

## Epigenetic profiling of ADHD symptoms trajectories: A prospective, methylome-wide study

Esther Walton, PhD\*,<sup>1</sup> Jean-Baptiste Pingault, PhD\*,<sup>2,3</sup> Charlotte AM Cecil, PhD,<sup>1</sup> Tom R. Gaunt,  
PhD,<sup>4</sup> Caroline Relton, PhD,<sup>4</sup> Jonathan Mill, PhD,<sup>3,5</sup> Edward D. Barker, PhD<sup>1</sup>

\*joint first authorship

<sup>1</sup> Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK.

<sup>2</sup> Division of Psychology and Language Sciences, University College London, UK.

<sup>3</sup> MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, UK

<sup>4</sup> Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol, UK.

<sup>5</sup> University of Exeter Medical School, University of Exeter, Exeter, UK.

**Short Title:** Epigenetic profiling of ADHD trajectories.

**Keywords:** DNA Methylation; methylome-wide; ADHD; longitudinal; ALSPAC.

**Word Count abstract:**229

**Word Count:** 3392

**No. Figures:** 3

**No. Tables:** 1

**No. Supplementary Information:** 1

Corresponding author: Edward D. Barker, Department of Psychology, Institute of Psychiatry, Psychology & Neuroscience, King's College London, De Crespigny Park, London, SE5 8AF, UK. Tel: +44 (0)207 848 0992, E-mail: [ted.barker@kcl.ac.uk](mailto:ted.barker@kcl.ac.uk).

## Abstract

Attention-deficit/hyperactivity disorder (ADHD) is a prevalent developmental disorder, associated with a range of long-term impairments. Variation in DNA methylation, an epigenetic mechanism, is implicated in both neurobiological functioning and psychiatric health. However, the potential role of DNA methylation in ADHD symptoms is currently unclear. In this study, we examined data from the Avon Longitudinal Study of Parents and Children (ALSPAC) – specifically the subsample forming the Accessible Resource for Integrated Epigenomics Studies (ARIES) – which includes (i) peripheral measures of DNA methylation (Illumina 450k) at birth (n=817, 49% male) and age 7 (n=892, 50% male) and (ii) trajectories of ADHD symptoms (7-15 yrs). We first employed a genome-wide analysis to test whether DNA methylation at birth associates with later ADHD trajectories; and then followed up at age 7 to investigate the stability of associations across early childhood. We found that DNA methylation at birth differentiated ADHD trajectories across multiple genomic locations, including probes annotated to *SKI* (involved in neural tube development), *ZNF544* (previously implicated in ADHD), *ST3GAL3* (linked to intellectual disability) and *PEX2* (related to peroxisomal processes). **None of these probes maintained an association with ADHD trajectories at age 7.** Findings lend novel insights into the epigenetic landscape of ADHD symptoms, highlighting the potential importance of DNA methylation variation in genes related to neurodevelopmental and peroxisomal processes, which play a key role in the maturation and stability of cortical circuits.

## Introduction

Attention-Deficit Hyperactivity Disorder (ADHD) is one of the most prevalent and disabling psychiatric conditions in childhood and adolescence,<sup>1</sup> often persisting into adulthood<sup>2</sup> and associating with a range of long-term impairments.<sup>3-6</sup> Like other psychiatric disorders, the aetiology of ADHD is complex and multifactorial. While epidemiological studies have identified numerous environmental risk factors,<sup>7,8</sup> family-based studies have also documented the importance of genetic factors in the development of ADHD symptomatology,<sup>9-11</sup> although genome-wide association studies have yet to find genetic variants robustly associated with ADHD.<sup>10,12,13</sup> The mechanisms underlying the association between these risk factors and the development of ADHD symptoms remain to be elucidated. In recent years, epigenetic processes, such as DNA methylation, regulating gene expression have emerged as candidate mechanisms as they have been associated with environmental/genetic risk as well as neurobiological functioning and psychiatric wellbeing.<sup>14-17</sup>

Recent research has begun to demonstrate the potential of epigenetic research for understanding ADHD.<sup>18,19</sup> However, current literature on the topic remains scant and presents a number of important limitations. First, the majority of existing studies have been based on candidate genes (e.g. dopaminergic genes), which precludes the identification of novel findings. A methylome-wide analysis, in contrast, is hypothesis-free with respect to which genes might be involved and hence has the potential to detect novel biological associations.<sup>20</sup> Second, studies have relied primarily on cross-sectional designs featuring DNA methylation at a single time point. This has precluded the possibility of examining whether altered DNA methylation patterns are a risk factor for and/or consequence of ADHD, as well as establishing the stability of associations across time.

To our knowledge, only one study to date has conducted a methylome-wide analysis of ADHD, implicating several potential biological pathways related to inflammatory mechanisms, such as homocysteine and fatty acid oxidation.<sup>21</sup> However, the study focussed on a relatively small sample (N = 105) of boys and relied in part on a priori information to identify differentially methylated sites. The cross-sectional design also meant that it was not possible to disentangle epigenetic predictors of ADHD from a posteriori markers of ADHD and/or associated characteristics (e.g. medication or stress resulting from ADHD symptoms). In the present study we aimed to address these gaps in the literature, by conducting the first methylome-wide study of ADHD symptomatology in a large population-based sample featuring repeated measures of DNA methylation and the use of a prospective design spanning birth to adolescence.

## Methods

### *Participants*

Participants were drawn from the Accessible Resource for Integrated Epigenomics Studies (ARIES, [www.ariesepigenomics.org.uk](http://www.ariesepigenomics.org.uk)),<sup>22</sup> containing DNA methylation data for a subset of 1018 mother-offspring pairs and nested within the *Avon Longitudinal Study of Parents and Children* (ALSPAC). ALSPAC is an ongoing epidemiological study of children born from 14,541 pregnant women residing in Avon, UK, with an expected delivery date between April 1991 and December 1992 (85% of eligible population).<sup>23</sup> **Informed consent was obtained from all ALSPAC participants and** ethical approval was obtained from the ALSPAC Law and Ethics Committee as well as Local Research Committees. The original ALSPAC sample is representative of the general population.<sup>24</sup> Please note that the study

website contains details of all the data that is available through a fully searchable data dictionary: <http://www.bris.ac.uk/alspac/researchers/dataaccess/data-dictionary/>.

For this study, we included youth from ARIES who had available data on ADHD symptomatology ratings (age 7-15) as well as epigenetic data at birth (n = 817, 49% male) and/or age 7 (n = 892, 50% male). **The overlap was n = 783 participants with DNA methylation at birth and age 7 as well as ADHD ratings (see Supplementary section 1.1 for details).**

### *Measures*

#### ADHD

ADHD symptomatology was assessed via maternal ratings at ages 7, 10, 13 and 15 years, using the well-validated Development and Well-Being Assessment interview (DAWBA).<sup>25</sup> The DAWBA was administered via computer, generating the following ‘probability bands’ (i.e. levels of prediction of the probability of disorder for a DSM-IV diagnosis of ADHD, ranging from 0 – very unlikely – to 5 – probable): 0: < 0.1% probability of children in this band having the disorder; 1: ~0.5%; 2: ~3%; 3: ~15%; 4: ~40%; 5: >70%, respectively. **See Supplementary section 1.2 for more details.**

#### DNA methylation data

500ng genomic DNA from blood (cord at birth; whole at age 7) was bisulfite-converted using the EZ-DNA methylation kit (Zymo Research, Orange, CA, USA). DNA methylation was quantified using the Illumina HumanMethylation450 BeadChip (HM450k; Illumina, USA) with arrays scanned using an Illumina iScan (software version 3.3.28). Samples ( $n_{\text{birth}} = 25$ ;  $n_{\text{age 7}} = 8$ ) or probes ( $n_{\text{birth}} = 7873$ ;  $n_{\text{age 7}} = 4861$ ) that failed quality control (>1% probes/samples with background detection p-value  $\geq 0.05$ ) were excluded from

further analysis. Sex checks were performed using X/Y chromosome methylation. Genotype probes on the HM450k were compared between samples from the same individual and against SNP-chip data to identify and remove any sample mismatches. Samples were quantile normalised using the *dasen* function within the *wateRmelon* package (version 1.4.0) in R. Normalization performance was evaluated using all three testing metrics in *wateRmelon* (*genki* assessing SNP-related probes, *dmrse* assessing imprinted probes and *seabi*, assessing gender differences). Methylation levels were then indexed by beta values (corresponding to the ratio of methylated signal divided by the sum of the methylated and unmethylated signal). **Probes known to be cross-reactive or polymorphic<sup>26,27</sup> and SNP (i.e. “rs”) probes were removed (n = 72,068).** We also removed participants with non-caucasian or missing ethnicity (based on self-reports; n = 61), leaving a total of 828 (cord) and 903 samples (age 7) after quality control. Cell type proportions (CD8 T lymphocytes, CD4 T lymphocytes, natural killer cells, B lymphocytes, monocytes and granulocytes) for each participant were estimated using the reference-based approach detailed in Houseman et al.<sup>28</sup> As a final step, we regressed out chip and cell type to remove potentially confounding effects. For more information, see Supplementary section 1.3.

## *Analyses*

### ADHD trajectories

Trajectories of hyperactivity/inattention across 7 to 15 years were estimated using k-means for longitudinal data (Package *KmL*).<sup>29-31</sup> This non-parametric procedure classifies participants into developmental trajectories, i.e. homogenous subgroups following similar developmental patterns. **Each participant is first assigned arbitrarily to one initial trajectory. Next, the center (mean) of each trajectory is calculated and each participant**

is reassigned to the closest trajectory. The operation is repeated until convergence (i.e. until no further mean adjustment change occurs in the trajectories). The process from assignment to convergence is then repeated (500 times in the present study) to make sure that the solution is not dependent on the initial assignment. The best solution is determined by a criterion that maximizes a ratio computed by dividing the trace of the between-variance by the trace of the within-variance (i.e. maximizing the differences between trajectories and maximizing the homogeneity within trajectories). Additional details on the procedure and the choice of the best solution are provided in Supplementary sections 1.4 and 1.5. The final sample consisted of  $n_{\text{birth}} = 777$  and  $n_{\text{age7}} = 842$  in the low trajectory group against  $n_{\text{birth}} = 40$  and  $n_{\text{age7}} = 50$  in the high trajectory group (see results section for further details).

#### Methylome-wide analysis

Methylome-wide association analysis between DNA methylation (407,462 probes, cell type and batch-corrected) and ADHD trajectories were performed at birth, adjusting for sex, using a **general linear model**. Differentially methylated probes (DMPs) surviving a False Discovery Rate (FDR) correction of  $q < 0.05$  were **visually inspected to assess equal variance between trajectories and** analysed further to determine whether they were also **nominal**ly significant at age 7 (i.e. follow-forward approach, **FDR-corrected  $q < 0.05$** ) and whether the direction of association was consistent across time points. All analyses were performed in R (version 3.0.2) using package IMA.<sup>32</sup> ~~with default parameters.~~

#### Code availability

Computer code used in our analyses is available from the authors upon request.

## Network analysis

To further analyse underlying genetic networks of ADHD trajectory-associated DMPs, we imported and analysed all genes related to FDR-corrected DMPs at birth using GeneMANIA (<http://www.genemania.org>) bioinformatics software with default parameters. **For more information on GeneMANIA methods, see Supplementary section 2.2 and references.**<sup>33,34</sup>

## **Results**

Trajectory analyses of the DAWBA ADHD scores yielded a 3 trajectory solution (Figure 1 and Supplementary section 1.5). Participants in the high trajectory (6.1%) had DAWBA band scores of around 3 across time points. Conversely, participants in the null trajectory (67.4%) and in the low trajectory (26.5%), had stable scores of around 0 and 1 respectively, corresponding to a close to zero probability of being ADHD cases. In the methylome-wide analyses, we therefore used a binary variable grouping the null and the low trajectories (called low thereafter) ( $n_{\text{birth}} = 777$  and  $n_{\text{age7}} = 842$ ) against the high trajectory ( $n_{\text{birth}} = 40$  and  $n_{\text{age7}} = 50$ ).

### *Methylome-wide analysis of ADHD trajectories*

At birth, 13 probes were differentially methylated between ADHD trajectories after FDR correction ( $q < 0.05$ ; Table 1 and Figure 2A). Inspection of the QQ-plot (Supplementary Figure 3) and a lambda statistic of 1.056 provided little evidence of inflation of test statistics. **Additionally, visual inspection of boxplots gave no strong indication for a violation of equal variance assumption between trajectories for any of these 13 probes (see**



**Supplementary section 2.1).** Cg24481594, the most significant DMP ( $\beta_{\text{stdn}} = -0.198$ ;  $p = 1.51 \times 10^{-8}$ ;  $q = 0.006$ ), was hypo-methylated in the high trajectory and is annotated to *SKI*, a gene related to Transforming Growth Factor-beta (TGF-beta) signalling and neural tube development.<sup>35,36</sup> Other DMPs of interest were located in genes such as (i) ***EPX*** (cg27469152;  $\beta_{\text{stdn}} = -0.181$ ;  $p = 2.26 \times 10^{-7}$ ;  $q = 0.031$ ; **Figure 2B**), a member of the peroxidase gene family, and ***PEX2*** (cg16290904;  $\beta_{\text{stdn}} = 0.173$ ;  $p = 7.35 \times 10^{-7}$ ;  $q = 0.048$ ; **Figure 2C**), a peroxisomal membrane protein gene involved in myelin production and fatty-acid metabolism; (ii) *ST3GAL3* (cg09989037;  $\beta_{\text{stdn}} = -0.172$ ;  $p = 9.46 \times 10^{-7}$ ;  $q = 0.048$ ), a gene linked to mental retardation;<sup>37</sup> (iii) *FBXW5* (cg13714586;  $\beta_{\text{stdn}} = 0.170$ ;  $p = 1.30 \times 10^{-6}$ ;  $q = 0.048$ ), associated with interleukin-1B signalling; (iv) *ELF3* (cg05653018;  $\beta_{\text{stdn}} = 0.169$ ;  $p = 1.43 \times 10^{-6}$ ;  $q = 0.048$ ), involved in preimplantation development; and (v) *ZNF544* (cg26263766;  $\beta_{\text{stdn}} = 0.173$ ;  $p = 8.72 \times 10^{-7}$ ;  $q = 0.048$ ; **Figure 2D**), implicated in transcriptional regulation and previously shown to associate with ADHD.<sup>38</sup> Absolute mean percent methylation difference between the high and low trajectory group for the 13 DMPs passing FDR-correction was 2.3% (range 0.6 – 4.8).

#### *Network analysis of FDR-corrected probes*

All genes annotated to FDR-corrected DMPs ( $n = 13$ ) were entered into a network analysis using GeneMANIA, which is based on known genetic and physical interactions, shared pathways and protein domains as well as protein co-expression data. The analysis showed that these genes form a compact cluster network (see **Figure 3**). The most enriched biological functions related to peroxisomal processes (**involving the genes *PEX2*, *PEX10* and *PEX12***;  $p_{\text{FDR}}$  ranging from 0.008 - 0.012; **Supplementary Table 3**), followed by functions related to transcription activity ( $p_{\text{FDR}} = 0.019$ ). **However, it should be noted that only *PEX2***

was directly measured, while the involvement of *PEX10* and *PEX12* was indirectly inferred based on shared protein domains and physical interactions.

#### *Follow-forward analysis at age 7*

As a last step, we investigated whether **any of the 13 DMPs** that differentiated ADHD trajectories at birth also did so at age 7. **While two of the 13 DMPs at birth also showed nominally significant effects at age 7** (cg27469152 located in the gene *EPX*;  $\beta_{\text{stdn}} = -0.090$ ;  $p = 0.008$ ; and cg16290904 located in the gene *PEX2*;  $\beta_{\text{stdn}} = -0.077$ ;  $p = 0.022$ ; Table 1), **no probe remained significant after FDR-correction for 13 test**. The direction of effect was consistent for cg27469152 (decreased methylation in the high trajectory group, both at birth and at age 7), but not for cg16290904.

#### *Additional analyses*

**Given that a minority** (see Supplementary Figure 1) **of participants had DNA methylation data only at birth or age 7**, we repeated the analyses in a subsample with complete DNA methylation data at both time points ( $n = 783$ ). **In agreement with our previous findings, effect sizes at birth were consistent in direction and size. Mean percent variation compared to the original effect sizes was 5.5% (standard error  $\pm 2.6$ ).** **Similar to the main analysis, no marker replicated at a FDR-corrected significance level at age 7 years (Supplementary section 2.3).**

## Discussion

In this study, we employed a methylome-wide prospective analysis (birth, age 7) with trajectories of ADHD symptoms (7-15 years). We identified 13 probes at birth that were differentially methylated between high and low trajectories of ADHD symptoms, **none** of which continued to be differentially methylated at age 7. Detected probes were located in the vicinity of genes implicated in peroxisomal processes, neural tube development and mental retardation, as well as one gene previously associated with ADHD. **We first discuss our findings in light of the few previous epigenetic studies on ADHD, before turning to potentially relevant biological mechanisms suggested by our findings.** We also discuss differences in results between birth and age 7 years from a developmental perspective.

**Our findings both contrast and support previous epigenetic studies of ADHD.** With regard to contrasting results, we found no evidence of differential methylation in *VIPR2* (a gene linked to mood disorders and circadian rhythm regulation), which was identified as a top hit by the only other published methylome-wide study on ADHD.<sup>21</sup> We also found no evidence of differential methylation in *DRD4*, which was reported as significantly associated with ADHD in two previous candidate gene studies.<sup>18,19</sup> A number of factors may explain these discrepancies, including differences in: (i) samples (population versus clinical); (ii) assessment (diagnosis versus continuous ratings); (iii) analytical methods (candidate genes vs. methylome-wide); (iv) design (prospective/longitudinal vs cross-sectional); and (v) developmental period examined (birth vs childhood). As a result, consideration of these differences should be applied when interpreting the findings of the present study.

Conversely, our findings regarding peroxisomal processes lend support to the involvement of mechanisms related to fatty acid oxidation in ADHD, as previously

reported by Wilmot et al.<sup>21</sup> Peroxisomes are cell components that play a key role in the metabolism of essential fatty acids from the omega-3 family. In particular, dietary alpha-linolenic acids are transformed into docosahexaenoic acid (DHA) through a final  $\beta$ -oxidation reaction that takes place in the peroxisomes.<sup>40</sup> ~~Peroxisomal processes are also implicated in the process leading to the formation of plasmalogens, a phospholipid constituting the major part of myelin sheath tissue.~~<sup>41</sup> DHA is particularly relevant as it accumulates in brain tissue at a rapid rate during the third trimester of pregnancy and continues to do so throughout early childhood and adolescence, playing an essential role in the maturation and stability of cortical circuits as well as in several other processes (e.g. implication in neurotransmitter systems including dopamine and serotonin).<sup>42</sup> ~~In addition to the myelination of axons, DHA is implicated in several aspects of brain maturation such as increasing growth factors (e.g. *NGF*; *BDNF*), promoting neurite outgrowth as well as being implicated in neurotransmitter systems including dopamine, serotonin and acetylcholine.~~<sup>42</sup> Deficits in DHA have been related to several psychiatric disorders, including ADHD.<sup>42,43</sup> **For instance, a recent meta-analysis of case control studies has demonstrated that ADHD is associated with robust blood DHA deficits.**<sup>44</sup> Furthermore, meta-analyses of randomized controlled trials in ADHD patients have shown a small but statistically significant effect of omega-3 fatty acid supplementation on ADHD symptoms<sup>43</sup> as well as some effects on cognition in people with low dietary omega-3 intake.<sup>45</sup> ~~Interestingly, the only other methylome-wide study on ADHD to date also identified hypomethylation in processes related to fatty acid oxidation.~~<sup>21</sup> Consequently, our findings, together with those of previous research,<sup>21</sup> suggest that early life methylation patterns in the peroxisomal network contribute to ADHD symptomatology in childhood and adolescence, possibly through disruptive effects on DHA synthesis. **If the association between DNA methylation patterns in this network and ADHD is indeed mediated by peroxisome abnormalities leading to DHA deficits, then**

**future studies should investigate whether dietary intake of preformed DHA, which bypasses peroxisome biosynthesis, mitigates the effect of altered DNA methylation patterns.**

In addition to peroxisomal processes, significant probes pointed towards several other processes associated with the development of the central nervous system. In particular, the most significant probe at birth was linked to SKI proto-oncogene, which functions as a suppressor of transforming growth factor-beta signalling, and is implicated in neural tube development and myelination.<sup>35,36,46</sup> **A study by Atanasoski et al.<sup>46</sup> showed that overexpression of *Ski* in cultured Schwann cells – the main glial cells in the peripheral nervous system - causes upregulation of myelin protein genes, indicating that *Ski* is an essential regulator that controls Schwann cell myelination. Interestingly, children with neurofibromatosis type 1 – a genetic disorder that can result in neurofibromas, which primarily affect Schwann cells<sup>47</sup> – are three times as likely to meet DSM-IV diagnostic criteria for ADHD compared to their unaffected siblings.<sup>48</sup> An additional probe, which we identified, was linked to *ST3GAL3*, a gene encoding a membrane protein involved in cellular communication located on chr1p34.1. This region (and *ST3GAL3* itself) had been previously associated with intellectual disability using linkage analysis, chromosome sorting and next-generation sequencing.<sup>37,49,50</sup> Furthermore, we identified two probes that are linked to *ZNF544* and *ZNF454*, respectively. Both genes belong to the same zinc finger family (C2H2-type) and are involved in gene transcription. Strikingly, *ZNF544* - along with several other *ZNF* genes - was previously flagged up in GWAS analysis of ADHD.<sup>38</sup> Using data from 376 family trios and a composite quantitative phenotype of ADHD symptoms based on DSM-IV criteria, the authors of that study identified, among others, a SNP linked to *ZNF544*. This marker is around**

**30,000 basepairs away from the CpG site identified in this study, which is associated with the same gene.**

Among the 13 probes identified at birth, **none** were still associated at age 7 years with ADHD trajectories **after FDR-correction**. A number of non-exclusive factors may drive this non-replication. First, a number of the 13 probes detected at birth may have been false positives. Although we had a much larger sample than any previous study, false positives are still possible. **Second, early life methylation patterns may be particularly important for the development of ADHD.** For example, one probe was linked to a gene involved in preimplantation development whose differential methylation may be only detectable at birth. In addition, it has been shown that DHA accrual in the brain is particularly important in the third trimester of pregnancy and the first year of life.<sup>42</sup> These and other processes that are more salient during early development may lead to enduring individual differences (e.g. in brain structure) without the epigenetic association being maintained. **Third, there is mounting evidence that DNA methylation patterns change considerably across development.<sup>51,52</sup> Consequently, a large proportion of epigenetic effects may be specific to certain developmental epochs. In the present case, differences in methylation patterns between birth and age 7 may reflect differences in environmental exposures at both ages and/or developmental genetic influences that have been reported for ADHD symptoms.<sup>11</sup> Importantly, given the current lack of similar prospective epigenetic studies in ADHD, it is not yet clear what should be expected in terms of continuity vs. discontinuity in epigenetic patterns. Therefore, the above considerations remain inevitably speculative and necessitate further investigation.**

### *Limitations*

The present findings should be interpreted in light of a number of limitations. Because of the use of a population-based sample, the proportion of youth showing severe ADHD symptomatology across time was relatively small and necessitates replication using larger clinical samples. **Furthermore, although the DAWBA is a well-validated and extensively used measure based on DSM-IV criteria, it is not, per se, a clinical diagnostic tool. As such, it will be important in future to replicate findings in clinical populations where diagnostic assessments are available. As psychostimulants can influence DNA methylation,<sup>53</sup> the use of clinical samples would also facilitate analyses investigating the effect of psychostimulant treatment on DNA methylation, which we were unable to do in the present sample.** The use of a larger number of cases vs controls will also make it possible to examine potential sex differences in the association between DNA methylation and ADHD symptoms.

Findings were based on DNA methylation from blood samples. Given that methylation patterns can be tissue-specific, the extent to which these changes reflect changes in the brain will need to be established. This is particularly relevant given that markers identified were related to genes involved in neural function and development. The analysis of transcriptomic data will also be important for assessing the functional significance of DNA methylation changes to gene expression levels.

**The markers identified in this study were not validated using alternative methods such as bisulfite-pyrosequencing. While we note that our previous work with bisulfite-pyrosequencing has demonstrated that the Illumina HM450 is a robust and sensitive platform for the detection of DNA methylation differences,<sup>54,55</sup> future work is needed to validate our findings further.**

## *Conclusion*

In conclusion, the present findings suggest that epigenetic mechanisms related to specific neurodevelopmental processes, such as neural tube development and peroxisomal mechanisms, are implicated in ADHD symptomatology. These results lend novel insights into longitudinal epigenetic risk markers for ADHD, pinpointing specific targets for further interrogation.

## **Acknowledgements**

We are extremely grateful to all the families who took part in this study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. With regard to the ALSPAC DNA methylation, we thank all involved, particularly the laboratory scientists and bioinformaticians who contributed considerable time and expertise to the data in this paper. The UK Medical Research Council and the Wellcome Trust (Grant ref: 092731) and the University of Bristol provide core support for ALSPAC. This work was funded by the National Institute of Child and Human Development grant (R01HD068437). ARIES was funded by the BBSRC (BBI025751/1 and BB/I025263/1). ARIES is maintained under the auspices of the MRC Integrative Epidemiology Unit at the University of Bristol (MC\_UU\_12013/2 and MC\_UU\_12013/8). Dr Walton is supported by the German Research Foundation. Dr Pingault is supported by a European Commission Marie Curie Intra-European Fellowship [N° 330699]. Dr Cecil is supported by the Economic and Social Research Council (grant ref: ES/N001273/1).



## **Conflict of Interest**

The authors declare that they have no conflict of interest.

## References

- 1 Willcutt EG. The prevalence of DSM-IV attention-deficit/hyperactivity disorder: a meta-analytic review. *Neurotherapeutics* 2012; **9**: 490–499.
- 2 Franke B, Faraone SV, Asherson P, Buitelaar J, Bau CHD, Ramos-Quiroga JA *et al.* The genetics of Attention Deficit/Hyperactivity Disorder in adults, a review. *Mol Psychiatry* 2012; **17**: 960–987.
- 3 Mannuzza S, Klein RG, Moulton III JL. Lifetime criminality among boys with attention deficit hyperactivity disorder: A prospective follow-up study into adulthood using official arrest records. *Psychiatry Res* 2008; **160**: 237–246.
- 4 Polderman TJC, Boomsma DI, Bartels M, Verhulst FC, Huizink AC. A systematic review of prospective studies on attention problems and academic achievement. *Acta Psychiatr Scand* 2010; **122**: 271–284.
- 5 Pingault J-B, Tremblay RE, Vitaro F, Carbonneau R, Genolini C, Falissard B *et al.* Childhood trajectories of inattention and hyperactivity and prediction of educational attainment in early adulthood: a 16-year longitudinal population-based study. *Am J Psychiatry* 2011; **168**: 1164–1170.
- 6 Pingault J-B, Côté S, Galéra C, Genolini C, Falissard B, Vitaro F *et al.* Childhood trajectories of inattention, hyperactivity and oppositional behaviors and prediction of substance abuse/dependence: a 15-year longitudinal population-based study. *Mol Psychiatry* 2013; **18**: 806–812.

- 7 Foulon S, Pingault J-B, Larroque B, Melchior M, Falissard B, Côté SM. Developmental predictors of inattention-hyperactivity from pregnancy to early childhood. *PLoS ONE* 2015; **10**: e0125996.
- 8 Galéra C, Côté SM, Bouvard MP, Pingault J-B, Melchior M, Michel G *et al.* Early risk factors for hyperactivity-impulsivity and inattention trajectories from age 17 months to 8 years. *Arch Gen Psychiatry* 2011; **68**: 1267–1275.
- 9 Chang Z, Lichtenstein P, Asherson PJ, Larsson H. Developmental twin study of attention problems: high heritabilities throughout development. *JAMA Psychiatry* 2013; **70**: 311–318.
- 10 Merwood A, Asherson P. Attention Deficit Hyperactivity Disorder: a lifespan genetic perspective. *Adv Ment Health Intellect Disabil* 2011; **5**: 33–46.
- 11 Pingault J-B, Viding E, Galéra C, Greven CU, Zheng Y, Plomin R *et al.* Genetic and Environmental Influences on the Developmental Course of Attention-Deficit/Hyperactivity Disorder Symptoms From Childhood to Adolescence. *JAMA Psychiatry* 2015; **72**: 651–658.
- 12 Neale BM, Medland SE, Ripke S, Asherson P, Franke B, Lesch K-P *et al.* Meta-analysis of genome-wide association studies of attention-deficit/hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry* 2010; **49**: 884–897.
- 13 Yang L, Neale BM, Liu L, Lee SH, Wray NR, Ji N *et al.* Polygenic transmission and complex neuro developmental network for attention deficit hyperactivity disorder: genome-wide association study of both common and rare variants. *Am J Med Genet B Neuropsychiatr Genet* 2013; **162B**: 419–430.

- 14 Cecil C a. M, Lysenko LJ, Jaffee SR, Pingault J-B, Smith RG, Relton CL *et al.* Environmental risk, Oxytocin Receptor Gene (OXTR) methylation and youth callous-unemotional traits: a 13-year longitudinal study. *Mol Psychiatry* 2014; **19**: 1071–1077.
- 15 Mill J, Petronis A. Pre- and peri-natal environmental risks for attention-deficit hyperactivity disorder (ADHD): the potential role of epigenetic processes in mediating susceptibility. *J Child Psychol Psychiatry* 2008; **49**: 1020–1030.
- 16 Lutz P-E, Turecki G. DNA methylation and childhood maltreatment: From animal models to human studies. *Neuroscience* 2014; **264**: 142–156.
- 17 McRae AF, Powell JE, Henders AK, Bowdler L, Hemani G, Shah S *et al.* Contribution of genetic variation to transgenerational inheritance of DNA methylation. *Genome Biol* 2014; **15**: 1–10.
- 18 Van Mil NH, Steegers-Theunissen RPM, Bouwland-Both MI, Verbiest MMPJ, Rijlaarsdam J, Hofman A *et al.* DNA methylation profiles at birth and child ADHD symptoms. *J Psychiatr Res* 2014; **49**: 51–59.
- 19 Xu Y, Chen X-T, Luo M, Tang Y, Zhang G, Wu D *et al.* Multiple epigenetic factors predict the attention deficit/hyperactivity disorder among the Chinese Han children. *J Psychiatr Res* 2015; **64**: 40–50.
- 20 Sullivan PF. The Psychiatric GWAS Consortium: Big Science Comes to Psychiatry. *Neuron* 2010; **68**: 182–186.
- 21 Wilmot B, Fry R, Smeester L, Musser ED, Mill J, Nigg JT. Methylomic analysis of salivary DNA in childhood ADHD identifies altered DNA methylation in VIPR2. *J Child Psychol Psychiatry* 2015. doi:10.1111/jcpp.12457.

- 22 Relton CL, Gaunt T, McArdle W, Ho K, Duggirala A, Shihab H *et al.* Data Resource Profile: Accessible Resource for Integrated Epigenomic Studies (ARIES). *Int J Epidemiol* 2015; : dyv072.
- 23 Fraser A, Macdonald-Wallis C, Tilling K, Boyd A, Golding J, Davey Smith G *et al.* Cohort Profile: The Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *Int J Epidemiol* 2013; **42**: 97–110.
- 24 Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J *et al.* Cohort Profile: the ‘children of the 90s’--the index offspring of the Avon Longitudinal Study of Parents and Children. *Int J Epidemiol* 2013; **42**: 111–127.
- 25 Goodman A, Heiervang E, Collishaw S, Goodman R. The ‘DAWBA bands’ as an ordered-categorical measure of child mental health: description and validation in British and Norwegian samples. *Soc Psychiatry Psychiatr Epidemiol* 2011; **46**: 521–532.
- 26 Chen Y, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW *et al.* Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics* 2013; **8**: 203–209.
- 27 Price EM, Cotton AM, Lam LL, Farré P, Emberly E, Brown CJ *et al.* Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium HumanMethylation450 BeadChip array. *Epigenetics & Chromatin* 2013; **6**: 4.
- 28 Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH *et al.* DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 2012; **13**: 86.

- 29 Genolini C, Falissard B. KmL: a package to cluster longitudinal data. *Comput Methods Programs Biomed* 2011; **104**: e112–121.
- 30 Hartigan JA, Wong MA. Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society Series C (Applied Statistics)* 1979; **28**: 100–108.
- 31 Genolini C, Falissard B. KmL: k-means for longitudinal data. *Computational Statistics* 2010; **25**: 317–328.
- 32 Wang D, Yan L, Hu Q, Sucheston LE, Higgins MJ, Ambrosone CB *et al.* IMA: An R package for high-throughput analysis of Illumina’s 450K Infinium methylation data. *Bioinformatics* 2012; : bts013.
- 33 Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P *et al.* The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res* 2010; **38**: W214–220.
- 34 Mostafavi S, Ray D, Warde-Farley D, Grouios C, Morris Q. GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. *Genome Biol* 2008; **9 Suppl 1**: S4.
- 35 Amaravadi LS, Neff AW, Sleeman JP, Smith RC. Autonomous neural axis formation by ectopic expression of the protooncogene c-ski. *Dev Biol* 1997; **192**: 392–404.
- 36 Chen D, Lin Q, Box N, Roop D, Ishii S, Matsuzaki K *et al.* SKI knockdown inhibits human melanoma tumor growth in vivo. *Pigment Cell Melanoma Res* 2009; **22**: 761–772.
- 37 Hu H, Eggers K, Chen W, Garshasbi M, Motazacker MM, Wrogemann K *et al.* ST3GAL3 Mutations Impair the Development of Higher Cognitive Functions. *The American Journal of Human Genetics* 2011; **89**: 407–414.

- 38 Lasky-Su J, Neale BM, Franke B, Anney RJL, Zhou K, Maller JB *et al.* Genome-wide association scan of quantitative traits for attention deficit hyperactivity disorder identifies novel associations and confirms candidate gene associations. *Am J Med Genet* 2008; **147B**: 1345–1354.
- 39 Cecil C a. M, Walton E, Viding E. Epigenetics of addiction: Current knowledge, challenges and future directions. *J Stud Alcohol Drugs* 2016; : In Press.
- 40 Masters C. Omega-3 fatty acids and the peroxisome. *Mol Cell Biochem* 1996; **165**: 83–93.
- 41 Farooqui AA, Horrocks LA. Plasmalogens, phospholipase A2, and docosahexaenoic acid turnover in brain tissue. *J Mol Neurosci* 2001; **16**: 263–272; discussion 279–284.
- 42 McNamara RK, Vannest JJ, Valentine CJ. Role of perinatal long-chain omega-3 fatty acids in cortical circuit maturation: Mechanisms and implications for psychopathology. *World J Psychiatry* 2015; **5**: 15–34.
- 43 Bloch MH, Qawasmi A. Omega-3 fatty acid supplementation for the treatment of children with attention-deficit/hyperactivity disorder symptomatology: systematic review and meta-analysis. *J Am Acad Child Adolesc Psychiatry* 2011; **50**: 991–1000.
- 44 Hawkey E, Nigg JT. Omega-3 fatty acid and ADHD: blood level analysis and meta-analytic extension of supplementation trials. *Clin Psychol Rev* 2014; **34**: 496–505.
- 45 Cooper RE, Tye C, Kuntsi J, Vassos E, Asherson P. Omega-3 polyunsaturated fatty acid supplementation and cognition: A systematic review and meta-analysis. *J Psychopharmacol (Oxford)* 2015; **29**: 753–763.

- 46 Atanasoski S, Notterpek L, Lee H-Y, Castagner F, Young P, Ehrenguber MU *et al.* The protooncogene Ski controls Schwann cell proliferation and myelination. *Neuron* 2004; **43**: 499–511.
- 47 Williams VC, Lucas J, Babcock MA, Gutmann DH, Korf B, Maria BL. Neurofibromatosis Type 1 Revisited. *Pediatrics* 2009; **123**: 124–133.
- 48 Hyman SL, Shores A, North KN. The nature and frequency of cognitive deficits in children with neurofibromatosis type 1. *Neurology* 2005; **65**: 1037–1044.
- 49 Najmabadi H, Motazacker MM, Garshasbi M, Kahrizi K, Tzschach A, Chen W *et al.* Homozygosity mapping in consanguineous families reveals extreme heterogeneity of non-syndromic autosomal recessive mental retardation and identifies 8 novel gene loci. *Hum Genet* 2007; **121**: 43–48.
- 50 Kuss AW, Garshasbi M, Kahrizi K, Tzschach A, Behjati F, Darvish H *et al.* Autosomal recessive mental retardation: homozygosity mapping identifies 27 single linkage intervals, at least 14 novel loci and several mutation hotspots. *Hum Genet* 2011; **129**: 141–148.
- 51 Numata S, Ye T, Hyde TM, Guitart-Navarro X, Tao R, Winer M *et al.* DNA methylation signatures in development and aging of the human prefrontal cortex. *Am J Hum Genet* 2012; **90**: 260–272.
- 52 Spiers H, Hannon E, Schalkwyk LC, Smith R, Wong CCY, O’Donovan MC *et al.* Methylomic trajectories across human fetal brain development. *Genome Res* 2015; **25**: 338–352.



- 53 McCowan T, Dhasarathy A, Carvelli L. The Epigenetic Mechanisms of Amphetamine. *J Addiction Prevention* 2015. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.681.8437&rep=rep1&type=pdf> (accessed 18 Feb2016).
- 54 Pidsley R, Viana J, Hannon E, Spiers H, Troakes C, Al-Saraj S *et al.* Methylomic profiling of human brain tissue supports a neurodevelopmental origin for schizophrenia. *Genome Biol* 2014; **15**. doi:10.1186/s13059-014-0483-2.
- 55 Dempster EL, Wong CCY, Lester KJ, Burrage J, Gregory AM, Mill J *et al.* Genome-wide Methylomic Analysis of Monozygotic Twins Discordant for Adolescent Depression. *Biol Psychiatry* 2014; **76**: 977–983.

## Tables and Figures

**Figure 1.** Trajectory analyses of the DAWBA ADHD scores yielded a 3 trajectory solution. Trajectories were estimated using k-means for longitudinal data, a non-parametric procedure, which classifies participants into homogenous subgroups following similar developmental patterns (trajectories). See method and results section as well as Supplementary sections 1.4 and 1.5 for statistical details.

**Figure 2.** Manhattan plot and bar graphs of methylome-wide results at birth. A) CpG chromosome positions are plotted against  $-\log_{10}$  p-values. The red line indicates FDR-corrected significance threshold. Results were derived using a general linear model between DNA methylation (407,462 probes at birth, cell type and batch-corrected) and ADHD trajectories, adjusting for sex. See method section for further statistical details. B-D) Bar graphs (mean  $\pm$  s.e.m.) of three DMPs associated with ADHD trajectories.

**Figure 3.** Gene network analysis using GeneMANIA. Darker circles represent genes associated with the 13 probes found to be related to ADHD trajectories in the methylome-wide analysis at birth. Lighted circles represent additional genes predicted by GeneMANIA based on genetic and physical interactions, shared pathways and protein domains as well as protein co-expression data. The gene network analysis demonstrates that, rather than being isolated (e.g. *FBXW5*), these genes clustered into a complex interconnected network. For more information on GeneMANIA methods, see Supplementary section 2.2.

**Table 1.** FDR-corrected probes that associate with ADHD trajectory. Estimates given for birth and age 7, ranked by birth p-values. ~~Probes that were also nominally significant at age 7, are marked in bold.~~ Chr, chromosome; StdB, standardized regression beta (negative values indicate hypomethylation in the high trajectory group); p, uncorrected p-value; q, FDR-corrected value; s.d., standard deviation.

CpG	Gene	Chr	Position	Birth					Age 7				
				StdB	p	q	mean±s.d. low trajectory group	mean±s.d. high trajectory group	StdB	p	q	mean±s.d. low trajectory group	mean±s.d. high trajectory group
cg24481594	<i>SKI</i>	1	2190850	-0.198	1.51E-08	0.006	<b>0.828 (0.026)</b>	<b>0.805 (0.034)</b>	0.005	0.876	<b>0.963</b>	<b>0.777 (0.025)</b>	<b>0.777 (0.024)</b>
cg03905179	<i>MAFK</i>	7	1582588	-0.182	2.00E-07	0.031	<b>0.761 (0.040)</b>	<b>0.727 (0.076)</b>	-0.002	0.951	<b>0.963</b>	<b>0.753 (0.034)</b>	<b>0.752 (0.028)</b>
cg27469152	<i>EPX</i>	17	56282313	-0.181	2.26E-07	0.031	<b>0.793 (0.024)</b>	<b>0.771 (0.032)</b>	-0.090	0.008	<b>0.104</b>	<b>0.772 (0.032)</b>	<b>0.758 (0.030)</b>
cg15096815	<i>JUN</i>	1	59249838	-0.178	3.53E-07	0.036	<b>0.105 (0.010)</b>	<b>0.097 (0.009)</b>	0.002	0.963	<b>0.963</b>	<b>0.104 (0.010)</b>	<b>0.104 (0.010)</b>
cg01324543	<i>CCDC30</i>	1	42999439	-0.174	7.21E-07	0.048	<b>0.871 (0.025)</b>	<b>0.851 (0.036)</b>	0.023	0.497	<b>0.881</b>	<b>0.876 (0.020)</b>	<b>0.877 (0.022)</b>
cg16290904	<i>PEX2</i>	8	77912348	0.173	7.35E-07	0.048	<b>0.055 (0.006)</b>	<b>0.061 (0.017)</b>	-0.077	0.022	<b>0.143</b>	<b>0.061 (0.011)</b>	<b>0.058 (0.013)</b>
cg26263766	<i>ZNF544</i> <i>ST3GAL</i>	19	58739734	0.173	8.72E-07	0.048	<b>0.064 (0.019)</b>	<b>0.079 (0.032)</b>	0.012	0.729	<b>0.948</b>	<b>0.063 (0.020)</b>	<b>0.064 (0.025)</b>
cg09989037	<i>3</i> <i>CDADC</i>	1	44300942	-0.172	9.46E-07	0.048	<b>0.443 (0.055)</b>	<b>0.402 (0.069)</b>	-0.056	0.099	<b>0.322</b>	<b>0.463 (0.044)</b>	<b>0.454 (0.048)</b>
cg18587973	<i>I</i>	13	49822535	0.170	1.20E-06	0.048	<b>0.076 (0.034)</b>	<b>0.108 (0.092)</b>	-0.015	0.656	<b>0.948</b>	<b>0.086 (0.029)</b>	<b>0.085 (0.039)</b>
cg22193912	<i>MAFG</i>	17	79881523	0.169	1.28E-06	0.048	<b>0.171 (0.056)</b>	<b>0.218 (0.082)</b>	-0.041	0.224	<b>0.570</b>	<b>0.311 (0.053)</b>	<b>0.303 (0.063)</b>
cg13714586	<i>FBXW5</i>	9	139838358	0.170	1.30E-06	0.048	<b>0.049 (0.005)</b>	<b>0.055 (0.026)</b>	-0.038	0.263	<b>0.570</b>	<b>0.050 (0.007)</b>	<b>0.049 (0.005)</b>
cg05653018	<i>ELF3</i>	1	201979533	0.169	1.43E-06	0.048	<b>0.799 (0.045)</b>	<b>0.833 (0.035)</b>	0.059	0.082	<b>0.322</b>	<b>0.841 (0.027)</b>	<b>0.847 (0.021)</b>
cg24843380	<i>ZNF454</i>	5	178367827	0.168	1.56E-06	0.049	<b>0.075 (0.010)</b>	<b>0.087 (0.053)</b>	0.021	0.542	<b>0.881</b>	<b>0.093 (0.021)</b>	<b>0.096 (0.015)</b>