

EXPLORING GENETIC VARIATION THAT INFLUENCES BRAIN METHYLATION IN ATTENTION-DEFICIT/HYPERACTIVITY DISORDER

Running title: Allele-specific methylation and ADHD

Laura Pineda-Cirera^{1,2,3,4,*}, Anu Shivalikanjli^{1,3,4,*}, Judit Cabana-Domínguez^{1,2,3,4}, Ditte Demontis, Ph. D.^{5,6,7}, Veera M. Rajagopal^{5,6,7}, Anders D Børghlum, Ph. D., M.D.^{5,6,7}, Stephen V. Faraone, Ph. D.⁸, Bru Cormand, Ph. D.^{1,2,3,4,8}, Noèlia Fernández-Castillo, Ph. D.^{1,2,3,4,8}

¹ Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Catalonia, Spain

² Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Spain

³ Institut de Biomedicina de la Universitat de Barcelona (IBUB), Catalonia, Spain

⁴ Institut de Recerca Sant Joan de Déu (IR-SJD), Esplugues de Llobregat, Barcelona, Catalonia Spain

⁵ The Lundbeck Foundation Initiative for Integrative Psychiatric Research, iPSYCH, Denmark

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

⁷ Department of Biomedicine - Human Genetics, Aarhus University, Aarhus, Denmark

⁸ Departments of Psychiatry and Neuroscience and Physiology, SUNY Upstate Medical University, Syracuse, NY, USA.

* Joint first authors that equally contributed

⁸ Senior authors that equally contributed

Corresponding authors:

Noèlia Fernández Castillo: Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Catalonia, Spain, Avinguda Diagonal 643, edifici Prevoiti, 1^a

planta, sala 134. 08028, Barcelona, Catalonia, Spain. Tel.: (+34) 93 4037082. Fax: (+34) 93 4034420; Email: noefernandez@ub.edu

Bru Cormand: Departament de Genètica, Microbiologia i Estadística, Facultat de Biologia, Universitat de Barcelona, Catalonia, Spain, Avinguda Diagonal 643, edifici Prevosti, 3^a planta, 08028, Barcelona, Catalonia, Spain. Tel.: (+34) 93 4021013. Fax: (+34) 93 4034420; Email: bcormand@ub.edu

ABSTRACT

Attention-deficit/hyperactivity disorder (ADHD) is a neurodevelopmental disorder caused by an interplay of genetic and environmental factors. Epigenetics is crucial to lasting changes in gene expression in the brain. Recent studies suggest a role for DNA methylation in ADHD. We explored the contribution to ADHD of allele-specific methylation (ASM), an epigenetic mechanism that involves SNPs correlating with differential levels of DNA methylation at CpG sites. We selected 3,896 tagSNPs reported to influence methylation in human brain regions and performed a case-control association study using the summary statistics from the largest GWAS meta-analysis of ADHD, comprising 20,183 cases and 35,191 controls. We identified associations with eight tagSNPs that were significant at a 5% False Discovery Rate (FDR). These SNPs correlated with methylation of CpG sites lying in the promoter regions of six genes. Since methylation may affect gene expression, we inspected these ASM SNPs together with 52 ASM SNPs in high LD with them for eQTLs in brain tissues and observed that the expression of three of those genes was affected by them. ADHD risk alleles correlated with increased expression (and decreased methylation) of *ARTN* and *PIDD1* and with a decreased expression (and increased methylation) of *C2orf82*. Furthermore, these three genes were predicted to have altered expression in ADHD, and genetic variants in *C2orf82* correlated with brain volumes. In summary, we followed a systematic approach to identify risk variants for ADHD that correlated with differential cis-methylation, identifying three novel genes contributing to the disorder.

INTRODUCTION

Attention-deficit/hyperactivity disorder (ADHD) is a common neurodevelopmental disorder with a worldwide prevalence of around 5% ¹. Its main symptoms include inattention and/or hyperactivity-impulsivity (DSM-V) ². ADHD is among the most heritable psychiatric disorders, with about 76% of its etiology accounted for by genetic risk factors ³ and with single-nucleotide polymorphisms (SNPs) explaining around 22% of the phenotypic variance ⁴. Furthermore, there is molecular evidence of shared genetic risk factors across the main five psychiatric disorders (ADHD, autism spectrum disorders (ASD), schizophrenia, bipolar disorder (BD) and major depression disorder (MDD)) ⁵. In ADHD, a recent genome-wide association study meta-analysis of 12 sample groups unraveled some of the specific genetic underpinnings of this polygenic disorder for the first time ⁴. One of the challenges of genome-wide association studies (GWAS) is to establish the causal relationship between the associated genetic variants, especially those located outside genes, and the disorder. In this regard, the use of epigenetic information can improve the interpretation of functionality of non-coding genetic variation ^{5,6}. In addition, some studies have hypothesized the importance of sub-threshold variants derived from GWAS ⁷, particularly those located in enhancer regions, with a potential impact on gene regulation ^{8,9}.

DNA methylation is one of the most stable epigenetic mechanisms, involving methylation mainly at cytosines of CpG dinucleotides. This mechanism plays an important role in the regulation of neurogenesis, differentiation and brain development ¹⁰. Furthermore, epigenetic alterations have been hypothesized to contribute to neurodevelopmental disorders ¹¹, including ADHD ¹², ASD ^{13,14}, or borderline personality disorder ¹⁵.

DNA methylation can be complementary if it involves both alleles or non-complementary when it affects only one allele, as in chromosome X inactivation in females or allele-specific methylation (ASM) ⁶. ASM is a common mechanism by which

single nucleotide variants determine differential methylation levels of CpG sites. ASM can alter promoter activity, leading to allele-specific expression ¹⁶ in combination with other still quite unknown factors, such as environmental effects ⁶. It is quantitative and heterogeneous across tissues and individuals ⁶. The environment affects DNA methylation leading to changes in gene regulation, although the underlying mechanism is still not well understood ¹⁷. It has been suggested that, during embryonic development, ASM regions could be especially sensitive to environmental effects ⁶. Investigating SNPs that display allele-specific methylation could help identify risk variants for common diseases, including neuropsychiatric disorders ¹⁸, as shown by recent studies of BD and schizophrenia ^{9,19}.

The present study investigated the possible contribution of ASM to ADHD using data from the largest GWAS meta-analysis performed to date in ADHD ⁴. We assessed its possible effect on gene expression and on brain volumes to identify new genes contributing to the disorder.

MATERIALS AND METHODS

Selection of Allele-Specific Methylation SNPs

The SNP selection was made based on the results of two previous studies^{20,21}, which identified ASM variants in multiple brain regions of *post-mortem* human samples. Gibbs *et al.*, 2010 considered four brain regions (cerebellum, frontal cortex, caudal pons and temporal cortex) of 150 subjects and Zhang *et al.*, 2010 used only the cerebellum of 153 subjects.

In the study by Zhang *et al.*, 2010 a total of 12,117 SNP-CpG pairs associations were reported in cerebellum, and Gibbs *et al.*, 2010 listed a total of 12,135 SNP-CpG pairs in frontal cortex, 11,374 in caudal pons, 16,734 in temporal cortex and 12,102 in cerebellum (Fig. 1). We combined the information from both studies and obtained a total of 43,132 SNP-CpG pairs involving 33,944 different SNPs and 5,306 CpG sites (Fig. 1). We considered all the ASM SNPs from all the tissues in the two studies, as there are multiple SNP-CpG pairs in common between them (Fig. S1).

We subsequently applied different filters to generate a sub-list of 3,896 SNPs out of these 33,944 variants to minimize redundancy: associations in *cis* between the SNP and the CpG site, correlation of the SNP with methylation levels of the CpG ($R^2 \geq 0.2$), as performed in Gibbs *et al.* 2010²⁰. We considered only autosomal SNPs and selected tagSNPs for each CpG site ($r^2 \geq 0.85$), by assessing linkage disequilibrium (LD) with Haploview software²² using the Central European (CEU) reference panel from 1000 Genomes Project Phase 3²³ (Fig. 1).

Case-control GWAS datasets

We explored the selected ASM SNPs in the summary statistics from a meta-analysis of 11 independent GWAS of ADHD conducted by the Psychiatric Genomics Consortium (PGC) and iPSYCH. This case-control study investigated 8,047,420 markers in 20,183

cases and 35,191 controls from Europe, USA, Canada, and China, with patients diagnosed according to the criteria detailed in Demontis *et al.*, 2017 ⁴.

Statistical analysis

From our initial selection of 3,896 ASM SNPs, information of 3,771 SNPs could be retrieved (96.8%) from the summary statistics of the ADHD GWAS meta-analysis ⁴. False Discovery Rate (FDR) was applied to correct for multiple testing. We used the q-value package for R ²⁴ and obtained a threshold p-value of 6.78e-05 corresponding to a 5% FDR. CpG sites highlighted by SNPs that were significant at this FDR threshold were followed-up in further analyses (Fig. 1). We also performed Bonferroni correction for multiple testing ($p \leq 1.32e-05$; $0.05/3771$ SNPs) although it was too stringent in our study as the considered SNPs are not completely independent. Finally, we considered and retrieved p-values of those SNPs in high LD ($r^2 \geq 0.85$) with the previous ones that correlate in *cis* with the methylation levels of the same CpG sites ($R^2 \geq 0.2$).

Functional annotation of associated ASM SNPs

We applied four methods to obtain information about the possible functional impact of the ASM SNPs that were associated with ADHD. First, we evaluated the presence of possible enhancer or promoter regions using the Haploreg v4.1 tool ²⁵. To do this, we considered histone modifications related to enhancer regions (H3K4me1 and H3K27ac) and promoters (H3K4me3 and H3K9ac) of 10 different brain regions (hippocampus middle, substantia nigra, anterior caudate, cingulate gyrus, inferior temporal lobe, angular gyrus, dorsolateral prefrontal cortex, germinal matrix and male and female fetal brain). Second, we evaluated the effect on gene expression through an eQTL analysis using GTEx data (Release V7) ²⁶. We considered eQTL information for all available brain tissues: amygdala, anterior cingulate cortex (BA24), caudate basal ganglia, cerebellar hemisphere, cerebellum, cortex, frontal cortex (BA9), hippocampus, hypothalamus, nucleus accumbens basal ganglia, putamen basal

ganglia, spinal cord cervical c-1 and substantia nigra. Third, we considered all the SNPs, not only ASMs, located within ± 1 Mb from the transcription start site (TSS) of each gene to infer if the genetically determined expressions of genes of interest correlated with ADHD. This analysis was carried out using MetaXcan²⁷, the input being the summary statistics of the ADHD GWAS meta-analysis⁴ and prediction models trained with RNA-Seq data of 10 GTEx²⁶ brain tissues and CommonMind²⁸ dorsolateral prefrontal cortex. The SNP covariance matrices were generated using the 1000 Genomes Project Phase 3²³ EUR genotypes of the prediction model SNPs. Bonferroni correction for multiple testing was considered ($p \leq 2.27\text{e-}03$; $0.05/22$ SNPs). Finally, we examined the possible influence of our identified variants on subcortical brain structures. We obtained the summary statistics of a GWAS meta-analysis of eight MRI volumetric measures (nucleus accumbens, amygdala, caudate nucleus, hippocampus, pallidum, putamen, thalamus and intracranial volume (ICV)) produced by the Enhancing Neuro Imaging Genetics through Meta-Analysis (ENIGMA) consortium²⁹. This ENIGMA2 discovery sample included 13,171 subjects of European ancestry and contained association results between seven million markers and variance in the volumes of the mentioned structures²⁹; we applied the Bonferroni correction ($p \leq 1\text{e-}03$; $0.05/50$ SNPs)

RESULTS

We investigated the possible association with ADHD of SNPs that show allele-specific methylation in brain regions. Starting from two previous studies^{20,21} that describe ASM in brain tissues we obtained 43,132 SNP-CpG pairs involving 33,944 SNPs and 5,306 CpG sites (Fig. 1). We observed some overlaps and redundancies between studies and tissues (Fig. S1), so we performed a selection process ending up with a list of 3,896 ASM tagSNPs (Fig. 1). Eight ASM tagSNPs were significantly associated with ADHD after correcting for multiple comparisons (5% FDR, $p \leq 6.78e-05$). These eight SNPs correlated with differential methylation at six CpG sites in cis (three for cg20225915, two for both cg22930187 and cg06207804, and one for each of cg13047596, cg11554507 and cg04464446) in different brain areas (Figs. 1 to 4). Although the explored SNPs are not completely independent in terms of linkage disequilibrium (LD), we also applied the Bonferroni correction ($p \leq 1.32e-05$; $0.05/3771$ SNPs) and three of the eight ASM tagSNPs remained statistically significant, all correlating with differential methylation at the cg20225915 site (Table 1). As considering only tagSNPs may overlook true causal variants, we retrieved association results from all the 52 ASM SNPs tagged by the previous ones (LD; $r^2 \geq 0.85$), ending up with 60 variants in eight LD blocks that show association with ADHD and correlate with methylation levels at six CpG sites (Table S1). Regional association plots of the eight tagSNPs can be found in Figs. S2 to S9 and LD patterns among the 60 variants are described in Figs. S10 to S14.

The ADHD risk alleles of the 60 SNPs and their effects on CpG methylation levels are reported in Figs. 2 to 4 and Table S1. Consistently, the direction of the effect of the risk alleles on methylation levels is the same for all the SNPs in each LD block, correlating with the same CpG site. Thus, the risk alleles correlate with decreased methylation of cg22930187, cg06207804, cg11554507 and cg20225915 and with increased methylation of cg13047596 and cg04464446^{20,21} (Figs. 2, 3, 4 and Table S1).

All six identified CpG sites are located in possible promoter regions (less than 5,000 bp upstream from a TSS) of six genes: *ARTN* (cg22930187 and cg06207804), *C2orf82* (cg13047596), *NEUROD6* (cg11554507), *PIDD1* (cg20225915), *RPLP2* (cg20225915) and *GAL* (cg04464446), all of them expressed in brain. Furthermore, most of the ASM SNPs have histone marks in different brain areas related to enhancer or promoter regions (Figs. 2 to 4 and Tables S2 to S6).

We subsequently assessed the possible effect of those 60 SNPs on gene expression and observed that all the SNPs from six LD blocks are eQTLs for three of the six genes, where the differentially methylated CpG sites lie in the possible promoter regions: *ARTN*, *C2orf82* and *PIDD1* (Figs. 2 to 4 and Table S1). The ADHD risk alleles are associated with increased expression of *ARTN* (in caudate basal ganglia, cerebellar hemisphere and cerebellum) and *PIDD1* (in cerebellar hemisphere, cerebellum and cortex) and with decreased expression of *C2orf82* (in amygdala, anterior cingulate cortex, caudate basal ganglia, cerebellar hemisphere, cerebellum, cortex, frontal cortex, hippocampus, hypothalamus, nucleus accumbens basal ganglia, putamen basal ganglia and spinal cord cervical c-1) (Figs. 2 to 4).

Consistently, all risk variants for ADHD showed opposite directions for methylation and gene expression levels, as expected for CpG sites lying in promoter regions (Table S1). Also, all the ADHD risk alleles from different LD blocks within each gene contribute in the same direction to its methylation and to its expression in all the investigated tissues (Figs. 2 to 4 and Table S1).

Furthermore, we assessed whether the effect on gene expression predicted for these genes correlate with ADHD by considering all variants within ± 1 MB from the TSS (and not only the ASM SNPs) using MetaXcan. We found significant associations of gene expression with ADHD for the same three genes in multiple brain tissues: *ARTN*, *PIDD1* showed increased expression ($3.57 < Z\text{-score} < 4.19$ and $3.57 < Z\text{-score} < 5.37$ respectively) and *C2orf82* with a decreased expression ($-3.64 < Z\text{-score} < -3.07$) (Table

S7), all of them surviving the Bonferroni correction for multiple testing ($p \leq 2.27e-03$). Consistently, for these genes, the predicted direction of the effects for the gene expression changes according to MetaXcan are the same as the ones observed in the eQTL analysis.

We also evaluated the correlation of the 60 ADHD-associated SNPs with subcortical brain volume changes in ENIGMA2 data, where we could retrieve information for 50 SNPs. We found that most the SNPs from three LD blocks are nominally significant in at least one of these three subcortical brain regions: accumbens, caudate and thalamus (Table S8). Most of the SNPs correlating with cg13047596 (near *C2orf82*) and with cg04464446 (near *GAL*), correlate with both accumbens and caudate volumes, while the only SNP correlating with cg11554507 (near *NEUROD6*) present in ENIGMA2 correlates with thalamus volume (Table S8).

DISCUSSION

This study is the first comprehensive assessment of the contribution to ADHD of genetic variants altering methylation in brain. We identified a total of 60 variants from eight LD blocks associated with ADHD that correlated with differential levels of methylation at six different CpG sites ^{20,21} (Tables 1 and S1). All the variants from six out of the eight LD blocks alter the methylation of CpG sites lying in the possible promoter regions and are also eQTLs for one of the following three genes in brain: *ARTN*, *C2orf82* and *PIDD1* (Figs. 2 to 4 and Table S1). It is well known that DNA methylation in promoter regions inversely correlates with levels of gene expression ¹⁷, and all the ASM variants that we found associated with ADHD in our study are concordant with this statement.

The *ARTN* gene, highlighted by two tagSNPs, encodes Artemin, a ligand of the *GDNF* family (glial cell line-derived neurotrophic factor). Artemin supports the survival of sensory and sympathetic peripheral neurons in culture by interacting with GFR α 3-RET and possibly also of dopaminergic neurons of the ventral mid-brain through activation of GFR α 1-RET complex ³⁰. Gene Ontology (GO) pathways link it to key neurodevelopmental functions: axon guidance (GO:0007411), neuroblast proliferation (GO:0007405) and peripheral nervous system development (GO:0007422). Risk alleles for ADHD lead to an overexpression of *ARTN*. Previously, overexpression of *ARTN* has been studied in transgenic mice and been linked to an increase of neuron excitability that leads to hypersensitivity ^{31,32}. Another study in *ARTN* knockout mice reported aberrations in the sympathetic nervous system related to migration and axonal projection ³³. The *C2orf82* gene (also known as *SNORC*) was highlighted by one tagSNP and it encodes a proteoglycan transmembrane protein that is expressed in brain more than in other tissues ²⁶. Little is known about its function. Finally, *PIDD1* was highlighted by three tagSNPs. It is a cell life regulator gene and it has been linked to apoptotic and anti-apoptotic pathways. The PIDD protein initiates apoptosis as a component of the PIDDosome together with RAIDD (RIP-associated ICH-1/ECD3-

homologous protein with a death domain) and procaspase-2³⁴ and it also activates an anti-apoptotic pathway involving the transcription factor NF-κB in response to genotoxic stress³⁵

The ASM SNPs from six LD blocks associated with ADHD are eQTLs for one of the following genes: *ARTN*, *C2orf82* and *PIDD1* in several brain regions (Figs. 2 to 4). Also, when we considered all the variants in these genes (± 1 MB from the TSS), *ARTN*, *C2orf82* and *PIDD1* were also predicted to be differentially expressed in the same direction in ADHD by MetaXcan (Table S7). The cerebellum is an expression site common to all these eQTLs, which can be explained in part by the fact that the initial SNP selection is enriched in ASM SNPs of this brain region (Fig. S1). Structural anomalies in cerebellum have been reported in ADHD individuals through neuroimaging studies^{36–38}. ADHD is suggested to be related to problems in processing temporal information, and cerebellum, basal ganglia and prefrontal cortex are involved in the temporal information processing³⁹. Cerebellar developmental trajectories and hippocampal volumes are linked to the severity of ADHD symptoms^{40–42}. Structural and functional abnormalities in cerebellum and basal ganglia have further been associated with motor impairments⁴³, which are frequent in nearly half of ADHD cases⁴⁴. Other areas identified through our eQTL analysis have also been related to ADHD, for instance: i) Remarkable differences in the shapes of caudate-putamen basal ganglia and smaller volumes have been reported in ADHD boys in comparison to typically developing individuals^{45–48}; ii) in adult males with ADHD, right caudate volume correlates with poor accuracy on sensory selection tasks⁴⁹ and also with hyperactivity/impulsivity^{50,51}; iii) nucleus accumbens is one of the areas, together with the caudate nucleus, putamen, amygdala, and hippocampus, reported to be structurally altered in the brain of ADHD patients⁵². A delayed cortical development e.g. in prefrontal regions has been linked to ADHD^{53,54} with below median intelligence quotient (IQ)⁵⁵, with the exception of primary motor cortex which shows slightly earlier

maturation⁵⁴; and iv) there is speculation about the role of cortical thickness⁵⁶, cortical volume⁵⁷ and functional connectivity in the anterior cingulate cortex⁵⁸, the region for cognitive control, attention, affect and drive⁵⁹ in ADHD⁵⁹⁻⁶². These fronto-subcortical structures (lateral prefrontal cortex, dorsal anterior cingulate cortex, caudate, and putamen) and pathways are rich in catecholamines which are involved in the pharmacological treatments for ADHD^{44,48,53,63}. Remarkably, most of the ASM SNPs in the LD block for *C2orf82*, also nominally correlate with increased volumes of accumbens and caudate subcortical regions (Table S8). It is interesting that these SNPs have the higher effect sizes as eQTLs in caudate basal ganglia (Fig. 3), one of the subcortical regions which volume correlates with genotype variation at these SNPs.

Interestingly, the methylation of cg20225915 has also been associated with *PIDD1* expression in peripheral blood⁶⁴, turning it into a good candidate as a biomarker. Also, two of the three highlighted genes had previously been related to other psychiatric diseases. The expression of *ARTN* and *C2orf82* was found to be altered in blood of MDD⁶⁵ and schizophrenia patients^{66,67}, respectively. The fact that these genes have previously been related to neuropsychiatric disorders that are often comorbid with ADHD⁶⁸ make them appealing candidates to be pursued.

ARTN is the only gene highlighted in our study that is present in one of the top regions reported in the ADHD GWAS meta-analysis⁴, although it did not contain SNPs surviving genome-wide significance. The GWAS findings in the region could be accounted for by one of several genes: *ST3GAL3*, *PTPRF*, *KDM4A*, *RP11-184I16.4*, *XR_246316.1*, *KDM4A-AS1* and *SLC6A9*. *ST3GAL3* had the most signals. Although two of our ASM variants associated with ADHD are intronic to *ST3GAL3*, this gene was not highlighted in our study as none of the associated variants correlated with differential methylation of CpG sites near the *ST3GAL3* transcription start site (distance from the nearest CpG site: 197 Kb) or were eQTLs for the gene in brain tissues. Instead, these SNPs correlated with a nearby gene, *ARTN*, both in terms of

methylation and gene expression. This suggests the importance of finding functional connections between disease-associated SNPs and genes, besides considering the genes in the physical vicinity of variants. Furthermore, another of our highlighted genes, *PIDD1*, although not being among the top findings in the ADHD GWAS meta-analysis⁴, it is pointed out by the gene-based association analysis performed in the same study.

Genetic variants surpassing genome-wide significance in GWAS explain only a small part of the SNP-based heritability and associations not reaching the significance threshold also contribute to disease susceptibility^{4,8}. An omnigenic model has been recently proposed suggesting that these subthreshold variants could point at regulatory elements of core genes⁷. Indeed, a previous study on a cardiovascular cardiac phenotype reported that nominally significant associations are enriched in enhancer regions⁸. Therefore, although none of the variants that we identified in our study display genome-wide significant association with ADHD, they may contribute to the susceptibility to ADHD, as they do have a functional impact (methylation, expression and in some cases brain structure) on genes that are expressed in brain. It is also supported by the data shown in Tables S2, S3, S4, S5 and S6, showing that most of our SNPs are related to different histone marks linked to possible promoter or enhancer regions. Brain-specific ASM information has also been utilized to detect key genes and pathways in bipolar disorder¹⁹. Also, enrichment of brain ASM in a schizophrenia GWAS in comparison to non-psychiatric GWAS reinforces the rationale of utilizing ASM SNPs to highlight genes that are relevant to psychiatric disorders from GWAS data⁹. None of our highlighted SNPs or CpG sites overlap with the ones described in a previous study relating ASM to schizophrenia⁶⁹.

There are some strengths and limitations in our study that should be discussed
Strengths: i) We used the largest meta-analysis of ADHD GWAS performed so far, including around 20,000 cases and 35,000 controls. ii) Two of the highlighted genes,

ARTN and *C2orf82*, had previously been associated with other psychiatric disorders. iii) For two of the genes there is more than one LD block showing the same effect on CpG site methylation. iv) Our results are concordant with eQTL information that had been assessed in an independent sample, with all the SNP showing the opposite effect on methylation of the promoter region and on the expression of a given gene in brain (more promoter methylation and less gene expression or vice versa), even for the different LD blocks included in each region. Limitations: i) We did not perform a follow-up study to replicate our association findings in an independent sample. ii) The previous studies that we used for the selection of ASM SNPs were performed on different genotyping platforms that do not include all the existing SNPs in the genome, and therefore we could not test all possible ASMs. iii) We only considered cis-associated ASM variants, which are the vast majority, although non-cis ASM also occurs. iv) There is an overrepresentation of ASM SNPs from cerebellum compared to the other studied tissues.

To conclude, our study points to the *ARTN*, *C2orf82* and *PIDD1* genes as potential contributors to ADHD susceptibility. The identified risk variants regulate the methylation levels of different CpG sites located in promoter regions and they inversely correlate with expression of the corresponding genes in brain. This finding is supported by a prediction of increased expression of *ARTN* and *PIDD1* and a decreased expression of *C2orf82* in ADHD. Moreover, most of the variants correlating with methylation at cg13047596 (near *C2orf82*) have an influence on accumbens and caudate volumes. Further studies are required to elucidate the mechanisms by which these genes contribute to ADHD.

ACKNOWLEDGMENTS

Major financial support for this research was received by BC from the Spanish 'Ministerio de Economía y Competitividad' (SAF2015-68341-R) and AGAUR, 'Generalitat de Catalunya' (2017-SGR-738). The research leading to these results has also received funding from the European Union Seventh Framework Programme [FP7/2007-2013] under grant agreement n° 602805 and from the European Union H2020 Program [H2020/2014-2020] under grant agreements n° 667302 and n° 643051, the latter supporting the contract of AS. LP-C and JC-D were supported by 'Generalitat de Catalunya' (2016 FI_B 00728 and 2015 FI_B 00448, respectively). LP-C was also supported by 'Ministerio de Educación, Cultura y Deporte' (FPU15/03867). AS was supported by the European Union H2020 Program [H2020/2014-2020] under grant agreements n° 643051 by the NF-C was supported by contracts of the 'Centro de Investigación Biomédica en Red de Enfermedades Raras' (CIBERER). VR was supported by the Graduate School of Health from the University of Aarhus. The iPSYCH team acknowledges support from the Lundbeck Foundation. Finally, SF was supported by the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 602805, the European Union's Horizon 2020 research and innovation programme under grant agreements n° 667302 and n° 728018 and NIMH grants 5R01MH101519 and U01 MH109536-01.

We are thankful to Roser Corominas (Universitat de Barcelona, Barcelona) for helpful advice. We are also grateful to the ADHD Working Group of the Psychiatric Genomics Consortium (PGC) and the iPSYCH team for distributing the summary statistics of the ADHD GWAS meta-analysis. Access to the PGC ADHD data was obtained through dbGaP project number 10608 that includes the following datasets: phs000016.v2.p2, phs000407.v1.p1, phs000358.v1.p1 and phs000490.v1.p1. We thank the ENIGMA

consortium for sharing the summary statistics of genome-wide association meta-analyses of MRI phenotypes.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- 1 Polanczyk G, De Lima MS, Horta BL, Biederman J, Rohde LA. The worldwide prevalence of ADHD: A systematic review and metaregression analysis. *Am J Psychiatry* 2007; **164**: 942–948.
- 2 American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 2013.
- 3 Faraone S V. et al. Molecular Genetics of Attention-Deficit/Hyperactivity Disorder. *Biol Psychiatry* 2005; **57**: 1313–1323.
- 4 Demontis D et al. Discovery Of The First Genome-Wide Significant Risk Loci For ADHD. *bioRxiv* 2017; : 145581.
- 5 Cross-Disorder Group of the Psychiatric Genomics Consortium et al. Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* 2013; **45**: 984–994.
- 6 Meaburn EL, Schalkwyk LC, Mill J. Allele-specific methylation in the human genome: Implications for genetic studies of complex disease. *Epigenetics* 2010; **5**: 578–582.
- 7 Boyle EA, Li YI, Pritchard JK. An Expanded View of Complex Traits: From Polygenic to Omnigenic. *Cell* 2017; **169**: 1177–1186.
- 8 Wang X et al. Discovery and validation of sub-threshold genome-wide association study loci using epigenomic signatures. *Elife* 2016; **5**: e10557.
- 9 Gagliano SA et al. Allele-Skewed DNA Modification in the Brain: Relevance to a Schizophrenia GWAS. *Am J Hum Genet* 2016; **98**: 956–962.
- 10 Ladd-Acosta C et al. DNA Methylation Signatures within the Human Brain. *Am J Hum Genet* 2007; **81**: 1304–1315.

- 11 Jang HS, Shin WJ, Lee JE, Do JT. CpG and Non-CpG Methylation in Epigenetic Gene Regulation and Brain Function. *Genes (Basel)* 2017; **8**: E148.
- 12 Xu Y et al. Multiple epigenetic factors predict the attention deficit/hyperactivity disorder among the Chinese Han children. *J Psychiatr Res* 2015; **64**: 40–50.
- 13 Schanen NC. Epigenetics of autism spectrum disorders. *Hum Mol Genet* 2006; **15**: R138–R150.
- 14 Grayson DR, Guidotti A. Merging data from genetic and epigenetic approaches to better understand autistic spectrum disorder. *Epigenomics* 2016; **8**: 85–104.
- 15 Dammann G et al. Increased DNA methylation of neuropsychiatric genes occurs in borderline personality disorder. *Epigenetics* 2011; **6**: 1454–1462.
- 16 Gaur U, Li K, Mei S, Liu G. Research progress in allele-specific expression and its regulatory mechanisms. *J Appl Genet* 2013; **54**: 271–283.
- 17 Li E, Zhang Y. DNA methylation in mammals. *Cold Spring Harb Perspect Biol* 2014; **6**: a019133.
- 18 Do C et al. Genetic-epigenetic interactions in cis: a major focus in the post-GWAS era. *Genome Biol* 2017; **18**: 120.
- 19 Chuang L-C, Kao C-F, Shih W-L, Kuo P-H. Pathway analysis using information from allele-specific gene methylation in genome-wide association studies for bipolar disorder. *PLoS One* 2013; **8**: e53092.
- 20 Gibbs JR et al. Abundant quantitative trait loci exist for DNA methylation and gene expression in Human Brain. *PLoS Genet* 2010; **6**: e1000952.
- 21 Zhang D et al. Genetic Control of Individual Differences in Gene-Specific Methylation in Human Brain. *Am J Hum Genet* 2010; **86**: 411–419.
- 22 Barrett JC. Haploview: Visualization and Analysis of SNP Genotype Data. *Cold*

- Spring Harb Protoc* 2009; **4**: 1–5.
- 23 The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* 2015; **526**: 68–74.
 - 24 John D. Storey, with contributions from Andrew J. Bass AD and, Robinson D. qvalue: Q-value estimation for false discovery rate control. R package version 2.6.0. 2015.<http://github.com/jdstorey/qvalue>.
 - 25 Ward LD, Kellis M. HaploReg v4: systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Res* 2016; **44**: D877–D881.
 - 26 GTEx Consortium et al. The Genotype-Tissue Expression (GTEx) project. *Nat Publ Gr* 2013; **45**: 580–585.
 - 27 Barbeira A et al. MetaXcan: Summary Statistics Based Gene-Level Association Method Infers Accurate PrediXcan Results. *bioRxiv* 2016; : 045260.
 - 28 Fromer M et al. Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nat Neurosci* 2016; **19**: 1442–1453.
 - 29 Hibar DP et al. Common genetic variants influence human subcortical brain structures. *Nature* 2015; **520**: 224–229.
 - 30 Baloh RH et al. Artemin, a novel member of the GDNF ligand family, supports peripheral and central neurons and signals through the GFR α 3-RET receptor complex. *Neuron* 1998; **21**: 1291–1302.
 - 31 Elitt CM et al. Artemin overexpression in skin enhances expression of TRPV1 and TRPA1 in cutaneous sensory neurons and leads to behavioral sensitivity to heat and cold. *J Neurosci* 2006; **26**: 8578–8587.
 - 32 Elitt CM, Malin SA, Koerber HR, Davis BM, Albers KM. Overexpression of artemin in the tongue increases expression of TRPV1 and TRPA1 in trigeminal

- afferents and causes oral sensitivity to capsaicin and mustard oil. *Brain Res* 2008; **1230**: 80–90.
- 33 Honma Y et al. Artemin is a vascular-derived neurotrophic factor for developing sympathetic neurons. *Neuron* 2002; **35**: 267–282.
- 34 Tinel A, Tschopp J. The PIDDosome, a Protein Complex Implicated in Activation of Caspase-2 in Response to Genotoxic Stress. *Science (80-)* 2004; **304**: 843–846.
- 35 Janssens S, Tinel A, Lippens S, Tschopp J. PIDD Mediates NF-kappaB Activation in Response to DNA Damage. *Cell* 2005; **123**: 1079–1092.
- 36 Berquin PC et al. Cerebellum in attention-deficit hyperactivity disorder: a morphometric MRI study. *Neurology* 1998; **50**: 1087–1093.
- 37 Castellanos FX et al. Quantitative brain magnetic resonance imaging in attention-deficit hyperactivity disorder. *Arch Gen Psychiatry* 1996; **53**: 607–616.
- 38 Castellanos FX et al. Quantitative brain magnetic resonance imaging in girls with attention-deficit/hyperactivity disorder. *Arch Gen Psychiatry* 2001; **58**: 289–295.
- 39 Toplak ME, Dockstader C, Tannock R. Temporal information processing in ADHD: findings to date and new methods. *J Neurosci Methods* 2006; **151**: 15–29.
- 40 Mackie S et al. Cerebellar Development and Clinical Outcome in Attention Deficit Hyperactivity Disorder. *Am J Psychiatry* 2007; **164**: 647–655.
- 41 Plessen KJ et al. Hippocampus and Amygdala Morphology in Attention-Deficit/Hyperactivity Disorder. *Arch Gen Psychiatry* 2006; **63**: 795–807.
- 42 Friedman LA, Rapoport JL. Brain development in ADHD. *Curr Opin Neurobiol* 2015; **30**: 106–111.

- 43 Pasini A, D'agati E. Pathophysiology of NSS in ADHD. *World J Biol Psychiatry* 2009; **10**: 495–502.
- 44 Curatolo P, D'Agati E, Moavero R. The neurobiological basis of ADHD. *Ital J Pediatr* 2010; **36**: 79.
- 45 Qiu A et al. Basal ganglia volume and shape in children with attention deficit hyperactivity disorder. *Am J Psychiatry* 2009; **166**: 74–82.
- 46 Greven CU et al. Developmentally Stable Whole-Brain Volume Reductions and Developmentally Sensitive Caudate and Putamen Volume Alterations in Those With Attention-Deficit/Hyperactivity Disorder and Their Unaffected Siblings. *JAMA Psychiatry* 2015; **72**: 490–499.
- 47 Sobel LJ et al. Basal ganglia surface morphology and the effects of stimulant medications in youth with attention deficit hyperactivity disorder. *Am J Psychiatry* 2010; **167**: 977–986.
- 48 Rubia K, Alegria AA, Brinson H. Brain abnormalities in attention-deficit hyperactivity disorder: a review. *Rev Neurol* 2014; **58 Suppl 1**: S3–S16.
- 49 Casey BJ et al. Implication of Right Frontostriatal Circuitry in Response Inhibition and Attention-Deficit/Hyperactivity Disorder. *J Am Acad Child Adolesc Psychiatry* 1997; **36**: 374–383.
- 50 Singh A, Yeh CJ, Verma N, Das AK. Overview of attention deficit hyperactivity disorder in young children. *Heal Psychol Res* 2015; **3:2115**: 23–35.
- 51 Onnink AMH et al. Brain alterations in adult ADHD: Effects of gender, treatment and comorbid depression. *Eur Neuropsychopharmacol* 2014; **24**: 397–409.
- 52 Hoogman M et al. Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis. *The lancet Psychiatry* 2017; **4**: 310–319.

- 53 Cortese S. The neurobiology and genetics of Attention-Deficit/Hyperactivity Disorder (ADHD): What every clinician should know. *Eur J Paediatr Neurol* 2012; **16**: 422–433.
- 54 Shaw P et al. Attention-deficit/hyperactivity disorder is characterized by a delay in cortical maturation. *Proc Natl Acad Sci U S A* 2007; **104**: 19649–19654.
- 55 de Zeeuw P et al. Differential Brain Development with Low and High IQ in Attention-Deficit/Hyperactivity Disorder. *PLoS One* 2012; **7**: e35770.
- 56 Shaw P et al. Longitudinal Mapping of Cortical Thickness and Clinical Outcome in Children and Adolescents With Attention-Deficit/Hyperactivity Disorder. *Arch Gen Psychiatry* 2006; **63**: 540–549.
- 57 Ambrosino S, de Zeeuw P, Wierenga LM, van Dijk S, Durston S. What can Cortical Development in Attention-Deficit/Hyperactivity Disorder Teach us About the Early Developmental Mechanisms Involved? *Cereb Cortex* 2017; **27**: 4624–4634.
- 58 Castellanos FX, Aoki Y. Intrinsic Functional Connectivity in Attention-Deficit/Hyperactivity Disorder: A Science in Development. *Biol psychiatry Cogn Neurosci neuroimaging* 2016; **1**: 253–261.
- 59 Makris N et al. Anterior cingulate volumetric alterations in treatment-naïve adults with ADHD: A pilot study. *J Atten Disord* 2010; **13**: 407–413.
- 60 Makris N et al. Cortical Thinning of the Attention and Executive Function Networks in Adults with Attention-Deficit/Hyperactivity Disorder. *Cereb Cortex* 2007; **17**: 1364–1375.
- 61 Yang X-R, Carrey N, Bernier D, MacMaster FP. Cortical Thickness in Young Treatment-Naive Children With ADHD. *J Atten Disord* 2015; **19**: 925–930.
- 62 Qiu M et al. Changes of Brain Structure and Function in ADHD Children. *Brain*

- Topogr* 2011; **24**: 243–252.
- 63 Pliszka SR, Lancaster J, Liotti M, Semrud-Clikeman M. Volumetric MRI differences in treatment-naïve vs chronically treated children with ADHD. *Neurology* 2006; **67**: 1023–1027.
 - 64 Wu Y et al. Integrative analysis of omics summary data reveals putative mechanisms underlying complex traits. *Nat Commun* 2018; **9**. doi:10.1038/s41467-018-03371-0.
 - 65 Otsuki K et al. Altered expression of neurotrophic factors in patients with major depression. *J Psychiatr Res* 2008; **42**: 1145–1153.
 - 66 Schizophrenia Working Group of the Psychiatric Genomics. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 2014; **511**: 421–427.
 - 67 Ripke S et al. Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat Genet* 2013; **45**: 1150–1159.
 - 68 Brown TE. *ADHD Comorbidities: Handbook for ADHD Complications in Children and Adults*. 2009.
 - 69 Numata S, Ye T, Herman M, Lipska BK. DNA methylation changes in the postmortem dorsolateral prefrontal cortex of patients with schizophrenia. *Front Genet* 2014; **5**: 280.

FIGURE LEGENDS

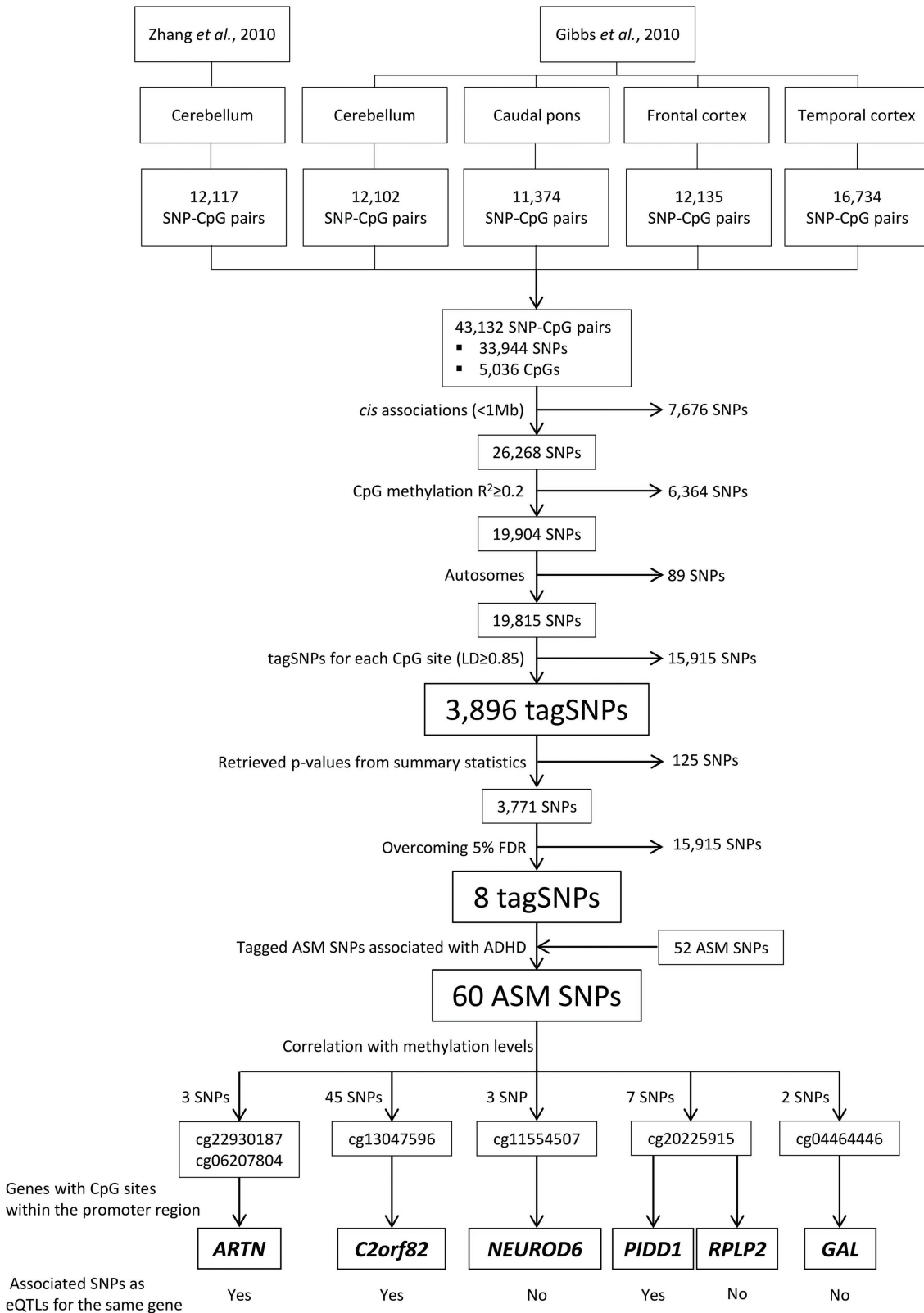
Fig. 1. Selection of allele-specific methylation (ASM) SNPs from two previous studies and association results obtained for ASM variants in ADHD. SNPs tested in the ADHD GWAS meta-analysis and multiple testing correction. SNPs correlating with differential methylation of CpG sites and eQTLs in brain regions (only for genes in which the CpG site lies < 5 kb from the transcription start site) are depicted.

Fig. 2. Genomic context of ASM variants, and methylation and eQTL information for cg22930187 and cg06207804. Genes are depicted in dark blue, showing the direction of transcription with an arrow; CpG sites inspected in the reference studies appear in brown; framed CpG sites indicate those sites showing differential levels of methylation for the associated ASM SNPs, and brown arrows indicate the effect on methylation of the ADHD risk variants, with indication of the brain regions where the ASMs were described. The tagSNPs are underscored. The colored rhombuses show the LD blocks present in each region. The colored dots for eQTLs indicate the effect on gene expression of the ADHD risk allele, according to the legend (red: over-expression, blue: under-expression). The number of enhancer (H3K4me1 and H3K27ac) and promoter (H3K4me3 and H3K9ac) histone marks found in the different brain areas are displayed for each SNP. ‘-’ indicates no known enhancer or promoter histone marks.

Fig. 3. Genomic context of ASM variants, and methylation and eQTL information for cg13047596. Genes are depicted in dark blue, showing the direction of transcription with an arrow; CpG sites inspected in the reference studies appear in brown; framed CpG sites indicate those sites showing differential levels of methylation for the associated ASM SNPs, and brown arrows indicate the effect on methylation of the ADHD risk variants, with indication of the brain regions where the ASMs were described. The tagSNPs are underscored. The colored rhombuses show the different

LD blocks present in each region. The colored dots for eQTLs indicate the effect on gene expression of the ADHD risk allele, according to the legend (red: over-expression, blue: under-expression). The number of enhancer (H3K4me1 and H3K27ac) and promoter (H3K4me3 and H3K9ac) histone marks found in the different brain areas are displayed for each SNP. '-' indicates no known enhancer or promoter histone marks.

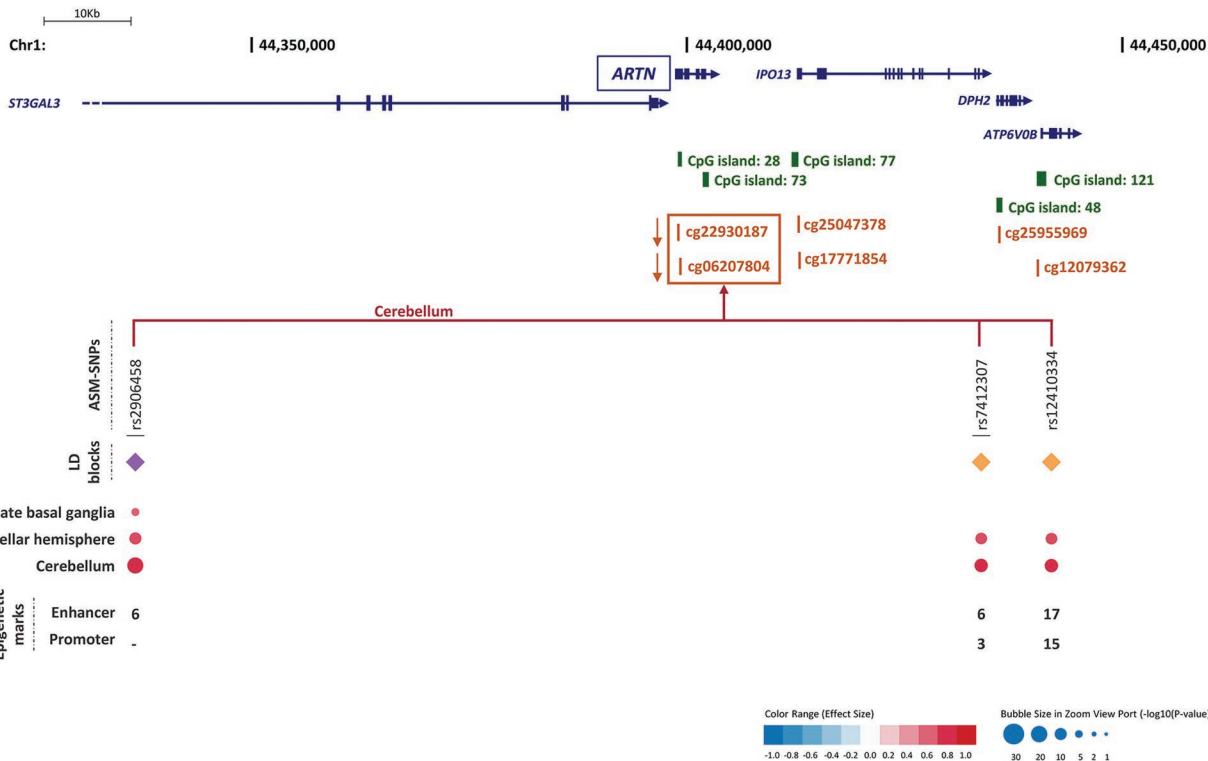
Fig. 4. Genomic context of ASM variants, and methylation and eQTL information for cg20225915. Genes are depicted in dark blue, showing the direction of transcription with an arrow; CpG sites inspected in the reference studies appear in brown; framed CpG sites indicate those sites showing differential levels of methylation for the associated ASM SNPs, and brown arrows indicate the effect on methylation of the ADHD risk variants, with indication of the brain regions where the ASMs were described. The tagSNPs are underscored. The colored rhombuses show the LD blocks present in each region. The colored dots for eQTLs indicate the effect on gene expression of the ADHD risk allele, according to the legend (red: over-expression, blue: under-expression). The number of enhancer (H3K4me1 and H3K27ac) and promoter (H3K4me3 and H3K9ac) histone marks found in the different brain areas are displayed for each SNP. '-' indicates no known enhancer or promoter histone marks.

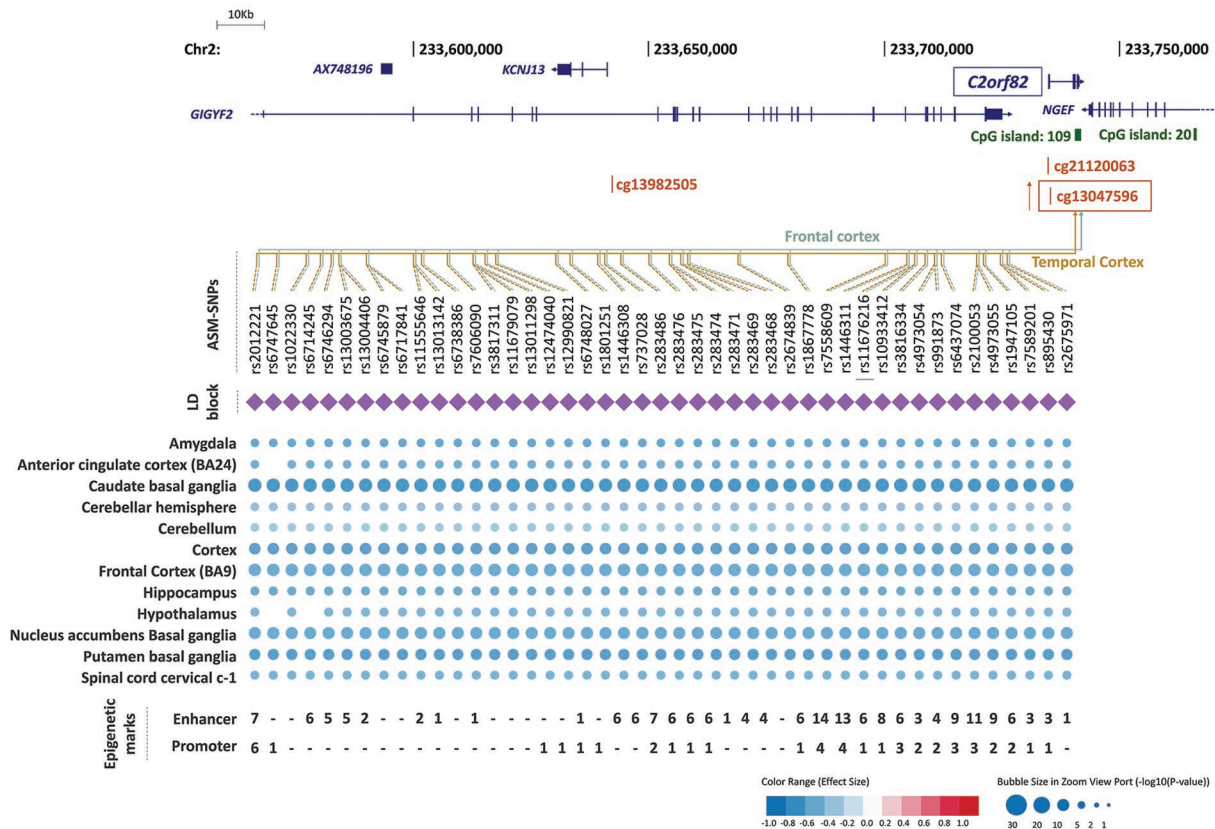


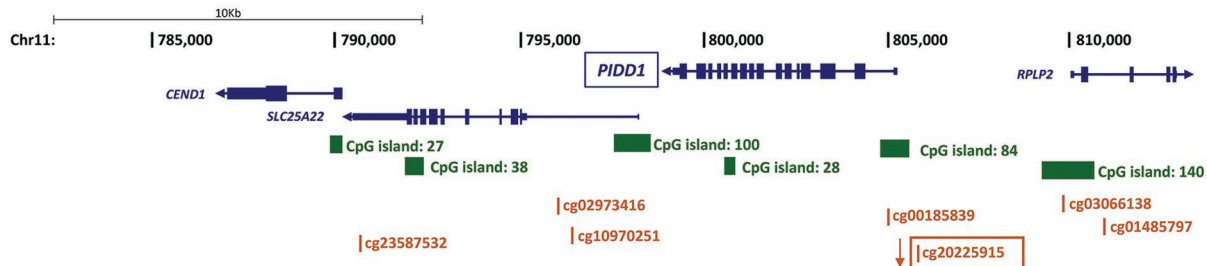
ASM SNP selection

ADHD
association
study

Follow-up study

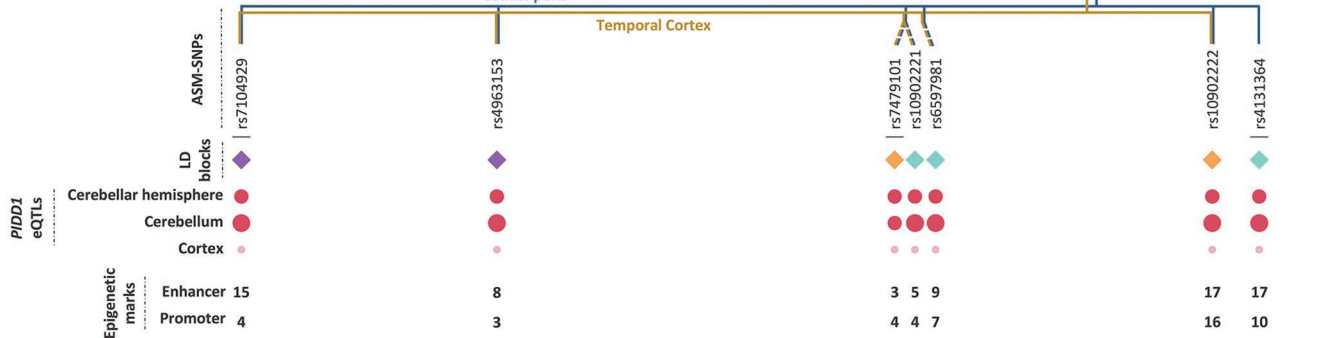






Caudal pons

Temporal Cortex



Color Range (Effect Size)



Bubble Size in Zoom View Port (-log10(P-value))



Table 1. Allele-Specific Methylation (ASM) tagSNPs associated with ADHD.

SNP	Chr	Pos	Alleles		OR	Freq A1		Meta-analysis p
			A1	A2		Cases	Controls	
rs2906458	1	44336389	A	<u>G</u>	0.94	0.74	0.756	3.01E-05
rs7412307	1	44433864	<u>C</u>	G	1.07	0.185	0.172	2.82E-05
rs11676216	2	233706368	T	<u>C</u>	0.95	0.646	0.654	6.78E-05
rs4140961	7	31349352	<u>A</u>	G	1.06	0.597	0.592	6.05E-05
rs7104929	11	784340	C	<u>G</u>	0.94	0.512	0.526	7.89E-06
rs7479101	11	802115	A	<u>G</u>	0.93	0.317	0.33	5.90E-06
rs4131364	11	812188	<u>A</u>	G	1.07	0.517	0.502	1.60E-06
rs11600377	11	68785803	<u>A</u>	G	1.06	0.731	0.72	4.38E-05

SNP: Single Nucleotide Polymorphism; Chr: Chromosome; Pos: Position (build hg19); A1: Allele 1; A2: Allele 2; OR: Odds Ratio (calculated on A1); Freq A1: Frequency of allele 1; In bold: Significant tagSNPs overcoming Bonferroni correction for multiple testing; Underlined allele: risk allele for ADHD; meta-analysis p: p-value obtained from the PGC+iPSYCH meta-analysis (Demontis *et al.*, 2017).