

Refining Attention-Deficit/Hyperactivity Disorder and Autism Spectrum Disorder Genetic Loci by Integrating Summary Data From Genome-wide Association, Gene Expression, and DNA Methylation Studies

Anke R. Hammerschlag, Enda M. Byrne, eQTLGen Consortium, BIOS Consortium, Meike Bartels, Naomi R. Wray, and Christel M. Middeldorp

ABSTRACT

BACKGROUND: Recent genome-wide association studies (GWASs) identified the first genetic loci associated with attention-deficit/hyperactivity disorder (ADHD) and autism spectrum disorder (ASD). The next step is to use these results to increase our understanding of the biological mechanisms involved. Most of the identified variants likely influence gene regulation. The aim of the current study is to shed light on the mechanisms underlying the genetic signals and prioritize genes by integrating GWAS results with gene expression and DNA methylation (DNAm) levels.

METHODS: We applied summary-data-based Mendelian randomization to integrate ADHD and ASD GWAS data with fetal brain expression and methylation quantitative trait loci, given the early onset of these disorders. We also analyzed expression and methylation quantitative trait loci datasets of adult brain and blood, as these provide increased statistical power. We subsequently used summary-data-based Mendelian randomization to investigate if the same variant influences both DNAm and gene expression levels.

RESULTS: We identified multiple gene expression and DNAm levels in fetal brain at chromosomes 1 and 17 that were associated with ADHD and ASD, respectively, through pleiotropy at shared genetic variants. The analyses in brain and blood showed additional associated gene expression and DNAm levels at the same and additional loci, likely because of increased statistical power. Several of the associated genes have not been identified in ADHD and ASD GWASs before.

CONCLUSIONS: Our findings identified the genetic variants associated with ADHD and ASD that likely act through gene regulation. This facilitates prioritization of candidate genes for functional follow-up studies.

Keywords: eQTL, Fetal brain, mQTL, Pleiotropy, Psychiatric disorders, SMR

<https://doi.org/10.1016/j.biopsych.2020.05.002>

Attention-deficit/hyperactivity disorder (ADHD) and autism spectrum disorder (ASD) are two prevalent neurodevelopmental disorders (1,2). ADHD is characterized by a persistent pattern of inattention and/or impulsiveness and hyperactivity (3). ASD is characterized by impaired social and communication skills together with repetitive and restrictive behavior (3). Heritability estimates are around 80% (4), and recent genome-wide association studies (GWASs) identified the first genetic loci associated with ADHD (5) and ASD (6). The next step is to understand the underlying biological mechanisms that drive these findings. This is frequently not straightforward, because the mapping resolution of GWAS is limited by linkage disequilibrium (LD) of the genome. Because of this and abundant long-range gene regulation (7), the causal gene is not necessarily the most proximal to the lead single nucleotide polymorphism (SNP) of an identified locus (8).

In silico analytical approaches can aid to prioritize plausible causal genes and regulatory elements at risk loci for further functional follow-up studies. It has been observed that most of the variants identified in GWASs are located in enhancers and regions of open chromatin (9,10), suggesting that they influence gene regulation, rather than affect the coding sequences of transcribed proteins (11). Genetic variants that are associated with gene expression levels are called expression quantitative trait loci (eQTLs). It has been shown that DNA methylation (DNAm) quantitative trait loci (mQTLs) colocalize with eQTLs (12,13). Hence, it is likely that mQTLs mediate the effects of genetic variants on expression levels of genes, mechanisms that might contribute to neurodevelopmental disorders including ADHD (14) and ASD (15,16). The aim of the current study is to integrate genetic variants from ADHD and ASD GWASs with eQTL and mQTL

SEE COMMENTARY ON PAGE 438

variants to shed light on the mechanisms underlying the genetic signals. Because ADHD and ASD show strong comorbidity (17), which seems to be partly explained by shared genetic factors (6,17–21), we also investigated whether there are shared biological mechanisms.

Summary-data-based Mendelian randomization (SMR) can be used to investigate if effects of genetic variants are mediated by gene expression levels (22) or DNAm levels (23) by utilizing summary-level data from independent GWASs and eQTL or mQTL studies. The SMR test is followed by a heterogeneity in dependent instruments (HEIDI) test to investigate if the association is due to a shared causal variant and is not a consequence of the widespread LD in the genome. The latter may lead to association owing to linkage, i.e., the colocalized GWAS variant and eQTL/mQTL may tag different causal variants that are in LD. This ability to distinguish a causal model from a linkage model is not featured in other methods that integrate GWAS and eQTL or mQTL datasets (22). However, a nonsignificant HEIDI test does not exclude the possibility of horizontal pleiotropy (the same variant controls both disease risk and DNAm/gene expression independently) instead of causality. We therefore use “pleiotropy” or “pleiotropic association” to describe associations identified by the HEIDI test to acknowledge that the associations cannot be interpreted with certainty as causality. The SMR approach also provides an opportunity to map DNAm sites to gene expression through shared genetic variants, in this way clarifying which genes are regulated by DNAm sites (24).

Neurodevelopmental disorders likely originate during prenatal life (25,26). Studies have observed an enrichment among high-confidence fetal brain eQTLs of genetic variants associated with ADHD and ASD (27), and colocalization of autism-associated genes with clusters of differentially methylated sites in fetal brain tissue (28), suggesting an important prenatal genetic component for both disorders. In the current study, we performed SMR analyses integrating ADHD and ASD GWAS findings with eQTL and mQTL studies of fetal brain tissue to interpret the functional consequences of common genetic variation associated with the two disorders. While previous studies mainly focused on one type of omics data (i.e., eQTLs or mQTLs), we compared results of both datatypes and subsequently used the SMR approach to identify pleiotropic relationships between DNAm and gene expression levels. We also investigated if the identified genes and DNAm sites for ADHD and ASD overlap to shed light on the shared genetic factors of the two disorders, as these are currently unknown. Because the fetal brain datasets are still relatively limited in number of samples, which reduces statistical power to identify QTLs, we took advantage of the growing number of eQTL and mQTL datasets for brain and blood and repeated our analyses using these datasets. It has been reported that eQTLs and mQTLs correlate highly between independent brain and blood samples (29), suggesting that these datasets provide additional insight by a gain in power for gene discovery. We compared our findings in fetal brain, brain, and blood to examine whether they point to the same genes.

METHODS AND MATERIALS

GWAS Datasets

We used the most recent publicly available genome-wide meta-association results for ADHD (5) and ASD (6). The ADHD study included 3 sets of data: 1) clinical ADHD diagnosis in 20,183 individuals and 35,191 control subjects that were collected by iPSYCH and Psychiatric Genomics Consortium (PGC) projects, 2) a continuous ADHD scale measured in 17,666 children of EAGLE (EARly Genetics and Lifecourse Epidemiology) and 2798 adolescents of QIMR (Queensland Institute of Medical Research), and 3) self-reported ADHD diagnosis in 5857 individuals and 70,393 control subjects part of 23andMe (Sunnyvale, CA). We applied the same approach of dataset inclusion as the ADHD GWAS of Demontis *et al.* (5) and performed our main analysis on dataset 1. Subsequently, we evaluated the results of a GWAS meta-analysis of datasets 1 and 2 and of datasets 1, 2, and 3. Details of these meta-analyses are described in Demontis *et al.* (5). Datasets 2 and 3 were not analyzed individually owing to limited statistical power. The ASD data included 18,381 ASD cases and 27,969 control subjects collected by the iPSYCH and PGC projects.

eQTL and mQTL Datasets

We used summary statistics from multiple *cis*-eQTL and *cis*-mQTL datasets measured in fetal brain, adult brain, and blood. All datasets are described in Table S1. For brain, we generated one input dataset for the SMR analyses of eQTLs and one for mQTLs, by meta-analyzing 4 brain eQTL datasets and 3 mQTL datasets. Because *cis*-eQTLs and *cis*-mQTLs are highly correlated in different brain regions and samples, ranging from 0.82 to 0.99 (29), this is the most efficient use of the available data, as it yields the highest statistical power. The meta-analyses were performed in MeCS (meta-analyze *cis*-eQTL data in correlated samples) (29), which is implemented in the SMR software. MeCS requires summary-level data in the *cis*-regions and accounts for sample overlap. Before meta-analyzing, we standardized betas and standard errors of eQTLs and mQTLs (22).

SMR Analysis to Identify Relationships Between Gene Expression and ADHD/ASD

We applied SMR (<https://cns.genomics.com/software/smr/>) (22) to identify genetic signals associated with both phenotypic and gene expression variation. This approach makes use of the concept of MR, a technique aimed at detection of causal effects. MR analysis utilizes genetic variants as an instrumental variable to test for the causative effect of an exposure on an outcome. The first step of the SMR method is an MR analysis in which the genetic variant is defined as the instrumental variable, the gene expression level as exposure, and the trait as outcome. This analysis is a two-step least-squares approach, including the effect size of the top *cis*-eQTL SNP (2-kb window) and its corresponding effect in the GWAS. We only included probes for which the top associated *cis*-eQTL had $p < 5 \times 10^{-8}$, because one of the assumptions for MR analysis is that the instrumental variable has a strong effect on

the exposure (in this case, gene expression). SMR *p* values were Bonferroni corrected for the number of genes tested.

In the second step of the SMR method, a HEIDI test is performed to investigate the presence of heterogeneity in the SMR association statistics. This reflects linkage (i.e., overlapping signals that are introduced by LD) as opposed to pleiotropy (i.e., the same variant influences both outcomes, through a causal or pleiotropic model), the latter being of more biological interest. The HEIDI test repeats the analysis with SNPs in LD with the top associated eQTL ($LD < 0.9$ and > 0.05) and an eQTL $p < 1.57 \times 10^{-3}$. Under a pleiotropic model, it is expected that the SMR association statistics are similar for SNPs in LD with the causal variant. This is the null hypothesis of the HEIDI test. Hence, associations with $p < .05$ were rejected. This is a conservative threshold for gene discovery because fewer genes are retained compared with a correction for multiple testing. LD was estimated from the Health and Retirement Study (30) data imputed to the 1000 Genomes Project (31).

In total, we performed 6 SMR analyses investigating overlapping genetic association signals for gene expression levels in fetal brain, brain, and blood, and ADHD and ASD (Figure 1A). To compare the results of the ADHD and ASD analyses, we estimated for each tissue type the Spearman correlation between the SMR betas matched by gene probe.

SMR Analysis to Identify Relationships Between DNAm and ADHD/ASD

Similar to the SMR analyses including eQTLs, we performed 6 SMR analyses investigating the overlapping genetic association signals of DNAm levels of fetal brain, brain, and blood, and of ADHD and ASD (Figure 1A). In these MR analyses, the instrumental variable is defined as the genetic variant, the DNAm level as exposure, and the trait as outcome. We followed the exact same steps as described above for the eQTL analyses, with the same corrections of multiple testing.

SMR Analysis to Identify Relationships Between DNAm and Gene Expression

The identified DNAm sites for ADHD and ASD resulting from the previous analysis step do not directly relate to the expression levels of specific genes. To gain more insight in the genes whose expression may be influenced by these associations, we applied the SMR approach to identify overlapping genetic signals associated with variation in both DNAm levels and gene expression levels (Figure 1B). In this MR analysis, the instrumental variable is defined as the genetic variant, the DNAm level as exposure, and the expression level of a gene as outcome. We applied this analysis to the blood mQTL and blood eQTL data in order to gain most statistical power. We

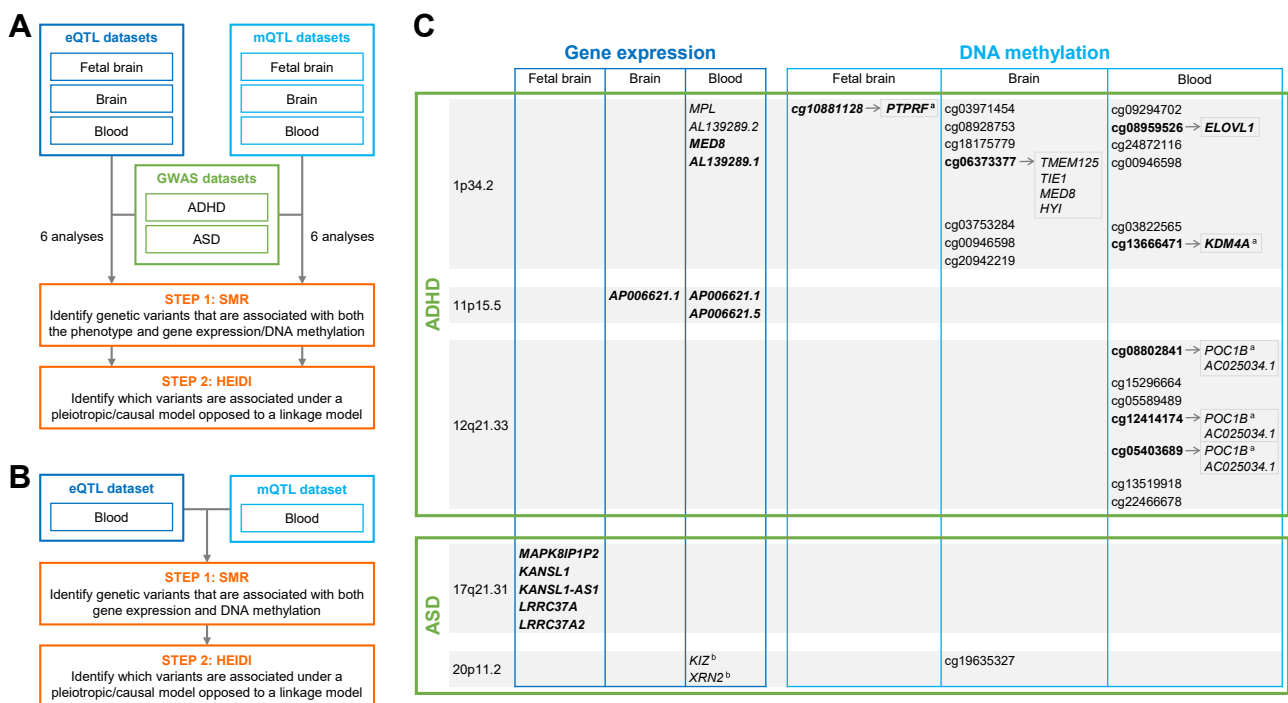


Figure 1. Flowchart of SMR analyses of eQTL, mQTL, and GWAS datasets and the analysis results. **(A)** SMR analyses performed for ADHD and ASD using eQTL and mQTL datasets in 3 tissue types. **(B)** SMR analyses performed for mQTLs that showed significant pleiotropic associations in panel **(A)** to relate them to gene expression levels. **(C)** Identified associations among gene expression, DNA methylation, and ADHD and ASD. Bold genes/DNA methylation sites passed the HEIDI test ($p \geq .05$), i.e., associations under a pleiotropic model opposed to a linkage model. The genes listed next to the DNA methylation probes are associated by the analyses performed in panel **(B)**. See Table S1 for accompanying summary statistics. ^aGene reported by the original ADHD GWAS (5). ^bGene reported by the original ASD GWAS (6). ADHD, attention-deficit/hyperactivity disorder; ASD, autism spectrum disorder; eQTL, expression quantitative trait loci; GWAS, genome-wide association study; HEIDI, heterogeneity in dependent instruments; mQTL, methylation quantitative trait loci; SMR, summary-data-based Mendelian randomization.

tested the *cis*-mQTLs located within 2-Mb distance in either direction of a gene expression probe, and only included probes with a top associated *cis*-mQTL of $p < 5 \times 10^{-8}$. SMR p values were Bonferroni corrected for the number of tested pairs of DNAm sites and genes per genomic region that we investigated. Subsequently, HEIDI tests were performed to distinguish pleiotropic associations from linkage, following the same steps as described above for the eQTL analyses.

RESULTS

Relationships Between Gene Expression/DNAm and ADHD

We applied SMR and HEIDI to test for pleiotropic associations between fetal brain gene expression levels and ADHD and between fetal brain DNAm levels and ADHD (Figure 1C). We could not detect significantly associated gene expression levels in the fetal brain data, but we did identify one fetal brain DNAm site (cg10881128) that passed the HEIDI test (i.e., pleiotropic association) at chromosome 1p34.2 (Figure 2 and Table S2).

We continued with SMR analyses of eQTLs and mQTLs measured in brain and blood (Figure 1C). The locus 1p34.2

showed associations in both the brain and blood mQTL datasets as well (Table S2 and Figure S1A, B), although different probes passed the HEIDI test (cg06373377 in brain; cg08959526 and cg13666471 in blood; cg10881128 did show strong associations with ADHD in both tissues). Furthermore, the SMR analysis of blood eQTL data showed pleiotropic associations for 2 genes, *MED8* and antisense gene *AL139289.1*, at the same locus (Table S2 and Figure S1C). Table S3 describes gene functions and associated phenotypes of these genes, and the genes identified by other SMR analyses of this study.

Besides the locus at chromosome 1, we identified multiple other loci in the brain and blood datasets (Table S2). Two related novel transcripts (*AP006621.1* and *AP006621.5*) that map to 11p15.5 were identified for ADHD in the analyses of brain and blood eQTL data and showed evidence of pleiotropy (Figure S2A, B). Furthermore, 7 DNAm sites in blood at 12q21.33 were identified, of which 3 showed evidence for pleiotropy (Figure S3).

In addition to the ADHD analyses described above, we repeated the SMR analyses using the GWAS meta-analysis data including PGC/iPSYCH and EAGLE/QIMR, and the GWAS meta-analysis data including PGC/iPSYCH, EAGLE/

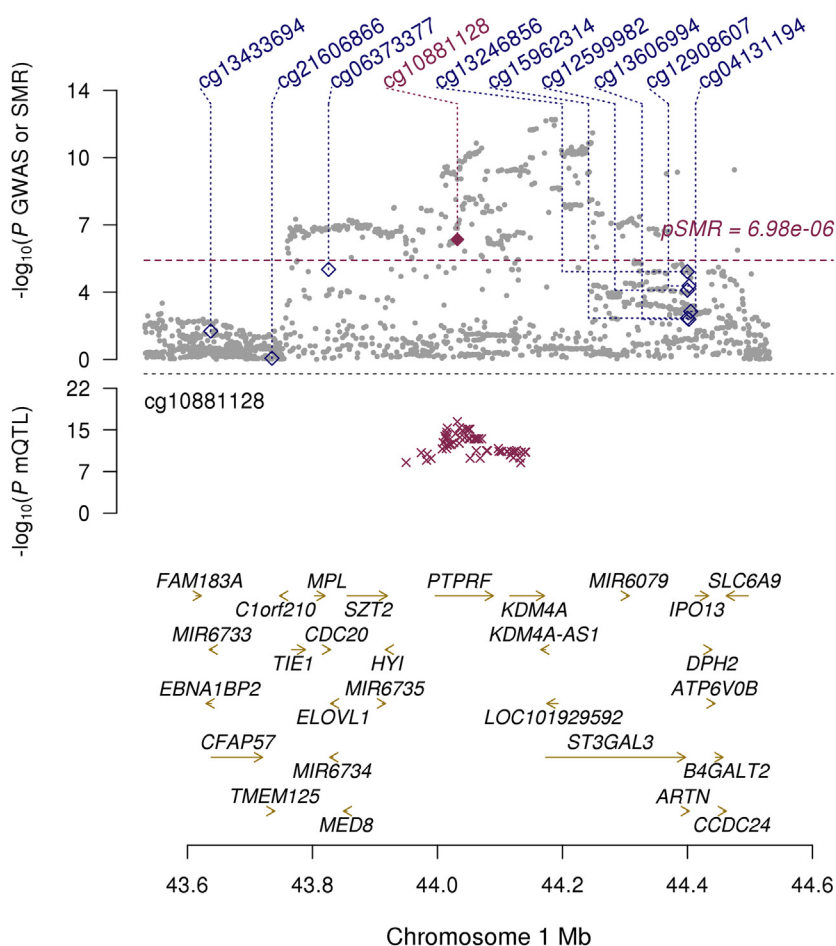


Figure 2. SMR analysis results of attention-deficit/hyperactivity disorder and fetal brain mQTLs at 1p34.2. In the lower panel, red crosses are the mQTLs for the probe displayed in the left upper corner. In the upper panel, gray dots are single nucleotide polymorphisms from the attention-deficit/hyperactivity disorder GWAS, blue diamonds are the tested DNA methylation sites in the region, and closed red diamonds are the significantly associated DNA methylation sites after Bonferroni correction (p_{SMR}) (horizontal dotted red line) that passed the heterogeneity in dependent instruments test (i.e., pleiotropic associations). GWAS, genome-wide association study; mQTL, methylation quantitative trait loci; SMR, summary-data-based Mendelian randomization.

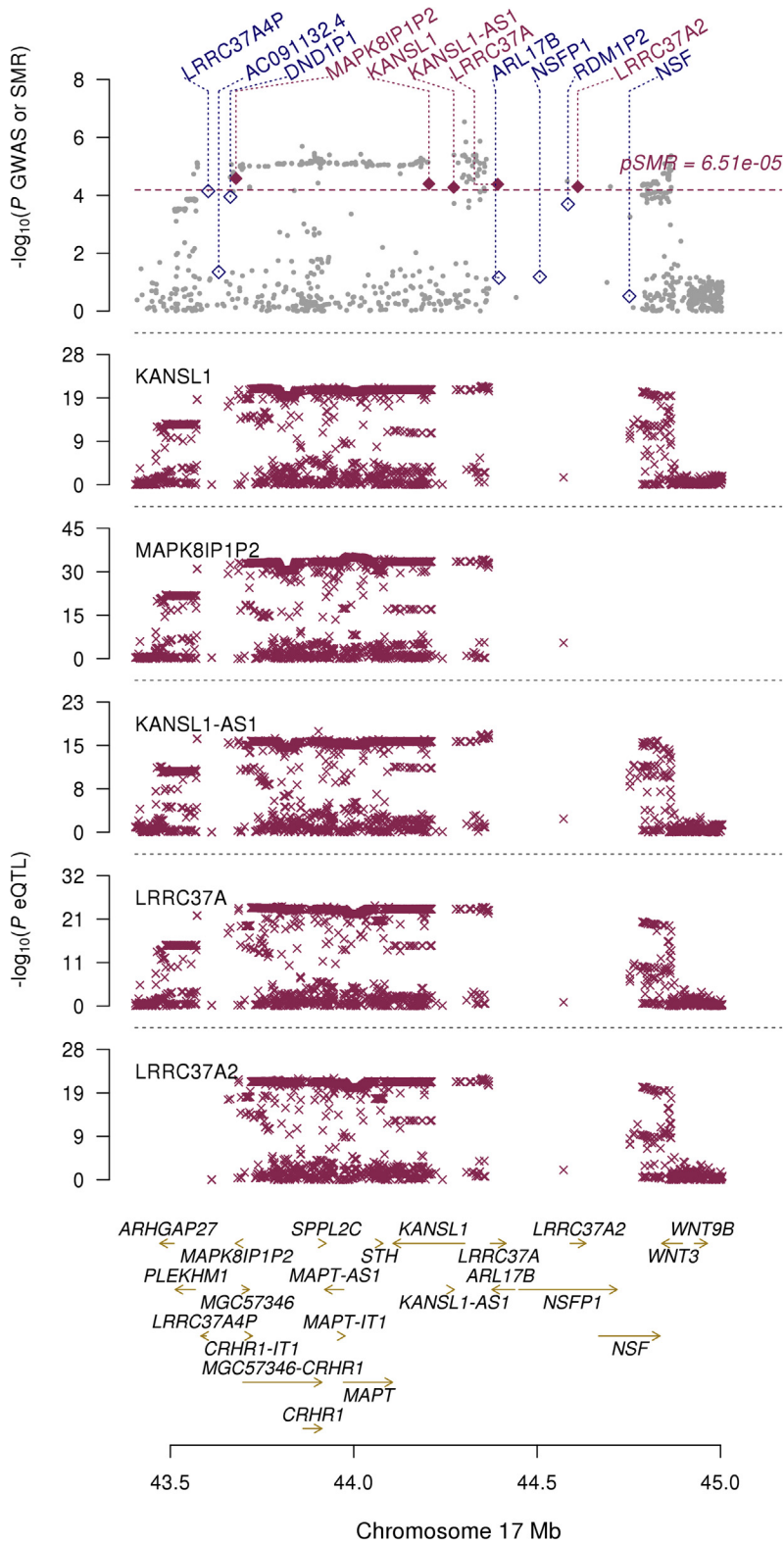


Figure 3. SMR analysis results of autism spectrum disorder and fetal brain eQTLs at 17q21.31. In the lower panel, red crosses are the eQTLs for the probe displayed in the left upper corner. In the upper panel, gray dots are single nucleotide polymorphisms from the autism spectrum disorder GWAS, blue diamonds are the tested genes in the region, and closed red diamonds are the significantly associated genes after Bonferroni correction (p_{SMR}) (horizontal dotted red line) that passed the heterogeneity in dependent instruments test (i.e., pleiotropic associations). eQTL, expression quantitative trait loci; GWAS, genome-wide association study; SMR, summary-data-based Mendelian randomization.

QIMR, and 23andMe. These results showed an increased number of associated gene expression and DNAm levels (Table S4), although fewer SMR associations passed the HEIDI test, which could be explained by increased heterogeneity in LD patterns when meta-analyzing different datasets. Hence, gene prioritization from these results is more challenging, and we focus on the initial analysis results.

Relationships Between Gene Expression/DNAm and ASD

We identified 5 significantly associated gene expression levels in fetal brain associated with ASD at 17q21.31 that all passed the HEIDI test (*KANSL1*, *KANSL1-AS1*, *LRRRC37A*, *LRRRC37A2*, and pseudogene *MAPK8IP1P2*) (Figure 1C and Table S2). Figure 3 shows that these eQTLs are located in a large region of 1 Mb that includes multiple genes.

In the subsequent brain and blood analyses, an additional locus at 20p11.2 was identified in the brain mQTL and blood eQTL data (Table S2 and Figure S4A, B). However, the identified methylation and expression probes showed evidence for linkage ($p_{\text{HEIDI}} < .05$), suggesting that other causal variants are present in this locus.

Relationships Between DNAm and Gene Expression

To gain more insight in the genes whose expression may be influenced by the DNAm sites identified in the analyses described above, we performed additional SMR analyses to link the DNAm levels to gene expression levels. This includes DNAm sites identified at 1p34.2 (Table S5) and 12q21.33 (Table S6) for ADHD. For ASD, we were not able to identify DNAm sites with pleiotropic associations. At locus 1p34.2, cg10881128 identified in fetal brain showed evidence of pleiotropic associations with *PTPRF* ($p_{\text{SMR}} = 2.3 \times 10^{-6}$, $p_{\text{HEIDI}} = .19$). Probe cg06373377, identified in brain, showed associations with 4 additional genes (*TMEM125*, *TIE1*, *MED8*, *HYI*), although all four associations showed evidence of linkage (analyzed in brain owing to absence in blood data). The two probes identified in blood, cg08959526 and cg13666471, were related to *ELOVL1* ($p_{\text{SMR}} = 9.2 \times 10^{-7}$, $p_{\text{HEIDI}} = .057$) and *KDM4A* ($p_{\text{SMR}} = 1.9 \times 10^{-6}$, $p_{\text{HEIDI}} = .070$), respectively. At locus 12q21.33, we identified three methylation probes that all showed associations for the same two genes: *POC1B* and *AC025034.1*, a novel transcript that is antisense to *ATP2B1*. However, these associations were not pleiotropic, as the HEIDI tests were significant.

Overlap Between ADHD and ASD

The associated gene expression and DNAm levels for ADHD and those for ASD did not overlap. Nevertheless, shared associated genes and DNAm sites may have remained unidentified because of limited power. To further investigate possible shared biology between the two disorders, we estimated in each tissue type the Spearman correlations between the SMR betas, which represent the effect of the association between the gene/DNAm site (when analyzing eQTLs or mQTLs, respectively) and the phenotype. Note that in the ADHD and ASD datasets, a different eQTL or mQTL can be associated with the same probe because the strongest eQTL/mQTL for each gene/DNAm site is selected in each dataset.

The identified correlations were moderate, with 0.40, 0.32, and 0.29 (based on 758, 11,173, and 15,497 matched genes, respectively) for the eQTL data of fetal brain, brain, and blood, respectively, and 0.31, 0.30, and 0.29 (based on 7163, 91,054, and 123,210 matched DNAm sites, respectively) for the mQTL data of these three tissues, respectively.

DISCUSSION

In this study, we integrated ADHD and ASD GWAS results with eQTL and mQTL variants measured in fetal brain, brain, and blood to investigate if the effects of common genetic variants related to ADHD and ASD are mediated by gene expression and DNAm levels. We identified multiple genes and DNAm sites located at three genomic loci (1p34.2, 12q21.33, and 11p15.5) that showed association through pleiotropy at a shared genetic variant with ADHD. For ASD, we identified several gene expression levels at 17q21.31 that showed pleiotropic associations. These results facilitate the prioritization of candidate genes within these loci. In addition, some of the associated loci have not been associated in GWASs of ADHD and ASD before, showing the ability of the current method to identify novel genes related to these disorders.

We were especially interested in gene expression and DNAm levels measured in fetal brain, because prenatal mechanisms likely contribute to the development of ADHD and ASD (25,26). The analyses of brain and blood revealed additional associations, supporting that the higher statistical power in these datasets can provide further insight in biological mechanisms driving the genetic associations (29,32). A second strength of our current approach is the ability to test if the associations are pleiotropic or induced by linkage, while other methods like transcriptome-wide association study (TWAS) cannot make this distinction, enabling us to further reduce the number of candidate genes for follow-up functional studies.

For ADHD, we prioritized multiple genes mapped to 1p34.2, suggesting that several genes at this locus might be involved in the development of ADHD. This locus includes the most strongly associated genetic variants in the ADHD GWAS (5) and is complicated by broad LD. Three other studies applied different approaches to integrate the ADHD GWAS with gene expression data (among them SMR, but applied to smaller datasets) and identified multiple other genes at this locus. The single gene in common by all studies was *MED8*, which we identified in blood (no strong eQTLs present in [fetal] brain datasets) and the other studies in the dorsolateral prefrontal cortex (32,33), brain (34), adrenal gland (32), and blood (34) as well. Liao *et al.* (33) showed in a TWAS of multiple brain regions that *MED8* association was specific to the dorsolateral prefrontal cortex, which might explain the absence of strong eQTLs for this gene in our meta brain data. Although eQTLs highly correlate across tissue types (29) and joint tissue approaches can improve gene prioritization for psychiatric disorders (32), the observed discrepancies might point to tissue specificity of some expression regulation mechanisms. Moreover, a recent study reported strong cell type-specific effects of eQTLs related to schizophrenia (35), which might become diluted in analyses of bulk brain tissue. This suggests that future studies of single-cell analyses might provide further

insight in regulatory mechanisms at 1p34.2. Finally, we cannot rule out that the discrepancies outlined above may be driven by data-specific effects and differences in sample size.

At a second dense gene locus, 11p15.5, we prioritized two long noncoding RNAs in blood eQTL data, of which *AP006621.5* has been related to ADHD before in analyses of brain expression data (32,33) (close to significance in our brain analysis; no strong eQTLs present in fetal brain). Besides our eQTL findings, we identified DNAm sites at 12q21.33, a locus that included associated variants in the ADHD GWAS (5), although previous studies have not prioritized genes yet. We now show that the associated variants likely influence DNAm levels at several sites that are all related to the two genes, *POC1B* and *AC025034.1* (antisense transcript for *ATP2B1*), although these associations might be induced by linkage. The genes described for the three loci above are good candidates for future functional studies to investigate their potential role in the development of ADHD.

For ASD, the 5 genes identified in fetal brain gene expression data map to 17q21.31, suggesting multiple causal signals at this locus. The locus harbors a common inversion polymorphism containing many genes (36) and has been related to a wide range of neurological disorders and traits, among them intellectual disability (37), Alzheimer's disease (38), Parkinson's disease (39), dyslexia (40), neuroticism (41), and intracranial volume (42). All 5 gene expression levels were close to significance in brain and blood, except blood expression of *KANSL1* and *LRRC37A*, despite increased statistical power of these data, suggesting possible brain-specific effects of these eQTLs. Although this locus was not identified in the ASD GWAS, a TWAS integrating ASD GWAS and brain expression levels reported 10 genes in this region (43). Furthermore, a TWAS of the PsychENCODE study (44), which measured gene expression levels in postmortem brain from individuals diagnosed with autism, revealed an association with *LRRC37A* only, although 9 additional genes were differentially expressed after applying a different strategy utilizing ASD polygenic scores. At 20p11.2, a locus including genome-wide significant genetic variants for ASD (6), we identified associations between ASD and expression levels of *XRN2*, *KIZ*, and a DNAm site. Nevertheless, other genes may be causally involved because our results suggest that these associations were introduced by linkage. Interestingly, a different gene at this locus, *NKX2-2*, has been identified by the TWAS in postmortem autism brain (44) (*NKX2-2* was only present in our brain data, $p = 4.19 \times 10^{-4}$), although their SMR analyses could not find this association. For both loci, future studies are required to investigate the relation of the prioritized genes and ASD in more detail.

Although our findings provide insight into a possible mechanism in which genetic variations exert their effects on ADHD and ASD, our results showed that comparing analyses in different tissues can be challenging. At several loci, the analyses of the three tissue types prioritized different genes. A tissue-specific nature of some eQTL and mQTL effects could underlie these discrepancies, as we described above. However, it is important to note that differences, not only within our analyses, but also with other studies, might be driven by data-specific effects and sample size of the datasets. Previous studies applying SMR to refine the genetic loci related to a

phenotype typically investigated one tissue type, and our study demonstrates that these results should be interpreted with caution, as possible discrepancies between different eQTL and mQTL datasets remain unidentified.

Even though we found associations of both gene expression and DNAm levels at the same regions, the SMR analyses linking the identified DNAm levels to gene expression levels did not reveal consistent association signals, potentially because of limited statistical power, or because of complex genetic associations within a genetic locus harboring high LD. Related to the latter, the statistical power of the HEIDI test is limited by LD structure, because local LD is used to distinguish pleiotropy from linkage. Hence, HEIDI is not always able to separate the two models when multiple causal variants are present that are in high LD. Several of the loci that include our identified genes harbor broad LD and have a high gene density. It may well be that these regions contain multiple association signals that might be incorrectly detected as heterogeneity by the HEIDI test. More research is needed to address these complex genetic loci that might harbor multiple signals that could act through different mechanisms. New opportunities lie not only within denser genotyping, fine-mapping, computational approaches such as haplotype and conditional analyses, and inclusion of more samples or ancestry groups, but also in the emergence of high-throughput assays and genome-editing experiments like CRISPR (clustered regularly interspaced short palindromic repeats)-Cas9 to evaluate the biological effects, considering different cell types and cellular context (45).

One of the aims of our study was to investigate if shared mechanisms contribute to both ADHD and ASD, as the disorders have overlapping genetic factors (17–20). We could not identify significantly associated shared genes or DNAm sites, but this may have been due to limited power. By correlating the SMR associations, we identified moderate overlap in effects between ADHD and ASD. Notably, these correlations were of equal strength as the estimated genetic correlation of 0.36 (6) between the disorders. Although these are different types of correlations, it might suggest that part of the genetic correlation may be explained by variants acting through the regulation of gene expression. Future studies are needed to investigate the underlying overlapping biological pathways in more detail.

In conclusion, we conducted an integrative analysis of GWAS and eQTL and mQTL datasets for ADHD and ASD. We identified several genetic variants that show pleiotropic associations with one of the disorders and gene expression levels or DNAm levels, indicating that associated genetic variation likely affects gene regulation. These results can facilitate the prioritization of candidate genes implicated in disease etiology and can inform functional follow-up studies that could potentially lead to therapeutic strategies.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by strategic funding from the University of Queensland and the Children's Hospital Foundation (to ARH); the European Union Horizon 2020 Research and Innovation Programme under Marie Skłodowska-Curie Grant No. 721567 (CAPICE [Childhood and Adolescence Psychopathology: unraveling the complex etiology by a large Interdisciplinary Collaboration in Europe] [to CMM]); National Health and Medical

Gene Regulation of ADHD and ASD Genetic Variants

Research Council Grant Nos. 1113400 (to NRW) and 1078901 (to NRW); European Research Council Consolidator Grant No. ERC-COG WELL-BEING 771057 (to MB); and the National Health and Medical Research Council of Australia Grant Nos. 1087889, 1145645, 1113400, 1078901, and 1078037 (to EMB).

We thank the research participants and employees of 23andMe, Inc., for contributing to this study. We thank T. Qi for providing analysis scripts for MeCS (meta-analyze cis-eQTL data in correlated samples) analyses.

The authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Child Health Research Centre (ARH, CMM), Institute for Molecular Bioscience (EMB, NRW), and Queensland Brain Institute (NRW), University of Queensland; and Child and Youth Mental Health Service (CMM), Children's Health Queensland Hospital and Health Services, Brisbane, Queensland, Australia; and the Department of Biological Psychology (ARH, MB, CMM), Vrije Universiteit Amsterdam; and Amsterdam Public Health Research Institute (ARH, MB), Amsterdam University Medical Centers, Amsterdam, the Netherlands.

eQTLGen Consortium (Author list is ordered alphabetically):

Mawussé Agbessi¹, Habibur Ahsan², Isabel Alves¹, Anand Andiappan³, Wibowo Arindarto⁴, Philip Awadalla¹, Alexis Battle^{5,6}, Frank Beutner⁷, Marc Jan Bonder^{8,9,10}, Dorret I Boomsma¹¹, Mark Christiansen¹², Anniqe Clarinbouild⁸, Patrick Deelen^{8,13}, Tõnu Esko¹⁴, Marie-Julie Favé¹, Lude Franke⁸, Timothy Frayling¹⁵, Sina A. Gharib^{16,12}, Gregory Gibson¹⁷, Bastiaan T. Heijmans⁴, Gibran Hemani¹⁸, Rick Jansen¹⁹, Mika Kähönen²⁰, Anette Kalnapenkis¹⁴, Silva Kasela¹⁴, Johannes Kettunen²¹, Yungil Kim^{6,22}, Holger Kirsten²³, Peter Kovacs²⁴, Knut Krohn²⁵, Jaanika Kronberg-Guzman¹⁴, Viktorija Kukushkina¹⁴, Zoltan Kutalik²⁶, Bernett Lee³, Terho Lehtimäki²⁷, Markus Loeffler²³, Urko M. Marigorta¹⁷, Hailang Mei⁴, Lili Milani¹⁴, Grant W. Montgomery²⁸, Martina Müller-Nurasyid^{29,30,31}, Matthias Nauck^{32,33}, Michel Nivard¹¹, Brenda Penninx¹⁹, Markus Perola³⁴, Natalia Pervjakova¹⁴, Brandon L. Pierce², Joseph Powell³⁵, Holger Prokisch^{36,37}, Bruce M. Psaty^{12,38,39}, Olli T. Raitakari⁴⁰, Samuli Ripatti⁴¹, Olaf Rottschke³, Sina Rüeger²⁶, Ashis Saha⁶, Markus Scholz²³, Katharina Schramm^{29,30,31}, Ilkka Seppälä²⁷, Eline P. Slagboom⁴, Coen D.A. Stehouwer⁴², Michael Stummvoll⁴³, Patrick Sullivan⁴⁴, Peter A.C. 't Hoen⁴⁵, Alexander Teumer⁴⁶, Joachim Thiery⁴⁷, Lin Tong², Anke Tönjes⁴³, Jenny van Dongen¹¹, Maarten van Iterson⁴, Joyce van Meurs⁴⁸, Jan H. Veldink⁴⁹, Joost Verloouw⁴⁸, Peter M. Visscher²⁸, Uwe Völker⁵⁰, Urmo Vösa^{3,14}, Harm-Jan Westra⁵, Cisca Wijmenga⁸, Hanieh Yaghootkar¹⁵, Jian Yang^{28,51}, Biao Zeng¹⁷, Futao Zhang²⁸

1. Computational Biology, Ontario Institute for Cancer Research, Toronto, Canada.
2. Department of Public Health Sciences, University of Chicago, Chicago, Illinois.
3. Singapore Immunology Network, Agency for Science, Technology and Research, Singapore, Singapore.
4. Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the Netherlands.
5. Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland.
6. Department of Computer Science, Johns Hopkins University, Baltimore, Maryland.
7. Heart Center Leipzig, Universität Leipzig, Leipzig, Germany.
8. Department of Genetics, University Medical Centre Groningen, Groningen, the Netherlands.
9. European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, United Kingdom.
10. Genome Biology Unit, European Molecular Biology Laboratory, Heidelberg, Germany.
11. Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands.
12. Cardiovascular Health Research Unit, University of Washington, Seattle, Washington.
13. Genomics Coordination Center, University Medical Centre Groningen, Groningen, the Netherlands.

14. Estonian Genome Center, Institute of Genomics, University of Tartu, Tartu, Estonia.

15. Exeter Medical School, University of Exeter, Exeter, United Kingdom.

16. Department of Medicine, University of Washington, Seattle, Washington.

17. School of Biological Sciences, Georgia Tech, Atlanta, Georgia.

18. MRC Integrative Epidemiology Unit, University of Bristol, Bristol, United Kingdom.

19. Department of Psychiatry and Amsterdam Neuroscience, Amsterdam UMC, Vrije Universiteit, Amsterdam, the Netherlands.

20. Department of Clinical Physiology, Tampere University Hospital and Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland.

21. Centre for Life Course Health Research, University of Oulu, Oulu, Finland.

22. Genetics and Genomic Science Department, Icahn School of Medicine at Mount Sinai, New York, New York.

23. Institut für Medizinische Informatik, Statistik und Epidemiologie, LIFE – Leipzig Research Center for Civilization Diseases, Universität Leipzig, Leipzig, Germany.

24. IFB Adiposity Diseases, Universität Leipzig, Leipzig, Germany.

25. Interdisciplinary Center for Clinical Research, Faculty of Medicine, Universität Leipzig, Leipzig, Germany.

26. Institute of Social and Preventive Medicine, Lausanne University Hospital, Lausanne, Switzerland.

27. Department of Clinical Chemistry, Fimlab Laboratories and Finnish Cardiovascular Research Center-Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland.

28. Institute for Molecular Bioscience, University of Queensland, Brisbane, Australia.

29. Institute of Genetic Epidemiology, Helmholtz Zentrum München – German Research Center for Environmental Health, Neuherberg, Germany.

30. IBE, Faculty of Medicine, Ludwig-Maximilians-University, München, Germany.

31. Department of Internal Medicine I (Cardiology), Hospital of the Ludwig-Maximilians-University Munich, München, Germany

32. Institute of Clinical Chemistry and Laboratory Medicine, University Medicine Greifswald, Greifswald, Germany.

33. DZHK (German Centre for Cardiovascular Research), partner site Greifswald, University Medicine Greifswald, Greifswald, Germany

34. National Institute for Health and Welfare, University of Helsinki, Helsinki, Finland.

35. Garvan Institute of Medical Research, Garvan-Weizmann Centre for Cellular Genomics, Sydney, Australia.

36. Institute of Human Genetics, Helmholtz Zentrum München, Neuherberg, Germany.

37. Institute of Human Genetics, Technical University Munich, Munich, Germany.

38. Departments of Epidemiology, Medicine, and Health Services, University of Washington, Seattle, Washington.

39. Kaiser Permanente Washington Health Research Institute, Seattle, Washington.

40. Centre for Population Health Research, Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital and University of Turku, Turku, Finland.

41. Statistical and Translational Genetics, University of Helsinki, Helsinki, Finland.

42. Department of Internal Medicine, Maastricht University Medical Centre, Maastricht, the Netherlands.

43. Department of Medicine, Universität Leipzig, Leipzig, Germany.

44. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.

45. Center for Molecular and Biomolecular Informatics, Radboud Institute for Molecular Life Sciences, Radboud University Medical Center Nijmegen, Nijmegen, the Netherlands.

46. Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany.

47. Institute for Laboratory Medicine, LIFE – Leipzig Research Center for

Civilization Diseases, Universität Leipzig, Leipzig, Germany.

48. Department of Internal Medicine, Erasmus Medical Centre, Rotterdam, the Netherlands.

49. Department of Neurology, University Medical Center Utrecht, Utrecht, the Netherlands.

50. Interfaculty Institute for Genetics and Functional Genomics, University Medicine Greifswald, Greifswald, Germany.

51. Institute for Advanced Research, Wenzhou Medical University, Wenzhou, Zhejiang, China.

BIOS Consortium (Biobank-based Integrative Omics Study):

Management Team: Bastiaan T. Heijmans (Chair)¹, Peter A.C. 't Hoen², Joyce van Meurs³, Aaron Isaacs⁴, Rick Jansen⁵, Lude Franke⁶.

Cohort Collection: Dorret I. Boomsma⁷, René Pool⁷, Jenny van Dongen⁷, Jouke J. Hottenga⁷ (Netherlands Twin Register); Marleen MJ van Greevenbroek⁸, Coen D.A. Stehouwer⁸, Carla J.H. van der Kallen⁸, Casper G. Schalkwijk⁸ (Cohort study on Diabetes and Atherosclerosis Maastricht); Gisca Wijmenga⁶, Lude Franke⁶, Sasha Zernakova⁶, Ettje F. Tigchelaar⁶ (LifeLines Deep); P. Eline Slagboom¹, Marian Beekman¹, Joris Deelen¹, Diana van Heemst⁹ (Leiden Longevity Study); Jan H. Veldink¹⁰, Leonard H. van den Berg¹⁰ (Prospective ALS Study Netherlands); Cornelia M. van Duijn⁴, Bert A. Hofman¹¹, Aaron Isaacs⁴, André G. Uitterlinden³ (Rotterdam Study).

Data Generation: Joyce van Meurs (Chair)³, P. Mila Jhamai³, Michael Verbiest³, H. Eka D. Suchiman¹, Marijn Verkerk³, Ruud van der Breggen¹, Jeroen van Rooij³, Nico Lakenberg¹. Data management and computational infrastructure Hailiang Mei (Chair)¹², Maarten van Iterson¹, Michiel van Galen², Jan Bot¹³, Daria V. Zernakova⁶, Rick Jansen⁵, Peter van 't Hof¹², Patrick Deelen⁶, Irene Nooren¹³, Peter A.C. 't Hoen², Bastiaan T. Heijmans¹, Matthijs Moed¹.

Data Analysis Group: Lude Franke (Co-Chair)⁶, Martijn Vermaat², Daria V. Zernakova⁶, René Luijk¹, Marc Jan Bonder⁶, Maarten van Iterson¹, Patrick Deelen⁶, Freerk van Dijk¹⁴, Michiel van Galen², Wibowo Arindrarto¹², Szymon M. Kielbasa¹⁵, Morris A. Swertz¹⁴, Erik W. van Zwet¹⁵, Rick Jansen⁵, Peter-Bram 't Hoen (Co-Chair)², Bastiaan T. Heijmans (Co-Chair)¹.

1. Molecular Epidemiology Section, Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, the Netherlands.

2. Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands.

3. Department of Internal Medicine, ErasmusMC, Rotterdam, the Netherlands.

4. Department of Genetic Epidemiology, ErasmusMC, Rotterdam, the Netherlands.

5. Department of Psychiatry, VU University Medical Center, Neuroscience Campus Amsterdam, Amsterdam, the Netherlands.

6. Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.

7. Department of Biological Psychology, VU University Amsterdam, Neuroscience Campus Amsterdam, Amsterdam, the Netherlands.

8. Department of Internal Medicine and School for Cardiovascular Diseases (CARIM), Maastricht University Medical Center, Maastricht, the Netherlands.

9. Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands.

10. Department of Neurology, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, the Netherlands.

11. Department of Epidemiology, ErasmusMC, Rotterdam, the Netherlands.

12. Sequence Analysis Support Core, Leiden University Medical Center, Leiden, the Netherlands.

13. SURFsara, Amsterdam, the Netherlands.

14. Genomics Coordination Center, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands.

15. Medical Statistics Section, Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, the Netherlands.

Address correspondence to Christel M. Middeldorp, M.D., Ph.D., University of Queensland, Child Health Research Centre, 62 Graham Street, South Brisbane, QLD 4101, Australia; E-mail: c.middeldorp@uq.edu.au.

Received Dec 19, 2019; revised Apr 9, 2020; accepted May 2, 2020.

Supplementary material cited in this article is available online at <https://doi.org/10.1016/j.biopsych.2020.05.002>.

REFERENCES

- Polanczyk GV, Willcutt EG, Salum GA, Kieling C, Rohde LA (2014): ADHD prevalence estimates across three decades: An updated systematic review and meta-regression analysis. *Int J Epidemiol* 43:434–442.
- Christensen DL, Baio J, Van Naarden Braun K, Bilder D, Charles J, Constantino JN, *et al.* (2016): Prevalence and characteristics of autism spectrum disorder among children aged 8 years - Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2012. *MMWR Surveill Summ* 65:1–23.
- American Psychiatric Association (2013): *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed. Arlington, VA: American Psychiatric Publishing.
- Polderman TJC, Benyamin B, de Leeuw CA, Sullivan PF, van Bochoven A, Visscher PM, Posthuma D (2015): Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat Genet* 47:702–709.
- Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E, *et al.* (2019): Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nat Genet* 51:63–75.
- Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H, *et al.* (2019): Identification of common genetic risk variants for autism spectrum disorder. *Nat Genet* 51:431–444.
- Schoenfelder S, Fraser P (2019): Long-range enhancer–promoter contacts in gene expression control. *Nat Rev Genet* 20:437–455.
- Wu Y, Zheng Z, Visscher PM, Yang J (2017): Quantifying the mapping precision of genome-wide association studies using whole-genome sequencing data. *Genome Biol* 18:86.
- Schaub MA, Boyle AP, Kundaje A, Batzoglou S, Snyder M (2012): Linking disease associations with regulatory information in the human genome. *Genome Res* 22:1748–1759.
- Ernst J, Kheradpour P, Mikkelson TS, Shores N, Ward LD, Epstein CB, *et al.* (2011): Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* 473:43–49.
- Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, *et al.* (2012): Systematic localization of common disease-associated variation in regulatory DNA. *Science* 337:1190–1195.
- Wagner JR, Busche S, Ge B, Kwan T, Pastinen T, Blanchette M (2014): The relationship between DNA methylation, genetic and expression inter-individual variation in untransformed human fibroblasts. *Genome Biol* 15:R37.
- Hannon E, Spiers H, Viana J, Pidsley R, Burrage J, Murphy TM, *et al.* (2016): Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci. *Nat Neurosci* 19:48–54.
- Xu Y, Chen X-T, Luo M, Tang Y, Zhang G, Wu D, *et al.* (2015): Multiple epigenetic factors predict the attention deficit/hyperactivity disorder among the Chinese Han children. *J Psychiatr Res* 64:40–50.
- Grayson DR, Guidotti A (2016): Merging data from genetic and epigenetic approaches to better understand autistic spectrum disorder. *Epigenomics* 8:85–104.
- Schanen NC (2006): Epigenetics of autism spectrum disorders. *Hum Mol Genet* 15:R138–R150.
- Rommelse NNJ, Franke B, Geurts HM, Hartman CA, Buitelaar JK (2010): Shared heritability of attention-deficit/hyperactivity disorder and autism spectrum disorder. *Eur Child Adolesc Psychiatry* 19:281–295.
- Miller M, Musser ED, Young GS, Olson B, Steiner RD, Nigg JT (2019): Sibling recurrence risk and cross-aggregation of attention-deficit/hyperactivity disorder and autism spectrum disorder. *JAMA Pediatr* 173:147–152.
- Ghirardi L, Brikell I, Kuja-Halkola R, Freitag CM, Franke B, Asherson P, *et al.* (2018): The familial co-aggregation of ASD and ADHD: A register-based cohort study. *Mol Psychiatry* 23:257–262.
- Musser ED, Hawkey E, Kachan-Liu SS, Lees P, Roulet J-B, Goddard K, *et al.* (2014): Shared familial transmission of autism

Gene Regulation of ADHD and ASD Genetic Variants

- spectrum and attention-deficit/hyperactivity disorders. *J Child Psychol Psychiatry* 55:819–827.
21. Martin J, Cooper M, Hamshere ML, Pocklington A, Scherer SW, Kent L, *et al.* (2014): Biological overlap of attention-deficit/hyperactivity disorder and autism spectrum disorder: Evidence from copy number variants. *J Am Acad Child Adolesc Psychiatry* 53:761–770.e26.
 22. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, *et al.* (2016): Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet* 48:481–487.
 23. Hannon E, Weedon M, Bray N, O'Donovan M, Mill J (2017): Pleiotropic effects of trait-associated genetic variation on DNA methylation: Utility for refining GWAS loci. *Am J Hum Genet* 100:954–959.
 24. Wu Y, Zeng J, Zhang F, Zhu Z, Qi T, Zheng Z, *et al.* (2018): Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. *Nat Commun* 9:918.
 25. Doherty JL, Owen MJ (2014): Genomic insights into the overlap between psychiatric disorders: Implications for research and clinical practice. *Genome Med* 6:29.
 26. Sullivan PF, Geschwind DH (2019): Defining the genetic, genomic, cellular, and diagnostic architectures of psychiatric disorders. *Cell* 177:162–183.
 27. O'Brien HE, Hannon E, Hill MJ, Toste CC, Robertson MJ, Morgan JE, *et al.* (2018): Expression quantitative trait loci in the developing human brain and their enrichment in neuropsychiatric disorders. *Genome Biol* 19:194.
 28. Spiers H, Hannon E, Schalkwyk LC, Smith R, Wong CCY, O'Donovan MC, *et al.* (2015): Methylopic trajectories across human fetal brain development. *Genome Res* 25:338–352.
 29. Qi T, Wu Y, Zeng J, Zhang F, Xue A, Jiang L, *et al.* (2018): Identifying gene targets for brain-related traits using transcriptomic and methylomic data from blood. *Nat Commun* 9:2282.
 30. Sonnega A, Faul JD, Ofstedal MB, Langa KM, Phillips JWR, Weir DR (2014): Cohort profile: The Health and Retirement Study (HRS). *Int J Epidemiol* 43:576–585.
 31. 1000 Genomes Project Consortium, Abecasis GR, Auton A, Brooks LD, DePristo MA, Durbin RM, *et al.* (2012): An integrated map of genetic variation from 1,092 human genomes. *Nature* 491:56–65.
 32. Gamazon ER, Zwinderman AH, Cox NJ, Denys D, Derks EM (2019): Multi-tissue transcriptome analyses identify genetic mechanisms underlying neuropsychiatric traits. *Nat Genet* 51:933–940.
 33. Liao C, Laporte AD, Spiegelman D, Akçimen F, Joob R, Dion PA, Rouleau GA (2019): Transcriptome-wide association study of attention deficit hyperactivity disorder identifies associated genes and phenotypes. *Nat Commun* 10:4450.
 34. Fahira A, Li Z, Liu N, Shi Y (2019): Prediction of causal genes and gene expression analysis of attention-deficit hyperactivity disorder in the different brain region, a comprehensive integrative analysis of ADHD. *Behav Brain Res* 364:183–192.
 35. Schrode N, Ho S-M, Yamamuro K, Dobbyn A, Huckins L, Matos MR, *et al.* (2019): Synergistic effects of common schizophrenia risk variants. *Nat Genet* 51:1475–1485.
 36. Zody MC, Jiang Z, Fung H-C, Antonacci F, Hillier LW, Cardone MF, *et al.* (2008): Evolutionary toggling of the MAPT 17q21.31 inversion region. *Nat Genet* 40:1076–1083.
 37. Koolen DA, Kramer JM, Neveling K, Nillesen WM, Moore-Barton HL, Elmslie FV, *et al.* (2012): Mutations in the chromatin modifier gene KANSL1 cause the 17q21.31 microdeletion syndrome. *Nat Genet* 44:639–641.
 38. Myers AJ, Kaleem M, Marlowe L, Pittman AM, Lees AJ, Fung HC, *et al.* (2005): The H1c haplotype at the MAPT locus is associated with Alzheimer's disease. *Hum Mol Genet* 14:2399–2404.
 39. Spencer CCA, Plagnol V, Strange A, Gardner M, Paisan-Ruiz C, Band G, *et al.* (2011): Dissection of the genetics of Parkinson's disease identifies an additional association 5' of SNCA and multiple associated haplotypes at 17q21. *Hum Mol Genet* 20:345–353.
 40. Veerappa AM, Saldanha M, Padakannaya P, Ramachandra NB (2014): Family based genome-wide copy number scan identifies complex rearrangements at 17q21.31 in dyslexics. *Am J Med Genet B Neuro-psychiatr Genet* 165B:572–580.
 41. Okbay A, Baselmans BML, De Neve J-E, Turley P, Nivard MG, Fontana MA, *et al.* (2016): Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nat Genet* 48:624–633.
 42. Ikram MA, Fornage M, Smith AV, Seshadri S, Schmidt R, Debette S, *et al.* (2012): Common variants at 6q22 and 17q21 are associated with intracranial volume. *Nat Genet* 44:539–544.
 43. Pain O, Pocklington AJ, Holmans PA, Bray NJ, O'Brien HE, Hall LS, *et al.* (2019): Novel insight into the etiology of autism spectrum disorder gained by integrating expression data with genome-wide association statistics. *Biol Psychiatry* 86:265–273.
 44. Gandal MJ, Zhang P, Hadjimichael E, Walker RL, Chen C, Liu S, *et al.* (2018): Transcriptome-wide isoform-level dysregulation in ASD, schizophrenia, and bipolar disorder. *Science* 362:eaat8127.
 45. Cannon ME, Mohlke KL (2018): Deciphering the emerging complexities of molecular mechanisms at GWAS loci. *Am J Hum Genet* 103:637–653.