1 DNA methylation analysis of amplicons from individuals exposed to maternal
2 tobacco use during pregnancy, and offspring conduct problems in childhood and
3 adolescence

4

5 Alexandra J. Noble¹, John F. Pearson², Alasdair D. Noble³ Joseph M. Boden⁴, L.
6 John Horwood⁴, Martin A. Kennedy², Amy J. Osborne¹

7

8

9 ¹School of Biological Sciences, University of Canterbury, Christchurch, New Zealand

10 ²Department of Pathology and Biomedical Sciences, University of Otago,
11 Christchurch, New Zealand

12 ³AgResearch, Lincoln Research Centre, Christchurch, New Zealand

13 ⁴Department of Psychological Medicine, University of Otago, Christchurch, New
14 Zealand

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30 Abstract

31 Maternal tobacco smoking during pregnancy is a large driver of health inequalities
 32 and a higher prevalence of conduct problem has been observed in exposed
 33 offspring. Further, maternal tobacco use during pregnancy can also alter offspring
 34 DNA methylation. However, currently, limited molecular evidence have been found to
 35 support this observation. Thus we aim to examine the association between maternal
 36 tobacco use in pregnancy and whether offspring Conduct problems is mediated by
 37 tobacco exposure-induced via DNA methylation differences. Understanding the
 38 etiology of the causal link will be crucial in the early identification and treatment of
 39 CP in children and adolescents. DNA was sourced from the Christchurch Health and
 40 Development Study, a longitudinal birth cohort studied for over 40 years in New
 41 Zealand. Bisulfite-based amplicon sequencing of 10 loci known to play a role in
 42 neurodevelopment, or with associations with CP phenotypes, was undertaken. We
 43 identified nominally significant differential DNA methylation at specific CpG sites in
 44 *CYP1A1*, *ASH2L* and *MEF2C* in individuals with Conduct problems who were
 45 exposed to tobacco *in utero*. We conclude that environmentally-induced DNA
 46 methylation differences could play a role in the observed link between maternal
 47 tobacco use during pregnancy and childhood/adolescent Conduct problems
 48 However, larger sample sizes are needed to produce an adequate amount of power
 49 to investigate this interaction further.

50 Introduction

51

52 Mothers who smoked tobacco during pregnancy have a higher prevalence of
 53 offspring developing a conduct disorder phenotype compared to mothers who did not
 54 smoke (Wakschlag et al. 1997). This association has been proven in several
 55 different cohort studies and the observation have remained following adjustment for
 56 various other confounding factors, for example, socio economic status, maternal
 57 age, substance abuse, parental anti-social personality, and maladaptive parenting
 58 (Wakschlag et al. 1997; Joelsson et al. 2016). However, there is limited molecular
 59 evidence to suggest a link between *in utero* tobacco exposure and offspring conduct
 60 disorder, thus a direct link between *in utero* tobacco exposure and Conduct problems
 61 (CP) remains elusive. Previously we conducted a pilot study assessing differential
 62 DNA methylation in a small cohort of individuals who were exposed to tobacco *in*
 63 *utero*, with sub-groups of individuals defined as having high conduct disorder scores
 64 (Noble et al. 2021). We found nominally significant DNA methylation changes with
 65 this interaction in several genes associated with neurodevelopment (cite epic in utero
 66 paper). Due to the limitations of using a small sample size combined with an array
 67 containing a large number of loci, results were underpowered, therefore observations
 68 were unable to reach genome wide significance. However, the biological relevance
 69 of these nominally significant CpG loci to the CP phenotype, combined with further
 70 research which has suggested an epigenetic link between *in utero* tobacco exposure
 71 and ADHD (Sengupta et al. 2017), implies that the link between DNA methylation
 72 and CP development in tobacco-exposed offspring should be investigated more fully.

73

74 Here, we will further pursue this hypothesis, by exploring differential methylation in
 75 genes that have known roles during *in utero* neurodevelopment and CP phenotypes,
 76 to understand whether DNA methylation may help explain the relationship between
 77 *in utero* tobacco exposure and development of CP in offspring. We applied a
 78 targeted approach via bisulfite-based amplicon sequencing (BSAS) of regions of
 79 genes involved in neurodevelopment. Amplicon sequencing has the ability to
 80 interrogate a region of the genome, therefore specifically targeting consecutive CpG
 81 sites in a row. We then assessed differential methylation in the DNA of participants
 82 from the Christchurch Health and Development Study (CHDS) whose mothers
 83 consumed tobacco during pregnancy, with high and low CP scores, and compared

84 this to controls who were not exposed. This approach allows us to specifically ask
85 whether DNA methylation at genes involved in neurodevelopment and CP
86 phenotypes are specifically differentially methylated in the DNA of offspring with CP,
87 who were exposed to tobacco *in utero*. A significant interaction here would provide
88 further support for a role for DNA methylation in the link between *in utero* exposure
89 and CP development, something which has so far proved elusive.

90

91

92 **Methods**

93 **Sample**

94 A sub-group of individuals from the CHDS were selected for this study (Table 1).
95 This longitudinal study originally included 97% of all the children (N= 1265) born in
96 the Christchurch, New Zealand urban region during a period in mid-1977 and has
97 been studied at 24 time points from birth to age 40 (N= 987 at age 30). All
98 participants were aged between 28-30 when blood samples and DNA was extracted.

99 For the subsets studied in this report, CHDS participants were chosen based on their
100 *in utero* tobacco exposure status, their adult smoking status, and their CP scores.
101 Group 1 consisted of individuals who were exposed *in utero* to tobacco smoke, and
102 never smokers at the time blood samples were taken (N= 32). Group 2 consisted of
103 individuals who were exposed *in utero* to tobacco smoke and were themselves
104 regular smokers at the time the blood was taken (N= 32). Group 3 consisted of
105 individuals who were not exposed to tobacco *in utero*, and never smokers at the time
106 blood was taken (N= 32). *In utero* tobacco exposure was defined as 10+ cigarettes
107 per day throughout pregnancy. Within each group of 32, 16 individuals were selected
108 with a 'high' score on a measure of childhood CP at age 7-9 years and 16 with a
109 'low' score. Severity of childhood CP was assessed using an instrument that
110 combined selected items from the Rutter and Conners child behaviour checklists
111 (Rutter M 1970; Conners 1970, 1969; Fergusson, Horwood, and Lloyd 1991) as
112 completed by parents and teachers at annual intervals from 7-9 years. Parental and
113 teacher reports were summed and averaged over the three years (Fergusson,
114 Horwood, and Ridder 2005) to derive a robust scale measure of the extent to which
115 the child exhibited conduct disordered/oppositional behaviours (mean (SD)=

50.1(7.9) ; range 41-97). For the purposes of this report a 'high' score was defined as falling into the top quartile of the score distribution (scores >53) and a 'low' score was defined as scores < 46.

Bisulfite-based amplicon sequencing

Bisulfite-based amplicon sequencing (BSAS) was carried out as described (Noble et al. 2020). DNA was extracted from whole blood samples using the Kingfisher Flex System (Thermo Scientific, Waltham, MA USA). DNA was quantified via nanodrop (Thermo Scientific, Waltham, MA USA). Bisulfite treatment was carried out using the EZ DNA Methylation-Gold kit (Zymo Research, Irvine, CA, USA) as per the manufacturer's instructions. DNA samples were then diluted to a final concentration of 100 ng/μl.

Amplicons for sequencing (Table 2 and Supplementary Table 1) were picked based upon several criteria: i) previously published differential DNA methylation in response to *in utero* tobacco smoking; ii) known associations with *in utero* brain development, and; iii) known associations with CP phenotypes. Primers were then designed to flank the CpG sites of interest, ~350 base pairs (bp) in total, or to amplify the promoter region of the gene if a specific CpG site was not known. Multiple pairs of primers were designed to amplify larger regions.

Bisulfite-converted DNA was amplified via PCR, using KAPA Taq HotStart DNA Polymerase (Sigma, Aldrich) under the following conditions: 95 °C for 10 min, 95 °C for 30 sec, 59 °C for 20 sec, 72 °C for 7 min, and held at 4 C° using the Mastercycler Nexus (Eppendorf, Australia). This was then cycled a total of 40 times. PCR products were purified with the Zymo DNA Clean & Concentrator Kit™ (Zymo Research, Irvine, CA, USA).

Following PCR, DNA was cleaned up with Agencourt® AMPure® XP beads (Beckman Coulter) and washed with 80% ethanol and allowed to air-dry. DNA was then eluted with 52.5 μl of 10 mM Tris pH 8.5 before being placed back into the magnetic stand. Once the supernatant had cleared, 50 μl was aliquoted for the experiment. DNA samples were quantified using the Quant-iT™ PicoGreen™ dsDNA Assay kit (Thermo Fisher) using the FLUORstar® Omega (BMG Labtech).

Samples were processed using the Illumina MiSeq™ 500 cycle Kit V2 and sequenced on the Illumina MiSeq™ system by Massey Genome Service (Palmerston North). Illumina MiSeq™ sequences were trimmed using SolexaQA++ software (Cox, Peterson, and Biggs 2010) and aligned to FASTA bisulfite converted reference sequences using the package Bowtie2 (version 2.3.4.3) Each individual read was then aligned to all reference sequences using the methylation-specific package Bismark (Krueger and Andrews 2011).

Statistics

Differential DNA methylation was assessed using the package edgeR (Chen et al.). Coverage level was set to greater or equal to “8” across unmethylated and methylated counts, as recommended by (Chen et al.). Two models were used – the first was a bivariate model, to assess differences between the *in utero* exposed to tobacco compared to the non-exposed control group (model 1).

$$Y \sim U + AS + e$$

The second was a multiple regression to assess the interaction term *in utero* maternal smoke exposure and offspring conduct problem score (high or low, model 2).

$$Y \sim U + C + AS + U:C + e$$

Where, Y is defined as the methylation M ratio, U is the exposed/unexposed *in utero* to maternal smoking, C is conduct problem score with high conduct problem score < 53 and low conduct problem core < 46, AS adult smoker/non-smoker and e is the unexplained variation or error tem.

This model was fitted with both ANOVA parameters and with contrasts between *in utero* exposure groups (exposed – non-exposed) within CP score levels. Top tables were constructed using the topTags function in edgeR, Log fold change, average log counts per million, and in some cases F statistic and were calculated and nominal significance was given for $P < 0.05$, these were then corrected using FDR. Scatter plots with the inclusion of confidence intervals were constructed from log transformed normalised methylated and unmethylated counts.

Results

Here we assessed DNA methylation within 10 separate genes (Table 2). DNA sequence data for 15 amplicons from these 10 genes (Supplementary Table 1) was generated, comprising a total of 280 CpG sites. These CpG sites included a combination of sites previously identified as differentially methylated, as well as amplification of all CpGs within the promoter region of genes associated with *in utero* neurodevelopment and CP phenotypes (Table 2). Differential methylation across these CpG sites was calculated to address whether any were specifically differentially methylated in individuals with CP, in response to *in utero* tobacco exposure.

Quantification of DNA methylation at previously reported CpG sites in response to *in utero* exposure to tobacco

Initially, we attempted to validate in our cohort (age ~28-30 years) five CpG sites which have been previously reported to be differentially methylated in the DNA of cord blood from newborns, and whole blood from children and adolescents (ages newborn to 17) in response to *in utero* tobacco exposure (Table 1). Data were partitioned into those individuals exposed *in utero*, and those who were not (model 1), to assess whether or not BSAS could detect previously reported CpG sites (Table 3).

AHRR (cg05575921) displayed a 3.1% decrease in DNA methylation between exposed and non-exposed individuals, at a nominal P value of 0.02. This site has been previously identified as hypomethylated in adult tobacco smokers, as well as in postnatal cord blood samples between *in utero* tobacco-exposed and non-exposed individuals. The probe cg05549655 in the gene *CYP1A1* displayed a 5.19% increase in DNA methylation in the *in utero* exposed group, however, this site did not reach nominal statistical significance in our cohort. Cg09935388 and cg09662411 in *GFI1* were unable to be replicated as differentially methylated between the exposed and the non-exposed groups (no significant change in β values). Both CpG sites show

hypomethylation, supporting previous observations of differential methylation within this gene. *CNTNAP2* (cg2594950) was similarly unable to be validated in our cohort using the method BSAS.

Differentially methylated CpGs under the interaction of *in utero* tobacco exposure and CP

Differential methylation dependent on both *in utero* exposure and CP score was found at 10 loci in six genes at nominal significance level, none were significant after correcting for false discovery rate (Table 4).

Of these CpG sites, five of the 10 CpG sites were found in the following genes: *CYP1A1*, *GFI1*, *ASH2L* and *GRIN2b*. Differential methylation between *in utero* exposed and non-exposed associated with for high conduct scores. No nominal significance from the interaction observed in association with low conduct scores. The top three CpG sites with nominal significance under the interaction are displayed in Figure 1. Here, differential methylation is found in response to high CP score and no differences are seen between the exposed and non-exposed low CP groups (Figure 1).

Discussion

In utero tobacco exposure is known to alter DNA methylation at the genome-wide level in offspring (Joubert, Haberg, et al. 2012; Joubert, Felix, Yousefi, Bakulski, Just, Breton, Reese, Markunas, Richmond, Xu, et al. 2016) (Joubert Bonnie et al. 2012; Richmond et al. 2014). The later-life implications of these tobacco-induced DNA methylation changes are unclear, however, an association between *in utero* tobacco exposure and CP has previously been observed (Sengupta et al. 2017). Given the complex etiology of CP phenotypes (Acosta, Arcos-Burgos, and Muenke 2004; Beaver et al. 2007; Salvatore and Dick 2018) and the vast array of socioeconomic variables associated with tobacco use (Lantz et al. 1998), proving a

causal link between maternal smoking and offspring CP is inherently challenging. Previously we quantified tobacco-induced DNA methylation changes that associate with CP phenotypes in offspring exposed to tobacco *in utero* (via maternal smoking) using the Illumina EPIC array, with results indicating that methylation was altered at genes that may have roles in neurodevelopment and CP phenotypes. However, due to a combination of a comparatively small sample size relative to the number of loci on the array, only nominal significance was observed. The data suggested a role for DNA methylation in the link between exposure and CP, so here we chose a panel of genes with known roles in these things, to see if methylation is changed at these loci too. Due to our small sample size we need to try an alternative approach. We've shown that BSAS previously is very good for targeting differential methylated so here we use BSAS at some targeted genes to see if we can detect differential methylation specific to the interaction between high CP score and exposure?

Validation of previously identified differentially methylated CpG from *in utero* tobacco exposure

First, we asked whether differentially methylated CpGs that have been previously associated with *in utero* tobacco exposure were supported by this cohort. Here, we present validation of differential methylation of a CpG site within the gene *AHRR* (cg05575921). *AHRR* is a well-defined tobacco smoking gene, which is consistently represented in tobacco methylation data. *AHRR* has previously been found to be differentially methylated in response to *in utero* tobacco exposure (Richmond, Simpkin, Woodward, Gaunt, Lyttleton, McArdle, Ring, Smith, Timpson, Tilling, Davey Smith, et al. 2015; Joubert, Håberg, Nilsen, Wang, Vollset, Murphy, Huang, Hoyo, Midttun, Cupul-Uicab, et al. 2012; de Vocht et al. 2015). Four other CpG sites investigated here due to previous association with *in utero* tobacco exposure were not differentially methylated in our data. However, the direction of methylation change was supported at all five sites investigated (Tehraniifar et al. 2018; Rauschert et al. 2019; Rotroff et al. 2016). We suggest that further investigation in a larger cohort may lead to nominal significance at the sites in *CYP1A1*, *CNTNAP2*, and *GFI1*.

Identification of *in utero* exposure-related differentially methylated CpG sites that are specific to individuals with high CP scores

Epidemiological data suggests that there is an increased association between *in utero* tobacco exposure and behavioural disorder in children and adolescents (Carter et al. 2008; Mick et al. 2002). Thus, here, we investigated DNA methylation changes induced by *in utero* tobacco exposure as a potential molecular mechanism of dysfunction that could link the phenotypic trait of CP to maternal tobacco use during pregnancy. We therefore analysed DNA methylation patterns within our gene panel in response to *in utero* tobacco exposure and its interaction with CP status. A total of 10 CpG sites in seven genes were found to display nominal significance in DNA methylation in response to *in utero* tobacco exposure and CP in this cohort (Table 4).

In the 10 CpG sites we identified under the interaction, *CYP1A1* showed greater magnitude differential methylation in high CP scores (exposed *in utero* vs. non-exposed with high CPS), with reduced reversed or no evidence of differential methylation at the same sites with low CP score. This indicates that within the observed nominal methylation changes the interaction was being driven in the high CP score group. One gene (*ASH2L*), contained three nominally significantly differentially methylated CpG sites, and *CYP1A1* and *MEF2C* both had two.

CYP1A1 (Cytochrome P450 family 1 subfamily A member 1) is a well-established marker for *in utero* tobacco smoke exposure (Richmond et al. 2018; Lee Ken et al. 2015; Richmond et al. 2014; Tehranifar et al. 2018). Neither of the two sites we observed under the interaction have probes at these locations on the Illumina array system, thus emphasising a benefit of amplicon sequencing compared to an array-based method. Variant differences in *CYP1A1* have previously been associated with child behavioural problems at age 2, from prenatal maternal environmental tobacco smoke (Hsieh et al. 2010). This highlights the need for this gene to be further investigated for its role in the development of conduct problems following *in utero* tobacco exposure.

Three CpG sites from the gene *ASH2L* (ASH2 like histone lysine methyltransferase complex subunit) showed in consistent levels of differential methylation in response to *in utero* tobacco exposure and CP, with two displaying hyper- and one hypomethylation. *ASH2L* has been found to interact with *MEF2C* (Myocyte enhancer

factor 2C) to mediate changes in histone 3 lysine 4 trimethylation (H3K4me3) (Jung et al.). Here, we detected nominal significance at two CpG sites within *MEF2C* (chr5, 88179596 and 88179541). Both of these sites were associated with a greater level of hypomethylation in participants who were exposed to tobacco *in utero* with high CP scores in this cohort, although not at the FDR significance level. *MEF2C* plays a role in neural crest formation during development, where tissue-specific inactivation of the gene results in embryonic lethality (Verzi et al. 2007). Further, *MEF2* interacts with oxytocin, which is affiliated with prosocial behaviours (Kosfeld et al. 2005; Zak, Stanton, and Ahmadi 2007). Alterations to oxytocin have been shown to change the morphology of neurons via *MEF2A* (Meyer et al. 2018; Meyer et al. 2020). Functional roles of the gene in relation to early neuronal development still remain unclear, however it is thought to play a crucial role (Harrington et al. 2016). Recent research in animal models suggests that nicotine-dependent induction of the *ASH2L* and *MEF2C* complex during development induces alterations that could lead to fundamental changes in the brain (Jung et al.).

While we cannot assert causality, our targeted approach shows that *in utero* tobacco exposure may be altering methylation at CpG sites associated with neural phenotypes which persist into adulthood and are then associated with increased risk of high CP.

Conclusion

Here we have presented preliminary data to suggest that the association between maternal tobacco use during pregnancy and the development of CP in children and adolescents may in part be mediated by altered DNA methylation, induced by *in utero* tobacco exposure during development, at genes that have roles in *in utero* brain development and CP phenotypes. We acknowledge the limitations of this study described above, however, the data presented here are suggestive of a role for DNA methylation in the link between *in utero* tobacco exposure and offspring CP. Our findings should stimulate further study using larger sample sizes.

Abbreviations

CP Conduct problems

330 CHDS Christchurch health and development study
 331 BSAS Bisulfite based amplicon sequencing
 332 SIDS Sudden infant death syndrome
 333 ADHD Attention-deficit hyperactivity disorder
 334 DOHaD Developmental origins of human health and disease
 335 GFI1 Growth Factor Independent one transcriptional repressor
 336 CPS Conduct disorder score
 337 AHRR Aryl hydrocarbon receptor repressor
 338 ASH2L ASH2 like histone lysine methyltransferase complex subunit
 339 BDNF Brain-derived neurotrophic factor,
 340 CNTNAP2 Contactin associated protein 2
 341 CYP1A1 Cytochrome P450 Family 1 Subfamily A Member 1
 342 DUSP6 Dual specificity phosphatase 6
 343 GRIN2b Glutamate Ionotropic Receptor NMDA Type Subunit 2B
 344 MEF2C Myocyte enhancer factor 2C
 345 PRDM8 PR/SET Domain 8
 346 FC Fold change
 347 CPM Counts per million
 348 FDR False discovery rate

349

350 **Funding**

351 Funding for this study came from the Maurice and Phyllis Paykel Trust. CHDS was
 352 funded by the Health Research Council of New Zealand (Programme Grant 16/600).
 353 CMRF supplied funding for the manuscript to be written.

354 **Availability of data**

355 Upon request.

356 **Contributions**

357 AJN-molecular lab work, data analysis, and major contributor to manuscript. JFP-
 358 study design, data analysis, and major contributor to manuscript. ADN- data

359 analysis. JMB and LJH study design, provided DNA samples via CHDS. MAK- study
360 design and over view. AJO- study design, molecular lab work, major contributor to
361 manuscript and source of funding. All authors read and approved the final
362 manuscript.

363 **Acknowledgements**

364 Not applicable

365 **Ethics declarations**

366 All aspects of the study were approved by the Southern Health and Disability Ethics
367 Committee, under application number CTB/04/11/234/AM10 "Collection of DNA in
368 the Christchurch Health and Development Study".

369 **Consent for publication**

370 Not applicable

371 **Competing interests**

372 The authors declare that they have no competing interests.

References

- Acosta, Maria Teresa, Mauricio Arcos-Burgos, and Maximilian Muenke. 2004. 'Attention deficit/hyperactivity disorder (ADHD): Complex phenotype, simple genotype?', *Genetics in Medicine*, 6: 1-15.
- Beaver, Kevin M., John Paul Wright, Matt DeLisi, Anthony Walsh, Michael G. Vaughn, Danielle Boisvert, and Jamie Vaske. 2007. 'A gene $\times$ gene interaction between DRD2 and DRD4 is associated with conduct disorder and antisocial behavior in males', *Behavioral and Brain Functions*, 3: 30.
- Carter, Sarnia, Janis Paterson, Wanzhen Gao, and Leon Iusitini. 2008. 'Maternal smoking during pregnancy and behaviour problems in a birth cohort of 2-year-old Pacific children in New Zealand', *Early Human Development*, 84: 59-66.
- Chen, Yunshun, Bhupinder Pal, Jane E. Visvader, and Gordon K. Smyth. 2017. 'Differential methylation analysis of reduced representation bisulfite sequencing experiments using edgeR', *F1000Res*, 6: 2055-55.
- Conners, C. Keith. 1969. 'A teacher rating scale for use in drug studies with children', *The American Journal of Psychiatry*, 126: 884-88.
- . 1970. 'Symptom patterns in hyperkinetic, neurotic, and normal children', *Child development*, 41: 667-82.
- Cox, Murray P., Daniel A. Peterson, and Patrick J. Biggs. 2010. 'SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data', *BMC Bioinformatics*, 11: 485.
- de Vocht, Frank, Andrew J Simpkin, Rebecca C Richmond, Caroline Relton, and Kate Tilling. 2015. 'Assessment of offspring DNA methylation across the lifecourse associated with prenatal maternal smoking using Bayesian Mixture Modelling', *International journal of environmental research and public health*, 12: 14461-76.
- Demontis, Ditte, Raymond K Walters, Joanna Martin, Manuel Mattheisen, Thomas Damm Als, Esben Agerbo, Rich Belliveau, Jonas Bybjerg-Grauholm, Marie Bækved-Hansen, and Felecia Cerrato. 2017. 'Discovery of the first genome-wide significant risk loci for ADHD', *bioRxiv*: 145581.
- Fergusson, D. M., L. J. Horwood, and M. Lloyd. 1991. 'Confirmatory factor models of attention deficit and conduct disorder', *Journal of child psychology and psychiatry, and allied disciplines*, 32: 257-74.
- Fergusson, D. M., L. J. Horwood, and E. M. Ridder. 2005. 'Show me the child at seven: the consequences of conduct problems in childhood for psychosocial functioning in adulthood', *Journal of child psychology and psychiatry, and allied disciplines*, 46: 837-49.
- Harrington, Adam J., Aram Raissi, Kacey Rajkovich, Stefano Berto, Jaswinder Kumar, Gemma Molinaro, Jonathan Raduazzo, Yuhong Guo, Kris Loerwald, Genevieve Konopka, Kimberly M. Huber, and Christopher W. Cowan. 2016. 'MEF2C regulates cortical inhibitory and excitatory synapses and behaviors relevant to neurodevelopmental disorders', *eLife*, 5: e20059.
- Hsieh, Chia-Jung, Suh-Fang Jeng, Yi-Ning Su, Hua-Fang Liao, Wu-Shiun Hsieh, Kuen-Yuh Wu, and Pau-Chung Chen. 2010. 'CYP1A1 Modifies the Effect of Maternal Exposure to Environmental Tobacco Smoke on Child Behavior', *Nicotine & Tobacco Research*, 12: 1108-17.
- Jiao, S. S., L. L. Shen, C. Zhu, X. L. Bu, Y. H. Liu, C. H. Liu, X. Q. Yao, L. L. Zhang, H. D. Zhou, D. G. Walker, J. Tan, J. Götz, X. F. Zhou, and Y. J. Wang. 'Brain-derived neurotrophic factor protects against tau-related neurodegeneration of Alzheimer's disease'.
- Joelsson, Petteri, Roshan Chudal, Ardesheer Talati, Auli Suominen, Alan S. Brown, and Andre Sourander. 2016. 'Prenatal smoking exposure and neuropsychiatric comorbidity of ADHD: a finnish nationwide population-based cohort study', *BMC Psychiatry*, 16: 306.
- Joubert, B. R., J. F. Felix, P. Yousefi, K. M. Bakulski, A. C. Just, C. Breton, S. E. Reese, C. A. Markunas, R. C. Richmond, C. J. Xu, L. K. Kupers, S. S. Oh, C. Hoyo, O. Gruzjeva, C. Soderhall, L. A. Salas, N. Baiz, H. Zhang, J. Lepeule, C. Ruiz, S. Ligthart, T. Wang, J. A. Taylor, L. Duijts, G. C. Sharp, S.

423 A. Jankipersadsing, R. M. Nilsen, A. Vaez, M. D. Fallin, D. Hu, A. A. Litonjua, B. F. Fuemmeler,
424 K. Huen, J. Kere, I. Kull, M. C. Munthe-Kaas, U. Gehring, M. Bustamante, M. J. Saurel-
425 Coubizolles, B. M. Quraishi, J. Ren, J. Tost, J. R. Gonzalez, M. J. Peters, S. E. Haberg, Z. Xu, J. B.
426 van Meurs, T. R. Gaunt, M. Kerkhof, E. Corpeleijn, A. P. Feinberg, C. Eng, A. A. Baccarelli, S. E.
427 Benjamin Neelon, A. Bradman, S. K. Merid, A. Bergstrom, Z. Herceg, H. Hernandez-Vargas, B.
428 Brunekreef, M. Pinart, B. Heude, S. Ewart, J. Yao, N. Lemonnier, O. H. Franco, M. C. Wu, A.
429 Hofman, W. McArdle, P. Van der Vlies, F. Falahi, M. W. Gillman, L. F. Barcellos, A. Kumar, M.
430 Wickman, S. Guerra, M. A. Charles, J. Holloway, C. Auffray, H. W. Tiemeier, G. D. Smith, D.
431 Postma, M. F. Hivert, B. Eskenazi, M. Vrijheid, H. Arshad, J. M. Anto, A. Dehghan, W.
432 Karmaus, I. Annesi-Maesano, J. Sunyer, A. Ghantous, G. Pershagen, N. Holland, S. K. Murphy,
433 D. L. DeMeo, E. G. Burchard, C. Ladd-Acosta, H. Snieder, W. Nystad, G. H. Koppelman, C. L.
434 Relton, V. W. Jaddoe, A. Wilcox, E. Melen, and S. J. London. 2016. 'DNA Methylation in
435 Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis',
436 *Am J Hum Genet*, 98: 680-96.

437 Joubert, B. R., S. E. Håberg, R. M. Nilsen, X. Wang, S. E. Vollset, S. K. Murphy, Z. Huang, C. Hoyo, Ø
438 Midttun, L. A. Cupul-Uicab, P. M. Ueland, M. C. Wu, W. Nystad, D. A. Bell, S. D. Peddada, and
439 S. J. London. 2012. '450K epigenome-wide scan identifies differential DNA methylation in
440 newborns related to maternal smoking during pregnancy', *Environ Health Perspect*, 120:
441 1425-31.

442 Joubert, B. R., S. E. Haberg, R. M. Nilsen, X. Wang, S. E. Vollset, S. K. Murphy, Z. Huang, C. Hoyo, O.
443 Midttun, L. A. Cupul-Uicab, P. M. Ueland, M. C. Wu, W. Nystad, D. A. Bell, S. D. Peddada, and
444 S. J. London. 2012. '450K epigenome-wide scan identifies differential DNA methylation in
445 newborns related to maternal smoking during pregnancy', *Environ Health Perspect*, 120:
446 1425-31.

447 Joubert, Bonnie R, Janine F Felix, Paul Yousefi, Kelly M Bakulski, Allan C Just, Carrie Breton, Sarah E
448 Reese, Christina A Markunas, Rebecca C Richmond, and Cheng-Jian Xu. 2016. 'DNA
449 methylation in newborns and maternal smoking in pregnancy: genome-wide consortium
450 meta-analysis', *The American Journal of Human Genetics*, 98: 680-96.

451 Joubert, Bonnie R, Siri E Håberg, Roy M Nilsen, Xuting Wang, Stein E Vollset, Susan K Murphy,
452 Zhiqing Huang, Cathrine Hoyo, Øivind Midttun, and Lea A Cupul-Uicab. 2012. '450K
453 epigenome-wide scan identifies differential DNA methylation in newborns related to
454 maternal smoking during pregnancy', *Environmental health perspectives*, 120: 1425.

455 Joubert Bonnie R., E. Håberg Siri, M. Nilsen Roy, Xuting Wang, E. Vollset Stein, K. Murphy Susan,
456 Zhiqing Huang, Cathrine Hoyo, Øivind Midttun, A. Cupul-Uicab Lea, M. Ueland Per, C. Wu
457 Michael, Wenche Nystad, A. Bell Douglas, D. Peddada Shyamal, and J. London Stephanie.
458 2012. '450K Epigenome-Wide Scan Identifies Differential DNA Methylation in Newborns
459 Related to Maternal Smoking during Pregnancy', *Environmental Health Perspectives*, 120:
460 1425-31.

461 Jung, Yonwoo, Lawrence S. Hsieh, Angela M. Lee, Zhifeng Zhou, Daniel Coman, Christopher J. Heath,
462 Fahmeed Hyder, Yann S. Mineur, Qiaoping Yuan, David Goldman, Angelique Bordey, and
463 Marina R. Picciotto. 'An epigenetic mechanism mediates developmental nicotine effects on
464 neuronal structure and behavior'.

465 Kosfeld, Michael, Markus Heinrichs, Paul J. Zak, Urs Fischbacher, and Ernst Fehr. 2005. 'Oxytocin
466 increases trust in humans', *Nature*, 435: 673-76.

467 Krueger, Felix, and Simon R. Andrews. 2011. 'Bismark: a flexible aligner and methylation caller for
468 Bisulfite-Seq applications', *Bioinformatics*, 27: 1571-72.

469 Lantz, Paula M., James S. House, James M. Lepkowski, David R. Williams, Richard P. Mero, and
470 Jieming Chen. 1998. 'Socioeconomic Factors, Health Behaviors, and Mortality Results From a
471 Nationally Representative Prospective Study of US Adults', *JAMA*, 279: 1703-08.

472 Lee Ken, W. K., Rebecca Richmond, Pingzhao Hu, Leon French, Jean Shin, Celine Bourdon, Eva
473 Reischl, Melanie Waldenberger, Sonja Zeilinger, Tom Gaunt, Wendy McArdle, Susan Ring,

474 Geoff Woodward, Luigi Bouchard, Daniel Gaudet, Davey Smith George, Caroline Relton,
475 Tomas Paus, and Zdenka Pausova. 2015. 'Prenatal Exposure to Maternal Cigarette Smoking
476 and DNA Methylation: Epigenome-Wide Association in a Discovery Sample of Adolescents
477 and Replication in an Independent Cohort at Birth through 17 Years of Age', *Environmental*
478 *Health Perspectives*, 123: 193-99.

479 Li, Liang, Xiangbin Ruan, Chang Wen, Pan Chen, Wei Liu, Liyuan Zhu, Pan Xiang, Xiaoling Zhang,
480 Qunfang Wei, Lin Hou, Bin Yin, Jiangang Yuan, Boqin Qiang, Pengcheng Shu, and Xiaozhong
481 Peng.

482 Meyer, M., I. Berger, J. Winter, and B. Jurek. 2018. 'Oxytocin alters the morphology of hypothalamic
483 neurons via the transcription factor myocyte enhancer factor 2A (MEF-2A)', *Mol Cell*
484 *Endocrinol*, 477: 156-62.

485 Meyer, Magdalena, Kerstin Kuffner, Julia Winter, Inga D. Neumann, Christian H. Wetzel, and
486 Benjamin Jurek. 2020. 'Myocyte Enhancer Factor 2A (MEF2A) Defines Oxytocin-Induced
487 Morphological Effects and Regulates Mitochondrial Function in Neurons', *International*
488 *Journal of Molecular Sciences*, 21: 2200.

489 Mick, Eric, Joseph Biederman, Stephen V. Faraone, Julie Sayer, and Seth Kleinman. 2002. 'Case-
490 Control Study of Attention-Deficit Hyperactivity Disorder and Maternal Smoking, Alcohol
491 Use, and Drug Use During Pregnancy', *Journal of the American Academy of Child &*
492 *Adolescent Psychiatry*, 41: 378-85.

493 Noble, Alexandra J, John F Pearson, Joseph M Boden, L. John Horwood, Martin A Kennedy, and Amy J
494 Osborne. 2021. 'Hypomethylation in FASTKD1 detected in the association between in utero
495 tobacco exposure and conduct problem in a New Zealand longitudinal study':
496 2021.04.08.438710.

497 Noble, Alexandra, John Pearson, Joseph Boden, John Horwood, Neil Gemmell, Martin Kennedy, and
498 Amy Osborne. 2020. 'A validation of Illumina EPIC array system with bisulfite-based
499 amplicon sequencing', *bioRxiv*: 2020.05.25.115428.

500 Rauschert, Sebastian, Phillip E. Melton, Graham Burdge, Jeffrey M. Craig, Keith M. Godfrey, Joanna
501 D. Holbrook, Karen Lillycrop, Trevor A. Mori, Lawrence J. Beilin, Wendy H. Oddy, Craig
502 Pennell, and Rae-Chi Huang. 2019. 'Maternal Smoking During Pregnancy Induces Persistent
503 Epigenetic Changes Into Adolescence, Independent of Postnatal Smoke Exposure and Is
504 Associated With Cardiometabolic Risk', *Frontiers in genetics*, 10: 770-70.

505 Richmond, R. C., A. J. Simpkin, G. Woodward, T. R. Gaunt, O. Lyttleton, W. L. McArdle, S. M. Ring,
506 Adac Smith, N. J. Timpson, K. Tilling, G. D. Smith, and C. L. Relton. 2015. 'Prenatal exposure
507 to maternal smoking and offspring DNA methylation across the lifecourse: findings from the
508 Avon Longitudinal Study of Parents and Children (ALSPAC)', *Human molecular genetics*, 24:
509 2201-17.

510 Richmond, Rebecca C, Matthew Suderman, Ryan Langdon, Caroline L Relton, and George Davey
511 Smith. 2018. 'DNA methylation as a marker for prenatal smoke exposure in adults',
512 *International Journal of Epidemiology*, 47: 1120-30.

513 Richmond, Rebecca C., Andrew J. Simpkin, Geoff Woodward, Tom R. Gaunt, Oliver Lyttleton, Wendy
514 L. McArdle, Susan M. Ring, Andrew D. A. C. Smith, Nicholas J. Timpson, Kate Tilling, George
515 Davey Smith, and Caroline L. Relton. 2015. 'Prenatal exposure to maternal smoking and
516 offspring DNA methylation across the lifecourse: findings from the Avon Longitudinal Study
517 of Parents and Children (ALSPAC)', *Human molecular genetics*, 24: 2201-17.

518 Richmond, Rebecca C., Andrew J. Simpkin, Geoff Woodward, Tom R. Gaunt, Oliver Lyttleton, Wendy
519 L. McArdle, Susan M. Ring, Andrew D.A.C. Smith, Nicholas J. Timpson, Kate Tilling, George
520 Davey Smith, and Caroline L. Relton. 2014. 'Prenatal exposure to maternal smoking and
521 offspring DNA methylation across the lifecourse: findings from the Avon Longitudinal Study
522 of Parents and Children (ALSPAC)', *Human Molecular Genetics*, 24: 2201-17.

523 Riva, Valentina, Marco Battaglia, Maria Nobile, Francesca Cattaneo, Claudio Lazazzera, Sara
524 Mascheretti, Roberto Giorda, Chantal Mérette, Claudia Émond, and Michel Maziade. 2015.

'GRIN2B predicts attention problems among disadvantaged children', *European child & adolescent psychiatry*, 24: 827-36.

Rotroff, Daniel M, Bonnie R Joubert, Skylar W Marvel, Siri E Häberg, Michael C Wu, Roy M Nilsen, Per M Ueland, Wenche Nystad, Stephanie J London, and Alison Motsinger-Reif. 2016. 'Maternal smoking impacts key biological pathways in newborns through epigenetic modification in Utero', *BMC genomics*, 17: 976.

Rutter M, Tizard J, Whitmore K. 1970. 'Education, Health and Behaviour', *London: Longmans*.

Rzehak, Peter, Richard Saffery, Eva Reischl, Marcela Covic, Simone Wahl, Veit Grote, Annick Xhonneux, Jean-Paul Langhendries, Natalia Ferre, Ricardo Closa-Monasterolo, Elvira Verduci, Enrica Riva, Piotr Socha, Dariusz Gruszfeld, Berthold Koletzko, and group European Childhood Obesity Trial Study. 2016. 'Maternal Smoking during Pregnancy and DNA-Methylation in Children at Age 5.5 Years: Epigenome-Wide-Analysis in the European Childhood Obesity Project (CHOP)-Study', *PLOS ONE*, 11: e0155554.

Salvatore, Jessica E., and Danielle M. Dick. 2018. 'Genetic influences on conduct disorder', *Neuroscience & Biobehavioral Reviews*, 91: 91-101.

Sengupta, Sarojini M., Alicia K. Smith, Natalie Grizenko, and Ridha Joobar. 2017. 'Locus-specific DNA methylation changes and phenotypic variability in children with attention-deficit hyperactivity disorder', *Psychiatry Research*, 256: 298-304.

Skogstrand, Kristin, Christian Munch Hagen, Nis Borbye-Lorenzen, Michael Christiansen, Jonas Bybjerg-Grauholm, Marie Bækvad-Hansen, Thomas Werge, Anders Børghlum, Ole Mors, Merethe Nordentoft, Preben Bo Mortensen, and David Michael Hougaard. 'Reduced neonatal brain-derived neurotrophic factor is associated with autism spectrum disorders'.

Suter, Melissa, Adi Abramovici, Lori Showalter, Min Hu, Cynthia Do Shope, Michael Varner, and Kjersti Aagaard-Tillery. 2010. 'In utero tobacco exposure epigenetically modifies placental CYP1A1 expression', *Metabolism-Clinical and Experimental*, 59: 1481-90.

Tehranifar, Parisa, Hui-Chen Wu, Jasmine A. McDonald, Farzana Jasmine, Regina M. Santella, Irina Gurvich, Julie D. Flom, and Mary Beth Terry. 2018. 'Maternal cigarette smoking during pregnancy and offspring DNA methylation in midlife', *Epigenetics*, 13: 129-34.

van Otterdijk, Sanne D, Alexandra M Binder, and Karin B Michels. 2017. 'Locus-specific DNA methylation in the placenta is associated with levels of pro-inflammatory proteins in cord blood and they are both independently affected by maternal smoking during pregnancy', *Epigenetics*, 12: 875-85.

Verzi, Michael P., Pooja Agarwal, Courtney Brown, David J. McCulley, John J. Schwarz, and Brian L. Black. 2007. 'The Transcription Factor MEF2C Is Required for Craniofacial Development', *Developmental Cell*, 12: 645-52.

Wakschlag, L. S., B. B. Lahey, R. Loeber, S. M. Green, R. A. Gordon, and B. L. Leventhal. 1997. 'Maternal smoking during pregnancy and the risk of conduct disorder in boys', *Arch Gen Psychiatry*, 54: 670-6.

Zak, Paul J., Angela A. Stanton, and Sheila Ahmadi. 2007. 'Oxytocin increases generosity in humans', *PLOS ONE*, 2: e1128-e28.

Table 1 - CHDS subsets selected for analysis. The range of conduct problem scores in each category is indicated in brackets. A score of 53 or more is the top quartile for CP, a score of 60 or more the top decile for CP.

	Group 1	Group 2	Group 3
	Exposed <i>in utero</i> and never smokers	Exposed <i>in utero</i> and a regular smoker	Not exposed <i>in utero</i> and never smokers
	n= 32	n= 32	n= 32
Sex			
Male	69%	72%	60%
Female	31%	28%	40%
Tobacco smoking status at the time of blood collection			
Never	100%	0%	100%
Regular	0%	100%	0%
Conduct problem score (CPS)			
Low CPS (<46)	n= 16 (42-46)	n= 16 (42-46)	n= 16 (41-43)
High CPS (>53)	n= 16 (53-75)	n= 16 (60-85)	n= 16 (53-68)

591 **Table 2 - Genes selected to investigate the link between in utero tobacco exposure**
592 **and CP.**

Gene	Function	Significance
AHRR (Joubert, Håberg, Nilsen, Wang, Vollset, Murphy, Huang, Hoyo, Middtun, and Cupul-Uicab 2012; Richmond, Simpkin, Woodward, Gaunt, Lyttleton, McArdle, Ring, Smith, Timpson, Tilling, Smith, et al. 2015; de Vocht et al. 2015; van Otterdijk, Binder, and Michels 2017; Rotroff et al. 2016)	Mediates toxicity of dioxin (found in cigarette smoke)	Hypomethylated in tobacco smokers and their offspring
ASH2L (Li et al.)	Histone lysine methyltransferase	Associated with schizophrenia
BDNF (Skogstrand et al. ; Jiao et al.)	Nerve growth factor	Promotes neuronal survival. Implicated in neurodegenerative disease
CNTNAP2 (Joubert, Felix, Yousefi, Bakulski, Just, Breton, Reese, Markunas, Richmond, and Xu 2016; Richmond, Simpkin, Woodward, Gaunt, Lyttleton, McArdle, Ring, Smith, Timpson, Tilling, Smith, et al. 2015; Rzehak et al. 2016)	Neurexin family – functions in vertebrate nervous system	Implicated in schizophrenia, autism, ADHD, intellectual disability. Hypomethylated in offspring of maternal smoking
CYP1A1 (Suter et al. 2010; Rotroff et al. 2016; Joubert, Håberg, Nilsen, Wang, Vollset, Murphy, Huang, Hoyo, Middtun, and Cupul-Uicab 2012; Richmond, Simpkin, Woodward, Gaunt, Lyttleton, McArdle, Ring, Smith, Timpson, Tilling, Smith, et al. 2015;	Monooxygenase – expression is induced by hydrocarbons found in cigarette smoke	Hypomethylated in offspring of maternal smoking

de Vocht et al. 2015; van Otterdijk, Binder, and Michels 2017)	DUSP6	Protein phosphatase, cellular proliferation and differentiation	Regulates neurotransmitter homeostasis
(Demontis et al. 2017)			
GFI1 (Rotroff et al. 2016; Joubert, Håberg, Nilsen, Wang, Vollset, Murphy, Huang, Hoyo, Midttun, and Cupul- Uicab 2012; van Otterdijk, Binder, and Michels 2017)		Zinc finger protein - transcriptional repressor	Part of a complex that controls histone modifications and gene silencing. Hypermethylated in offspring of maternal smoking
GRIN2B (Riva et al. 2015)		Glutamate receptor – expressed early in the brain and is required for normal brain development	Mutations associated with autism, ADHD, schizophrenia
MEF2C (Demontis et al. 2017)		MEF2C is associated with hippocampal-dependent learning and memory	MEF2C is crucial for normal neuronal development. Associated with ADHD
PRDM8 (Joubert, Felix, Yousefi, Bakulski, Just, Breton, Reese, Markunas, Richmond, and Xu 2016)		Histone methyltransferase - Controls expression of genes involved in neural development and neuronal differentiation	Hypomethylated in offspring of maternal smoking

593

594

595

596

597

598

599

600

601

602 *Table 3* - Previously reported CpG sites showing differential DNA methylation in
603 response to *in utero* tobacco exposure, and their average methylation values in
604 individuals from this cohort.

Gene	Illumina ID	Exposed <i>in</i>	Non-	β difference	P value
------	-------------	-------------------	------	--------------------	---------

		<i>utero</i> β value	exposed <i>in</i> <i>utero</i> β value		
<i>AHRR</i>	cg05575921	72.287	75.448	-3.161	0.022
<i>CNTNAP2</i>	cg2594950	3.845	3.860	-0.014	0.991
<i>CYP1A1</i>	cg05549655	26.894	21.699	5.195	0.425
<i>GFI1</i>	cg09935388	75.151	75.330	-0.582	0.055
<i>GFI1</i>	cg09662411	95.837	97.400	-1.583	0.274

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626 *Table 4* - CpG sites where differential methylation between conduct problem scores
627 differs with *in utero* exposure at $P < 0.05$. Log Fold Change (FC) and P values
628 (unadjusted) from log ratio tests for the effect on normalized methylation ratios of: (1)
629 interaction between *in utero* exposure and Conduct Problem Score, (2) *In utero*
630 exposed - non-exposed contrast within Low CPS and (3) within High CPS

631 participants. Loci with nominally significant ($P < 0.05$) interaction shown, all FDR P
 632 values > 0.05 .

633

Gene	CpG location	Interaction ⁽¹⁾		Low CPS ⁽²⁾		High CPS ⁽³⁾	
		<i>Log FC</i>	<i>P value</i>	<i>Log FC</i>	<i>P value</i>	<i>Log FC</i>	<i>P value</i>
<i>CYP1A1</i>	Chr15, 75019290	-2.013	0.010	0.344	0.493	-1.669	0.005
<i>GFI1</i>	Chr1, 92947705	-0.957	0.011	0.002	0.992	-0.955	0.001
<i>ASH2L</i>	Chr8, 37962878	1.257	0.024	-0.447	0.253	0.811	0.042
<i>MEF2C</i>	Chr5, 88179596	-1.679	0.040	0.678	0.174	-1.000	0.122
<i>DUSP6</i>	Chr12, 89746588	-1.444	0.041	0.864	0.107	-0.580	0.204
<i>ASH2L</i>	Chr8, 37962657	-0.199	0.042	0.052	0.455	-0.147	0.033
<i>CYP1A1</i>	Chr15, 75019127	-1.221	0.045	0.403	0.319	-0.819	0.072
<i>ASH2L</i>	Chr8, 37962901	1.250	0.046	-0.561	0.205	0.688	0.121
<i>GRIN2b</i>	Chr12, 14133359	2.711	0.048	0.121	0.903	2.832	0.004
<i>MEF2C</i>	Chr5, 88179541	-1.336	0.050	0.615	0.139	-0.720	0.190

634