Epigenetic signatures associated with the observed link between maternal tobacco use during pregnancy, and offspring conduct problems in childhood and adolescence

Alexandra J. Noble[1], John F. Pearson[2], Alasdair D. Noble[3] Joseph M. Boden[4], L. John Horwood[4], Martin A. Kennedy[2], Amy J. Osborne[1]

[1]School of Biological Sciences, University of Canterbury, Christchurch, New Zealand

[2]Department of Pathology and Biomedical Sciences, University of Otago, Christchurch, New Zealand

[3]AgResearch, Lincoln Research Centre, Christchurch, New Zealand

[4]Department of Psychological Medicine, University of Otago, Christchurch, New Zealand

Abstract

Background

Metastable epialleles (MEs) are loci at which epigenetic regulation is established during development and subsequently maintained throughout life. Consequently, individuals can have the same genetic sequence, yet their epigenetic regulation of the underlying sequence can vary. MEs can be independent of genetic variation and may be induced by environmental exposures. Maternal tobacco smoking during pregnancy can alter offspring DNA methylation, hence there is potential for MEs to be induced during development in response to maternal tobacco smoking during pregnancy. Furthermore, associations between maternal tobacco use during pregnancy and conduct problems (CP) in offspring exposed to tobacco smoke *in utero*, have been observed. However, currently, we do not know what molecular mechanism may link these associations.

Results

We investigated the observed link between maternal tobacco use during pregnancy and CP outcomes in exposed offspring. Individuals who were exposed to tobacco *in utero* via maternal smoking, and unexposed individuals, both with and without CP, were selected from the Christchurch Health and Development Study (CHDS), a longitudinal birth cohort studied for over 40 years in New Zealand. Bisulfite-based amplicon sequencing (BSAS) was used to investigate DNA methylation differences and potential MEs between the groups at high risk loci. We identified nominally significant differential DNA methylation at specific CpG sites in individuals with CP who were exposed to tobacco *in utero*. This differential methylation appears to be specific to *in utero* tobacco exposure, and interacts with CP. Given its association with the *in utero* environment we hypothesise that these methylation changes may represent MEs.

Conclusion

We conclude that environmentally-induced DNA methylation differences may play a role in the observed link between maternal smoking during pregnancy and childhood/adolescent CP, potentially via the generation of MEs. Larger sample sizes and a genome-wide approach are required to investigate this association further.

Introduction

The use of tobacco during pregnancy is one of the leading causes of perinatal compromise for developing offspring, and one of the most preventable [1]. For example, low birth weight [2], congenital heart anomalies [3], asthma/respiratory illness [4, 5], and sudden infant death syndrome (SIDS)[6] are all associated with maternal tobacco use during pregnancy, the rate of which remains relatively high in New Zealand (18.4% [7]), despite declining tobacco use rates overall [8].

While immediate perinatal compromise in infants due to maternal smoking is well documented, the long term effects into later childhood, adolescence and adulthood are not understood. There is increasing evidence of linkages between maternal tobacco use in pregnancy and later risks of mental health and related adjustment problems in childhood and adolescence. In particular, there is evidence that maternal smoking during pregnancy is associated with increased risks of conduct disorders and antisocial behaviours in offspring [9] [10-12]. This association is not explained by post-natal environment [13]. Further associations have been identified between maternal tobacco use during pregnancy and the increased risk of cardiometabolic disease [14], and the development of attention-deficit hyperactivity disorder (ADHD) [15]. Also affected are offspring neurodevelopment and behaviour, suggesting that poor behavioural adjustment (often termed 'conduct problems', CP) can be considered a consequence of maternal smoking during pregnancy [9]. While these traits in themselves can be linked to other societal risk factors such as low socioeconomic status and early-life adversity [16], their association with maternal tobacco use during pregnancy is intriguing. Understanding the link between exposures such as tobacco use during pregnancy and the association with conduct disorder is crucial to further our understanding the paradigm of the developmental origins of human health and disease (DOHaD) [17].

Recent research has demonstrated links between prenatal tobacco exposure and specific DNA methylation patterns of newborn offspring [18-21]. Tobacco-induced DNA methylation changes can persist into adolescence [22] [21, 23] with potential for these unexplained marks to be inherited by future generation of offspring of exposed individuals [24]. Further, meta-analyses of multiple CpG sites in the gene, *GFI1* (Growth Factor Independent one transcriptional repressor) were found to be

differentially methylated in adult offspring in response to being exposed to tobacco *in utero*, at multiple sites within the gene [25]. However, these studies are limited in their scope - they provide evidence for differential DNA methylation induced in both children and adults by tobacco exposure *in utero*, but do not relate these DNA methylation changes to a phenotype that is associated with *in utero* tobacco exposure. Thus, while limited preliminary work has been carried out, in which three loci which indicated modest DNA methylation changes in response to maternal smoking during pregnancy and CP phenotypes [26], the etiology of this link has not been fully explored. One potential mechanism is that differential DNA methylation caused during the *in utero* time period is playing a role later in life of the affected offspring via the *in utero* generation of metastable epialleles (MEs). Evidence at this stage has largely come from animal studies, where *in utero* exposures cause the development of MEs [27-29]. Potentially these *in utero* exposures can generate permanent epigenetic changes to the genome [30] that may contribute to an individual's phenotype later in life [29-32]

Thus, given: i) the fact that maternal tobacco smoking during pregnancy is linked to offspring conduct problems during early childhood and adolescence, and; ii) that maternal tobacco use during pregnancy can affect DNA methylation of offspring through to adolescence and adulthood, and; iii) that *in utero* exposures can create permanent epigenetic changes that can affect health in later life, here we hypothesise that DNA methylation is altered at genes involved in *in utero* brain development, and in those that associate with CP phenotypes, in the adult offspring of individuals who were exposed to tobacco *in utero*.

To test this hypothesis, we quantified DNA methylation at a suite of genes with known roles in *in utero* neurodevelopment and CP phenotypes, to assess whether DNA methylation may be implicated in the interaction between maternal tobacco use during pregnancy and the development of conduct problems in offspring.

We applied a targeted approach via bisulfite-based amplicon sequencing (BSAS) of each gene in our panel, to interrogate differential methylation in the DNA of participants from the Christchurch Health and Development Study (CHDS) whose

mothers consumed tobacco during pregnancy (offspring DNA collected from whole blood at ~28 years of age). We partitioned the analyses into multiple subsets (Table 1) to show that differential methylation in this panel of genes is most fully explained by the specific analysis of individuals exposed to tobacco *in utero* with CP, vs. those that were not exposed *in utero*. Further, we show that these specific methylation changes, in response to *in utero* tobacco exposure are independent of adult smoking status, implying that the changes we observe in methylation status of these genes are permanent epigenetic changes induced during *in utero* exposure, indicating their potential as metastable epialleles.

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 Exposed *in utero* and never smokers	Group 2 Exposed *in utero* and a regular smoker	Group 3 Not *exposed in utero* 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)

Results

Here we assessed DNA methylation in 10 separate genes. DNA sequence data for 15 amplicons (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 fitted to four separate statistical models (see Methods) and calculated to address whether any were specifically differentially methylated in individuals with CP, in response to *in utero* tobacco exposure.

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

Gene	Function	Significance
AHRR [33-37]	Mediates toxicity of dioxin (found in cigarette smoke)	Hypomethylated in tobacco smokers and their offspring
ASH2L [38]	Histone lysine methyltransferase	Associated with schizophrenia
BDNF [39, 40]	Nerve growth factor	Promotes neuronal survival. Implicated in neurodegenerative disease
CNTNAP2 [34, 41, 42]	Neurexin family – functions in vertebrate nervous system	Implicated in schizophrenia, autism, ADHD, intellectual disability. Hypomethylated in offspring of maternal smoking
CYP1A1 [33-37, 43]	Monooxygenase – expression is induced by hydrocarbons found in cigarette smoke	Hypomethylated in offspring of maternal smoking
DUSP6 [44]	Protein phosphatase, cellular proliferation and differentiation	Regulates neurotransmitter homeostasis
GFI1 [33, 36, 37]	Zinc finger protein - transcriptional repressor	Part of a complex that controls histone modifications and gene silencing. Hypermethylated in offspring of maternal smoking
GRIN2B [45]	Glutamate receptor – expressed early in the brain and is required for normal brain development	Mutations associated with autism, ADHD, schizophrenia
MEF2C [44]	MEF2C is associated with hippocampal-dependent learning and memory	MEF2C is crucial for normal neuronal development. Associated with ADHD
PRDM8 [41]	Histone methyltransferase - Controls expression of genes involved in neural development and neuronal differentiation	Hypomethylated in offspring of maternal smoking

Quantification of DNA methylation at previously reported CpG sites in response to *in utero* exposure to tobacco – Model 1

Initially, we attempted to validate in our cohort (age ~28-30 years) 5 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 3). Data were partitioned into those individuals exposed *in utero*, and those who were not (Model 1, Methods).

Table 3 - Previously reported CpG sites showing differential DNA methylation in response to in utero tobacco exposure, and their average methylation values in individuals from this cohort (Model 1).

Gene	Illumina ID	Exposed *in utero* β value	Non-exposed *in utero* β value	β difference	P Value
AHRR	cg05575921	72.287	75.448	-3.161	0.022
CNTNAP2	cg2594950	3.8457	3.8600	-0.014	0.991
CYP1A1	cg05549655	26.894	21.699	5.195	0.425
GFI1	cg09935388	75.151	75.330	-0.582	0.055
GFI1	cg09662411	95.837	97.400	-1.583	0.274

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

Differentially methylated CpGs by *in utero* tobacco exposure status

Data were partitioned according to *in utero* exposure status only (exposed vs. unexposed) using Model 1 (Methods). Of the 10 genes (encompassing a total of 280 CpG sites) selected for BSAS, 6 genes showed nominally significant differential methylation between *in utero*-exposed and non-exposed controls, across 22 different CpG sites that resided in those regions: *AHRR2*, *GRIN2b*, *GFI1*, *BDNF*, *ASH2L* and *DUSP6* (Table 4). The remaining genes, *CNTNAP2*, *MEF2C*, *SLC9A9* and *CYP1A1*, showed no differential methylation across the region in response to *in utero* tobacco exposure alone.

Table 4 - Top CpG sites found to be nominally significantly differentially methylated (unadjusted P < 0.05) in response to in utero *tobacco exposure (Model 1). Asterisk, *, indicates CpG sites in genes identified as differentially methylated in response to adult smoking status (Supplementary Table 3). Abbreviations: FC, fold change; CPM, counts per million; FDR, FDR-corrected P value.*

Gene	Illumina ID, CpG site location	Log FC	Average Log CPM	P Value	FDR
**AHRR*	Chr5, 373398	-0.369	12.699	0.0009	0.187
**GFI1*	Chr1, 92946546	-0.588	12.284	0.002	0.192
**BDNF*	Chr11, 27743856	-1.323	10.237	0.004	0.192
**GRIN2b*	Chr12, 14133243	2.100	10.113	0.004	0.192
**GFI1*	Chr1, 92947559	-0.507	9.0675	0.005	0.192
**GFI1*	Chr1, 92947752	-0.433	9.8441	0.006	0.192
GRIN2b	Chr12, 14133359	1.789	10.523	0.007	0.192
**GFI1*	Chr1, 92946452	-0.374	12.211	0.008	0.192
**GIF1*	Chr1, 92946429	-0.558	12.163	0.009	0.192
BDNF	Chr11, 27743594	-0.773	11.078	0.010	0.192
GFI1	Chr1, 92946514	-0.477	10.053	0.011	0.200
**BDNF*	Chr11, 27743729	-1.266	8.550	0.016	0.262
GFI1	Chr1, 92946568	-0.339	12.218	0.019	0.284
**AHRR*	cg05575921	-0.270	12.687	0.022	0.291
AHRR	Chr5, 373355	-0.228	12.749	0.022	0.291
**GIF1*	Chr1, 92946418	-0.512	12.160	0.030	0.365
DUSP6	Chr12, 89746641	-0.635	10.060	0.033	0.371
GFI1	Chr1, 92946434	-0.314	12.193	0.035	0.371
GFI1	Chr1, 92946340	-0.368	12.360	0.047	0.413
**GFI1*	Chr1, 92946132	-0.420	12.295	0.048	0.413
DUSP6	Chr12, 89746479	0.813	10.285	0.049	0.413
ASH2L	Chr8, 37962720	0.692	11.626	0.049	0.413

The top log fold changes (2.1 and 1.78) in differential methylation between *in utero* exposed individuals verses non-exposed individuals both come from CpG sites in *GRIN2b* (Chr12: 14133243 and Chr12: 14133359), followed by two further larger log fold changes in two CpG sites in *BDNF* (Chr11:, 27743857 and Chr11:, 27743730).

A MA plot of the log average difference between individuals exposed *in utero*, and non-exposed individuals (Figure 1,Table 4) indicates those sites with the highest log fold changes, and demonstrates the direction of change in methylation of the 22 nominally significantly differentially methylated CpGs ($P < 0.05$); 4 are hypermethylated (pink) and 18 are hypomethylated (cyan).

Data were then partitioned based upon CP and non-CP status (Model 2). A total of nine CpG sites were found to be differentially methylated (Supplementary Table 2). Four CpG sites were independent of *in utero* exposure, and the remaining five were also differentially methylated in response to *in utero* exposure.

Differential methylation in response to adult smoking status (Model 3)

Smoking in adulthood was assessed for its confounding effect on DNA methylation across the amplicons of genes of interest. The data was partitioned into those individuals who were tobacco smokers in adulthood, and those who were never smokers. When differential methylation was calculated in smokers vs. never smokers, 26 out of 280 CpG sites in total were identified as significantly differentially methylated (nominal $P < 0.05$, Supplementary Table 3). These loci were in general hypomethylated, consistent with the literature for the same or near sites with the only hypermethylated site located in the GRIN2b promoter. There were a total of 12 CpG sites that were also found to be differentially methylated in response to both of the univariate analyses of adult smoking status and *in utero* exposure (* in Table 4). 14 CpG sites were found solely to be differentially methylated in response to adult smoking status and 10 CpG sites differentially methylated only in response to *in utero* exposure.

Differentially methylated CpGs dependent on both *in utero* tobacco exposure and CP (Model 4)

Differential methylation dependent on both in utero exposure and conduct problem score was found at 10 loci in 6 genes at nominal significance level, none were significant after correcting for false discovery rate (Table 5).

For 9 out of the 10 sites (all except *DUSP6*) there was more differential methylation between *in utero* exposure states for high conduct scores than low conduct scores. CpG sites within *CYP1A1*, *GFI1*, ASH2L, *and GRIN2b* were nominally significantly differentially methylated $P < 0.05$ in the DNA of *in utero*-exposed individuals with high CP score (Model 4, Table 5).

Table 5 - CpG sites where differential methylation between conduct problem scores differs with in utero exposure at P<0.05. Log Fold Change (FC) and P values (unadjusted) from log ratio tests for the effect on normalized methylation ratios of: (1) interaction between *in utero* exposure and Conduct Problem Score, (2) In utero exposed - non-exposed contrast within Low CPS and (3) within High CPS participants. Loci with nominally significant (P<0.05) interaction shown, all FDR P values > 0.05.

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

Negative log fold change values for the significantly differentially methylated sites within the high conduct problem score group correspond to hypomethylation within the exposed group, whereas positive log fold changes correspond to hypermethylation in the *in utero* exposed group as the log normalized ratios are

negative, three examples are shown in Figure 2. These associations were not detected when data was partitioned and analysed to assess the impact of CP only on DNA methylation (Model 2, Supplementary Table 2).

Discussion

In utero tobacco exposure is known to alter DNA methylation at the genome-wide level in offspring [18, 19] [20, 21]. 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 [26]. Given the complex etiology of conduct disorder phenotypes [46-48] and the vast array of socioeconomic variables associated with tobacco use [49], proving a causal link between maternal smoking and offspring CP is inherently challenging. However, here we provide preliminary evidence of tobacco-induced DNA methylation changes that associate with conduct problem phenotypes in offspring exposed to tobacco *in utero* (via maternal smoking), within a panel of genes that have known roles in *in utero* brain development and conduct problem phenotypes.

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 [22, 35, 50]. We find that this particular CpG within *AHRR* remains differentially methylated in response to *in utero* tobacco exposure in our adult cohort at age ~28-30 (Table 3). However, in this study, differential methylation at this CpG site was also explained by adult smoking status (Supplementary Table 3). 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 [37, 51, 52]. We suggest that further investigation in a larger cohort may lead to nominal significance at the sites in *CYP1A1*, *CNTNAP2,* and *GFI1*.

Given this finding, we further propose that the differential methylation at *AHRR* identified here, in adults, that specifically associates with *in utero* tobacco exposure, may represent a metastable epiallele; the stable integration of differential methylation during fetal development, in response to *in utero* maternal tobacco exposure at specific CpG sites, which remain stable through adulthood. Given the nature of methylation at this site, it is possible that differential methylation at *AHRR* represents a permanent epigenetic change in offspring. The status of metastable epialleles at *AHRR* is supported by the finding that *AHRR* is differentially methylated in cord blood of newborns as well as blood samples from childhood and adolescence [14, 21, 51], and through to adulthood as identified in this study. Further exploration of *AHRR* methylation status in a longitudinal cohort with access to DNA from multiple timepoints would allow a more robust assessment of this question.

Identification of *in utero* exposure-related differentially methylated CpGs

Next, we compared all individuals exposed to tobacco *in utero*, to individuals not exposed to tobacco *in utero,* and we identified a large number of differentially methylated CpG sites (22, Table 4). Of these, 20 represent novel sites, which are not target CpG sites in the Illumina EPIC array system. Thus, these sites were unable to be previously identified as differentially methylated in response to *in utero* tobacco exposure. This highlights the benefits of the BSAS method, which enables estimates of differential methylation of all CpGs within a particular amplicon [53]. Further, the novel CpG sites we identify here are all in relatively close proximity to one another, suggesting that these sites may represent differentially methylation regions. Differentially methylation regions have important roles in regulating gene expression, thus potentially leading to changes in phenotype that could have detrimental health outcomes [54]. The identification here of differential methylation across multiple CpG sites within genes, rather than just one specific CpG site, illustrates that a gene regulation process directly related to *in utero* tobacco exposure may be associated

with CP. None of the 22 CpG sites identified as being differentially methylated in response to *in utero* tobacco exposure remained significantly differentially methylated after FDR correction, which was expected because of small sample size. However, while our data are nominally significant, it does suggest that *in utero* tobacco exposure may be affecting DNA methylation at CpG sites within genes that had no overlap with adult smoking status in this study. Nevertheless, we note that the CpG sites that were identified in both the *in utero* exposure model (Model 1) and the adult smoking model (Model 3) cannot yet be ruled out as being differentially methylated as a consequence of *in utero* exposure. This is because the identification of differential methylation in both models here may be due to either the *in utero* environment, or the effect of adult smoking on DNA methylation.

Some changes in response to adult smoking status and *in utero* exposure unable to be differentiated

We assessed what effect adult smoking status was having on differential methylation within well studied genes, in order to determine differential methylation patterns specifically impacted by *in utero* tobacco exposure. The premise here was that CpG sites which were not identified in response to adult smoking status would indicate that the differential methylation we identify was much more likely to be induced during development, and not a by-product of adult smoking status. When the data were partitioned based on adult smoking status (Model 3), we identified 26 differentially methylated CpGs (Supplementary Table 3). Of these, 12 CpG sites overlapped with the CpG sites found to be differentially methylated when the data was partitioned based upon *in utero* tobacco exposure status (Table 4, Model 1). This indicates that differential DNA methylation identified in genes which overlap between Models 1 and 3 may be explained by adult smoking status, or *in utero* exposure. However, the remaining ten CpG sites observed in our panel of genes are not explained by adult smoking status. This implies that differential methylation at these CpG sites is explained more fully by *in utero* tobacco exposure, and provides confidence that the differential methylation we observe within these genes is more likely due to the *in utero* environment, than to adult smoking. We cannot ignore the fact that adult tobacco smoking may still be playing a role in differential DNA

methylation at these sites, but it does not appear to explain the variation in methylation we observe at the sites investigated in this study as fully as the *in utero* environment.

Differential methylation within *AHRR* (cg05575921) was explained by adult smoking status in this study (Supplementary Table 3). This was an expected result as this site is one of the most pronounced and associated sites found to be differentially methylated in tobacco smoking [55, 56]. This site, however, also showed nominal significance in response to *in utero* maternal tobacco exposure. The reason for this may be due to the study design; this study was limited by sample size, and as such, distinguishing between adult smoking status and *in utero* tobacco exposure is difficult; CpG sites which could show differences in response to both variables may have skewed the results when independently assessing them within this relatively small sample.

Tobacco smoking is known to greatly affect DNA methylation, and because the DNA samples used in this study are from individuals who were between 28 and 30 years old, adult smoking is closer temporally than *in utero* exposure. Thus we hypothesise that the data used in the *in utero* exposure model could be expected to be confounded to some extent by adult smoking status, meaning that, in these data, differential methylation at certain sites can be explained independently by both *in utero* tobacco exposure and adult smoking status. Further investigations in larger cohorts, preferably at the genome-wide level, are required. To further rule out adulthood smoking status as an explanatory factor in the differential methylation we observe within our panel of brain development and CP genes, this study should be expanded to include an additional group of individuals that were not exposed to tobacco *in utero*, but are smokers as adults.

Identification of *in utero* exposure-related differentially methylated CpGs that are specific to individuals with conduct problem (CP)

An overwhelming amount of epidemiological data has shown an increased association between *in utero* tobacco exposure and behavioural disorder in children and adolescents [57, 58]. 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 conduct problem 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 six genes were found to display nominal significance in DNA methylation in response to *in utero* tobacco exposure and CP in this cohort (Table 5, Model 4). Differential methylation at none of these CpG sites could not be explained by adult smoking status.

The candidate genes explored here have been shown to be differentially methylated in response to both adult smoking and *in utero* smoking. We observed that when *in utero* smoking and conduct problem score were considered together, differential methylation attributed to *in utero* exposure was significantly different in those with high conduct problem scores than in those with low conduct problem scores. In the 10 loci we identified with interactive differential methylation, all but the loci in *DUSP6* showed greater magnitude differential methylation in high conduct problem 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 conduct problem score . While we cannot assert causality, our results are consistent with *in utero* smoking altering methylation at loci associated with neural phenotypes which persist into adulthood and are then associated with increased risk of conduct problems.

Our results indicate that *in utero* tobacco exposure is associated with a greater level of *MEF2C* hypomethylation in participants who were exposed to tobacco *in utero* with CP in this cohort, although not at the FDR significance level. We identified differential methylation at two CpG sites that are located next to each other within the gene *MEF2C* (chr5, 88179596 and 88179541). *MEF2C* (Myocyte enhancer factor 2C) is a transcription factor which regulates gene expression for development and maintenance in a variety of tissues [59]. It has been shown to play an important role in the brain [60-64], particularly, in neuronal migration and neuronal differentiation [65-67]. More so, *MEF2C* in plays a role in neural crest formation during development, where tissue-specific inactivation of the gene results in embryonic lethality [68]. Further, *MEF2* interacts with oxytocin, which is affiliated with prosocial behaviours [69, 70]. Alterations to oxytocin have been shown to change the morphology of neurons via *MEF2A* [71, 72]. Functional roles of the gene in relation

to early neuronal development still remain unclear, however it is thought to play crucial role [73].

Three CpG sites from the gene *ASH2L* (ASH2 like histone lysine methyltransferase complex subunit) were also found to display differential methylation in response to *in utero* tobacco exposure and CP. *ASH2L* has been found to interact with *MEF2C* to mediate changes in histone 3 lysine 4 trimethylation (H3K4me3 [74]). 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. These consist of dendritic branching and hypersensitive passive avoidance behaviour which is a consequence of developmental nicotine exposure [74]. Our findings support this hypothesis by providing molecular evidence of CpG site alterations in these genes via *in utero* tobacco exposure in individuals with high CP score.

However, these sites were not differentially methylated in response to CP vs non-CP alone (Model 2, Supplementary Table 2), suggesting that DNA methylation changes in developmental genes are both induced by maternal tobacco use during pregnancy, and involved in pathways in development of CP phenotypes. Further, the persistence of specific *in utero* related DNA methylation changes into adulthood, as identified here, indicates that methylation differences at these genes may be induced during development and stable over the life course, potentially indicating the presence of metastable epialleles within these genes.

Although adult smoking status was the only other variable able to control for in this study we cannot account for many other confounding variables when assessing *in utero* effects. Other genetic factors such as sex and ethnicity, as well as social interactions of economic status are all confounding variables. Ideally, this study should be repeated in a larger cohort to further for assess these confounding variables on *in utero* tobacco exposure.

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 a larger sample size, preferably with analysis at the genome-wide level.

Abbreviations

CP Conduct problems

CHDS Christchurch health and development study

BSAS Bisulfite based amplicon sequencing

SIDS Sudden infant death syndrome

ADHD Attention-deficit hyperactivity disorder

DOHaD Developmental origins of human health and disease

GFI1 Growth Factor Independent one transcriptional repressor

CPS Conduct disorder score

AHRR Aryl hydrocarbon receptor repressor

ASH2L ASH2 like histone lysine methyltransferase complex subunit

BDNF Brain-derived neurotrophic factor,

CNTNAP2 Contactin associated protein 2

CYP1A1 Cytochrome P450 Family 1 Subfamily A Member 1

DUSP6 Dual specificity phosphatase 6

GRIN2b Glutamate Ionotropic Receptor NMDA Type Subunit 2B

MEF2C Myocyte enhancer factor 2C

PRDM8 PR/SET Domain 8

FC Fold change

CPM Counts per million

FDR False discovery rate

Methods

Sample

A sub-group of individuals from the CHDS were selected for this study (Table 1). This longitudinal originally included 97% of all the children (n = 1265) born in the Christchurch, New Zealand urban region during a period in mid-1977 and has been studied at 24 time points from birth to age 40 (n = 987 at age 30). All participants were aged between 28-30 when blood samples and DNA was extracted. For the subsets studied in this report, CHDS participants were chosen based on their *in utero* tobacco exposure status, their adult smoking status, and their CP scores. Group 1 consisted of individuals who were exposed *in utero* to tobacco smoke, and never smokers at the time blood samples were taken (n=32). Group 2 consisted of individuals who were exposed *in utero* to tobacco smoke and were themselves regular smokers at the time the blood was taken (n=32). Group 3 consisted of individuals who were not exposed to tobacco *in utero*, and never smokers at the time blood was taken (n=32). *In utero* tobacco exposure was defined as 10+ cigarettes per day throughout pregnancy. Within each group, 16 individuals were selected with a 'high' score on a measure of childhood conduct problems at age 7-9 years and 16 with a 'low' score. Severity of childhood conduct problems was assessed using an instrument that combined selected items from the Rutter and Conners child behaviour checklists [75-78] as completed by parents and teachers at annual intervals from 7-9 years. Parental and teacher reports were summed and averaged over the three years [79] to derive a robust scale measure of the extent to which 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 [53]. 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) to 100 ng/µl. 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 (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 ~350bp of 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 FLUROstar® 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 SolexQA++ software [80] 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 [81].

Statistics

Differential DNA methylation was assessed using the package edgeR [82]. MA plots were carried out for clustering based on group and for the top differentially methylated sites via edgeR. The following models were fitted to the data:

Univariate regression:

Model 1 - effect of in utero *tobacco exposure on DNA methylation (Table 3 and Table 4)*

$$\mathrm{Y} \sim \mathrm{U} + e$$

Model 2 - effect of conduct problem on DNA methylation (Supplementary Table 2)

$$\mathrm{Y} \sim \mathrm{C} + e$$

Model 3 - effect of adult smoking on DNA methylation (fitted on Exposed participants only, Supplementary Table 3)

$$\mathrm{Y} \sim \mathrm{AS} + e$$

Multiple Regression:

Model 4 - effect of in utero *tobacco exposure and conduct problem on DNA methylation (Table 5)*

$$\mathrm{Y} \sim \mathrm{U} + \mathrm{C} + \mathrm{U{:}C} + e$$

Where:

Y = methylation M ratio

U = Exposed/Unexposed *in utero* to maternal smoking

C = Conduct problem/Non-conduct problem

e ~ N(0,s)

AS = Adult smoking/Non-adult smoking

U:C is interaction term between U and C

Model 4 was fitted with both anova parameters and with contrasts between *in utero* exposure groups (exposed – non-exposed) within conduct problem 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. Co-variates such as ethnicity and sex were not corrected for. Box plots were constructed from log transformed normalized methylated and unmethylated counts. .

Funding

Funding for this study came from the Maurice and Phyllis Paykel Trust.

CHDS was funded by the Health Research Council of New Zealand (Programme Grant 16/600) and the Canterbury Medical Research Foundation (New Zealand).

Availability of data

Upon request.

Contributions

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

Acknowledgements

Not applicable

Ethics declarations

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

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

References

1. Castles, A., et al., *Effects of smoking during pregnancy. Five meta-analyses.* Am J Prev Med, 1999. **16**(3): p. 208-15.
2. Blatt, K., et al., *Association of reported trimester-specific smoking cessation with fetal growth restriction.* Obstet Gynecol, 2015. **125**(6): p. 1452-9.
3. Baardman, M.E., et al., *Combined adverse effects of maternal smoking and high body mass index on heart development in offspring: evidence for interaction?* Heart, 2012. **98**(6): p. 474-479.
4. Hayatbakhsh, M.R., et al., *Maternal smoking during and after pregnancy and lung function in early adulthood: a prospective study.* Thorax, 2009. **64**(9): p. 810-814.
5. Beyer, D., H. Mitfessel, and A. Gillissen, *Maternal smoking promotes chronic obstructive lung disease in the offspring as adults.* European Journal of Medical Research, 2009. **14**(4): p. 27.
6. Kandall, S.R. and J. Gaines, *Maternal substance use and subsequent Sudden Infant Death Syndrome (SIDS) in offspring.* Neurotoxicology and Teratology, 1991. **13**(2): p. 235-240.
7. Andrews, A., et al., *Smoking prevalence trends: An analysis of smoking at pregnancy registration and at discharge from a midwife Lead Maternity Carer, 2008 to 2010.* New Zealand College of Midwives Journal, 2014. **49**.
8. Health, M.o., *Annual Update of Key Results 2017/18*, N.Z.H. Survey, Editor. 2018, In, Ministry of Health: Wellington, Ministry of Health.
9. Fergusson, D.M., L.J. Woodward, and L.J. Horwood, *Maternal Smoking During Pregnancy and Psychiatric Adjustment in Late Adolescence.* Archives of General Psychiatry, 1998. **55**(8): p. 721-727.
10. Pratt, T.C., J.M. McGloin, and N.E. Fearn, *Maternal Cigarette Smoking During Pregnancy and Criminal/Deviant Behavior: A Meta-Analysis.* International Journal of Offender Therapy and Comparative Criminology, 2006. **50**(6): p. 672-690.
11. Brennan, P.A., E.R. Grekin, and S.A. Mednick, *Maternal smoking during pregnancy and adult male criminal outcomes.* Arch Gen Psychiatry, 1999. **56**(3): p. 215-9.
12. Roza, S.J., et al., *Maternal smoking during pregnancy and child behaviour problems: the Generation R Study.* International Journal of Epidemiology, 2008. **38**(3): p. 680-689.
13. Gaysina, D., et al., *Maternal smoking during pregnancy and offspring conduct problems: evidence from 3 independent genetically sensitive research designs.* JAMA psychiatry, 2013. **70**(9): p. 956-963.
14. Rauschert, S., et al., *Maternal smoking during pregnancy induces persistent epigenetic changes into adolescence, independent of postnatal smoke exposure and is associated with cardiometabolic risk.* Frontiers in Genetics, 2019. **10**(770).
15. Langley, K., et al., *Maternal smoking during pregnancy as an environmental risk factor for attention deficit hyperactivity disorder behaviour. A review.* Minerva Pediatr, 2005. **57**(6): p. 359-71.
16. Kazdin, A.E., *Practitioner review: Psychosocial treatments for conduct disorder in children.* Child Psychology & Psychiatry & Allied Disciplines, 1997. **38**(2): p. 161-178.
17. Gluckman, P.D., M.A. Hanson, and T. Buklijas, *A conceptual framework for the developmental origins of health and disease.* Journal of Developmental Origins of Health and Disease, 2009. **1**(1): p. 6-18.
18. Joubert, B.R., et al., *450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy.* Environ Health Perspect, 2012. **120**(10): p. 1425-31.
19. Joubert, B.R., et al., *DNA Methylation in Newborns and Maternal Smoking in Pregnancy: Genome-wide Consortium Meta-analysis.* Am J Hum Genet, 2016. **98**(4): p. 680-96.

650 20. Joubert Bonnie, R., et al., *450K Epigenome-Wide Scan Identifies Differential DNA Methylation*
651 *in Newborns Related to Maternal Smoking during Pregnancy.* Environmental Health
652 Perspectives, 2012. **120**(10): p. 1425-1431.
653 21. Richmond, R.C., et al., *Prenatal exposure to maternal smoking and offspring DNA*
654 *methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and*
655 *Children (ALSPAC).* Human Molecular Genetics, 2014. **24**(8): p. 2201-2217.
656 22. Richmond, R.C., et al., *Prenatal exposure to maternal smoking and offspring DNA*
657 *methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and*
658 *Children (ALSPAC).* Human molecular genetics, 2015. **24**(8): p. 2201-2217.
659 23. Lee, K.W., Richmond, R., Hu, P., French, L., Shin, J., Bourdon, C., … Pausova, Z. (2015).
660 Prenatal exposure to maternal cigarette smoking and DNA methylation: epigenome-wide
661 association in a discovery sample of adolescents and replication in an independent cohort at
662 birth through 17 years of age. Environmental health perspectives, 123(2), 193–199.
663 doi:10.1289/ehp.1408614.
664 24. Osborne, A.J., et al., *Chapter 4 - Epigenetics and the Maternal Germline*, in *Transgenerational*
665 *Epigenetics*, T. Tollefsbol, Editor. 2014, Academic Press: Oxford. p. 27-41.
666 25. Parmar, P., et al., *Association of maternal prenatal smoking GFI1-locus and cardio-metabolic*
667 *phenotypes in 18,212 adults.* EBioMedicine, 2018. **38**: p. 206-216.
668 26. Sengupta, S.M., et al., *Locus-specific DNA methylation changes and phenotypic variability in*
669 *children with attention-deficit hyperactivity disorder.* Psychiatry Research, 2017. **256**: p. 298-
670 304.
671 27. Dolinoy, D.C., D. Huang, and R.L. Jirtle, *Maternal nutrient supplementation counteracts*
672 *bisphenol A-induced DNA hypomethylation in early development.* 2007. **104**(32): p. 13056-
673 13061.
674 28. Waterland, R.A., et al., *Maternal methyl supplements increase offspring DNA methylation at*
675 *Axin fused.* 2006. **44**(9): p. 401-406.
676 29. Waterland, R.A. and R.L. Jirtle, *Transposable Elements: Targets for Early Nutritional Effects*
677 *on Epigenetic Gene Regulation.* 2003. **23**(15): p. 5293-5300.
678 30. Waterland, R.A., et al., *Season of conception in rural gambia affects DNA methylation at*
679 *putative human metastable epialleles.* PLoS genetics, 2010. **6**(12): p. e1001252-e1001252.
680 31. Harris, R.A., D. Nagy-Szakal, and R. Kellermayer, *Human metastable epiallele candidates link*
681 *to common disorders.* Epigenetics, 2013. **8**(2): p. 157-163.
682 32. Kühnen, P., et al., *Interindividual Variation in DNA Methylation at a Putative POMC*
683 *Metastable Epiallele Is Associated with Obesity.* Cell Metabolism, 2016. **24**(3): p. 502-509.
684 33. Joubert, B.R., et al., *450K epigenome-wide scan identifies differential DNA methylation in*
685 *newborns related to maternal smoking during pregnancy.* Environmental health
686 perspectives, 2012. **120**(10): p. 1425.
687 34. Richmond, R.C., et al., *Prenatal exposure to maternal smoking and offspring DNA*
688 *methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and*
689 *Children (ALSPAC).* Human Molecular Genetics, 2015. **24**(8): p. 2201-2217.
690 35. de Vocht, F., et al., *Assessment of Offspring DNA Methylation across the Lifecourse*
691 *Associated with Prenatal Maternal Smoking Using Bayesian Mixture Modelling.* International
692 journal of environmental research and public health, 2015. **12**(11): p. 14461-14476.
693 36. van Otterdijk, S.D., A.M. Binder, and K.B. Michels, *Locus-specific DNA methylation in the*
694 *placenta is associated with levels of pro-inflammatory proteins in cord blood and they are*
695 *both independently affected by maternal smoking during pregnancy.* Epigenetics, 2017.
696 **12**(10): p. 875-885.
697 37. Rotroff, D.M., et al., *Maternal smoking impacts key biological pathways in newborns through*
698 *epigenetic modification in Utero.* BMC Genomics, 2016. **17**(1): p. 976.
699 38. Li, L., et al.

700 39. Skogstrand, K., et al., *Reduced neonatal brain-derived neurotrophic factor is associated with*
701 *autism spectrum disorders.*
702 40. Jiao, S.S., et al., *Brain-derived neurotrophic factor protects against tau-related*
703 *neurodegeneration of Alzheimer's disease.*
704 41. Joubert, B.R., et al., *DNA methylation in newborns and maternal smoking in pregnancy:*
705 *genome-wide consortium meta-analysis.* The American Journal of Human Genetics, 2016.
706 **98**(4): p. 680-696.
707 42. Rzehak, P., et al., *Maternal smoking during pregnancy and DNA-methylation in children at*
708 *age 5.5 years: epigenome-wide-analysis in the European Childhood Obesity Project (CHOP)-*
709 *study.* PloS one, 2016. **11**(5): p. e0155554.
710 43. Suter, M., et al., *In utero tobacco exposure epigenetically modifies placental CYP1A1*
711 *expression.* Metabolism-Clinical and Experimental, 2010. **59**(10): p. 1481-1490.
712 44. Demontis, D., et al., *Discovery of the first genome-wide significant risk loci for ADHD.* BioRxiv,
713 2017: p. 145581.
714 45. Riva, V., et al., *GRIN2B predicts attention problems among disadvantaged children.* European
715 child & adolescent psychiatry, 2015. **24**(7): p. 827-836.
716 46. Acosta, M.T., M. Arcos-Burgos, and M. Muenke, *Attention deficit/hyperactivity disorder*
717 *(ADHD): Complex phenotype, simple genotype?* Genetics in Medicine, 2004. **6**(1): p. 1-15.
718 47. Beaver, K.M., et al., *A gene × gene interaction between DRD2 and DRD4 is associated with*
719 *conduct disorder and antisocial behavior in males.* Behavioral and Brain Functions, 2007.
720 **3**(1): p. 30.
721 48. Salvatore, J.E. and D.M. Dick, *Genetic influences on conduct disorder.* Neuroscience &
722 Biobehavioral Reviews, 2018. **91**: p. 91-101.
723 49. Lantz, P.M., et al., *Socioeconomic Factors, Health Behaviors, and MortalityResults From a*
724 *Nationally Representative Prospective Study of US Adults.* JAMA, 1998. **279**(21): p. 1703-
725 1708.
726 50. Joubert, B.R., et al., *450K epigenome-wide scan identifies differential DNA methylation in*
727 *newborns related to maternal smoking during pregnancy.* Environ Health Perspect, 2012.
728 **120**(10): p. 1425-31.
729 51. Tehranifar, P., et al., *Maternal cigarette smoking during pregnancy and offspring DNA*
730 *methylation in midlife.* Epigenetics, 2018. **13**(2): p. 129-134.
731 52. Rauschert, S., et al., *Maternal Smoking During Pregnancy Induces Persistent Epigenetic*
732 *Changes Into Adolescence, Independent of Postnatal Smoke Exposure and Is Associated With*
733 *Cardiometabolic Risk.* Frontiers in genetics, 2019. **10**: p. 770-770.
734 53. Noble, A., et al., *A validation of Illumina EPIC array system with bisulfite-based amplicon*
735 *sequencing.* bioRxiv, 2020: p. 2020.05.25.115428.
736 54. Tobi, E.W., et al., *DNA methylation signatures link prenatal famine exposure to growth and*
737 *metabolism.* Nature Communications, 2014. **5**(1): p. 5592.
738 55. Philibert, R.A., et al., *Changes in DNA methylation at the aryl hydrocarbon receptor repressor*
739 *may be a new biomarker for smoking.* Clinical Epigenetics, 2013. **5**(1): p. 19.
740 56. Lee, D.-H., et al., *Performance of urine cotinine and hypomethylation of AHRR and F2RL3 as*
741 *biomarkers for smoking exposure in a population-based cohort.* PLOS ONE, 2017. **12**(4): p.
742 e0176783.
743 57. Carter, S., et al., *Maternal smoking during pregnancy and behaviour problems in a birth*
744 *cohort of 2-year-old Pacific children in New Zealand.* Early Human Development, 2008. **84**(1):
745 p. 59-66.
746 58. Mick, E., et al., *Case-Control Study of Attention-Deficit Hyperactivity Disorder and Maternal*
747 *Smoking, Alcohol Use, and Drug Use During Pregnancy.* Journal of the American Academy of
748 Child & Adolescent Psychiatry, 2002. **41**(4): p. 378-385.
749 59. Pon, J.R. and M.A. Marra, *MEF2 transcription factors: developmental regulators and*
750 *emerging cancer genes.* Oncotarget, 2016. **7**(3): p. 2297-2312.

751 60. Li, H., et al., *Transcription factor MEF2C influences neural stem/progenitor cell differentiation*
752 *and maturation in vivo.* Proceedings of the National Academy of Sciences, 2008.
753 **105**(27): p. 9397-9402.
754 61. Li, Z., et al., *Myocyte Enhancer Factor 2C as a Neurogenic and Antiapoptotic Transcription*
755 *Factor in Murine Embryonic Stem Cells.* The Journal of Neuroscience, 2008. **28**(26): p. 6557-
756 6568.
757 62. Chen, S.X., et al., *The transcription factor MEF2 directs developmental visually driven*
758 *functional and structural metaplasticity.* Cell, 2012. **151**(1): p. 41-55.
759 63. Brusco, J. and K. Haas, *Interactions between mitochondria and the transcription factor*
760 *myocyte enhancer factor 2 (MEF2) regulate neuronal structural and functional plasticity and*
761 *metaplasticity.* The Journal of Physiology, 2015. **593**(16): p. 3471-3481.
762 64. Shalizi, A., et al., *A Calcium-Regulated MEF2 Sumoylation Switch Controls Postsynaptic*
763 *Differentiation.* Science, 2006. **311**(5763): p. 1012-1017.
764 65. Gong, X., et al., *Cdk5-Mediated Inhibition of the Protective Effects of Transcription Factor*
765 *MEF2 in Neurotoxicity-Induced Apoptosis.* Neuron, 2003. **38**(1): p. 33-46.
766 66. Mao, Z., et al., *Neuronal Activity-Dependent Cell Survival Mediated by Transcription Factor*
767 *MEF2.* Science, 1999. **286**(5440): p. 785-790.
768 67. Okamoto, S.-i., et al., *Dominant-interfering forms of MEF2 generated by caspase cleavage*
769 *contribute to NMDA-induced neuronal apoptosis.* Proceedings of the National Academy of
770 Sciences, 2002. **99**(6): p. 3974-3979.
771 68. Verzi, M.P., et al., *The Transcription Factor MEF2C Is Required for Craniofacial Development.*
772 Developmental Cell, 2007. **12**(4): p. 645-652.
773 69. Kosfeld, M., et al., *Oxytocin increases trust in humans.* Nature, 2005. **435**(7042): p. 673-676.
774 70. Zak, P.J., A.A. Stanton, and S. Ahmadi, *Oxytocin increases generosity in humans.* PloS one,
775 2007. **2**(11): p. e1128-e1128.
776 71. Meyer, M., et al., *Oxytocin alters the morphology of hypothalamic neurons via the*
777 *transcription factor myocyte enhancer factor 2A (MEF-2A).* Mol Cell Endocrinol, 2018. **477**: p.
778 156-162.
779 72. Meyer, M., et al., *Myocyte Enhancer Factor 2A (MEF2A) Defines Oxytocin-Induced*
780 *Morphological Effects and Regulates Mitochondrial Function in Neurons.* International
781 Journal of Molecular Sciences, 2020. **21**(6): p. 2200.
782 73. Harrington, A.J., et al., *MEF2C regulates cortical inhibitory and excitatory synapses and*
783 *behaviors relevant to neurodevelopmental disorders.* eLife, 2016. **5**: p. e20059.
784 74. Jung, Y., et al., *An epigenetic mechanism mediates developmental nicotine effects on*
785 *neuronal structure and behavior.*
786 75. Rutter M, T.J., Whitmore K, *Education, Health and Behaviour.* London: Longmans, 1970.
787 76. Conners, C.K., *Symptom patterns in hyperkinetic, neurotic, and normal children.* Child
788 Development, 1970. **41**(3): p. 667-682.
789 77. Conners, C.K., *A teacher rating scale for use in drug studies with children.* The American
790 Journal of Psychiatry, 1969. **126**(6): p. 884-888.
791 78. Fergusson, D.M., L.J. Horwood, and M. Lloyd, *Confirmatory factor models of attention deficit*
792 *and conduct disorder.* J Child Psychol Psychiatry, 1991. **32**(2): p. 257-74.
793 79. Fergusson, D.M., L.J. Horwood, and E.M. Ridder, *Show me the child at seven: the*
794 *consequences of conduct problems in childhood for psychosocial functioning in adulthood.* J
795 Child Psychol Psychiatry, 2005. **46**(8): p. 837-49.
796 80. Cox, M.P., D.A. Peterson, and P.J. Biggs, *SolexaQA: At-a-glance quality assessment of*
797 *Illumina second-generation sequencing data.* BMC Bioinformatics, 2010. **11**(1): p. 485.
798 81. Krueger, F. and S.R. Andrews, *Bismark: a flexible aligner and methylation caller for Bisulfite-*
799 *Seq applications.* Bioinformatics, 2011. **27**(11): p. 1571-1572.
800 82. Chen, Y., et al., *Differential methylation analysis of reduced representation bisulfite*
801 *sequencing experiments using edgeR.* F1000Research, 2017. **6**: p. 2055-2055.

802

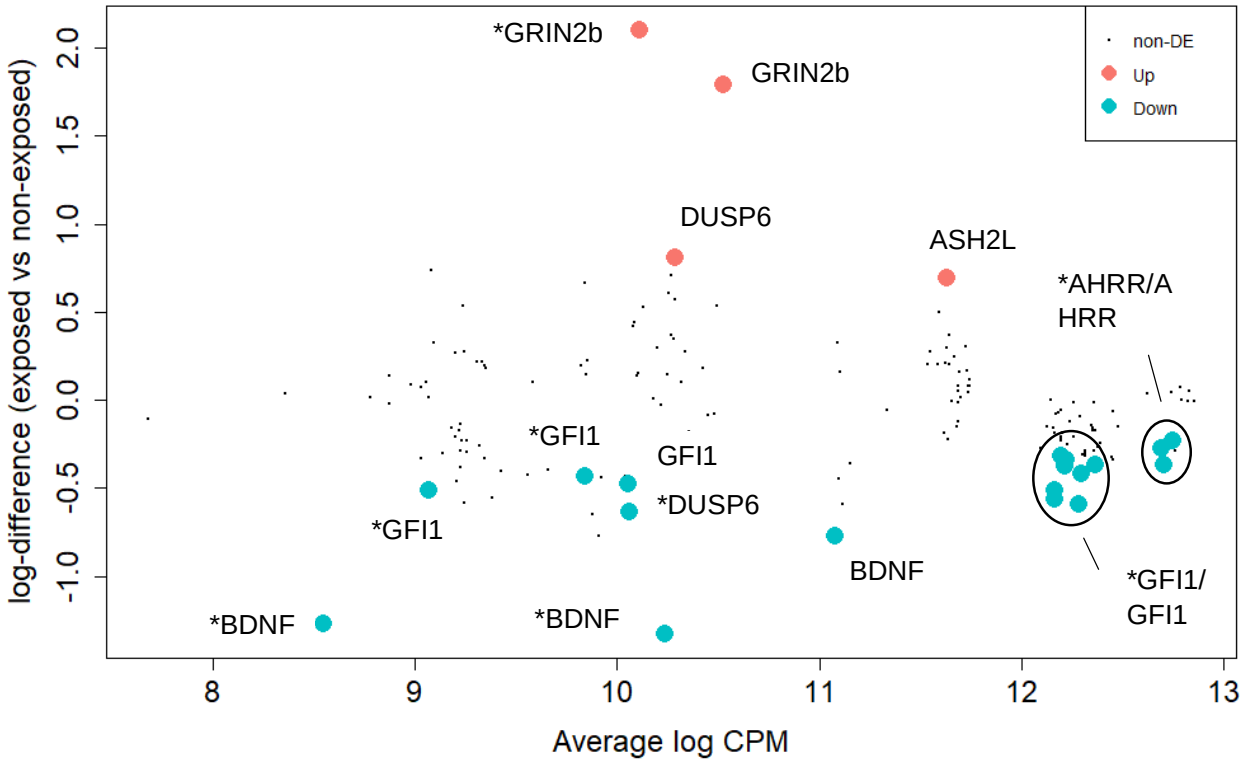
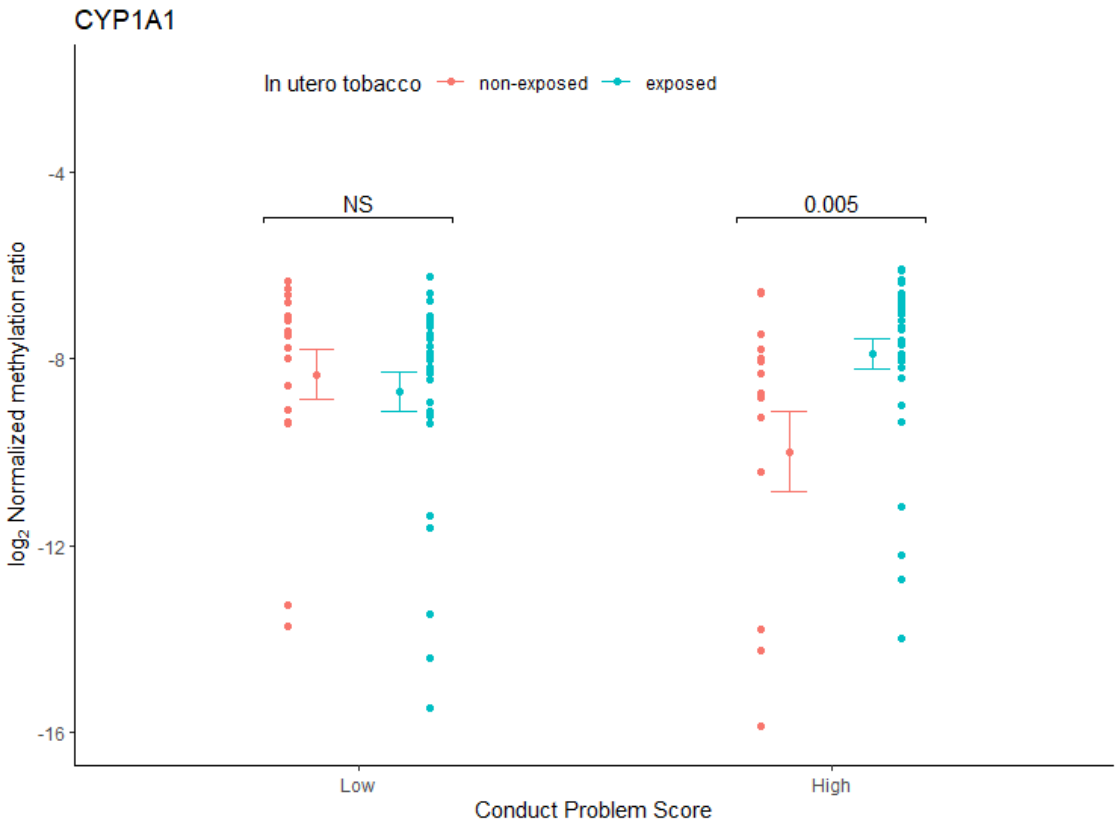
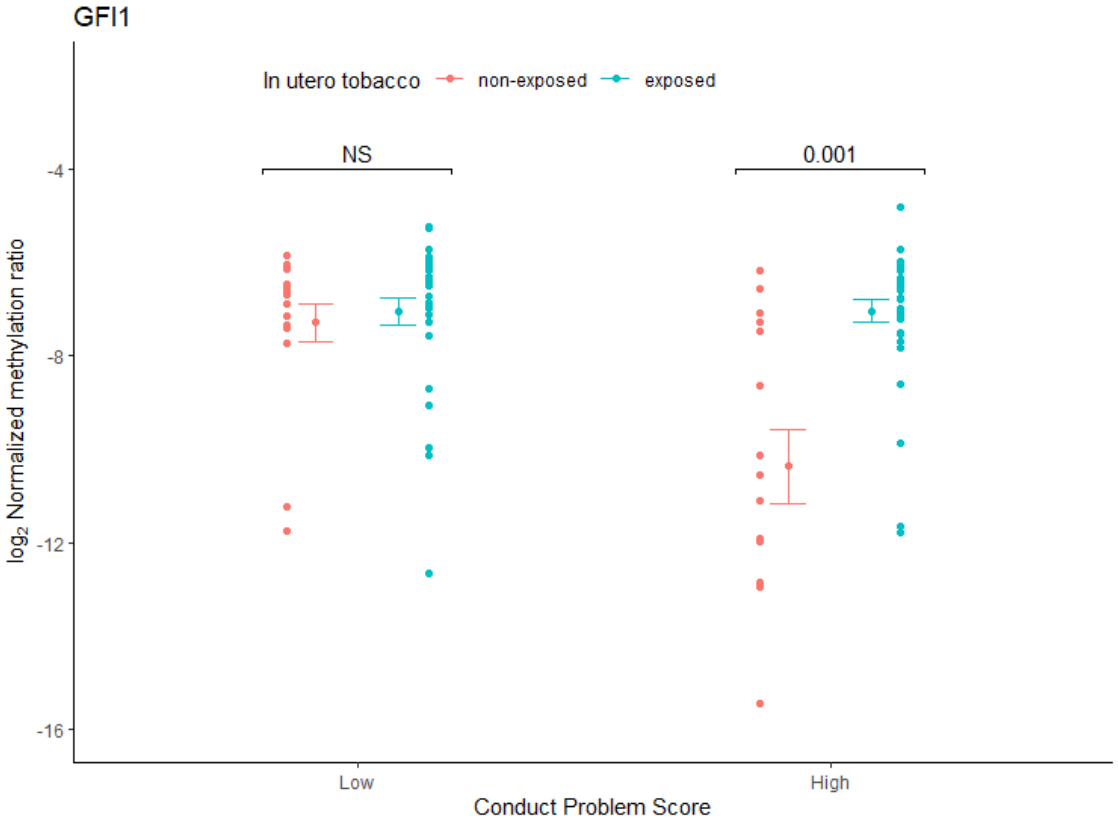


Figure 1- Differential DNA methylation of individuals exposed to tobacco in utero *vs non-exposed* in utero *individuals. Dots that are displayed in colour represent those that are significantly differentially methylated at a nominal P < 0.05: cyan, hypomethylation; pink, hypermethylation; black, non-significantly differentially methylated sites. *previously shown to be differentially methylated in response to adult smoking status.*

A


CYP1A1
In utero tobacco
non-exposed
exposed
NS
0.005
-4
-8
-12
-16
log2 Normalized methylation ratio
Low
High
Conduct Problem Score

B


GFI1
In utero tobacco
non-exposed
exposed
NS
0.001
-4
-8
-12
-16
log2 Normalized methylation ratio
Low
High
Conduct Problem Score

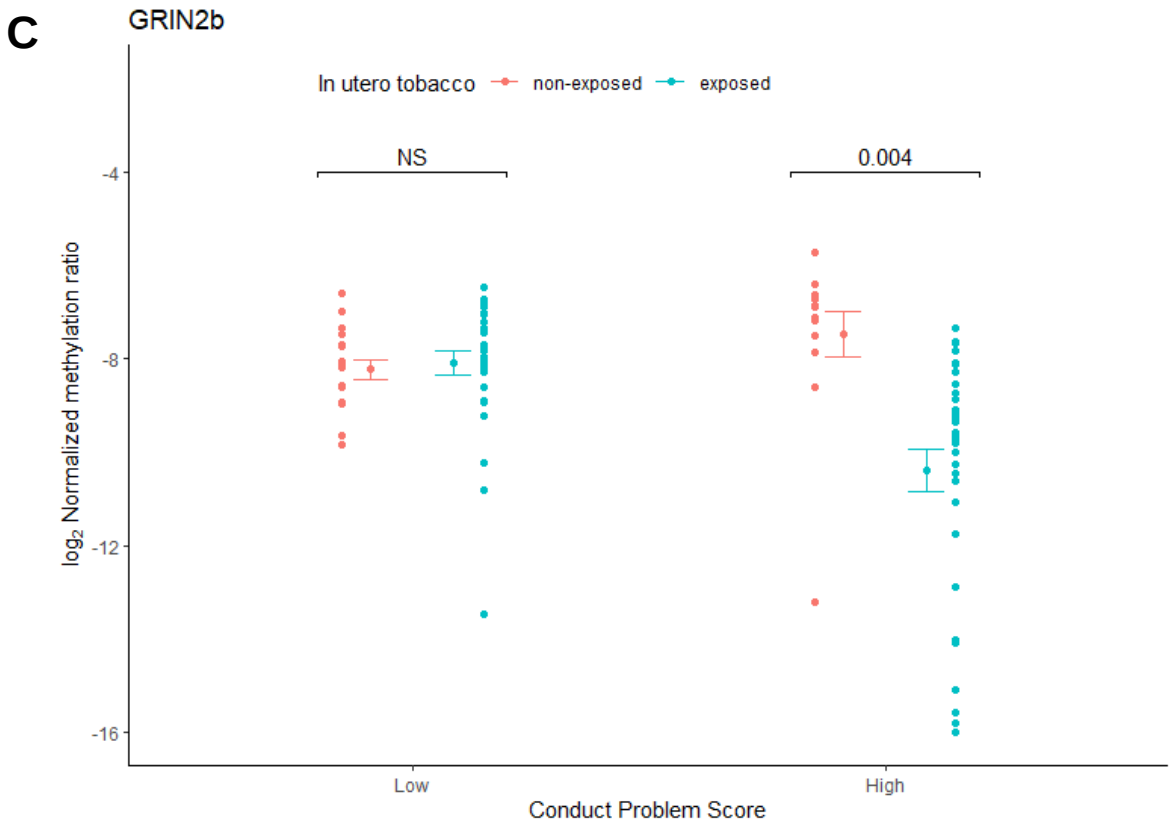


Figure 2- Differential methylation with *in utero* tobacco exposure for individuals with high conduct problem score that is not observed in individuals with low conduct problem score. A- *CYP1A1* (Chr15, 75019290), B- GFI1 (Chr1, 92947705) and C- , GRIN2b (Chr12, 14133359).