. Mapping DNA methylation across development, genotype and schizophrenia in the human frontal cortex. Nat Neurosci. 2016;19:40–7.

67. Zhang Y, Sloan SA, Clarke LE, Caneda C, Plaza CA, Blumenthal PD, et al. Purification and characterization of progenitor and mature human astrocytes reveals transcriptional and functional differences with mouse. Neuron 2016;89:37–53.

68. Johnson MB, Kawasawa YI, Mason CE, Krsnik Z, Coppola G, Bogdanovic D, et al. Functional and evolutionary insights into human brain development through global transcriptome analysis. Neuron 2009;62:494–509.

69. Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu XM, Li MF, et al. Spatio-temporal transcriptome of the human brain. Nature 2011;478:483–9.

70. Colantuoni C, Lipska BK, Ye TZ, Hyde TM, Tao R, Leek JT, et al. Temporal dynamics and genetic control of transcription in the human prefrontal cortex. Nature 2011;478:519–U117.

71. Zhao Y, Cui D, Lu W, Li H, Zhang H, Qiu J. Aberrant gray matter volumes and functional connectivity in adolescent patients with ADHD. J Magn Reson Imaging. 2020;51:719–26.

72. Shaw P, Malek M, Watson B, Sharp W, Evans A, Greenstein D. Development of cortical surface area and gyrification in attention-deficit/hyperactivity disorder. Biol Psychiat. 2012;72:191–7.

73. Rubia K. Neuro-anatomic evidence for the maturational delay hypothesis of ADHD. Proc Natl Acad Sci USA. 2007;104:19663–4.

74. Shaw P, Eckstrand K, Sharp W, Blumenthal J, Lerch JP, Greenstein D, et al. Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation. P Natl Acad Sci USA. 2007;104:19649–54.

75. Krain AL, Castellanos FX. Brain development and ADHD. Clin Psychol Rev. 2006;26:433–44.

76. Takeuchi H, Taki Y, Nouchi R, Yokoyama R, Kotozaki Y, Nakagawa S, et al. Global associations between regional gray matter volume and diverse complex cognitive functions: evidence from a large sample study. Sci Rep. 2017. https://doi.org/10.1038/s41598-017-10104-8.

77. Goghari VM, MacDonald AW, Sponheim SR. Relationship between prefrontal gray matter volumes and working memory performance in schizophrenia: A family study. Schizophr Res. 2014;153:113–21.

78. O'Connell KS, Shadrin A, Smeland OB, Bahrami S, Frei O, Bettella F, et al. Identification of genetic loci shared between ADHD, intelligence and educational attainment. Biol Psychiat. 2020;87:1052–62.

79. Howard DM, Adams MJ, Shirali M, Clarke T-K, Marioni RE, Davies G, et al. Genome-wide association study of depression phenotypes in UK Biobank (n = 322,580) identifies the enrichment of variants in excitatory synaptic pathways. bioRxiv. 2017. https://doi.org/10.1038/s41467-018-03819-3.

80. Li H, Radford JC, Ragusa MJ, Shea KL, McKercher SR, Zaremba JD, et al. Transcription factor MEF2C influences neural stem/progenitor cell differentiation and maturation in vivo. Proc Natl Acad Sci USA. 2008;105:9397–402.

81. Li Z, McKercher SR, Cui J, Nie ZG, Soussou W, Roberts AJ, et al. Myocyte enhancer factor 2C as a neurogenic and antiapoptotic transcription factor in murine embryonic stem cells. J Neurosci. 2008;28:6557–68.

82. Tang XL, Wang XM, Gong XM, Tong M, Park D, Xia ZG, et al. Cyclin-dependent kinase 5 mediates neurotoxin-induced degradation of the transcription factor myocyte enhancer factor 2. J Neurosci. 2005;25:4823–34.

83. Adachi M, Lin PY, Pranav H, Monteggia LM. Postnatal loss of Mef2c results in dissociation of effects on synapse number and learning and memory. Biol Psychiat. 2016;80:140–8.

84. Klein M, Singgih EL, van Rens A, Demontis D, Borglum AD, Mota NR, et al. Contribution of intellectual disability-related genes to ADHD risk and to locomotor activity in drosophila. Am J Psychiatry. 2020;177:526–36.

85. Savage JE, Jansen PR, Stringer S, Watanabe K, Bryois J, de Leeuw CA, et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. Nat Genet. 2018;50:912–9.

86. 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;49:1107–12.

87. Yingjun X, Haiming Y, Mingbang W, Liangying Z, Jiaxiu Z, Bing S, et al. Copy number variations independently induce autism spectrum disorder. Biosci Rep. 2017. https://doi.org/10.1042/BSR20160570.

88. Ripke S, Neale BM, Corvin A, Walters JTR, Farh KH, Holmans PA, et al. Biological insights from 108 schizophrenia-associated genetic loci. Nature 2014;511:421–7.

89. Lambert JC, Ibrahim-Verbaas CA, Harold D, Naj AC, Sims R, Bellenguez C, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet. 2013;45:1452–8.

90. Davies G, Lam M, Harris SE, Trampush JW, Luciano M, Hill WD, et al. Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. Nat Commun. 2018;9:2098.

91. Hill WD, Marioni RE, Maghzian O, Ritchie SJ, Hagenaars SP, McIntosh AM, et al. A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. Mol Psychiatr. 2019;24:169–81.

92. Ardlie KG, DeLuca DS, Segre AV, Sullivan TJ, Young TR, Gelfand ET, et al. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science 2015;348:648–60.

93. Albayrak O, Putter C, Volckmar AL, Cichon S, Hoffmann P, Nothen MM, et al. Common obesity risk alleles in childhood attention-deficit/hyperactivity disorder. Am J Med Genet B Neuropsychiatr Genet. 2013;162B:295–305.

94. Bonora E, Graziano C, Minopoli F, Bacchelli E, Magini P, Diquigiovanni C, et al. Maternally inherited genetic variants of CADPS2 are present in autism spectrum disorders and intellectual disability patients. EMBO Mol Med. 2014;6:795–809.

95. Velez JI, Chandrasekharappa SC, Henao E, Martinez AF, Harper U, Jones M, et al. Pooling/bootstrap-based GWAS (pbGWAS) identifies new loci modifying the age of onset in PSEN1 p.Glu280Ala Alzheimer's disease. Mol Psychiatry. 2013;18:568–75.

96. Ibrahim-Verbaas CA, Bressler J, Debette S, Schuur M, Smith AV, Bis JC, et al. GWAS for executive function and processing speed suggests involvement of the CADM2 gene. Mol Psychiatry. 2016;21:189–97.

97. Zhou H, Sealock JM, Sanchez-Roige S, Clarke TK, Levey DF, Cheng ZS, et al. Genome-wide meta-analysis of problematic alcohol use in 435,563 individuals yields insights into biology and relationships with other traits. Nat Neurosci. 2020;23:809.

98. Nagel M, Watanabe K, Stringer S, Posthuma D, van der Sluis S. Item-level analyses reveal genetic heterogeneity in neuroticism. Nat Commun. 2018. https://doi.org/10.1038/s41467-018-03242-8.

99. Linner RK, Biroli P, Kong E, Meddens FW, Wedow R, Fontana MA, et al. Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. Nat Genet. 2019;51:245.

AUTHOR CONTRIBUTIONS
JL conceptualized the study. KD performed data analysis and wrote the manuscript. JL, JC, VDC, and ES helped interpret the results and revise the manuscript. BF, JKB, MH, JO, PJH, DH, CAH, and AA-V contributed to data collection and data management. JL, JAT, AA-V, JKB and BF contributed to funding acquisition. All authors contributed to interpreting the results and revising the manuscript.

COMPETING INTERESTS
JKB has been a consultant in the past 3 years to/member of the advisory board of/ and/or speaker for Janssen Cilag BV, Eli Lilly, Lundbeck, Shire, Roche, Medice, Novartis, and Servier. He has received research support from Roche and Vifor. He is not an employee of any of these companies and not a stock shareholder of any of these companies. He has no other financial or material support, including expert testimony, patents, or royalties. BF has received educational speaking fees from Medice. The other authors declare no competing interests.

ADDITIONAL INFORMATION
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41398-023-02349-x.

Correspondence and requests for materials should be addressed to Jingyu Liu.

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

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