DNA METHYLATION AND STRESS IN CHILD DEVELOPMENT

A POPULATION-BASED APPROACH

ROSA H. MULDER

DNA methylation and stress in child development

A population-based approach

DNA-methylering en stress in de ontwikkeling van het kind

Een epidemiologische benadering

© copyright Rosa H. Mulder, 2020

Cover image: Christa Rijneveld
Printing: ProefschriftMaken || www.proefschriftmaken.nl

ISBN: 978-94-6423-035-2

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission of the author or the copyright-owning journals for previous published Chapters.

DNA methylation and stress in child development
A population-based approach

DNA-methylering en stress in de ontwikkeling van het kind
Een epidemiologische benadering

Proefschrift
ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
rector magnificus

Prof. dr. F.A. van der Duijn Schouten

en volgens besluit van het College voor Promoties.
De openbare verdediging zal plaatsvinden op
dinsdag 19 januari 2021 om 15:30 uur

door

Rosa Henriëtte Mulder
geboren te Amsterdam

Erasmus University Rotterdam

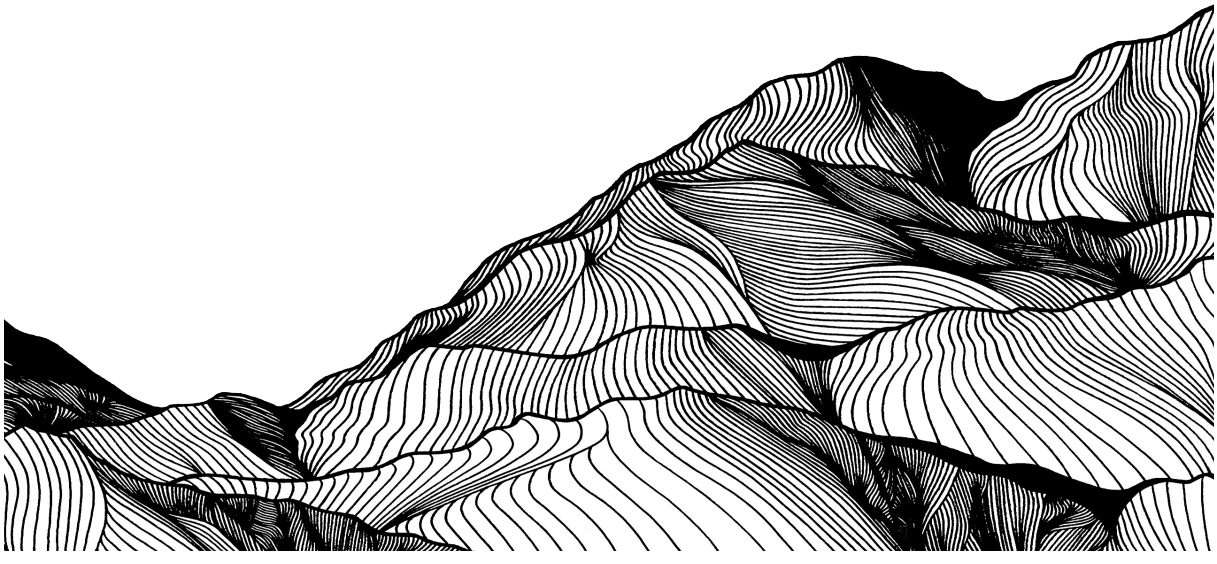
PROMOTIECOMMISSIE

Promotoren	Prof. dr. M. H. van IJzendoorn
	Prof. dr. H. Tiemeier
Overige leden	Prof. dr. P. W. Jansen
	Prof. dr. P. E. Slagboom
	Prof. dr. C. L. Relton
Copromotor	Prof. dr. M. J. Bakermans-Kranenburg

Paranimfen	Desi Koçevska
	Stefanie Schenk

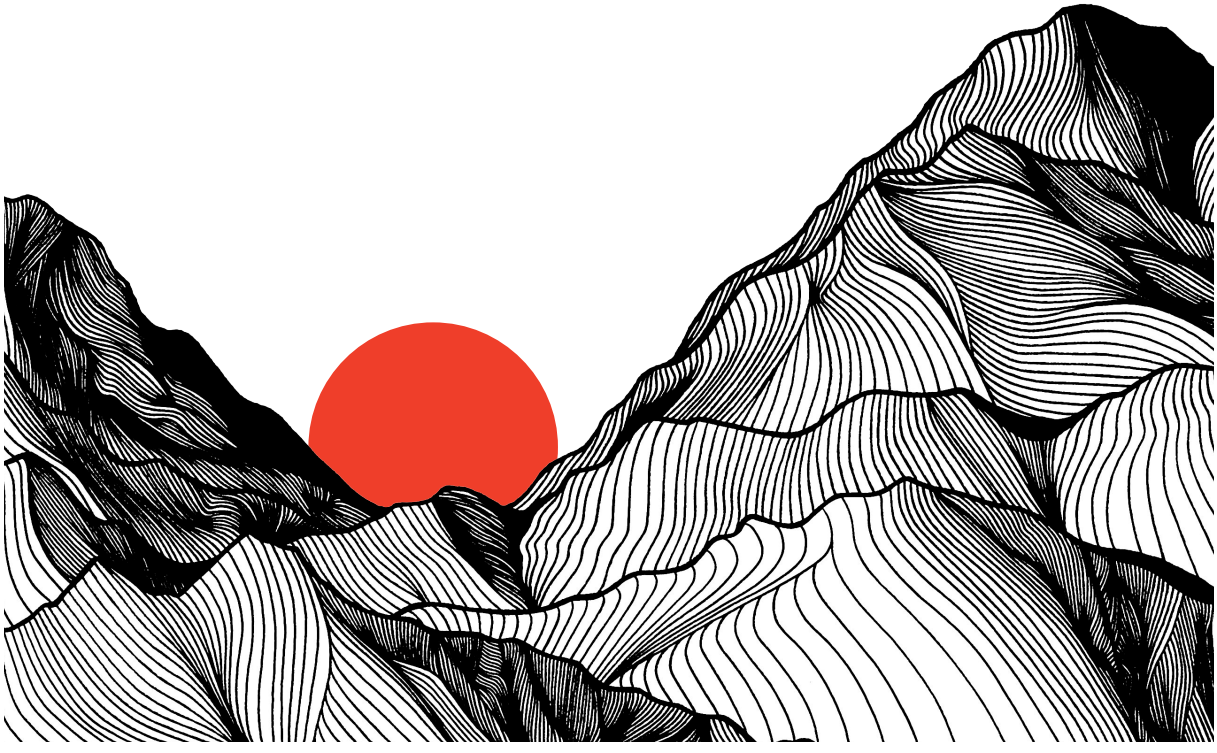
Index

Chapter I	Introduction	9
Chapter II	DNA methylation as a mediator between parenting stress and adverse child development	21
Chapter III	Methylation matters: FK506 binding protein 51 (*FKBP5*) methylation moderates the associations of *FKBP5* genotype and resistant attachment with stress regulation	51
Chapter IV	Epigenome-wide associations between observed maternal sensitivity and offspring DNA methylation: A population-based prospective study in children	79
Chapter V	Genome-wide DNA methylation patterns associated with sleep and mental health in children: A population-based study	111
Chapter VI	Facing ostracism: Micro-coding facial expressions in the Cyberball social exclusion paradigm	137
Chapter VII	Epigenomics of being bullied: Changes in DNA methylation following bullying exposure	165
Chapter VIII	Epigenome-wide change and variation in DNA methylation from birth to late adolescence	203
Chapter IX	Discussion	245
Chapter X	Summary / Samenvatting	261
Summary	Samenvatting	265
Addendum	Authors and affiliations	273
	Publications	275
	About the author	277
	PhD portfolio	278
	Dankwoord (words of gratitude)	282
	Acknowledgements	286



CHAPTER I

Introduction



"Such gene–environment interactions could be explained by epigenetic effects.."[1] is one of many similar statements in the Discussion section of research articles, aiming to explain associations between genetic variations, environment and psychosocial wellbeing, or lack thereof [2-18]. Indeed, life events may affect small molecular structures on and around the DNA, thereby adding a layer of information on top of the genetic code itself and affecting gene functioning. An enticing example of the environment affecting the genetic functioning was given by Weaver, Szyf, and Meaney[19]. They categorized mother rats by the amount of caretaking they did for their pups, and saw epigenetic differences between pups that had received more versus less caretaking. These epigenetic differences coincided with higher stress reactivity in the pups who were less taken care of by their mother and ultimately, when the female pups became mothers themselves they showed less caretaking behavior. These findings could have major implications if similar processes can be identified in humans, since they may explain how stress can get 'under the skin', and, as the authors illustrated, they may explain transgenerational transmission of stress and associated psychopathology. Finding such epigenetic patterns associated with stress may ultimately help understand how to intervene before stress develops into psychopathology, or such patterns may be used as 'markers' (biomarkers) to identify individuals exposed to stress, or sensitive to stressors. In the current thesis, we therefore set out to detect associations between DNA methylation and stress in a population-based sample of children.

Epigenetics

'Epigenetics' was first introduced as a term by Conrad Waddington, in 1939[20]. He envisioned an 'epigenetic landscape' – which he portrayed as a mountainous landscape – formed by genes and 'the chemical tendencies which the gene produce', and affected by environmental stimuli. The developing organism – portrayed as a ball rolling down the ridges and valleys of the landscape – was envisioned to have its phenotype formed by this landscape[21, 22]. In present times, now that scientific advances have made it possible to measure the DNA structure and the chemical compounds around it, Waddington's 'epigenetic landscape' does not seem far from reality.

Many different forms of epigenetics have been identified and generally they affect the accessibility of the genome to transcription. Our human chromosomes, the 23 pairs of DNA strands, are altogether about 2 meters long, but fit within a cell nucleus, which only has a diameter of about 10 μm[23]. This is possible because the chromosomes are coiled and tightly packed by proteins called histones. For a gene to be 'read' and turned into a functioning protein, the DNA structure needs to uncoil, activators need to attach to enhancers, transcription factors need to attach, the DNA structure needs to bend over, so that finally RNA polymerase can attach to the promoter of that gene, transcribe a copy of it, and translate it into a protein[24] (Figure 1). This is exactly the point at which epigenetics can alter gene functioning; by changing the three-dimensional structure of the DNA so that the gene becomes more or less accessible

to activators and/or RNA polymerase and can thereby be more or less transcribed. Known forms of epigenetics often have to do with modification of the histones around which the DNA is packed (e.g. histone acetylation, histone methylation, histone phosphorylation)[25]. Another form is DNA methylation, in which case a methyl group (one carbon atom with three hydrogen atoms) is attached directly to the DNA, on the phosphor bridge between a cytosine and guanine nucleotide (cytosine-phosphor-guanine site; CpG site). This is a more stable form of epigenetic modification and is most frequently studied. For example, when a methyl group is bound to a CpG site located on the promoter of a gene, it is thought that DNA methylation can block RNA polymerase, which would result in less transcribed protein from that gene[26]. DNA methylation, however, can also bind at other parts on or around the genes and there are also examples of DNA methylation íncreasing the transcription, or DNA methylation on one gene affecting the transcription of another gene, sometimes even on a different chromosome, probably due to changing the three-dimensional structure and the consequential alignment of different chromosomes[27-29]. With these characteristics, we know that DNA methylation affects basic developmental processes, such as cell differentiation[30, 31], X-chromosomal inactivation[32], and genome stability[33]. We further know that DNA methylation is influenced by the genes themselves[34], as well as by environmental aspects, such as smoking[35,

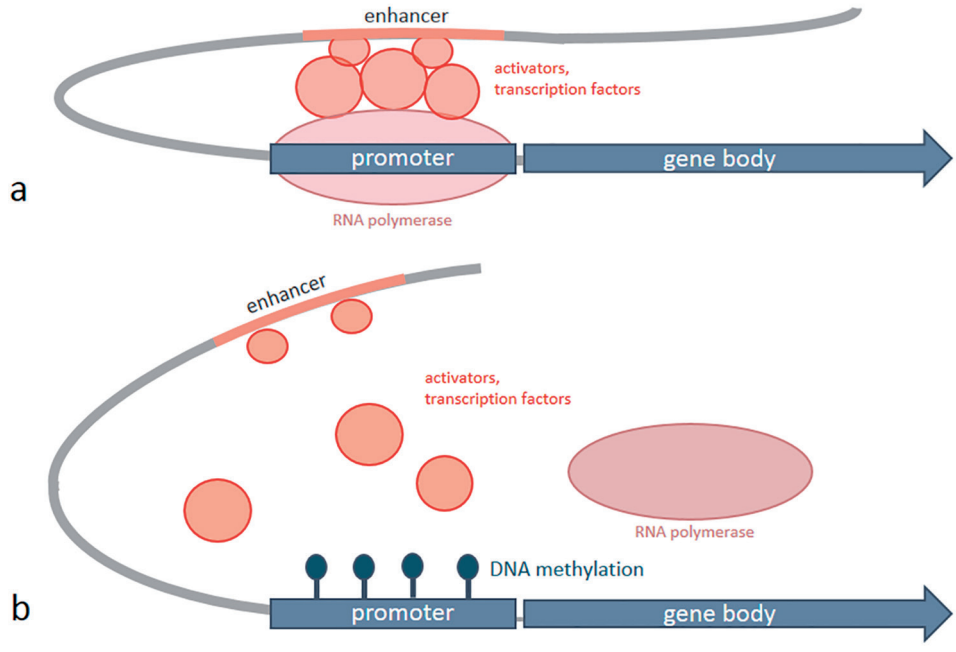


Figure 1. Binding of RNA polymerase to gene promoter (a) and model of how DNA methylation might block this binding (b)

[36]. Here we will try to specify whether we can measure associations between DNA methylation and stress exposure and–outcomes in children.

Measurement of epigenetics

In the example of the research on rats given above, DNA methylation was studied on the promoter of a single gene. Such an approach is often used with so-called candidate genes, genes that are suspected to be relevant for phenotype of interest. A candidate-gene approach can function as a 'proof-of-principle', showing that DNA methylation affects the functioning of a familiar gene. To be able to find novel biological pathways through epigenetic analyses, however, one needs to cast a wider net. In recent years, epigenome-wide array testing has become increasingly available. With epigenome-wide association studies (EWASs), one can measure DNA methylation at hundreds of thousands of CpG sites widespread over the genome. Here, we will use both candidate epigenetic research as well as epigenome-wide approaches to study DNA methylation in blood tissue. Further, while studies on animals or candidate-gene studies have shown a critical role of DNA methylation in development, changes in DNA methylation in blood during childhood are not well-characterized. This lack of knowledge is impeding interpretation of findings in current EWASs. As one of the goals of the current thesis, we aimed to form an encompassing epigenome-wide characterization of DNA methylation throughout development.

Stress

All people encounter stress, some more than others. Our body is adapted to deal with stress through the hypothalamic-pituitary-adrenal axis, or HPA axis. When a stressor occurs and this is registered in the brain, the hypothalamus produces corticotropin-releasing hormones, which signals to the pituitary to produce adrenocorticotropic hormone, and this in turn to the adrenal cortices to produce cortisol. Amongst others, cortisol helps activate the hippocampus to encode the event into a memory. When cortisol feeds back to the hippocampus, hypothalamus, and pituitary, it signals to reduce the production of corticotropin-releasing hormone and adrenocorticotropic hormone in a negative feedback loop[37]. Cortisol levels vary throughout the day, are related to the sleep-wake cycle, and variations in cortisol facilitate the consolidation of memories in the hippocampus during sleep[38, 39]. Because DNA methylation might impact the genetic expression of cortisol or related hormones, we will study sleep in association with DNA methylation.

It is thought that childhood experiences can have long-term effects on HPA axis functioning. Early family experiences and attachments to the parents are important in the development of the child's cognition representation of how safe the world is, and how to deal with stressors[40, 41]. When children start moving into adolescence, peers take on a more formative role in their development[42]. We will therefore study attachment and stress within the family setting in early childhood, and interpersonal stress with peers in early adolescence.

Aim

In this thesis, we will study associations between stress and DNA methylation in the developing child.

Setting

The main study population in the current thesis are the children in the Generation R Study. As part of this study, pregnant women residing in Rotterdam, the Netherlands, with an expected delivery date between April 2002 and January 2006 were invited to enroll[43]. Children of the Generation R Study are ethnically diverse, DNA methylation was however measured in a subsample of children, all of which have parents born in the Netherlands. DNA methylation was measured in these children at birth, 6 years, and 10 years. Further, hands-on measurements, such as parent-child observations and sleep measures, were also conducted in a subsample of children with parents born in the Netherlands. Two studies in this thesis additionally make use of data from the Avon Longitudinal Study of Parents and Children (ALSPAC). Here, pregnant women residing in the study area of former county Avon, in the United Kingdom, with an expected delivery date between April 1991 and December 1992 were invited to enroll[44, 45]. In a subsample of children, DNA methylation was taken at birth, 7 years, and 17 years of age.

Outline

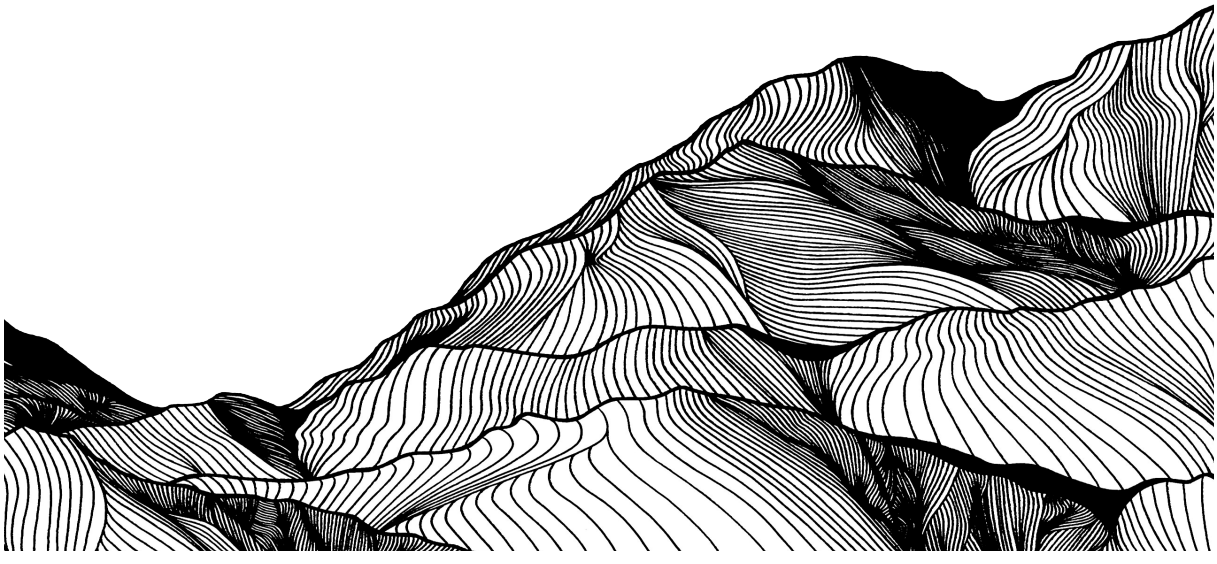
We will start this thesis by giving an overview of current literature on DNA methylation and stress in the family in Chapter II. Here, we will highlight several shortcomings in these previous studies, such as the use of small sample sizes, a lack of effort to replicate results, and a lack of knowledge on longitudinal characteristics of DNA methylation, which are issues that we try to address in the following Chapters. In Chapter III, we will perform a proof of principle, studying if DNA methylation of a single candidate gene previously shown to be involved in HPA functioning also in our data, affects cortisol reactivity. From Chapter IV onwards, we will extend our approach towards the whole genome to examine if maternal sensitivity is associated with DNA methylation and, in another study, with sleep. Subsequently, we will study the associations of interpersonal stress in Chapter V and VI . In Chapter V, we will present our application of a social exclusion paradigm and a new method of micro-coding facial expressions, showing how stressful social exclusion can be. In Chapter VI, using a longitudinal design we will study if bullying is related to change in DNA methylation, both in children of Generation R as well as in children of ALSPAC. Since epigenome-wide longitudinal studies are becoming more prevalent, but basic information on which methylated CpGs change is currently lacking, in our final study in Chapter VII we will detail the epigenome-wide change from birth to late adolescence. This again is a study with participants from Generation R and ALSPAC. Finally, in Chapter VIII, we will discuss our findings on DNA methylation and stress in the developing child.

References

1. Tollenaar MS, Molendijk ML, Penninx BWJH, Milaneschi Y, Antypa N. The association of childhood maltreatment with depression and anxiety is not moderated by the oxytocin receptor gene. *European Archives of Psychiatry and Clinical Neuroscience.* 2017;267:517-26.
2. Bergen SE, Petryshen TL. Genome-wide association studies (GWAS) of schizophrenia: does bigger lead to better results? *Current Opinion in Psychiatry.* 2012;25:76.
3. Breen ME, Seifuddin F, Zandi PP, Potash JB, Willour VL. Investigating the role of early childhood abuse and HPA axis genes in suicide attempters with bipolar disorder. *Psychiatric Genetics.* 2015;25:106.
4. Buckholtz JW, Meyer-Lindenberg A. MAOA and the neurogenetic architecture of human aggression. *Trends in Neuroscience.s* 2008;31:120-9.
5. Chen J, Yu J, Liu Y, Zhang L, Zhang J. BDNF Val66Met, stress, and positive mothering: Differential susceptibility model of adolescent trait anxiety. *Journal of Anxiety Disorders.* 2015;34:68-75.
6. Clarke MC, Tanskanen A, Huttunen MO, Clancy M, Cotter DR, Cannon M. Evidence for shared susceptibility to epilepsy and psychosis: a population-based family study. *Biological Psychiatry*. 2012;71:836-9.
7. Cohen-Woods S, Fisher HL, Ahmetspahic D, et al. Interaction between childhood maltreatment on immunogenetic risk in depression: discovery and replication in clinical case-control samples. *Brain, Behavior, and Immunity.* 2018;67:203-10.
8. Erhardt A, Lucae S, Unschuld PG, et al. Association of polymorphisms in P2RX7 and CaMKKb with anxiety disorders. *Journal of Affective Disorders* 2007;101:159-68.
9. Frodl T, Reinhold E, Koutsouleris N, et al. Childhood stress, serotonin transporter gene and brain structures in major depression. *Neuropsychopharmacology*. 2010;35:1383-90.
10. Gallardo-Pujol D, Andrés-Pueyo A, Maydeu-Olivares A. MAOA genotype, social exclusion and aggression: An experimental test of a gene–environment interaction. *Genes, Brain and Behavior.* 2013;12:140-5.
11. Gerritsen L, Milaneschi Y, Vinkers CH, et al. HPA axis genes, and their interaction with childhood maltreatment, are related to cortisol levels and stress-related phenotypes. *Neuropsychopharmacology.* 2017;42:2446-55.
12. Holz N, Boecker R, Buchmann AF, et al. Evidence for a sex-dependent MAOA× childhood stress interaction in the neural circuitry of aggression. *Cerebral Cortex.* 2016;26:904-14.
13. Lahti J, Ala-Mikkula H, Kajantie E, Haljas K, Eriksson JG, Räikkönen K. Associations between self-reported and objectively recorded early life stress, FKBP5 polymorphisms, and depressive symptoms in midlife. *Biological Psychiatry.* 2016;80:869-77.
14. Mbarek H, Milaneschi Y, Fedko IO, et al. The genetics of alcohol dependence: Twin and SNP-based heritability, and genome-wide association study based on AUDIT scores. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics.* 2015;168:739-48.
15. Mitchell C, Notterman D, Brooks-Gunn J, et al. Role of mother's genes and environment in postpartum depression. *Proceedings of the National Academy of Sciences.* 2011;108:8189-93.

16. Nigg J, Nikolas M, Burt SA. Measured gene-by-environment interaction in relation to attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child & Adolescent Psychiatry.* 2010;49:863-73.
17. van Os J, Rutten BPF, Poulton R. Gene-environment interactions in schizophrenia: review of epidemiological findings and future directions. *Schizophrenia Bulletin.* 2008;34:1066-82.
18. Zhang W, Cao C, Wang M, Ji L, Cao Y. Monoamine oxidase a (MAOA) and catechol-o-methyltransferase (COMT) gene polymorphisms interact with maternal parenting in association with adolescent reactive aggression but not proactive aggression: evidence of differential susceptibility. *Journal of Youth and Adolescence.* 2016;45:812-29.
19. Weaver ICG, Szyf M, Meaney MJ. From maternal care to gene expression: DNA methylation and the maternal programming of stress responses. *Endocrine Research.* 2002;28:699-0.
20. Waddington CH. *An Introduction to Modern Genetics.* 1939.
21. Waddington C. *The Strategy of the Genes.* Allen & Unwin, London, 1957.
22. Waddington CH. *Organisers and Genes.* 1940.
23. Dundr M, Misteli T. Functional architecture in the cell nucleus. *Biochemical Journal.* 2001;356:297-310.
24. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. *Molecular Biology of the Cell.* Garland Science. New York 2007;1392.
25. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Research.* 2011;21:381-95.
26. Métivier R, Gallais R, Tiffoche C, et al. Cyclical DNA methylation of a transcriptionally active promoter. *Nature.* 2008;452:45-50.
27. Ball MP, Li JB, Gao Y, et al. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nature Biotechnology.* 2009;27:361.
28. Jones PA. Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nature Reviews Genetics.* 2012;13:484.
29. Li G, Reinberg D. Chromatin higher-order structures and gene regulation. *Current Opinion in Genetics & Development.* 2011;21:175-86.
30. Khavari DA, Sen GL, Rinn JL. DNA methylation and epigenetic control of cellular differentiation. *Cell Cycle.* 2010;9:3880-3.
31. Meissner A, Mikkelsen TS, Gu H, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature.* 2008;454:766-70.
32. Chow J, Heard E. X inactivation and the complexities of silencing a sex chromosome. *Current Opinion in Cell Biology.* 2009;21:359-66.
33. Eden A, Gaudet F, Waghmare A, Jaenisch R. Chromosomal instability and tumors promoted by DNA hypomethylation. *Science.* 2003;300:455-0.
34. Gaunt TR, Shihab HA, Hemani G, et al. Systematic identification of genetic influences on methylation across the human life course. *Genome Biology.* 2016;17:61.
35. Zeilinger S, Kühnel B, Klopp N, et al. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS ONE.* 2013;8:e63812.

36. Bojesen SE, Timpson N, Relton C, Smith GD, Nordestgaard BG. AHRR (cg05575921) hypomethylation marks smoking behaviour, morbidity and mortality. *Thorax* 2017;72:646-53.
37. Wolf OT. HPA axis and memory. *Best Practice & Research Clinical Endocrinology & Metabolism.* 2003;17:287-99.
38. Zeiders KH, Doane LD, Adam EK. Reciprocal relations between objectively measured sleep patterns and diurnal cortisol rhythms in late adolescence. *Journal of Adolescent Health.* 2011;48:566-71.
39. Wagner U, Born J. Memory consolidation during sleep: interactive effects of sleep stages and HPA regulation. *Stress.* 2008;11:28-41.
40. Shonkoff JP, Garner AS, Siegel BS, et al. The lifelong effects of early childhood adversity and toxic stress. *Pediatrics.* 2012;129:e232-e46.
41. Mesman J, van IJzendoorn MH, Bakermans-Kranenburg MJ. Unequal in opportunity, equal in process: Parental sensitivity promotes positive child development in ethnic minority families. *Child Development Perspectives.* 2012;6:239-50.
42. Forbes EE, Dahl RE. Pubertal development and behavior: hormonal activation of social and motivational tendencies. *Brain and Cognition.* 2010;72:66-72.
43. Kooijman MN, Kruithof CJ, van Duijn CM, et al. The Generation R Study: design and cohort update 2017. *European Journal of Epidemiology.* 2016;31:1243-64.
44. Boyd A, Golding J, Macleod J, et al. Cohort profile: the 'children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology.* 2013;42:111-27.
45. Fraser A, Macdonald-Wallis C, Tilling K, et al. Cohort profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *International Journal of Epidemiology.* 2012;42:97-110.

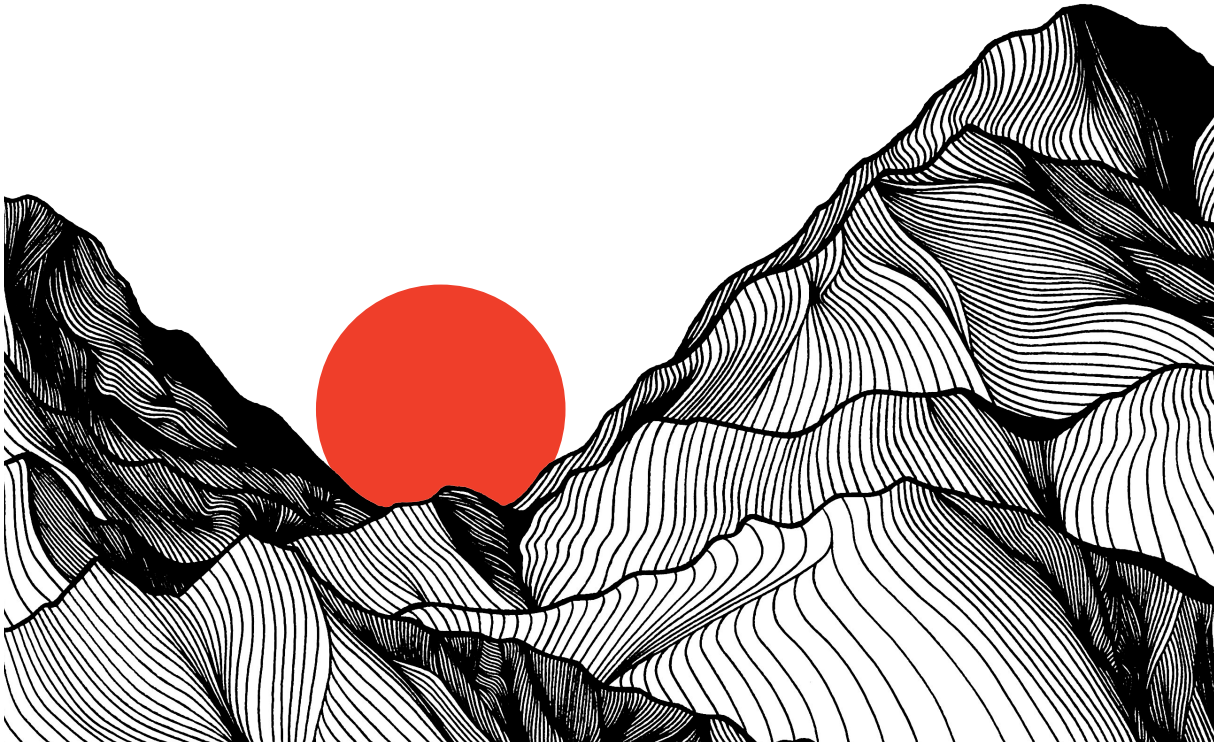


CHAPTER II

DNA methylation as a mediator between parenting stress and adverse child development

Rosa H. Mulder, Jolien Rijlaarsdam, & Marinus H. van IJzendoorn

In Parental Stress and Early Child Development (2017), pp. 157-180. Springer, Cham.



Abstract

In this Chapter, we provide an overview of empirical studies evaluating the role of epigenetics in mediating the association between parenting stress and adverse child development. We focus on DNA methylation, as this epigenetic mechanism is most often studied in humans. Here, parenting stress will be defined as including both prenatal stressors (e.g. maternal psychopathology during pregnancy) as well as postnatal maladaptive parenting (e.g. harsh discipline). We define adverse child development in terms of biological (e.g. cortisol reactivity, brain morphology), as well as psychological outcomes.

Most epigenetic research focuses on *either* the association between parenting stress and DNA methylation *or* on the association between DNA methylation and child outcomes. In the current Chapter, we emphasize the mediation of the association between parenting stress and adverse child outcomes *via* epigenetics. We conclude with some caveats that should be considered when conducting or reading epigenetic studies on parenting and child development, and we discuss future research avenues.

Introduction

In the contemporary version of the nature versus nurture debate it is taken for granted that the (parental) environment *as well as* the genetic make-up determine the behavior of a developing child, with the child's genome being differentially open to environmental influences. For example, in their ground-breaking gene-by-environment (G × E) study, Caspi et al.[1] found that individuals who had experienced stressful life events were more often depressed when they carried one or two short alleles of the serotonin transporter gene (*5-HTT* or *SLC6A4*) in the serotonin-transporter-linked polymorphic region (*5-HTTLPR*). Likewise, in a first randomized controlled G × E trial, Bakermans-Kranenburg et al.[2] showed that changing sensitive parenting and limit setting only influenced the externalizing behavior if the child was a carrier of the dopamine D4 receptor (*DRD4*) 7-repeat allele.

However, in G × E studies it remains unknown where and how genetics and the environment exactly interact. The field of epigenetics might suture this gap between nature and nurture[3]. 'Epigenetics' is a term coined by the embryologist Conrad Waddington[4, 5], who used it to describe the interplay of genes and external cues in the development of the omnipotent cell into a fully specialized one. A related term, epigenesis, was later used by Gilbert Gottlieb to emphasize how variation in the DNA does not simply lead to variation in functioning proteins in a one-to-one fashion, but rather contributes in a bidirectional manner with several layers to the developmental system, going from the genetic level, via the neural and behavior level all the way to the environmental level[6]. Indeed, with modern lab technologies, different epigenetic mechanisms have been identified through which the environment can get 'under the skin' and act upon genetic variation to affect the transcriptional and translational processes to form genes' main product: proteins.

One of these epigenetic mechanisms is DNA methylation, involving a methyl group (CH_3) that attaches to the cytosine nucleotide in the DNA, in places where the cytosine nucleotide is situated alongside a guanine nucleotide, connected via a phosphate bridge (hence cytosine-phosphate-guanine site, or CpG site)[7, 8]. The human genome has millions of CpG sites where a methyl group might be attached, which has been found to affect the three-dimensional DNA formation so that it may hinder or facilitate transcription of the DNA[9]. Other mechanisms work at the level of the histones, proteins around which the chromatin is packaged. Examples would be histone acetylation, or histone methylation[10, 11]. Again, these histone-based mechanisms change the accessibility of the gene for transcriptional processes. Epigenetic mechanisms might also take place further into the translational process, for example, in the form of small non-coding RNA, which can affect splicing variants[12].

Since DNA methylation takes place through the covalent binding of the methyl group to the cytosine nucleotide, this is the most physically stable form of epigenetics and most likely to

survive the chemical treatment that takes place in the lab. Therefore, it is the most frequently studied form of epigenetics. It has been shown that DNA methylation indeed makes the genome more dynamic and is involved, as postulated by Waddington, in cell differentiation[13, 14], as well as in X-chromosomal inactivation in female mammals[15, 16] and in aberrant cell functioning such as cancer[17]. The downstream effects of DNA methylation are complex: it might functionally silence a gene by decreasing its accessibility by DNA polymerase, promote gene transcription by increasing its accessibility but could also, for example, indirectly affect transcription of genes by altering accessibility of distal regulatory regions such as enhancers or silencers[18, 19].

Importantly, it seems that DNA methylation can be affected by life events. In a series of experiments on rodents, Weaver and colleagues showed that early life stress, for example maternal separation, is related to altered stress reactivity in the adult offspring, and that this effect seemed to be mediated by methylation of the promoter of the glucocorticoid receptor gene (*NR3C1* or *GR*) in the hippocampus[20]. The binding of corticosterone (rodents) or cortisol (humans) to the glucocorticoid receptors causes negative feedback to the hypothalamic-pituitary-adrenal axis (HPA-axis) and is necessary to control stress reactivity. Intriguingly, the results of Weaver, Szyf and Meaney imply not only that DNA methylation is affected by life events, but also that it could influence gene transcription to the extent that it changes behavior into adulthood. Moreover, Weaver et al.[21] showed that normal variation in maternal caretaking, as measured by the amount of licking and grooming, could alter methylation of the *NR3C1* promoter.

In humans, it has also been shown that major life events can modify outcomes in later life, possibly via DNA methylation. Examples can be found in the Dutch Hunger Winter Families Studies[22], which focused on offspring conceived in the winter of 1944-45 during the Second World War, a period in which food was extremely scarce and starvation ubiquitous. In these studies, it was found that fetuses who were exposed to famine in the first trimester after conception had less methylation of the insulin-like growth factor II gene (*IGF2*)[23], resulting in lower birth weights, and LDL cholesterol[24] in adulthood.

In this Chapter, we examine whether DNA methylation mediates the relation between parenting stress and child development. Parenting stress is typically indicated by the recording of actual stressors, of parental psychological affliction such as depression or anxiety, and/or of a history of abuse in the child. Such stress might occur during pregnancy, as well as postnatally. Child development may be operationalized as psychological, hormonal, or neurological development. Importantly, throughout the Chapter, several methodological issues will be touched upon as behavioral epigenetics is an emerging field facing a large number of problems and pitfalls.

In the following section, we review studies on the association between parenting stress and DNA methylation, prenatally and postnatally. Effects of DNA methylation cannot be

separated from the genes they act upon, and we will elaborate on such epi-allelic interactions. Subsequently, we consider research on the association between DNA methylation and adverse child development, with a special emphasis on the mediation of the association between parenting stress and child development *via* epigenetics. In a final discussion section, we summarize our findings, and address some caveats.

Epigenetic Signatures as Biomarkers of Exposure

Candidate Epi-Gene Approaches

Adversities and related stress (e.g., maternal depression and anxiety in the prenatal period) have been suggested to affect epigenetic patterns in the neonate, and differences in epigenetic signatures have been speculated to be markers of prenatal programming for postnatal life circumstances (see also Neuenschwander and Oberlander[25]), in accordance with the Barker hypothesis[26]. Several studies examined the association between prenatal stress and methylation state of the *NR3C1* promotor region of the offspring. *NR3C1* has been found to co-regulate secretion and re-uptake of cortisol and might thus be important for regulation of stress. In a ground-breaking study building on earlier work by Weaver and colleagues[21] on rodents, McGowan et al.[27] investigated the postmortem hippocampal brain tissues of male suicide victims with (*n* = 12) and without (*n* = 12) a history of child abuse and those of matched controls who died in car accidents (*n* = 12). They found that suicide victims with a history of child abuse had less GR expression and more methylation of *NR3C1* than did suicide victims without a history of child abuse or controls, whereas no significant difference was found between suicide victims without a history of abuse and controls. Specifically, DNA hypermethylation was found in 3 out of 38 measured CpG sites. Moreover, it was found that within the group of suicide victims with child abuse, more DNA methylation was associated with less GR messenger RNA, as well as less GR messenger RNA overall (messenger RNA triggers the production of associated proteins downstream). These findings indicate that childhood abuse is related to DNA methylation, which decreases *NR3C1* transcription. This might lead to aberrant HPA-axis functioning and dysfunctional stress regulation, rendering the affected individual more susceptible to the development of psychopathologies such as depression and anxiety, ultimately increasing the risk of suicide.

DNA methylation might also be a mechanism through which the intergenerational transmission of stress dysregulation takes place (see also Mileva-Seitz and Fleming[28]). This hypothesis was tested by Yehuda et al.[29], who examined *NR3C1* promoter methylation in a sample of adult offspring (without PTSD) with at least one Holocaust survivor parent (with or without PTSD) (n=80) and demographically matched participants without parental Holocaust exposure or PTSD (n=15). Yehuda et al.[29] found an interaction between maternal and paternal PTSD in the prediction of offspring *NR3C1* promoter methylation. Specifically, only in the absence

of maternal PTSD, offspring exposed to paternal PTSD had higher levels of *NR3C1* promoter methylation. Offspring exposed to both maternal and paternal PTSD showed lower levels of *NR3C1* promoter methylation. Interestingly, *NR3C1* promoter methylation negatively correlated with *NR3C1* expression. Furthermore, stronger cortisol suppression was related to lower DNA methylation. Replication of the rather complicated interactions in a relatively small sample is of course needed, and the results should be considered potentially fruitful hypotheses about the biological underpinnings of the intergenerational transmission of posttraumatic stress.

Thus far, we primarily discussed the association between postnatal parental stress and DNA methylation in the child. It is theorized however, that prenatal parenting, may it be through the intake of harmful agents or through psychological stress, can have a lasting harmful impact on the child[30, 31]. Below, we discuss two studies on how prenatal psychological stress may affect *NR3C1* methylation.

In a study of 83 pregnant women, Hompes et al.[32, 33] assessed maternal stress each trimester and found it to be significantly associated with methylation of one specific CpG site of the *NR3C1* promoter in the cord blood of their newborns. Also, several dimensions of pregnant women's anxiety about their impending delivery predicted methylation of various CpG sites of the nerve growth factor inducible protein A (NGFI-A) binding sites of *NR3C1*[32]. The study was meant to replicate the earlier results of a pioneering study by Oberlander et al[34]. who found no multivariate association between the methylation state of 13 CpG sites in *NR3C1* with several measures for prenatal depression and anxiety in 82 mothers (*n* = 46 depressed), but did find that the methylation of 3 CpG sites were correlated with some prenatal depression and anxiety indicators. Oberlander et al.'s results were not replicated by Hompes et al.[32] who conducted statistical analyses with corrections for multiple testing and found associations during different time windows, on different CpG sites and with different directions. In spite of these inconsistencies, it seems likely that maternal stress during pregnancy is capable of altering gene expression in offspring in ways that increase the risk of stress dysregulation at future points in their development (see also Neuenschwander and Oberlander[25]).

In another related study, 23 mother-child dyads were assessed with retrospective reports of intimate partner violence during mothers' pregnancy and DNA methylation was extracted from blood samples when the children were 10-19 years old[35]. These authors found a significantly higher mean DNA methylation percentage in 10 CpG sites of the promotor region of *NR3C1* in those adolescents whose mothers had experienced intimate partner violence during pregnancy. However, the small number of subjects from various ethnic backgrounds and the relatively large number of statistical tests (not corrected for multiple testing) might make replication of these results difficult. Together, the results of Radtke et al. and Hompes et al. show that stress during pregnancy might affect *NR3C1* methylation of the fetus in a lasting way, but replication is needed.

Taking into account all aforementioned studies, it seems that the effect of stress on *NR3C1* promoter methylation that was initially found in rats, translates into studies on humans. Following, we will briefly discuss some studies that also focus on methylation of genes other than *NR3C1*.

In a study on 57 mothers and their offspring Braithwaite, Kundakovic, Ramchandani, Murphy, and Champagne[36] studied the association between 2nd and 3rd trimester depressive symptoms in the mother and methylation of *NR3C1* and *BDNF* in 2 months old offspring, while controlling for postnatal maternal depressive symptoms. They found that prenatal depressive symptoms were associated with neonatal increased *NR3C1* DNA methylation in male infants, and they also found decreased methylation of an exon upstream of the brain-derived neurotrophic factor gene (*BDNF*) in both male and female infants. In an earlier study on prenatal depression in 82 pregnant women Devlin, Brain, Austin, and Oberlander[37] showed associations with methylation status of *5-HTT*, but in contrast to Braithwaite et al.[36], they did not find associations with methylation of *BDNF*.

Using a sample of 152 females, Vijayendran, Beach, Plume, Brody, and Philibert[38] examined the associations between childhood sexual abuse and DNA methylation at 16 sites across the *5-HTT* gene in females. One out of the 16 measured CpG sites was positively associated with both genotype and sexual abuse, whereas DNA methylation of another CpG site was associated solely with sexual abuse. In a cross-sectional study, Unternaehrer et al.[39] investigated the association between maternal care and DNA methylation of *BDNF* (one sequence including 7 CpG sites) and the oxytocin receptor gene (*OXTR*; two sequences including 6 and 17 CpG sites, respectively). They showed that university students reporting low maternal care in childhood and adolescence (n = 45) had higher levels of DNA methylation in the *BDNF* target sequence than students reporting high maternal care (n = 40). Similarly, students reporting low maternal care had higher levels of DNA methylation in the first *OXTR* target sequence but not in the second target sequence.

Together, these studies suggest that candidate genes involved in stress regulation as well those affecting other regulators of the central nervous system are affected by parenting stress. However, research driven by a priori hypotheses on genes involved can form an 'information bottleneck'[40], as it is unlikely to reveal new genes or mechanisms. Like genome-wide association studies (GWASs), epigenome-wide association studies (EWASs) are hypothesis free, and cover the length of the whole genome. With the latest arrays, EWASs can gauge up to 850,000 CpG sites, in locations such as the promoter, intergenic regions, and intragenic regions. In the following paragraph, we will discuss studies that relate stressful parenting to epigenome-wide DNA methylation.

Epigenome-Wide Association Studies

In developmental and psychiatric epigenetics, the dominant approach is based on methylation patterns of candidate genes and their promotor areas. Epigenome-wide association studies (or EWASs) seem less often used, presumably because the sample sizes involved in this type of research are too small to offer sufficient power for the large numbers of CpG sites to be examined. The Illumina Infinium 450K HumanMethylation array is often used to assess DNA methylation at 485,577 CpG sites. The array is considered a highly suitable platform for large-scale studies, but it still targets only < 2% of the CpG sites present in the human genome. Nonetheless, some rather small EWASs have been conducted on pregnant women with psychiatric symptoms and possible epigenetic alterations in infant cellular function.

In a prospective study on 201 pregnant women suffering from (mainly depressive) psychiatric illness and using various medications, Schroeder et al.[41] found no significant methylation effects across 27,578 CpG sites in the newborn cord blood. However, the authors did find an average methylation rate difference of 3 percent at 2 loci, tumor necrosis factor receptor subfamily 21 (*TNFRSF21*) and cholinergic receptor, nicotinic, α1 (*CHRNA2*), for use of antidepressant medication. In contrast, Non, Binder, Kubzansky, and Michels[42] compared cord blood DNA methylation in newborns of mothers not medicated during pregnancy (*n* = 13), of newborns of mothers using SSRIs during pregnancy (*n* = 22), and of unexposed newborns (*n* = 23), and did not find DNA methylation effects as a result of maternal depression that was treated with SSRIs. On the other hand, non-medicated prenatal depression was associated with 10 differentially methylated CpG sites, most of which had slightly lower DNA methylation rates, compared to non-depressed controls in genes clusters involved in regulation, translation and cell division processes.

Labonté et al.[43] took an epigenome-wide approach in brain tissue, studying DNA methylation of 400K promoters of 25 suicide completers with a history of childhood abuse and of 16 control subjects. They found 362 promoters to be differentially methylated, about two-thirds of which were hypermethylated. In a subsample (13 suicide with abuse and 9 controls), these hypermethylated CpG sites were shown to be related to decreased expression levels. Moreover, it seemed that most of the differentially methylated promoters were in the neuronal, rather than the glial tissue of the hippocampus and that most genes of affected promoters were involved in neuronal plasticity.

Nemoda et al.[44] also studied DNA methylation using the Illumina 450K array, and compared EWAS hits from cord blood with DNA methylation ratios in brain tissue, in children of mothers who had experienced depression. They compared the DNA methylation level of T-cells in cord blood of 15 newborns with mothers with current depression, 14 with mothers with past (but not during pregnancy) depression, and 15 newborns of mothers without any history of depression. Differences of the separate depression groups versus control group were

negligible, but when the two depression groups were taken together and compared with the control group, 145 differentially methylated CpG sites were found. In a comparison of hippocampal tissue of 12 males with a history of maternal depression with 50 males without a history of maternal depression, some genes were found to be differentially methylated in the brain that were also differentially methylated in the cord blood. These genes were often associated with immune function.

One of the largest studies on epigenome-wide DNA methylation patterns in newborns to date (*N* = 912 mother-newborn dyads) was conducted by our research group as part of the Generation R cohort study[45, 46], with a replication in the Avon Longitudinal Study of Parents and Children[47]. The aim of this study was to examine the association between a composite score of prenatal exposure to maternal stress and offspring genome-wide cord blood methylation using meta-analysis, follow-up pathway analyses, and differentially methylated regions (DMRs) analyses. The composite measure of prenatal maternal stress was based on maternal reports at several points in time during pregnancy, covering four stress domains[48]: (i) life stress (e.g., death in family, illness, work problems), (ii) contextual stress (e.g., financial difficulties, housing problems), (iii) personal stress (e.g., psychopathology, substance abuse including alcohol and drugs), and (iv) interpersonal stress (e.g., family relationship difficulties, arguments with partner).

It was remarkable that the large meta-analysis (total *N* = 1,740) across the two studies revealed no epigenome-wide associations of prenatal maternal stress exposure with neonatal differential DNA methylation. Follow-up analyses of the top hits derived from the epigenome-wide meta-analysis indicated an over-representation of the methyltransferase activity pathway. Methyltransferases are important in regulating gene expression and might therefore form an efficient system for feedback regulation of the response to initial environmental pressures and stress might decrease the plasticity of the genomic regulation of protein levels[48]. However, we identified no DMRs associated with prenatal maternal stress exposure. When the two extreme top and bottom 10% scoring respondents on the prenatal stress composite were compared, no significant DNA methylation differences emerged. Three marginally significant DMRs in Generation R were not replicated in ALSPAC. Concluding, combining data from two independent population-based samples in an epigenome-wide meta-analysis, Rijlaarsdam, Pappa, et al.[48] did not find large, replicable effects of prenatal maternal stress exposure on neonatal DNA methylation.

To summarize, candidate epi-gene studies indicate that parenting stress is associated with DNA methylation in the child. However, EWASs do not confirm that methylation of genes such as *NR3C1* is associated with parenting stress and employ rather lenient corrections for multiple comparisons to find associations with methylation of other genes. Here of course null findings trigger a large number of alternative interpretations related to the normalcy of

the samples, the self-reported strains and stresses in specific periods of pregnancy, but fact is that in this study state-of-the-art methods were used, and a built-in replication effort was conducted. Although they might disappoint high but premature expectations of significant hits in earlier, smaller studies such replication efforts are essential in the search for robust associations, whether derived from candidate gene methylation or epigenome-wide studies. This is reason why Rijlaarsdam, Pappa, et al.[48] sub-titled the paper: 'A model approach for replication'. Myriad of problems and pitfalls are inherent to EWAS including limited coverage of the genome and extremely large numbers of tests. In addition previous studies found small effect sizes in small samples without replication in independent samples or animal model systems, which raise concerns regarding the reproducibility of the epigenetic findings in the behavioral sciences.

In summary, it is likely that a global environmental influence such as parenting stress has a global effect on many CpG sites adjacent to many genes, instead of a very localized effect on a few CpG sites. This makes it a challenge to pinpoint where parenting stress exactly affects DNA methylation. Moreover, child development is expected to be influenced by many small, pleiotropic DNA methylation effects. Furthermore, these effects on and of DNA methylation are unlikely to stand alone. Rather, it is expected that they interact with the underlying genetic code. These issues will be discussed below.

Bidirectional Effects of the Genome and Epigenome

When considering literature on the effect of the environment on DNA methylation, one should bear in mind that in some cases, DNA methylation patterns and associations may be allele-specific[49]. Hence, DNA methylation, or the environmental effects on DNA methylation, might be affected by the genome itself. For example, Van der Knaap et al.[50] showed in 939 adolescents that stressful life events were positively associated with methylation of *5-HTTLPR* for those with the protective *ll* variant, but not among those with the *sl*/*ss* variants. Van IJzendoorn et al.[51] reported that methylation of the *5-HTT* gene at *5-HTTLPR* was positively associated with risk of unresolved loss or trauma in the *5-HTTLPR ll* variant but not in the *sl* and *ss* variants in 143 adoptees. The authors observed this gene by DNA methylation interaction in the absence of (epi)genetic main effects, suggesting that opposing associations cancelled each other out. Together, these studies provide suggestive evidence that DNA methylation might be allele-specific, masking or revealing associations between genotype and stress exposure.

Similar to associations between stress exposure and DNA methylation, associations between DNA methylation and psychological outcomes (e.g., emotional and behavioral problems) might be allele-specific. Hence, the effect of DNA methylation on child outcomes should not be seen separately from the genome it acts upon. Ziegler et al.[52] compared *OXTR* methylation

in unmedicated 110 social anxiety patients and matched 110 controls, taking into account *OXTR* rs53576 allelic variation. They showed that *OXTR* methylation was predominant in social anxiety patients carrying the *OXTR* rs53576 A-allele. Similarly, Reiner et al. reported that, in their sample of 43 clinically depressed women and 42 healthy controls, *OXTR* rs53576 clinically depressed A-allele carriers, but not G-allele homozygotes, exhibited significantly increased *OXTR* methylation levels.

In a population-based study on 298 mother-child dyads[53], we showed that cord blood methylation patterns of the *FKBP5* gene, which is involved in hypothalamic-pituitary-adrenal (HPA) axis functioning, increased cortisol reactivity of 14-month old infants. This association was especially present when the infants were also T-allele carriers of rs1360780 *FKBP5, and* when infants had an insecure-resistant attachment to their mother. While the temporal organization of the study did not allow for examination of potential environmental effects on DNA methylation, this Gene × Methylation × Environment (G ×M × E) study does expose some of the complexities that are involved in the study of epigenetics.

In all, we discussed how the association between parenting stress and DNA methylation may be modified by the genetic variance of the child. Furthermore, it seems that the effect of DNA methylation on child outcomes might be dependent on the genetic code as well. We will encounter more epi-allelic effects in the following section, as we discuss studies that take into account the suspected antecedents as well as the consequences of DNA methylation.

DNA Methylation as Mediation

Candidate (Epi-) Genomic Approaches

Whereas studies discussed above imply that the family environment can affect DNA methylation and that DNA methylation may influence child outcomes, studies that incorporate both the presumed precursors as well as the consequences of DNA methylation are needed to confirm that DNA methylation is a true mediator of parenting stress and child development. An early example of this approach is the Oberlander et al.[34] study showing that maternal depressed/anxious prenatal mood was associated with methylation of NGFI-A binding site of the *NR3C1* gene, which was in turn associated with increased salivary cortisol. An important caveat, however, is that no formal mediation testing was conducted, which leaves open whether mediation was only partial or complete.

Using a longitudinal design embedded in in the Avon Longitudinal Study of Parents and Children (ALSPAC), Cecil et al.[54] demonstrated that neonates (N = 84) who were exposed to maternal stress (e.g., maternal psychopathology, criminal behaviors, substance use) in the prenatal period had higher methylation levels of the oxytocin receptor (*OXTR*) gene than

non-exposed neonates. Higher neonatal *OXTR* methylation, in turn, showed temporal stability (from birth to 9 years of age) and was associated with callous-unemotional (CU) traits at age 13 years independent of postnatal stress exposure and associated *OXTR* methylation. Interestingly, these associations were observed exclusively in early-onset persistent (EOP) conduct problem (CP) youth with low internalizing problems versus EOP CP youth with high internalizing problems, suggesting distinct developmental pathways to CU. However, despite this innovative path analytic model that incorporated stress exposure, *OXTR* methylation and CU traits, no formal mediation analysis was presented.

Using data from the Generation R Study, our research group[55], examined *OXTR* rs53576 allele-specific sensitivity for neonatal *OXTR* methylation in relation to both prenatal maternal stress exposure and child autistic traits at age 6 in 743 children. Specifically, we investigated the extent to which prenatal maternal stress exposure was predicted by of *OXTR* methylation variation among neonates, while taking into account *OXTR* rs53576 genotype. In addition, we investigated the extent to which prenatal maternal stress exposure and neonatal *OXTR* methylation combined either additively or interactively with *OXTR* rs53576 genotype to influence child autistic traits. We demonstrated that prenatal maternal stress exposure, but not *OXTR* rs53576 genotype and *OXTR* methylation, showed a main effect on child autistic traits. Because prenatal maternal stress exposure and *OXTR* DNA methylation were unrelated across both *OXTR* rs53576 G-allele homozygous children and A-allele carriers, findings argued against a mediating role of *OXTR* methylation in the association between prenatal maternal stress exposure and child autistic traits. However, we did observe a significant *OXTR* rs53576 genotype by *OXTR* methylation interaction for child autistic traits in general and social communication problems in particular. More specifically, *OXTR* methylation levels were positively associated with social problems for *OXTR* rs53576 G-allele homozygous children but not for A-allele carriers. These results highlight the importance of incorporating epi-allelic information and support the role of both stress exposure and *OXTR* methylation in child autistic traits.

Elevated methylation of the *OXTR* CpG island is expected to decrease gene expression[56] and subsequently levels of circulating oxytocin[57]. Evidence also suggests, that the *OXTR* rs53576 A-allele is a “risk allele” for autistic traits[58-60]. Thus, *OXTR* methylation may decrease the expression of the otherwise protective *OXTR* rs53576 GG-allele and elevate the risk for emotional or behavioral problems. Consequently, one would expect the emotional or behavioral problems of G-allele homozygous children to more closely resemble those of A-allele carriers. Together, these findings suggest that DNA methylation might (i) nullify the effect of the protective allele, resulting in a functionality equivalent to the risk allele or (ii) mask the effect of risk alleles[50-52, 55, 61].

By means of a formal mediation test, another longitudinal study embedded in ALSPAC Rijlaarsdam, Cecil, et al.[62] also highlights the importance of the prenatal environment. The

authors examined, for youth with early-onset persistent (EOP, *n* = 83) versus low conduct problems (CP, *n* = 81), the extent to which high-fat and -sugar diet (prenatal, postnatal) associates with ADHD symptoms (age 7-13) via DNA methylation of the insulin-like growth factor 2 gene (*IGF2*; birth, age 7, collected from blood). Results showed a positive association between prenatal high-fat and –sugar diet with *IGF2* DNA methylation at birth across both EOP and low CP youth. However, only for EOP youth, higher *IGF2* DNA methylation at birth predicted ADHD symptoms. Interestingly, only for EOP youth, the association between prenatal high-fat and –sugar diet and higher ADHD symptoms was mediated by *IGF2* DNA methylation at birth independent of postnatal diet and associated *IGF2* methylation. Together, these studies support ideas focusing on prenatal maternal health as an important risk for postnatal child disease vulnerability[26, 63]. For example, a prenatal maternal high-fat and -sugar diet may alter the DNA methylation status of the *IGF2* gene at birth, which in turn, may increase risk for psychiatric and health disorders –as was illustrated dramatically in the pioneering Dutch Hunger Winter study[23].

IGF2 was also targeted in our prospective Generation R study by Bouwland-Both et al.[64] focusing on the influence of prenatal maternal smoking on newborn birthweight via *IGF2* methylation in 506 newborns. Prenatal maternal smoking should in fact be considered a risky type of prenatal parenting that in the population-based cohort of Generation R was shown by an impressive 25% of the pregnant women who reported on their tobacco smoking habits at three time-points before the birth of their child. Continued maternal prenatal smoking was inversely related to the level of DNA methylation in a differentially methylated region of *IGF2* , in a dose-response manner. A formal mediation test showed that prenatal maternal smoking led to lower birthweight via lower *IGF2* DMR methylation levels, which explained part of the variance in weight (partial mediation). Paternal tobacco smoking did not show a similar cascade of effects[64].

We have seen that postnatal stressors might also leave their traces in epigenetic signatures. Klengel et al.[65] found that trauma in childhood (*n* = 30; versus *n* = 46 controls) was related to *FKBP5* demethylation, which was exclusively the case for T-carriers of the *FKBP5* rs1360780 SNP. Importantly, adult trauma did not seem to be related to *FKBP5* methylation in either the childhood trauma group, or the control group, indicating that it was especially childhood trauma and not later trauma that affected *FKBP5* methylation. Investigating the effects of *FKBP5* methylation, Klengel et al.[65] also found that methylation of the *FKBP5* gene attenuated the response to dexamethasone administration, indicating that methylation of the *FKBP5* gene can affect stress reactivity. This study shows us both sides of the equation: childhood trauma may affect DNA methylation, and DNA methylation might have long term effects on psychobiological functioning. However, no formal test of mediation was conducted.

Demonstrating the feasibility of DNA methylation mediation testing, Beach, Brody, Todorov, Gunter, and Philibert[66] examined in 155 women whether methylation of the *5-HTT* promoter mediated the association between childhood sexual abuse and symptoms of antisocial personality disorder in adulthood, by contrasting models with direct and indirect pathways between the three variables. First, they found that childhood sexual abuse was related to antisocial personality disorder, that childhood sexual abuse was related to *5-HTT* promoter hypermethylation, and that *5-HTT* hypermethylation was associated with antisocial personality disorder. Importantly, in a second step, they showed that a model with a direct path from sexual abuse to antisocial personality disorder differed significantly from a model with only the indirect paths, via *5-HTT* methylation, included. Therefore, it was concluded that the association between childhood sexual abuse and antisocial personality disorder was mediated by *5-HTT* promoter methylation.

In summary, these candidate epi-gene studies substantiate the idea that DNA methylation can be a mediator between parenting stress and child outcomes and that its role is often dependent upon the genetic code itself. In the next paragraph, we will explore whether such candidate epi-gene associations also emerge in EWASs.

Epigenome-Wide Association Studies (EWASs)

In an EWAS on 169 participants with and without PTSD Mehta et al.[67] found that PTSD patients with a history of childhood trauma (n = 32) and PTSD patients without childhood trauma but otherwise matched on adult trauma (n = 29) had dissimilar genome expression profiles, suggesting that converging clinical syndromes can arise from different genetic transcription profiles. Further analysis showed that the PTSD group with child abuse especially had differential DNA methylation in gene expression networks involved in CNS development, amongst others, while the PTSD group without child abuse especially had differential methylation in gene expression networks involved in apoptosis and growth rate. Importantly, the genes associated with these two expression profiles were tested for DNA methylation within each group versus controls (PTSD but no trauma, or trauma but no PTSD, respectively). It was found that much more (up to 12 times) of the variance of the genetic transcripts was explained by variance in DNA methylation in the PTSD group that had experienced childhood trauma than in the PTSD group that had only experienced trauma in adulthood. It seems that childhood abuse may have a long lasting effect on psychosocial functioning, possibly through the effect on DNA methylation (see also McGowan et al.[27]) and that the traumatic experiences associated with the development of PTSD are in particular related to methylomic changes when they happen early in life. However, formal mediation tests were not reported.

In another small EWAS on 83 males who were 60 years or older, Khulan et al.[68] studied DNA methylation differences between participants who were separated from their families for about two years during the Second World War at the age of 5 years, and a group of non-separated

men. Ten years later, a psychological follow-up was performed. Earlier research in the Helsinki Birth Cohort Study already had shown that separated individuals have a higher prevalence of psychological problems, altered cortisol reactivity, and poorer cognitive control[69-71]. However, no difference in DNA methylation was found between separated and non-separated men.

Earlier, we discussed how allelic variation should be taken into account when investigating associations of DNA methylation with child development. In an EWAS, this would of course lead to major statistical power issues. However, Chen et al.[72] did take into account the variation of one particular SNP in their EWAS in the Singaporean GUSTO birth cohort (*N* = 237). In this study, Chen et al. examined the associations between prenatal maternal anxiety, epigenome-wide methylation and neonatal brain volumes, while taking *BDNF* genotype into account. Maternal prenatal anxiety was found to be related to methylation of a SNP dependent way; for infants with methione (Met/Met) genotype, methylation of more CpG sites was related to maternal prenatal anxiety than in infants with Met/valine (Val) and Val/Val genotypes. In a second step, they examined the association between epigenome-wide methylation and neonatal brain volumes. It was found that DNA methylation was associated with the volumetrics of several brain areas, again in a *BDNF* SNP dependent way. Unfortunately, it remains unclear to what extent CpG sites implicated in prenatal maternal anxiety corresponded to the CpG sites related to neonatal brain volumes, thereby precluding strong inferences on the role of DNA methylation as a mediator between prenatal maternal anxiety and neonatal brain volumes.

Altogether, the results from candidate epi-gene studies and EWASs offer support for the notion that epigenetics, in the form of DNA methylation, can mediate the association between parenting stress and child outcomes. Interestingly, genes that appear differentially methylated in candidate epi-gene studies, do not necessarily appear among the hits in the EWASs discussed. One reason for this might be that EWASs are still underpowered to find the effects that are observed in candidate studies. However, this discrepancy might also confirm the idea the hypothesis-driven approach of candidate epi-gene studies creates an 'information bottleneck'. The human DNA contains over 20,000 genes and focusing on the DNA methylation of only a few seems far-fetched. These and other methodological issues, will be elaborated upon in the following section, before coming to a final conclusion.

Caveats and Conclusions

Reliability and Validity of DNA Methylation Measurement

While the number of studies on DNA methylation in developmental and family psychology is increasing, pivotal questions regarding the reliability and validity of DNA methylation indicators in human research remain unanswered. In fact, basic research on these essential characteristics of any adequate measure has been neglected. Several issues should be mentioned here.

First, it is not clear which markers of DNA methylation are stable over what periods of time (trait-like indicators) and which markers can change rapidly depending on momentary endogenous or exogenous changes (state-like markers). For parenting and developmental studies, this is crucial information, as we are mostly interested in influences of parenting on long-term and more persistent, trait-like changes in the child's development. Regarding epigenome-wide array analyses, large parts of the epigenome as assessed by the Illumina approach is stable by definition because it pertains to CpG sites that show no methylation at all or, in contrast, show maximum methylation (with a confidence interval indicating imprecision of measurement) which may inflate epigenomic stability figures. Nevertheless, Lévesque et al.[73] found that more than half of the probes measured with the 450K Illumina were unstable over a 3 to 6 months' time period in young adolescents. In contrast, Wang et al.[74] analyzed the methylome of newborns and found that only 5% of CpG sites made a true shift from methylated to unmethylated, or vice versa, within the first 2 years of life.

CpG sites of interest to developmentalists can potentially vary due to environmental pressures but at the same time should not show short-term volatility. In a small sample of adults we found that at some genes, such as *DRD4* or *5-HTT*, almost all indicators of reliability across time were satisfactory. In contrast, at *BDNF*, many probes showed poor reliability especially in blood spots[75]. Talens et al.[76] found some evidence for stable DNA methylation patterns in peripheral blood over a period of one to two decades in CpG sites of eight genes, of young to middle-aged individuals. Taken together, these results seem to indicate that DNA methylation can be stable over a prolonged period of time, but the disparity in age range, array methods, and definition of temporal stability makes it impossible to draw firm conclusions before more systematic reliability studies become available.

Second, tissue is the issue. The central question for parenting and developmental research is the link between DNA methylation markers derived from peripheral tissue and methylation patterns in behaviorally relevant regions of the brain. Because in humans brain DNA methylation patterns are nearly inaccessible ante mortem, very few studies have looked into the association with peripheral DNA methylation, with somewhat disappointing results. For example, Hannon, Lunnon, Schalkwyk, and Mill[77] examined inter-individual methylomic variation across blood, cortex, and cerebellum and found that the majority of DNA methylation derived from whole blood was not a strong predictor of variation in the brain, although the relation with cortical regions appeared to be stronger than with the cerebellum. DNA methylation of only about 1% of CpG sites were strongly correlated between blood and brain, and about 6% are moderately correlated.

DNA methylation patterns derived from blood may however not be the most valid indicator of methylation in the brain, not only because of the blood-brain barrier but also because of the heterogeneity of cellular composition of blood samples that might be corrected for in

various ways[78]. Buccal tissue has been used rather frequently as a source of information about methylation levels because it is less invasive compared to blood sampling. Buccal epithelial tissue has been argued to be less heterogeneous than blood cells and to be 'closer' to brain tissue in a developmental sense[79]. Without going into technical details, we found better test-retest reliability figures for DNA methylation levels established in buccal cells than for blood or blood spots[75]. Of course, some part of DNA methylation stability may depend on the heritability of DNA methylation levels as suggested by rather strong associations between some genotypes and DNA methylation level[80] but it would be premature to conclude that DNA methylation is in fact determined by variations in structural DNA and thus potentially completely genotypic instead of (endo-)phenotypic.

Reproducibility

In the behavioral and biomedical sciences, the problem of reproducibility of research has been discussed quite vigorously, starting with the (in-)famous Ioannidis[81] paper on 'Why most published research findings are false'. The replication controversies around candidate G × E studies seem to be still ongoing[82-85] with considerable emphasis on the need for large sample sizes and built-in replication or meta-analysis efforts, parallel to current practices in the GWAS area. For DNA methylation studies based on selection of one or few candidate genes for which DNA methylation data are collected, the issue of statistical power, and thus the problem of reproducibility may be aggravated because the number of tests might increase considerably compared to the few tests with bi-allelic candidate genes. The advantage, though, which also comes up in the candidate gene versus GWAS debate, is the theory-guided hypothesis testing approach for which the functionality of hyper- or hypo-methylation might already have been established[67]. Factor analysis to reduce the number of CpG sites to a few dimensions might also be helpful[53, 62]. Last but not least, meta-analyses to synthesize a large number of small candidate gene methylation studies are indispensable.

Epigenome-wide association studies (EWASs) suffer from power issues similar to the hypothesis-free approach of GWAS, and certainly even more so in comparison to candidate epigenetic approaches, simply because of the large number of CpG sites interrogated by the standard Illumina 450K, a problem that might become exacerbated by the new Illumina 850K chip. Alternative methods of significance testing (e.g., permutation testing) are important as well as robust statistical corrections for multiple testing, and analyses that account for dependence of CpG sites in differentially methylated regions[86] and through bump hunter[87, 88] or block finder[89]. Nevertheless, EWAS results will be difficult to replicate because of the small effect sizes to be expected, and the large number of tests on a relatively small number of subjects[48]. Therefore, replication of EWAS findings in independent studies are needed before more definite conclusions can be reached and large consortia like the Pregnancy And Childhood Epigenetics (PACE) consortium are badly needed. As an example, in the PACE consortium with 16 cohort studies Joubert et al.[90] identified more than 6,000 differentially methylated

CpG sites in newborns in relation to prenatal maternal smoking, with nearly half of the sites not previously associated with smoking and DNA methylation in either newborns or adults.

Causality

The large majority of human DNA methylation studies are correlational–whether conducted with a retrospective, concurrent or prospective design with regard to timing of assessment of biological tissue, and the predictors and outcomes of interest. This design is beset with problems of confounding, spurious associations because of unmeasured third factors, and reversed causality. In this respect, epigenetics is not different from any other epidemiological approach[91]. One of the statistical means to address the question of causality is through mediation modeling, preferably based on longitudinal study designs[92, 93]. Full mediation provides insights into the cascade from environmental determinants through mediating DNA methylation changes to some developmental outcome. In the foregoing sections we emphasized the need for mediation analyses to shed light on the role of DNA methylation in the association between environmental input and behavioral output, but only few studies successfully probed this mediation mechanism. DNA methylation should be considered a mechanism instead of endpoint of child development determined by prenatal adversities or the ultimate determinant of later developmental outcomes. Crucial is the combination of stressful (prenatal) parenting influencing child development mediated by DNA methylation.

Of course, for inferring causality, no design can beat real experiments with randomized assignment of subjects to experimental manipulation and control group (for an excellent example on nonhuman primates, see Provençal et al.[94]). The number of quasi-experimental studies on DNA methylation is increasing as methylation signatures have been used as indicators of symptom improvement in psychotherapy of combat veterans with PTSD[95], as markers of therapeutic success in a matched-controls design with clinically depressed in-patients[61] or in a pre- posttest only design[96], in studies of cognitive behavior therapy with anxious individuals[96, 97], and in studies of foster care based on a randomized control design turned into a correlational approach[98].

These are important exploratory advances in the field of therapygenetics[96] because they suggest that DNA methylation may not only change for the worse, but with the right intervention, also for the better. At the least, these preliminary results do suggest the usefulness of DNA methylation as a biomarker, with which one might be able to gauge responsiveness to therapy. However, the quasi-experiments need to be followed by genuine experiments similar to randomized G × E interventions. We initiated experimental human candidate G × E research more than 10 years ago[99] and recently showed that the statistical power of experimental candidate G × E is some 13 times higher than the regular correlational candidate G × E studies[100]. Such experiments also adequately prevent gene-environment correlations from confounding the G × E results. The same might be true for DNA methylation studies in which DNA methylation

changes are considered proximal outcomes or mediators of subsequent behavioral changes as a consequence of the intervention.

DNA methylation may be a crucial component of genetic differential susceptibility that explains why interventions usually show disappointingly small effects on child development. Differential susceptibility theory suggests that some children are more susceptible to the environment, for better *and* for worse, than their peers without a susceptible genotype. We found that dopamine- and serotonin-related genes are involved in differential susceptibility to parenting and speculated that DNA methylation might play a role in G × E interactions leading to Gene × Methylation × Environment interplay[100] (G × M × E). DNA methylation may modulate adaptation to a changing environment and make the organism less dependent on its structural DNA. A prime example is the crucial role of DNA methylation in genetically identical apomictic dandelions that adapt to a great variety of ecological niches only due to epigenetic changes superimposed on an asexually inherited fixed genotype[101]. Genetic markers of differential susceptibility in humans might in part reflect allelic differences coding for degree of plasticity of DNA methylation that make some individuals less adaptive to adverse (prenatal) circumstances and thus dispose them to sub-optimal development, whereas these same individuals may benefit more from benign environments compared to their peers with more flexible epi-allelic characteristics.

Conclusion and Future Directions

We presented some evidence for the exciting hypothesis that distressed parenting can affect DNA methylation of the offspring, which in turn through regulating the expression of genes may influence behavioral development. DNA methylation is one of the epigenetic mechanisms that holds great promise for the unification of the fields of nature- and nurture-centered research. We also argued, however, that research on behavioral epigenetics in humans often fails to ascertain the reproducibility of its results, using unreliable and invalid measures and samples that are too small, often also failing to address the question of mediation and causality. We did not touch on the million-dollar question of the transgenerational transmission (see also Mileva-Seitz and Fleming[28]) of DNA methylation –epigenetic heritability through the germline– potentially explaining the transmission of (abusive or positive) parenting across generations. Whereas in some plant species[101, 102] and in rodents[103] this transgenerational transmission of epigenetic signatures has been experimentally illustrated, in humans it still is one of the most challenging and outstanding issues to be addressed in a rigorous manner. For a Lamarckian cry of victory it is way too early.

Looking ahead, whereas many questions regarding DNA methylation specifically and epigenetics in general have been left unanswered, its possible applications are titillating. Since the research discussed in the current Chapter seems to indicate that stressful parenting can affect the epigenetic profile of the developing child in a detrimental way, one wonders whether positive

parenting or intervention might influence the child's epigenetics beneficially. In a study in adult rats performed by Weaver et al.[104], methyl supplementation was able to reverse *NR3C1* methylation and stress responses induced by experienced maternal stress early in life. Also, in a study performed by Roberts et al.[97], it was shown that children with anxiety disorders who responded well to cognitive behavior therapy, had an increase in DNA methylation of a CpG site upstream of *5-HTT*, whereas methylation of this CpG site decreased in non-responding children. Even though we cannot be certain that the change in DNA methylation is a functional one, its possible use as a diagnostic tool is intriguing. We are evidently a long way from using epigenetics validly in a therapeutic setting, but a continuous investment in epigenetic research may bring us closer to understanding the intricacies of the interplay of genes, the environment, and the developing child.

References

1. Caspi A, Sugden K, Moffitt TE, et al. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science.* 2003;301:386-9.
2. Bakermans-Kranenburg MJ, van IJzendoorn MH, Pijlman FTA, Mesman J, Juffer F. Experimental evidence for differential susceptibility: dopamine D4 receptor polymorphism (DRD4 VNTR) moderates intervention effects on toddlers' externalizing behavior in a randomized controlled trial. *Developmental Psychology.* 2008;44:293.
3. Meaney MJ. Epigenetics and the biological definition of gene X environment interactions. *Child Devevelopment.* 2010;81:41-79.
4. Waddington CH. *Principles of Embryology.* 1956. In: London, Allen & Unwin Ltd, 2017.
5. Waddington CH. *An Introduction to Modern Genetics.* 1939.
6. Gottlieb G, Halpern CT. A relational view of causality in normal and abnormal development. *Development and Psychopathology.* 2002;14:421-35.
7. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics.* 2010;11:204-20.
8. Klose RJ, Bird AP. Genomic DNA methylation: the mark and its mediators. *Trends in Biochemicals.* 2006;31:89-97.
9. Li G, Reinberg D. Chromatin higher-order structures and gene regulation. *Current Opinion in Genetics & Development.* 2011;21:175-86.
10. Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Research.* 2011;21:381-95.
11. Berger A. *Individual Differences in Self-Regulation.* Berger, Andrea 2011:61-90.
12. Stefani G, Slack FJ. Small non-coding RNAs in animal development. *Nature Reviews Molecular Cell Biology.* 2008;9:219-30.
13. Holliday R, Pugh JE. DNA modification mechanisms and gene activity during development. *Science.* 1975;187:226-32.
14. Meissner A, Mikkelsen TS, Gu H, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells*. Nature.* 2008;454:766-70.
15. Chow J, Heard E. X inactivation and the complexities of silencing a sex chromosome. *Current Opinion in Cell Biology.* 2009;21:359-66.
16. Riggs AD. X inactivation, differentiation, and DNA methylation. *Cytogenetic and Genome Research.* 1975;14:9-25.
17. Esteller M, Herman JG. Cancer as an epigenetic disease: DNA methylation and chromatin alterations in human tumours. *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland.* 2002;196:1-7.
18. Aran D, Sabato S, Hellman A. DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes*. Genome Biology.* 2013;14:1-14.
19. Moore LD, Le T, Fan G. DNA methylation and its basic function. *Neuropsychopharmacology.* 2013;38:23-38.

20. Weaver ICG, Szyf M, Meaney MJ. From maternal care to gene expression: DNA methylation and the maternal programming of stress responses. *Endocrine Research.* 2002;28:699-0.
21. Weaver ICG, Cervoni N, Champagne FA, et al. Epigenetic programming by maternal behavior. *Nature Neuroscience.* 2004;7:847-54.
22. Lumey LH, Stein AD, Kahn HS, et al. Cohort profile: the Dutch Hunger Winter families study. *International Journal of Epidemiology.* 2007;36:1196-204.
23. Heijmans BT, Tobi EW, Stein AD, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proceedings of the National Academy of Sciences.* 2008;105:17046-9.
24. Tobi EW, Goeman JJ, Monajemi R, et al. DNA methylation signatures link prenatal famine exposure to growth and metabolism. *Nature Communications.* 2014;5:1-14.
25. Neuenschwander R, Oberlander TF. Developmental origins of self-regulation: prenatal maternal stress and psychobiological development during childhood. *Parental Stress and Early Child Development: Springer*, 2017:127-56.
26. Barker DJ. The fetal and infant origins of adult disease. *BMJ: British Medical Journal* 1990;301:1111.
27. McGowan PO, Sasaki A, D'Alessio AC, et al. Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nature Neuroscience.* 2009;12:342-8.
28. Mileva-Seitz VR, Fleming AS. How being mothered affects the development of mothering. *Parental Stress and Early Child Development: Springer*, 2017:235-59.
29. Yehuda R, Daskalakis NP, Lehrner A, et al. Influences of maternal and paternal PTSD on epigenetic regulation of the glucocorticoid receptor gene in Holocaust survivor offspring. *American Journal of Psychiatry.* 2014;171:872-80.
30. Jacobson SW, Chiodo LM, Sokol RJ, Jacobson JL. Validity of maternal report of prenatal alcohol, cocaine, and smoking in relation to neurobehavioral outcome. *Pediatrics.* 2002;109:815-25.
31. Huizink AC, Mulder EJH, Buitelaar JK. Prenatal stress and risk for psychopathology: specific effects or induction of general susceptibility? *Psychological Bulletin.* 2004;130:115.
32. Hompes T, Izzi B, Gellens E, Morreels M. Investigating the influence of maternal cortisol and emotional state during pregnancy on the DNA methylation status of the glucocorticoid receptor gene (NR3C1). *Journal of Psychiatric Research.* 2013;47:880-891.
33. Hompes T, Izzi B, Gellens E, et al. Corrigendum to "Investigating the influence of maternal cortisol and emotional state during pregnancy on the DNA methylation status of the glucocorticoid receptor gene (NR3C1) promoter region in cord blood" [J Psychiatr Res 47 880-91]. *Journal of Psychiatric Research.* 2014;56:165-7.
34. Oberlander TF, Weinberg J, Papsdorf M, Grunau R, Misri S, Devlin AM. Prenatal exposure to maternal depression, neonatal methylation of human glucocorticoid receptor gene (NR3C1) and infant cortisol stress responses. *Epigenetics.* 2008;3:97-106.
35. Radtke KM, Ruf M, Gunter HM, et al. Transgenerational impact of intimate partner violence on methylation in the promoter of the glucocorticoid receptor. *Translational Psychiatry.* 2011;1:e21.

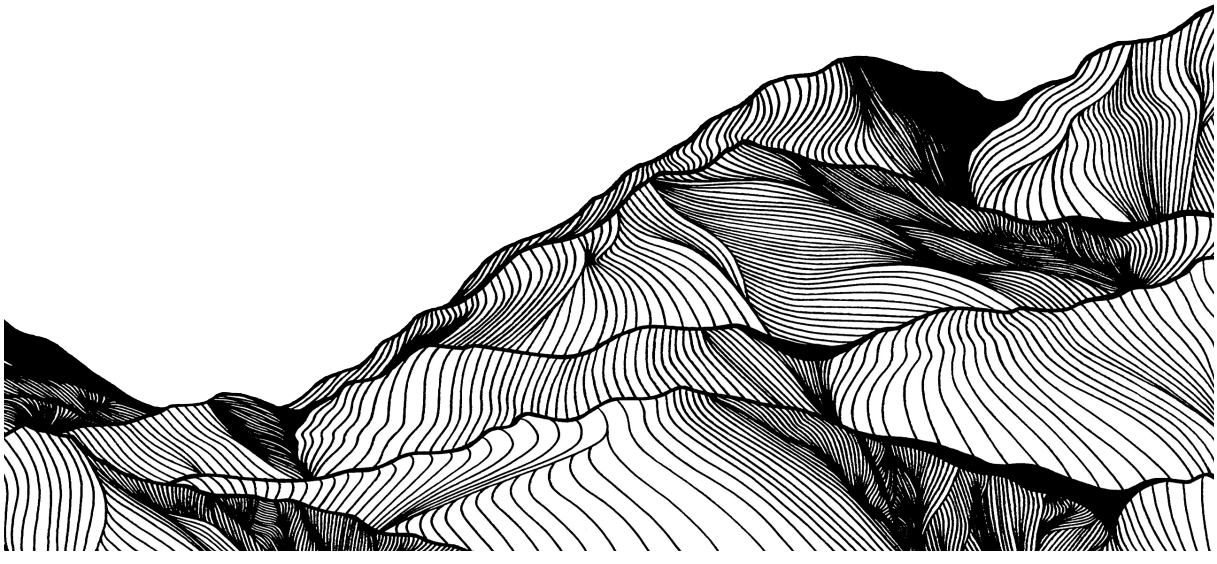
36. Braithwaite EC, Kundakovic M, Ramchandani PG, Murphy SE, Champagne FA. Maternal prenatal depressive symptoms predict infant NR3C1 1F and BDNF IV DNA methylation. *Epigenetics.* 2015;10:408-17.
37. Devlin AM, Brain U, Austin J, Oberlander TF. Prenatal exposure to maternal depressed mood and the MTHFR C677T variant affect SLC6A4 methylation in infants at birth. *PLoS ONE* 2010;5.
38. Vijayendran M, Beach SR, Plume JM, Brody GH, Philibert RA. Effects of genotype and child abuse on DNA methylation and gene expression at the serotonin transporter. *Frontiers in Psychiatry.* 2012;3:55.
39. Unternaehrer E, Meyer AH, Burkhardt SC, et al. Childhood maternal care is associated with DNA methylation of the genes for brain-derived neurotrophic factor (BDNF) and oxytocin receptor (OXTR) in peripheral blood cells in adult men and women. *Stress* 2015:1-11.
40. Zhu M, Zhao S. Candidate gene identification approach: progress and challenges. *International Journal of Biological Sciences.* 2007;3:420.
41. Schroeder JW, Smith AK, Brennan PA, et al. DNA methylation in neonates born to women receiving psychiatric care. *Epigenetics.* 2012;7:409-14.
42. Non AL, Binder AM, Kubzansky LD, Michels KB. Genome-wide DNA methylation in neonates exposed to maternal depression, anxiety, or SSRI medication during pregnancy. *Epigenetics.* 2014;9:964-72.
43. Labonté B, Suderman M, Maussion G, et al. Genome-wide methylation changes in the brains of suicide completers. *American Journal of Psychiatry.*2013;170:511-20.
44. Nemoda Z, Massart R, Suderman M, et al. Maternal depression is associated with DNA methylation changes in cord blood T lymphocytes and adult hippocampi. *Translational Psychiatry.* 2015;5:e545-e.
45. Jaddoe VW, van Duijn CM, Franco OH, et al. The Generation R Study: design and cohort update 2012. *European Journal of Epidemiology.* 2012;27:739-56.
46. Kruithof CJ, Kooijman MN, van Duijn CM, et al. The generation R study: biobank update 2015. *European Journal of Epidemiology.* 2014;29:911-27.
47. Fraser A, Macdonald-Wallis C, Tilling K, et al. Cohort Profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *International Journal of Epidemiology.* 2013;42:97-110.
48. Rijlaarsdam J, Pappa I, Walton E, et al. An epigenome-wide association meta-analysis of prenatal maternal stress in neonates: A model approach for replication. *Epigenetics.* 2016;11:140-9.
49. Meaburn EL, Schalkwyk LC, Mill J. Allele-specific methylation in the human genome: implications for genetic studies of complex disease. *Epigenetics.* 2010;5:578-82.
50. van der Knaap LJ, Riese H, Hudziak JJ, et al. Adverse life events and allele-specific methylation of the serotonin transporter gene (SLC6A4) in adolescents: the TRAILS study. *Psychosomatic Medicine.* 2015;77:246-55.
51. van IJzendoorn MH, Caspers K, Bakermans-Kranenburg MJ, Beach SRH, Philibert R. Methylation matters: Interaction between methylation density and serotonin transporter genotype predicts unresolved loss or trauma. *Biological Psychiatry.* 2010;68:405-7.
52. Ziegler C, Dannlowski U, Bräuer D, et al. Oxytocin receptor gene methylation: converging multilevel evidence for a role in social anxiety. *Neuropsychopharmacology.* 2015;40:1528-38.

53. Mulder RH, Rijlaarsdam J, Luijk MPCM, et al. Methylation matters: FKBP506 binding protein 51 (FKBP5) methylation moderates the associations of FKBP5 genotype and resistant attachment with stress regulation. *Development and Psychopathology.* 2017;29:491.
54. Cecil CAM, Lysenko LJ, Jaffee SR, et al. Environmental risk, Oxytocin Receptor Gene (OXTR) methylation and youth callous-unemotional traits: a 13-year longitudinal study. *Molecular Psychiatry.* 2014.
55. Rijlaarsdam J, van IJzendoorn MH, Verhulst FC, et al. Prenatal stress exposure, oxytocin receptor gene (OXTR) methylation, and child autistic traits: the moderating role of OXTR rs53576 genotype. *Autism Research.* 2017;10:430-8.
56. Kusui C, Kimura T, Ogita K, et al. DNA methylation of the human oxytocin receptor gene promoter regulates tissue-specific gene suppression. *Biochemical and Biophysical Research Communications.* 2001;289:681-6.
57. Dadds MR, Moul C, Cauchi A, et al. Methylation of the oxytocin receptor gene and oxytocin blood levels in the development of psychopathy. *Devevlopment and Psychopathology.* 2014;26:33-40.
58. Wermter AK, Kamp-Becker I, Hesse P, Schulte-Körne G, Strauch K, Remschmidt H. Evidence for the involvement of genetic variation in the oxytocin receptor gene (OXTR) in the etiology of autistic disorders on high-functioning level. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics.* 2010;153:629-39.
59. Liu X, Kawamura Y, Shimada T, et al. Association of the oxytocin receptor (OXTR) gene polymorphisms with autism spectrum disorder (ASD) in the Japanese population. *Journal of Human Genetics.* 2010;55:137-41.
60. Wu S, Jia M, Ruan Y, et al. Positive association of the oxytocin receptor gene (OXTR) with autism in the Chinese Han population. *Biological Psychiatry.* 2005;58:74-7.
61. Reiner I, Van IMH, Bakermans-Kranenburg MJ, Bleich S, Beutel M, Frieling H. Methylation of the oxytocin receptor gene in clinically depressed patients compared to controls: The role of OXTR rs53576 genotype. *Journal of Psychiatric Research.* 2015;65:9-15.
62. Rijlaarsdam J, Cecil CAM, Walton E, et al. Prenatal unhealthy diet, insulin-like growth factor 2 gene (IGF 2) methylation, and attention deficit hyperactivity disorder symptoms in youth with early-onset conduct problems. *Journal of Child Psychology and Psychiatry.* 2017;58:19-27.
63. Barker DJP. The developmental origins of adult disease. *Journal of the American College of Nutrition.* 2004;23:588S-95S.
64. Bouwland-Both MI, van Mil NH, Tolhoek CP, et al. Prenatal parental tobacco smoking, gene specific DNA methylation, and newborns size: the Generation R study. *Clinical Epigenetics.* 2015;7:83.
65. Klengel T, Mehta D, Anacker C, et al. Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. *Nature Neuroscience.* 2013;16:33-41.
66. Beach SRH, Brody GH, Todorov AA, Gunter TD, Philibert RA. Methylation at 5HTT mediates the impact of child sex abuse on women's antisocial behavior: An examination of the iowa adoptee sample. *Psychosomatic Medicine.* 2011;73:83-7.
67. Mehta D, Klengel T, Conneely KN, et al. Childhood maltreatment is associated with distinct genomic and epigenetic profiles in posttraumatic stress disorder. *Proceedings of the National Academy of Sciences.* 2013;110:8302-7.

68. Khulan B, Manning JR, Dunbar DR, et al. Epigenomic profiling of men exposed to early-life stress reveals DNA methylation differences in association with current mental state. *Translational Psychiatry.* 2014;4.
69. Pesonen A-K, Eriksson JG, Heinonen K, et al. Cognitive ability and decline after early life stress exposure. *Neurobiology of Aging.* 2013;34:1674-9.
70. Pesonen A-K, Räikkönen K, Feldt K, et al. Childhood separation experience predicts HPA axis hormonal responses in late adulthood: a natural experiment of World War II. *Psychoneuroendocrinology.* 2010;35:758-67.
71. Räikkönen K, Lahti M, Heinonen K, et al. Risk of severe mental disorders in adults separated temporarily from their parents in childhood: the Helsinki birth cohort study. *Journal of Psychiatric Research.* 2011;45:332-8.
72. Chen L, Pan H, Tuan TA, et al. Brain-derived neurotrophic factor (BDNF) Val66Met polymorphism influences the association of the methylome with maternal anxiety and neonatal brain volumes. *Development and Psychopathology.* 2015;27:137-50.
73. Lévesque ML, Casey KF, Szyf M, et al. Genome-wide DNA methylation variability in adolescent monozygotic twins followed since birth. *Epigenetics.* 2014;9:1410-22.
74. Wang D, Liu X, Zhou Y, et al. Individual variation and longitudinal pattern of genome-wide DNA methylation from birth to the first two years of life. *Epigenetics.* 2012;7:594-605.
75. Forest M, O'Donnell KJ, Voisin G, et al. Agreement in DNA methylation levels from the Illumina 450K array across batches, tissues, and time. *Epigenetics.* 2018;13:19-32.
76. Talens RP, Boomsma DI, Tobi EW, et al. Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. *The FASEB Journal.* 2010;24:3135-44.
77. Hannon E, Lunnon K, Schalkwyk L, Mill J. Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics.* 2015;10:1024-32.
78. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics.* 2012;13:86.
79. Lowe R, Gemma C, Beyan H, et al. Buccals are likely to be a more informative surrogate tissue than blood for epigenome-wide association studies. *Epigenetics.* 2013;8:445-54.
80. Gaunt TR, Shihab HA, Hemani G, et al. Systematic identification of genetic influences on methylation across the human life course. *Genome Biology.* 2016;17:1-14.
81. Ioannidis JPA. Why most published research findings are false. *PLoS Medicine.* 2005;2:e124.
82. Bosker FJ, Hartman CA, Nolte IM, et al. Poor replication of candidate genes for major depressive disorder using genome-wide association data. *Molecular Psychiatry.* 2011;16:516-32.
83. Caspi A, Hariri AR, Holmes A, Uher R, Moffitt TE. Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. *American Journal of Psychiatry.* 2010;167:509-27.
84. Duncan LE, Keller MC. A critical review of the first 10 years of candidate gene-by-environment interaction research in psychiatry. *American Journal of Psychiatry.* 2011;168:1041-9.

85. Gage SH, Davey Smith G, Ware JJ, Flint J, Munafo MR. G= E: What GWAS can tell us about the environment. *PLoS Genetics*. 2016;12:e1005765.
86. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. *Nature Reviews Genetics*. 2011;12:529-41.
87. Choudalakis G. On hypothesis testing, trials factor, hypertests and the BumpHunter. arXiv preprint *arXiv:*1101.0390 2011.
88. Jaffe AE, Murakami P, Lee H, et al. Bump hunting to identify differentially methylated regions in epigenetic epidemiology studies. *International Journal of Epidemiology*. 2012;41:200-9.
89. Hansen KD, Timp W, Bravo HC, et al. Increased methylation variation in epigenetic domains across cancer types. *Nature Genetics.* 2011;43:768.
90. Joubert BR, Felix JF, Yousefi P, et al. DNA methylation in newborns and maternal smoking in pregnancy: genome-wide consortium meta-analysis. *The American Journal of Human Genetics.* 2016;98:680-96.
91. Mill J, Heijmans BT. From promises to practical strategies in epigenetic epidemiology. *Nature Reviews Genetics.* 2013;14:585-94.
92. MacKinnon DP, Fairchild AJ, Fritz MS. Mediation analysis. *Annual Reviews of Psychology.* 2007;58:593-614.
93. Selig JP, Preacher KJ. Mediation models for longitudinal data in developmental research. *Research in Human Development.* 2009;6:144-64.
94. Provençal N, Suderman MJ, Guillemin C, et al. The signature of maternal rearing in the methylome in rhesus macaque prefrontal cortex and T cells. *Journal of Neuroscience.* 2012;32:15626-42.
95. Yehuda R, Daskalakis NP, Desarnaud F, et al. Epigenetic biomarkers as predictors and correlates of symptom improvement following psychotherapy in combat veterans with PTSD. *Frontiers in Psychiatry.* 2013;4.
96. Eley TC, Hudson JL, Creswell C, et al. Therapygenetics: the 5HTTLPR and response to psychological therapy. *Molecular Psychiatry.* 2012;17:236-7.
97. Roberts S, Lester KJ, Hudson JL, et al. Serotonin tranporter methylation and response to cognitive behaviour therapy in children with anxiety disorders. *Translational Psychiatry.* 2014;4:e444-e.
98. Non AL, Hollister BM, Humphreys KL, et al. DNA methylation at stress-related genes is associated with exposure to early life institutionalization. *American Journal of Physical Anthropology.* 2016;161:84-93.
99. Bakermans-Kranenburg MJ, van IJzendoorn MH. Gene-environment interaction of the dopamine D4 receptor (DRD4) and observed maternal insensitivity predicting externalizing behavior in preschoolers. *Developmental Psychobiology: The Journal of the International Society for Developmental Psychobiology.* 2006;48:406-9.
100. Bakermans-Kranenburg MJ, van IJzendoorn MH. The hidden efficacy of interventions: Gene× environment experiments from a differential susceptibility perspective. *Annual Review of Psychology.* 2015;66:381-409.
101. Verhoeven KJF, Jansen JJ, Van Dijk PJ, Biere A. Stress-induced DNA methylation changes and their heritability in asexual dandelions. *New Phytologist.* 2010;185:1108-18.

102. Holeski LM, Jander G, Agrawal AA. Transgenerational defense induction and epigenetic inheritance in plants. *Trends in Ecology & Evolution.* 2012;27:618-26.
103. Bohacek J, Farinelli M, Mirante O, et al. Pathological brain plasticity and cognition in the offspring of males subjected to postnatal traumatic stress. *Molecular Psychiatry.* 2015;20:621-31.
104. Weaver ICG, Champagne FA, Brown SE, et al. Reversal of maternal programming of stress responses in adult offspring through methyl supplementation: altering epigenetic marking later in life. *Journal of Neuroscience.* 2005;25:11045-54.

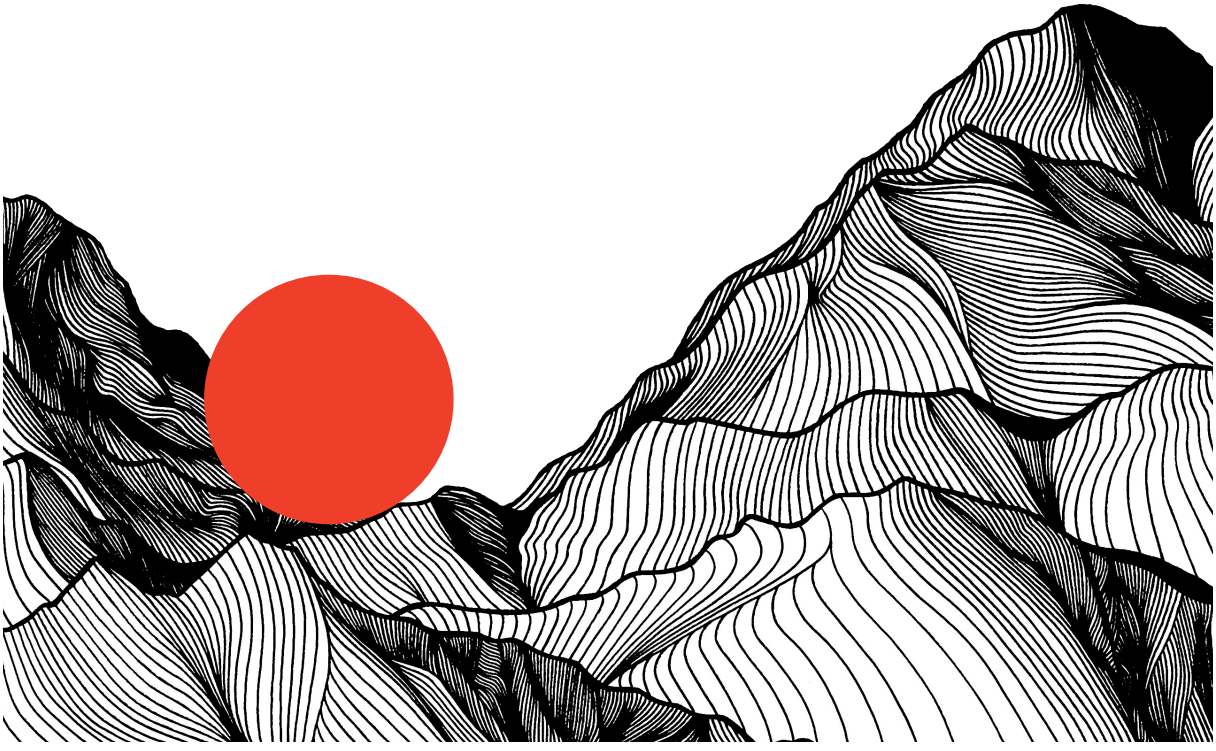


CHAPTER III

Methylation matters: FK506 binding protein 51 (*FKBP5*) methylation moderates the associations of *FKBP5* genotype and resistant attachment with stress regulation

Rosa H. Mulder, Jolien Rijlaarsdam, Maartje P. C. M. Luijk, Frank C. Verhulst, Janine F. Felix, Henning Tiemeier, Marian J. Bakermans-Kranenburg, & Marinus H. van IJzendoorn

Development and Psychopathology, 29(2), (2017) 491.



Abstract

The parent-child attachment relationship plays an important role in the development of the infant's stress regulation system. However, genetic and epigenetic factors such as FK506 binding protein 51 (*FKBP5*) genotype and DNA methylation have also been associated with hypothalamic-pituitary-adrenal (HPA) axis functioning. In the current study, we examined how parent-child dyadic regulation works in concert with genetic and epigenetic aspects of stress regulation. We study the associations of attachment, extreme maternal insensitivity, *FKBP5* SNP 1360780, and *FKBP5* methylation, with cortisol reactivity to the Strange Situation Procedure in 298 14-month-old infants. Results indicate that *FKBP5* methylation moderates the associations of *FKBP5* genotype and resistant attachment with cortisol reactivity. We conclude that the inclusion of epigenetics in the field of developmental psychopathology may lead to a more precise picture of the interplay between genetic make-up and parenting in shaping stress reactivity.

Introduction

The attachment relationship between infant and parent is important in shaping the development of the child's stress regulation system[1]. In the first year of life, human infants are dependent on protective caregivers to regulate their temperature, food and fluid intake, and also to regulate stress in the face of threats and dangers[2]. Sensitive parents, who promptly and adequately respond to their infants' distress signals, help to create a safe haven from which the child can freely explore the environment[3]. These infants are more likely to develop a secure attachment relationship and the associated expectation that, in times of need, their parent will be available to protect them[4]. Insensitive parents, however, may be less prompt and effective in buffering stressful events and settings for their infant. In turn, their infants will be less likely to develop trust and the expectancy of reassuring parental support in times of illness, threat, anxiety and other stressful situations. These infants are also more likely to develop an insecure attachment relationship and a more tenuous style of coping with stressors, potentially resulting in a more reactive hormonal stress system[5]. Stress regulation takes place via the hypothalamic-pituitary-adrenal axis (HPA axis) and one of the crucial hormones involved is cortisol. Therefore, cortisol reactivity to stressors is usually considered to be a measure of the amount of stress experienced by children when confronted with challenges such as separation from the parent or entering an unknown environment or meeting with a stranger[6, 7].

Extreme insensitivity of a parent, including displays of fright because of memories of traumatic experiences, or other threatening behaviors towards the infant, like physical abuse, may elicit even more disturbed attachment behaviors. In particular, extreme parental insensitivity or otherwise frightening behaviors may lead to disorganized/disoriented attachment, reflected in infant behavior, for example, in prolonged stilling, rapid approach–avoidance vacillation, sudden unexplained affect changes, severe distress followed by avoidance, or expressions of fear or disorientation upon return of a parent who has been away for a couple of minutes. Disorganized attachments are overrepresented in clinical samples and in samples with a high prevalence of child maltreatment and family violence[7-9]. Dysregulation of the hormonal stress system has been noted in infants with a disorganized attachment relationship to the parent[10, 11].

In the current study, we examined how attachment and extreme insensitivity interact with infants' stress-related genetics to explain variability in their stress regulation. Specifically, we focus on the FK506 binding protein 51 (*FKBP5*) gene. *FKBP5* has been shown to impede negative feedback of the HPA axis[12], and variants, amongst which the rs1360780 SNP in the *FKBP5* gene has been related to recovery from psychosocial stress[13]. Moreover, it was found that rs1360780 interacts with child abuse in the prediction of later development of posttraumatic stress disorder[14, 15] and of attempt of suicide after childhood trauma[16]. In a subsample with European ancestry of the Generation R cohort, we previously found that rs1360780 interacts with variations in attachment quality in the prediction of stress reactivity.

More specifically, infants with an insecure-resistant attachment to their mother—but not those with an insecure-disorganized attachment—had heightened cortisol reactivity to a mildly stressful situation (the Strange Situation Procedure, SSP[4]), especially if these children were carriers of the T allele in the rs1360780 SNP[17]. Here we aim at extending our previous study in the Generation R subsample, by including extreme maternal insensitivity as an indicator of atypical parental caregiving behavior, as well as by taking DNA methylation into account.

Epigenetics is a relatively new avenue in the field of developmental psychopathology. One of the most often studied epigenetic processes in cohort studies is DNA methylation, where a methyl group attaches to a cytosine nucleotide located next to a guanine in the DNA at a so-called CpG (cytosine-phosphate-guanine) site. Methylation can change the three-dimensional formation of the chromatin[18], and subsequently affect gene transcription. DNA methylation is thought to be influenced by prenatal[19-21] and postnatal life events[22-24], as well as by genetic background. It can therefore be seen as the dynamic interface between genes and the environment[25, 26]. These genotype-by-methylation patterns may in turn affect associations between environmental factors and developmental outcomes[27]. Hence, SNP associations with phenotypes such as stress reactivity may become more clearly apparent when DNA methylation is included in the analysis.

In rodents, maternal separation has been related to differential DNA methylation in a variety of HPA-axis related genes and altered stress-responsiveness[24, 28, 29]. In humans, similar results have been found. For example, in individuals who were adopted after stressful early life experiences, the short variant of the serotonin-transporter-linked polymorphic region predicted more unresolved loss or trauma, but only if methylation was low[27]. Another study showed that prenatal exposure to maternal depressed mood was associated with nuclear receptor subfamily 3, group C, member 1 (*NR3C1*) gene methylation, which was in turn related to increased cortisol reactivity in 3-month-old infants[30]. The *NR3C1* gene codes for the glucocorticoid receptor (GR), and methylation is presumed to impede transcription of the *NR3C1* gene into the GR protein, decreasing HPA-axis negative feedback through corticosteroid binding.

For the *FKBP5* gene, which is associated with the binding of cortisol to the GR, Klengel et al.[15] found that experienced early trauma was related to methylation of *FKBP5*, especially in carriers of the rs1360780 T-allele. The T-allele of rs1360780 facilitates gene transcription, which would lead to less sensitive GRs and ultimately to more or prolonged cortisol reactivity. Functionally, Klengel et al.[15] showed that *FKBP5* methylation affected cortisol reactivity as well and, in a separate sample, they found that GR sensitivity was especially affected in T-carriers of rs1360780 that had also experienced childhood abuse. Although these findings are elucidating, we do not know whether they generalize to the general population, where early traumatic experiences are relatively uncommon. Paquette et al.[31] analyzed placental samples of the general population and infant neurodevelopment. They found an rs1360780

dependent effect of methylation of *FKBP5* on mRNA expression in placental cells. Moreover, higher levels of placental *FKBP5* methylation were found to be related to more arousal in 3-year-olds. However, it should be noted that arousal does not necessarily equate to cortisol regulation.

The goal for this report was to further explore the relationship between extreme maternal insensitivity, attachment, and cortisol reactivity, for the first time including both genetic and epigenetic factors. In the findings of Luijk, Velders, et al.[17], it remained puzzling why insecure-resistant attached infants seemed most affected by the SSP in terms of their cortisol reactivity, more so than disorganized infants. Resistant attachment behavior is usually accompanied with explicit signs of distress such as crying and the display of anger to the parent on return after a brief separation. As a result, children with insecure-resistant attachments might show higher cortisol stress reactivity to this challenge than securely attached infants. But infants with insecure-disorganized attachments might have even more difficulties with coping, and may be more dysregulated than insecure-resistant children because their previous experiences with extremely insensitive and frightening parental behaviors may have made them hypersensitive to stress and to lack of parental support when badly needed[32, 33]. Including genetic as well as epigenetic factors influencing the expression of the *FKBP5* gene might be necessary to uncover the associations between parenting, attachment, and allelic differences. G × E interactions might emerge more clearly when epigenetic variance is taken into account.

In sum, in this study, we aim to clarify if DNA methylation interacts with genetic effects and parenting on cortisol reactivity. We expand the study by Luijk, Velders, et al.[17] by investigating if and how *FKBP5* methylation affects the rs1360780 SNP-by-resistant attachment interaction reported in that study. Moreover, by including extreme maternal insensitivity, we take a broader perspective on the caregiver-child interaction. We hypothesize that the group with the highest risk for increased stress reactivity includes infants who show resistant or disorganized attachment behaviors, whose mothers display signs of extreme insensitive parenting, who are rs1360780 T-carriers, and who have the highest levels of *FKBP5* methylation.

Methods

Setting

The current study is embedded in Generation R, a prospective population-based cohort from fetal life onwards. Pregnant women living in the study area of Rotterdam, the Netherlands, with an expected delivery date between April 2002 and January 2006 were invited to participate. A more detailed description of the Generation R Study can be found elsewhere[34, 35]. In a randomly assigned subgroup of European pregnant women and their infants, detailed assessments were performed, including the Strange Situation Procedure (SSP). This subgroup

is ethnically homogenous (all with European ancestry) to exclude confounding or ethnic stratification effects. The Generation R Study is conducted in accordance with the World Medical Association Declaration of Helsinki and has been approved by the Medical Ethics Committee of the Erasmus Medical Center, Rotterdam. Written informed consent was obtained from the parents of all participating infants.

Study population

DNA was collected from cord blood samples at birth. Information on rs1360780 genotype and *FKBP5* methylation levels was available for 956 infants. At the age of 14 months (mean = 14.58, *SD* = 0.87), 568 of them participated in a lab visit, during which the SSP, extreme maternal insensitivity, and salivary cortisol samples were obtained. We were able to retrieve salivary cortisol samples from a total of 298 of these infants. This sample is nearly identical to the sample used by Luijk, Velders, et al.[17] (*n* = 310), with the discrepancy primarily caused by missing *FKBP5* methylation data. Unsuccessful cortisol sampling was mainly due to the infants' unwillingness to chew on the cotton swabs, and was especially seen in infants who were unfamiliar with pacifiers or who ceased using them. Sample characteristics are presented in Table 1. Excluded infants (i.e., infants without data on salivary cortisol, *N* = 270) did not differ from included children (*N* = 298) on resistant behavior during the SSP ($t(566) = 0.16$, $p = 0.87$, $d = 0.01$), disorganized attachment behavior ($t(566) = 1.06$, $p = 0.29$, $d = 0.09$), extreme maternal insensitivity ($t(513) = -0.87$, $p = 0.39$, $d = 0.08$) or maternal smoking during pregnancy ($\chi^2(1) = 0.06$, $p = 0.80$, $d = 0.02$). However, excluded infants differed from included infants in terms of age at the time of the SSP ($t(566) = 2.37$, $p = 0.02$, $d = 0.20$), sex ($\chi^2(1) = 7.98$, $p < 0.01$, $d = 0.24$) and maternal education ($\chi^2(1) = 4.64$, $p = 0.03$, $d = 0.19$). Specifically, infants with successful cortisol sampling were younger (mean age was 14.6 years in the included group, versus 14.8 years in the excluded group), were more often boys (57.0% in the included group, versus 45.2% in the excluded group) and their mothers were more often lower educated (39.5% of the mothers in included group had no formal higher education, versus 30.9% of the excluded group).

In 12 of the 298 infants for whom cortisol samples were available, observations of extreme maternal insensitivity were missing, due to procedural problems. To avoid reducing the group size of infants with the rs1360780 TT genotype (the hypothesized risk group), extreme maternal insensitivity scores were imputed using the expectation-maximization (EM) algorithm, using all other variables as well as prenatal maternal lifetime depression and breastfeeding at 6 months. Imputation with the EM algorithm was also performed to impute two missing values on the amount of crying during the SSP. Results remained essentially unchanged when rerunning the analyses using listwise deletion.

Table 1. Sample Characteristics (*N* = 298)

Variable	Mean (*SD*)
Infant characteristics	
Age at assessment of SSP, months	14.6 (0.9)
Sex, % girls	43.0
FKBP5 rs1360780 variant, %	
CC	47.0
CT	45.0
TT	8.1
FKBP5 methylation factor 1, score	0.15 (0.02)
FKBP5 methylation factor 2, score	0.31 (0.04)
Resistant behavior, continuous score	2.2 (1.3)
Resistant attachment, % resistant	24.5
Disorganized attachment behavior, score	3.4 (1.8)
Cortisol reactivity in Δ nmol/L	0.7 (6.2)
Mother characteristics	
Educational level, % lower	39.9
Smoking during pregnancy, % yes	11.4
Extreme insensitivity, continuous score	1.4 (1.0)
Extreme insensitive behaviors, % one or more	16.1

Note. SSP = Strange Situation Procedure

Measures

Genotyping

Cord blood DNA was genotyped for the rs1360780 single nucleotide polymorphism (SNP) of the FK506 Binding Protein 5 (FKBP5) with the TaqMan allelic discrimination assay (Applied Biosystems, Foster City, CA) and Abgene QPCR mix (Abgene, Hamburg, Germany). Polymerase chain reaction (PCR) took place with the GeneAmp® PCR system 9600, at 95 °C for 15 min, followed by 40 cycles of 94 °C for15s and 60 °C for 1 min. The 7900HT Fast Real-Time PCR System (Applied Biosystems) was used for fluorescence detection and genotypes were determined with SDS software (version 2.3, Applied Biosystems).

Contamination with the mother's blood was checked for the boys, by examination of the sex chromosomes. Samples in which contamination had occurred were excluded (< 1%). Furthermore, genotyping of the *FKBP5* SNP was successful in 97-99% of the cases and reanalysis of 276 randomly selected samples showed an error rate of < 1%. Genotype frequencies were in Hardy-Weinberg equilibrium ($\chi^2 = 1.07$, $p = 0.30$).

DNA methylation

Per sample, 500 nanogram of leukocyte DNA was extracted from cord blood and underwent bisulfite conversion with the EZ-96 DNA Methylation kit (Shallow) (Zymo Research Corporation, Irvine, USA). Methylation was analyzed with the Illumina Infinium Human Methylation 450K BeadChip (Illumina Inc., San Diego, USA). Quality control of samples was performed using standardized criteria. Samples were checked for < 99% call rate (six samples were excluded), color balance > 3, staining efficiency, extension efficiency, hybridization performance, stripping efficiency after extension (no samples excluded in each case), and bisulfite conversion (one sample excluded). Also, two samples were removed due to a sex mismatch, leaving a total of 969 samples that passed quality control. Dasen normalization was ran using a pipeline adapted from Touleimat and Tost[36], as described by Pidsley et al.[37], and samples were dye bias corrected.

We extracted the beta values of 32 CpGs that mapped to the *FKBP5* gene or overlapping regions adjacent to *FKBP5* (i.e. position 35543611 to 35697760; see Figure 1). Beta values represent the ratio of methylated signal relative to the sum of the methylated and unmethylated signals, per CpG. To avoid multiple testing issues due to the large number of CpG beta values, we decided to examine the dimensional structure of the data (mean r = 0.01, range r [-0.63, 0.77]) by using factor analysis in *M*Plus Version 7.31[38]. The factor analysis took place in the full DNA methylation sample, using the 29 CpG beta values with sufficient variation (*SD* > 0.01). Factor analysis proceeded in two steps. First, exploratory factor analysis (EFA) was performed. The optimal number of underlying factors was assessed by inspecting the Scree-plot and by comparing fit statistics between models estimating one to five factors. CpGs with a Geomin (oblique) rotated absolute loading of > 0.40 to one of the factors were included. Model fit

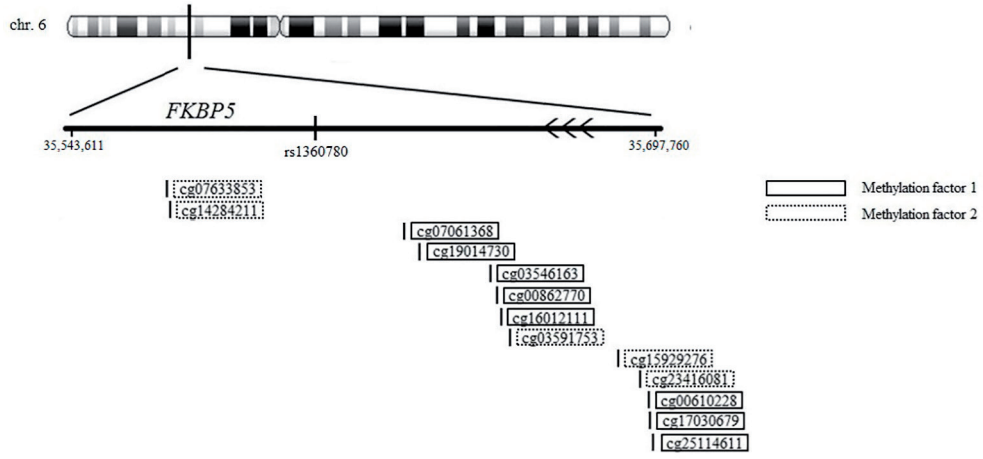


Figure 1. Beta values of 32 CpGs that mapped to the *FKBP5* gene or overlapping regions adjacent to *FKBP5*

was established using the chi-square statistic. In the event of significant chi-square values, we further examined relative fit indices, including the mean square error of approximation (RMSEA; acceptable fit ≤ 0.08), as well as the comparative fit index and the Tucker-Lewis index (CFI and TLI respectively; acceptable fit ≥ 0.90). A two-factor model was identified ($\chi^2[53] = 367.56$, $p < 0.001$; RMSEA = 0.078; TLI = 0.949; CFI = 0.926). The first factor had an eigenvalue of 5.2 and contained eight CpGs, of which five had positive and three had negative factor loadings. The second factor had an eigenvalue of 2.4 and contained five CpGs, all of which had positive factor loadings (see Figure 1). In the second step, we used confirmatory factor analysis (CFA) to validate the two-factor model ($\chi^2(64) = 438.60$; RMSEA = 0.077; CFI = 0.913; TLI = 0.894). For each *FKBP5* methylation factor, we computed average methylation scores based on the relevant CpGs, using reversed scores for those with negative loadings on factor 1. These average methylation factor scores were used throughout. In an exploratory analysis, regression analyses were repeated for each CpG inidividually, to gauge if our main finding was caused by only one or a few CpGs, or rather by the combined effect of all CpGs.

Attachment

Mother-infant dyads were observed in the Strange Situation Procedure (SSP). During the SSP, mild stress evokes attachment behavior in the infant by the unfamiliar lab environment, a stranger entering the room and engaging with the infant, and the parent briefly leaving the room twice. The total procedure consists of seven three-minute episodes, with the preseparation and separation in our study shortened by 1 min each, keeping the critical reunion episodes intact[39, 40].

Two reliable coders, trained at the University of Minnesota, coded the SSP recordings, according to the Ainsworth et al.[4] and Main and Solomon[33] coding systems. For each of two reunions with the mother, the infant received a resistant behavior score ranging from 1 to 7. These scores were averaged to create a resistant behavior score. Examples of resistant behavior include (i) a struggle against being held or (ii) throwing away toys that are handed to the infant. Intercoder reliability (intraclass correlation or ICC, single measure, absolute agreement) for resistant behavior was 0.86 ($n = 70$). For a sensitivity analysis (see below), a resistant attachment classification was derived from a pattern of attachment behaviors during the reunion periods. A typically resistant infant actively seeks proximity to the mother and tries to maintain contact with her, while at the same time showing obvious signs of resistance to her attempts of consolidation. Intercoder agreement for resistant attachment was 77% ($\kappa = 0.63$, $n = 70$). Resistant behavior in the reunion episodes and resistant attachment classification were strongly correlated ($r = 0.78$, $p < 0.01$). Disorganization of attachment behavior was rated using the 9-point Main and Solomon[33] coding system. Examples of disorganized/disoriented behaviors are prolonged stilling, rapid approach–avoidance vacillation, sudden unexplained affect changes, severe distress followed by avoidance, and expressions of fear or disorientation upon return of mother. The ICC for the disorganization rating scale was 0.88 ($n = 70$)[41].

Extreme maternal insensitivity

Extreme maternal insensitivity was observed during the psychophysiological assessment and during the break of the 14-month lab visit and was rated by coders unaware of the attachment coding. During the psychophysiological assessment, the child had ECG-measurement equipment attached while sitting on the mothers lap and watching an episode of the Teletubbies© (BBC/ Ragdoll Limited). The break was unstructured, and mother and child interacted freely. The extreme maternal insensitivity scale includes: (1) withdrawal and neglect; and (2) intrusive, negative, aggressive or otherwise harsh parental behaviors[42]. Extremely insensitive behaviors were coded on a 9-point scale, with higher scores indicating more extreme insensitivity. The ICC was 0.63 (n = 36).

Cortisol reactivity

Saliva samples were taken during the 14-month lab visit with Salivette sampling devices (Sarstedt, Rommelsdorf, Germany). Samples were centrifuged and frozen at -80 °C and analyzed by the Kirschbaum laboratory (Technical University of Dresden, Biological Psychology, Germany). Salivary cortisol concentrations were assessed with a chemiluminescence imunnoassay (CLIA; IBL Hamburg, Germany). Intra- and interassay coefficients of variation were below 7% and below 9%, respectively. Cortisol concentrations above the 99th percentile (>200 nmol/L) were excluded (n = 12) from the analyses. Cortisol reactivity was determined by calculation of the difference between cortisol concentration 15 minutes after the SSP (post-SSP cortisol) and cortisol concentrations prior to the SSP (pre-SSP cortisol). Mean sampling time of pre-SSP cortisol was 11:26 a.m. (SD = 2:01 h), mean sampling time of post-SSP cortisol was 12:22 p.m. (SD = 2:00 h). We had information on corticosteroid medication for 248 infants. None of these infants used systemic corticosteroid medication, but five infants used other corticosteroid-containing medication. Since these infants did not differ significantly in cortisol reactivity from infants without corticosteroid-containing medication ($t(246) < 0.01$, $p > 0.99$, $d < 0.01$), they were included in all further analyses.

Covariates

Information on family background characteristics was obtained by questionnaire during pregnancy. We included as covariates infant's age at the SSP, infant sex, mothers' highest attained educational level (no formal higher education versus higher vocational training or higher academic education), maternal smoking during pregnancy (never smoked or quit when pregnancy was known versus continued smoking during pregnancy), technical covariates (sample array number and position on the array, and leukocyte cell type proportions [CD4+ T-lymphocytes, CD8+ T-lymphocytes, natural killer cells, B-lymphocytes, monocytes, and granulocytes][43]. To account for the negative association (which might be interpreted as a ceiling effect) between the initial cortisol value and the slope of the cortisol reactivity, the cortisol concentration prior to the SSP (pre-SSP cortisol) was also included as a covariate. Finally, to exclude the possibility that resistant behavior and cortisol reactivity are related through

the physiologically arousing nature of crying that often accompanies resistant behavior, we performed the regression analysis with and without the inclusion of percentage of crying time during the SSP as a covariate.

Statistical analyses

Hierarchical linear regressions were performed using SPSS version 23 (IBM Corporation, Chicago, USA) to examine the associations of *FKBP5* rs1360780, *FKBP5* methylation and attachment (resistant or disorganized) with infant cortisol reactivity during the SSP. These regression analyses were performed separately for the two *FKBP5* methylation factors and for each of the two attachment variables.

In the first step of the regression equation, *FKBP5* rs1360780, *FKBP5* methylation, attachment, extreme maternal insensitivity, and the covariates were entered. In the second step, all two-way interactions between *FKBP5* rs1360780, *FKBP5* methylation, attachment, and extreme maternal insensitivity were entered. In the third step, all three-way interactions were entered. In the interest of statistical power, the four-way interaction with all possible predictors was not included. When one of the main predictors was not found to have a significant main or interaction effect on cortisol reactivity, the steps were repeated excluding this variable.

To reduce the influence of extreme scores on the results, two outliers (z-score >3.29) for *FKBP5* methylation factor 1, four for *FKBP5* methylation factor 2, six for cortisol reactivity, and 10 for extreme insensitivity were winsorized (i.e. transformed to match the next highest value). *FKBP5* rs1360780, the *FKBP5* average methylation factors, resistant and disorganized behavior, and extreme maternal insensitivity, were mean-centered in order to reduce collinearity due to the scaling of variables.

Sensitivity analyses

Two sensitivity analyses were performed. First, in order to examine whether associations were dependent on the continuous resistance scale, we also used the resistant versus non-resistant attachment classification[17] as a predictor instead of the continuous resistant behavior score. Second, since most mothers had the lowest possible score on extreme insensitivity, which resulted in a skewed distribution of scale scores, we performed a sensitivity analysis with a dichotomized extreme insensitivity variable. Mothers not showing any extremely insensitive behaviors were contrasted with mothers presenting one or more extremely insensitive behaviors.

Results

Extreme maternal insensitivity

As can be seen in Table 2, none of the main predictors were correlated, with the exception of *FKBP5* methylation factors 1 and 2 ($r = 0.34$, $p < 0.01$), and disorganized and resistant behavior ($r = 0.19$, $p < 0.01$). The regression analyses did not show an association of cortisol reactivity with extreme maternal insensitivity ($\beta = -0.04$, $p = 0.43$) in the first step, nor a with a two-way interaction of extreme maternal insensitivity and *FKBP5* rs1360780, *FKBP5* methylation factor 1, or resistant behavior (strongest interaction with *FKBP5* rs1360780: $\beta = -0.02$, $p = 0.69$) in the second step, nor with a three-way interaction with extreme maternal insensitivity and any combination of these predictors (strongest interaction with *FKBP5* methylation factor 1 and resistant behavior: $\beta = -0.04$, $p = 0.51$) in the final step. This was also the case for the analyses with *FKBP5* methylation factor 2.

Resistant attachment, FKBP5 rs1360780, and FKBP5 methylation

Table 3 shows that both *FKBP5* rs1360780 ($\beta = 0.13$, $p < 0.01$) and resistant behavior ($\beta = 0.30$, $p < 0.01$), but not *FKBP5* factor 1 methylation ($\beta = -0.06$, $p = 0.40$), were positively associated with infant cortisol reactivity. The two-way interaction of rs1360780 and *FKBP5* methylation factor 1 was significant ($\beta = 0.11$, $p = 0.03$), as was the three-way interaction *FKBP5* rs1360780 × *FKBP5* methylation factor 1 × resistant behavior ($\beta = 0.14$, $p < 0.01$). T-allele carriers of *FKBP5* rs1360780 with high *FKBP5* methylation factor 1 scores and high levels of resistant behavior had the highest cortisol reactivity.

Similarly, *FKBP5* methylation factor 2 was unrelated to cortisol reactivity ($\beta = 0.04$, $p = 0.77$; Table 4). The interaction between *FKBP5* rs1360780 and *FKBP5* methylation factor 2 did not reach significance, but there was again a positive association between cortisol reactivity and resistant behavior ($\beta = 0.28$, $p < 0.01$) and a significant three-way interaction of *FKBP5* rs1360780 × *FKBP5* methylation factor 2 × resistant behavior ($\beta = 0.13$, $p = 0.01$), again suggesting that

Table 2. Pearson correlations ($N = 298$)

	1	2	3	4	5	6	7	8
1. *FKBP5* rs1360780								
2. *FKBP5* methylation factor 1	< -0.01							
3. *FKBP5* methylation factor 2	-0.03	0.34***						
4. Resistant behavior	-0.02	-0.03	-0.05					
5. Disorganized behavior	0.01	-0.02	< 0.01	0.19**				
6. Extreme insensitivity	-0.01	0.05	0.03	0.01	0.02			
7. Crying	0.02	0.01	-0.03	0.55***	0.01	0.02		
8. Cortisol reactivity	0.14*	-0.04	< 0.01	0.26***	-0.07	< 0.01	0.36***	

Note. FKBP5 rs1360780: CC = 0, CT = 1, TT = 2.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Table 3. Associations between *FKBP5* rs1360780, *FKBP5* methylation factor 1, and resistant behavior on cortisol reactivity during the Strange Situation Procedure (*N* = 298)

Model	*B (95% CI)*	β
FKBP5 rs1360780	1.13 (0.30; 1.97)	0.13**
FKBP5 methylation factor 1	-12.99 (-43.50; 17.53)	-0.06
Resistant behavior	1.82 (1.20; 2.45)	0.30***
FKBP5 rs1360780 × *FKBP5* methylation factor 1	39.11 (3.38; 74.85)	0.11*
FKBP5 rs1360780 × resistant behavior	0.89 (-0.14; 1.93)	0.09
FKBP5 methylation factor 1 × resistant behavior	10.43 (-19.47; 40.32)	0.04
FKBP5 rs1360780 × *FKBP5* methylation factor 1 × resistant behavior	64.22 (17.34; 111.10)	0.14**

Note. FKBP5 rs1360780: CC = 0, CT = 1, TT = 2. Analyses are adjusted for technical methylation covariates, cell type proportions of DNA methylation sample, infant age at assessment of SSP, infant sex, educational level of the mother and maternal smoking during pregnancy. The statistics are derived from the final block of the regression model.
*p < 0.05, **p < 0.01, ***p < 0.001.

Table 4. Associations between *FKBP5* rs1360780, *FKBP5* methylation factor 2, and resistant behavior on cortisol reactivity during the Strange Situation Procedure (*N* = 298)

Model	*B (95% CI)*	β
FKBP5 rs1360780	1.14 (0.30; 1.98)	0.13**
FKBP5 methylation factor 2	5.19 (-29.48; 39.86)	0.04
Resistant behavior	1.75 (1.14; 2.36)	0.28***
FKBP5 rs1360780 × *FKBP5* methylation factor 2	9.69 (-14.00; 33.37)	0.04
FKBP5 rs1360780 × resistant behavior	0.41 (-0.58; 1.39)	0.04
FKBP5 methylation factor 2 × resistant behavior	-9.44 (-23.80; 4.91)	-0.07
FKBP5 rs1360780 × *FKBP5* methylation factor 2 × resistant behavior	31.06 (6.60; 55.51)	0.13*

Note. FKBP5 rs1360780: CC = 0, CT = 1, TT = 2. Analyses are adjusted for technical methylation covariates, cell type proportions of DNA methylation sample, infant age at assessment of SSP, infant sex, educational level of the mother and maternal smoking during pregnancy. The statistics are derived from the final block of the regression model.
*p < 0.05, **p < 0.01, ***p < 0.001.

T-allele carriers of rs1360780, with high *FKBP5* methylation factor 2 levels and high resistant behavior had the highest cortisol reactivity to the SSP.

Although resistant behavior was positively correlated with crying (r = 0.55, p < 0.001), adding crying as a covariate to the model did not meaningfully change the results. The three-way interactions of (i) *FKBP5* rs1360780 × *FKBP5* methylation factor 1 × resistant behavior (β = 0.14, p < 0.01) and (ii) *FKBP5* rs1360780 × *FKBP5* methylation factor 2 × resistant behavior (β = 0.11, p = 0.03) remained significant.

Finally, to explore whether the results for the methylation factor scores were localized in just one or a few CpGs, or were based on the combined effect of all CpGs, the analyses were

repeated for each CpG separately. For methylation factor 1, three out of eight CpGs were associated with cortisol reactivity in the *FKBP5* rs1360780 × *FKBP5* CpG × resistant behavior interaction at the $p < 0.05$ level (Table 5). However, two of the other five CpGs may also have contributed to the *FKBP5* methylation factor 1 involvement in the three-way interaction, as the interaction terms for two CpGs were associated with cortisol reactivity at $p < 0.10$. For methylation factor 2, four out of five CpGs were associated on the $p < 0.05$ with cortisol reactivity in interaction with *FKBP5* rs1360780 and resistant behavior (Table 6). For CpGs of both *FKBP5* methylation factors, no clear localization pattern of $p < 0.05$ results could be distinguished, as they were relatively scattered over the *FKBP5* gene.

Table 5. Characteristics of the individual *FKBP5* methylation factor 1 CpGs and β and *p*-value of the *FKBP5* rs1360780 x *FKBP5* CpG beta value x resistant behavior in a regression analysis of the associations between *FKBP5* rs1360780, *FKBP5* CpG, and resistant behavior on cortisol reactivity during the Strange Situation Procedure

				Three-way interaction values with cortisol reactivity	
CpG	**mean beta value**	***(SD)***	**Factor 1 loadings**	**β**	***p*-value**
cg07061368	.89	(.03)	-0.58	-0.01	0.811
cg19014730	.80	(.05)	-0.76	-0.14	0.007
cg03546163	.79	(.06)	-0.71	-0.12	0.017
cg00862770	.07	(.01)	0.72	0.09	0.071
cg16012111	.10	(.01)	0.63	0.04	0.453
cg00610228	.14	(.04)	0.93	0.09	0.092
cg17030679	.07	(.02)	0.55	0.07	0.210
cg25114611	.35	(.04)	0.73	0.12	0.014

Note. FKBP5 rs1360780: CC = 0, CT = 1, TT = 2. Analyses are adjusted for technical methylation covariates, cell type proportions of DNA methylation sample, infant age at assessment of SSP, infant sex, educational level of the mother and maternal smoking during pregnancy. The statistics are derived from the final block of the regression model.

Table 6. Characteristics of the individual *FKBP5* methylation factor 2 CpGs and β and *p*-value of the *FKBP5* rs1360780 x *FKBP5* CpG beta value x resistant behavior in a regression analysis of the associations between *FKBP5* rs1360780, *FKBP5* CpG, and resistant behavior on cortisol reactivity during the Strange Situation Procedure

				Three-way interaction values with cortisol reactivity	
CpG	**mean beta value**	***(SD)***	**Factor 2 loadings**	**β**	***p*-value**
cg07633853	.33	(.07)	0.63	0.05	0.313
cg14284211	.26	(.06)	0.90	0.14	0.006
cg03591753	.55	(.03)	0.66	0.11	0.036
cg15929276	.13	(.04)	0.47	0.13	0.008
cg23416081	.27	(.05)	0.86	0.11	0.033

Note. FKBP5 rs1360780: CC = 0, CT = 1, TT = 2. Analyses are adjusted for technical methylation covariates, cell type proportions of DNA methylation sample, infant age at assessment of SSP, infant sex, educational level of the mother and maternal smoking during pregnancy. The statistics are derived from the final block of the regression model.

Disorganized attachment, FKBP5 rs1360780, and FKBP5 methylation

When resistant behavior was replaced by disorganized attachment behavior in the regression analyses including *FKBP5* methylation factor 1 and maternal extreme insensitivity, neither an association between cortisol reactivity and attachment disorganization (β = -0.03, *p* = 0.56), nor any two- or three-way interactions (strongest interaction with *FKBP5* methylation factor 1: β = -0.06, *p* = 0.31) with disorganized behavior was found. Results were found to be similarly non-significant for the analyses with *FKBP5* methylation factor 2 (see also Supplemental Table 1 and Supplemental Table 2, respectively). Moreover, a z-test indicated that the main effect for disorganized attachment behavior and the FKBP5 rs1360780 × FKBP5 methylation factor 1 × disorganized attachment behavior interaction differed significantly from the main effect for resistant behavior (z = 5.11, p < 0.01) and the FKBP5 rs1360780 × FKBP5 methylation factor 1 × resistant behavior interaction (z = 2.78, p < 0.01).

Sensitivity Analyses

When repeating the analyses using the resistant attachment classification variable(categorical, resistant versus nonresistant) instead of the continuous resistant behavior score, a similar pattern of findings for *FKBP5* methylation factor 1 was observed. That is, we observed a significant three-way interactions of *FKBP5* rs1360780 × *FKBP5* methylation factor 1 × resistant attachment (β = 0.19, *p* < 0.01) in the prediction of infant cortisol reactivity (Table 7). The three-way interaction of *FKBP5* rs1360780 × *FKBP5* methylation factor 2 × resistant attachment was also significant (Table 8; β = 0.13, *p* = 0.04).

Last, we inserted the dichotomous extreme insensitivity variable in the model instead of the continuous variant. This yielded no significant additive or interactive associations of variables with the dichotomous extreme insensitivity variable involved with cortisol reactivity. This was the case for the analysis with *FKBP5* methylation factor 1 as well as for the analysis with *FKBP5* methylation factor 2.

Table 7. Associations between *FKBP5* rs1360780, *FKBP5* methylation factor 1, and resistant attachment classification on cortisol reactivity during the Strange Situation Procedure (*N* = 298)

Model	*B*	*(95% CI)*	β
FKBP5 rs1360780	1.72	(0.79; 2.65)	0.20***
FKBP5 methylation factor 1	-4.14	(-39.10; 30.81)	-0.02
Resistant attachment	1.86	(1.26; 2.47)	0.30***
FKBP5 rs1360780 × *FKBP5* methylation factor 1	72.95	(32.75; 113.15)	0.20***
FKBP5 rs1360780 × resistant attachment	1.12	(0.20; 2.05)	0.13*
FKBP5 methylation factor 1 × resistant attachment	18.15	(-10.02; 46.32)	0.08
FKBP5 rs1360780 × *FKBP5* methylation factor 1 × resistant attachment	69.22	(29.10; 109.34)	0.19**

Note. FKBP5 rs1360780: CC = 0, CT = 1, TT = 2. Analyses are adjusted for technical methylation covariates, cell type proportions of DNA methylation sample, infant age at assessment of SSP, infant sex, educational level of the mother and maternal smoking during pregnancy. The statistics are derived from the final block of the regression model.
*p < 0.05, **p < 0.01, ***p < 0.001.

Table 8. Associations between *FKBP5* rs1360780, *FKBP5* methylation factor 2, and resistant attachment classification on cortisol reactivity during the Strange Situation Procedure (*N* = 298)

Model	*B*	*(95% CI)*	β
FKBP5 rs1360780	1.61	(0.66; 2.56)	0.19**
FKBP5 methylation factor 2	1.64	(-34.35; 37.63)	0.01
Resistant attachment	1.80	(1.19; 2.41)	0.29***
FKBP5 rs1360780 × *FKBP5* methylation factor 2	25.82	(-3.06; 54.70)	0.11
FKBP5 rs1360780 × resistant attachment	0.99	(0.04; 1.94)	0.12*
FKBP5 methylation factor 2 × resistant attachment	-8.64	(-25.60; 8.33)	-0.06
FKBP5 rs1360780 × *FKBP5* methylation factor 2 × resistant attachment	30.07	(1.63; 58.52)	0.13*

Note. FKBP5 rs1360780: CC = 0, CT = 1, TT = 2. Analyses are adjusted for technical methylation covariates, cell type proportions of DNA methylation sample, infant age at assessment of SSP, infant sex, educational level of the mother and maternal smoking during pregnancy. The statistics are derived from the final block of the regression model.
*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Discussion

In this population-based cohort study, we found that resistant attachment behavior and *FKBP5* rs1360780 genotype were associated with cortisol reactivity both in an additive and in an interactive manner. Methylation of the *FKBP5* gene moderated the relationship between *FKBP5* rs1360780 genotype and cortisol reactivity, in that rs1360780 T-carriers had an even higher chance of increased cortisol reactivity, when they also had a high *FKBP5* methylation factor 1 score. This might suggest that DNA methylation patterns affect transcription of the *FKBP5* gene to influence subsequent stress responses. The modification of the rs1360780 association with cortisol reactivity by *FKBP5* methylation factor 1 score, seemed especially pronounced in infants who displayed resistant attachment behavior towards their mother. Results for the analysis with methylation factor 2 and with the resistant attachment classification corroborated these findings. It is noteworthy that the interaction between rs1360780 and methylation was specifically modified by resistant attachment behavior, and not by disorganized attachment.

Although our study shows promising findings and does support the potentially important role of DNA methylation in infant cortisol reactivity, it should be emphasized that the study must be firmly placed in the context of discovery[44]. For several reasons it is too early for this and related human development studies on DNA methylation to provide more definite confirmation or falsification of hypotheses or theories in the context of justification. First, little is known about the metric qualities of DNA methylation indices. For example, stability of DNA methylation across time has not yet been examined thoroughly for most genes and developmental periods (see Wong et al.[45], for an exception). Second, it is still not fully clear whether and how strongly DNA methylation patterns in blood and brain regions are associated[26]. Some recent studies show significant convergence between *FKBP5* methylation

derived from peripheral blood and brain tissue[46, 47], but more research is certainly needed. Third, most studies on DNA methylation in the domain of developmental psychopathology are severely underpowered with potentially quite a few false positive findings that may turn out to be impossible to replicate. In an epigenome-wide study of a large Generation R sample of 912 families, we were unable to replicate our suggestive findings on the association between maternal prenatal stress and neonatal DNA methylation in another large sample of 828 families, the Avon Longitudinal Study of Parents and Children, ALSPAC[48]. Our current study on almost 300 children is one of the largest candidate-(epi)gene studies on DNA methylation, but still underpowered in view of the Gene × Methylation × Environment[27, 49] (G × M × E) three-way interactions. Independent replication is therefore badly needed[48].

It is somewhat assuring that our findings are in line with Klengel et al.[15], who found an association between *FKBP5* methylation and HPA axis regulation, particularly in *FKBP5* rs1360780 T allele carriers. However, their sample size for these analyses was only 76 with 30 highly traumatized cases and 46 controls in one of their central epigenetic analyses. Paquette et al.[31] also found an association between placental *FKBP5* methylation and postnatal infant arousal, but this was specific for infants with the *FKBP5* rs1360780 CC genotype. One explanation for this diverging result might be that arousal is regulated by the autonomic nerve system, which is related to the HPA axis, but does not completely overlap in its function and activity. Also, whereas Klengel et al.[15] and Paquette et al.[31] specifically found effects of methylation of CpG sites in intron 7 of the *FKBP5* gene, we considered methylation of all *FKBP5* CpG sites that contributed meaningfully to one of two factors (as in Philibert et al.[50]). Since these factors, which included CpG sites with positive as well as negative factor loadings, were found to be associated with cortisol reactivity, it might be that the effects of DNA methylation are less unidirectional than assumed previously. Exploratory analyses also showed that the individual CpGs contributing to the *FKBP5* methylation factor scores were quite scattered along the *FKBP5* gene, rather than being localized in a specific part. Unfortunately, the different methodologies for DNA methylation detection employed do not allow for direct comparison of our approach with those of Klengel et al.[15] and Paquette et al.[31].

The interaction between *FKBP5* rs1360780 and *FKBP5* methylation was only found in children with resistant but not with disorganized attachment behavior. This is somewhat unexpected, since disorganized attachment has been related to dysregulation of the HPA axis functioning in a number of studies[10, 11, 51]. Another remarkable result is the negligible role of maternal extreme insensitivity in the prediction of cortisol reactivity, since we had expected that infants of mothers displaying extreme insensitive parenting behaviors would show increased cortisol reactivity. Perhaps our relatively brief observation in a lab setting was not optimal to register maternal extreme insensitivity. Moreover, the non-clinical nature of the sample may also have contributed to the skewedness of the distribution, thereby hindering detection of associations with maternal extreme insensitivity. This might also explain the lack of association

between maternal extreme insensitivity and disorganized attachment. Future research on maternal extreme insensitivity therefore might include more high-risk populations than the one examined here, which could possibly also help in further exploring the (epi-)genetic differences in stress regulation between children with disorganized and resistant attachments.

Some limitations of the current study should be mentioned. First, DNA methylation levels were measured in cord blood at birth, whereas cortisol reactivity was measured at 14 months. Neonatal DNA methylation might be influenced by prenatal environmental factors such as maternal smoking[19, 52] or prenatal stress[53]. An important question that remains unanswered, is whether DNA methylation levels are stable between birth and our behavioral observations at 14 months. However, based on Klengel et al.'s[15] finding that trauma during childhood affects *FKBP5* methylation in a way that dysregulates HPA axis functioning, one might speculate that insecure-resistant mother-child attachment—although not traumatizing in and of itself of course—could affect *FKBP5* methylation over the first year of life, so that its associations with cortisol reactivity would have been even stronger with *FKBP5* methylation measured at 14 months than at birth. In order to attain a more complete picture of the role of epigenetics in shaping the relations between parenting, attachment and stress regulation, longitudinal and experimental research is needed to test whether the quality of parenting (i.e., sensitivity) and the attachment relationship in itself can affect DNA methylation. Longitudinal data on DNA methylation of stress-related genes at multiple time-points may be informative, as well as pre- and posttest assessments of DNA methylation patterns in randomized controlled trials aiming at enhancing the quality of parent-child interactions and relationship[54].

Second, another limitation may be the candidate-(epi-)gene approach that limits the analysis to one specific gene, i.e. the *FKBP5* gene, in combination with a single SNP, i.e. rs1360780, which was a logical follow-up on the study performed by Luijk, Velders, et al.[17]. It would be interesting to try and obtain a more complete picture of DNA methylation in stress regulation by including more SNPs of the *FKBP5* gene as well as other genes related to cortisol reactivity (e.g., *NR3C1*[53]*;* or *KITLG*[55]). Combinations of such HPA-axis related genes into a genetic pathway might provide a better basis for a wider epigenetic search into the influence of DNA methylation patterns on stress regulation. It should be noted, however, that focusing on methylation patterns of a single gene with documented functionality for the phenotype of interest has the advantage of better localization of the effect and of optimizing the statistical power that is often lacking in hypothesis-free approaches. Nevertheless, our results are based on a complicated three-way interaction (G × M × E) and should be replicated in independent samples. In such studies, the factor-analytic method to examine the dimensionality of an interrelated set of CpG beta values may reduce the number of tests that otherwise would lower statistical power. More specifically, this G × M × E study shows that stress regulation in an infant with a resistant attachment to their mother is more likely to be problematic

when the infant is a *FKBP5* rs1360780 T-carrier and even more so when it also has a higher methylation factor score.

In sum, the current findings are a valuable extension of our earlier results on attachment, *FKBP5* rs1360780 and cortisol reactivity in that genetic effects on child outcomes may be better specified when DNA methylation is taken into account. Moreover, whereas most research on *FKBP5* methylation focuses on extreme circumstances in early life, this study reveals that DNA methylation plays a role in coping with everyday stressors in a non-clinical population. Although we emphasize that epigenetic studies on (child developmental) psychopathology are still in an exploratory stage, neglect of DNA methylation and other regulatory mechanisms in molecular genetic studies increases the risk that an incomplete picture of associations between genes, environment and development is created[25]. The study of epigenetics is therefore an important asset to the field of developmental psychopathology, and a crucial move to the biological level of gene-by-environment interplay.

References

1. Gunnar MR, Brodersen L, Nachmias M, Buss K, Rigatuso J. Stress reactivity and attachment security. *Developmental Psychobiology.* 1996;29:191-204.
2. Bowlby J. Attachment, Vol. 1 of *Attachment and Loss*. In: New York: Basic Books*,* 1969.
3. Cassidy J. *The nature of the child's ties*. 2008.
4. Ainsworth MDS, Blehar MC, Waters E, Wall SN. *Patterns of attachment: A psychological study of the strange situation*. Psychology Press, 2015.
5. Fox NA, Hane AA. Studying the biology of human attachment. *Handbook of attachment: Theory, Research, and Clinical Applications.* 2008;2:217-40.
6. Doom JR, Gunnar MR. Stress physiology and developmental psychopathology: Past, present and future. *Development and Psychopathology.* 2013;25.
7. Carlson V, Cicchetti D, Barnett D, Braunwald K. Disorganized/disoriented attachment relationships in maltreated infants. *Developmental Psychology.* 1989;25:525.
8. Cyr C, Euser EM, Bakermans-Kranenburg MJ, van IJzendoorn MH. Attachment security and disorganization in maltreating and high-risk families: A series of meta-analyses. *Development and Psychopathology.* 2010;22:87-108.
9. Lyons-Ruth K, Alpern L, Repacholi B. Disorganized infant attachment classification and maternal psychosocial problems as predictors of hostile-aggressive behavior in the preschool classroom. *Child Development.* 1993;64:572-85.
10. Hertsgaard L, Gunnar M, Erickson MF, Nachmias M. Adrenocortical responses to the strange situation in infants with disorganized/disoriented attachment relationships. *Child Development.* 1995;66:1100-6.
11. Spangler G, Grossmann KE. Biobehavioral organization in securely and insecurely attached infants. *Child Development.* 1993;64:1439-50.
12. Binder EB. The role of FKBP5, a co-chaperone of the glucocorticoid receptor in the pathogenesis and therapy of affective and anxiety disorders. *Psychoneuroendocrinology.* 2009;34:S186-S95.
13. Ising M, Depping AM, Siebertz A, et al. Polymorphisms in the FKBP5 gene region modulate recovery from psychosocial stress in healthy controls. *European Journal of Neuroscience.* 2008;28:389-98.
14. Binder EB, Bradley RG, Liu W, et al. Association of FKBP5 polymorphisms and childhood abuse with risk of posttraumatic stress disorder symptoms in adults. *Jama.* 2008;299:1291-305.
15. Klengel T, Mehta D, Anacker C, et al. Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. *Nature Neuroscience.* 2013;16:33-41.
16. Roy A, Gorodetsky E, Yuan Q, Goldman D, Enoch M-A. Interaction of FKBP5, a stress-related gene, with childhood trauma increases the risk for attempting suicide. *Neuropsychopharmacology.* 2010;35:1674-83.
17. Luijk MP, Velders FP, Tharner A, et al. FKBP5 and resistant attachment predict cortisol reactivity in infants: gene–environment interaction. *Psychoneuroendocrinology.* 2010;35:1454-61.
18. Li G, Reinberg D. Chromatin higher-order structures and gene regulation. *Current Opinion in Genetics & Development.* 2011;21:175-86.

19. Bouwland-Both MI, van Mil NH, Tolhoek CP, et al. Prenatal parental tobacco smoking, gene specific DNA methylation, and newborns size: the Generation R study. *Clinical Epigenetics.* 2015;7:1.
20. Cao-Lei L, Massart R, Suderman MJ, et al. DNA methylation signatures triggered by prenatal maternal stress exposure to a natural disaster: Project ice storm. *PLoS ONE.* 2014;9.
21. Mychasiuk R, Ilnytskyy S, Kovalchuk O, Kolb B, Gibb R. Intensity Matters: Brain, Behaviour and the Epigenome of Prenatally Stressed Rats. *Neuroscience.* 2011;180:105-10.
22. Hughes LA, van den Brandt PA, De Bruïne AP, et al. Early life exposure to famine and colorectal cancer risk: a role for epigenetic mechanisms. *PLoS ONE.* 2009;4:e7951.
23. Mehta D, Klengel T, Conneely KN, et al. Childhood maltreatment is associated with distinct genomic and epigenetic profiles in posttraumatic stress disorder. *Proceedings of the National Academy of Sciences.* 2013;110:8302-7.
24. Murgatroyd C, Patchev AV, Wu Y, et al. Dynamic DNA methylation programs persistent adverse effects of early-life stress. *Nature Neuroscience.* 2009;12:1559-66.
25. Meaney MJ. Epigenetics and the biological definition of gene X environment interactions. *Child Devevelopment.* 2010;81:41-79.
26. van IJzendoorn MH, Bakermans-Kranenburg MJ, Ebstein RP. Methylation Matters in Child Development: Toward Developmental Behavioral Epigenetics. *Child Development Perspectives.* 2011;5:305-10.
27. van IJzendoorn MH, Caspers K, Bakermans-Kranenburg MJ, Beach SR, Philibert R. Methylation matters: interaction between methylation density and serotonin transporter genotype predicts unresolved loss or trauma. *Biological Psychiatry.* 2010;68:405-7.
28. Kember R, Dempster E, Lee T, Schalkwyk LC, Mill J, Fernandes C. Maternal separation is associated with strain-specific responses to stress and epigenetic alterations to Nr3c1, Avp, and Nr4a1 in mouse. *Brain and Behavior.* 2012;2:455-67.
29. Wu Y, Patchev AV, Daniel G, Almeida OF, Spengler D. Early-life stress reduces DNA methylation of the Pomc gene in male mice. *Endocrinology.* 2014;155:1751-62.
30. Oberlander TF, Weinberg J, Papsdorf M, Grunau R, Misri S, Devlin AM. Prenatal exposure to maternal depression, neonatal methylation of human glucocorticoid receptor gene (NR3C1) and infant cortisol stress responses. *Epigenetics.* 2008;3:97-106.
31. Paquette AG, Lester BM, Koestler DC, Lesseur C, Armstrong DA, Marsit CJ. Placental FKBP5 genetic and epigenetic variation is associated with infant neurobehavioral outcomes in the RICHS cohort. *PLoS ONE.* 2014;9.
32. Hesse E, Main M. Frightened, threatening, and dissociative parental behavior in low-risk samples: Description, discussion, and interpretations. *Development and Psychopathology* 2006;18:309-43.
33. Main M, Solomon J. Procedures for identifying infants as disorganized/disoriented during the Ainsworth Strange Situation. *Attachment in the preschool years: Theory, research, and intervention.* 1990;1:121-60.
34. Jaddoe VW, van Duijn CM, Franco OH, et al. The Generation R Study: design and cohort update 2012. *European Journal of Epidemiology.* 2012;27:739-56.
35. Kruithof CJ, Kooijman MN, van Duijn CM, et al. The generation R study: Biobank update 2015. *European Journal of Epidemiology.* 2014;29:911-27.

36. Touleimat N, Tost J. Complete pipeline for Infinium® Human Methylation 450K BeadChip data processing using subset quantile normalization for accurate DNA methylation estimation. *Epigenomics.* 2012;4:325-41.
37. Pidsley R, Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics.* 2013;14:293.
38. Muthén LK, Muthén BO. *Mplus User's Guide* (7th edn). In: Muthén and Muthén Los Angeles, CA, 2012.
39. Kok R, van IJzendoorn MH, Linting M, et al. Attachment insecurity predicts child active resistance to parental requests in a compliance task. *Child: Care, Health and Development.* 2013;39:277-87.
40. Luijk MP, Saridjan N, Tharner A, et al. Attachment, depression, and cortisol: Deviant patterns in insecure-resistant and disorganized infants. *Developmental Psychobiology.* 2010;52:441-52.
41. Luijk MP, Roisman GI, Haltigan JD, et al. Dopaminergic, serotonergic, and oxytonergic candidate genes associated with infant attachment security and disorganization? In search of main and interaction effects. *Journal of Child Psychology and Psychiatry.* 2011;52:1295-307.
42. Out D, Bakermans-Kranenburg MJ, van IJzendoorn MH. The role of disconnected and extremely insensitive parenting in the development of disorganized attachment: Validation of a new measure. *Attachment & Human Development.* 2009;11:419-43.
43. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics.* 2012;13:1.
44. Popper KR. *The logic of scientific discovery.* London: Hutchinson 1959.
45. Wong CCY, Caspi A, Williams B, et al. A longitudinal study of epigenetic variation in twins. *Epigenetics.* 2010;5:516-26.
46. Ewald ER, Wand GS, Seifuddin F, et al. Alterations in DNA methylation of Fkbp5 as a determinant of blood–brain correlation of glucocorticoid exposure. *Psychoneuroendocrinology.* 2014;44:112-22.
47. Hannon E, Lunnon K, Schalkwyk L, Mill J. Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics.* 2015;10:1024-32.
48. Rijlaarsdam J, Pappa I, Walton E, et al. An epigenome-wide association meta-analysis of prenatal maternal stress in neonates: A model approach for replication. *Epigenetics.* 2016;11:140-9.
49. Bakermans-Kranenburg MJ, van IJzendoorn MH. The hidden efficacy of interventions: Gene× environment experiments from a differential susceptibility perspective. *Annual Review of Psychology.* 2015;66:381-409.
50. Philibert RA, Beach SR, Gunter TD, Brody GH, Madan A, Gerrard M. The effect of smoking on MAOA promoter methylation in DNA prepared from lymphoblasts and whole blood. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics.* 2010;153:619-28.
51. Bernard K, Dozier M. Examining infants' cortisol responses to laboratory tasks among children varying in attachment disorganization: stress reactivity or return to baseline? *Developmental Psychology.* 2010;46:1771.

52. Richmond RC, Simpkin AJ, Woodward G, et al. Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Human Molecular Genetics.* 2014;24:2201-17.
53. Mulligan CJ, D'Errico NC, Stees J, Hughes DA. Methylation changes at NR3C1 in newborns associate with maternal prenatal stress exposure and newborn birth weight. *Epigenetics.* 2012;7:853-7.
54. Bakermans-Kranenburg MJ, van IJzendoorn MH, Mesman J, Alink LRA, Juffer F. Effects of an attachment-based intervention on daily cortisol moderated by dopamine receptor D4: A randomized control trial on 1-to 3-year-olds screened for externalizing behavior. *Development and Psychopathology.* 2008;20:805-20.
55. Houtepen LC, Vinkers CH, Carrillo-Roa T, et al. Genome-wide DNA methylation levels and altered cortisol stress reactivity following childhood trauma in humans. *Nature Communications.* 2016;7.

Supplemental Tables

Supplemental Table 1. Associations between *FKBP5* rs1360780, *FKBP5* methylation factor 1, disorganized behavior, and extreme insensitivity on cortisol reactivity during the Strange Situation Procedure (*N* = 298)

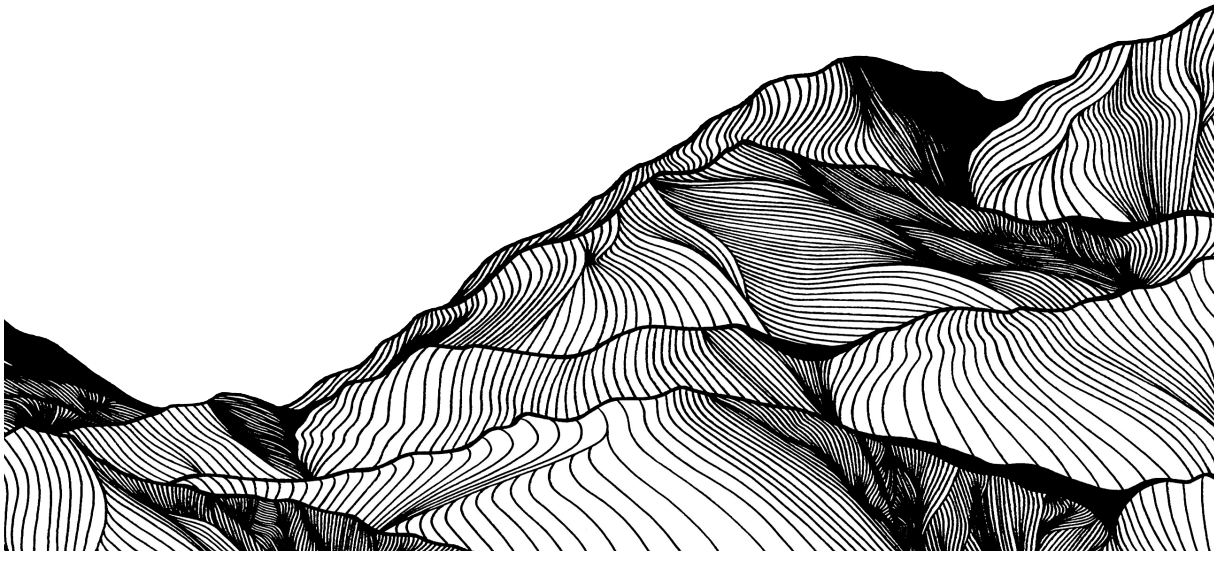
Model	*B*	*(95% CI)*	β
FKBP5 rs1360780	1.05	(0.15; 1.95)	0.12*
FKBP5 methylation factor 1	-23.11	(-55.80; 9.59)	-0.10
Disorganized behavior	-0.10	(-0.42; 0.23)	-0.03
Extreme insensitivity	-0.26	(-1.06; 0.53)	-0.04
FKBP5 rs1360780 × *FKBP5* methylation factor 1	35.39	(-3.35; 74.13)	0.10
FKBP5 rs1360780 × disorganized behavior	0.16	(-0.37; 0.68)	0.03
FKBP5 methylation factor 1 × disorganized behavior	-7.26	(-21.36; 6.83)	-0.06
FKBP5 rs1360780 × extreme insensitivity	-0.86	(-2.19; 0.48)	-0.07
FKBP5 methylation factor 1 × extreme insensitivity	-2.83	(-30.48; 24.82)	-0.01
Disorganized behavior × extreme insensitivity	-0.22	(-0.77; 0.33)	-0.05

Note. FKBP5 rs1360780: CC = 0, CT = 1, TT = 2. Analyses are adjusted for technical methylation covariates, cell type proportions of DNA methylation sample, infant age at assessment of SSP, infant sex, educational level of the mother and maternal smoking during pregnancy. The statistics are derived from the final block of the regression model.
None of the possible three-way interactions was significantly associated with cortisol reactivity.
*$p < 0.05$.

Supplemental Table 2. Associations between *FKBP5* rs1360780, *FKBP5* methylation factor 2, disorganized behavior, and extreme insensitivity on cortisol reactivity during the Strange Situation Procedure (*N* = 298)

Model	*B*	*(95% CI)*	β
FKBP5 rs1360780	1.02	(0.11; 1.93)	0.12*
FKBP5 methylation factor 2	3.16	(-34.33; 40.64)	0.02
Disorganized behavior	-0.04	(-0.36; 0.27)	-0.02
Extreme insensitivity	-0.22	(-0.98; 0.55)	-0.03
FKBP5 rs1360780 × *FKBP5* methylation factor 2	7.68	(-18.39; 33.75)	0.03
FKBP5 rs1360780 × disorganized behavior	0.11	(-0.42; 0.65)	0.02
FKBP5 methylation factor 2 × disorganized behavior	1.66	(-6.96; 10.27)	0.02
FKBP5 rs1360780 × extreme insensitivity	-0.84	(-2.22; 0.55)	-0.07
FKBP5 methylation factor 2 × extreme insensitivity	3.84	(-19.22; 26.91)	0.02
Disorganized behavior × extreme insensitivity	-0.27	(-0.81; 0.27)	-0.06

Note. FKBP5 rs1360780: CC = 0, CT = 1, TT = 2. Analyses are adjusted for technical methylation covariates, cell type proportions of DNA methylation sample, infant age at assessment of SSP, infant sex, educational level of the mother and maternal smoking during pregnancy. The statistics are derived from the final block of the regression model.
None of the possible three-way interactions was significantly associated with cortisol reactivity.
*$p < 0.05$.



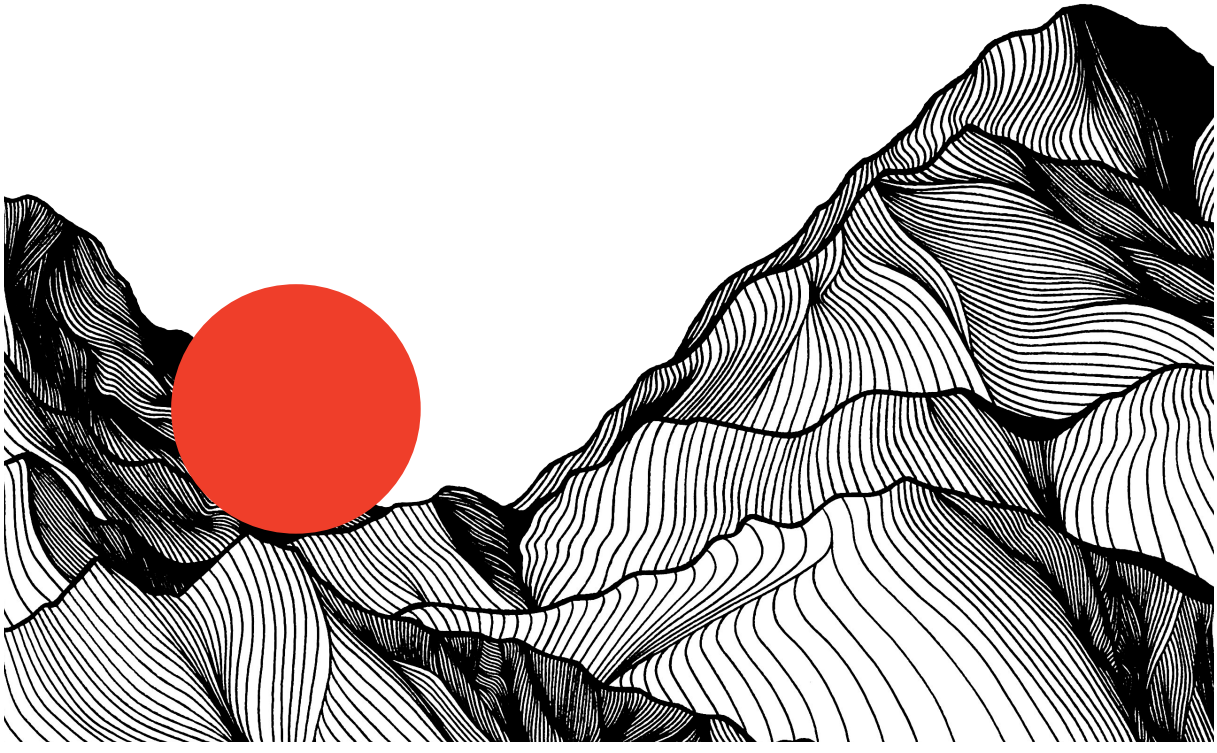
CHAPTER IV

Epigenome-wide associations between observed maternal sensitivity and offspring DNA methylation: A population-based prospective study in children

Lorenza Dall'Aglio, Jolien Rijlaarsdam*, Rosa H. Mulder*, Alexander Neumann, Janine F. Felix, Rianne Kok, Marian J. Bakermans-Kranenburg, Marinus H. van IJzendoorn, Henning Tiemeier, & Charlotte A. M. Cecil

*authors contributed equally

Psychological Medicine, accepted for publication.



Abstract

Background: Experimental work in animals has shown that DNA methylation, an epigenetic mechanism regulating gene expression, is influenced by typical variation in maternal care. While emerging research in humans supports a similar association, studies to date have been limited to candidate gene and cross-sectional approaches, with a focus on extreme deviations in the caregiving environment.

Methods: Here, we investigated the prospective association between typical variation in maternal sensitivity and offspring epigenome-wide DNA methylation, in a population-based cohort of children (*N* = 235). Maternal sensitivity was observed when children were three- and four-years-old. DNA methylation, quantified with the Infinium 450K array, was extracted at age six (whole blood). The influence of methylation quantitative trait loci (mQTLs), DNA methylation at birth (cord blood), and confounders (socioeconomic status, maternal psychopathology) was considered in follow-up analyses.

Results: Genome-wide significant associations between maternal sensitivity and offspring DNA methylation were observed at 13 regions ($p < 1.06x10^{-07}$), but not at single sites. Follow-up analyses indicated that associations at these regions were in part related to genetic factors, confounders, and baseline DNA methylation levels at birth, as evidenced by the presence of mQTLs at five regions and estimate attenuations. Robust associations with maternal sensitivity were found at four regions, annotated to *ZBTB22*, *TAPBP*, *ZBTB12*, and *DOCK4*.

Conclusions: These findings provide novel leads into the relationship between typical variation in maternal caregiving and offspring DNA methylation in humans, highlighting robust regions of associations, previously implicated in psychological and developmental problems, immune functioning, and stress responses.

Introduction

Parental sensitivity, i.e. the responsiveness to children's signals and communications, is an important predictor of developmental outcomes across the behavioral, emotional, and cognitive domains[1–3]. Low sensitivity of primary caregivers–typically mothers–has been associated with a host of negative outcomes, including higher risk for child psychopathology[4,5], externalizing and internalizing problems[1,6], and lower cognitive abilities[7]. This influence can be long-lasting, as shown by prospective human studies[8,9] and experimental work in animals[10]. Yet, the molecular mechanisms underlying the enduring effects of maternal care on neurodevelopmental and behavioral outcomes in humans remain unclear.

Previous studies have provided initial support for DNA methylation–an epigenetic modification regulating gene expression–as a mechanism of interest for these processes[11–13]. DNA methylation involves the addition of a methyl group to DNA base pairs, primarily to the 5-carbon of cytosine nucleotides, resulting in 5-methylcytosine. DNA methylation is sensitive to both environmental and genetic influences[13–15], with the latter being evidenced by changes in the methylome due to DNA variation, named methylation quantitative trait loci (mQTLs)[16]. Further, DNA methylation plays an essential role in healthy development and functioning by modulating the programming of wider biological systems (e.g. neural and immune functioning) and by coordinating key cellular processes (e.g. tissue differentiation)[17]. DNA methylation might thus represent a mechanism by which genetic and environmental factors, including the early caregiving environment, jointly and/or independently predict developmental outcomes[14].

Most evidence of maternal care effects on DNA methylation comes from animal models. In a seminal study by Weaver et al.[13], low levels of maternal care in the first week of life – operationalized as the frequency of licking/grooming and arched-back nursing behaviors– altered DNA methylation patterns in offspring at the glucocorticoid receptor (*GR*, also known as *NR3C1*), a key regulator of stress response[18]. Importantly, these epigenetic changes were long-lasting, but could be reversed via cross-fostering or chemical interventions, leading to a normalization of physiological and behavioral responses to stress[13,19]. These findings generated widespread interest, as they indicated (i) a causal role of maternal care on offspring's epigenetic dysregulation and downstream phenotypes, independent of genetic liability, and (ii) the possibility of influencing developmental trajectories through environmental interventions, mediated by DNA methylation. Since this initial work, other studies have replicated the effects of maternal care on gr methylation in rodents[20] and extended findings to demonstrate DNA methylation changes in other tissues and genes[21–23] (e.g. brain derived neurotrophic factor (*BDNF*) and oxytocin receptor (*OXTR*)) as well as in other species such as rhesus macaques[24].

Although rodents and primates widely differ from humans in their caregiving, a number of similarities in maternal-infant relationships have been observed across mammalian species[25,26].

Parallels at the sensory, hormonal, behavioral, and brain circuit levels have been noted[25–27], including the touch-based behavior characterizing rodents, primates, and humans in the early caregiving and the involvement of the limbic network in maternal-infant relationships[25]. Guided by the animal literature, a growing number of studies have sought to determine the extent to which different forms of caregiving and adversities affect DNA methylation in humans.

Human studies have focused on different forms of adversities[28] including poly-victimization[29], and on extreme deviations in the early caregiving environment, such as maltreatment[30–34], institutionalization[35], and maternal psychopathology[36]. Generally, literature focusing on the caregiving environment has provided preliminary support in line with animal findings, identifying, for example, similar increases in GR methylation in both *postmortem* hippocampal tissue and peripheral tissues from individuals exposed to childhood maltreatment or early-life stress[20]. Studies also indicate that epigenetic patterns associated with the caregiving environment extend beyond *GR*, implicating other genes related to, among other processes, neurodevelopment and stress, such as *OXTR* and *BDNF*. Moreover, by leveraging epigenome-wide DNA methylation, novel genes were identified (e.g. *KCNQ2*, *miR124-3*) in relation to maltreatment and child abuse in individuals with post-traumatic stress disorder[32], borderline personality disorder[33], and depression[34].

While these results are promising and suggest a role of the caregiving environment in the human methylome, the current evidence in humans is limited in a number of key ways. First, since research has mostly focused on extreme deviations in the caregiving environment in selected samples, little is known about how typical variation in maternal sensitivity associates with offspring DNA methylation in the general population. Second, while studies on extreme deviations in maternal care have leveraged epigenome-wide approaches, literature on normative variation in maternal care has solely focused on candidate genes. This has impeded the identification of novel DNA methylation loci associated with maternal sensitivity, which might instead be detected with a hypothesis-free approach by performing an epigenome-wide association study (EWAS). Third, studies have typically relied on cross-sectional designs, in which the early caregiving environment is measured retrospectively via the use of questionnaires, raising doubts about the directionality of observed associations and about the validity of measurements, which may be prone to recall bias[37,38]. Moreover, previous studies rarely investigated whether the identified associations may be confounded by genetic background shared between parents and offspring. The examination of the relationship between maternal care and DNA methylation might indeed capture intergenerational genetic transmission. Lastly, the influence on offspring DNA methylation of factors preceding postnatal maternal care, including the prenatal environment, remains unexplored.

To address these gaps, we firstly examined how typical variation in observed maternal sensitivity prospectively associates with epigenome-wide DNA methylation patterns in a

general population of children. Secondly, with a series of follow-up analyses, we explored the extent to which associations reflected genetic influences as well as confounding by "baseline" DNA methylation levels at birth, which precede exposure to postnatal maternal care and might constitute a biological indicator of the prenatal environment as well as of genetic effects on the methylome.

Materials and Methods

Participants

The present research was embedded in the Generation R Study, a prospective population-based cohort study from fetal life onwards in Rotterdam, The Netherlands[39] (Supplemental Information 1). Ethical approval was obtained from the Medical Ethics Committee of Erasmus MC, University Medical Center Rotterdam. For the purposes of this study, children within the Generation R Study with data on maternal sensitivity (at three and/or four years) and DNA methylation (at six years) were selected (*N* = 235). Since 5 sibling-pairs were present, we later excluded one sibling per pair (*N* = 230) to ensure genetic relatedness did not impact results.

Participant characteristics are shown in Supplemental Table 1. Participants with data on both maternal sensitivity and DNA methylation (age six) differed from participants invited to the age six assessment in gestational age at birth ($M_{subsample}$ = 40.3 weeks (*SD* = 1.4), $M_{fullsample}$ = 39.8 (*SD* = 1.9), *t* = 5.6, *p* = $6.50x10^{-08}$), but not other covariates.

Measures

Maternal sensitivity

Maternal sensitivity was assessed at ages three and four years through observations of mother-child interactions during teaching tasks too complex for the age of the child. These involved (i) building a tower and (ii) completing an etch-a-sketch drawing. Mother-child interactions were recorded and subsequently coded, according to the revised Erickson 7-point rating scales 40, based on two interdependent subscales: intrusiveness (IN) and supportive presence (SP), which together form the maternal sensitivity construct. Inter-coder reliability amounted to 0.81 at age three and 0.84 at age four 41.

Eight measures of maternal sensitivity (i.e. IN and SP scales x two tasks x two time-points) were available. IN scores were reversed, and both IN and SP scores were standardized. An overall maternal sensitivity score was calculated, for participants with data at age three and/or four, by averaging such standardized measures[42]. This was done in line with previous literature[41], due to the stability of the maternal sensitivity scores between age three and four years 1, the temporality of these assessments, which both precede DNA methylation at age

six, and to maximize our sample size. Cronbach's alpha reliability of the obtained measure was acceptable (Cronbach's $\alpha = 0.70$)[43].

DNA methylation

DNA methylation in whole blood at age six was used for our epigenome-wide analyses. This was selected due to it being the closest DNA methylation assessment after maternal sensitivity observations (age three and four years), and to test the prospective association of maternal sensitivity with DNA methylation. Based on previous studies in animals, which found maternal care to have long-lasting influences on the methylome[13], we expected for maternal care effects to endure in early childhood.

To obtain DNA methylation data, DNA extraction and bisulfite conversion via the EZ-96 DNA Methylation kit (Shallow) (Zymo Research Corporation, Irvine, USA) were performed, and samples were processed with the Illumina Infinium HumanMethylation450 BeadChip (Infinium 450K), which measures 485 577 CpGs. The incorporating control probe adjustment and reduction of global correlation pipeline[44] was employed for the preparation and normalization of the data using R. Firstly, the minfi package[45] in R was used to read the idat files. Probes that had a detection p-value above background (based on the sum of methylated and unmethylated intensity values) ≥ 1E-16 were set to missing per array. Next, the intensity values were stratified by autosomal and non-autosomal probes and quantile normalized for each of the six probe-type categories separately: type II red/green, type I methylated red/green and type I unmethylated red/green. For each probe, DNA methylation levels were indexed by beta values (i.e. the ratio of methylated signal divided by the sum of the methylated and unmethylated signal [M / (M + U + 100)]). Quality control procedures were additionally performed (e.g. check for sex mismatch). Only arrays with a call rate above 95% per sample were considered for additional processing. DNA methylation data was winsorized (> 3 *SD*) to reduce the influence of potential outliers. In total, we obtained information on 457,872 autosomal sites in 493 six-year-olds.

We additionally used DNA methylation data collected at birth in cord blood for a follow-up analysis. This was subject to the same pipeline as the DNA methylation data at age six and was also measured based on the Infinium 450K BeadChip. Only CpGs identified as significant or within DNA methylation significant regions were selected for these analyses.

Covariates

All analyses were adjusted for a key set of potential covariates guided by previous literature[46–49], including batch effects (plate number), sex, gestational age at birth, maternal smoking during pregnancy (never smoked, smoked until pregnancy known, continued during pregnancy), and estimated cell-type proportions[50] (Supplemental Information 1). We additionally adjusted for two sets of covariates: (i) maternal education (highest level completed) as proxy for

socioeconomic status, and postnatal maternal psychopathology (Brief Symptom Inventory), and (iii) DNA methylation levels at birth (cord blood tissue), together with respective cell-type and batch effect adjustments (Supplemental Information 1).

Statistical Analyses

Analyses were performed in R (version 4.0.0) and are described in-depth in Supplemental Information 1. A *probe-level EWAS* (multiple linear regression models) was run with the CpGassoc R package[51], to test for associations of maternal sensitivity with each DNA methylation site individually (Bonferroni epigenome-wide significance threshold: $p < 1.09x10^{-07}$). To account for potential bias and inflation, the BACON R package[52] was used.

Moreover, to capture correlations across CpGs, reduce data dimensionality, and attenuate the multiple testing burden, a *regional-level EWAS* was performed by using the R package DMRff[53]. This estimates correlations across nominally-significant probes within a 500 bp window (default setting) and combines the EWAS summary statistics of such neighboring CpGs to identify differentially methylated regions while accounting for multiple testing with a Bonferroni procedure in both gene regions and sub-regions[54].

A *candidate gene look-up* was also performed to maximize comparability with previously reported DNA methylation-maternal care associations. To date, DNA methylation levels of four genes have been associated with maternal care in humans[55–58], by at least one study: *GR, BDNF*, the serotonin receptor (*SLC6A4*), and *OXTR*. We looked-up the EWAS results for probes located within these genes, as annotated in the HumanMethylation450 v1.2 Manifest File. Following previous studies[29,59], gene-level Bonferroni correction was used as significance threshold (i.e. $p < 0.05$/number of annotated probes).

To identify enriched biological pathways, we performed an in-house *gene ontology (GO) analysis*[59–61] on sites with $p < 0.001$ in the probe-level EWAS, in line with previous literature[59,60,62,63]. We performed *p*-value adjustments based on default procedures 61. Enriched pathways were confirmed by an independent GO approach from the missMethyl R package[64] ($p < 0.05$).

Finally, a series of *follow-up analyses* were run. Firstly, the influence of genetic factors on our top hits (i.e. Bonferroni-significant sites or sites within Bonferroni-significant DNA methylation regions) was assessed by drawing on an *mQTL database*[16] (www.mqtldb.org). We examined whether hits were associated with known mQTLs during childhood, based on the results from a genome-wide complex trait conditional analysis. Secondly, we explored the robustness of top hits to *additional adjustments* for (i) postnatal maternal education and maternal psychopathology ($N = 223$) and (ii) pre-exposure DNA methylation ($N = 226$). The latter was done to account for the effect of DNA methylation at birth on DNA methylation at age six and to capture potential pre-existing influences (e.g. intrauterine exposures) on DNA

methylation in childhood. Spearman correlations between DNA methylation at birth and age six were also calculated, per CpG. Thirdly, based on a list of our CpG hits, the in-house *gene ontology analysis* and missMethyl validation were run, with the same procedures as the main GO analysis specified above. Finally, to understand the relevance of our findings to the brain, which is linked to the caregiving environment[13,41], we looked up *brain-blood concordance* values for our top hits using the BECon online tool (https://redgar598.shinyapps.io/BECon/)[65].

Results

Probe-level EWAS

Maternal sensitivity was not associated with any single CpGs at age six, after genome-wide correction ($p < 1.09x10^{-07}$) (Figure 1, Supplemental Table 2). BACON analysis revealed a normal lambda (λ = 1.00), minimal bias (Bayesian estimate of bias = –0.002) and deflation in the test results – indicative of low power (Bayesian inflation factor = 0.925) (Supplemental Figure 1). Following BACON correction for deflation, one intergenic CpG reached genome-wide significance: cg25628898 (estimate = –0.008; SE = 0.002; $p = 1.03x10^{-07}$) (Supplemental Table 2).

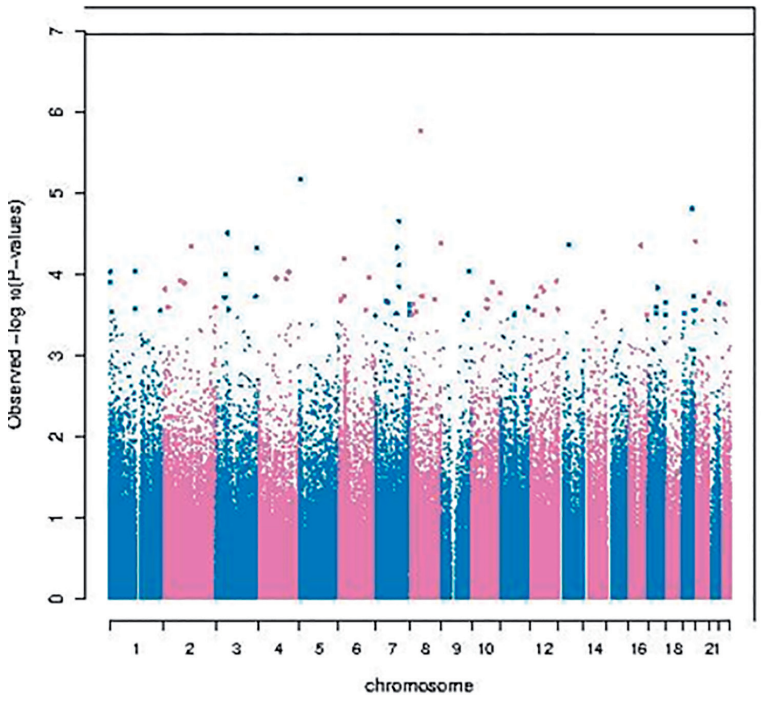


Figure 1. Manhattan plot of CpG sites associated with maternal sensitivity
Note. The Manhattan plot displays the log *p*-values for each site tested in association with maternal sensitivity in the EWAS, across autosomal chromosomes. No genome-wide significant association was observed ($p < 1.09x10^{-07}$).

Regional-level EWAS

With a regional-level EWAS, we identified 13 DNA methylation regions associated with maternal sensitivity ($p < 1.06x10^{-07}$; $\alpha = 0.05$) (Table 1, Figure 2, Supplemental Table 3), spanning 143 CpGs. The top three DNA methylation regions coincided with the *ANKMY1*, *RNF39*, and *ZBTB22* and *TAPBP* genes (Table 1). The largest estimates were shown at regions encompassing *COLEC11* and *DOCK4*. None of the CpGs within our significant regions was related to prenatal maternal smoking, based on previous research in neonates and children[66,67], suggesting adjustments in the EWAS accounted for its confounding role. When siblings ($N = 230$) were excluded all but one region (annotated to *RNF5P1*, *RNF5*, *AGPAT1*) remained significantly associated with maternal sensitivity.

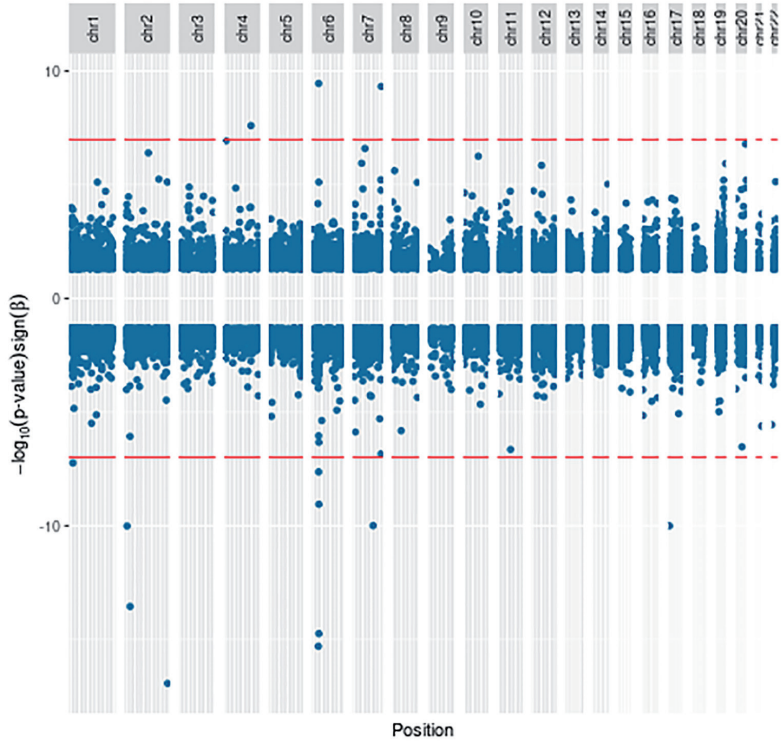


Figure 2. Miami plot of DNA methylation regions associated with maternal sensitivity
Note. The Miami plot displays the log *p*-values and estimates direction for each DNA methylation region tested in association with maternal sensitivity, across autosomal chromosomes. Thirteen regions were Bonferroni significant, three of which showed a positive relation with maternal sensitivity and 10 a negative one.

Table 1. DNA methylation regions significantly associated with maternal sensitivity from the regional-level EWAS

DNAm region location	Annotated gene(s)	N CpGs included	Estimate	Standard error	Raw *p*-value	Bonferroni Adj. *p*-value
chr2: 241458886-241460002	*ANKMY1*	8	0.365	0.043	1.17E-17	5.61E-12
chr6: 30039027-30039600	*RNF39*	22	-0.227	0.028	5.03E-16	2.42E-10
chr6: 33282879-33283184	*ZBTB22; TAPBP*	17	-0.215	0.027	1.83E-15	8.77E-10
chr2: 21266727-21267334	*APOB*	10	-0.302	0.040	2.83E-14	1.36E-08
chr2: 3642629-3642867	*COLEC11*	6	-0.875	0.135	9.80E-11	4.71E-05
chr17: 6797034-6797771	*ALOX12P2*	6	-0.571	0.088	1.00E-10	4.80E-05
chr7: 111368367-111368847	*DOCK4*	4	-0.822	0.127	1.02E-10	4.90E-05
chr6: 32145383-32146595	*RNF5P1; RNF5; AGPAT1**	27	0.047	0.007	3.55E-10	0.000171
chr7: 158749953-158751591	*Non-annotated region*	8	0.558	0.090	4.80E-10	0.000231
chr6: 33280149-33280436	*TAPBP*	9	-0.282	0.046	8.89E-10	0.000427
chr6: 31867757-31868169	*ZBTB12*	19	-0.100	0.018	2.35E-08	0.011285
chr4: 147164778-147165097	*Non-annotated*	4	0.427	0.077	2.53E-08	0.012128
chr1: 11714218-11714254	*FBXO44; FBXO2*	3	-0.439	0.081	5.82E-08	0.027955

DNAm region location = Genomic location of the DNA methylation region (chromosome, start position, and end position)
Annotated gene(s) = Gene(s) annotated to the CpGs within the DNA methylation region
N CpGs included = Number of CpGs included in the DNA methylation region
Estimate = Estimate for the association of maternal sensitivity with DNA methylation at a region
Standard error = Standard error for the association of maternal sensitivity with DNA methylation at a region
Raw *p*-value = Unadjusted *p*-value for the association of maternal sensitivity with DNA methylation at a region
Bonferroni adj. *p*-value = *P*-value adjusted for multiple testing with Bonferroni correction
*This region was not genome-wide significant when siblings were excluded from the sample.

Candidate gene look-up

The candidate gene look-up showed that, of the four selected genes (*NR3C1*, *BDNF*, *SLC6A4*, *OXTR*), which included 14 to 74 sites, no CpG met Bonferroni-adjusted gene-wide significance in association with maternal sensitivity (Table 2, Supplemental Table 4). Only three sites reached nominal significance ($p < 0.05$).

Gene ontology

The in-house GO analysis, based on sites with $p < 0.001$ in the probe-level EWAS, revealed enrichment for 148 pathways. Yet, this threshold might have been overinclusive. Thirty-nine of the 148 pathways were confirmed by the missMethyl GO method ($p < 0.05$) (Supplemental Table 5). Both methods indicated enrichment for, among others, calcium ion channels functioning, phosphorylation, and tissue and cell polarity.

Follow-up analyses

Firstly, an mQTL search revealed that five of the 13 significant DNA methylation regions contained at least one CpG associated with one or more known SNPs (Table 3, Supplemental

Table 6). Eight regions, including *ZBTB22*/*TAPBP* (one of our top regions), did not present any mQTLs. Of the 143 sites within the 13 significant regions, 22% (*n* = 31) associated with one or more known SNPs. All associations were in cis.

Secondly, after additional adjustments for socioeconomic status and maternal psychopathology, associations attenuated at seven regions (median = -1%, range = –44%–13%). Regions which did not decrease in effect were *TAPBP*, *RNF39*, two non-annotated regions, *ANKMY1*, and *ALOX12P2* (Supplemental Table 7). When adjusting for pre-exposure DNA methylation levels, (Supplemental Table 8), associations attenuated at ten regions (median = -45%, range = -97%–17%), with *RNF39* being the most affected. Regions whose estimates did not decrease were *ZBTB12*, *FBXO44*/*FBXO2*, and a non-annotated region (chromosome 7). The median correlation between each CpG DNA methylation levels at birth and age six was of Rho = 0.43 (range: 0.11 – 0.86) (Supplemental Table 9).

Thirdly, in a follow-up GO analysis, based on the sites within the significant DNA methylation regions (*n* = 143), enrichment was found at 63 pathways (in-house method). Of these, 33 were validated by missMethyl ($p < 0.05$). Both methods indicated enrichment for, among others, several lipoprotein processes (e.g. particle remodeling), and peptide binding (Supplemental Table 10).

Lastly, of the 13 significant DNA methylation regions, six contained half or more sites with greater than average blood-brain tissue concordance 65 in at least one brain tissue (for BA7 $r > |.36|$, for BA10 $r > |.40|$, for BA20 $r > |.33|$), for a total of 67 sites (Supplemental Table 11) (not empirically tested).

Table 2. The association between maternal sensitivity and DNA methylation: Candidate gene look-up

Gene	Chr	N probes	Gene-level sign.	Nominal sign.	Estimate range	%Positive associations	% Negative associations
NR3C1	5	40	No	Yes (cg17342132)	-0.004–0.006	65%	35%
BDNF	11	74	No	Yes (cg26840770)	-0.010–0.005	50%	50%
SLC6A4	17	14	No	Yes (cg06841846)	-0.004–0.005	29%	71%
OXTR	3	18	No	No	-0.006–0.006	56%	44%

Gene = Candidate gene
Chr = Chromosome
N probes = Number of probes annotated to the gene (based on the Infinium 450K)
Gene-level sign. = Gene-level Bonferroni significance in any of the probes annotated to the candidate gene ($p < 0.05$/number of annotated probes)
Nominal sign. = Nominal significance in any of the probes annotated to the candidate gene ($p < 0.05$)
Estimate range = Range of estimates for the probes annotated to the candidate genes
% positive associations = Percentage of probes with a positive association with maternal sensitivity
% negative associations = Percentage of probes with a negative association with maternal sensitivity

Table 3. mQTLs within the statistically significant DNA methylation regions

DNAm region location	Annotated gene(s)	N CpGs included	N mQTL associations	N CpGs with mQTLs	%CpGs with mQTLs
chr2: 241458886-241460002	*ANKMY1*	8	16	7	88%
chr6: 30039027-30039600	*RNF39*	22	0	0	0%
chr6: 33282879-33283184	*ZBTB22; TAPBP*	17	0	0	0%
chr2: 21266727-21267334	*APOB*	10	19	10	100%
chr2: 3642629-3642867	*COLEC11*	6	6	6	100%
chr17: 6797034-6797771	*ALOX12P2*	6	0	0	0%
chr7: 111368367-111368847	*DOCK4*	4	0	0	0%
chr6: 32145383-32146595	*RNF5P1; RNF5; AGPAT1*	27	0	0	0%
chr7: 158749953-158751591	*Non-annotated region*	8	5	5	63%
chr6: 33280149-33280436	*TAPBP*	9	0	0	0%
chr6: 31867757-31868169	*ZBTB12*	19	0	0	0%
chr4: 147164778-147165097	*Non-annotated region*	4	0	0	0%
chr1: 11714218-11714254	*FBXO44; FBXO2*	3	3	3	100%
Total		**143**	**49**	**31**	**22%**

DNAm region location = Genomic location of the DNA methylation region (chromosome, start position, and end position)
Annotated gene(s) = Gene(s) annotated to the DNA methylation region
N CpGs included = Number of CpGs included in the DNA methylation region
N mQTL associations = Number of SNPs – DNA methylation associations at a region
N CpGs with mQTLs = Number of CpGs presenting one or more mQTL(s) at a region
% CpGs with mQTLs = Percentage of CpGs presenting one or more mQTL(s) at a region

Discussion

This is the first epigenome-wide study investigating the prospective association between typical variation in maternal sensitivity (observed) and offspring DNA methylation, in a general population of children. Genome-wide significant associations were observed at 13 DNA methylation regions, four of which did not contain mQTLs and were minimally affected by adjustments for postnatal confounders and by pre-exposure DNA methylation levels, thus showing robustness in associations.

Summary of Key Findings

Our first aim was to examine the prospective relationship between maternal sensitivity and child DNA methylation using complementary approaches. Firstly, no individual CpG was identified in the probe-level EWAS after genome-wide correction. This might indicate that associations at a site-level are subtle and challenging to identify, especially considering this study assessed typical variation in maternal care as opposed to extreme deviations (e.g. abuse). The high multiple testing correction burden that probe-level EWASs entail may also impede the detection of single sites of small effect, which could be uncovered with larger

samples. For instance, with our sample (N = 235) and model (multiple linear regression, 10 predictors), 80% power, and a genome-wide threshold, only moderate estimates (as small as 0.27) could be detected.

When employing a regional approach, which can detect weaker but more widespread signals by accounting for correlations across CpGs, 13 DNA methylation regions were significantly associated with maternal sensitivity ($p < 1.06x10^{-07}$, $\alpha = 0.05$). These findings support the presence of offspring methylomic signatures of maternal care, which may be best uncovered through hypothesis-free approaches with methods capturing the correlational patterns of DNA methylation. Yet, replication of these findings is needed, and the possibility of false-positive findings should not be excluded. Notably, when considering a more stringent significance threshold ($p < 2.18x10^{-09}$; $\alpha = 0.001$), as suggested to reduce false-positive rates 68, most of the regions (77%, N = 10) remained significantly related to maternal sensitivity.

Further, we failed to detect an association between maternal sensitivity and DNA methylation variation at candidate genes previously identified by studies of maternal care in humans[55–58]. Inconsistencies may reflect several factors, including differences in sample characteristics (e.g. psychiatric vs. population-based samples), maternal care assessments (retrospective vs. prospective reports) and analysis (e.g. gene regions covered by pyrosequencing vs. Infinium 450K). Lastly, candidate gene studies may be particularly vulnerable to false positives, as shown in the genetic field[69].

As a second aim, we explored whether identified maternal sensitivity-DNA methylation associations may be influenced by genetic factors, based on mQTL mapping. Twenty-two percent of the sites in our significant regions were linked to known SNPs. This suggests that associations for those sites may be in part confounded by genetic factors and corroborates previous research highlighting DNA methylation responsiveness to both external exposures and genetic variation[14]. However, the presence of mQTLs alone does not preclude environmental effects. Indeed, recent studies have found that interindividual variability in DNA methylation is primarily explained by gene-environment combinations (additive and interactive effects)[70,71]. Moreover, mQTLs were identified based on a publicly available database, as our sample was underpowered to directly test for genetic confounding. Future studies employing genetically-sensitive designs could more precisely quantify the effect of maternal sensitivity on DNA methylation by directly modeling genetic influences.

When exploring the robustness of findings to additional adjustments, we observed attenuations at half of the regions, after controlling for socioeconomic status and maternal psychopathology. When considering pre-exposure DNA methylation levels, estimates attenuated at most regions. Although neonatal methylomic patterns were measured in cord blood at birth and not in peripheral blood (used at age six), which may lead to additional differences, these findings

indicate that associations partly reflected pre-existing DNA methylation levels. This was clearly exemplified by *RNF39*, a region strongly associated with sensitivity, robust to postnatal confounders, and genetic influences. After adjustments, its estimate reduced by 97%, showing that associations did not result from postnatal caregiving, as they were already present at baseline (birth). These findings cast doubts on previous studies of caregiving which did not consider pre-exposure DNA methylation levels, and raise questions on the directionality of associations between maternal care and DNA methylation, as well as on the potential role of other confounders affecting child DNA methylation at birth and in childhood, and maternal sensitivity (e.g. shared genetics, maternal distress).

Here, we highlight four "high-confidence" associations with maternal caregiving, which were not linked to any mQTLs, and were most robust to adjustments for confounders and pre-exposure DNA methylation levels. These spanned (i) *ZBTB22*/*TAPBP*, (ii) *ZBTB12*, (iii) *DOCK4*, and (iv) a non-annotated region in chromosome four. All four genes are protein-coding 18. *DOCK4* is implicated in neuronal processes, such as neuronal migration, and dendritic arborization[72] and its DNA methylation region presented higher than average blood-BA10 concordance in this study. *ZBTB22* and *ZBTB12* are involved in transcriptional regulation and nuclear chromatin localization[73]. These two genes, together with *TAPBP*, are within the Major Histocompatibility Complex (MHC). While these associations should be carefully interpreted as the MHC is characterized by extensive linkage disequilibrium[74], this genomic region plays an important role in immune functioning and has been implicated in neuronal plasticity[75,76]. *TAPBP* specifically is involved in MHC class I protein complex assembly, gene expression regulation, and immunodeficiency[73]. In this study, enrichment for MHC class I protein assembly and peptide binding was found for maternal sensitivity, suggesting that such exposure might enact on *TAPBP*-related functions via DNA methylation.

Generally, our high-confidence genes have been previously associated with psychological and developmental problems, inflammation, and stress-responses. Molecular changes were shown at *TAPBP* for major depressive disorder and suicide[77], *TAPBP* and *DOCK4* for schizophrenia[78–80], *ZBTB22* for intellectual disability[73] and psychopathologies following hypercortisolism[81], and *DOCK4* for autism and dyslexia[82,83]. Enrichment for pathways including Dock4 has been repeatedly associated with stress-related responses in mice[84–86], while *ZBTB12* DNA methylation is related to markers of inflammation (e.g. white blood cell counts)[87].

Limitations and Suggestions for Future Research

Our findings should be interpreted in light of several limitations. Firstly, identified associations may have been influenced by additional parental factors that we could not control for in the present study, either because this information was not available (e.g. parental temperament, parental genotype) or due to the low number of cases (e.g. maternal medication and substance use in pregnancy). Nevertheless, we did control for the most important maternal confounders

(smoking during pregnancy, socioeconomic status, psychopathology). Secondly, if unmeasured changes in maternal sensitivity and covariates occurred during the two-to-three-year time-lag between our exposure and outcome, noise would be introduced in the identified associations. A prospective design, as opposed to a cross-sectional one, remains however preferable due to the possibility to better understand the directionality of associations. Nonetheless, repeated postnatal measurements of both DNA methylation and maternal sensitivity would be ideal to longitudinally examine how associations change over time and disentangle directionality. Thirdly, we did not have information on whether the mothers included in this study were primary or secondary caregivers (at four years only). Yet, within Generation R, most mothers are primary caregivers[88]. Additionally, while the use of the Infinium 450K provided novel insights into the genes affected by maternal sensitivity, future research should employ, when possible, the EPIC 850K array due to its wider and more diverse genomic coverage[89]. Lastly, our investigation solely focused on the association of maternal sensitivity on the child methylome. Related molecular signatures, such as transcription changes and epigenetic clocks, could be examined in future research to better understand the biological consequences of maternal care.

In conclusion, this population-based study supports a prospective association of typical variation in maternal sensitivity with epigenome-wide DNA methylation in children. We highlight four DNA methylation regions that showed the strongest associations with maternal sensitivity as well as minimal evidence of genetic and pre-exposure influences, and which should thus be prioritized in future research. These results permit further delineation of the relationship between DNA methylation and maternal care in humans and warrant confirmation by future research with large, longitudinal, and genetically-sensitive studies.

References

1. Kok R, Linting M, Bakermans-Kranenburg MJ, et al. Maternal sensitivity and internalizing problems: Evidence from two longitudinal studies in early childhood. *Child Psychiatry & Human Development.* 2013;44(6):751-765.
2. Madigan S, Prime H, Graham SA, et al. Parenting behavior and child language: A meta-analysis. *Pediatrics.* 2019;144(4).
3. Thomas JC, Letourneau N, Campbell TS, Tomfohr-Madsen L, Giesbrecht GF. Developmental origins of infant emotion regulation: Mediation by temperamental negativity and moderation by maternal sensitivity. *Developmental Psychology.* 2017;53(4):611-628.
4. Kimbrel NA, Nelson-Gray RO, Mitchell JT. Reinforcement sensitivity and maternal style as predictors of psychopathology. *Personal Individual Differences.* 2007;42(6):1139-1149.
5. Haltigan JD, Roisman GI, Fraley RC. The predictive significance of early caregiving experiences for symptoms of psychopathology through midadolescence: Enduring or transient effects? *Development and Psychopathology.* 2013;25(1):209-221.
6. Rijlaarsdam J, Stevens GWJM, Jansen PW, et al. Maternal childhood maltreatment and offspring emotional and behavioral problems: Maternal and paternal mechanisms of risk transmission. *Child Maltreatment.* 2014;19(2):67-78.
7. Bernier A, Carlson SM, Deschênes M, Matte-Gagné C. Social factors in the development of early executive functioning: a closer look at the caregiving environment. *Developmental Science.* 2012;15(1):12-24.
8. Raby KL, Roisman GI, Fraley RC, Simpson JA. The enduring predictive significance of early maternal sensitivity: Social and academic competence through age 32 years. *Child Development.* 2015;86(3):695-708.
9. Stams G-JJM, Juffer F, van IJzendoorn MH. Maternal sensitivity, infant attachment, and temperament in early childhood predict adjustment in middle childhood: The case of adopted children and their biologically unrelated parents. *Devevlopmental Psychology.* 2002;38(5):806-821.
10. Meaney MJ. Maternal care, gene expression, and the transmission of individual differences in stress reactivity across generations. *Annual Review of Neuroscience.* 2001;24(1):1161-1192.
11. Mulder RH, Rijlaarsdam J, van IJzendoorn MH. DNA methylation: A mediator between parenting stress and adverse child development? In: Deater-Deckard K, Panneton R, eds. *Parental Stress and Early Child Development: Adaptive and Maladaptive Outcomes.* Springer International Publishing; 2017:157-180.
12. Szyf M. The genome- and system-wide response of DNA methylation to early life adversity and its implication on mental health. *The Canadian Journal of Psychiatry.* 2013;58(12):697-704.
13. Weaver ICG, Cervoni N, Champagne FA, et al. Epigenetic programming by maternal behavior. *Nature Neuroscience.* 2004;7(8):847-854.
14. Ladd-Acosta C, Fallin MD. The role of epigenetics in genetic and environmental epidemiology. *Epigenomics.* 2016;8(2):271-283.

15. Smith AK, Kilaru V, Kocak M, et al. Methylation quantitative trait loci (meQTLs) are consistently detected across ancestry, developmental stage, and tissue type. *BMC Genomics.* 2014;15(1):145.
16. Gaunt TR, Shihab HA, Hemani G, et al. Systematic identification of genetic influences on methylation across the human life course. *Genome Biology.* 2016;17:61.
17. Carey N. Life as we know it now. In: *The Epigenetics Revolution.* Columbia University Press; 2012:72-73.
18. Geer LY, Marchler-Bauer A, Geer RC, et al. The NCBI BioSystems database*. Nucleid Acids Research.* 2010;38:D492-6.
19. Weaver ICG, Champagne FA, Brown SE, et al. Reversal of maternal programming of stress responses in adult offspring through methyl supplementation: Altering epigenetic marking later in life. *Journal of Neuroscience.* 2005;25(47):11045-11054.
20. Turecki G, Meaney MJ. Effects of the social environment and stress on glucocorticoid receptor gene methylation: A systematic review. *Biological Psychiatry.* 2016;79(2):87-96.
21. Doherty TS, Forster A, Roth TL. Global and gene-specific DNA methylation alterations in the adolescent amygdala and hippocampus in an animal model of caregiver maltreatment. *Behavioural Brain Research.* 2016;298(0 0):55-61.
22. Beery AK, McEwen LM, MacIsaac JL, Francis DD, Kobor MS. Natural variation in maternal care and cross-tissue patterns of oxytocin receptor gene methylation in rats. *Hormones and Behavior.* 2016;77:42-52.
23. Blaze J, Asok A, Borrelli K, et al. Intrauterine exposure to maternal stress alters Bdnf IV DNA methylation and telomere length in the brain of adult rat offspring. *International Journal of Developmental Neuroscience.* 2017;62:56-62.
24. Provençal N, Suderman MJ, Guillemin C, et al. The signature of maternal rearing in the methylome in rhesus macaque prefrontal cortex and t cells. *Journal of Neuroscience.* 2012;32(44):15626-15642.
25. Feldman R. The neurobiology of mammalian parenting and the biosocial context of human caregiving. *Hormones Behavior.* 2016;77:3-17.
26. Knop J, Joëls M, van der Veen R. The added value of rodent models in studying parental influence on offspring development: opportunities, limitations and future perspectives. *Current Opinion in Psychology.* 2017;15:174-181.
27. Glynn LM, Baram TZ. The influence of unpredictable, fragmented parental signals on the developing brain. *Frontiers in Neuroendocrinology.* 2019;53:100736.
28. Daskalakis NP, Yehuda R. Site-specific methylation changes in the glucocorticoid receptor exon 1F promoter in relation to life adversity: systematic review of contributing factors. *Frontiers in Neuroscience.* 2014;8.
29. Marzi SJ, Sugden K, Arseneault L, et al. Analysis of DNA methylation in young people: Limited evidence for an association between victimization stress and variation in blood. *American Journal of Psychiatry.* 2018;175(6):517-529.
30. Cecil CAM, Smith RG, Walton E, Mill J, McCrory EJ, Viding E. Epigenetic signatures of childhood abuse and neglect: Implications for psychiatric vulnerability. *Journal of Psychiatric Research.* 2016;83:184-194.

31. Gouin JP, Zhou QQ, Booij L, et al. Associations among oxytocin receptor gene (OXTR) DNA methylation in adulthood, exposure to early life adversity, and childhood trajectories of anxiousness. *Scientific Reports.* 2017;7(1):7446.
32. Mehta D, Klengel T, Conneely KN, et al. Childhood maltreatment is associated with distinct genomic and epigenetic profiles in posttraumatic stress disorder. *Proceedings of the National Academy of Science.* 2013;110(20):8302-8307.
33. Stenz L, Prados J, Courtet P, et al. Borderline personality disorder and childhood maltreatment: A genome-wide methylation analysis. *European Psychiatry.* 2016;33(S1):S183-S183.
34. Weder N, Zhang H, Jensen K, et al. Child Abuse, Depression, and Methylation in Genes Involved With Stress, Neural Plasticity, and Brain Circuitry. *Journal of American Academy of Child & Adolescent Psychiatry.* 2014;53(4):417-424.e5.
35. Naumova OYu, Lee M, Koposov R, Szyf M, Dozier M, Grigorenko EL. Differential patterns of whole-genome DNA methylation in institutionalized children and children raised by their biological parents. *Development and Psychopathology.* 2012;24(01):143-155.
36. Oberlander TF, Weinberg J, Papsdorf M, Grunau R, Misri S, Devlin AM. Prenatal exposure to maternal depression, neonatal methylation of human glucocorticoid receptor gene (NR3C1) and infant cortisol stress responses. *Epigenetics.* 2008;3(2):97-106.
37. Reuben A, Moffitt TE, Caspi A, et al. Lest we forget: Comparing retrospective and prospective assessments of adverse childhood experiences in the prediction of adult health. *Journal of Child Psychology and Psychiatry.* 2016;57(10):1103-1112.
38. Baldwin JR, Reuben A, Newbury JB, Danese A. Agreement between prospective and retrospective measures of childhood maltreatment: A systematic review and meta-analysis. *JAMA Psychiatry.* 2019;76(6):584-593.
39. Kooijman MN, Kruithof CJ, van Duijn CM, et al. The Generation R Study: design and cohort update 2017. *European Journal of Epidemiology*. 2016;31(12):1243-1264. doi:10.1007/s10654-016-0224-9
40. Egeland B, Erickson MF, Clemenhagen-Moon J, Hiester MK, Korfmacher J. 24 months tools coding manual. *Project STEEP-revised 1990 from Mother-Child project scales.* Published online 1990.
41. Kok R, Thijssen S, Bakermans-Kranenburg MJ, et al. Normal variation in early parental sensitivity predicts child structural brain development. *Journal of American Academy of Child & Adolescent Psychiatry.* 2015;54(10):824-831.e1.
42. Cents RAM, Kok R, Tiemeier H, et al. Variations in maternal 5-HTTLPR affect observed sensitive parenting. *Journal of Child Psychology and Psychiatry.* 2014;55(9):1025-1032.
43. Cortina JM. What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology.* 1993;78(1):98-104.
44. Lehne B, Drong AW, Loh M, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biology.* 2015;16(1):37.
45. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics.* 2014;30(10):1363-1369.

46. Birney E, Smith GD, Greally JM. Epigenome-wide association studies and the interpretation of disease -omics. *PLoS Genetics.* 2016;12:e1006105.
47. Breton CV, Byun H-M, Wenten M, Pan F, Yang A, Gilliland FD. Prenatal tobacco smoke exposure affects global and gene-specific DNA methylation. *American Journal of Respiratory Critical Care Medicine.* 2009;180(5):462-467.
48. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. *Nature Reviews Genetics.* 2011;12(8):529-541.
49. Liang L, Cookson WOC. Grasping nettles: cellular heterogeneity and other confounders in epigenome-wide association studies. *Human Molecular Genetics.* 2014;23(R1):R83-R88.
50. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics.* 2012;13(1):86.
51. Barfield RT, Kilaru V, Smith AK, Conneely KN. CpGassoc: an R function for analysis of DNA methylation microarray data. *Bioinformatics.* 2012;28(9):1280-1281.
52. van Iterson M, van Zwet EW, Bios Consortium, Heijmans BT. Controlling bias and inflation in epigenome- and transcriptome-wide association studies using the empirical null distribution. *Genome Biology.* 2017;18:19.
53. Suderman M, Staley JR, French R, Arathimos R, Simpkin A, Tilling K. dmrff: identifying differentially methylated regions efficiently with power and control. *bioRxiv.* Published online December 31, 2018:508556.
54. Suderman M, Staley JR, French R, Arathimos R, Simpkin A, Tilling K. Dmrff: Identifying differentially methylated regions efficiently with power and control. *Bioinformatics*; 2018.
55. Conradt E, Hawes K, Guerin D, et al. The contributions of maternal sensitivity and maternal depressive symptoms to epigenetic processes and neuroendocrine functioning. *Child Development.* 2016;87(1):73-85.
56. Unternaehrer E, Meyer AH, Burkhardt SCA, et al. Childhood maternal care is associated with DNA methylation of the genes for brain-derived neurotrophic factor (BDNF) and oxytocin receptor (OXTR) in peripheral blood cells in adult men and women. *Stress.* 2015;18(4):451-461.
57. Provenzi L, Fumagalli M, Giorda R, et al. Maternal sensitivity buffers the association between SLC6A4 methylation and socio-emotional stress response in 3-month-old full term, but not very preterm infants. *Frontiers in Psychiatry.* 2017;8.
58. Bosmans G, Young JF, Hankin BL. NR3C1 methylation as a moderator of the effects of maternal support and stress on insecure attachment development. *Developmental Psychology.* 2018;54(1):29-38.
59. Cecil CAM, Walton E, Jaffee SR, et al. Neonatal dna methylation and early-onset conduct problems: A genome-wide, prospective study. First Posting Jun 9, 2017. *Developmental Psychopathology* 2017 Pp No Pagination Specified Jun Dev Psychopathol. Published online 2017.
60. Cecil CAM, Walton E, Pingault J-B, et al. DRD4 methylation as a potential biomarker for physical aggression: An epigenome-wide, cross-tissue investigation. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics.* 2018;177(8):746-764.

61. Hannon E, Dempster E, Viana J, et al. An integrated genetic-epigenetic analysis of schizophrenia: evidence for co-localization of genetic associations and differential DNA methylation. *Genome Biology.* 2016;17:176.
62. Mulder RH, Walton E, Neumann A, et al. Epigenomics of being bullied: changes in DNA methylation following bullying exposure. *Epigenetics*. 2020;15(6-7):750-764.
63. Roberts S, Suderman M, Zammit S, et al. Longitudinal investigation of DNA methylation changes preceding adolescent psychotic experiences. *Translational Psychiatry.* 2019;9(1):1-12.
64. Phipson B, Maksimovic J, Oshlack A. missMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. *Bioinformatics.* 2016;32(2):286-288.
65. Edgar RD, Jones MJ, Meaney MJ, Turecki G, Kobor MS. BECon: a tool for interpreting DNA methylation findings from blood in the context of brain. *Translational Psychiatry.* 2017;7(8):e1187.
66. Joubert BR, Felix JF, Yousefi P, et al. DNA methylation in newborns and maternal smoking in pregnancy: Genome-wide consortium meta-analysis. *American Journal of Human Genetics.* 2016;98(4):680-696.
67. Rzehak P, Saffery R, Reischl E, et al. 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.* 2016;11(5).
68. Colquhoun D. An investigation of the false discovery rate and the misinterpretation of p-values. *Royal Society Open Science*. 2014;1(3):140216.
69. Sullivan PF. Spurious Genetic Associations. *Biological Psychiatry.* 2007;61(10):1121-1126.
70. Czamara D, Eraslan G, Page CM, et al. Integrated analysis of environmental and genetic influences on cord blood DNA methylation in new-borns. *Nature Communications.* 2019;10(1):1-18.
71. Teh AL, Pan H, Chen L, et al. The effect of genotype and in utero environment on interindividual variation in neonate DNA methylomes. *Genome Research.* 2014;24(7):1064-1074.
72. Shi L. Dock protein family in brain development and neurological disease. *Communicative & Integrative Biology.* 2013;6(6):e26839.
73. Agapite J, Albou L-P, Aleksander S, et al. Alliance of Genome Resources Portal: unified model organism research platform. *Nucleic Acids Res.* 2020;48(D1):D650-D658. doi:10.1093/nar/gkz813
74. Price AL, Weale ME, Patterson N, et al. Long-range LD can confound genome scans in admixed populations. *American Journal of Human Genetics*. 2008;83(1):132-135.
75. Shatz CJ. MHC class I: An unexpected role in neuronal plasticity. *Neuron.* 2009;64(1):40-45.
76. Sobue A, Ito N, Nagai T, et al. Astroglial major histocompatibility complex class I following immune activation leads to behavioral and neuropathological changes. *Glia.* 2018;66(5):1034-1052.
77. Murphy TM, Crawford B, Dempster EL, et al. Methylomic profiling of cortex samples from completed suicide cases implicates a role for PSORS1C3 in major depression and suicide. *Translational Psychiatry*. 2017;7(1):e989.
78. Alkelai A, Lupoli S, Greenbaum L, et al. DOCK4 and CEACAM21 as novel schizophrenia candidate genes in the Jewish population. *International Journal of Neuropsychopharmacology.* 2012;15(4):459-469.

79. Lee, Kim J-H, Song GG. Pathway analysis of a genome-wide association study in schizophrenia. *Gene.* 2013;525(1):107-115.
80. Zhang Y, You X, Li S, et al. Peripheral blood leukocyte RNA-Seq identifies a set of genes related to abnormal psychomotor behavior characteristics in patients with schizophrenia. *Medical Science Monitor: International Medical Journal of Experimental Clinical Research.* 2020;26:e922426-1-e922426-31.
81. Glad CAM, Andersson-Assarsson JC, Berglund P, Bergthorsdottir R, Ragnarsson O, Johannsson G. Reduced DNA methylation and psychopathology following endogenous hypercortisolism – a genome-wide study. *Scientific Reports.* 2017;7:44445.
82. Liang, Wang X, Zou M, et al. Family-based association study of ZNF533, DOCK4 and IMMP2L gene polymorphisms linked to autism in a northeastern Chinese Han population. *Journal of Zhejiang University Science B.* 2014;15(3):264-271.
83. Maestrini E, Pagnamenta AT, Lamb JA, et al. High-density SNP association study and copy number variation analysis of the AUTS1 and AUTS5 loci implicate the IMMP2L–DOCK4 gene region in autism susceptibility. *Molecular Psychiatry.* 2010;15(9):954-968. doi:10.1038/mp.2009.34
84. Lee, Chang D-E, Yeom M, et al. Gene expression profiling in hypothalamus of immobilization-stressed mouse using cDNA microarray. *Molecular Brain Research.* 2005;135(1):293-300.
85. Lisowski P, Juszczak GR, Goscik J, Wieczorek M, Zwierzchowski L, Swiergiel AH. Effect of chronic mild stress on hippocampal transcriptome in mice selected for high and low stress-induced analgesia and displaying different emotional behaviors. *European Neuropsychopharmacology*. 2011;21(1):45-62.
86. Papale LA, Madrid A, Li S, Alisch RS. Early-life stress links 5-hydroxymethylcytosine to anxiety-related behaviors. *Epigenetics.* 2017;12(4):264-276.
87. Noro F, Gianfagna F, Gialluisi A, et al. ZBTB12 DNA methylation is associated with coagulation- and inflammation-related blood cell parameters: findings from the Moli-family cohort. *Clinical Epigenetics.* 2019;11(1):74.
88. White T, Muetzel RL, El Marroun H, et al. Paediatric population neuroimaging and the Generation R Study: the second wave. *European Journal of Epidemiology.* 2018;33(1):99-125.
89. Illumina. Infinium MethylationEPIC Kit | Methylation profiling array for EWAS. Published 2020. Accessed July 6, 2020. https://www.illumina.com/products/by-type/microarray-kits/infinium-methylation-epic.html

Supplemental Information on the Methods

Participants

Generation R was designed to shed light into environmental, genetic, and other pathways involved in (ab)normal development. For the purposes of this study, children within Generation R with data on maternal sensitivity (at three and/or four years) and DNA methylation (at six years) were selected. Maternal sensitivity assessments at ages three (*N* = 1247) and four (*N* = 752) years were both considered. This was done in line with previous literature[1], due to the stability of the maternal sensitivity scores between age 3 and 4 years[2], the temporality of these assessments, which both precede DNA methylation at age 6, and to maximize our sample size. Amongst children with maternal sensitivity data at either time point, 235 also had DNA methylation information at age six (i.e. the closest prospective DNA methylation assessment). Of note, this sample included 5 sibling-pairs and was the sample used for the main analyses. To ensure genetic relatedness did not impact results, one sibling per pair was later excluded in a sensitivity analysis. The excluded sibling presented the least covariate data.

In follow-up analyses, where additional adjustments for covariates were made, the main sample was further restricted to complete cases. For adjustments for maternal education and psychopathology, 223 children had data. For adjustments for DNA methylation levels at birth, 226 children had data.

Covariates

The main model in the epigenome-wide association study (EWAS) was adjusted for cell types, batch effects, sex, gestational age at birth, and maternal smoking during pregnancy. Further information on these variables is shown below.

Cell-type and batch effects

Cell-type adjustments were performed, for analyses with DNA methylation at age six, for the following cell types: CD4+ T-lymphocytes, CD8+ T-lymphocytes, B-lymphocytes, monocytes, natural killer cells. Of note, granulocyte cells were excluded due to multicollinearity. The sample plate was used as a measure of batch effects. This variable presented 17 levels.

Sex and gestational age at birth

Sex and gestational age were measured at child-birth. Sex was coded binarily into males and females. Gestational age at birth was measured continuously.

Maternal smoking during pregnancy

We analyzed maternal smoking during pregnancy as a three-level variable: (i) did not smoke during pregnancy, (ii) smoked until pregnancy was known, (iii) smoked throughout pregnancy. This was based on previous work from Joubert et al.[3] showing that sustained smoking throughout

pregnancy has the strongest associations with offspring DNA methylation, with any smoking in pregnancy also showing significant associations, although not as strong. To ensure such variable was not subject to important bias, we additionally examined whether any of our hits (i.e. significant sites or sites within significant regions) overlapped with CpGs related to smoking, based on previous literature. Given the sample at hand, we used the Pregnancy and Childhood Epigenetics (PACE) consortium prenatal smoking exposure reference[3]. In this publication, 6,074 genome-wide significant CpGs were identified in association with maternal smoking during pregnancy in cord blood. Additionally, since tissue- and age-specific effects might be present, we considered another EWAS of smoking carried out in childhood (age 5.5) in whole blood[4], which identified five genome-wide significant probes.

Maternal education and maternal psychopathology

In follow-up analyses, we additionally adjusted for maternal education and maternal psychopathology. Maternal education was coded into low, medium, and high, respectively denoting primary, secondary, and tertiary education levels. Maternal psychopathology (postnatal: child age six months) was measured according to the Beck Symptoms Inventory (BSI) which presents information on the total maternal psychopathology symptoms.

DNA methylation at birth

In another follow-up analysis, additional adjustments for DNA methylation levels at birth were performed. This was done for top hits only. Covariates which are key to appropriately measure DNA methylation levels were also included: batch effects (measured by sample plate) and cell types (CD4+ T-lymphocytes, CD8+ T-lymphocytes, B-lymphocytes, monocytes, natural killer cells, and nucleated red blood cells – a cell type present only in cord blood).

Statistical Analyses

Regional-level EWAS

The dmrff approach, based on simulations, performs better compared to other regional methods in terms of false positive control, statistical power, and replicability across datasets[5]. Of note, the probe- and regional-level EWASs were rerun after one sibling per sibling-pair was excluded.

Candidate gene look-up

For the candidate gene look-up, we selected genes based on previous literature. We searched the PubMed and Google Scholar engines by using a combination of the following terms: “maternal care” or “maternal sensitivity” with “DNA methylation”. Only studies in humans were considered. Review articles were excluded. Both epigenome-wide association studies (probe- and regional-level EWASs) as well as candidate gene studies were considered, yet, no EWAS had been performed to date on normative maternal care. Of the identified candidate

gene studies, only those with statistically significant results were included, for a total of four publications[6–9]. Genes significantly related with maternal care/sensitivity included *NR3C1*, *BDNF*, *SLC6A4*, *OXTR*, and *11B-HSD2*. Due to methylomic values not being available in our sample for 11B-HSD2, such gene was excluded. Overall, four genes were selected based on previous literature.

Gene Ontology: In-House method

In this method, genes in the test list were tested in relation to pathway membership, with a logistic regression approach. We controlled for the number of probes annotated to each gene in the test list. The Gene Ontology website was utilized to obtain pathways. Genes annotated to parent terms were used too. A gene list was formed based on the probes associated with maternal sensitivity at a *p*-value threshold < 0.001, based on the probe-level EWAS. The Illumina UCSC gene annotation permitted the annotation of probes to genes. Genes were considered if they were included in, at minimum, one gene ontology pathway and presented at least one annotated probe. Pathways were considered if including from 10 to 2000 genes. Once this method was used for all pathways, the significant ones with overlapping genes were retested. Associations were retested in all significant pathways, after adjusting for the most significant term. In case the associations at such pathways were no longer significant, the most significant pathway was considered as explaining the relationship. In such situation, pathways were grouped together. This process was repeated, with the next most significant pathway being adjusted for, till all pathways were considered as the most significant one or were identified as pertaining to a more significant pathway. A minimum of two genes was necessary for GO terms to be interpreted.

Follow-up analyses

Firstly, the influence of genetic factors on DNA methylation was examined, based on an openly-accessible mQTL database. The database mQTL information was based on the results from the Accessible Resource for Integrative Epigenomics Studies (ARIES). The ARIES mQTL database includes data on the single nucleotide polymorphisms (SNPs) significantly affecting DNA methylation levels in cis or trans (p < 1E-14, 1Mb window), at several lifespan stages, based on the Infinium 450K array. Here, we selected information for children, based on the results from a genome-wide complex trait conditional analysis.

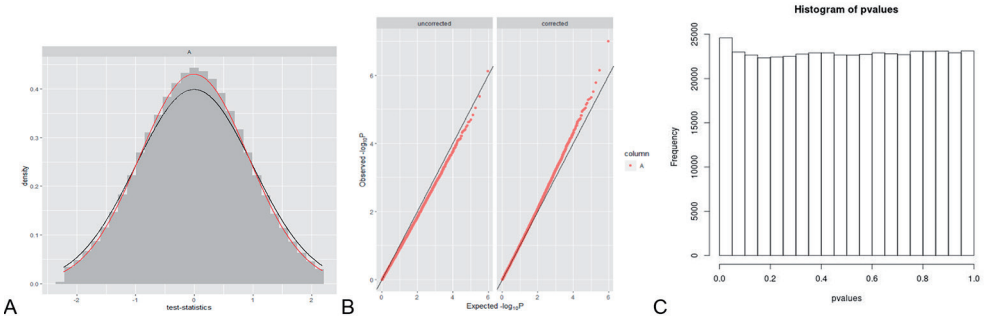
Secondly, we adjusted for an additional set of parental confounders, maternal education and psychopathology, in a subsample of children with such information (N = 223). This was done on top hits only. To ensure estimate changes resulted from adjustments as opposed to the restriction to the subsample, firstly, we ran a multiple linear regression on our top hits, within the subsample of children with data on such confounders (Model A, N = 223). This model was still unadjusted for maternal education and psychopathology. Subsequently, adjustments for maternal education and psychopathology were performed on top hits, within

the subsample (Model B, *N* = 223, adjusted model), with a multiple linear regression. The newly-obtained site summary statistics for both models were then inputted in DMRff, where the function dmrff.stats enables the recalculation of the statistics per DNA methylation region. The percent estimate change was then calculated ((estimate after adjustments – estimate before adjustments) / estimate before adjustments * 100).

The same procedure was employed for adjustments for DNA methylation levels at birth. Therefore, a multiple linear regression where, for each site, its own DNA methylation levels at baseline were used as covariates, was tested in association with maternal sensitivity: site DNA methylation at six ~ maternal sensitivity + main set of covariates + site DNA methylation at birth (Model B, *N* = 226). This was compared to a restricted unadjusted model (Model A, *N* = 226). Site statistics were inputted in dmrff.stats to obtain regional-level statistics.

Lastly, the BECon online tool used here includes information on tissue concordance between DNA methylation in blood and Brodmann Areas (BA) 7, 10 and 20, based on brain postmortem samples from 16 subjects.

Supplemental Figure



Supplemental Figure 1 A-C. BACON-corrected probe-level EWAS of the association between maternal sensitivity and DNA methylation

Note. Panel A shows the test statistic histogram. This indicates minimal bias, as shown by the proximity of the black and red lines. Panel B depicts uncorrected and BACON-corrected QQ-plots. The uncorrected QQ-plot is from the main EWAS and shows no significant associations. The BACON-corrected QQ-plot suggests that values were originally deflated in the EWAS and that, after corrections, associations are present. Panel C is a histogram of *p*-values after BACON-correction, showing that there is a greater proportion of sites with low *p*-values compared to the proportion of sites with higher *p*-values, further suggesting that associations of maternal sensitivity with DNA methylation are present.

Supplemental Tables

Supplemental Table 1. Sample characteristics

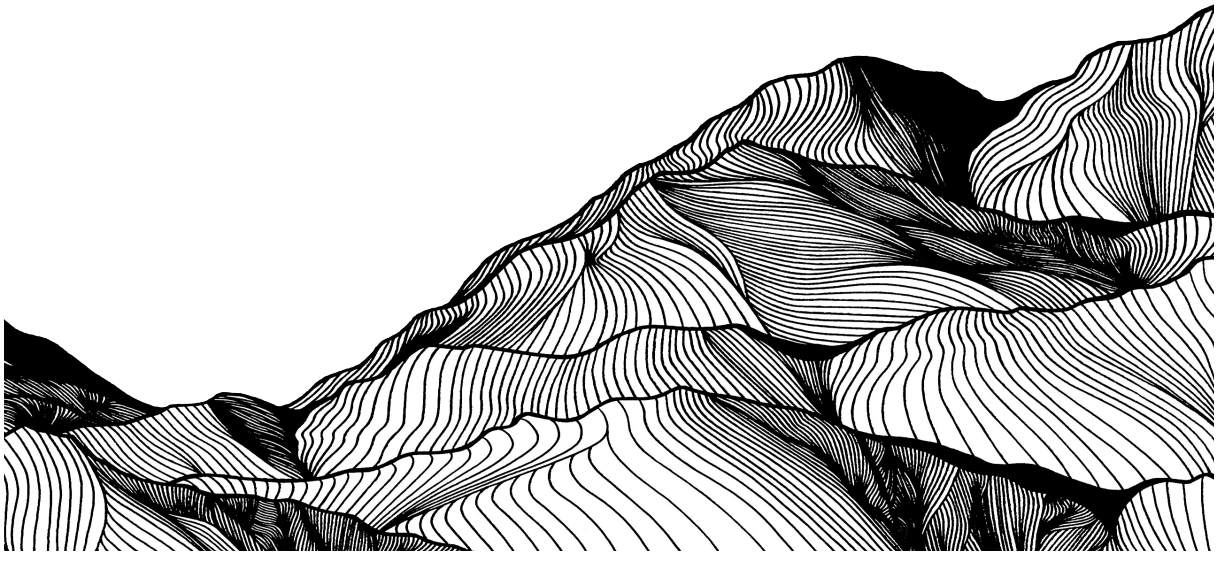
Child and maternal characteristics	Percentage/Mean
Child Characteristics	
Sex–females	49%
Child Ethnicity–Dutch	100%
Maternal Characteristics	
Observed maternal sensitivity	0.05
Maternal Education*	
Primary education	8%
Secondary education	21%
Higher education	71%
Gestational age at birth	40.31
Maternal prenatal smoking	
Never smoked	75%
Quit when pregnancy known	11%
Continued during pregnancy	14%
Maternal psychopathology*	0.15

Note. % = percentage
*Information was available only for a subsample (n = 223)

Supplemental Table 2-11 will be available online in the published article–https://www.cambridge.org/core/journals/psychological-medicine

Supplemental References

1. Kok R, Thijssen S, Bakermans-Kranenburg MJ, et al. Normal variation in early parental sensitivity predicts child structural brain development. *Journal of American Academy of Child & Adolescent Psychiatry*. 2015;54(10):824-831.e1.
2. Kok R, Linting M, Bakermans-Kranenburg MJ, et al. Maternal sensitivity and internalizing problems: Evidence from two studies in early childhood. *Child Psychiatry & Human Development.* 2013;44(6):751-765.
3. Joubert BR, Felix JF, Yousefi P, et al. DNA methylation in newborns and maternal smoking in pregnancy: Genome-wide consortium meta-analysis. *American Journal of Human Genetics.* 2016;98(4):680-696.
4. Rzehak P, Saffery R, Reischl E, et al. 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.* 2016;11(5).
5. Suderman M, Staley JR, French R, Arathimos R, Simpkin A, Tilling K. Dmrff: Identifying differentially methylated regions efficiently with power and control. *Bioinformatics*; 2018.
6. Bosmans G, Young JF, Hankin BL. NR3C1 methylation as a moderator of the effects of maternal support and stress on insecure attachment development. *Developmental Psychology.* 2018;54(1):29-38. doi:10.1037/dev0000422
7. Conradt E, Hawes K, Guerin D, et al. The contributions of maternal sensitivity and maternal depressive symptoms to epigenetic processes and neuroendocrine functioning. *Child Development.* 2016;87(1):73-85.
8. Provenzi L, Fumagalli M, Giorda R, et al. Maternal sensitivity buffers the association between SLC6A4 methylation and socio-emotional stress Response in 3-month-old full term, but not very preterm infants. *Frontiers in Psychiatry.* 2017;8.
9. Unternaehrer E, Meyer AH, Burkhardt SCA, et al. Childhood maternal care is associated with DNA methylation of the genes for brain-derived neurotrophic factor (BDNF) and oxytocin receptor (OXTR) in peripheral blood cells in adult men and women. *Stress.* 2015;18(4):451-461.

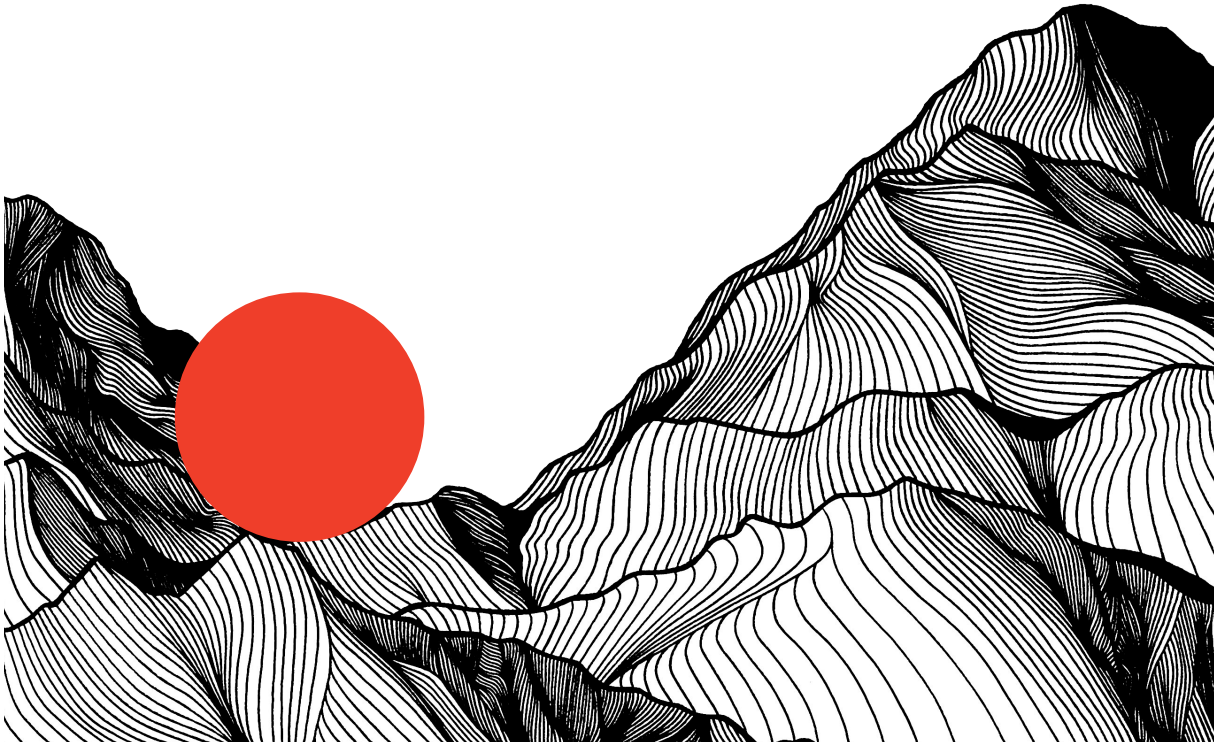


CHAPTER V

Genome-wide DNA methylation patterns associated with sleep and mental health in children: A population-based study

M. Elisabeth Koopman-Verhoeff, Rosa H. Mulder, Jared M. Saletin, Irwin Reiss, Gijsbertus T.J. van der Horst, Janine F. Felix, Mary A. Carskadon, Henning Tiemeier, Charlotte A. M. Cecil

Journal of Child Psychology and Psychiatry (2020).



Abstract

Objective: DNA methylation has been implicated in the biology of sleep. Yet, how DNA methylation patterns across the genome relate to different sleep outcomes, and whether these associations overlap with mental health is currently unknown. Here, we investigated associations of DNA methylation with sleep and mental health in a pediatric population.

Method: This cross-sectional study included 465 10-year-old children (51.3% female) from the *Generation R Study*. Genome-wide DNA methylation levels were measured using the Illumina450K array (peripheral blood). Sleep problems were assessed from self-report, and mental health outcomes from maternal questionnaires. Wrist actigraphy was used in 188 11-year-old children to calculate sleep duration and midpoint sleep. Weighted gene co-expression network analysis was used to identify highly co-methylated DNA methylation 'modules', which were tested for associations with sleep and mental health outcomes.

Results: We identified 64 DNA methylation modules, one of which associated with sleep duration after covariate and multiple-testing adjustment. This module included CpG sites spanning 9 genes on chromosome 17, including *MAPT*– a key regulator of Tau proteins in the brain involved in neuronal function – as well as genes previously implicated in sleep duration. Follow-up analyses suggested that DNA methylation variation in this region is under considerable genetic control and shows strong blood-brain concordance. DNA methylation modules associated with sleep did not overlap with those associated with mental health.

Conclusion: We identified one DNA methylation region associated with sleep duration, including genes previously reported by recent GWAS studies. Further research is warranted to examine the functional role of this region and its longitudinal association with sleep.

Introduction

Sleep is increasingly recognized as an important factor in child mental health. Sleep disturbances, such as short sleep and shifted circadian rhythm, often develop in late childhood and have been implicated in mental health problems[1, 2]. While poor sleep can exacerbate mental health difficulties[3, 4], mental health problems can also precede and worsen sleep[5]. Thus, the association between sleep and mental health is complex and likely bidirectional[1]. The mechanisms underlying this association, however, remain unknown.

Complex traits, including sleep, result from the interplay of genetic and environmental influences[6]. How these factors jointly influence normative sleep, or the development of sleep problems, is currently unclear. Epigenetic processes such as DNA methylation have been proposed as a mechanism of interest[7, 8]. Differential DNA methylation has been linked to a broad range of developmental outcomes, including sleep, as well as mental and physical health problems[9, 10]. Most research on this topic emerges from animal models[11], with only a handful of studies examining DNA methylation and sleep in humans. These have typically relied on small samples of adults with dysregulated sleep (e.g. shift workers) and utilized a candidate gene approach focusing primarily on 'clock' genes: genes driving circadian rhythms in metabolism, physiology and behaviour[11-13]. In contrast, we are aware of only two epigenetic studies during development, both of which examined adolescence. One reported an association in 18-19 year-olds between sleep duration and DNA methylation of *DOCK1*, a gene influenced by circadian rhythmicity[14]. The second found that higher DNA methylation in metabolic genes *PPARA* and *HSD11B2* was associated with shorter sleep, specifically in girls[15] 0.

Despite these promising preliminary findings, existing research has been limited in four key ways, namely (i) the use of small samples of adults or older adolescents; (ii) a focus on a candidate gene approach; (iii) the lack of multi-modal assessments of sleep, making it unclear whether associations between sleep and DNA methylation differ between self-report and objective measures (e.g., actigraphy); and (iv) despite evidence showing that mental health is related to DNA methylation alterations[9] and sleep[1], no study has examined these factors jointly.

To address these gaps, we examined the relationship between genome-wide DNA methylation, sleep and mental health in a general population sample of 10-year old children – an important period for development of sleep and mental health problems alike. The aims of our study were two-fold: first, to characterize cross-sectional associations of DNA methylation with reported (i.e. dyssomnia symptoms) and actigraphy-assessed (i.e. sleep duration and midpoint) sleep using both a genome-wide approach and an targeted approach focusing on well-characterized clock genes to maximize comparability with existing studies; and second, to investigate whether sleep-associated DNA methylation patterns are also associated with common mental health problems. Findings were tested for consistency in a small independent sample.

Methods

Participants

This cross-sectional study included 10-year-old children of European ancestry (51.3% female) from the *Generation R Study*, a prospective population-based cohort from foetal life onward. Pregnant women (expected delivery date April 2002–January 2006) living in Rotterdam, the Netherlands, were invited to participate[16]. The current analyses are based on children who had DNA methylation data and subjectively assessed sleep (n = 410). Of these, 188 also had actigraphy data. Written informed consent was obtained for all participants. The Medical Ethical Committee of the Erasmus MC, University Medical Center Rotterdam approved the study.

Measures

DNA methylation

Five-hundred nanograms of DNA were extracted from peripheral blood at age 10 and underwent bisulfite conversion with the EZ-96 DNA Methylation kit (Shallow) (Zymo Research Corporation, Irvine, USA). Samples were plated onto 96-well plates in no specific order. DNA methylation was analyzed with the Illumina Infinium Human Methylation 450K BeadChip (Illumina Inc., San Diego, USA). Quality control of samples was performed using standardized criteria using the CPACOR workflow[17]. Probes with a detection p-value above background ≥ 1E-16 were set to missing per array. Arrays with observed technical problems including failed bisulfite conversion, hybridization or extension, and arrays with a mismatch between child sex and sex determined by the chr X and Y probe intensities were removed. Nonautosomal probes were excluded. Additionally, only arrays with a call rate > 95% per sample were processed further. Methylation beta values outside a range of the 25th percentile minus 3*interquartile range to the 75th percentile plus 3*interquartile range were set to missing. The final dataset contained 425 samples, analyzing 458,563 CpG sites. For our targeted approach, we examined DNA methylation levels of CpG sites that were annotated to well-characterized clock-related genes (939 CpG sites across 39 genes[18]) (Supplemental Table 1). For each CpG site, Beta values represent the ratio of methylated signal relative to the sum of (the methylated and unmethylated signals plus 100).

Child-reported dyssomnia symptoms

At age 10 years, children completed six questions of the Sleep Disturbance Scale for Children[19] about perceived sleep, for example, “Do you find it difficult to fall asleep?”; “If you wake up at night, do you find it difficult to fall asleep again?”; “Do you feel rested when you wake in the morning?” (previously described[20]). The questions were rephrased for our paediatric population. Responses were scored on a three-point Likert scale (“No”, “Sometimes” or “Yes”; α = 0.64). Items were summed; higher scores indicate greater sleep problems.

Actigraphy-estimated sleep

Sleep patterns were estimated with wrist tri-axial actigraphy (GENEActiv) on the non-dominant wrist for 5 consecutive school nights in 188 children at age 11 (i.e. after DNA sampling)[20, 21]. The Geneactiv accelerometers were set a frequency of 50 Hz. The binary files were processed with the R-package GGIR[22]. Accompanying sleep diaries were collected and used to guide actigraphy analyses. Sleep duration was estimated as the total time scored sleep between falling asleep and final waking. Sleep midpoint was estimated as the halfway point between sleep onset and final waking. Sleep duration and midpoint were averaged across the week, excluding weekends to best approximate typical school-day sleep patterns and to minimize the influence of atypical weekend events.

Child psychopathology

The Child Behavior Checklist 6-18 (CBCL/6-18) was assessed using maternal-reports at age 10 to derive broadband Internalizing and Externalizing problem-scales[23]. The CBCL/6-18 is widely used internationally and has been found to be generalizable across 23 societies, including the Netherlands[24]. Mothers rated various emotional and behavioural problems of the child in the previous six months on a three-point scale (0 = not true, 1 = somewhat true, 2 = very true).

Covariates

Sex of the child was obtained from medical records and maternal characteristics by questionnaires. Maternal education was defined by the highest attained educational level and classified into two categories (higher vocational education and university: yes or no). Correction for sample plate and cell type proportions was also applied. We used the Houseman method[25] to estimate relative proportions of six white blood cell subtypes (CD4+ T-lymphocytes, CD8+ T-lymphocytes, NK (natural killer) cells, B-lymphocytes, monocytes and granulocytes), based on a standard reference population[26].

Statistical analysis

We had nearly complete cases, with four participants missing data on maternal education (defined as highest educational level). These participants were excluded from the analysis. Statistical analyses were performed in R[27], following three steps:

Step 1. Associations between DNA methylation and sleep

We applied weighted gene co-expression network analysis[28] (WGCNA) – a system-level data reduction approach – to reduce the dimensionality of the data and identify clusters (so called 'modules') of highly co-methylated DNA methylation sites across genome. As such, rather than focusing on individual sites or genes, WGCNA enables utilization of correlation patterns between sites to identify wider DNA methylation networks, which may also be functionally related[29]. Block-wise network construction was run using default settings (power threshold of 6; minimal module size of 30 sites; merge cut height of 0.25). Each derived module was colored

by size automatically and summarized by a 'module eigengene' (ME) value, the first principal component of the given module. We numbered the derived modules by significance with outcome for simplicity. CpG sites that do not co-methylate were assigned to an 'unclassified' module. WGCNA analyses were performed twice: first based on the entire genome-wide data (i.e. hypothesis free; *N* = 458,563 sites), and second based on the subset of clock genes (i.e., targeted approach, *n* = 939 CpG sites).

Next, we tested bivariate correlations between the co-methylated modules and the three sleep outcomes (i.e. child-reported dyssomnia symptoms, actigraphy-estimated sleep duration and midpoint sleep). We selected modules that were associated with sleep outcomes after Bonferroni correction for multiple testing (0.05/n modules*3 sleep measures)[30]. These modules were further examined using linear regression models controlling for batch, cell-types, child sex and age, and maternal education.

Modules that were significantly associated with sleep were examined further using publicly available resources to characterize (i) their genomic location; (ii) potential genetic influences, by checking whether the CpG sites included in the modules are known to be polymorphic[31] (i.e. overlapping with single nucleotide polymorphisms(SNPs)) linked to methylation quantitative trait loci[32] (mQTLs; i.e. SNPs that associate with DNA methylation levels, either in cis or in trans; http://www.mqtldb.org/; GCTA set) or heritable, based on twin data[33] (i.e. explained by additive genetic influences as opposed to shared and non-shared environmental influences); and (iii) blood-brain concordance, based on postmortem data from 122 individuals with DNA methylation from whole blood and four brain regions (the prefrontal cortex, entorhinal cortex, superior temporal gyrus, and cerebellum[34] (https://epigenetics.essex.ac.uk/bloodbrain/).

Step 2: Testing the overlap of associations with mental health

Bivariate correlations between the co-methylated modules, sleep and mental health measures were examined to establish whether associations of DNA methylation and sleep are co-localized on the genome with associations of DNA methylation and internalizing and externalizing problems.

Step 3. Generalizability in independent sample

Associations identified in Steps 1 and 2 were estimated in an independent sample of 63 older adolescents (14.5±0.3 years, 54% girls) of the Generation R Study to judge generalizability of results, with information on DNA methylation available at 10 years and actigraphy-assessed sleep at 14 years (i.e. prospective association). The children in this sample were recruited for a second actigraphy study at a later age than the first study described above due to logistic reasons (no repeated measurements).

Results

Characteristics of the study sample are presented in Table 1. For correlations across sleep and mental health variables, see Supplemental Table 2. The average midpoint sleep was 2:49 (*SD* = 35min) and the mean sleep duration was 7:36 (*SD* = 40min).

1. Are DNA methylation patterns associated with sleep outcomes in children?

Genome-wide analyses

We identified 64 co-methylated modules, containing between 30 and 65,804 CpG sites (Supplemental Table 3). The majority of sites were unclassified (*n* = 261,374), suggesting they did not correlate strongly enough to form modules. Two modules correlated with sleep after Bonferroni correction for multiple testing (0.05/64 modules * 3 outcomes = 0.00026042) – both of which associated with sleep duration (module1 *r* = -0.18, *p* = 0.00006, module2 *r* = -0.18, *p* = 0.0001) (Table 2), but not with sleep midpoint or dyssomnia symptoms. Only the association between module1 and sleep duration remained significant in a regression model adjusting for covariates (ß = -0.22, 95% CI = -0.37–-0.07, *p* = 0.004). As a sensitivity analysis we replaced the missing values (*n* = 4) on maternal highest educational level attained by

Table 1. Sample characteristics

	Reported dyssomnia symptoms (*N* = 410)	**Actigraphic sleep (*N* = 188)**
Demographics		
Sex, female, %	234 (50.3%)	93 (49.5%)
Age (years)	9.8±0.3	11.7±0.1
Maternal education, %		
Low & Intermediate	152 (32.7%)	64 (34.0%)
High	308 (66.2%)	121 (64.4%)
Dyssomnia symptoms, self-reported (score;range)	10.80 (8.00-18.00)	10.86 (6.00-17.00)
Sleep duration, actigraphy (hours:minutes)	-	7:35±0:44
Midpoint sleep, actigraphy, time (hours:minutes)	-	02:48±0:35
Internalizing problems, mother-reported, mean(*SD*)	4.16 (4.38)	4.03 (4.28)
Externalizing problems, mother-reported, mean(*SD*)	3.41 (4.25)	3.16 (3.82)

Table 2. Associations between DNA methylation modules and actigraphy-derived sleep duration in children (*N* = 188)

	A. Correlations of the WGCNA modules with sleep duration				**B. Standardized regression coefficients**		
Module	***r***	***p*-value**	***N* cpgs**	***N* genes**	**β**	**CI**	***p*-value**
Module1	**-0.18**	**0.00006**	**32**	**9**	**-0.22**	**-0.37– -0.07**	**0.004**
Module2	-0.18	0.0001	5845	3462	-0.14	-0.54–36	0.07

maternal highest educational level, yielding highly consistent results. Additionally, as time of blood sampling corrected for the time of habitual awakening could be of influence, we re-ran analyses adjusting for these variables, and found that results remained highly consistent (ß = -0.19, 95% CI = -0.34–-0.05, *p* = 0.008). Lastly, as cell proportions are estimated, rather than derived from actual cell counts, we re-ran analyses without cell type correction to test stability of associations, and found that results were highly consistent (ß = -0.22, 95% CI = -0.36–-0.07, *p* = 0.004).

Targeted circadian clock CpG site analyses.

The targeted WGCNA approach containing exclusively clock-related genes identified 5 modules (ranging from 19-300 CpG sites over 10-39 genes), each including CpG sites spanning multiple genes, as opposed to clustering by gene. The majority of the CpG sites were unclassified (*n* = 540). No modules were associated with sleep outcomes after multiple testing correction.

Functional characterization of module1 (Table S4).

Annotation to genes and genomic region

Module1 contained 32 sites spanning 9 genes. The largest number of sites (n = 6) were annotated to the Microtubule-Associated Protein Tau (*MAPT*) gene. The CpCs of module1

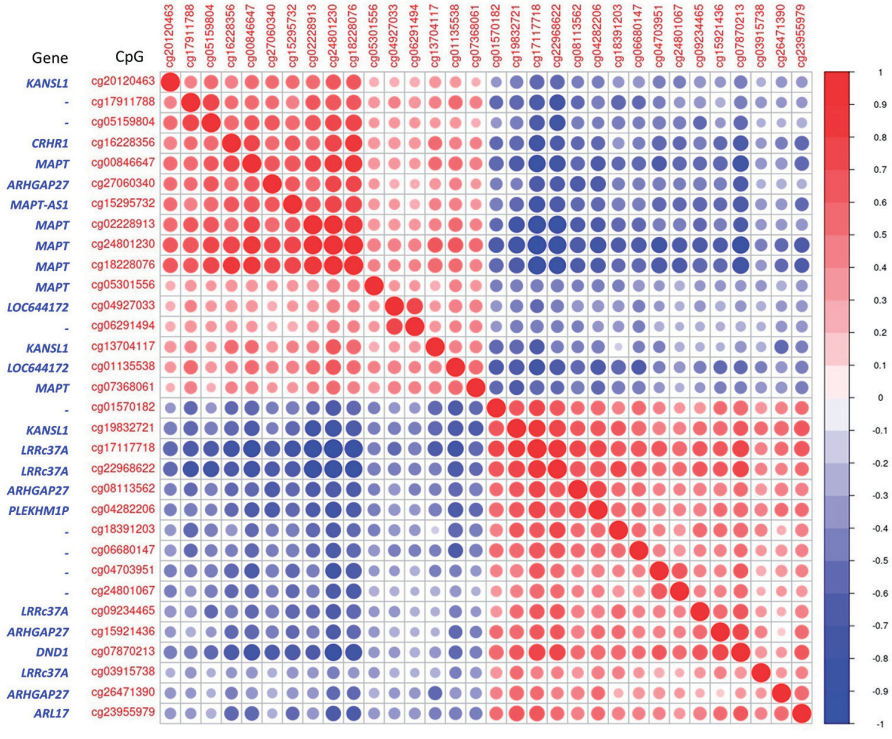


Figure 1. Intercorrelations between CpG sites in module1

were highly correlated with each other (Figure 1), as well as with sleep duration and were all located in the chromosome 17q21.31 region, chr17:43502999-62843696, with the exception of one CpG site on chromosome 5.

Genetic influences

Six of the CpGs included in module1 were previously identified as polymorphic (three of which in *MAPT*), and twelve (37.5%) were found to be associated to mQTLs on chromosome 17, with a total of 71 associations (between 4 and 10 associations per CpG). The CpG site located on chromosome 5 (cg07870213) associated with both mQTLs on chromosome 5 in cis as well as chromosome 17 in trans, all of which were located in the module1 region (chr17:41993881-44852612). Finally, 10 of the 32 CpG sites in module1 had twin heritability estimates available, all of which showed moderate to strong genetic influences (r = 0.34–1.00).

Blood-brain concordance

For all but one of the CpG sites in module1, DNA methylation levels in blood correlated significantly with DNA methylation levels in at least one brain region. The three MAPT CpG sites that associated most strongly with sleep duration showed high blood-brain correlations (Figure S1). Of these, cg24801230 (one of the sites found to be polymorphic) showed an almost perfect correlation (r = 0.99) between blood and brain, with DNA methylation levels across tissues clustering into three alleles (Supplemental Figure 1).

2. Are DNA methylation-sleep associations overlapping with child psychiatric symptoms?

No modules were associated with internalizing and externalizing problems after correction for multiple testing. Generally, we found weak associations between the DNA methylation modules and internalizing (strongest association: r = 0.15, p = 0.001) and externalizing problems (strongest association: r = 0.14, p = 0.002). Associated modules did not overlap with those identified for sleep duration (Supplemental Figure 2).

3. Are results consistent in an independent sample?

The association between module1 and sleep duration was tested in an independent sample of older children, in order to test for consistency across developmental stage. Results from a regression analysis, controlling for covariates, yielded a highly comparable effect size (Discovery: ß = -0.22, 95% CI = -0.37–-0.07, p = 0.004; Generalization sample: ß = -0.23, 95% CI = -0.50–0.04, p = 0.09), although the association was not statistically significant, likely due to the larger confidence intervals resulting from the use of a smaller sample (1/3 of discovery sample).

Discussion

The current study utilized a network-based approach to investigate associations between genome-wide DNA methylation, sleep, and mental health in a pediatric population. We highlight here two key findings. First, we found that DNA methylation patterns associated with sleep duration, but not with other sleep parameters. Specifically, our hypothesis-free analyses identified one DNA methylation module associated with actigraphy-assessed sleep duration. This module (i) contained 32 sites annotated to multiple genes previously linked to sleep duration in GWASes, including *MAPT*; (ii) showed strong evidence of genetic influences based on molecular and twin data; and (iii) showed cross-tissue concordance between blood and brain. In contrast, hypothesis-driven analyses did not reveal associations between DNA methylation in clock genes and sleep parameters. Second, we found that DNA methylation patterns were only weakly associated with mental health outcomes. These associations did not overlap with those identified for sleep outcomes, suggesting co-methylation modules associated with sleep and mental health are largely independent.

Self-reported and actigraphic sleep assess distinct sleep domains[1, 35], as reflected in the weak correlations between these metrics found in the present study. Of note, self-reported measures capture sleep perception and reports may be biased by subject characteristics. Interestingly, we found here that DNA methylation associated with actigraphic sleep duration but not with self-reported dyssomnia. This could be due to the fact that actigraphic sleep shows greater variability in the general population and has less measurement error[36]. Furthermore, we did not find associations between DNA methylation and actigraphic determined midpoint sleep. Nights assessed in our sample have been constrained by school schedules, limiting variability in midpoint. Since circadian preference changes during adolescence[37] future research should study the longitudinal association between DNA methylation, and sleep and circadian rhythm across this age period.

Most epigenetic research on sleep in humans has focused on sleep deprivation[11]. In this study, we show that DNA methylation patterns associate with typical variation in sleep in 10-year old children. Specifically, one DNA methylation module was found to associate with actigraphic sleep duration. This association was generalizable to a smaller, independent sample of Generation R participants at age 14 years. The lack of significance could be due to low power in this smaller sample. The fact that we found a generally comparable effect size supports the robustness of our findings.

The sleep-associated module contained 32 CpG sites spanning a large region on chromosome 17. Based on accessible databases, we found that several of the sites in the module were located directly on SNPs, and over a third were linked to known mQTLs. Intriguingly, the one CpG site in this module on chromosome 5 was associated with multiple mQTLs located within

the chromosome 17 region, supporting a genetically-driven link in DNA methylation patterns between these two chromosomal regions. Genetic influences were further corroborated by twin data showing moderate-to-high heritability estimates for DNA methylation sites in this module. Together, these findings suggest that underlying genetic variation might largely account for observed associations between DNA methylation in this region and sleep duration. This is in line with existing literature indicating that variation in DNA methylation is best explained by genetic influences and gene-environment interactions, as opposed to environmental main effects[38, 39]. Finally, DNA methylation variability in the identified module showed high blood-brain concordance, highlighting that the signals currently found in blood might be useful proxies for DNA methylation status in the brain. Future studies will need to test concordance with other brain areas implicated in sleep duration, e.g. the hypothalamus, and establish whether the degree of correspondence differs across specific cell-types in the brain.

Of the 9 genes annotated to our module, several stood out for their role in brain-related processes and previous links to sleep outcomes based on GWAS data. Specifically, a single-nucleotide polymorphism (SNP) in *MAPT* was recently identified as a top GWAS hit for self-reported sleep duration[40] and SNPs in *MAPK81P1P2* and *KANSL1-AS1* were identified as top hits in a GWAS on accelerometer-based sleep duration[41]. Additionally, a study based on UK Biobank and 23andMe data indicated that variants in *ARHGAP27*, *LRRC37A*, *CRHR1*, *MAPT*, and *KANSL1* associated with various self-reported sleep traits, including sleep duration[42]. These findings further support genetic influences on DNA methylation and sleep duration in this region.

The most strongly associated probe in module1 was annotated to the *MAPT* antisense RNA 1, a non-protein coding RNA gene identified as epigenetic regulator of *MAPT* expression[43], while six sites where annotated to the *MAPT* gene itself. *MAPT* encodes the Tau protein, which is important for neuronal stabilization. Its aberrant aggregation has been frequently linked to Alzheimer's disease and other neurodegenerative diseases[44] as well as neurodevelopmental disorders[45]. A recent study suggested the involvement of Tau proteins and sleep in the pathogenesis of neurodegenerative diseases, though this process is not yet fully understood[46]. Another gene annotated to module1 was *CRHR1* (corticotropin-releasing hormone receptor 1), a pivotal player in hypothalamic-pituitary-adrenal axis functioning[47] as well as sleep[48]. Our study adds to this growing body of evidence by showing for the first time that, in childhood, epigenetic variation in *MAPT* and surrounding regions are associated with sleep duration.

The epigenetic patterns associated with sleep in this study did not overlap with those associated with mental health. This may be due to several reasons. First, although the link between sleep and mental health is well-established[1], it is possible that such associations may not be epigenetically-mediated. Second, associations between sleep and mental health tend to be stronger for self-report than objective measures[1]. As such, there might be different

underlying biological correlates driving the associations between mental health and reported sleep and actigraphic derived sleep. For example, cortisol levels, associated with anxiety and depression, have been linked to self-reported sleep quality but not to actigraphy-derived sleep quantity[49]. Third, our population-based cohort may have lacked psychiatric severity to detect shared associations. Future studies are needed to clarify the mechanisms underlying associations between sleep and mental health.

Limitations and future directions

This study has several limitations. First, from our cross-sectional data, we are unable to determine the direction of effect for the association between DNA methylation and sleep regulation, and we cannot exclude that the observed association may result from a common influence (e.g. environmental or genetic modulation). In the future, the use of longitudinal data on DNA methylation and sleep, the application of advanced causal inference methods (e.g. two-step Mendelian randomization), as well as integration with genetic data will mark important steps for furthering our understanding of DNA methylation-sleep associations. Second, the sample was based on participants of European ancestry. Studies including other ethnicities are necessary to investigate the generalizability of our findings. Third, our independent sample was smaller, limiting statistical power. Fourth, our measure of midpoint sleep, derived from actigraphy, is constrained by school schedules. Studying free nights may better describe underlying circadian processes. Fifth, while we assume that focusing on modules as opposed to single sites may help us to identify broader, functionally meaningful DNA methylation networks associated with sleep, (a) this does not preclude that there may be important sleep-associated single CpG sites, which might have been missed by using this approach; and (b) integration with gene expression data will be necessary to establish the extent to which the identified module may play a regulatory role, which we could not do in our study. In addition to the clock genes tested in the current study, it would be interesting to examine associations with CpG sites annotated to genes that have been previously implicated in other sleep parameters, such as sleep duration or chronotype (e.g. by GWAS studies). Sixth, the blood-brain concordance tool we used is based on an elderly population. As such, it is unclear to what extent the identified pattern of concordance extends to the pediatric population, for which there are currently no available tools. Finally, it is unclear whether identified DNA methylation patterns are functionally relevant. The use of experimental models could inform the biological consequences of these associations. Additionally, it is important to see in future studies whether DNA methylation levels at these sites change across development. If there is no change in DNA methylation levels over time, this could indicate that a regulatory process is acting from birth, whereas an epigenetic mark that changes throughout life might indicate that it may be responsive to environmental stimuli.

Conclusion

In summary, the preliminary results of the current study show promising sleep-associated DNA methylation patterns in the pediatric population. Specifically, we identified an association between sleep duration and DNA methylation in the 17q21.31 region, spanning multiple genes previously linked to sleep by GWAS studies, including *MAPT*. These epigenetic patterns did not overlap with those associated with self-reported sleep problems, midpoint sleep or mental health. Future studies are needed to replicate our findings and establish causality. Overall, our findings offer novel insights into epigenetic patterns associated with typical variation in sleep duration in children.

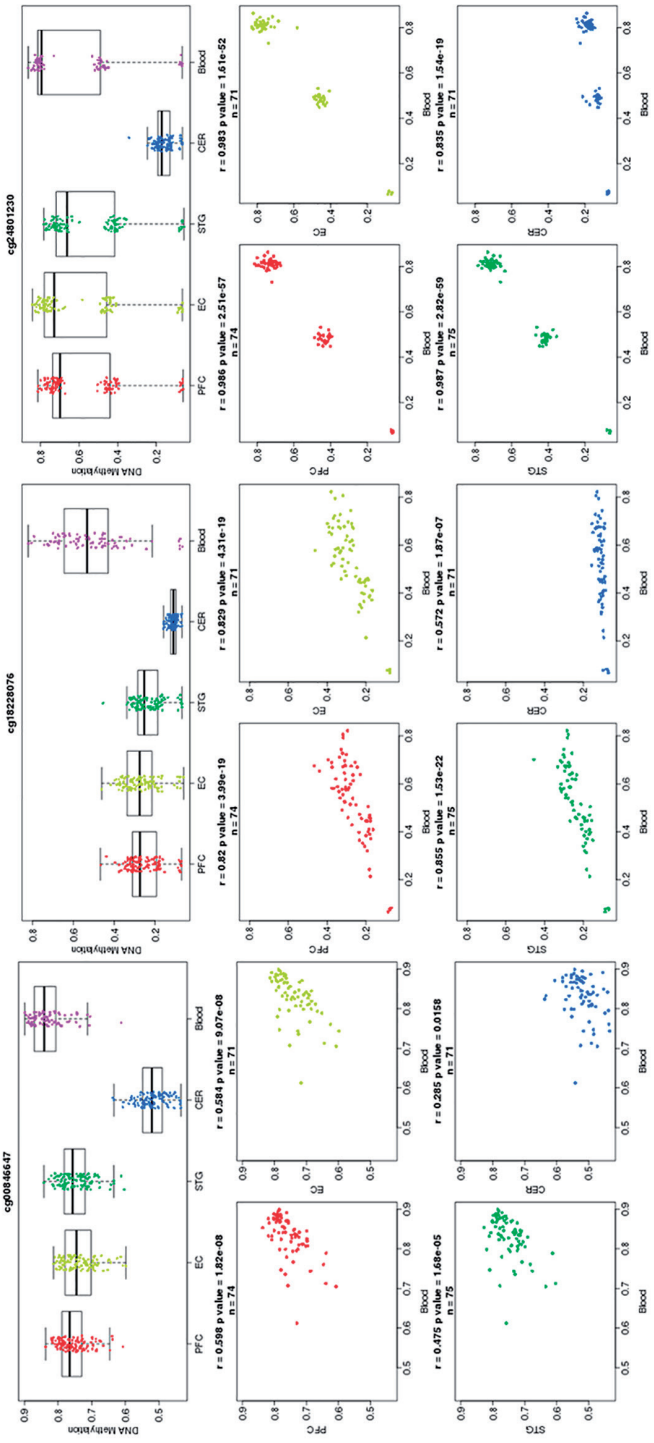
References

1. Gregory AM, Sadeh A. Annual research review: sleep problems in childhood psychiatric disorders–a review of the latest science. *Journal of Child Psychology and Psychiatry.* 2016;57:296-317.
2. Wolfson AR, Carskadon MA. Sleep schedules and daytime functioning in adolescents. *Child Development.* 1998;69:875-87.
3. Lovato N, Gradisar M. A meta-analysis and model of the relationship between sleep and depression in adolescents: recommendations for future research and clinical practice. *Sleep Medicine Reviews.* 2014;18:521-9.
4. Owens JA. The ADHD and sleep conundrum: a review. *Journal of Developmental & Behavioral Pediatrics.* 2005;26:312-22.
5. Verhoeff ME, Blanken LME, Kocevska D, et al. The bidirectional association between sleep problems and autism spectrum disorder: a population-based cohort study. *Molecular Autism.* 2018;9:8.
6. Romens SE, McDonald J, Svaren J, Pollak SD. Associations between early life stress and gene methylation in children. *Child Development.* 2015;86:303-9.
7. Massart R, Freyburger M, Suderman M, et al. The genome-wide landscape of DNA methylation and hydroxymethylation in response to sleep deprivation impacts on synaptic plasticity genes. *Translational Psychiatry.* 2014;4:e347-e.
8. Morales-Lara D, De-la-Peña C, Murillo-Rodríguez E. Dad's snoring may have left molecular scars in your DNA: the emerging role of epigenetics in sleep disorders. *Molecular Neurobiology.* 2018;55:2713-24.
9. Barker ED, Walton E, Cecil CAM. Annual Research Review: DNA methylation as a mediator in the association between risk exposure and child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry.* 2018;59:303-22.
10. Breton CV, Marsit CJ, Faustman E, et al. Small-magnitude effect sizes in epigenetic end points are important in children's environmental health studies: the children's environmental health and disease prevention research center's epigenetics working group. *Environmental Health Perspectives.* 2017;125:511-26.
11. Gaine ME, Chatterjee S, Abel T. Sleep deprivation and the epigenome. *Frontiers in Neural Circuits.* 2018;12:14.
12. Cedernaes J, Osler ME, Voisin S, et al. Acute sleep loss induces tissue-specific epigenetic and transcriptional alterations to circadian clock genes in men. *The Journal of Clinical Endocrinology & Metabolism.* 2015;100:E1255-E61.
13. Lahtinen A, Puttonen S, Vanttola P, et al. A distinctive DNA methylation pattern in insufficient sleep. *Scientific Reports.* 2019;9:1-9.
14. Huang H, Zhu Y, Eliot MN, et al. Combining human epigenetics and sleep studies in Caenorhabditis elegans: a cross-species approach for finding conserved genes regulating sleep. *Journal of Sleep and Sleep Disorders Research.* 2017;40:zsx063.
15. Jansen EC, Dolinoy DC, O'Brien LM, et al. Sleep duration and fragmentation in relation to leukocyte DNA methylation in adolescents. *Sleep.* 2019;42:zsz121.

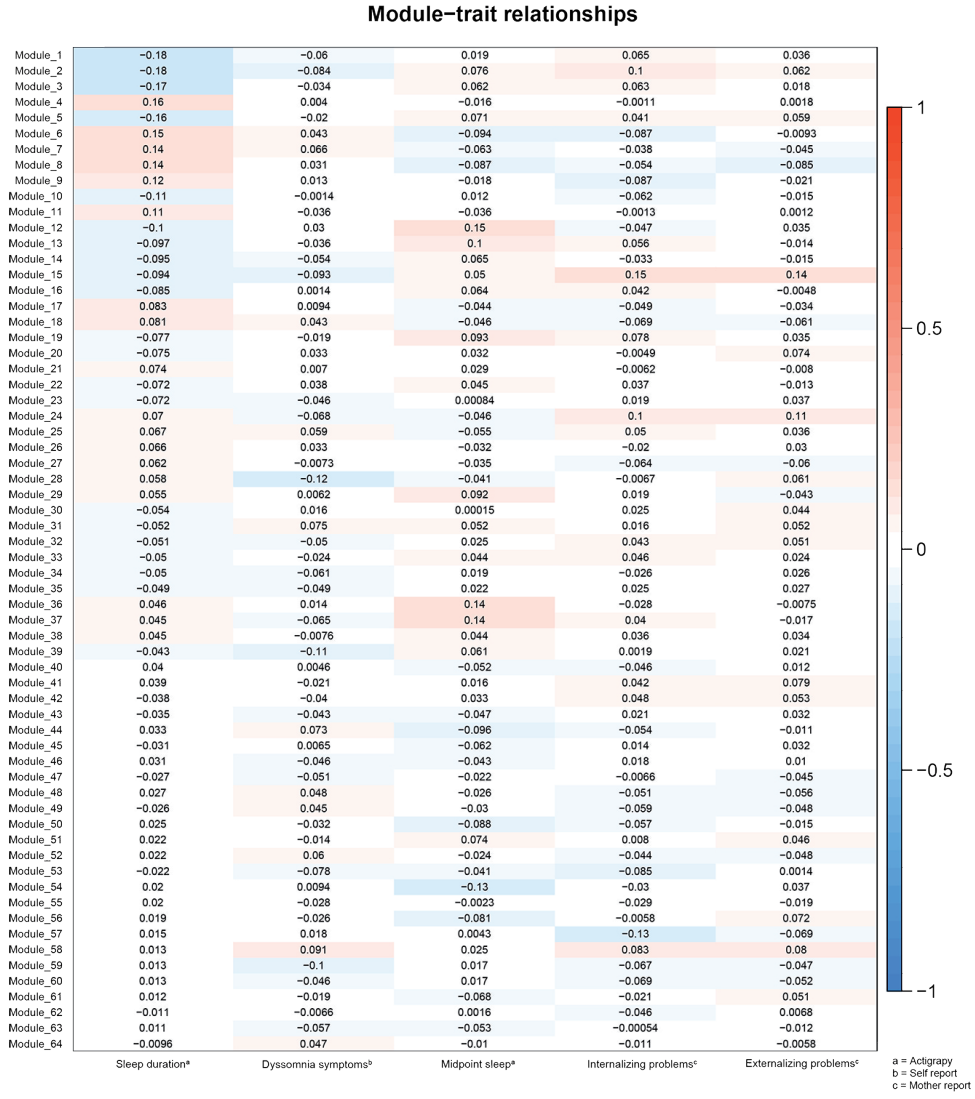
16. Kooijman MN, Kruithof CJ, van Duijn CM, et al. The Generation R Study: design and cohort update 2017. *European Journal of Epidemiolog.y* 2016;31:1243-64.
17. Lehne B, Drong AW, Loh M, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biology.* 2015;16:37.
18. van den Berg CB, Chaves I, Herzog EM, Willemsen SP, van der Horst GTJ, Steegers-Theunissen RPM. Early-and late-onset preeclampsia and the DNA methylation of circadian clock and clock-controlled genes in placental and newborn tissues. *Chronobiology International*. 2017;34:921-32.
19. Bruni O, Ottaviano S, Guidetti V, et al. The Sleep Disturbance Scale for Children (SDSC) Construct ion and validation of an instrument to evaluate sleep disturbances in childhood and adolescence. *Journal of Sleep Research.* 1996;5:251-61.
20. Koopman-Verhoeff ME, Serdarevic F, Kocevska D, et al. Preschool family irregularity and the development of sleep problems in childhood: a longitudinal study. *Journal of Child Psychology and Psychiatry.* 2019;60:857-65.
21. Koopman-Verhoeff ME, Bolhuis K, Cecil CAM, et al. During day and night: Childhood psychotic experiences and objective and subjective sleep problems. *Schizophrenia Research.* 2019;206:127-34.
22. van Hees VT, Fang Z, Langford J, et al. Autocalibration of accelerometer data for free-living physical activity assessment using local gravity and temperature: an evaluation on four continents. *Journal of Applied Physiology.* 2014;117:738-44.
23. Achenbach TM, Rescorla LA. Manual for the ASEBA School-Age Forms & Profiles Burlington. *In: VT, USA: University of Vermont, Research Center for Children, Youth, & Families*, 2001.
24. Ivanova MY, Achenbach TM, Rescorla LA, et al. Preschool psychopathology reported by parents in 23 societies: testing the seven-syndrome model of the child behavior checklist for ages 1.5–5. *Journal of the American Academy of Child & Adolescent Psychiatry.* 2010;49:1215-24.
25. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics.* 2012;13:1.
26. Reinius LE, Acevedo N, Joerink M, et al. Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS ONE.* 2012;7:e41361.
27. R Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available 2013.
28. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics.* 2008;9:559.
29. Botía JA, Vandrovcova J, Forabosco P, et al. An additional k-means clustering step improves the biological features of WGCNA gene co-expression networks. *BMC Systems Biology.* 2017;11:47.
30. Chuang Y-H, Paul KC, Bronstein JM, Bordelon Y, Horvath S, Ritz B. Parkinson's disease is associated with DNA methylation levels in human blood and saliva. *Genome Medicine.* 2017;9:76.
31. Chen Y-a, Lemire M, Choufani S, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics.* 2013;8:203-9.
32. Gaunt TR, Shihab HA, Hemani G, et al. Systematic identification of genetic influences on methylation across the human life course. *Genome Biology.* 2016;17:1-14.

33. Hannon E, Knox O, Sugden K, et al. Characterizing genetic and environmental influences on variable DNA methylation using monozygotic and dizygotic twins. *PLoS Genetics.* 2018;14:e1007544.
34. Hannon E, Lunnon K, Schalkwyk L, Mill J. Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics.* 2015;10:1024-32.
35. Meltzer LJ, Biggs S, Reynolds A, Avis KT, Crabtree VM, Bevans KB. The Children's Report of Sleep Patterns–Sleepiness Scale: A self-report measure for school-aged children. *Sleep Medicine.* 2012;13:385-9.
36. Sadeh A. The role and validity of actigraphy in sleep medicine: an update. *Sleep Medicine Reviews.* 2011;15:259-67.
37. Crowley SJ, Wolfson AR, Tarokh L, Carskadon MA. An update on adolescent sleep: New evidence informing the perfect storm model. *Journal of Adolescence.* 2018;67:55-65.
38. Czamara D, Eraslan G, Page CM, et al. Integrated analysis of environmental and genetic influences on cord blood DNA methylation in new-borns. *Nature Communications.* 2019;10:1-18.
39. Teh AL, Pan H, Chen L, et al. The effect of genotype and in utero environment on interindividual variation in neonate DNA methylomes. *Genome Research.* 2014;24:1064-74.
40. Dashti HS, Jones SE, Wood AR, et al. Genome-wide association study identifies genetic loci for self-reported habitual sleep duration supported by accelerometer-derived estimates. *Nature Communications.* 2019;10:1-12.
41. Doherty A, Smith-Byrne K, Ferreira T, et al. GWAS identifies 14 loci for device-measured physical activity and sleep duration. *Nature Communications.* 2018;9:1-8.
42. Jansen PR, Watanabe K, Stringer S, et al. Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nature Genetics.* 2019;51:394-403.
43. Coupland KG, Kim WS, Halliday GM, Hallupp M, Dobson-Stone C, Kwok JBJ. Role of the long non-coding RNA MAPT-AS1 in regulation of microtubule associated protein tau (MAPT) expression in Parkinson's disease. *PLoS ONE.* 2016;11:e0157924.
44. Wang Y, Mandelkow E. Tau in physiology and pathology. *Nature Reviews Neuroscience.* 2016;17:22.
45. Rankovic M, Zweckstetter M. Upregulated levels and pathological aggregation of abnormally phosphorylated Tau-protein in children with neurodevelopmental disorders. *Neuroscience & Biobehavioral Reviews.* 2019;98:1-9.
46. Cantero JL, Hita-Yañez E, Moreno-Lopez B, Portillo F, Rubio A, Avila J. Tau protein role in sleep-wake cycle. *Journal of Alzheimer's Disease.* 2010;21:411-21.
47. Wasserman D, Wasserman J, Sokolowski M. Genetics of HPA-axis, depression and suicidality. *European Psychiatry.* 2010;25:278-80.
48. Romanowski CPN, Fenzl T, Flachskamm C, et al. Central deficiency of corticotropin-releasing hormone receptor type 1 (CRH-R1) abolishes effects of CRH on NREM but not on REM sleep in mice. *Sleep.* 2010;33:427-36.
49. Bassett SM, Lupis SB, Gianferante D, Rohleder N, Wolf JM. Sleep quality but not sleep quantity effects on cortisol responses to acute psychosocial stress. *Stress.* 2015;18:638-44.

Supplemental Figures



Supplemental Figure 1. Blood-brain associations of *MAPT* CpG sites



Supplemental Figure 2. Correlation matrix of DNA methylation modules, sleep and mental health

Supplemental Tables

Supplemental Table 1. Selection of clock and clock-related genes based on van den Berg et al., 2017 (n_{genesv}= 39; n_{CpGs} = 939)

Gene	CpG (*n*)	Gene	CpG (*n*)	Gene	CpG (*n*)
AKT1	41	*FOXO3*	44	*PRDX6*	15
ARNTL	23	*GSK3B*	19	*PRKAA1*	9
ARNTL2	13	*MAPK1*	18	*PRKAB1*	11
BHLHE40	19	*NPAS2*	34	*PRKACA*	23
BHLHE41	16	*NR1D1*	24	*PRKCA*	79
CLOCK	14	*NR1D2*	17	*RORA*	106
CRTC1	40	*PER1*	18	*RORB*	11
CRY1	14	*PER2*	24	*RORC*	13
CRY2	24	*PER3*	25	*SIRT1*	17
CSNK1E	24	*PRDX1*	17	*SIRT2*	18
DBP	20	*PRDX2*	12	*STRA13*	22
FGF21	10	*PRDX3*	16	*TIMELESS*	15
FOXO1	33	*PRDX5*	22	*WEE1*	19

Supplemental Table 2. WGCNA-derived modules and number of CpG sites

Module	*n* CpG	Module	*n* CpG	Module	*n* CpG	Module	*n* CpG
Module1	*32*	Module17	*669*	Module33	*963*	Module49	*33*
Module2	*5845*	Module18	*794*	Module34	*132*	Module50	*40*
Module3	*17473*	Module19	*1783*	Module35	*757*	Module51	*42*
Module4	*1952*	Module20	*122*	Module36	*30*	Module52	*65804*
Module5	*31*	Module21	*73*	Module37	*42*	Module53	*68*
Module6	*3845*	Module22	*21133*	Module38	*88*	Module54	*1242*
Module7	*47*	Module23	*33*	Module39	*36*	Module55	*448*
Module8	*33*	Module24	*32*	Module40	*448*	Module56	*191*
Module9	*31*	Module25	*31*	Module41	*786*	Module57	*41*
Module10	*35*	Module26	*2596*	Module42	*50775*	Module58	*33*
Module11	*61*	Module27	*125*	Module43	*35*	Module59	*32*
Module12	*31*	Module28	*5750*	Module44	*267*	Module60	*41*
Module13	*47*	Module29	*788*	Module45	*549*	Module61	*9177*
Module14	*30*	Module30	*86*	Module46	*37*	Module62	*33*
Module15	*218*	Module31	*32*	Module47	*63*	Module63	*41*
Module16	*1040*	Module32	*61*	Module48	*261374*	Module64	*56*

Supplemental Table 3. Functional characterization of module1 (part I)

CpG	CpG information								
	Genomic location	Proximity to CpG island	Position	Chromo-some	Gene	Std B	*p*-value	Mean	*SD*
cg15295732	Body		43942128	17	MAPT-L7:L31AS1	0.3018893	9.9E-05	0.70	0.05
cg00846647	Body	Island	44060252	17	MAPT	0.2465164	0.001391	0.85	0.05
cg18228076	5'UTR		43983362	17	MAPT	0.2514921	0.001591	0.45	0.15
cg23955979			45126661	17	[ARL17]	-0.259001	0.001847	0.57	0.05
cg07870213	Body	Island	140052090	5	DND1	-0.234355	0.002187	0.67	0.05
cg24801067			62843696	17	-	-0.236203	0.003393	0.9	0.03
cg17117718		Island	43663208	17	[LRRc37A]	-0.219581	0.003635	0.13	0.1
cg18391203		N_Shelf	44317291	17	-	-0.211059	0.004529	0.57	0.06
cg24801230	5'UTR	S_Shelf	43978533	17	MAPT	0.2130216	0.005724	0.68	0.23
cg16228356			43848958	17	[CRHR1]	0.1982664	0.012255	0.46	0.06
cg22968622		Island	43663579	17	[LRRc37A]	-0.178802	0.017754	0.22	0.21
cg13704117	Body		44207360	17	KANSL1	0.1851593	0.020294	0.78	0.07
cg19832721	TSS1500		44249866	17	KANSL1	-0.187245	0.02096	0.67	0.1
cg04703951			43578652	17	-	-0.176918	0.022176	0.6	0.06
cg20120463		N_Shore	44301886	17	[KANSL1]	0.1776023	0.025006	0.2	0.03
cg04282206	TSS1500	S_Shore	62833786	17	*PLEKHM1P*	-0.163108	0.029668	0.11	0.03
cg27060340	5'UTR	N_Shelf	43502999	17	*ARHGAP27*	0.1872267	0.032394	0.39	0.09
cg01135538	Body	N_Shore	43678735	17	*LOC644172/ MAPK8IP1P2*	0.1659441	0.033713	0.89	0.02
cg03915738			43651976	17	*[LRRc37A]*	-0.166783	0.034135	0.86	0.02
cg01570182		Island	44337453	17	-	-0.151754	0.050768	0.36	0.04
cg15921436		Island	44337874	17	*[ARHGAP27]*	-0.13892	0.064751	0.58	0.04
cg26471390	TSS1500	S_Shore	43511301	17	*ARHGAP27*	-0.156186	0.070719	0.64	0.04
cg17911788		Island	44343683	17	-	0.1410486	0.079972	0.12	0.05
cg05159804		Island	44343776	17	-	0.1404805	0.106976	0.22	0.06
cg02228913	Body	N_Shelf	44058016	17	*MAPT*	0.1238992	0.118991	0.85	0.08
cg06680147		S_Shore	44344931	17	-	-0.104775	0.162643	0.08	0.02
cg09234465		S_Shore	43664173	17	*[LRRc37A]*	-0.091065	0.24945	0.91	0.02
cg08113562	5'UTR	Island	43508428	17	*ARHGAP27*	-0.071283	0.346673	0.09	0.03
cg04927033	Body	Island	43679265	17	*LOC644172/ MAPK8IP1P2*	-0.062307	0.439858	0.88	0.02
cg07368061	Body		44090862	17	*MAPT*	0.0542763	0.519156	0.87	0.02
cg06291494		N_Shore	44321403	17	-	0.0399064	0.629656	0.88	0.02
cg05301556	TSS1500	N_Shore	43971177	17	*MAPT/MAPT-AS1*	0.0075591	0.922981	0.85	0.04

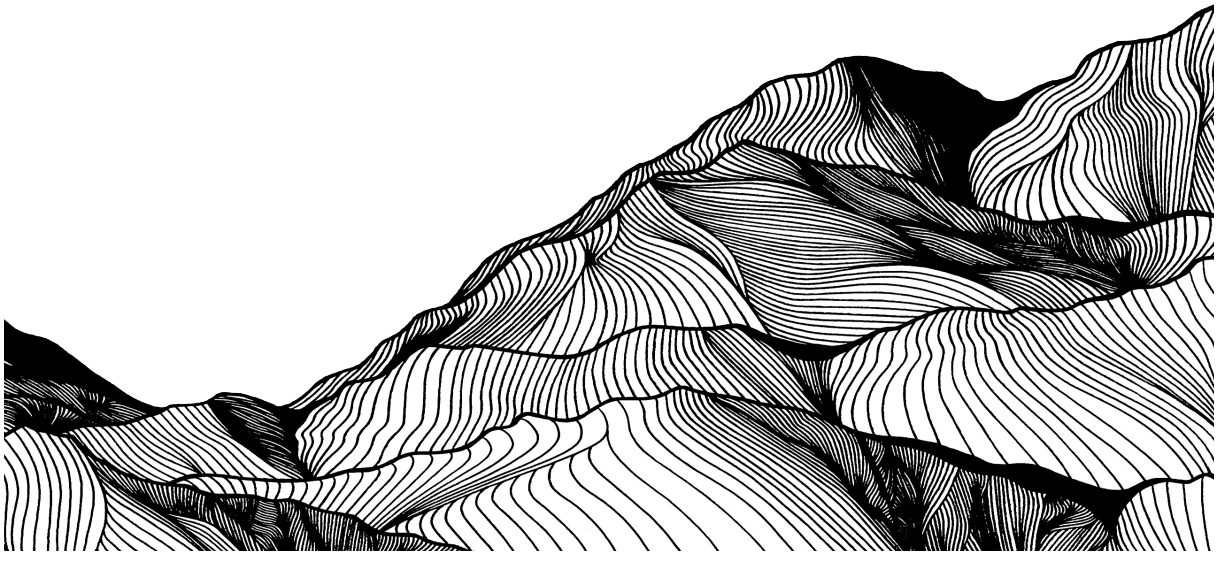
CpGs with a dash (-) are not annotated to genes (intergenic)
CpGs with a gene name are actually annotated/located in that gene
CpGs with a gene in brackets [] are located in proximity to these genes

Supplemental Table 3**.** Functional characterization of module1 (part II)

CpG	Blood-CNS concordance 1 (Hannon et al.. 2015)							
	PFC		STG		EC		CER	
	r	*p*-value	*r*	*p*-value	*r*	*p*-value	*r*	*p*-value
cg15295732	0.422	1.82E-04	0.535	1.53E-06	0.491	7.87E-06	0.291	1.38E-02
cg00846647	0.598	1.82E-08	0.584	9.07E-08	0.475	1.68E-05	0.285	1.58E-02
cg18228076	0.82	3.99E-19	0.829	4.31E-19	0.855	1.53E-22	0.572	1.87E-07
cg23955979	-0.09	4.17E-01	-0.12	3.18E-01	-0.00282	9.81E-01	0.542	1.02E-06
cg07870213	0.848	1.42E-21	0.753	3.47E-14	0.824	1.06E-19	0.652	7.08E-10
cg24801067	0.682	2.31E-11	0.741	1.45E-13	0.695	4.67E-12	0.707	5.44E-12
cg17117718	0.97	4.87E-46	0.961	4.44E-40	0.973	2.86E-48	0.921	6.23E-30
cg18391203	0.6595	6.45E-12	0.577	1.37E-07	0.684	1.37E-11	0.673	1.31E-10
cg24801230	0.986	2.51E-57	0.983	1.61E-52	0.987	2.82E-59	0.835	1.54E-19
cg16228356	0.243	3.73E-02	0.277	1.91E-02	0.262	2.30E-02	0.073	5.45E-01
cg22968622	0.995	1.03E-72	0.99	5.17E-61	0.996	7.64E-80	0.979	2.36E-49
cg13704117	0.81	2.40E-18	0.839	7.25E-20	0.82	1.89E-19	0.855	2.47E-21
cg19832721	0.08	4.87E-01	0.0396	7.43E-01	0.0387	7.41E-01	0.26	2.85E-02
cg04703951	0.717	6.67E-13	0.841	4.65E-20	0.812	9.15E-19	0.842	3.63E-20
cg20120463	0.358	1.73E-03	0.322	6.16E-03	0.211	6.93E-02	0.232	5.13E-02
cg04282206	*0.10*	0.417	*0.30*	1.10E-02	*0.25*	2.87E-02	*0.12*	3.26E-01
cg27060340	*0.71*	1.3E-07	*0.74*	2.58E-13	*0.73*	1.18E-13	*0.75*	5.79E-14
cg01135538	*0.53*	1.54E-06	*0.59*	5.05E-08	*0.52*	2.03E-06	*0.65*	8.70E-10
cg03915738	*-0.08*	0.502	*-0.07*	0.55	*-0.05*	6.72E-01	*0.30*	1.08E-02
cg01570182	*0.77*	1.17E-15	*0.80*	4.08E-17	*0.78*	3.56E-16	*0.62*	7.22E-09
cg15921436	*0.47*	2.83E-05	*0.50*	9.33E-06	*0.67*	4.48E-11	*0.57*	2.61E-07
cg26471390	*-0.23*	0.0229	*-0.41*	0.000352	*-0.52*	1.94E-06	*-0.66*	5.95E-10
cg17911788	*0.51*	3.45E-06	*0.50*	9.72E-06	*0.64*	9.25E-10	*0.17*	1.67E-01
cg05159804	*0.38*	0.00836	*0.22*	7.24E-02	*0.18*	1.21E-01	*-0.35*	3.20E-03
cg02228913	*0.99*	1.32E-61	*0.99*	8.07E-60	*0.99*	1.10E-58	*0.99*	1.27E-54
cg06680147	*0.56*	1.75E-07	*0.46*	6.28E-05	*0.52*	2.28E-06	*0.44*	1.51E-04
cg09234465	*0.41*	0.000242	*0.28*	1.99E-02	*0.44*	7.09E-05	*-0.19*	1.09E-01
cg08113562	*-0.25*	0.0355	*-0.30*	0.0122	*-0.17*	0.155	*-0.19*	0.109
cg04927033	*0.21*	0.0755	*0.22*	6.41E-02	*0.02*	8.44E-01	*0.16*	1.85E-01
cg07368061	*0.12*	0.323	*0.30*	1.09E-02	*0.30*	8.17E-03	*0.16*	1.80E-01
cg06291494	*0.01*	0.89	*0.19*	1.16E-01	*0.00*	9.90E-01	*0.12*	3.26E-01
cg05301556	*0.09*	0.433	*0.21*	7.79E-02	*0.01*	9.46E-01	*0.04*	7.38E-01

Supplemental Table 3. Functional characterization of module1 (part III)

CpG	Twin heritability estimates (Hannon et al.. 2018)			Genetic influences	
	Additive genetic (A)	Common environment (C)	Non-shared environment (E)	mQTLs (Gaunt et al.. 2016)	Polymorphic probes (Chen et al.. 2013)
	Mean	Mean	Mean		
cg15295732	0.66637303	5E-13	0.33362697		TRUE
cg00846647				TRUE	
cg18228076					TRUE
cg23955979					
cg07870213				TRUE	
cg24801067				TRUE	
cg17117718	0.969856123	5.42E-14	0.030143877	TRUE	
cg18391203					TRUE
cg24801230					TRUE
cg16228356	0.837183536	1.34E-13	0.162816464	TRUE	
cg22968622	0.996728437	5.32E-14	0.003271563	TRUE	
cg13704117					TRUE
cg19832721	0.880572078	2.06E-14	0.119427922		
cg04703951				TRUE	
cg20120463				TRUE	
cg04282206				TRUE	
cg27060340	0.712452588	0.061709448	0.225837964		TRUE
cg01135538					
cg03915738					
cg01570182				TRUE	
cg15921436				TRUE	
cg26471390					
cg17911788				TRUE	
cg05159804	0.787265816	2.07E-13	0.212734184		
cg02228913					TRUE
cg06680147					
cg09234465					
cg08113562	0.712193943	3.45E-12	0.287806057		
cg04927033					
cg07368061	0.344706126	0.068495184	0.58679869		
cg06291494					
cg05301556	0.357656393	0.059606674	0.582736933		

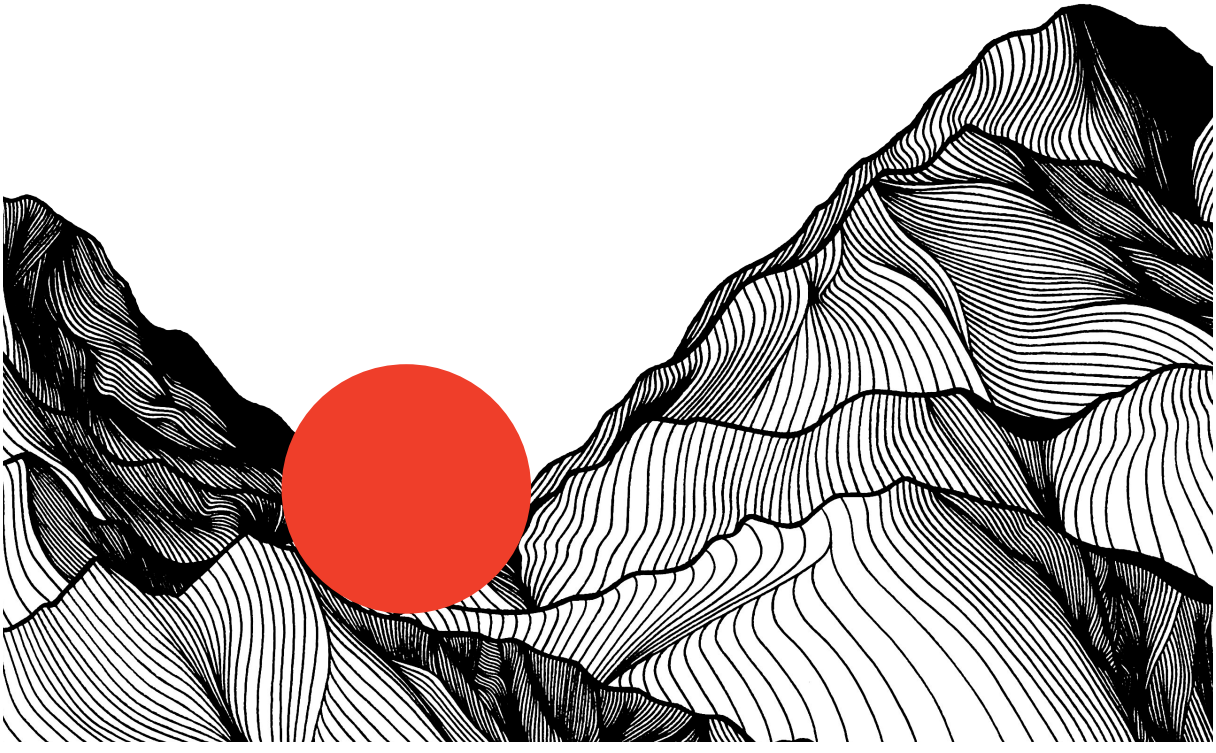


CHAPTER VI

Facing ostracism: Micro-coding facial expressions in the Cyberball social exclusion paradigm

Rosa H. Mulder, Marian J. Bakermans-Kranenburg, Johan Veenstra, Henning Tiemeier, & Marinus H. van IJzendoorn

Submitted.



Abstract

Humans are remarkably reactive to exclusion from social interaction. Social exclusion is often simulated in a lab setting with the Cyberball paradigm, a computerized ball-tossing game. Across the board, Cyberball induces negative feelings, but associations between reactions to Cyberball and individual characteristics have been inconsistent. One reason for this may be that most studies used self-report questionnaires, which leaves the data vulnerable to reporter bias.

Here, we video-recorded children during Cyberball and developed a real-time micro-coding method to measure facial expressions of anger, sadness and contempt during social exclusion in a multi-ethnic population-based sample of 4,813 10-year-olds of the Generation R Study. We tested the utility of these measures by estimating associations between facial expressions and self-reported responses in a post-Cyberball questionnaire. Exploratively, we studied associations between reactions to Cyberball and child characteristics.

We found that facial expressions of sadness and anger were associated with self-reported negative feelings *during* the game, but not with self-reported negative feelings *after* the game. Some associations between facial expressions during Cyberball and child characteristics were found. For example, boys showed more expressions of sadness during the game, whereas girls showed more anger.

We conclude that facial expressions are associated with self-reported negative feelings during Cyberball, but also convey information that is different from self-report. We discuss that reporter-bias can be minimized using facial expressions.

Introduction

Humans are a highly social species. As cooperation has been a necessary tool for survival throughout evolution[1, 2], social interaction is an integral part of human behavior. Not surprisingly, humans are remarkably reactive to exclusion from such interaction[3]. Reactions to social exclusion (or ostracism; being excluded or ignored[4]) are often studied in the lab with the Cyberball game. In this computerized ball-tossing game, the participant will be excluded from the game by two or more avatars[5, 6]. This paradigm has been shown to reliably induce an experience of social exclusion, even when participants are explicitly told that the avatars are not real, when the course of the game is preconceived[7], when they are explained that receiving the ball is tied to a monetary penalty[8], or when participants are led to believe they are excluded by members of a detested organization such as the KKK[9]. Across the board, Cyberball induces negative feelings and threatens fundamental needs of self-esteem, control, belonging, and meaningful existence[5, 6]. Due to its simplicity and effectiveness, Cyberball has proven to be a valuable tool in understanding behavioral and affective responses to social exclusion in humans.

Individual differences in responses to social exclusion

Individual differences in the reactions to Cyberball have been associated with differences in personal characteristics. For example, self-esteem after social exclusion was lower for younger than for older children[10], children with more internalizing problems reported higher threatened needs of belonging, self-esteem, and meaningful existence after Cyberball[11], and children with autism spectrum disorder did not show a decrease in mood after social exclusion during Cyberball, whereas control children did[12]. Associations between reactions to Cyberball and individual characteristics, however, have not always been found. In his review on ostracism, Williams mentions that the effects of ostracism seem very resilient to moderation by individual or situational factors[4]. In a meta-analysis of 120 Cyberball studies, a medium-sized moderation effect of personal or situational variables was found for the impact of social exclusion in Cyberball, but a considerable funnel-plot asymmetry suggested substantial publication bias[5].

Several possible reasons for these inconsistent results have been identified. First of all, it might simply be that the Cyberball paradigm is a victim to its own potency, that is, its strong aversive effect may drown out any associations between individual differences and the reaction to social exclusion. For example, in the study where participants were led to believe that they were excluded by the KKK, those excluded by this despised outgroup reported aversive effect sizes as large as those excluded by a rival outgroup, and as those excluded by in-group members[9]. Second, the research field of Cyberball and individual differences has been plagued by small sample sizes[12, 13], which implies that many studies may not have had the statistical power to detect associations between reactions to social exclusion and

differences in personality or psychopathology. Lastly and most importantly, many studies[11,14,15] used self-report questionnaires to measure the reaction to Cyberball as well as to assess individual characteristics. On the one hand, some individuals may not be able to recognize or verbally express their emotions and feelings, which could lead to deflated associations between reactions to Cyberball and individual characteristics. For example, men tend to be less accurate than women to identify and express their own emotions[16], and the same might be true for individuals in the autism spectrum who might have a lower ability to identify and verbalize emotions[17]. On the other hand, spurious associations might occur between response biases congruent between self-reported predictor and outcome; individuals with extreme responses on one questionnaire might also be extreme on another, whereas others may have a social desirability bias that covers up negative feelings or responses in any questionnaire, thereby inducing shared method variance bias. Together, this indicates that to establish reliable and valid associations between personal characteristics and reactions to Cyberball, one would preferentially use not only self-report data, but also observational data, and test associations in a large sample.

The current study

Here, we video-recorded the faces of children while they played the Cyberball game and developed a new micro-coding method to measure expressions of anger, sadness and contempt during social exclusion in a population-based sample of nearly 5,000 children. We investigate the utility of these measures by studying associations between facial expressions and conventional self-reported responses to Cyberball. Further, we exploratively study associations between reactions to Cyberball and child characteristics. We hypothesize that the micro-coded facial expressions and self-reported responses to Cyberball have shared variance. Second, we hypothesize that the facial expressions and self-reported responses each convey unique information and have distinctive associations with child characteristics.

Methods

Setting

This study is part of the Generation R Study, a prospective population-based birth cohort that follows children and their parents from pregnancy onwards. Pregnant women residing in the municipality of Rotterdam, the Netherlands, with an expected delivery date between April 2002 and January 2006 were invited to participate in the study. An extensive report on the design of the study can be found elsewhere[18]. The Generation R Study is conducted in accordance with the World Medical Association Declaration of Helsinki and has been approved by the Medical Ethics Committee of the Erasmus Medical Center, Rotterdam. Written informed consent was obtained from all parents and children over 12 years old, and informed assent was obtained from all children younger than 12.

Study Population

In the Generation R Study, 9,778 pregnant mothers had 9,749 live-born children. At the age of 10 years, 5,862 of these children visited the research center and 5,708 children performed the Cyberball task. The data of 5,214 children was usable; 586 had procedural issues (*n* = 22, e.g. the child did not understand the game, or was eating while playing the game), technical issues (*n* = 202, e.g. the game was not working, or the video recording was not working), or were excluded because the child was not visible for more than 30 seconds (*n* = 362). Of these children, 4,813 filled out the final version of the questionnaire as well (flow chart in Figure 1).

Attrition analyses (Supplemental Table 1) indicated that children in the initial sample (*n* = 9,749) more often had older siblings (SMD = 0.09, $p < 0.05$), their mothers less often had completed university (SMD = 0.05, $p < 0.05$), and less often had parents that were both born in the Netherlands (SMD = 0.14, $p < 0.05$) compared to the children in the final sample (*n* = 4,813). The selection effects are in line with earlier reports of lower follow-up rates among children with lower socio-economic status[18]. Further, children who visited the research center at 10 years (*n* = 5,862) did not differ from the final sample for the current paper (*n* = 4,813) on these characteristics, but were on average somewhat older than the children in the final sample (SMD = 0.05, $p < 0.05$).

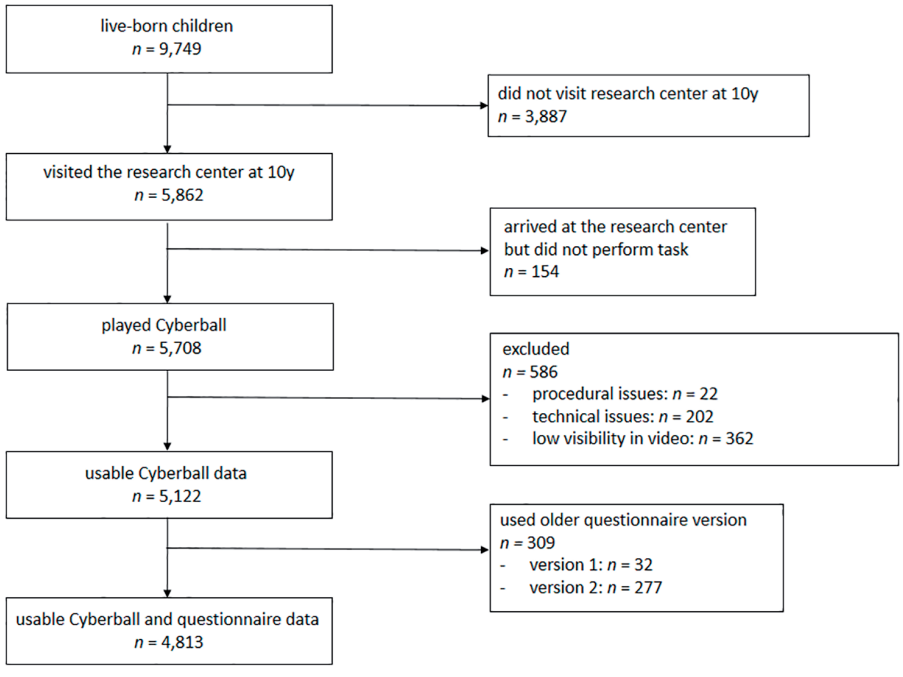


Figure 1. Flow chart of participant inclusion

Procedure

Cyberball was administered during the lab visit at the age of 9.8 (*SD* = 0.3) years. The child sat behind a computer and was told that they would play a ball-tossing game with two other children and asked to imagine that the game was happening in real life. The Cyberball program started with a screen that gave the impression that the Cyberball game was played online. Once the game started, the child saw two pictures of the two putative fellow players, which were pictures of children of the same sex as the participating child. The participating child was represented with a baseball glove and when they received the ball they could chose to which picture they tossed the ball by pressing one of two arrow keys. The game consisted of 42 continuous ball tosses. The Fair Period consisted of the first six tosses, during which the child received the ball twice (toss 3 and toss 6) and lasted on average 11.6 seconds (*SD* = 6.0). The Unfair Period seamlessly followed upon the Fair Period and consisted of 36 tosses which lasted on average 52.4 (*SD* = 2.8) seconds. During the Unfair Period, the child received the ball only twice (toss 15 and toss 25), which was done to retain the child's attention. Unbeknownst to the child, the webcam was turned on to allow recording of the facial expressions. After another, unrelated, dexterity task lasting two minutes, the child was asked to fill out the post-Cyberball questionnaire on their feelings on the computer. During both the Cyberball game and the questionnaire, the experimenter was absent. After completion of the questionnaire, the child was debriefed.

Facial expressions

From the webcam videos, the three most often facially expressed emotions; sadness, anger, and contempt, were micro-coded by four coders. We developed a method to micro-code the facial expressions in real-time. To this end, a Java program was developed[19] that picked up the motions of a joystick and throttle (Thrustmaster Inc) and represented them in two graphs alongside each video, which allowed the scoring of three emotions at the same time: sadness, anger, and contempt (code freely available at https://github.com/AVeenstra/JavaJoyMon). Similar methods have successfully been applied to other observational settings, in which dyadic interpersonal behaviors were coded[20]. With the joystick, a dot could be moved along an x- and y-axis in one graph, with the axes representing sadness and anger, respectively. Movement of the throttle affected the movement of a dot along a single y-axis in a second graph, representing contempt. This system allowed for the micro-coding of multiple emotions at the same time (Figure 2).

The micro-coding method for the emotions was adapted from Ekman's coding of discrete muscle movements[21]. Sadness was coded if the corners of the mouth were dragged downwards and/or the inner corners of the eyebrows upwards and inwards. Anger was coded if the eyebrows moved inwards and downwards and/or the lips were tightened. Contempt was coded if one of the corners of the mouth was pulled outwards. The intensity and combination of movements influenced the height of the score (0-5), which was continually scored by the

position of the joystick and throttle and allowed for the computation of an area under the curve (AUC), representing both the time and intensity of the emotion.

To determine inter-rater reliability, a subset of 129 videos were scored by all four coders. Intercoder reliability (ICC, single measure, absolute agreement) was fair for sadness (ICC = 0.44), and good for anger (ICC = 0.62), contempt (ICC = 0.70) and total emotion (ICC = 0.68)[22].

For the current study, area under the curve (AUC) was computed for the Fair and Unfair Period for each emotion (r = 0.15-0.22 Fair and Unfair Period). Data of children were only included if the child's face was visible for at least 8 seconds during the Fair Period and at least 30 seconds during the Unfair Period. Each AUC was corrected for the total amount of time the child was visible and represents the average AUC per 10 seconds. The Unfair Period score was corrected for the Fair Period score by residualizing the Unfair Period on the Fair Period. The facial expression scores thus reflect both the total time and intensity of emotion shown during the Unfair Period, corrected for emotional expressions during the Fair Period. A total emotion AUC was computed by summing the AUCs of sadness, anger, and contempt, within the Fair and Unfair Period and again residualizing the Unfair Period on the Fair Period. In case of uncertainty in the coding, a second opinion was asked from another coder (2% of videos). Since residualized AUC data for the individual emotions had a positive skew, the residualized AUCs for sadness, anger and total emotion were logarithmically scaled. Since the contempt AUCs had a clear bimodal distribution, with the largest peak with the lower scores of children who did not show any contempt, the score for contempt was dichotomized.

Post-Cyberball Questionnaire

The post-Cyberball questionnaire included 15 questions and was adapted from questionnaires used in previous Cyberball studies[6, 7, 23]. The first nine questions concerned the feelings of the child at the moment of reporting, i.e. after the game; the six subsequent questions concerned

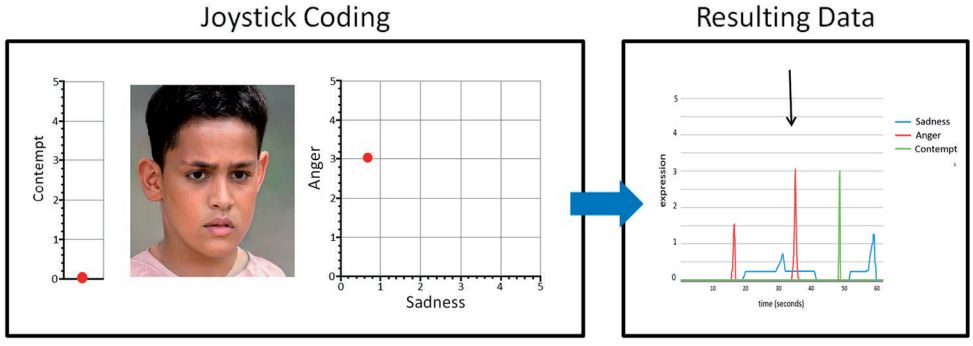


Figure 2. Visual representation of scoring system and resulting data. *Image of child obtained from generated. photos*.

the feelings of the child during the game. Questions concerned mood (e.g. 'I feel happy'; 'I was happy during the ball game on the computer') and self-esteem (e.g. 'My self-esteem is high'; 'My self-esteem was high during the ball game on the computer'). The questions regarding the feelings of the child during the game also included items on feelings of control (e.g. 'The others decided everything during the ball game on the computer') and belonging ('I felt like I belonged to the group during the ball game on the computer'), and the questions regarding the feelings of the child after the game also covered meaningful existence (e.g. 'I feel invisible'). Each item was scored on a Likert-scale from 1 ('Not at all') to 5 ('Very much').

To examine the dimensional structure of the post-Cyberball questionnaire an exploratory factor analysis was performed, which groups the individual items by their covariance structure. Exploratory factor analysis was performed with the *nFactor* R package[24] in R version 3.6.1[25]. Two items; 'I feel important' and 'I feel invisible' were excluded from the analyses due to low factor loadings (< 0.40) in the various factor solutions. A scree plot of the post-Cyberball questionnaire items showed that a three-factor solution was the most optimal, and an exploratory factor analysis with orthogonal rotations ('varimax') led to three factors. However, since the second and the third factor tapped into similar constructs and the third factor included only three items, we decided for a two-factor solution, which also showed a good fit (RMSEA = 0.086 and TLI = 0.82; Supplemental Table 1). The first factor included six items regarding the child's feelings during the game (Eigenvalue = 3.6, 19% explained variance), the second factor included seven items regarding the child's feelings after the game (Eigenvalue 2.2, 16% explained variance). Positively phrased items were reversed, and the subscales were labelled 'Negative feelings during the game' and 'Negative feelings after the game', respectively. Reliability of 'Negative feelings during the game' was good with a Cronbach's alpha of 0.81, and that of 'Negative feelings after the game' was acceptable with a Cronbach's alpha of 0.72.

Child characteristics

Child sex and date of birth were obtained from midwife- and hospital registries. Maternal education, parity, and child ethnicity were obtained via questionnaire at enrollment. Maternal education was used as a proxy of socio-economic status (SES) and dichotomized into children whose mothers have completed a university study (scored as 1) and those who have not (scored as 0). Definition of ethnicity of the child was based on the parent's country of birth[26]. The mothers' country of birth was used in cases where both parents were born abroad. Non-verbal intelligence (IQ) of the child was measured with the Snijders-Oomen Non-verbal intelligence test – Revised (SON-R 2.5-7)[27] at the mean age of 6.1 (*SD* = 0.4) years. The IQ-score was tailored to the age of the child[27, 28], and based on the scores from two subsets with each 15 items; 'Mosaics' for visuospatial abilities, and 'Categories' for abstract reasoning. Last, social unresponsiveness was measured with the Social Responsiveness Scale (SRS), which was completed by the mother when the children were 6.0 (*SD* = 0.4) years old. We

used a brief version of the SRS with 18 items, each rated on a 4-point scale (0 = 'not true',1 = 'sometimes true',2 = 'often true', 3 = 'almost always true') reporting over the past 6 months, covering three domains: Social Cognition (5 items, e.g. 'Is able to understand the meaning of other people's tone of voice and facial expressions'), Social Communication (8 items, e.g. 'Is slow or awkward in turn-taking interactions with peers'), and Autistic Mannerisms (5 items, 'Has repetitive odd behaviors such as hand flapping or rocking'). Weighted average scores allowed for 25% missingness. The SRS has been shown to be a reliable and valid measure[29] and has been related to autism diagnoses[30]. In the current sample, Cronbach's alpha was 0.77.

Statistical analyses

All analyses were performed in R version 3.6.1[25]. Missing data on child characteristics and predictors of interest (*n* = 163 for parity, *n* = 397 for education of the mother, *n* = 712 for non-verbal IQ, *n* = 114 for ethnicity, *n* = 1,270 autism-related traits) were imputed using the *mice* R package[31] using a maximum of 100 iterations creating 30 datasets. All results are based on pooled estimates of the multiply imputed sets.

Basic characteristics of the facial expressions were studied on the group level in three ways: (1) by examining the time-course of the facial expressions by computing a 3-second moving average, (2) by comparing the AUC of the Fair and Unfair Period, averaged over the duration of each respective period, for each child separately by computing the difference between the two AUCs and on a group level by paired T-tests of the two logarithmically scaled AUCs, and (3) by studying the maximum intensity of each micro-coded facial expression by computing the median highest peak for each child.

Associations between facial expressions (total emotion, sadness, anger, contempt) and self-reported feelings in the post-Cyberball questionnaire ('Negative feelings during the game' and 'Negative feelings after the game') were measured in linear regressions with the facial expression as the predictor and the post-Cyberball questionnaire scale as the outcome. Each combination of facial expressions and post-Cyberball questionnaire scale was adjusted for by child characteristics (child sex, age during Cyberball, parity, maternal education, non-verbal IQ, ethnicity, and social unresponsiveness). With a more explorative aim, associations between child characteristics and reactions to Cyberball (facial expressions and post-Cyberball questionnaire scales) were estimated using linear regressions with the child characteristics as the predictors and each reaction to Cyberball as the outcome. Last, in a series of sensitivity analyses, linear regressions were repeated within the non-imputed dataset (*n* = 3,041).

Since novel measurements were tested in this study, we did not correct for multiple testing and set α at 0.05. Together with the sample size of *N* = 4,813, this allowed us (statistical power = 0.95, number of predictors = 9) to detect small effect sizes of $f^2 = 0.005$ and above.

Results

Sample characteristics are displayed in Table 1. The time-course of the facial expressions was studied on a group level by computing a 3-second moving average. Figure 3 shows that the group average expression increased over the course of the Cyberball paradigm but that the expression was less intense each time the child received the ball. This pattern was present for each negative emotion coded. A comparison of the AUC score for the Unfair Period and the Fair Period showed that most children had more negative facial expression during the Unfair Period than during the Fair period, indicating that the social exclusion paradigm was effective in inducing negative emotions: 75.0% showed more sadness, 70.8% showed more anger, 24.0% showed more contempt, and 87.9% had higher scores on Total Emotion. Paired t-tests between the logarithmically scaled Unfair and Fair AUCs showed that the former was higher than the latter for sadness ($t[4812] = 58.82$, $p < 2.20 \times 10^{-16}$), anger, ($t[4812] = 55.28$, $p < 2.20 \times 10^{-16}$), contempt ($t[4812] = 21.20$, $p < 2.20 \times 10^{-16}$), and Total Emotion ($t[4812] = 73.40$, $p < 2.20 \times 10^{-16}$). The intensities of the facial expressions were modest: on a scale from 0 to 5, the median highest peak for sadness was 0.11 (*SD* = 0.93), for anger 0.78 (*SD* = 0.78), for contempt 1.76 (*SD* = 1.00), and for total emotion 1.18 (*SD* = 1.20).

The correlations between the AUCs for the different facial expressions are depicted in Table 2. Sadness and anger were positively associated ($r = 0.13$, $p = 2.07 \times 10^{-19}$), as were anger and contempt ($r = 0.03$, $p = 0.042$). No association was found between sadness and contempt. The two scales of the post-Cyberball questionnaire, 'Negative feelings during the game' and 'Negative feelings after the game' correlated positively ($r = 0.26$, $p < 2.40 \times 10^{-73}$).

Table 1. Sample characteristics

Sex, No. boys (%)	2368 (49.2)
Age at Cyberball in years, mean (*SD*)	9.80 (0.34)
Parity, mean number of older siblings (*SD*)	0.59 (0.83)
Education of the mother, No. university (%)	1282 (26.6)
non-verbal IQ, mean (*SD*)	101.70 (14.97)
Ethnicity, No. (%)	
Dutch	2856 (59.3)
non-Dutch, Western	433 (9.0)
non-Dutch, non-Western	1524 (31.7)
Social unresponsiveness, mean (*SD*)	0.23 (0.25)
Facial expressions – Total emotion, mean (*SD*)	1.46 (0.18)
Facial expressions – Sadness, mean (*SD*)	1.56 (0.15)
Facial expression – Anger, mean (*SD*)	1.24 (0.20)
Facial expression – Contempt, No. yes (%)	1228 (25.5)
Questionnaire – Negative feelings during the game, mean (*SD*)	2.85 (0.84)
Questionnaire – Negative feelings after the game , mean (*SD*)	1.57 (0.45)

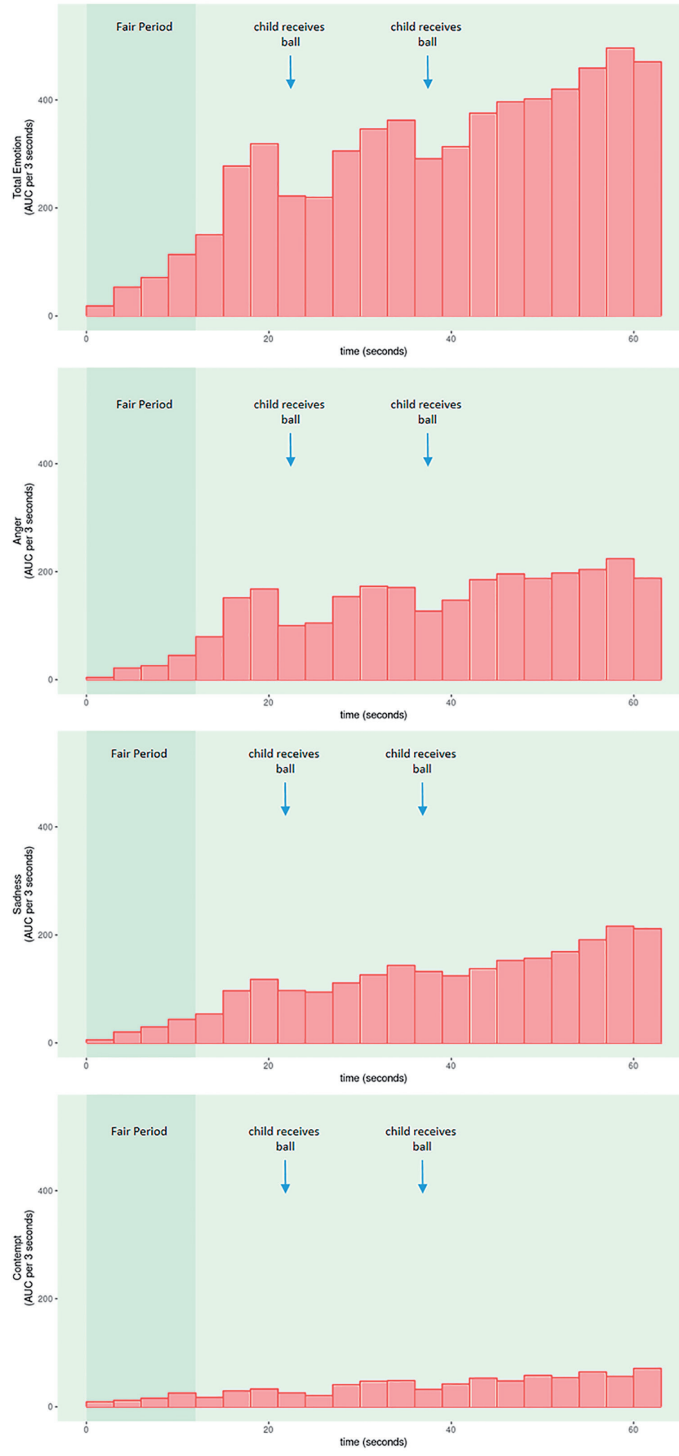


Figure 3. Mean area under the curve (AUC) over each three seconds of the Cyberball time course

Table 2. Correlations between facial expressions during Cyberball

	Sadness	Anger	Contempt
Total emotion	0.63**	0.75**	0.21**
Sadness		0.13**	-0.01
Anger			0.03*
Contempt			

*$p < 0.05$
**$p < 0.0001$

Facial expressions and self-reported feelings

Associations between the AUCs of the facial expressions and self-reported feelings in the post-Cyberball questionnaire were examined with linear regressions with the facial expressions as predictor and self-report as outcome. These showed that total emotion, sadness, and anger associated positively with 'Negative feelings during the game' (β = 0.09, β = 0.09, and β = 0.06, respectively, $p < 9.70x10^{-06}$) (Table 3). Contempt was not associated with 'Negative feelings during the game'. There were no associations between the facial expressions and 'Negative feelings after the game'.

Child characteristics and facial expressions

Associations between child characteristics and facial expressions are depicted in Table 4. Boys showed more sadness (β = 0.09, $p = 0.002$) than girls, whereas girls showed more anger than boys (β = -0.10, $p = 4.10x10^{-04}$). Children of non-Western ethnicity showed more Total Emotion (β = 0.07, $p = 0.042$) than those of Western ethnicity. Children from non-Dutch, Western ethnicity also showed more anger (β = 0.09, $p = 0.015$), as did children from non-Western ethnicity (β = 0.14, $p = 0.009$). Last, older children showed less contempt than younger children did (β = -0.10, $p = 0.006$). No other associations were found between child characteristics and facial expressions.

Child characteristics and self-reported feelings

Boys reported more negative feelings during the game (β = 0.18, $p = 9.05x10^{-10}$) than girls, as did children of mothers who completed university versus those whose mothers did not (β = 0.07, $p = 0.032$), and children with higher versus lower non-verbal IQ (β = 0.08, $p = 7.42x10^{-08}$)

Table 3. Associations between facial expressions during Cyberball and self-reported feelings in post-Cyberball questionnaire

	Negative feelings during the game			Negative feelings after the game		
	β	95% CI	*p*	β	95% CI	*p*
Total emotion	0.09	0.06; 0.12	$1.21x10^{-10}$	0.00	-0.03; 0.02	0.819
Sadness	0.09	0.06; 0.12	$7.35x10^{-10}$	0.00	-0.02; 0.03	0.805
Anger	0.06	0.04; 0.09	$9.70x10^{-06}$	0.00	-0.03; 0.03	0.861
Contempt	0.01	-0.06; 0.07	0.821	-0.01	-0.08; 0.05	0.713

Table 4. Associations between child characteristics and facial expressions during Cyberball

	Total emotion			Sadness			Anger			Contempt		
	β	**95% CI**	***p***	**β**	**95% CI**	***p***	**β**	**95% CI**	***p***	**β**	**95% CI**	***p***
Sex (boy)	-0.02	-0.08; 0.04	0.510	0.09	0.03; 0.15	0.002	-0.10	-0.16; -0.05	4.10×10^{-04}	0.05	-0.08; 0.18	0.446
Age	0.01	-0.02; 0.04	0.422	0.01	-0.02; 0.04	0.542	0.02	-0.01; 0.05	0.143	-0.10	-0.17; -0.03	0.006
Parity	-0.01	-0.03; 0.02	0.721	-0.01	-0.04; 0.02	0.653	0.00	-0.03; 0.03	0.940	-0.03	-0.10; 0.04	0.352
Maternal education (high)	-0.03	-0.10; 0.04	0.416	-0.03	-0.10; 0.04	0.390	-0.02	-0.09; 0.05	0.588	-0.10	-0.26; 0.06	0.229
Non-verbal IQ	0.02	-0.01; 0.05	0.235	0.01	-0.02; 0.05	0.357	0.01	-0.02; 0.05	0.419	0.01	-0.06; 0.09	0.699
Ethnicity – non-Dutch Western[a]	0.09	-0.01; 0.19	0.083	0.01	-0.10; 0.11	0.904	0.14	0.03; 0.23	0.009	-0.24	-0.49; 0.00	0.051
Ethnicity – non-Western[a]	0.07	0.00; 0.14	0.042	0.05	-0.02; 0.11	0.179	0.09	0.02; 0.15	0.015	-0.13	-0.29; 0.03	0.108
Social unresponsiveness	-0.03	-0.07; 0.00	0.091	-0.03	-0.07; 0.00	0.056	-0.03	-0.07; 0.00	0.087	-0.03	-0.04; 0.11	0.428

[a]Reference group: Dutch

Table 5. Associations between child characteristics and self-reported feelings in post-Cyberball questionnaire

	Negative feelings during the game			Negative feelings after the game		
	β	**95% CI**	***p***	**β**	**95% CI**	***p***
Sex (boy)	0.18	0.12; 0.23	$9.05x10^{-10}$	-0.05	-0.11; 0.01	0.097
Age	0.00	-0.03; 0.03	0.997	-0.03	-0.05; 0.00	0.078
Parity	-0.03	-0.06; -0.01	0.018	0.01	-0.02; 0.04	0.686
Maternal education (high)	0.07	0.01; 0.14	0.032	0.02	-0.04; 0.09	0.477
Non-verbal IQ	0.08	0.05; 0.12	$7.42x10^{-08}$	-0.09	-0.13; -0.06	$4.26x10^{-08}$
Ethnicity – non-Dutch Western[a]	0.02	-0.08; 0.12	0.663	0.04	-0.06; 0.14	0.442
Ethnicity – non-Western[a]	-0.10	-0.17; -0.03	0.004	-0.06	-0.13; 0.01	0.100
Social unresponsiveness	-0.02	-0.06; 0.01	0.169	0.06	0.02; 0.10	$9.06x10^{-04}$

[a]Reference group: Dutch

(Table 4). Children with more older siblings reported less negative feelings during the game than those with less older siblings (β = -0.03, *p* = 0.018), as did children of non-Western ethnicity (β = -0.10, *p* = 0.004). In contrast to the positive association between non-verbal IQ and negative feelings during the game, children with a higher non-verbal IQ reported less negative feelings after the game (β = -0.13, $p = 4.26x10^{-08}$) (Table 5). Last, children with more social unresponsiveness also reported more negative feelings after the game (β = 0.06, $p = 9.06x10^{-04}$).

Since non-Western children reported strikingly low levels of negative feelings during the game, we performed a follow-up analysis on the specific ethnic groups within our non-imputed sample (*n* = 4,699) for all reactions to Cyberball (Supplemental Table 3). Results showed that children of Turkish/Moroccan parents reported less negative feelings during the game, but did not show less negative facial expressions during the game. Following this up one step further, children of Moroccan parents in particular reported both less negative feelings during and after the game, but they did not show any differences in facial expressions (Supplemental Table 4).

Sensitivity analyses on non-imputed set

Analyses were repeated on a set where child characteristics were not imputed (*n* = 3,041). Results for associations between facial expressions and post-Cyberball questionnaire scales were confirmed in this set (Supplemental Table 6-8). Results for associations between child characteristics and facial expressions were largely in line with the imputed data.

Discussion

In the current study, we aimed to introduce and examine a system for the micro-coding of facial expressions in the Cyberball social exclusion paradigm, a frequently used computerized

paradigm simulating social exclusion. In a population-based sample of 10-year old children, we found that most children showed negative emotional responses to the paradigm via their facial expressions. The facial expressions of sadness and anger were associated with post-Cyberball self-report on negative feelings *during* the game, but not with negative feelings *after* the game. The facial expression of contempt was not associated with self-reported feelings during or after the game.

The discrepancy between the self-reported feelings *during* and *after* the game with respect to their associations with the facial expression of sadness and anger was very consistent across most measures. This might indicate that the two different assessments tap into two different stages of emotion processing in social exclusion. In previous research, two stages of reaction to social exclusion have been identified: 1) the reflexive stage, and 2) the reflective stage[32]. Reactions to Cyberball in the reflexive stage are immediate and thus measured ideally during the Cyberball game, though in practice often directly after the game via self-report[5]. Reactions in the reflective stage are thought to be subject to coping mechanisms and affected by the speed of recovery from the social exclusion experience, and are generally measured a few minutes to almost an hour after the end of the game[9, 33, 34]. Here, both were measured a few minutes after the game, thus even though the construct *Feelings during the game* may have been retrospectively reflexive rather than immediate reflexive, it was associated with the observed facial expressions, whereas the *Feelings after the game* may be more reflective. This might imply that a strong observed reaction during the game does not mean that the child still feels bad a few minutes after the game.

No association between facial expressions of contempt and self-reported feelings was found. The association between expressions of contempt and anger was not very strong, which is in line with existing literature[35]. In this series of studies, contempt was shown to be related to anger as both often co-occur in an attempt to reassert social status in a negative social interaction. Contempt co-occurred with interpersonal distancing, whereas anger was shown to be a more short-term emotion and associated with reconciliation. Anger may be more adaptive in a social situation as the one simulated in the Cyberball paradigm, and older children may have learned this, which would explain our finding that older children showed less contempt than younger children did. Research on contempt during development is however scarce, making the interpretation of this finding speculative, and pointing to the need of developmental research on (the expression of) contempt.

Responses to social exclusion and individual differences

The finding that boys showed more sadness and girls more anger contradicts other research, in which 8-year old boys displayed more anger during social rejection than girls[36]. A developmental perspective might explain the difference between studies; a meta-analytic review revealed that while boys express more anger and girls more sadness during middle childhood, during

adolescence girls showed more anger than boys[37]. Another reason for girls to display more anger might be that girls have a higher preference for equity in social situations than boys do[38]. The Cyberball paradigm violates this expectancy, which might frustrate girls more than boys.

Children of immigrant Moroccan parents showed the same amount of negative facial expressions as other children, but they reported to have less negative feelings during and after the game. It has been previously reported that children of Moroccan national origin reported to express less negative emotions towards their family members than children from Dutch origin[39]. These findings underline the importance of observational measures in research to counter reporter bias. Furthermore, these findings are relevant for teachers and physicians in the Netherlands, indicating that children of immigrant parents might not be as vocal about their problems and needs as other children are.

Last, children with mothers with a higher education and children with a higher non-verbal IQ reported to have had more negative feelings during the Cyberball paradigm, whereas children with a higher non-verbal IQ reported *less* negative feelings after the game. Children with lower socio-economic status tend to get bullied more often[40], which could mean that the Cyberball paradigm triggered memories of earlier exclusion events, which may have made them to take longer to recover from the social exclusion.

Strengths, limitations, and future directions

The size of the current population-based sample allowed for the detection of small effect sizes. Associations found between facial expressions and self-report were highly significant, which decreases the chance that these were false positives. Most of the reported associations detected in the explorative analyses between child characteristics and reactions to Cyberball were highly significant as well, but a few (for parity, maternal education) would not survive a multiple testing correction and warrant replication. Overall, we note that the large majority of the children in the current study reacted with facial expressions of negative emotions. This observation is in line with Williams[4] and Hartgerink et al.[5], who noted that the social exclusion paradigm has pervasive negative effects on participants. Further, Williams and Hartgerink observed that effects of Cyberball are relatively robust against moderation by personal or situational factors; the relatively small effect sizes in the current non-clinical cohort study are concordant with that observation. Yet, such effects are of interest because they may point at larger individual differences in children that are more sensitive to social exclusion because of their atypical (bullying or parenting) experiences or constitutional characteristics. In future research, non-verbal reactions to Cyberball in longitudinal samples with an over-representation of bullied or socially unresponsive children might shed some light on the origins and development of individual differences in nonverbal responses to social rejection.

Conclusion

In conclusion, this is the first time that this micro-coding method was employed to study facial expressions during Cyberball. Results showed concurrent validity with self-reported feelings during the game. Results also indicate that the real-time micro-coding system conveys information that is distinct from self-report. We suggest that self-report questionnaires may be subject to a social desirability response bias in specific groups of children, underlining the value of observational methods.

References

1. Bowles S, Gintis H. Origins of human cooperation. *Genetic and Cultural Evolution of Cooperation.* 2003;2003:429-43.
2. Boyd R, Richerson PJ. Solving the puzzle of human cooperation. *Evolution and Culture*. 2006:105-32.
3. Williams KD. Ostracism: The kiss of social death. *Social and Personality Psychology Compass*. 2007;1:236-47.
4. Williams KD. Ostracism. *Annual Reviews in Psychology*. 2007;58:425-52.
5. Hartgerink CHJ, Van Beest I, Wicherts JM, Williams KD. The ordinal effects of ostracism: A meta-analysis of 120 Cyberball studies. *PLoS ONE.* 2015;10:e0127002.
6. Williams KD, Cheung CKT, Choi W. Cyberostracism: Effects of being ignored over the Internet. *Journal of Personality and Social Psychology.* 2000;79:748.
7. Zadro L, Williams KD, Richardson R. How low can you go? Ostracism by a computer is sufficient to lower self-reported levels of belonging, control, self-esteem, and meaningful existence. *Journal of Experimental Social Psychology*. 2004;40:560-7.
8. Van Beest I, Williams KD. When inclusion costs and ostracism pays, ostracism still hurts. *Journal of Personality and Social Psychology*. 2006;91:918.
9. Gonsalkorale K, Williams KD. The KKK won't let me play: Ostracism even by a despised outgroup hurts. *European Journal of Social Psychology*. 2007;37:1176-86.
10. Abrams D, Weick M, Thomas D, Colbe H, Franklin KM. On-line ostracism affects children differently from adolescents and adults. *British Journal of Developmental Psychology*. 2011;29:110-23.
11. Hawes DJ, Zadro L, Iannuzzelli R, et al. Internalising problems and the effects of peer ostracism on children's primary needs. *International Journal of Developmental Science*. 2013;7:41-5.
12. Sebastian C, Blakemore S-J, Charman T. Reactions to ostracism in adolescents with autism spectrum conditions. *Journal of Autism and Developmental Disorders*. 2009;39:1122-30.
13. McPartland JC, Crowley MJ, Perszyk DR, et al. Temporal dynamics reveal atypical brain response to social exclusion in autism. *Developmental Cognitive Neuroscience*. 2011;1:271-9.
14. Levinson CA, Langer JK, Rodebaugh TL. Reactivity to exclusion prospectively predicts social anxiety symptoms in young adults. *Behavior Therapy*. 2013;44:470-8.
15. Ruggieri S, Bendixen M, Gabriel U, Alsaker F. Do victimization experiences accentuate reactions to ostracism? An experiment using Cyberball. *International Journal of Developmental Science.* 2013;7:25-32.
16. Levant RF, Hall RJ, Williams CM, Hasan NT. Gender differences in alexithymia. *Psychology of Men & Masculinity*. 2009;10:190.
17. Poquérusse J, Pastore L, Dellantonio S, Esposito G. Alexithymia and autism spectrum disorder: a complex relationship. *Frontiers in Psychology*. 2018;9:1196.
18. Kooijman MN, Kruithof CJ, van Duijn CM, et al. The Generation R Study: design and cohort update 2017. *European Journal of Epidemiology.* 2016;31:1243-64.

19. IntelliJ-IDEA. the most intelligent Java IDE. In: JetBrains [online].[cit. 2016-02-23]. Dostupné z: https://www. jetbrains. com/idea/# chooseYourEdition, 2011.
20. Sadler P, Ethier N, Gunn GR, Duong D, Woody E. Are we on the same wavelength? Interpersonal complementarity as shared cyclical patterns during interactions. *Journal of Personality and Social Psychology*. 2009;97:1005.
21. Ekman P. Methods for measuring facial action. *Handbook of Methods in Nonverbal Behavior Research.* 1982:45-90.
22. Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*. 1994;6:284.
23. Eisenberger NI, Lieberman MD, Williams KD. Does rejection hurt? An fMRI study of social exclusion. *Science.* 2003;302:290-2.
24. Raiche G, Magis D, Raiche MG. Package 'nFactors'. Parallel Analysis and Non Graphical Solutions to the Cattell Scree Test. Available at CRAN repository: http://cran. r-project. org/web/packages/n Factors/n Factors. pdf 2010.
25. R Core Team. *R: A language and environment for statistical computing*. 2013.
26. Statistics Netherlands. *Immigrants in the Netherlands 2004 (Allochtonen in Nederland 2004)*. Den Haag/Heerlen: Statistics Netherlands (Centraal Bureau voor de Statistiek) 2004.
27. Tellegen PJ, Winkel M, Wijnbergen-Williams BJ. *Snijders-Oomen Niet-verbale Intelligentietest-Revisie SON-R 2½–7.* In: Lisse: Swets & Zeitlinger, 1996.
28. Langeslag SJE, Schmidt M, Ghassabian A, et al. Functional connectivity between parietal and frontal brain regions and intelligence in young children: the Generation R study. *Human Brain Mapping.* 2013;34:3299-307.
29. Wigham S, McConachie H, Tandos J, Le Couteur AS, Gateshead Millennium Study core t. The reliability and validity of the Social Responsiveness Scale in a UK general child population. *Research in Developmental Disabilities*. 2012;33:944-50.
30. Constantino JN, Davis SA, Todd RD, et al. Validation of a brief quantitative measure of autistic traits: comparison of the social responsiveness scale with the autism diagnostic interview-revised. *Journal of Autism and Developmental Disorders*. 2003;33:427-33.
31. Buuren Sv, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*. 2010:1-68.
32. Williams KD. Ostracism: A temporal need-threat model. *Advances in Experimental Social Psychology*. 2009;41:275-314.
33. Gonsalkorale K, Carter-Sowell AR, Sloan LR, Williams KD. The KKK won't let Tyrone play; but inclusion hurts too. In: *Presentation at the annual meeting of the Society for Personality and Social Psychology*; 2008, 2008.
34. Zadro L, Boland C, Richardson R. How long does it last? The persistence of the effects of ostracism in the socially anxious. *Journal of Experimental Social Psychology*. 2006;42:692-7.
35. Fischer AH, Roseman IJ. Beat them or ban them: The characteristics and social functions of anger and contempt. *Journal of Personality and Social Psychology*. 2007;93:103.

36. Hubbard JA. Emotion expression processes in children's peer interaction: The role of peer rejection, aggression, and gender. *Child Development*. 2001;72:1426-38.
37. Chaplin TM, Aldao A. Gender differences in emotion expression in children: A meta-analytic review. *Psychological Bulletin*. 2013;139:735.
38. Meuwese R, Crone EA, de Rooij M, Güroğlu B. Development of equity preferences in boys and girls across adolescence. *Child Development*. 2015;86:145-58.
39. Novin S, Rieffe CJ. Het ervaren en uiten van emoties door Marokkaans-Nederlandse en autochtoon-Nederlandse kinderen. *Kind en Adolescent*. 2006;27:106.
40. Jansen PW, Verlinden M, Dommisse-van Berkel A, et al. Prevalence of bullying and victimization among children in early elementary school: Do family and school neighbourhood socioeconomic status matter? *BMC Public Health*. 2012;12:494.

Supplemental Tables

Supplemental Table 1. Post-Cyberball questionnaire and loadings on exploratory factor analysis

		factor 1 ***Feelings during the game (retrospectively reported)***	**factor 2** ***Feelings after the game***	**excluded**
1	I feel good about myself		X	
2	My self-esteem is high		X	
3	I feel invisible			X
4	I feel important			X
5	I feel bad		X	
6	I feel happy		X	
7	I dislike myself		X	
8	I feel sad		X	
9	I feel angry		X	
10	I felt like I belonged to the group during the ball game on the computer	X		
11	The others decided everything during the ball game on the computer	X		
12	The others let me decide during the ball game on the computer	X		
13	I was angry during the ball game on the computer	X		
14	I was happy during the ball game on the computer	X		
15	My self-esteem was high during the ball game on the computer	X		

Supplemental Table 2. Attrition analyses

	set 1 – live born children	**set 2 – visited research center age 10**	**set 3 – usable Cyberball data**	**set 1 vs set 3** ***p***[a]	**set 2 vs set 3** ***p***[a]
n	9749	5862	4813		
Sex, No. boys (%)	4940 (50.7)	2914 (49.7)	2368 (49.2)	0.098	0.614
Parity, mean number of older siblings (*SD*)	0.66 (0.90)	0.60 (0.83)	0.59 (0.83)	$2.44x10^{-05}$	0.561
Education of the mother, No. high (%)	2255 (23.1)	1551 (26.5)	1282 (26.6)	$3.84x10^{-06}$	0.853
Ethnicity, No. (%)				$3.42x10^{-13}$	0.973
Dutch	5163 (53.0)	3485 (59.5)	2856 (59.3)		
non-Dutch, Western	907 (9.0)	532 (9.1)	433 (9.0)		
non-Dutch, non-Western	3679 (37.7)	1845 (31.5)	1524 (31.7)		
Age at Cyberball in years, mean (*SD*)		9.81 (0.36)	9.80 (0.34)		0.035
non-verbal IQ, mean (*SD*)		101.58 (15.03)	101.70 (14.97)		0.679
Social unresponsiveness, mean (*SD*)		0.24 (0.26)	0.23 (0.25)		0.197

[a]*p*-value as based on T-test for numerical variables and X-square for categorical variables.

Supplemental Table 3. Differences in reactions to Cyberball for children of different ethnicities (n = 4,699)

		Total emotion			Sadness			Anger			Contempt			*Feelings during*			*Feelings after*		
Ethnicity[a]	***n***	**β**	**95% CI**	***p***	**β**	**95% CI**	***p***	**β**	**95% CI**	***p***	**β**	**95% CI**	***p***	**β**	**95% CI**	***p***	**β**	**95% CI**	***p***
other Western[b]	424	0.02	0.00;0.03	0.086	0.00	-0.01;0.02	0.901	0.03	0.01;0.05	0.010	-0.24	-0.49;0.00	0.052	0.02	-0.06;0.11	0.603	0.02	-0.03;0.06	0.470
Turkish/ Moroccan[c]	549	0.01	-0.01;0.02	0.417	-0.01	-0.02;0.01	0.263	0.02	0.00;0.04	0.076	0.02	-0.20;0.24	0.830	-0.14	-0.22;-0.06	$4.40x10^{-04}$	-0.05	-0.09;-0.01	0.021
African[d]	236	0.01	-0.02;0.03	0.645	0.02	-0.01;0.04	0.140	0.01	0.02;0.03	0.621	-0.25	-0.57;0.08	0.133	-0.12	-0.23;-0.01	0.035	0.00	-0.06;0.06	0.938
Caribbean[e]	464	0.02	0.00;0.04	0.022	0.02	0.00;0.03	0.050	0.02	0.00;0.15	0.017	-0.17	-0.41;0.06	0.151	-0.05	-0.13;0.04	0.278	0.00	-0.05;0.04	0.905
other[f]	214	0.01	-0.02;0.03	0.585	0.01	-0.01;0.04	0.187	0.00	-0.03;0.03	0.844	-0.27	-0.60;0.07	0.119	0.01	-0.11;0.12	0.931	-0.05	-0.11;0.01	0.100

[a]Reference group: Dutch (*n* = 2812)
[b]European (*n* = 354); American, Western (*n* = 29); Asian, Western (*n* = 6); Oceanian (*n* = 9); Indonesian (*n* = 26)
[c]Turkish (*n* = 311); Moroccan (*n* = 238)
[d]African (*n* = 100); Cape Verdian (*n* = 136)
[e]Dutch-Antillean (*n* = 130); Surinamese-Creole (*n* = 134); Surinamese-Hindustani (*n* = 137); Surinamese-Unspecified (*n* = 63)
[f]American, non-Western (*n* = 88); Asian, non-Western (*n* = 126)
Associations adjusted for sex, age, parity, maternal education, non-verbal IQ, and social unresponsiveness

Supplemental Table 4. Differences in reactions to Cyberball for Turkish and Moroccan children versus Dutch children

		Total emotion			Sadness			Anger			Contempt			*Feelings during*			*Feelings after*		
Ethnicity[a]	***n***	**β**	**95% CI**	***p***	**β**	**95% CI**	***p***	**β**	**95% CI**	***p***	**β**	**95% CI**	***p***	**β**	**95% CI**	***p***	**β**	**95% CI**	***p***
Turkish	238	0.01	-0.01;0.03	0.364	-0.01	-0.03;0.00	0.117	0.03	0.00;0.05	0.023	0.11	-0.16;0.38	0.418	-0.10	-0.19;0.00	0.054	0.02	-0.03;0.07	0.446
Moroccan	311	0.00	-0.02;0.03	0.786	0.00	-0.02;0.02	0.999	0.00	-0.02;0.03	0.795	-0.10	-0.42;0.22	0.538	-0.21	-0.32;0.09	$3.58x10^{-04}$	-0.15	-0.21;-0.09	$1.93x10^{-06}$

[a]Reference group: Dutch (*n* = 2812)
Associations adjusted for sex, age, parity, maternal education, non-verbal IQ, and social unresponsiveness

Supplemental Table 5. Associations between facial expressions during Cyberball and self-reported feelings in post-Cyberball questionnaire in non-imputed set (*n* = 3,041)

	Negative feelings during the game			Negative feelings after the game		
	β	**95% CI**	***p***	**β**	**95% CI**	***p***
Total emotion	0.08	0.05; 0.12	$3.30x10^{-06}$	-0.01	-0.05; 0.02	0.450
Sadness	0.09	0.05; 0.12	$8.83x10^{-07}$	0.00	-0.03; 0.04	0.846
Anger	0.05	0.02; 0.09	$3.01x10^{-03}$	0.00	-0.04; 0.03	0.850
Contempt	0.01	-0.03; 0.04	0.649	-0.01	-0.04; 0.03	0.780
Negative feelings during the game				0.26	0.22; 0.29	$9.28x10^{-45}$

Supplemental Table 6. Associations between child characteristics and facial expressions during Cyberball in non-imputed set (*n* = 3,041)

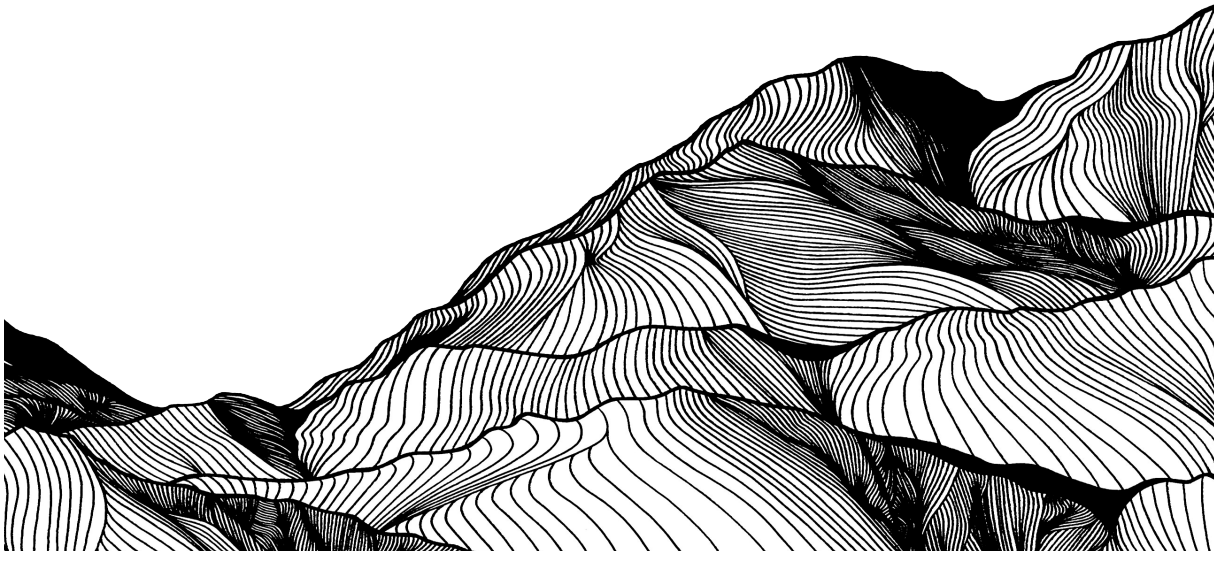
	Total emotion			Sadness			Anger			Contempt		
	β	**95% CI**	***p***	**β**	**95% CI**	***p***	**β**	**95% CI**	***p***	**β**	**95% CI**	***p***
Sex (boy)	0.01	-0.06; 0.09	0.695	0.12	0.04; 0.19	0.002	-0.08	-0.15; -0.01	0.030	0.09	-0.08; 0.25	0.309
Age	0.02	-0.02; 0.05	0.370	0.00	-0.03; 0.04	0.876	0.03	-0.01; 0.07	0.121	-0.06	-0.14; 0.03	0.211
Parity	-0.02	-0.06; 0.01	0.218	-0.01	-0.05; 0.03	0.645	-0.02	-0.06; 0.01	0.237	-0.03	-0.11; 0.06	0.521
Maternal education (high)	-0.04	-0.12; 0.04	0.363	-0.06	-0.15; 0.02	0.141	-0.01	-0.09; 0.07	0.803	-0.21	-0.40; -0.02	0.032
Non-verbal IQ	0.02	-0.01; 0.06	0.231	0.01	-0.03; 0.05	0.703	0.02	-0.02; 0.06	0.318	0.04	-0.05; 0.13	0.367
Ethnicity – other Western[a]	0.11	-0.02; 0.23	0.101	0.05	-0.08; 0.18	0.419	0.13	0.00; 0.26	0.045	-0.22	-0.52; 0.08	0.149
Ethnicity – non-Western[a]	0.02	-0.07; 0.11	0.669	0.03	-0.06; 0.13	0.486	0.02	-0.08; 0.11	0.723	-0.20	-0.41; 0.02	0.071
Social unresponsiveness	-0.03	-0.07; 0.01	0.086	-0.03	-0.07; 0.01	0.117	-0.05	-0.08; -0.01	0.019	-0.01	-0.09; 0.08	0.865

[a]Reference group: Dutch

Supplemental Table 7. Associations between child characteristics and self-reported feelings in post-Cyberball questionnaire in non-imputed set (*n* = 3,041)

	Negative feelings during the game			***Negative feelings after the game***		
	β	**95% CI**	***p***	**β**	**95% CI**	***p***
Sex (boy)	0.20	0.13; 0.28	$3.67x10^{-08}$	-0.06	-0.12; 0.02	0.126
Age	0.00	-0.04; 0.04	0.989	-0.03	-0.06; 0.01	0.151
Parity	-0.03	-0.07; 0.00	0.084	0.02	-0.02; 0.06	0.270
Maternal education (high)	0.08	0.00; 0.16	0.061	0.06	-0.02; 0.14	0.139
Non-verbal IQ	0.08	0.05; 0.12	$1.35x10^{-05}$	-0.08	-0.12; -0.04	$2.71x10^{-05}$
Ethnicity – other Western[a]	0.04	-0.08; 0.17	0.486	0.02	-0.11; 0.15	0.754
Ethnicity – non-Western[a]	-0.08	-0.17; 0.01	0.078	-0.08	-0.17; 0.01	0.078
Social unresponsiveness	-0.02	-0.06; 0.02	0.343	0.08	0.04; 0.12	$1.80x10^{-05}$

[a]Reference group: Dutch

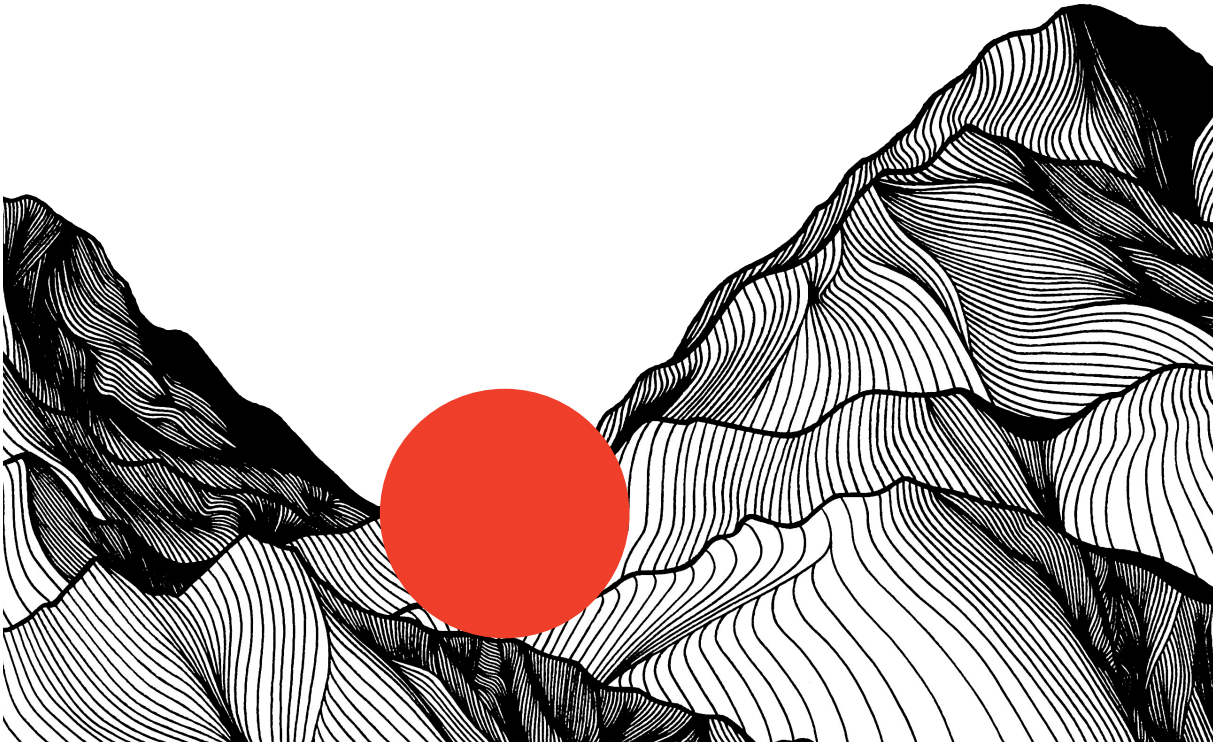


CHAPTER VII

Epigenomics of being bullied: Changes in DNA methylation following bullying exposure

Rosa H. Mulder, , Esther Walton, Alexander Neumann, Lotte C. Houtepen, Janine F. Felix, Marian J. Bakermans-Kranenburg, Matthew Suderman, Henning Tiemeier, Marinus H. van IJzendoorn, Caroline L. Relton, & Charlotte A. M. Cecil

Epigenetics (2020), 15(6-7), 750-764.



Abstract

Bullying among children is ubiquitous and associated with pervasive mental health problems. However, very little is known about the biological pathways that change after exposure to bullying.

Epigenome-wide changes in DNA methylation in peripheral blood were studied from pre- to post measurement of bullying exposure, in a longitudinal study of the population-based cohorts Generation R Study and Avon Longitudinal Study of Parents and Children (combined N = 1,352). Linear mixed-model results were meta-analyzed to estimate how DNA methylation changed as a function of exposure to bullying. Sensitivity analyses including co-occurring child characteristics and risks was performed, as well as a Gene Ontology analysis. A candidate gene follow-up was employed for CpG (cytosine-phosphate-guanine) sites annotated to *5-HTT* and *NR3C1*.

One CpG site, cg17312179, showed small changes in DNA methylation associated to bullying exposure ($b = 2.67x10^{-03}$, SE = $4.97x10^{-04}$, $p = 7.17x10^{-08}$). This site is annotated to *RAB14*, an oncogene related to Golgi apparatus functioning, and its methylation levels decreased for exposed whereas they increased for non-exposed. This result was consistent across sensitivity analyses. Enriched Gene Ontology pathways for differentially methylated sites included cardiac function and neurodevelopmental processes. Top CpG sites tended to have overall low levels of DNA methylation, decreasing in exposed, while increasing in non-exposed individuals. There were no gene-wide corrected findings for *5-HTT* and *NR3C1*.

This is the first study to identify changes in DNA methylation associated with bullying exposure at the epigenome-wide significance level. Consistent with other population-based studies, we do not find evidence for strong associations between bullying exposure and DNA methylation.

Introduction

The social environment is a major contributor to mental health. Bullying is a ubiquitous social stressor, with worldwide estimates ranging from one in ten to almost half of all children that are exposed[1]. Following Olweus' definition, a person is being bullied 'when he or she is exposed, repeatedly and over time, to negative actions on the part of one or more other persons'. Such negative actions should be intentional and performed by someone perceived be more powerful than the subject. Actions can include physical behaviors, such as hitting and kicking, verbal behaviors, such as calling names, as well as indirect or relational behaviors, such as social exclusion[2]. Bullying exposure (i.e. bullying victimization) has been associated with numerous mental health issues including behavioral problems, depressive symptoms [3-6] and suicidal ideation[7]. However, whereas a myriad of harmful and persistent psychiatric consequences of being bullied have been identified, the biological pathways that change after exposure to bullying remain largely uncharted. Identifying these pathways is a pivotal step in understanding how peer-inflicted stress affects the human body.

Research on other environmental stressors, such as parental abuse[8], prenatal maternal stress[9, 10] or childhood trauma in general[11-15] has incorporated epigenetic data to investigate the hypothesis that stressors affect the molecular configuration on and around the DNA, thereby influencing its functionality, with potential downstream effects on stress reactivity and mental health[16-19]. One often studied epigenetic mechanism is DNA methylation, in which a methyl-group binds to a cytosine nucleotide of the DNA (cytosine-phosphate-guanine site or CpG site). Whereas early epigenetic studies focused on DNA methylation of a single candidate gene, there has been an increase in hypothesis-free epigenome-wide methylation studies (EWASs) investigating DNA methylation levels of hundreds of thousands of CpG sites (CpGs) across the genome. One study[11], for example, found multiple epigenome-wide significant differentially methylated CpGs related to different types of childhood maltreatment.

In contrast to other forms of adversity, research on bullying exposure and epigenetics is markedly scarce. To the best of our knowledge, only three such studies have been performed. In 28 monozygotic twin pairs discordant for bullying exposure[20], increased levels of methylation were observed in the serotonin transporter gene (*5-HTT)* promoter region for the exposed twin siblings from 5 to 10 years, but not for the non-exposed twin siblings. Another study in 1,149 13 to 14 year old children found bullying exposure to be associated with increased methylation levels of exon 1F of the glucocorticoid receptor gene (*NR3C1*)[21]. Further, an EWAS was performed on bullying exposure in 1,658 twins[13], thereby expanding the search for differentially methylated sites beyond the 'usual suspects', i.e. candidate genes that have been firmly implicated in neurotransmitter and hormonal functions, to enable the identification of potentially novel biological pathways. Bullying exposure during childhood as reported by mother and child at age 7-12 years, and bullying exposure during adolescence, retrospectively

reported by the child at 18 years, were however not related to differential methylation. Given that DNA methylation is expected to change over time[22] due to both extrinsic as well as intrinsic factors, a model in which DNA methylation both before and after bullying exposure is taken into account should be more sensitive to the effects of exposure.

In the current study, we made use of two population-based cohorts featuring repeated measures of DNA methylation to characterize longitudinal epigenome-wide associations with bullying exposure. Longitudinal mixed models were performed separately in the two cohorts to identify associations between exposure to bullying and changes in DNA methylation from pre- to post bullying report. Results were then meta-analyzed to maximize statistical power and to evaluate coherence among the estimates derived from the two populations. Epigenome-wide associations with bullying were studied in a hypothesis-free analysis. In a secondary candidate gene follow-up analysis we examined DNA methylation at *5-HTT* and *NR3C1* for gene-wide associations with bullying.

Results

Sample characteristics

Sample characteristics are described in Table 1. In Generation R, bullying exposure was reported by the mother at the mean (*SD*) age of 8.1 (0.1) years, and DNA methylation was measured at 6.0 (0.3) years and 9.8 (0.3) years of age. In the Avon Longitudinal Study of Parents and Children (ALSPAC), bullying exposure was reported by the child at the mean (*SD*) age of 8.6 (0.2) years, and DNA methylation was measured at 7.5 (0.1) and 17.1 (1.0) years of age (Supplemental Figure 1). In the main analysis 45.5% of children in the Generation R sample and 39.4% of children in the ALSPAC sample were categorized as exposed to bullying victimization. In the sensitivity analysis with a more stringent definition of bullying, these numbers are 9.9% and 12.1%, respectively. The current selected samples for each cohort were compared with (i) a set of participants with complete data on covariates, and (ii) a set of participants with complete data on both covariates and bullying exposure (full details in Supplemental Table 1). This showed that children in the current sets had a higher gestational age and higher SES, had mothers who were older, had a higher non-verbal intelligence quotient (IQ) score and were older at bullying exposure report. In Generation R, but not ALSPAC, children in the selected sample also had a lower Body Mass Index (BMI), less behavioral problems prior to exposure, and had less reported stressful experiences other than bullying exposure. No differences were found for child sex, or bullying exposure.

Table 1. Sample characteristics

	Generation R (*n* = 506)		**ALSPAC (*n* = 846)**	
Age in years bullying exposure report (mean (*SD*))	8.1 (0.1)		8.6 (0.2)	
Sex (No. (%) boys)	251 (49.6)		407 (48.3)	
Gestational age in weeks (mean (*SD*))	40.2 (1.4)		39.6 (1.5)	
Maternal education (No. (%))				
Low	21 (4.2)		72 (8.5)	
Medium	101 (20.0)		336 (39.7)	
High	384 (75.9)		438 (51.8)	
Maternal age at delivery (mean (*SD*))	32.8 (3.9)		29.7 (4.4)	
Bullying exposure (No. (%) yes)	229 (45.3)		333 (39.4)	
Bullying exposure–sensitivity analysis (No. (%) yes)	50 (9.9)		102 (12.1)	
Behavioral problem score (mean (*SD*)) (GenR *n* = 451, ALSPAC *n* = 794)	17.3 (12.2)		6.9 (3.9)	
Intelligence quotient (mean (*SD*)) (GenR *n* = 465, ALSPAC *n* = 811)	107.3 (14.0)		102.6 (16.7)	
Other stressful experiences (mean (*SD*)) (GenR *n* = 482, ALSPAC *n* = 597)	3.7 (2.1)		1.5 (1.4)	
Alcohol use (mean (*SD*)) (ALSPAC *n* = 624)			8.1 (4.8)	
Methylation measurement	**T1 (*n* = 404)**	**T2 (*n* = 391)**	**T1 (*n* = 820)**	**T2 (*n* = 819)**
Age in years DNA methylation (mean (*SD*))	6.0 (0.3)	9.8 (0.3)	7.5 (0.1)	17.1 (1.0)
BMI in kg/m^2 (mean (*SD*))	15.9 (1.3	17.1 (2.0)	16.2 (2.0)	22.6 (3.6)
AHRR CpG quintiles (No. (%))				
0.769, 0.873	93 (23.0)	66 (16.9)	123 (15.0)	205 (25.0)
0.873, 0.891	79 (19.6)	80 (20.5)	146 (17.8)	182 (22.2)
0.891, 0.906	72 (17.8)	87 (22.3)	171 (20.9)	157 (19.2)
0.906, 0.920	73 (18.1)	86 (22.0)	186 (22.7)	142 (17.3)
0.920, 0.963	87 (21.5)	72 (18.4)	194 (23.7)	133 (16.2)

SD: standard deviation; No.: number; T1: time point 1; T2: time point 2; BMI: Body Mass Index

Comparison of longitudinal epigenome-wide association studies

The separate longitudinal EWASs for the two cohorts (Q-Q plots in Supplemental Figure 3), identified no epigenome-wide Bonferroni-significant associations, with a lowest obtained *p*-value of $p = 5.93 \times 10^{-06}$ (CpG site cg034529555, annotated to *NAV2*) for Generation R and of $p = 1.08 \times 10^{-06}$ (CpG site cg24506221, annotated to *GSTM1*) in ALSPAC. Estimates for bullying exposure among the top 1000 CpG sites in each cohort more often had a negative direction (79.8% in Generation R, 66.3% in ALSAPC) than would be expected by chance ($X^2(1) = 355.22$, $p < 2.20 \times 10^{-16}$ for Generation R, $X^2(1) = 106.28$, $p < 2.20 \times 10^{-16}$ for ALSPAC). Bullying exposure estimates among the top 1000 in Generation R were not significantly correlated with those in ALSPAC ($r(998) = -0.05$, $p = 1.09 \times 10^{-01}$). The correlation between the top 1000 in ALSPAC was slight but significantly negative ($r(998) = -0.07$, $p = 1.09 \times 10^{-02}$) with those in Generation R.

Meta-analysis

In the meta-analysis, one CpG site was significantly associated with bullying exposure: cg17312179 (b = -2.67x10^{-03}, SE = 4.97x10^{-04}, p = 7.17x10^{-08}; Supplemental Figure 4), a site in the leader sequence (5'UTR) of the *RAB14* gene, located on chromosome 9. In Generation R (b = -2.47x10^{-03}, SE = 6.17x10^{-04}, p = 6.25x10^{-05}), DNA methylation of this CpG site increased on average 0.13% from the mean age of 6.0 to 9.8 years in non-exposed, but decreased -0.12% in exposed. In ALSPAC (b = -3.05x10^{-03}, SE = 8.36x10^{-04}, p = 2.64x10^{-04}) methylation of this CpG site increased 0.09% from the mean age of 7.5 to 17.1 years in non-exposed, whereas it decreased -0.21% in exposed (Figure 1). DNA methylation differences at cg17312179 between the non-exposed and the exposed group were not present before exposure measurement, but were so after (Supplemental Analysis). The ten CpGs with the lowest p-values are shown in Table 2 (see Supplemental Table 2 for associated functions).

The top 1000 CpGs from the meta-analysis had a higher representation of CpGs with a negative estimate for bullying exposure (83.5%) than all other CpGs (58.2%, $X^2(1) = 261.07$, $p < 2.20x10^{-16}$). Moreover, the top 1000 CpGs had a higher representation of CpGs with a positive age estimate (representing change in non-exposed, 79.2%) than the other CpGs (65.1%, $X^2(1) = 87.29$, $p < 2.20x10^{-16}$), and a higher representation of CpGs with low levels of DNA methylation (mean β value < 0.2 in both Generation R and ALSPAC, threshold as elsewhere[36]) at both time points (55.7% versus 37.0%, $X^2(1) = 149.63$, $p < 2.20x10^{-16}$), and

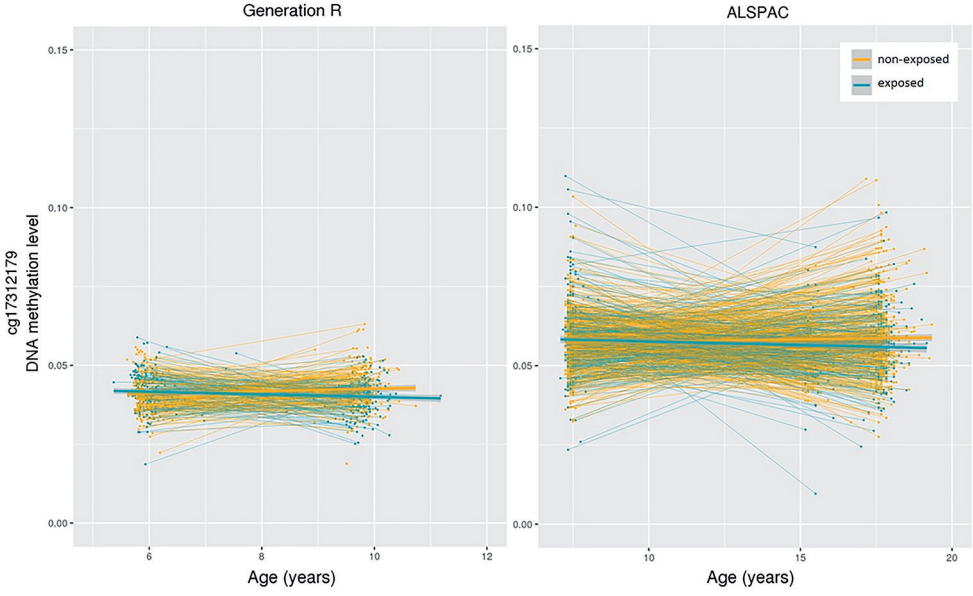


Figure 1. Change in DNA methylation pre- and post- bullying exposure measurement for exposed and non-exposed in Generation R and ALSPAC. Data are residualized for covariates present in linear mixed model

Table 2. Ten CpG sites with lowest *p*-values in meta-analysis of epigenome-wide associations with bullying exposure

CpG site	Gene	Chr	Relation to gene	Relation to CpG site	*B* (SE)	*p*-value	Direction of change exposed / non-exposed
cg17312179	*RAB14*	9	5'UTR	Island	-2.67e-03 (4.97e-04)	7.17e-08	-/+
cg09291817	*MAZ*	16	TSS1500	Island	-2.01e-03 (4.03e-04)	6.21e-07	-/+
cg11278602	*HCG4*	6	Body	Island	-2.98e-03 (6.41e-04)	3.35e-06	-/+
cg00911813	*TNRC18*	7	5'UTR	Island	-1.32e-03 (2.91e-04)	5.37e-06	-/+
cg08971637	*DGUOK*	2	TSS1500	N shore	-9.63e-03 (2.14e-03)	6.66e-06	-/+
cg12767834	*SNPH*	20	5'UTR	Island	-2.01e-03 (4.46e-04)	7.01e-06	-/+
cg26394220	*MIR375; CCDC108*	2	TSS1500; Body	Island	-2.40e-03 (5.40e-04)	8.92e-06	-/+
cg19790568	*PRX*	19	Body	Island	-1.49e-03 (3.36e-04)	9.08e-06	-/+
cg10929442	*ST8SIA4*	5	Body	N shore	-1.01e-03 (2.29e-04)	9.60e-06	-/+
ch.4.134822993R		4			-1.20e-03 (2.72e-04)	1.11e-05	-/+

SE: standard error; exposed/non-exposed; exposed to bullying victimization/not exposed to bullying victimization

Table 3. Characteristics of CpG sites selected for various levels of significance in meta-analysis

CpG sites	*n*	Negative bullying exposure coefficient (%)	Positive coefficient non-exposed (%)	Low methylation (%)	Promoter associated (%)	CpG island associated (%)
all	473864	58.3	65.1	37.0	20.1	30.8
$P < 0.1$	53168	71.2**	69.2**	44.1**	23.1**	34.6**
$P < 0.01$	5997	78.6**	75.3**	50.5**	26.3**	37.9**
$P < 0.001$	644	85.1**	79.7**	56.4**	28.1**	40.8**
$P < 0.0001$	66	86.4**	83.3**	62.1**	30.3**	47.0**
$P < 0.00001$	9	100.0**	100.0*	88.9**	44.4	77.8*

Note: CpG sites were classified as having low methylation if β value < 0.2 in Generation R and ALSPAC.
* $p < 0.05$ compared to all other CpG sites, ** $p < 0.001$ compared to all other CpG sites

more often located in promoters (28.8% versus 20.1%, $X^2(1) = 46.66$, $p = 8.45 \times 10^{-12}$) and CpG islands (41.8% versus 30.8%, $X^2(1) = 56.63$, $p = 5.25 \times 10^{-14}$). Together, this indicates that top sites tended to be located in promoters and CpG islands, and to have overall low levels of DNA methylation, decreasing in exposed, while increasing in non-exposed. See Table 3 for these characteristics at multiple *p*-thresholds.

Follow-up analyses

Sensitivity analyses

A series of sensitivity analyses were performed on cg17312179 in each cohort and then meta-analyzed. First, an analysis was performed with a more stringent definition of bulling

exposure. Second, we reran separate analyses or additionally adjusting for (i) BMI; (ii) pre-existing behavior problems; (iii) non-verbal intelligence quotient (IQ); (iv) stressful experiences other than bullying exposure; and (v) alcohol use. The bullying exposure coefficients from sensitivity analyses were not different from the bullying exposure coefficient from the main analysis (lowest $p = 7.42x10^{-02}$). 'Other stressful experiences' was the only added variable that independently associated with cg17312179 ($b = 2.20x10^{-04}$, SE = $9.96x10^{-05}$, $p = 2.73x10^{-02}$).

Genetic associations

A triad of look-ups did not show evidence of genetic associations with cg17312179 methylation. First, the probe was not present in a list of polymorphic probes[23]. Second, no *cis* or *trans* methylation quantitative trait loci (mQTLs) were found to associate with this probe[24], and third, low additive genetic influences ($1.79x10^{-10}$%) and high shared (34.9%) and non-shared (65.1%) environmental influences have been reported for this probe based on twin heritability analyses[25].

Look-up of previous findings in the literature

Results from eleven EWASs on childhood adversity[11, 13, 15, 16, 26-32] were searched for cg17312179 or other CpGs annotated to *RAB14.* No *RAB14*-associated probes were reported. Since cg17312179 was not reported in these studies, we could not establish if the direction of association was congruent with the one currently reported.

Candidate gene-wide analyses

Results from the meta-analysis were separately studied for probes annotated to *5-HTT* and *NR3C1* (Supplemental Figure 5). None of the probes reached gene-wide Bonferroni-significance (thresholds $p = 3.13x10^{-03}$ for *5-HTT* and $p = 1.22x10^{-03}$ for *NR3C1*).

Functional associations

Gene Ontology (GO) analysis on CpGs with $p < 0.001$ ($n = 644$ CpGs, $n = 396$ genes) yielded 126 pathways, 25 of which were confirmed by a GO analysis on CpGs with $p < 0.01$ ($n = 5997$ CpGs, $n = 3722$ genes) and 43 of which were confirmed by a GO analysis on CpGs with $p < 0.0001$ ($n = 66$ CpGs, $n = 53$ genes). Ryanodine-sensitive calcium-release channel activity as the most enriched ($p = 9.99x10^{-08}$) biological process (Supplemental Tables 4-6, Supplemental Figure 6). Three isoforms of the ryanodine receptors exist [33], RYR1, RYR2, and RYR3, each present in a different tissue. Here, *RYR2* was part of the GO pathway, a gene specifically active in the heart tissue. Other enriched terms for biological processes involve various neurodevelopmental processes, such as astrocyte differentiation and action potential regulation, as well as processes such as muscle fiber development.

Discussion

The current study is the first to characterize epigenome-wide intra-individual changes in DNA methylation related to bullying exposure. Our meta-analysis identified a CpG site with increasing levels of DNA methylation in non-exposed but decreasing levels in the exposed group. Other research [23-25] on this probe suggests that variance in DNA methylation at this CpG is primarily explained by environmental influences, with weak evidence of genetic effects. Sensitivity analyses showed that this association was not explained by co-occurring child characteristics, co-occurring risks, or consequences of bullying, including pre-existing behavioral problems, IQ, BMI, alcohol use or exposure to stressful experiences other than bullying. The site is located in the 5' untranslated region of *RAB14*, a member of the Ras oncogene family of GTPases. Ras GTPases are important in cellular signaling and RAB14 is involved in vesicle transport and Golgi apparatus functioning[34], and is expressed in multiple tissues (Supplemental Figure 7). No *RAB14*-associated probes have been reported in previous EWAS on childhood adversity.

RAB14 expression has however been associated to stress in different tissues. In rat hippocampus it was shown to be downregulated after prenatal stress[35] and upregulated after mild chronic stress in stress-resilient rats[36], possibly marking an adaptive response. In humans, its expression was found to be reduced in the prostate of men with prostate cancer after nutrition and lifestyle intervention focused on stress reduction[37]. Also, *RAB14* expression in human brain tissue has been linked to depression and suicide[38], as was Syntaphilin, a protein that regulates synaptic vesicle processing[39] and encoded by *SYNPH*, a gene associated to one of the top 10 CpG sites from the meta-analysis. How the observed changes in *RAB14* methylation might relate to expression levels in this gene and what the downstream effects of these changes might be, however, remains to be elucidated in future functional studies.

GO analysis showed enrichment of the biological process of ryanodine-sensitive calcium release channel activity; these channels are a pathway important in cardiac functioning and the fight-or-flight response[40]. This finding is congruent with several other enriched pathways associated with cardiac functioning, and fits with GO findings from other research on epigenetics and physical abuse [11]. Further, GO analysis showed many neurodevelopmental processes, such as neuron differentiation, a biological pathway in which two of the associated genes, *SYNPH* and *ST8SIA4,* were among the top meta-analysis hits. ST8SIA4 (CMP-N-acetylneuraminate-poly-alpha-2,8-sialyltransferase) is present in the Golgi apparatus, involved in neural plasticity[41], and *ST8SIA4* knockout mice have been shown to display a decreased motivation for social interaction[42]. Functioning of both genes has been associated with brain disorders, such as schizophrenia[43, 44] and Alzheimer's disease[45, 46]. Interestingly, in contrast to previous studies[20, 21], no associations with bullying exposure were found at candidate genes *5-HTT* and *NR3C1*. Failure to replicate candidate epigenetic studies with epigenome-wide analyses is not

uncommon[13]. This discrepancy may be explained by the stricter multiple testing correction applied in (candidate gene analyses as part of) epigenome-wide studies, or in the different specific regions tested by targeted gene approaches and microarray studies, rendering direct comparison unfeasible.

A pattern emerged of enrichment for CpGs with low overall methylation levels, increased over time, but decreased for exposed individuals. Furthermore, the top CpGs from the meta-analysis were more often located in CpG islands and promoter regions than would be expected by chance. Together this might indicate that bullying exposure is associated with an overall delayed downregulation of gene expression. However, promoter regions typically have low levels of methylation, and the enrichment of CpGs with low overall methylation levels in general seems to be more pronounced than the enrichment of promoter CpGs. In an EWAS on childhood abuse and promoter DNA methylation in adulthood[8], the stressor was also more often negatively than positively associated with DNA methylation. In another EWAS on childhood maltreatment and DNA methylation around 10 years of age (range 5-14 years)[29], researchers found an enrichment of CpGs with low methylation levels, often located in promoter regions as we did in the current study, but the association with maltreatment was more often positive. Unfortunately, direct comparison among studies is not straightforward because of timing differences in the measurement of DNA methylation, as well as the inherently unclear timing of often retrospectively reported stressors[32]. In the current study, effect estimates for the top ranking CpGs were incongruent and even seemed slightly oppositional between the two cohorts. One explanation might be the longer time period between bullying exposure and the DNA methylation measurement in ALSPAC. One study on timing differences in ALSPAC for example, found that recency of adversity exposure was more important in explaining DNA methylation levels than accumulation of adversity, regardless of timing[32]. On the other hand, even among top ranking CpGs associations were weak. Such effect sizes are in line however, with other epigenome-wide studies in population-based samples[13, 15, 47-49], where exposures are generally less extreme and abundant than in risk samples that typically encounter larger effect sizes[11, 29]. In any case, thorough knowledge of normative development of DNA methylation levels is currently lacking and needed to interpret dissimilar estimates in the face of different measurement periods. Regarding the interpretation of the top hit, we further highlight that while stringent significance thresholds were used to reduce the risk of false positives, our current results may still reflect a chance finding and will need to be replicated in future studies.

To facilitate harmonization of the bullying exposure measurements in both cohorts, bullying exposure was defined with a lenient threshold. This implies that the difference in DNA methylation found for the CpG in *RAB14* is associated with exposure to bullying that is prevalent for children in the normal population. A more stringent definition of bullying might have brought forward different results, but a larger sample would be preferential for such an analysis. Additionally, with the current design we were unable to control for bullying exposure

that participants might have been subjected to outside of the moment of measurement. More measurements of bullying exposure would likely lead to more precise estimates. Further, more questions on the different types of bullying in the Generation R Study would have permitted us to differentiate between specific bullying exposures. Multiple reporters of bullying would have been preferable as well, especially the current use of mother report in one cohort and child report in the other is suboptimal. For the *RAB14* CpG site, there was converging agreement however. Another constraint of the study was that the current selected samples were more affluent than the fuller populations of their respective cohorts, where ideally the full spectrum of characteristics for the children in our cohorts would be represented. Last, we do not know if changes in DNA methylation are the consequence of bullying exposure, or that such changes are associated with children who are more at risk of being bullied[50]. An experimental set-up, for example with an anti-bullying intervention[51], would shed more light on this.

In conclusion, the current study is the first to report an epigenome-wide hit related to bullying exposure. This CpG site is located in the *RAB14* gene and suggests that exposure bullying might be associated with Golgi apparatus functioning. The effect size was small, but in line with other population-based studies. Further, we found an enrichment for CpGs related to cardiac functioning and neurodevelopment, as well as for CpGs with low levels of methylation and sites for which DNA methylation decreased in exposed but increased in non-exposed. We believe that experimental and longitudinal research into DNA methylation is the path to a broader understanding of social stress and its effect on biological pathways.

Methods

Setting

Data were drawn from two population-based prospective birth cohorts, the Dutch Generation R Study (Generation R) and the British Avon Longitudinal Study of Parents and Children (ALSPAC). Pregnant women residing in the municipality of Rotterdam, the Netherlands, with an expected delivery date between April 2002 and January 2006 were invited to enroll in the Generation R Study. A more extensive description of the study can be found elsewhere[52]. The Generation R Study is conducted in accordance with the World Medical Association Declaration of Helsinki and has been approved by the Medical Ethics Committee of the Erasmus Medical Center, Rotterdam. Written informed consent was obtained for all participants.

Pregnant women residing in the study area of former county Avon, United Kingdom, with an expected delivery date between April 1991 and December 1992 were invited to enroll in the ALSPAC study. Detailed information on the study design has been published previously[53, 54]. The ALSPAC website contains details of all available data through a fully searchable data

dictionary and variable search tool (http://www.bristol.ac.uk/alspac/researchers/our-data/). Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Consent for biological samples has been collected in accordance with the Human Tissue Act (2004). Informed consent for the use of data collected via questionnaires and clinics was obtained from participants following the recommendations of the ALSPAC Ethics and Law Committee at the time.

In both cohorts DNA methylation was studied before and after reported bullying exposure. A timeline can be found in Supplemental Figure 1.

Study Population

In the Generation R Study, 9,778 pregnant mothers gave birth to 9,749 live-born children. For a subsample of 608 singletons DNA methylation data was collected at 6 and/or 10 years old (343 at both time points). Of these, 506 children had information available on bullying exposure and relevant covariates, including 289 children with DNA methylation available for both time points (Supplemental Figure 2a). This subsample consisted of participants with parents born in the Netherlands, with European ancestry confirmed based on genetic principle component analysis for all children with genetic data available (99.6% of the current sample).

In ALSPAC, the inclusion of 14,541 pregnant mothers resulted in 14,062 live births. DNA methylation was available at 7 and/or 17 years old for a subsample of 936 European singletons (877 at both time points) as part of the Accessible Resource for Integrated Epigenomic Studies (ARIES) study[55]. For 846 of these children data on bullying exposure and relevant covariates was available, including 793 children with DNA methylation data at both time points (Supplemental Figure 2b), leading to a combined sample size of 1,352 children in the meta-analysis. In each cohort, bullying exposure and covariates were compared between the selected sample and (i) a set of participants with complete data on covariates, irrespective of availability of data on bullying exposure or DNA methylation (n = 8,528 in Generation R and n = 12,393 in ALSPAC), and (ii) a set of participants with complete data on both covariates and bullying exposure, irrespective of availability of DNA methylation data (n = 4,336 in Generation R and n = 6,347 in ALSPAC).

Bullying exposure

In Generation R, mothers filled out a questionnaire (adapted[56]) containing three questions on bullying exposure in the past few months, covering physical ('In the past few months, how often has your child been bullied by way of spitting, hitting, kicking, or pinching?'), verbal ('In the past few months, how often has your child been bullied by insulting, calling names or laughed at?'), and relational bullying ('In the past few months, how often has your child been bullied by being excluded from activities, ignored by other children, or gossiped about?'). Items were rated on a 5-point scale (ranging from *never* to *several times a week*). In

ALSPAC, bullying exposure was measured through self-report with an adapted version of the Bullying and Friendship Interview Schedule (BFIS)[57]. Nine questions covered physical (being hit or beaten up/belongings taken), verbal (threatened or blackmailed/tricked/called nasty names), and relational forms of bullying (others would not play with them/being made to do things they did not want to do/others told lies or nasty things about them/had games spoilt) in the preceding six months on a 4-point scale (ranging from *never* to *at least once a week*). Internal reliability of both measures was acceptable (Generation R: $\alpha = 0.74$, ALSPAC: $\alpha = 0.73$). Scores were dichotomized to harmonize the two bullying scales and avoid issues arising from extreme skewness of the data. Children were classified as being exposed to bullying if they were bullied '*at least once or twice in the past few months*' on at least one of the items in Generation R, and at least '*1-3 times in the past six months*' in ALSPAC[58-61].

Variables sensitivity analyses

A more *stringent bullying exposure* variable was defined, in which children were classified as exposed when at least indicated to be bullied '*2 or 3 times a month'* in Generation R, and they at least indicated to be bullied '*about once a week'* in ALSPAC.

Body Mass Index (BMI) (kg/m^2) was measured at 6 and 10 years in the Generation R and 7 and 17 years in the ALSPAC. Values were standardized to *SD* scores, adjusted for age and sex.

Child behavioral problems were measured at age 3 years with the mother-reported Child Behavior Checklist for toddlers (CBCL1½-5)[62] in Generation R. Ninety-nine items were scored one a 3-point scale (range 0-2), regarding symptoms of anxiety, sadness, withdrawn behavior, attention problems, and aggressive behaviors ($\alpha = 0.92$). Items were summed into a weighed total problem behavior scale, with 25% missing allowed. In ALSPAC, the mother-reported Strengths and Difficulties Questionnaire (SDQ)[63] at 4 years was used. The scales for emotional, conduct, and hyperactivity problems were used as a total problem behavior score, consisting of 15 items, each rated on a 3-point scale (range 0-2) ($\alpha = 0.74$). The remaining problem scale of the SDQ, the 'peer problems' scale, was excluded from the total score due to content overlap with bullying exposure.

Child non-verbal intelligence quotient (IQ) was measured by testing visuospatial abilities (Mosaics) and abstract reasoning (Categories) with the Snijders-Oomen Niet-verbale Intelligentie Test-Revisie (SON-R 2½-7)[64] at age 6 years in Generation R. In ALSPAC, a shortened version of the Wechsler Intelligence Scale for Children (3rd UK edition (WISC-III))[65] was measured at age 9 years.

O*ther stressful experiences* were measured in Generation R with a major life events inventory[66], reported by the mother, when the child was 10 years. This inventory covers stressful life events spanning the lifetime of the child, such as physical abuse, sexual abuse, conflict in

the household, illness or death in the family, and parental separation. Three items related to bullying exposure were excluded, leaving 21 items (range 0-1). In ALSPAC, we used an Adverse Child Experiences (ACE) lifetime composite score[15, 67]. This score is based on 541 questions mapping on to 10 ACEs up to age 16 years. Participants were included if there was at least 50% of the data available for each ACE. We excluded the ACE for bullying, leaving physical, sexual, and emotional abuse, emotional neglect, substance use in the household, violence between parents, parental mental health, parent conflict, parent offence, and parental separation (each range 0-1).

Last, *alcohol use* was measured in ALSPAC with Alcohol Use Disorders Identification Test (AUDIT) at age 17. This tests consists of 10 items (range 0-4); a total score of 8 or more is considered hazardous[68].

DNA methylation

Both cohorts used the EZ-96 DNA Methylation kit (Shallow) (Zymo Research Corporation, Irvine, USA) for bisulfite conversion on the extracted DNA. DNA methylation profiles were generated using the Illumina Infinium HumanMethylation450 BeadChip (Illumina Inc., San Diego, USA). Quality control and normalization steps can be found in Supplemental Methods. Analyses were restricted to 473,864 autosomal CpGs. DNA methylation levels are characterized by beta values (β values), representing the ratio of methylated signal relative to the sum of methylated and unmethylated signal measured per CpG. Outlying data points outside the 3*interquartile range were winsorized to the nearest point for each CpG. White blood cell (WBC) composition was estimated using the reference-based Houseman method[69]. Batch effects and additional unknown confounding were estimated using surrogate variable analysis (SVA) in *meffil*[70, 71] in R version 3.4.3[72].

Statistical analyses

Associations between bullying exposure and changes in DNA methylation were analyzed with a linear mixed model:

$$M_{ij} = \beta_0 + u_{0i} + \beta_1 Age_{ij} + \beta_2 Bullied_{ij} + \text{covariates} + \epsilon_{ij}$$

Here, M denotes DNA methylation level, β_0 fixed intercept, u_{0i} random intercept, β_1 fixed age coefficient, β_2 fixed bullying exposure coefficient, and ϵ random error. The *Bullied* variable was set to 0 for the first DNA methylation measurement and to 1 or 0 for the second measurement depending on whether the participant had been exposed to bullying, or not, respectively, and random intercept u_{0i} allowed for inter-individual variation in DNA methylation at the first measurement. Participants are denoted by i and time points by j. Covariates included sex, gestational age, socio-economic status as indicated by highest attained educational level of the mother (low versus medium or high), surrogate variables (n = 20), WBCs (CD4+

T-lymphocytes, CD8+ T-lymphocytes, natural killer cells, B-lymphocytes, monocytes, and granulocytes). Current direct and second-hand smoking was adjusted for with the methylation level of *AHRR* cg05575921, which has proven to be a valid marker of tobacco exposure[73-76]. Methylation level for this CpG at both time points was entered into the equation, with levels divided into quintiles (as described elsewhere[77]) and lower levels indicating more smoking. Linear mixed models were applied using the *lme4* package[78].

To compare congruency between results from the two cohorts, estimates of the top 1000 CpGs in each cohort were correlated with estimates of those CpGs in the other cohort (as elsewhere[15]). Meta-analysis of estimates and standard errors of the two cohorts was performed using fixed models within the *metafor*[79] R package. To account for multiple testing (n = 473,864 CpGs), the significance threshold was set at a Bonferroni-corrected p-value of 1.06×10^{-07}.

Follow-up analyses

Sensitivity analyses

A series of sensitivity analyses were performed on CpGs with $p < 1.06 \times 10^{-07}$ in the meta-analysis. First, because the classification of bullying exposure is rather broad in the main analysis, we performed a sensitivity analysis with a more stringent dichotomization. Second, we reran analyses additionally adjusting for the following potential confounders or mediators that have been previously shown to associate with bullying exposure and DNA methylation. In each sensitivity analysis, one of the following variables was added to the main analysis: (i) Body Mass Index (BMI) (kg/m^2) *SD* scores, adjusted for age and sex, at 6 and 10 years in Generation R and 7 and 17 years in ALSPAC (full sample size available)[60, 80]; (ii) pre-existing behavioral problems (Generation R n = 451, ALSPAC n = 794)[5, 50]; (iii) child non-verbal IQ (Generation R n = 465, ALSPAC n = 811)[81]; (iv) stressful experiences other than bullying exposure (Generation R n = 482, ALSPAC n = 597)[82, 83]; and (v) alcohol use in ALSPAC (n = 624), where children are older[84, 85]. For each sensitivity analysis, the coefficient for bullying exposure was compared with that for the main analysis with a z-test[86, 87].

Genetic associations

DNA methylation for CpGs with $p < 1.06 \times 10^{-07}$ in the meta-analysis were tested for genetic associations in three ways. First, a look-up was performed in a list of CpGs located on a single nucleotide polymorphism (SNP), e.g. polymorphic CpGs[23]. Second, we tested for known associations with genetic variants, e.g. methylation quantitative trait loci, *in cis* (*cis* mQTLs) and *in trans* (*trans* mQTLs; http://www.mqtldb.org/; GCTA set)[24]. Third, we tested for additive genetic influences versus shared and unique environmental influences on the DNA methylation, as based on twin heritability analyses[25].

Look-up of previous findings in the literature

Previous studies on childhood adversity and DNA methylation, measured with the Illumina Infinium HumanMethylation450 BeadChip, were searched for current CpGs with $p < 1.06x10^{-07}$ and their associated genes in the meta-analysis. Eleven studies were selected[11, 13, 15, 16, 26-32]. All studies examined childhood abuse or trauma, on study additionally examined bullying[13].

Candidate gene-wide analyses

A candidate gene follow-up analysis was conducted on the results stemming from the meta-analysis, for sites annotated to *5-HTT* and *NR3C1*. The significance threshold was set at a Bonferroni gene-level corrected *p*-value of $3.13x10^{-03}$ (*n* = 16 CpGs) and $1.22x10^{-03}$ (*n* = 41 CpGs), respectively.

Functional associations

Enrichment of Gene Ontology (GO) pathways was tested for genes associated with CpGs with $p < 0.001$ in the meta-analysis (cutoff described elsewhere [11]), while adjusting for gene size and pruning for redundant terms (a full method description can be found elsewhere [11]). Terms with $p < 0.05$ and more than one associated genes are reported, and highlighted if confirmed by near-identical terms from GO analyses with CpGs $p < 0.01$ and $p < 0.001$ in the meta-analysis.

References

1. Craig W, Harel-Fisch Y, Fogel-Grinvald H, et al. A cross-national profile of bullying and victimization among adolescents in 40 countries. *International Journal of Public Health.* 2009;54:216-24.
2. Olweus D. Bullying at school. Aggressive behavior: Springer, 1994:97-130.
3. Ringoot AP, Jansen PW, Rijlaarsdam J, et al. Self-reported problem behavior in young children with and without a DSM-disorder in the general population. *European Psychiatry.* 2017;40:110-5.
4. Schoeler T, Duncan L, Cecil CM, Ploubidis GB, Pingault J-B. Quasi-experimental evidence on short- and long-term consequences of bullying victimization: A meta-analysis. *Psychological Bulletin.* 2018;144:1229.
5. Zwierzynska K, Wolke D, Lereya TS. Peer victimization in childhood and internalizing problems in adolescence: a prospective longitudinal study. *Journal of Abnormal Child Psychology.* 2013;41:309-23.
6. Copeland WE, Wolke D, Angold A, Costello EJ. Adult psychiatric outcomes of bullying and being bullied by peers in childhood and adolescence. *JAMA Psychiatry.* 2013;70:419-26.
7. Winsper C, Lereya T, Zanarini M, Wolke D. Involvement in bullying and suicide-related behavior at 11 years: a prospective birth cohort study. *Journal of the American Academy of Child & Adolescent Psychiatry.* 2012;51:271-82. e3.
8. Suderman M, Borghol N, Pappas JJ, et al. Childhood abuse is associated with methylation of multiple loci in adult DNA. *BMC Medical Genomics.* 2014;7:13.
9. Heijmans BT, Tobi EW, Stein AD, et al. Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proceedings of the National Academy of Sciences.* 2008;105:17046-9.
10. Rijlaarsdam J, Pappa I, Walton E, et al. An epigenome-wide association meta-analysis of prenatal maternal stress in neonates: A model approach for replication. *Epigenetics.* 2016;11:140-9.
11. Cecil CAM, Smith RG, Walton E, Mill J, McCrory EJ, Viding E. Epigenetic signatures of childhood abuse and neglect: implications for psychiatric vulnerability. *Journal of Psychiatric Research.* 2016;83:184-94.
12. Rusiecki JA, Byrne C, Galdzicki Z, et al. PTSD and DNA methylation in select immune function gene promoter regions: A repeated measures case-control study of U.S. military service members. *Frontiers in Psychiatry.* 2013;4:1-12.
13. Marzi SJ, Sugden K, Arseneault L, et al. Analysis of DNA methylation in young people: limited evidence for an association between victimization stress and epigenetic variation in blood. *American Journal of Psychiatry.* 2018;175:517-29.
14. Labonté B, Suderman M, Maussion G, et al. Genome-wide epigenetic regulation by early-life trauma. *Archives General Psychiatry.* 2012;69:722-31.
15. Houtepen LC, Hardy R, Maddock J, et al. Childhood adversity and DNA methylation in two population-based cohorts. *Translational Psychiatry.* 2018;8:266.
16. Houtepen LC, Vinkers CH, Carrillo-Roa T, et al. Genome-wide DNA methylation levels and altered cortisol stress reactivity following childhood trauma in humans. *Nature Communications.* 2016;7.

17. Jovanova OS, Nedeljkovic I, Spieler D, et al. DNA methylation signatures of depressive symptoms in Middle-aged and elderly persons: meta-analysis of multiethnic Epigenome-wide studies. *JAMA Psychiatry.* 2018;75:949-59.
18. Mulder RH, Rijlaarsdam J, Luijk MPCM, et al. Methylation matters: FK506 binding protein 51 (*FKBP5*) methylation moderates the associations of *FKBP5* genotype and resistant attachment with stress regulation. *Development and Psychopathology.* 2017;29:491-503.
19. van Dongen J, Zilhão NR, Sugden K, et al. Epigenome-wide Association Study of Attention-Deficit/Hyperactivity Disorder Symptoms in Adults. *Biological Psychiatry.* 2019.
20. Ouellet-Morin I, Wong CC, Danese A, et al. Increased serotonin transporter gene (SERT) DNA methylation is associated with bullying victimization and blunted cortisol response to stress in childhood: a longitudinal study of discordant monozygotic twins. *Psychological Medicine.* 2013;43:1813-23.
21. Efstathopoulos P, Andersson F, Melas PA, et al. NR3C1 hypermethylation in depressed and bullied adolescents. *Translational Psychiatry.* 2018;8:121.
22. Bjornsson HT, Sigurdsson MI, Fallin MD, et al. Intra-individual change over time in DNA methylation with familial clustering. *Jama.* 2008;299:2877-83.
23. Chen Y-a, Lemire M, Choufani S, et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics.* 2013;8:203-9.
24. Gaunt TR, Shihab HA, Hemani G, et al. Systematic identification of genetic influences on methylation across the human life course. *Genome Biology.* 2016;17:61.
25. Hannon E, Knox O, Sugden K, et al. Characterizing genetic and environmental influences on variable DNA methylation using monozygotic and dizygotic twins. *PLoS Genetics.* 2018;14:e1007544.
26. Mehta D, Klengel T, Conneely KN, et al. Childhood maltreatment is associated with distinct genomic and epigenetic profiles in posttraumatic stress disorder. *Proceedings of the National Academy of Science.* 2013;110:8302-7.
27. Prados J, Stenz L, Courtet P, et al. Borderline personality disorder and childhood maltreatment: A genome-wide methylation analysis. *Genes, Brain and Behavior.* 2015;14:177-88.
28. Suderman M, Pappas JJ, Borghol N, et al. Lymphoblastoid cell lines reveal associations of adult DNA methylation with childhood and current adversity that are distinct from whole blood associations. *International Journal of Epidemiology.* 2015;44:1331-40.
29. Yang BZ, Zhang H, Ge W, et al. Child abuse and epigenetic mechanisms of disease risk. *American Journal of Preventive Medicine.* 2013;44:101-7.
30. Kumsta R, Kreppner J, Kennedy M, Knights N, Rutter M, Sonuga-Barke E. Psychological Consequences of Early Global Deprivation An Overview of Findings From the English & Romanian Adoptees Study. *European Psychologist.* 2015;20:138-51.
31. Marinova Z, Maercker A, Küffer A, et al. DNA methylation profiles of elderly individuals subjected to indentured childhood labor and trauma. *BMC Medical Genetics.* 2017;18:21.
32. Dunn EC, Soare TW, Zhu Y, et al. Sensitive periods for the effect of childhood adversity on DNA methylation: Results from a prospective, longitudinal study. *Biological Psychiatry.* 2019;85:838-49.

33. Lanner JT, Georgiou DK, Joshi AD, Hamilton SL. Ryanodine receptors: structure, expression, molecular details, and function in calcium release. *Cold Spring Harbor Perspectives in Biology.* 2010;2:a003996.
34. Junutula JR, De Maziére AM, Peden AA, et al. Rab14 is involved in membrane trafficking between the Golgi complex and endosomes. *Molecular Biology of the Cell.* 2004;15:2218-29.
35. Biala YaN, Bogoch Y, Bejar C, Linial M, Weinstock M. Prenatal stress diminishes gender differences in behavior and in expression of hippocampal synaptic genes and proteins in rats. *Hippocampus.* 2011;21:1114-25.
36. Zhou J, Liu Z, Yu J, et al. Quantitative proteomic analysis reveals molecular adaptations in the hippocampal synaptic active zone of chronic mild stress-unsusceptible rats. *International Journal of Neuropsychopharmacology.* 2015;19:pyv100.
37. Ornish D, Magbanua MJM, Weidner G, et al. Changes in prostate gene expression in men undergoing an intensive nutrition and lifestyle intervention. *Proceedings of the National Academy of Sciences.* 2008;105:8369-74.
38. Sequeira A, Mamdani F, Ernst C, et al. Global brain gene expression analysis links glutamatergic and GABAergic alterations to suicide and major depression. *PLoS ONE.* 2009;4:e6585.
39. Das S, Boczan J, Gerwin C, Zald PB, Sheng Z-H. Regional and developmental regulation of syntaphilin expression in the brain: a candidate molecular element of synaptic functional differentiation. *Molecular Brain Research.* 2003;116:38-49.
40. Shan J, Kushnir A, Betzenhauser MJ, et al. Phosphorylation of the ryanodine receptor mediates the cardiac fight or flight response in mice. *The Journal of Clinical Investigation.* 2010;120:4388-98.
41. Oltmann-Norden I, Galuska SP, Hildebrandt H, et al. Impact of the polysialyltransferases ST8SiaII and ST8SiaIV on polysialic acid synthesis during postnatal mouse brain development. *Journal of Biological Chemistry.* 2008;283:1463-71.
42. Calandreau L, Márquez C, Bisaz R, Fantin M, Sandi C. Differential impact of polysialyltransferase ST8SiaII and ST8SiaIV knockout on social interaction and aggression. *Genes, Brain and Behavior.* 2010;9:958-67.
43. Park C, Lee S-A, Hong J-H, et al. Disrupted-in-schizophrenia 1 (DISC1) and Syntaphilin collaborate to modulate axonal mitochondrial anchoring. *Molecular Brain.* 2016;9:69.
44. Sato C, Kitajima K. Impact of structural aberrancy of polysialic acid and its synthetic enzyme ST8SIA2 in schizophrenia. *Frontiers in Cellular Neuroscience.* 2013;7:61.
45. Lin M-Y, Cheng X-T, Tammineni P, et al. Releasing syntaphilin removes stressed mitochondria from axons independent of mitophagy under pathophysiological conditions. *Neuron.* 2017;94:595-610. e6.
46. Mikkonen M, Soininen H, Tapiola T, Alafuzoff I, Miettinen R. Hippocampal plasticity in Alzheimer's disease: changes in highly polysialylated NCAM immunoreactivity in the hippocampal formation. *European Journal of Neuroscience.* 1999;11:1754-64.
47. Panni T, Mehta AJ, Schwartz JD, et al. Genome-wide analysis of DNA methylation and fine particulate matter air pollution in three study populations: KORA F3, KORA F4, and the normative aging study. *Environmental Health Perspectives.* 2016;124:983-90.

48. Green BB, Karagas MR, Punshon T, et al. Epigenome-wide assessment of DNA methylation in the placenta and arsenic exposure in the New Hampshire Birth Cohort Study (USA). *Environmental Health Perspectives.* 2016;124:1253-60.
49. Sharp GC, Arathimos R, Reese SE, et al. Maternal alcohol consumption and offspring DNA methylation: findings from six general population-based birth cohorts. *Epigenomics.* 2018;10:27-42.
50. Singham T, Viding E, Schoeler T, et al. Concurrent and longitudinal contribution of exposure to bullying in childhood to mental health: the role of vulnerability and resilience. *JAMA Psychiatry.* 2017;74:1112-9.
51. Waasdorp TE, Bradshaw CP, Leaf PJ. The impact of schoolwide positive behavioral interventions and supports on bullying and peer rejection: A randomized controlled effectiveness trial. *Archives of Pediatrics & Adolescent Medicine.* 2012;166:149-56.
52. Kooijman MN, Kruithof CJ, van Duijn CM, et al. The Generation R Study: design and cohort update 2017. *European Journal of Epidemiology.* 2016;31:1243-64.
53. Boyd A, Golding J, Macleod J, et al. Cohort profile: the 'children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology.* 2013;42:111-27.
54. Fraser A, Macdonald-Wallis C, Tilling K, et al. Cohort profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *International Journal of Epidemiology.* 2012;42:97-110.
55. Relton CL, Gaunt T, McArdle W, et al. Data resource profile: accessible resource for integrated epigenomic studies (ARIES). *International Journal of Epidemiology.* 2015;44:1181-90.
56. Perren S, Alsaker FD. Social behavior and peer relationships of victims, bully-victims, and bullies in kindergarten. *Journal of Child Psychology and Psychiatry.* 2006;47:45-57.
57. Wolke D, Woods S, Stanford K, Schulz H. Bullying and victimization of primary school children in England and Germany: Prevalence and school factors. *British Journal of Psychology.* 2001;92:673-96.
58. Horwood J, Waylen A, Herrick D, Williams C, Wolke D. Common visual defects and peer victimization in children. *Investigative Ophthalmology & Visual Science.* 2005;46:1177-81.
59. Paget A, Parker C, Heron J, et al. Which children and young people are excluded from school? Findings from a large B ritish birth cohort study, the Avon Longitudinal S tudy of P arents and C hildren (ALSPAC). *Child: Care, Health and Development.* 2018;44:285-96.
60. Jansen PW, Verlinden M, Dommisse-van Berkel A, et al. Teacher and peer reports of overweight and bullying among young primary school children. *Pediatrics.* 2014:peds. 2013-3274.
61. Verlinden M, Jansen PW, Veenstra R, et al. Preschool attention-deficit/hyperactivity and oppositional defiant problems as antecedents of school bullying. *Journal of the American Academy of Child & Adolescent Psychiatry.* 2015;54:571-9.
62. Achenbach TM, Rescorla LA. *Manual for the ASEBA preschool forms and profiles*: Burlington, VT: University of Vermont, Research center for children, youth , 2000.
63. Goodman R. The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry.* 1997;38:581-6.

64. Tellegen PJ, Winkel M, Wijnbergen-Williams BJ. *Snijders-Oomen Niet-verbale Intelligentietest-Revisie SON-R 2½–7*. In: Lisse: Swets & Zeitlinger, 1996.
65. Wechsler D. Wechsler Intelligence Scale for Children–3rd UK ed (WISC–III UK). *Sidcup, UK: The Psychological Corporation.* 1992.
66. Dunn EC, Nishimi K, Neumann A, et al. Time-dependent effects of exposure to physical and sexual violence on psychopathology symptoms in late childhood: in search of sensitive periods in development. *Journal of the American Academy of Child & Adolescent Psychiatry.* 2019.
67. Houtepen LC, Heron J, Suderman MJ, Tilling K, Howe LD. Adverse childhood experiences in the children of the Avon Longitudinal Study of Parents and Children (ALSPAC). *Wellcome Open Research*. 2018;3.
68. Saunders JB, Aasland OG, Babor TF, De la Fuente JR, Grant M. Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction.* 1993;88:791-804.
69. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics.* 2012;13:1.
70. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics.* 2012;28:882-3.
71. Leek JT, Storey JD. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genetics.* 2007;3:e161.
72. R Core Team. *R: A language and environment for statistical computing*. 2013.
73. Zeilinger S, Kühnel B, Klopp N, et al. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS ONE.* 2013;8:e63812.
74. Philibert R, Hollenbeck N, Andersen E, et al. Reversion of AHRR demethylation is a quantitative biomarker of smoking cessation. *Frontiers in Psychiatry.* 2016;7:55.
75. Reynolds LM, Magid HS, Chi GC, et al. Secondhand tobacco smoke exposure associations with DNA methylation of the aryl hydrocarbon receptor repressor. *Nicotine & Tobacco Research.* 2016;19:442-51.
76. Dogan MV, Lei M-K, Beach SRH, Brody GH, Philibert RA. Alcohol and tobacco consumption alter hypothalamic pituitary adrenal axis DNA methylation. *Psychoneuroendocrinology.* 2016;66:176-84.
77. Bojesen SE, Timpson N, Relton C, Smith GD, Nordestgaard BG. AHRR (cg05575921) hypomethylation marks smoking behaviour, morbidity and mortality. *Thorax.* 2017;72:646-53.
78. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *arXiv preprint* arXiv:1406.5823 2014.
79. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software.* 2010;36.
80. Fradin D, Boëlle P-Y, Belot M-P, et al. Genome-wide methylation analysis identifies specific epigenetic marks in severely obese children. *Scientific Reports.* 2017;7:46311.
81. Verlinden M, Veenstra R, Ghassabian A, et al. Executive functioning and non-verbal intelligence as predictors of bullying in early elementary school. *Journal of Abnormal Child Psychology.* 2014;42:953-66.

82. Jansen PW, Verlinden M, Dommisse-van Berkel A, et al. Prevalence of bullying and victimization among children in early elementary school: Do family and school neighbourhood socioeconomic status matter? *BMC Public Health.* 2012;12:494.
83. Holt MK, Kaufman Kantor G, Finkelhor D. Parent/child concordance about bullying involvement and family characteristics related to bullying and peer victimization. *Journal of School Violence.* 2008;8:42-63.
84. Rospenda KM, Richman JA, Wolff JM, Burke LA. Bullying victimization among college students: Negative consequences for alcohol use. *Journal of Addictive Diseases.* 2013;32:325-42.
85. Zhang R, Miao Q, Wang C, et al. Genome-wide DNA methylation analysis in alcohol dependence. *Addiction Biology.* 2013;18:392-403.
86. Paternoster R, Brame R, Mazerolle P, Piquero A. Using the correct statistical test for the equality of regression coefficients. *Criminology.* 1998;36:859-66.
87. Clogg CC, Petkova E, Haritou A. Statistical methods for comparing regression coefficients between models. *American Journal of Sociology.* 1995;100:1261-93.

Supplemental Methods

Quality control and normalization of DNA methylation

In Generation R, quality control was performed using the CPACOR workflow[1] on all 2,467 available DNA methylation samples, including cord blood samples (1,475 cord bloods, 500 peripheral bloods at 6 years, and 492 peripheral bloods at 10 years). Arrays with observed technical problems such as failed bisulfite conversion, hybridization or extension, as well as arrays with a mismatch between sex of the proband and sex determined by the chr X and Y probe intensities were removed from subsequent analyses. Additionally, only arrays with a call rate>95% per sample were processed further, resulting in 2,355 samples, 44 of which were sibling pairs. Hence 2,333 were carried forward into normalization.

In ALSPAC, quality control was performed on 6,057 samples (including 1,127 from cord blood, 1086 from peripheral blood at 7 years, and 1,073 from peripheral blood at 17 years from ALSPAC children and 2,771 peripheral bloods from their parents),using the *meffil* package[2]. After removing samples with mismatched genotypes, mismatched gender, incorrect relatedness, low concordance with samples collected at other time points, extreme dye bias, and poor probe detection, 5,337 samples remained, including 2,845 samples for 1,003 children.

To minimize cohort effects as much as possible, we normalized both cohorts together as a single dataset. In detail, functional normalization (10 control probe principal components, slide included as a random effect) was performed in R version 3.4.3[3] with the *meffil* package on a combined Generation R and ALSPAC set including cord and whole blood samples comprising a total of 5,178 samples for a total of 485,512 CpG sites.

Supplemental Analysis

Associations per time-point

The association between bullying exposure and DNA methylation at cg17312179 was analyzed separately for each time-point (pre- (T1) and post-measurement (T2) of bullying exposure), to understand if reported change stemmed from a larger difference between exposed and non-exposed children after bullying report, rather than beforehand. Associations between bullying exposure at 6 years in Generation R and at 7 years in ALSPAC (combined n = 1,224) were meta-analyzed, as were those at 10 years in Generation R and 17 years in ALSPAC (combined n = 1,210). Results showed that the association with DNA methylation change at cg17312179 was driven by a difference in exposed versus non-exposed individuals at T2 (b = -2.43x10^{-03}, SE = 5.20x10^{-04}, p = 3.04x10^{-06}), not at T1 (b = 7.24x10^{-04}, SE = 4.86x10^{-04}, p = 1.37x10^{-01}). This was confirmed by a z-test[4, 5] between the coefficients for the two time-points (p = 9.64x10^{-06}).

Supplemental Tables

Supplemental Table 1. Missingness analysis: Comparison of sample characteristics across different selected sample sets in Generation R and ALSPAC (part I)

		Generation R					ALSPAC				
		Set 1	Set 2	Set 3			Set 1	Set 2	Set 3		
		Participants with complete data on covariates[a]	Participants in Set 1 and complete data on bullying exposure	Participants in Set 2 and complete data on DNA methylation	Set 1 versus Set 3 *p*-value[b]	Set 2 versus Set 3 *p*-value[b]	Participants with complete data on covariates[a]	Participants in Set 1 and complete data on bullying exposure	Participants in Set 2 and complete data on DNA methylation	Set 1 versus Set 3 *p*-value[b]	Set 2 versus Set 3 *p*-value[b]
main analysis	*n*	8528	4336	506			12393	6347	846		
	Sex (No. (%) boys)	4315 (50.6)	2159 (49.8)	251 (49.6)	6.98e-01	9.75e-01	5996 (48.4)	3189 (50.2)	439 (51.9)	5.24e-02	3.88e-01
	Gestational age in weeks (mean (*SD*))	39.7 (2.0)	39.8 (1.8)	40.2 (1.4)	**3.68e-12**	**3.95e-08**	39.4 (1.9)	39.5 (1.9)	39.6 (1.5)	**5.66e-03**	**1.15e-02**
	Maternal education (No. (%))				**5.51e-36**	**8.87e-09**				**3.00e-27**	**2.15e-06**
	Low	2254 (26.4)	644 (14.9)	37 (7.3)			2501 (20.2)	815 (12.8)	71 (8.4)		
	Medium	2614 (30.7)	1205 (27.8)	112 (22.1)			5509 (44.5)	2751 (43.3)	334 (39.5)		
	High	3660 (42.9)	2487 (57.4)	357 (70.6)			4383 (35.4)	2781 (43.8)	441 (52.1)		
	Maternal age at delivery (mean (*SD*))	30.48 (5.26)	31.97 (4.52)	32.80 (3.92)	**1.12e-32**	**1.17e-05**	28.27 (4.86)	29.23 (4.51)	29.69 (4.35)	**8.45e-19**	**4.69e-03**
	Age in years bullying exposure report (mean (*SD*))	-	8.07 (0.12)	8.16 (0.23)	-	**1.15e-37**	-	8.62 (0.27)	8.56 (0.19)	-	**7.99e-14**
	Bullying exposure (No. (%) yes)	-	1828 (42.2)	229 (45.3)	-	1.98e-01	-	2500 (39.4)	333 (39.4)	-	1.00

Supplemental Table 1. Missingness analysis: Comparison of sample characteristics across different selected sample sets in Generation R and ALSPAC (part II)

		Generation R					ALSPAC				
		Set 1	Set 2	Set 3			Set 1	Set 2	Set 3		
		Participants with complete data on covariates[a]	Participants in Set 1 and complete data on bullying exposure	Participants in Set 2 and complete data on DNA methylation	Set 1 versus Set 3 *p*-value[b]	Set 2 versus Set 3 *p*-value[b]	Participants with complete data on covariates[a]	Participants in Set 1 and complete data on bullying exposure	Participants in Set 2 and complete data on DNA methylation	Set 1 versus Set 3 *p*-value[b]	Set 2 versus Set 3 *p*-value[b]
sensitivity analyses	*n*	-	4336	506			-	6347	846		
	Bullying exposure–sensitivity analysis (No. (%) yes)	-	462 (10.7)	50 (9.9)	-	6.46e-01	-	1293 (20.4)	174 (20.6)	-	9.31e-01
	n	6070	3926	404			7448	5802	820		
	BMI (kg/m^2) at T1 (mean (*SD*))	16.2 (1.9)	16.0 (1.7)	15.9 (1.3)	**1.11e-06**	**2.83e-02**	16.2 (2.0)	16.2 (2.0)	16.2 (2.0)	9.02e-01	6.67e-01
	n	5230	3664	391			4573	3654	819		
	BMI (kg/m^2) at T2 (mean (*SD*))	17.5 (2.7)	17.3 (2.5)	17.1 (2.0)	**9.60e-06**	5.52e-02	22.8 (4.2)	22.7 (4.1)	22.6 (3.6)	9.43e-02	3.77e-01
	n	4624	3399	451			9286	5695	794		
	Behavioral problem score (mean (*SD*))	20.7 (15.1)	19.8 (14.2)	17.3 (12.2)	**5.03e-08**	**7.18e-05**	6.9 (5.4)	6.8 (4.8)	6.9 (3.9)	7.80e-01	9.67e-01
	n	5558	3639	465			6339	5846	811		
	Nonverbal intelligence quotient (mean (*SD*))	101.3 (15.1)	103.5 (14.6)	107.3 (14.0)	**7.57e-18**	**4.60e-08**	100.0 (16.9)	100.3 (16.8)	102.6 (16.7)	**4.18e-05**	**2.62e-04**
	n	5045	3536	482			3646	3189	597		
	Other stressful life events score (mean (*SD*))	4.1 (2.4)	3.80 (2.2)	3.7 (2.1)	**1.15e-04**	1.63e-01	1.5 (1.4)	1.5 (1.4)	1.5 (1.4)	4.05e-01	7.44e-01
	n	-	-	-			3785	3105	624		
	Alcohol use (mean (*SD*))	-	-	-	-	-	8.0 (4.8)	8.1 (4.8)	8.1 (4.8)	4.78e-01	7.46e-01

SD: standard deviation; No.: number; T1: time point 1; T2: time point 2; BMI: Body Mass Index

[a]non-technical covariates, i.e.: child sex, gestational age, and education of the mother.

[b]*p*-value as based on T-test for numerical variables and X-square for categorical variables. Values in bold are $p < 0.05$.

Supplemental Table 2. Functions of genes associated with the ten CpG sites with the lowest *p*-values in the meta-analysis of epigenome-wide associations with bullying exposure

CpG site	Gene	Associated functions
cg17312179	*RAB14*	Golgi apparatus functioning, vesicle processing[6]
cg09291817	*MAZ*	Inflammation induced, regulates amyloid A protein[7]
cg11278602	*HCG4*	Unknown; located in major histocompatibility region, which is important for immune functioning[8]
cg00911813	*TNRC18*	Unknown; other trinucleotide repeat (TNR) genes associated to transcription regulation and neurological disorders[9]
cg08971637	*DGUOK*	Mitochondrial functioning[10]
cg12767834	*SNPH*	Synaptic vesicle processing[11]
cg26394220	*MIR375; CCDC108*	Micro RNA, pancreas functioning[12]
cg19790568	*PRX*	Peripheral myelin upkeep[13]
cg10929442	*ST8SIA4*	Golgi apparatus functioning, neural cell plasticity[14]
ch.4.134822993R		

Supplemental Table 3. Meta-analyzed results for cg17312179 from main analysis and sensitivity analyses

			Added variable in sensitivity analysis[b]		Bullying exposure[c]		*Bullying exposure:* main analysis vs sensitivity analysis[d]
		combined *n*[a]	*B* (SE)	*p*-value	*B* (SE)	*p*-value	*p*-value
Main analysis	Bullying exposure	1352	-	-	-2.67e-03 (4.97e-04)	7.17e-08	-
Sensitivity analyses	Bullying exposure (more stringent definition)	1352	-	-	-2.62e-03 (7.23e-04)	2.95e-04	9.46e-01
	BMI	1352	-3.34e-05 (2.03e-04)	8.69e-01	-2.76e-03 (5.11e-04)	6.62e-08	9.05e-01
	Behavioral problems	1245	-1.14e-05 (1.89e-05)	5.44e-01	-2.87e-03 (5.18e-04)	2.92e-08	7.85e-01
	Intelligence quotient	1276	-7.82e-06 (1.28e-05)	5.40e-01	-2.64e-03 (5.18e-04)	3.45e-07	9.61e-01
	Other stressful experiences	1079	2.20e-04 (9.96e-05)	2.73e-02	-2.51e-03 (5.39e-04)	3.18e-06	8.26e-01
	Alcohol use (only in ALSPAC)	1130	-1.09e-04 (7.33e-05)	1.38e-01	-2.91e-03 (5.23e-04)	2.56e-08	7.42e-01

SE: standard error; BMI: Body Mass Index

[a]Sample size of Generation R + ALSPAC

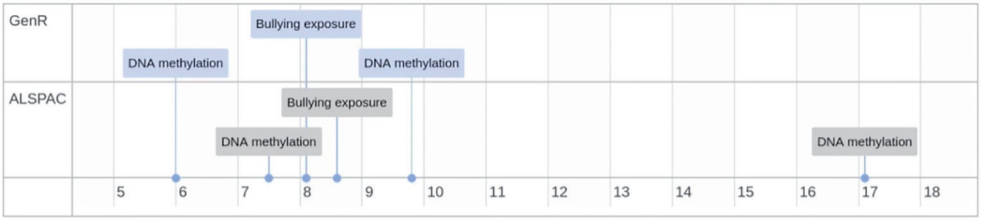
[b]Each variable was added separately to the main analysis, while more stringent definition of bullying exposure replaced the bullying exposure variable in the main analysis. This was done in each cohort and estimates are based on the meta-analysis, except for alcohol, which was only available to ALSPAC.

[c]Based on meta-analysis

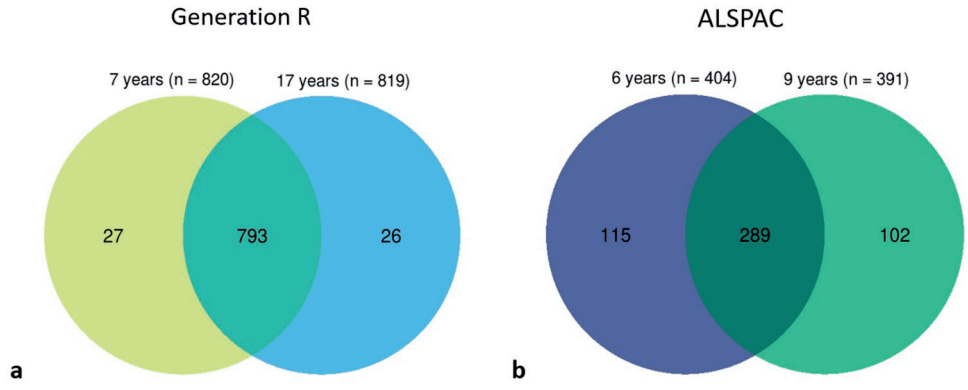
[d]Comparison performed with z-test on estimates of bullying exposure variable in main analysis versus bullying exposure variable in sensitivity analysis

Supplemental Table 4-6 are available online in the published article–https://www.tandfonline.com/doi/full/10.1080/15592294.2020.1719303

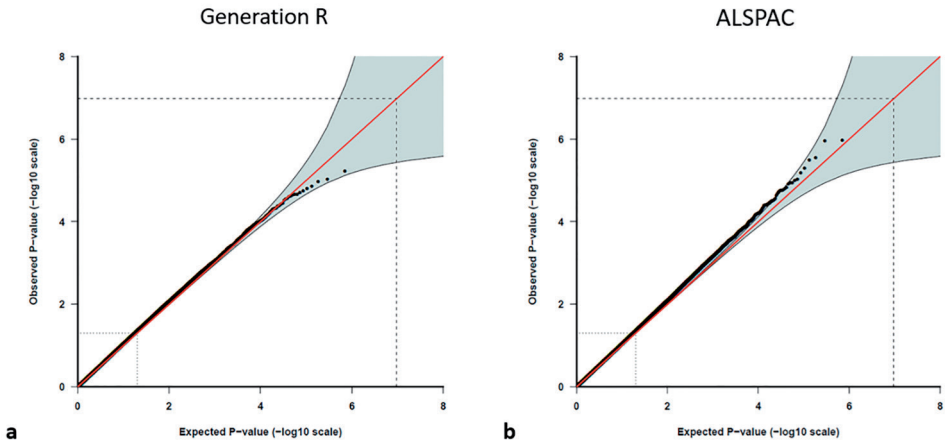
Supplemental Figures



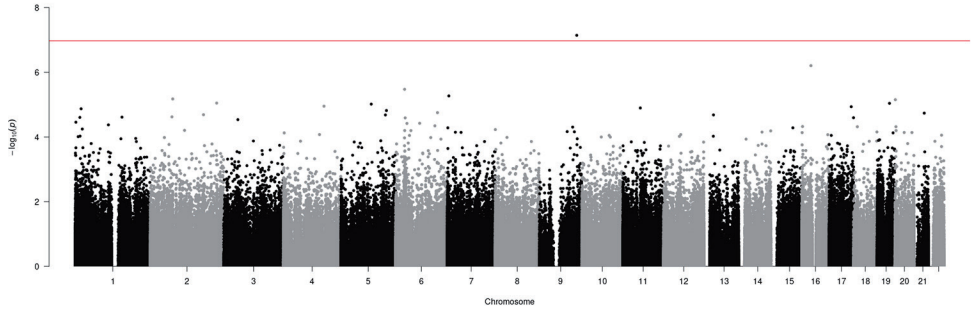
Supplemental Figure 1. Timeline of measurements in Generation R and ALSPAC



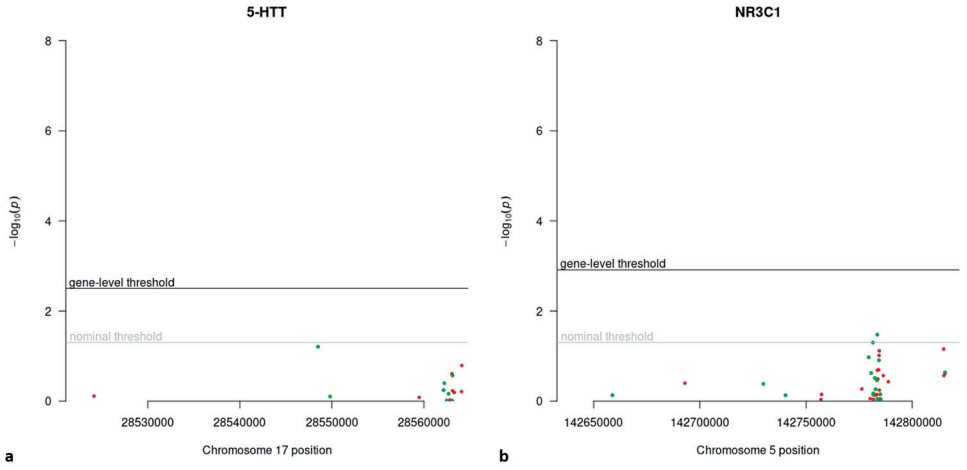
Supplemental Figure 2. Overlap samples sizes per time point (a) in Generation R (*n* data points = 795, *n* participants = 506) and (b) in ALSPAC (*n* data points = 1,639, *n* participants = 846)



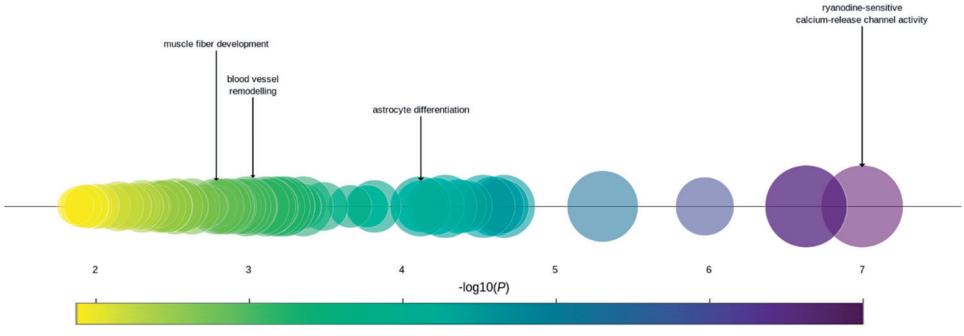
Supplemental Figure 3. QQ-plot of epigenome-wide study in (a) Generation R (λ = 1.063) and (b) QQ- ALSPAC (λ = 1.064)



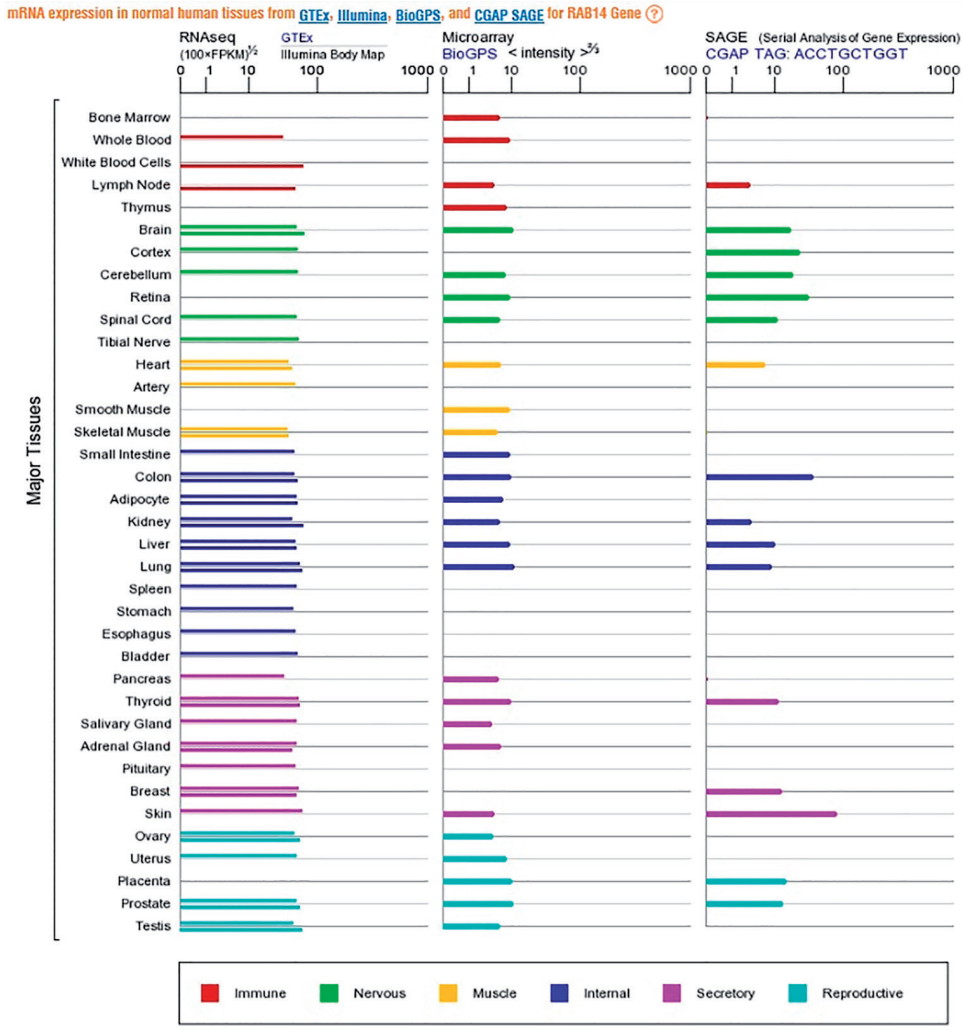
Supplemental Figure 4. Manhattan plot of -log10 *p*-values from meta-analysis



Supplemental Figure 5. Regional Manhattan plot of -log10 *p*-values from meta-analysis for candidate genes for the (a) serotonin transporter *5-HTT*, and (b) glucocorticoid receptor, *NR3C1*. Green points indicate a more positive change over time for the group exposed to bullying versus the non-exposed group, red points indicate a more negative change.



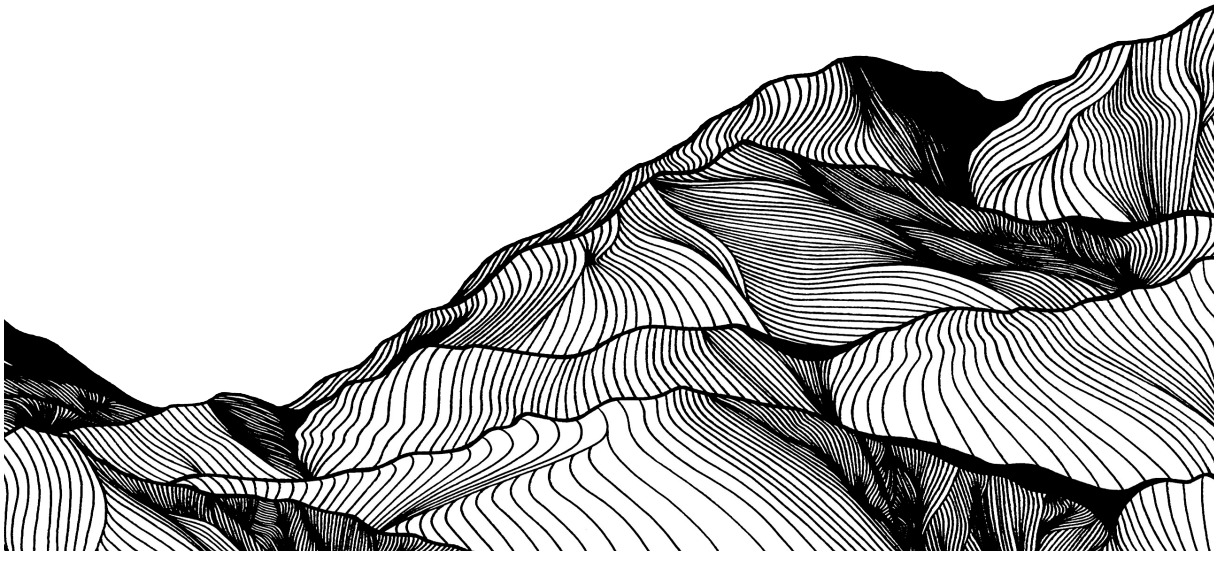
Supplemental Figure 6. Enriched biological processes ($p < 0.05$) in Gene Ontology analysis of CpG sites with $p < 0.001$ ($n = 644$ CpG sites, $n = 396$ genes) ordered by significance. Circle size represents percentage of genes represented in pathway versus all genes in that pathway.



Supplemental Figure 7. *RAB14* mRNA expression in human tissues as portrayed by https://www.genecards.org/cgi-bin/carddisp.pl?gene=RAB14[15]

Supplemental References

1. Lehne B, Drong AW, Loh M, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biology.* 2015;16:37.
2. Min JL, Hemani G, Davey Smith G, Relton C, Suderman M, Hancock J. Meffil: efficient normalization and analysis of very large DNA methylation datasets. *Bioinformatics.* 2018.
3. R Core Team. *R: A language and environment for statistical computing*. 2013.
4. Paternoster R, Brame R, Mazerolle P, Piquero A. Using the correct statistical test for the equality of regression coefficients. *Criminology.* 1998;36:859-66.
5. Clogg CC, Petkova E, Haritou A. Statistical methods for comparing regression coefficients between models. *American Journal of Sociology.* 1995;100:1261-93.
6. Junutula JR, De Maziére AM, Peden AA, et al. Rab14 is involved in membrane trafficking between the Golgi complex and endosomes. *Molecular Biology of the Cell.* 2004;15:2218-29.
7. Ray BK, Murphy R, Ray P, Ray A. SAF-2, a splice variant of SAF-1 acts as a negative regulator of transcription. *Journal of Biological Chemistry.* 2002.
8. Shiina T, Inoko H, Kulski JK. An update of the HLA genomic region, locus information and disease associations: 2004. *Tissue Antigens.* 2004;64:631-49.
9. Almeida B, Fernandes S, Abreu IA, Macedo-Ribeiro S. Trinucleotide repeats: a structural perspective. *Frontiers in Neurology.* 2013;4:76.
10. Ronchi D, Garone C, Bordoni A, et al. Next-generation sequencing reveals DGUOK mutations in adult patients with mitochondrial DNA multiple deletions. *Brain.* 2012;135:3404-15.
11. Das S, Boczan J, Gerwin C, Zald PB, Sheng Z-H. Regional and developmental regulation of syntaphilin expression in the brain: a candidate molecular element of synaptic functional differentiation. *Molecular Brain Research.* 2003;116:38-49.
12. Kloosterman WP, Lagendijk AK, Ketting RF, Moulton JD, Plasterk RHA. Targeted inhibition of miRNA maturation with morpholinos reveals a role for miR-375 in pancreatic islet development. *PLoS Biology.* 2007;5:e203.
13. Gillespie CS, Sherman DL, Blair GE, Brophy PJ. Periaxin, a novel protein of myelinating Schwann cells with a possible role in axonal ensheathment. *Neuron* .1994;12:497-508.
14. Oltmann-Norden I, Galuska SP, Hildebrandt H, et al. Impact of the polysialyltransferases ST8SiaII and ST8SiaIV on polysialic acid synthesis during postnatal mouse brain development. *Journal of Biological Chemistry.* 2008;283:1463-71.
15. Stelzer G, Rosen N, Plaschkes I, et al. The GeneCards suite: from gene data mining to disease genome sequence analyses. *Current Protocols in Bioinformatics.* 2016;54:1.30. 1-1.. 3.



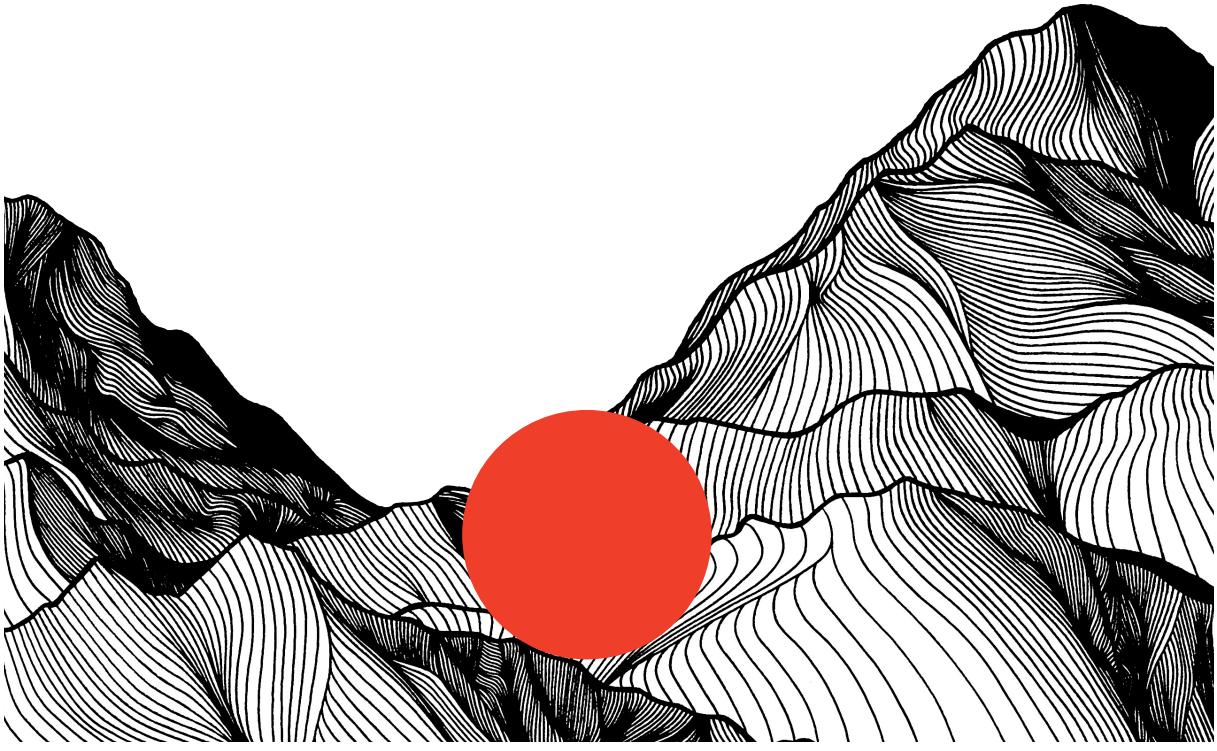
CHAPTER VIII

Epigenome-wide change and variation in DNA methylation from birth to late adolescence

Rosa H. Mulder, Alexander Neumann*, Charlotte A. M. Cecil*, Esther Walton, Lotte C. Houtepen, Andrew J. Simpkin, Jolien Rijlaarsdam, Bastiaan T. Heijmans, Tom R. Gaunt, Janine F. Felix, Vincent W. V. Jaddoe, Marian J. Bakermans-Kranenburg, Henning Tiemeier, Caroline L. Relton, Marinus H. van IJzendoorn, & Matthew Suderman

*authors contributed equally

Submitted. Preprint: bioRxiv (2020).
Results available: http://epidelta.mrcieu.ac.uk/.



Abstract

DNA methylation is known to play a pivotal role in childhood health and development, but a comprehensive characterization of genome-wide DNA methylation trajectories across this age period is currently lacking. We have therefore performed a series of epigenome-wide association studies in 5,019 blood samples collected at multiple time-points from birth to late adolescence from 2,348 participants of two large independent cohorts. DNA methylation profiles of autosomal CpG sites (CpGs) were generated using the Illumina Infinium HumanMethylation450 BeadChip. Change over time was widespread, observed at over one-half (53%) of CpGs. In most cases DNA methylation was decreasing (36% of CpGs). Inter-individual variation in linear trajectories was similarly widespread (27% of CpGs). Evidence for nonlinear change and inter-individual variation in nonlinear trajectories was somewhat less common (11% and 8% of CpGs, respectively). Very little inter-individual variation in change was explained by sex differences (0.4% of CpGs) even though sex-specific DNA methylation was observed at 5% of CpGs. DNA methylation trajectories were distributed non-randomly across the genome. For example, CpGs with decreasing DNA methylation were enriched in gene bodies and enhancers and were annotated to genes enriched in immune-developmental functions. By contrast, CpGs with increasing DNA methylation were enriched in promoter regions and annotated to genes enriched in neurodevelopmental functions. These findings depict a methylome undergoing widespread and often nonlinear change throughout childhood. They support a developmental role for DNA methylation that extends beyond birth into late adolescence and has implications for understanding life-long health and disease. DNA methylation trajectories can be visualized at http://epidelta.mrcieu.ac.uk.

Introduction

DNA methylation, an epigenetic process whereby DNA is modified by the addition of methyl groups, has gained increasing attention over the past few decades, due to its pivotal role in development. *In utero*, DNA methylation is involved in a range of essential processes including cell differentiation[1-3], X-chromosome inactivation[4] and fetal growth[5]. Its role extends well beyond birth, e.g. by maintaining cell type identity and genome stability[6-8], responding to environmental exposures[9-11], and its involvement in immune[12] and neural development[13]. Since it is influenced by *both* genetic and environmental factors[14, 15], DNA methylation has also emerged as a key mechanism of interest for understanding the gene-environmental interplay in normal ageing and disease development.

Numerous studies have identified strong associations between DNA methylation and age. While most have relied on cross-sectional data[16-18], but a few have utilized longitudinal measurements of DNA methylation within individuals[19-23]. Longitudinal measurements allow one to distinguish intra-individual change from inter-individual differences in change, thereby greatly improving the power to detect change over time and to identify differences between individuals[24]. Identifying and characterizing CpGs for which DNA methylation changes differently over time between individuals (i.e. inter-individual variation in change) is a necessary step in identifying genetic and environmental influences on the methylome as well as their potential impact on health outcomes[25]. Moreover, longitudinal designs facilitate the study of nonlinear trajectories[26, 27], which might help to identify sensitive periods for DNA methylation change in development. To date, the largest epigenome-wide longitudinal study on DNA methylation included 385 elderly individuals who were followed up to five times over a maximum period of 18 years, identifying DNA methylation change at 1,316 CpG (Cytosine-phosphate-Guanine) sites[19] and inter-individual variation at change at 570 CpGs[20]. Yet, little is known about DNA methylation trajectories across early development, as existing studies in *childhood* DNA methylation typically have been limited by small sample sizes[21, 23], short time-periods[22, 28] or focused on specific CpGs in relation to maternal smoking[29], birthweight[30], or maternal BMI[31].

In the current study, we aim to provide a benchmark of typical epigenome-wide age-related DNA methylation trajectories within individuals, spanning the first two decades of life. This study combines repeated measurements of DNA methylation at nearly half a million CpG sites across the genome from two large population-based cohorts, the Generation R Study and Avon Longitudinal Study of Parents and Children (ALSPAC), to form one integrated dataset with four time-points of measurement. In a series of three epigenome-wide mixed model analyses we study linear (Model 1), nonlinear (Model 2) and sex-related (Model 3) trajectories of change across development. Further, we aim to identify CpGs for which trajectories vary between individuals (Model 1 and 2). Results are interpreted in the context of CpG location

and biological pathways. The key findings are discussed here, full results per CpG can be freely accessed and visualized at http://epidelta.mrcieu.ac.uk/.

Results

Cross-cohort comparability

Sample characteristics of 1,399 Generation R participants (total DNA methylation samples = 2,333) and of 949 ALSPAC participants (total DNA methylation samples = 2,686; Figure 1) are provided in Supplemental Table 1. After the DNA methylation datasets of the two cohorts underwent joint functional normalization (see Supplemental Figure 1 for distributions of mean DNA methylation levels), within-cohort stability of DNA methylation at birth and 6 or 7 years (in Generation R and ALSPAC, respectively) was compared. Stability of DNA methylation at individual CpG sites (437,864 autosomal sites) was estimated in three ways: relative concordance using Spearman correlations between time points, absolute concordance using intraclass correlations between time points (children with data for both time points: *n* Generation R = 476, *n* ALSPAC = 826), and change over time using change estimates from a linear mixed model (Model 1, see Methods) applied within each cohort (children with data for at least one of the two time-points: *n* Generation R = 1,394, *n* ALSPAC = 944). Estimates of all stability measures for both cohorts are depicted in Figure 2. Next, agreement of these stability estimates between the two cohorts was estimated with the Spearman (ρ) or Pearson (*r*) correlation (depending on normality of the data) across all CpGs, between the datasets. The Spearman correlation of the relative concordance was ρ = 0.62, the Pearson correlation of the absolute concordance was ρ = 0.60, and the Pearson correlation of the change estimates was *r* = 0.86, indicating strong agreement between datasets. Based on these results the two

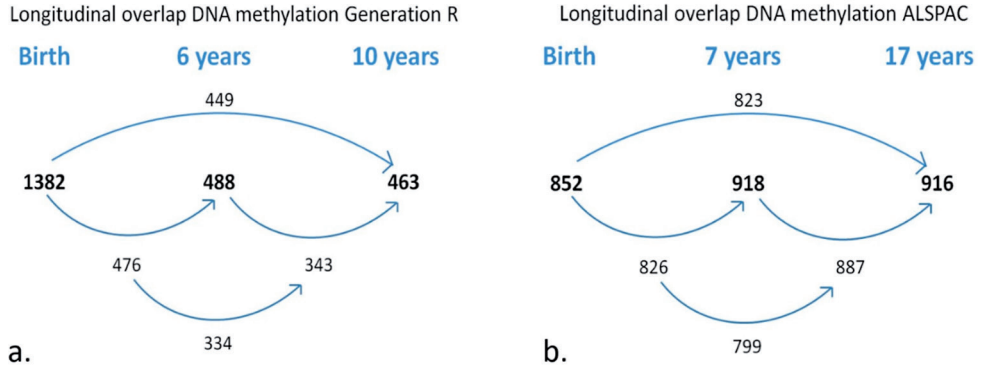


Figure 1. Longitudinal sample sizes. Sample sizes for (a) Generation R (N total children = 1,399, N total DNA methylation samples = 2,333); and (b) ALSPAC (N total children = 949, N total DNA methylation samples = 2,686). Bolded numbers represent total sample size at each time-point; non-bolded number refer to overlapping samples between time-points.

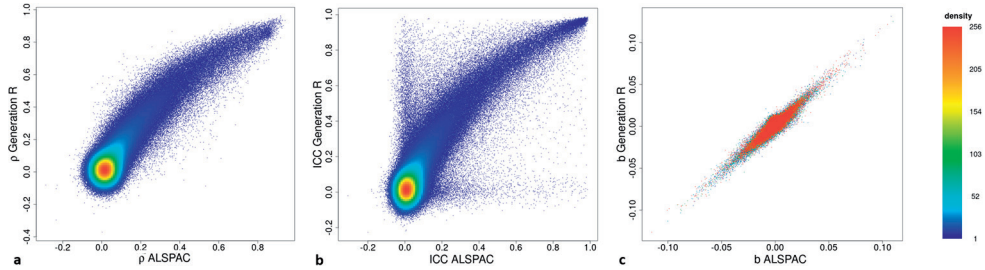


Figure 2. Scatterplots of within-cohort stability of DNA methylation. Showing (a) Spearman correlations, (b) intraclass correlation coefficients and (c) change estimates from birth to 6/7 years per CpG for Generation R and ALSPAC.

datasets were joined to form one set with four different time-points of DNA methylation (birth, age 6/7 years, 10 years, 17 years).

Linear DNA methylation change from birth to early adulthood

Estimates of overall change in DNA methylation from birth to early adolescence (Model 1; see Methods) indicated linear change at 51.6% of CpGs at a Bonferroni-corrected threshold ($P < 1x10^{-07}$) (Figure 3a and 3b). Specifically, DNA methylation decreased over time at 35.5% of all CpGs and increased at 16.0% (Figure 4). The mode intercept indicated that the decreasing CpGs were 88% methylated at birth (Figure 5). DNA methylation levels for increasing CpGs typically started at 5%.

The mode estimate DNA methylation change was $b = -9.24x10^{-04}$ (with corresponding mode $SE = 6.85x10^{-05}$), indicating an overall 0.09% DNA methylation decrease per year at a typical CpG site. This translates into a 1.66% decrease in DNA methylation over the course of 18 years. An example of a CpG site with a typical change in DNA methylation is depicted in Figure 3a. The largest observed absolute change in DNA methylation was $b = -3.47x10^{-02}$ ($SE = 3.65x10^{-04}$, $p < 9.88x10^{-324}$), indicating an overall DNA methylation decrease of 62.5% over 18 years (Figure 3b). Only twenty-two CpGs showed an absolute change > 50% over the course of 18 years (Supplemental Table 2). From this it follows that typically in (cord-/peripheral) blood tissue, DNA methylation levels for CpGs do not change from a fully unmethylated to fully methylated state, or vice versa, over the course of 18 years.

Further, we observed substantial inter-individual variation in linear DNA methylation changes over time at 27.4% of all CpGs (i.e. random slope variance was greater than zero at Bonferroni-corrected threshold $p < 1x10^{-07}$; Figure 3c). On average, this variation accounted for 2.7% (SD = 1.5%) of all estimated inter-individual variation (for intercept, age, batch, and residual) at these CpGs. At 17.3% of all CpGs, we observed both change and inter-individual variation in change.

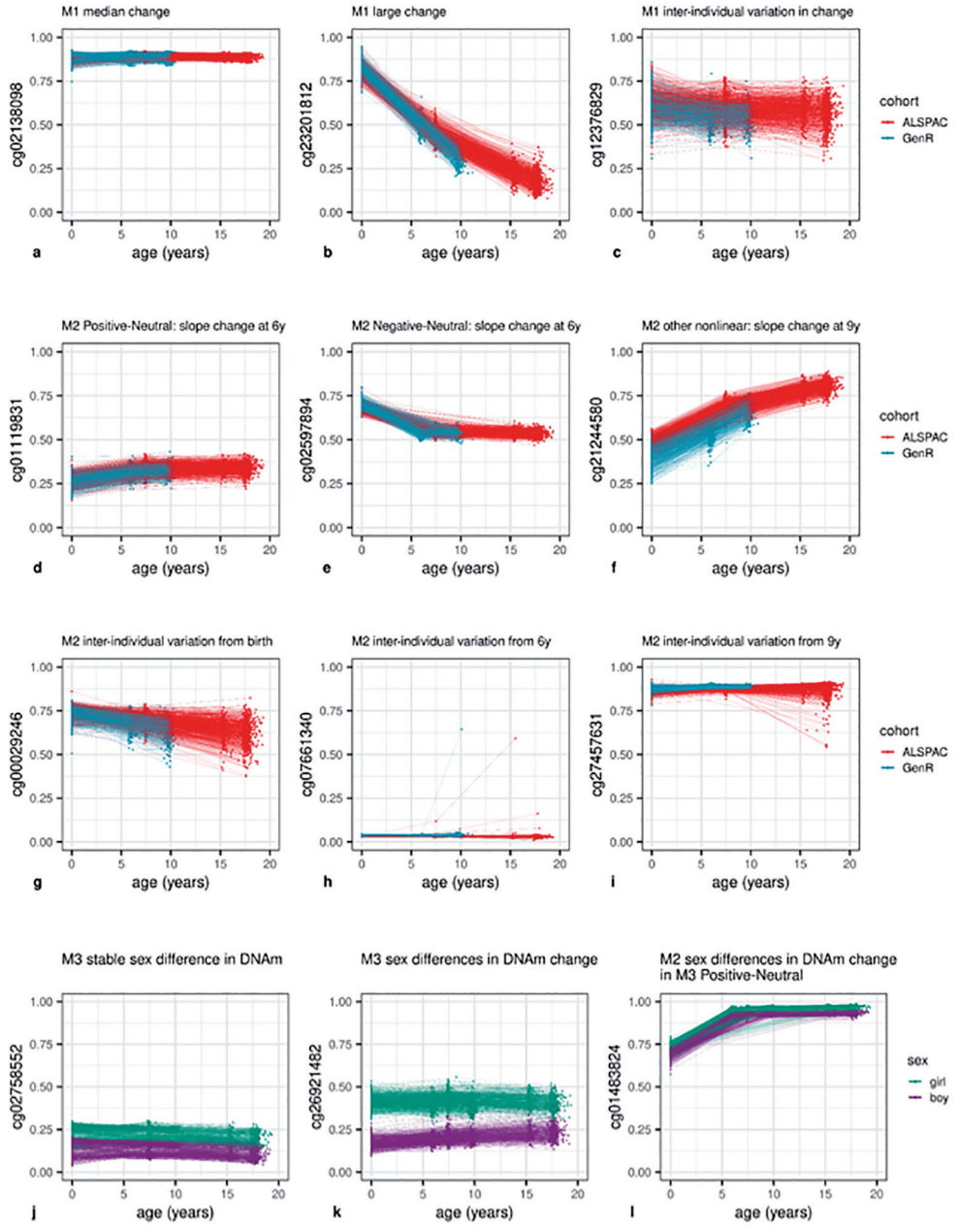


Figure 3. DNA methylation levels of selected CpG sites across childhood. Parts (a-c) show CpG sites with linear change over time (Model 1). A typical site is shown in (a), the site with the largest observed change in (b) and with inter-individual variation in DNA methylation change (c). Parts (d-f) show CpG sites with non-linear change (Model 2). A Positive-Neutral trajectory is shown in (d), a Negative-Neutral trajectory in (e) and a Positive-More Positive-Less Positive in (f). Parts (g-i) show CpG sites with inter-individual variation in change (Model 2). A site with slope variation from birth is shown in (g), slope change variation at 6 in (h) and slope change variation at 9 in (i). Parts (j-l) show CpG sites with sex-specific DNA methylation. A site with stable sex differences is shown in (Model 3) (j), sex-specific slope in (Model 3) (k) and sex-specific slope change at 6 in (Model 2) (l).

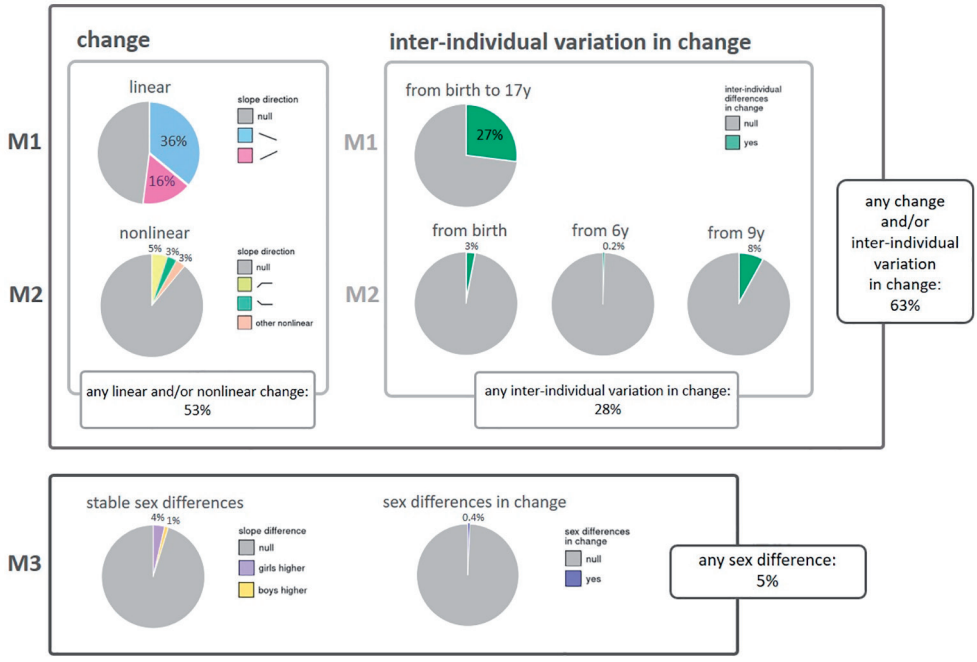


Figure 4. Overview of results from the three models. Model 1 (M1) was applied for overall change in DNA methylation and inter-individual variation in linear change; Model 2 (M2) for nonlinear change in DNA methylation and inter-individual variation in nonlinear change; and Model 3 (M3) for stable sex differences in DNA methylation and sex differences in change of DNA methylation (Sex by Time interaction). Percentages represent percentage of autosomal CpGs below Bonferroni-corrected threshold ($p < 1x10^{-07}$).

Nonlinear DNA methylation change

Model 2 (see Methods) was identical to Model 1 but permitted slope changes at ages 6 and 9 years to test for nonlinear DNA methylation trajectories. At 11.0% of CpGs a nonlinear trajectory was detected. Specifically, at 4.8% of all CpGs, DNA methylation increased from birth and remained stable from 6 onward (Positive-Neutral; Figure 3d). Second, at 3.1% of all CpGs, DNA methylation decreased from birth and then remained stable at 6 years (Negative-Neutral; Figure 3e). The remaining 3.0% of all CpGs followed other nonlinear trajectories (e.g. Figure 3f), with each trajectory observed in < 1.0% of all CpGs. Overall, linear and/or nonlinear changes in Model 1 or 2 were observed in 52.6% of CpGs (Figure 3), indicating that most nonlinear patterns were also detected as linear patterns in Model 1.

Inter-individual differences in change (i.e. random variance in slopes) from birth onward was detected at 3.4% of all sites (Figure 3g), inter-individual differences in slope change at 6 years in 0.2% (Figure 3h), and inter-individual differences in slope change at 9 years at 8.2% of CpGs (Figure 3i). Inter-individual differences in slope (change) at each time-point were detected more often at CpGs with an increasing rather than decreasing overall DNA methylation change in

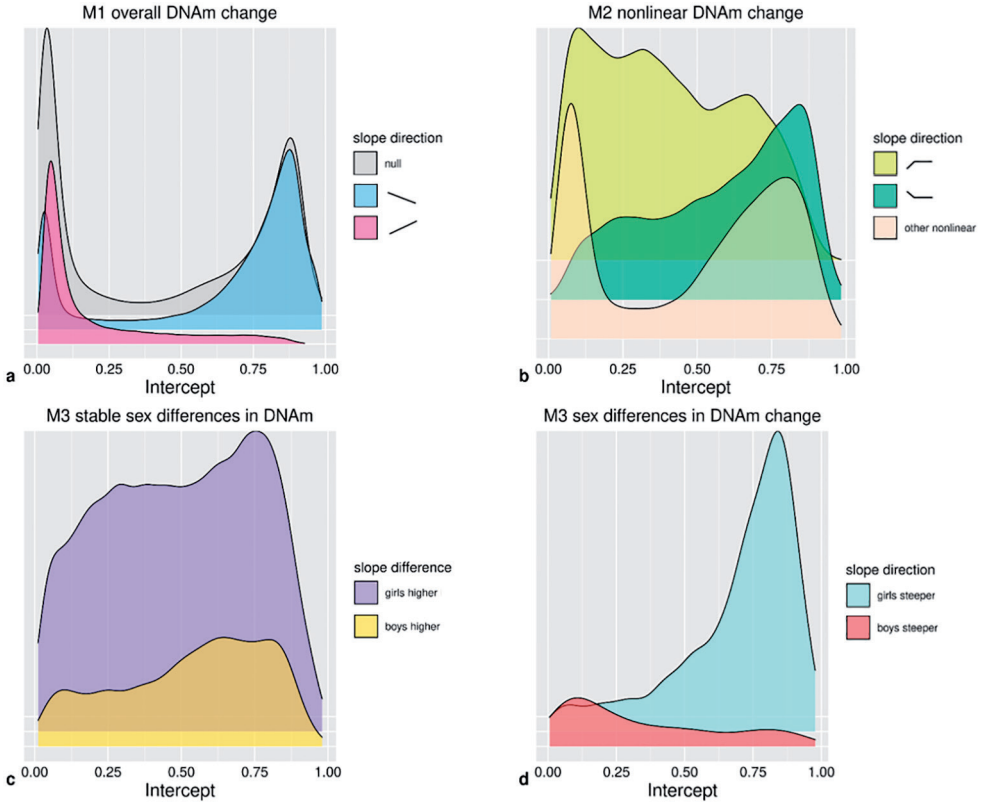


Figure 5. Density plots of intercepts of CpGs. Intercepts for CpGs with (a) directions of change in Model 1 (n = 473,864); (b) nonlinear trajectories in Model 2 (n = 52,043); (c) stable sex differences in Model 3 (n = 22,821); (d) sex differences in DNA methylation change in Model 3 (n = 1,768).

Model 1 ($p = 2.37 \times 10^{-144}$). Last, both Positive-Neutral and Negative-Neutral changes coincided more often with inter-individual variation from birth ($p < 9.88 \times 10^{-324}$). Any inter-individual differences in change, detected by Model 1 or 2, was observed at 27.9% of CpGs. In total, Models 1 and 2 detected age-related change whether linear, non-linear or inter-individual differences in change at 62.8% of all CpG sites (Figure 3).

Sex differences in longitudinal DNA methylation and DNA methylation change

According to Model 3 (see Methods), sex differences in DNA methylation were present at 4.9% of (autosomal) CpGs (Figure 3). Specifically, stable longitudinal sex differences (main sex effects) were observed at 4.8% of all (autosomal) CpGs (Figure 3j), and sex differences in DNA methylation change (sex by age interaction effects) were found at 0.4% of all (autosomal) CpGs (Figure 3k). At sites with stable sex differences, DNA methylation levels were higher in girls at 3.6% (Figure 3j) and lower at 1.2% of CpG sites. DNA methylation at sites with higher

DNA methylation in girls tended to increase over time, whereas DNA methylation at sites with higher DNA methylation in boys tended to decrease ($p = 4.20x10^{-205}$). Most commonly (at 0.2% of all CpGs), DNA methylation was higher in girls at birth but DNA methylation in boys increased at a higher rate.

Both CpGs with stable sex differences and those with sex differences in DNA methylation change were less likely to show inter-individual variation than other sites (20.8% versus 27.5% and 18.1% versus 27.3%; $p = 5.36x10^{-111}$ and $p = 7.57x10^{-18}$). Finally, CpGs with stable sex differences or sex differences in DNA methylation change detected in Model 3 were much more likely to follow an overall Positive-Neutral trajectory of DNA methylation change detected in Model 2 than other CpG sites were (24.2% of CpGs with stable sex differences followed a Positive-Neutral trajectory versus 3.8% of other CpGs and 53.9% of CpGs with sex differences in DNA methylation change followed a Positive-Neutral trajectory versus 4.6% of other CpGs; $p < 9.88x10^{-324}$, $p < 9.88x10^{-324}$; Figure 3l). Albeit less prominently so, CpGs with stable sex differences or sex differences in DNA methylation change also more often followed a Negative-Neutral trajectory than other CpGs did (stable sex differences: 5.0% versus 3.0%, $p = 5.43x10^{-62}$; sex differences in DNA methylation change: 7.7% versus 3.1%, $p < 7.11x10^{-28}$).

Follow-up analyses

Follow-up analyses were performed to understand how different types of age-related DNA methylation trajectories are distributed across the genome (Supplemental Tables 3-5). All reported enrichments have significance below a Bonferroni-corrected threshold of $p < 4.46x10^{-04}$, corrected for the number of chi-square tests ($n = 112$). We further report enrichment of Gene Ontology (GO) pathways (nominal $p < 0.05$) for genes annotated to CpG sites in each trajectory (Supplemental Tables 5-7). Last, we study enrichment of age-related DNA methylation trajectories in reported hits of different EWASs (Figure 6). All reported EWAS enrichments are below a Bonferroni-corrected threshold of $p < 1.38x10^{-04}$, corrected for the number of Fishers' exact tests ($n = 363$; Supplemental Table 8).

Patterns of DNA methylation change and CpG location

CpG sites with DNA methylation change associated patterns were labeled by gene associated regions, CpG island associated regions, as well as enhancer elements. Although many exceptions exist, low levels of DNA methylation in the promoter area but high levels of DNA methylation in the gene body are generally associated with increased gene transcription[32, 33]. CpGs annotated to TSS200 regions more often showed an overall DNA methylation increase (Model 1) than other CpGs (19.0% versus 15.6%), whereas CpGs annotated to the gene body more often showed an overall DNA methylation decrease than other sites (38.8% versus 33.7%). TSS200 CpGs showed less inter-individual variation in overall DNA methylation change than other sites (22.2% versus 28.1%), whereas gene body CpGs showed somewhat more inter-individual variation in overall DNA methylation change than other sites (28.9% versus 26.5%).

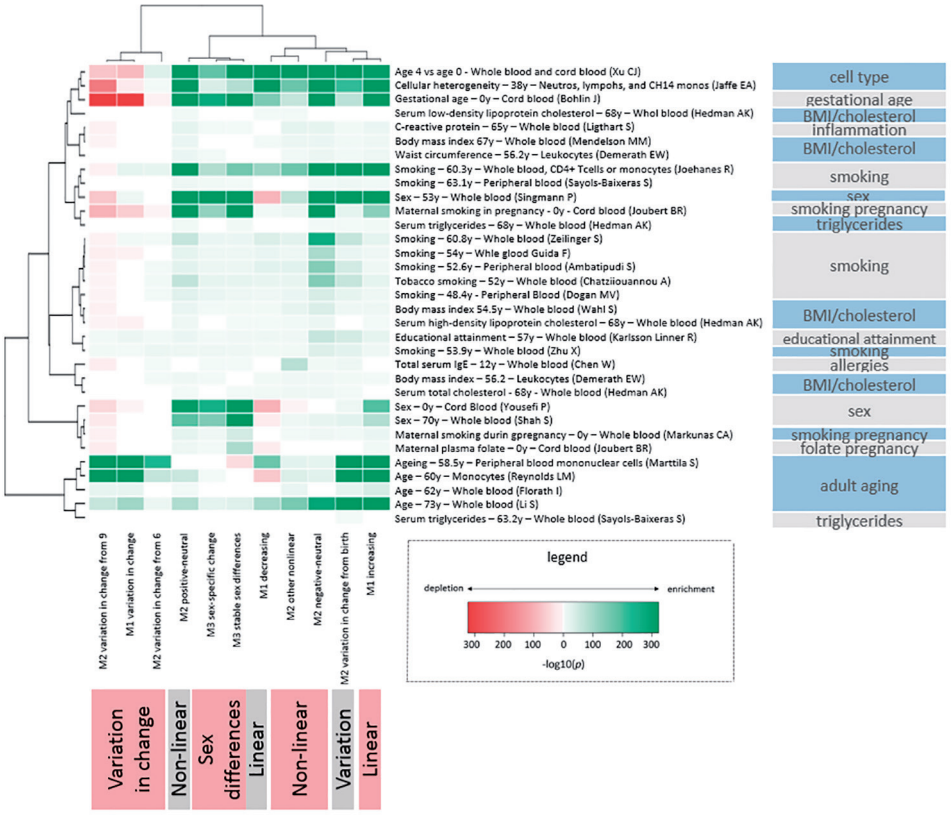


Figure 6. Enrichment of age-related trajectories in EWASs

Promoter areas often coincide with CpG islands[34]. Here, 63.3% of TSS200 CpGs were also annotated to CpG islands. As in TSS200 areas, CpGs annotated to CpG islands had lower DNA methylation levels (mode M1 intercept 2.4% (*SD* = 30.2%)), and more often showed an overall DNA methylation increase than other sites (25.2% versus 12.0%). DNA methylation sex differences were especially present in the shores of CpG islands compared to all other island associated regions (stable sex differences: 7.5% versus 4.0%, sex differences in DNA methylation change: 0.6% versus 0.3%).

Enhancers act on promoters to regulate gene transcription[35]. CpGs annotated to enhancer elements (2.0% of CpGs) tended to have low DNA methylation levels (mode M1 intercept 5.07%; *SD* = 31.4%) and then increased with age more than other CpGs (23.9% versus 15.9%). Inter-individual variation in change from birth was more common at enhancer sites than at other sites (5.6% versus 3.3%).

Functional associations

Enrichment of Gene Ontology categories was tested for genes linked to CpGs with different DNA methylation trajectories. In short, genes annotated to CpGs with overall decreasing DNA methylation levels were enriched in immune-developmental functions, whereas those annotated to CpGs with increasing levels were enriched in neurodevelopmental functions. This pattern seemed even more pronounced at genes annotated to nonlinear Negative-Neutral and Positive-Neutral CpGs, with the former more often associated to immune-development and the latter to neurodevelopment. Genes linked to CpGs with stable sex differences and sex differences in DNA methylation change were enriched in pathways associated with sexual development, such as genital development, as well as pathways associated with neurodevelopment. Genes linked to CpGs with sex difference in DNA methylation change were also enriched in functions related to tooth and hair development.

Enrichment in EWASs

We further investigated functional relevance of CpG sites with age-related DNA methylation trajectories by testing enrichment with published EWAS associations (Figure 6)[28, 36-61]. Unsupervised clustering of the enrichments shows that CpG sites with inter-individual variation in change over time have distinct enrichments and cluster differently from those with age-associated change that is consistent among individuals. The CpG sites of each age-associated DNA methylation trajectory were enriched with published age associations in adulthood. Multiple smoking EWAS clustered together with enrichment patterns exhibiting strongest enrichments among CpG sites with Negative-Neutral trajectories and mostly weak enrichments among CpG sites with inter-individual variation in change. Further, despite adjusting for cell count heterogeneity in our models, we observed enrichments of CpG sites that differ by white blood cell type among sites following nearly all age-associated trajectories. Finally, we observed enrichments of CpG sites associated with gestational age and prenatal smoking with sex-specific DNA methylation.

Discussion

In this study we described changes in DNA methylation levels through the first two decades of human life. We examined DNA methylation levels per CpG by their linear association with age, their nonlinear trajectories and inter-individual variation in change, as well as sex differences and CpG characteristics.

We found that about half of sites change: consistent linear and/or nonlinear DNA methylation change was found at 53% of sites. We further found that over a quarter of sites, 28%, were characterized by substantial inter-individual differences in the direction of this change. DNA

methylation sex differences were present, but not abundant: 5% of autosomal sites displayed different DNA methylation levels or differences in change over time for girls and boys.

Specifically, we determined that DNA methylation at 52% of the measured methylome have some form of linear change from birth to late adolescence, with DNA methylation decreasing at 36% and increasing at 16% of CpGs. CpGs with decreasing DNA methylation tended to have high levels of DNA methylation and were more often located in gene bodies. CpGs with increasing levels of DNA methylation tended to have low levels of DNA methylation and were more likely to be located in promoter regions and at enhancers. The predominance of decreasing CpGs is in agreement with literature on epigenome-wide DNA methylation and age in cross-sectional research on children and adults[18, 62], as well as with longitudinal research in adults[19].

Nonlinear DNA methylation trajectories were detected at 11% of CpGs, mostly involving changes in DNA methylation from birth to age 6 years, after which DNA methylation was more stable. We note that this could be due to cord blood being used to generate DNA methylation profiles at birth, whereas peripheral blood was used at later ages. A previous study[23] including eight children showed that the cord blood DNA methylation profile at birth clustered separately from later peripheral profiles, after which DNA methylation changed gradually from 1, to 2.5, to 5 years. Such differences between DNA methylation in cord and peripheral blood might be due to uncaptured differences in white blood cell composition, as well as to different gene-regulatory functioning in the intra-uterine versus extra-uterine environment. On the other hand, Gene Ontology analyses showed that functional associations for positive and negative *linear* DNA methylation patterns, which are unlikely to be affected by tissue type, were consistent with functional associations for nonlinear positive and negative patterns, respectively (e.g. positive and negative up to 6 years, and then no change up to 18 years).

Specifically, sites with decreasing levels of DNA methylation, both with or without slope changes around the age of 6 years, were functionally enriched for immune-developmental pathways, and sites with increasing levels of DNA methylation, both with or without slope changes, were enriched for neurodevelopmental pathways. Since these observations were based on blood DNA methylation, it remains to be studied what roles genes linked to neurodevelopmental pathways play in in blood, or to what extent DNA methylation trajectories in blood mirror those in neural tissue.

Inter-individual differences in linear DNA methylation trajectories were found at 27% of CpGs, indicating change at different rates or directions for different individuals. Such sites tended to have overall increasing rather than decreasing levels of DNA methylation from birth to 18 years. This observation is consistent with the only other large study to examine inter-individual differences in DNA methylation change[20], although we note that this study

included elderly subjects and detected a smaller number of relevant CpGs. We are the first to investigate inter-individual differences in nonlinear DNA methylation trajectories. These were most often found in the slope change at 9 years (8% of CpG sites), indicating that most inter-individual differences in DNA methylation emerge after the first decade of life. More research is needed to understand if the direction of change in this period is determined by stimuli during that period, or rather by preceding, perhaps cumulative, exposures. However, it is clear that, given the high proportion of CpG sites with inter-individual variation in DNA methylation change over time that we have observed, it is important to restrict the range of ages of children included a single EWAS. Specific limits should be discussed given the rapidly growing number of studies generating DNA methylation profiles across childhood[63].

Stable sex differences were found at 5% of autosomal CpGs, and sex differences in DNA methylation *change* were found at 0.4% of all CpGs. In general, if there were stable sex differences, girls had higher levels of DNA methylation (4% of all CpGs), in case of sex differences in DNA methylation change, boys had an accelerated upward change (0.2% of all CpGs). The direction of stable sex differences detected are congruent with a cross-sectional study on newborns, in which girls had higher DNA methylation levels than boys for the large majority of the 3031 significant autosomal CpGs[54]. Sex-discordant associations with age seemed to be more prevalent from birth to age 6 years than afterwards, suggesting that any phenotypic sex differences associated to DNA methylation would be established in early childhood. Their enrichment in the shores of CpG islands, areas at which DNA methylation has been associated with tissue differentiation and tissue-specific gene expression[64], is consistent with the critical role that these processes play in sexual differentiation. Studies into sex differences in epigenetic regulation might want to focus on these locations.

We also found the other DNA methylation trajectories to be arranged throughout the genome in a nonrandom fashion. Earlier studies[32, 65] have shown that, for active genes, lower DNA methylation towards the promoter area (TSS200) and higher DNA methylation in the gene body relate to increased gene transcription. Here we add the observation that promoter DNA methylation tends to increase and gene body DNA methylation tends to decrease with age. From this finding, one might infer that a downregulation of gene expression takes place from birth to late adolescence. Enrichment analyses of published EWAS associations further showed that different traits and exposures exhibited distinct enrichment patterns among DNA methylation trajectories. For example, there were clear differences between smoking and BMI-related traits. Enrichment of sites with DNA methylation sex differences in EWASs on prenatal maternal smoking is consistent with studies finding that prenatal smoking affects traits such as birth weight[66] , brain development[67, 68], and attention[69] differently in boys and girls. Clustering for prenatal maternal smoking EWASs also showed enrichment for CpGs with consistent change among individuals, not for CpGs with inter-individual variation in change. This may suggest a link with the well-known effects of prenatal smoking on childhood

development since consistent DNA methylation change is more likely related to development or aging programming than inter-individual variation. This may explain why changes associated to prenatal smoking persist throughout life[70]. Notably, this pattern of change without inter-individual variation is visible in cg05575921, the *AHRR* CpG site strongly and persistently associated with prenatal smoking[71, 72] (Supplemental Figure 2; http://epidelta.mrcieu.ac.uk/).

'Epigenetic age acceleration' is a term coined to indicate the deviation of chronological age from age as estimated by an 'epigenetic clock' and is associated with disease risk and mortality[73]. Existing clocks are all linear models based on DNA methylation. Consequently, one might expect that all CpGs included in the clock model change linearly with age. Furthermore, to detect age acceleration, one would expect that these CpG sites would also vary between individuals. Surprisingly, many CpG sites included in the most popular clocks do not match these expectations[74, 75] (Supplemental Tables 9, 10). For example, we observe that over one-quarter and nearly one-half of the CpG sites included in the Horvath and Hannum clocks, respectively, follow non-linear DNA methylation trajectories in childhood. Given the widespread use of clocks to investigate biological aging, further investigation is warranted to better understand how, and perhaps if, associations using these clocks should be interpreted in child DNA methylation profiles.

We note three main limitations of our findings. First, the use of different tissue types (cord blood and peripheral blood) could account for some of the differences between birth and later time points, e.g. sites that increased or decreased between birth and 6, but did not show change after that. Generation of DNA methylation profiles of a single tissue or cell type collected across childhood would be needed to disentangle this issue further. Unfortunately, such a dataset is not currently available as most cohorts have generated DNA methylation profiles from peripheral blood and cord blood[63]. Analysis of these complex tissues has nevertheless yielded many valuable insights. Second, since DNA methylation at 9 years was measured only in Generation R and at 17 years only in ALSPAC, DNA methylation differences from 9 to 17 may be to some extent driven by batch effects or cohort differences. This may explain some of the inter-individual differences in slope changes at 9 towards 17 years. However, the high level of agreement in both stability and change among the corresponding time points of the two cohorts is reassuring. Moreover, it is not entirely surprising that inter-individual variation in directionality of change was higher for the largest age interval. This interval, furthermore, encompasses the period of adolescent development, a time in which many inter-individual phenotypic differences arise. Finally, it should be noted that the current study only included children of European ancestry. Considerable DNA methylation differences have been found between populations[76-78], but research on age-associated DNA methylation differences is scarce. One study[79] reported evidence for overlap in age-associated CpGs in two African populations with studies on European-ancestry populations, but more research is needed to map the generalizability of longitudinal DNA methylation changes among different populations.

In conclusion, in the first comprehensive CpG-by-CpG characterization of DNA methylation from birth to late adolescence, we found that DNA methylation at more than half of the studied CpG sites changes consistently between individuals, and that considerable inter-individual variation in change exists. Further, characteristics such as child sex, CpG location, and environmental and disease traits have distinct associations with patterns of DNA methylation change. Further analysis of these patterns is made readily available at http://epidelta.mrcieu.ac.uk/, which we hope can be used in future studies to test developmental hypotheses that promote our understanding of the developmental nature of DNA methylation, its role in gene functioning, and the associated biological pathways leading to health and disease.

Methods

Setting

Data were obtained from two population-based prospective birth cohorts, the Dutch Generation R Study (Generation R) and the British Avon Longitudinal Study of Parents and Children (ALSPAC). Pregnant women residing in the study area of Rotterdam, the Netherlands, with an expected delivery date between April 2002 and January 2006 were invited to enroll in Generation R. A more extensive description of the study can be found elsewhere[80]. The Generation R Study is conducted in accordance with the World Medical Association Declaration of Helsinki and has been approved by the Medical Ethics Committee of the Erasmus Medical Center, Rotterdam. Informed consent was obtained for all participants.

Pregnant women residing in the study area of former county Avon, United Kingdom, with an expected delivery date between April 1991 and December 1992 were invited to enroll in the ALSPAC study. Detailed information on the study design can be found elsewhere[3, 81]. The ALSPAC website contains details of all available data through a fully searchable data dictionary and variable search tool (http://www.bristol.ac.uk/alspac/researchers/our-data/). Ethical approval for the study was obtained from the ALSPAC Ethics and Law Committee and the Local Research Ethics Committees. Consent for biological samples has been collected in accordance with the Human Tissue Act (2004). Informed consent for the use of data collected via questionnaires and clinics was obtained from participants following the recommendations of the ALSPAC Ethics and Law Committee at the time.

Study Population

In the Generation R Study, 9,778 pregnant mothers had 9,749 live-born children. For a subsample of 1,414 children DNA methylation data was collected at birth and/or 6 years and/or 10 years of age. This subsample consisted of participants with parents born in the Netherlands (European ancestry[82] confirmed for all children with genetic data available (95.4%)). Fifteen sibling pairs were present in the dataset. From each pair one sibling with the

lowest number of DNA methylation measurements, or otherwise randomly, was excluded, resulting in a sample with 1,399 children (with 2,333 DNA methylation samples; see below).

In the ALSPAC study, 15,247 pregnant mothers gave birth to 14,973 live-born children. DNA methylation at birth and/or 7 years and/or 17 years was available for a subsample of 1,003 children as part of the Accessible Resource for Integrated Epigenomic Studies (ARIES) study[83]. From this sample, 48 children with non-European ancestry as based on genetic principle component analysis and 6 children with missing data on gestational age were excluded, resulting in a sample of 949 children with DNA methylation data (with 2,686 DNA methylation samples; see below).

DNA methylation

Cord blood was drawn after birth for both cohorts, and peripheral blood was drawn at a mean age of 6.0 (*SD* = 0.47) and 9.8 (*SD* = 0.3) years for Generation R, and 7.5 (*SD* = 0.2) and 17.1 (*SD* = 1.0) years for ALSPAC. Both cohorts made use of the EZ-96 DNA methylation kit (shallow) (Zymo Research Corporation, Irvine, USA) to perform bisulfite conversion on the extracted leukocytic DNA. Samples were further processed with the Illumina Infinium HumanMethylation450 BeadChip (Illumina Inc., San Diego, USA) to analyze DNA methylation.

In Generation R, quality control was performed on all 2,467 available DNA methylation samples with the CPACOR workflow[84]. Arrays with observed technical problems such as failed bisulfite conversion, hybridization or extension, as well as arrays with a mismatch between sex of the proband and sex determined by the chromosome X and Y probe intensities were removed from subsequent analyses. Additionally, only arrays with a call rate >95% per sample were processed further, resulting in 2,355 samples, 22 of which belonged to half of an excluded sibling pair, hence 2,333 samples were carried forward into normalization.

In ALSPAC, quality control was performed on 6,057 samples (3,286 belonging to children, 2,771 to their mothers), using the *meffil* package[85] in R version 3.4.3[86]. After removing samples with mismatched genotypes, mismatched sex, incorrect relatedness, low concordance with samples collected at other time points, extreme dye bias, and poor probe detection, 5,337 samples remained, 2,845 of which belonging to children, used in the current study.

To minimize cohort effects as much as possible, we normalized both cohorts together as a single dataset. Functional normalization (10 control probe principal components, slide included as a random effect) was performed with the *meffil* package in R[85]. Normalization took place on the combined Generation R and ALSPAC set comprising a total of 5,178 samples for a total of 485,512 CpGs. One-hundred and fifty-nine ALSPAC samples belonging to non-European children or children with missing data on gestational age were excluded, leading to a final

ALSPAC set of 2,686 samples (for 949 children). Together with 2,333 samples for Generation R (of 1,399 children) they formed a combined set of 5,019 samples (of 2,348 children.)

Analyses were restricted to 473,864 autosomal CpGs. DNA methylation levels were operationalized as beta values (β values), representing the ratio of methylated signal relative to the sum of methylated and unmethylated signal measured per CpG.

Covariates

Sample plate number (*N* = 29 in Generation R and *N* = 31 in ALSPAC), was used to correct for batch effects, which was added as a random variable in the model (see below). White blood cell (WBC) composition was estimated with the reference-based Bakulski method[87] for cord blood and Houseman method[88] for peripheral blood (Supplemental Table 11). Nucleated red blood cells were not further analyzed due to its specificity to cord blood, leaving CD4+ T-lymphocytes, CD8+ T-lymphocytes, natural killer cells, B-lymphocytes, monocytes, and granulocytes. Other covariates included gestational age in weeks, sex of the child, and cohort.

Statistical analyses

Step 1: Assessing cross-cohort comparability in DNA methylation stability

To ascertain comparability amongst the two cohorts we compared within-cohort DNA methylation stability between the time points that were present in both cohorts – i.e. birth and 6/7 years (Generation R/ALSPAC, respectively).

Longitudinal stability per CpG within each cohort was assessed by studying estimates of concordance and change. For concordance, DNA methylation data was first residualized within each cohort for all variables present in the longitudinal models except the 'cohort' variable, in order to remove between-cohort differences due to other covariates. Concordance was then measured both with Spearman correlation (data at most CpGs is not normally distributed) as a measure of relative concordance, and with intra-class correlations as a measure of absolute concordance (children with data for both time points: *n* Generation R = 476, *n* ALSPAC = 826). Longitudinal change from birth to 6/7 years was assessed by studying the estimates of the change in DNA methylation per year by applying Model 1 (see below) within each cohort (children with data for at least one of the two time-points: *n* Generation R = 1,394, *n* ALSPAC = 944).

In a second step, cross-cohort comparability was assessed with Spearman (ρ) correlation of concordance estimates of the CpGs of each cohort (which were not normally distributed) and Pearson correlations (*r*) amongst the change estimates of the CpGs of each cohort (which were normally distributed).

Step 2: Longitudinal modelling of DNA methylation using combined Generation R and ALSPAC data

The combined Generation R and ALSPAC dataset had four time points of collection (birth, age 6/ 7 years, 10 years, and 17 years). We fit three linear mixed models to CpG site DNA methylation across the genome to identify (i) linear change over time (Model 1); (ii) nonlinear change over time (Model 2); and (iii) sex differences in change over time (Model 3). Both fixed and random effects were examined to allow for inter-individual variation in DNA methylation patterns over time. The models are described in detail below.

Model 1: Linear change. This model was applied to identify CpGs that show an overall change in DNA methylation from birth to 18 years (i.e. fixed age effect), as well as CpGs with inter-individual differences in change during that time (i.e. random age effect). The Model 1 is defined as:

M1: $$M_{ijk} = \beta_0 + u_{0i} + \beta_1 Age_{ij} + u_{1i} Age_{ij} + u_{0k} + \text{covariates} + \epsilon_{ijk}$$
$$\epsilon_{ijk} \sim N(0,\sigma_\epsilon^2)$$
$$u_{0i} \sim N(0,\sigma_{0i}^2)$$
$$u_{1i} \sim N(0,\sigma_{1i}^2)$$
$$u_{0k} \sim N(0,\sigma_{0k}^2)$$

Here, participants are denoted by *i*, time points by *j,* and sample plates by *k*. *M* denotes DNA methylation level, β_0 fixed intercept, u_{0i} random intercept, β_1 fixed age coefficient, u_{1i} random age coefficient, u_{0k} random intercept for sample plate. Hence, β_1 represents the average change in DNA methylation per one year. Variability in this change amongst individuals was captured with u_{1i}. To avoid problems with model identification, the random slope of age was uncorrelated to the random intercept (i.e. a diagonal random effects matrix was used).

Model 2: Nonlinear change. To identify nonlinear changes in DNA methylation, we extended Model 1 to allow slope changes at ages 6 and 9[30, 31]:

M2: $$M_{ijk} = \beta_0 + u_{0i} + \beta_1 Age_{ij} + \beta_2 (Age_{ij} - 6)^+ + \beta_3 (Age_{ij} - 9)^+ + u_{1i} Age_{ij} + u_{2i} (Age_{ij} - 6)^+ + u_{3i} (Age_{ij} - 9)^+ + u_{0k} + \text{covariates} + \epsilon_{ijk}$$
$$\epsilon_{ijk} \sim N(0,\sigma_\epsilon^2)$$
$$u_{0i} \sim N(0,\sigma_{0i}^2)$$
$$u_{1i} \sim N(0,\sigma_{1i}^2)$$
$$u_{2i} \sim N(0,\sigma_{2i}^2)$$
$$u_{3i} \sim N(0,\sigma_{3i}^2)$$
$$u_{0k} \sim N(0,\sigma_{0k}^2)$$

Where $a^+ = a$ if $a>0$ and 0 otherwise, so that β_2 represents the average change in DNA methylation per year from 6 years of age onward, after accounting for the change per year from birth onward, as denoted by β_1. Likewise, $\beta_3$represents the average change in DNA methylation per year from 9 years of age onward, after accounting for the change per year from 6 years of age onward. Hence, with those variables we are able to detect slope changes at 6 and 9 years old. These slope changes were used to identify different types of nonlinear patterns. With u_{2i} and u_{3i} the inter-individual variation in slope changes at 6 and 9 years were captured, respectively. General linear hypothesis testing[89] was applied to our fitted models to determine if there were changes in DNA methylation per year from 6-9 years and from 9-18 years.

Model 3: Sex differences in change: To identify CpGs for which DNA methylation changes differently over time for boys and girls, we applied the following model:

M3: $$M_{ijk} = \beta_0 + u_{0i} + \beta_1 Age_{ij} + u_{1i} Age_{ij} + \beta_2 Sex_i Age_{ij} + u_{0k} + \text{covariates} + \epsilon_{ijk}$$
$$\epsilon_{ijk} \sim N(0,\sigma_\epsilon^2)$$
$$u_{0i} \sim N(0,\sigma_u^2)$$
$$u_{1i} \sim N(0,\sigma_{1i}^2)$$
$$u_{0k} \sim N(0,\sigma_{0k}^2)$$

Here, Sex_i denotes the sex of child *i*. Both main and interaction effects for sex were studied.

The three mixed models were fitted using maximum likelihood estimation in *R* with the *lme4* package[90]. Continuous covariates (WBCs, gestational age) were z-score standardized. Random slopes were kept uncorrelated with random intercepts and the NLopt optimizer was used, enabling us to improve computational speed compared to the default settings. *P*-values for the fixed effects were computed with a z-test. *P*-values for random slopes of the Age effects were obtained by refitting the model without the random slope and comparing the fit estimates of the two models with a likelihood ratio test. Within each model, *P*-value thresholds were Bonferroni-corrected for the number of tested CpGs (i.e. to $p < 1x10^{-07}$).

Step 3: Functional characterization of probes with comparable patterns of change
To interpret the functionality of the age-related DNA methylation patterns from the three models, CpG sites adhering to 8 different age-related patterns (M1 linear change and inter-individual variation in linear change, M2 nonlinear trajectories, and inter-individual variation in change from birth, in slope change at 6 years, and in slope change at 9 years, and M3 stable sex differences and sex differences in DNA methylation change) were tested for enrichment in:

(i) gene-relative genomic regions (TSS1500, TSS200, 5'UTR, 1st exon, gene body, 3'UTR, and intergenic regions[64]),

(ii) CpG island-relative genomic regions (N shelf, N shore, CpG island, S shore, S shelf, and open sea regions[64]) as indicated by the Illumina HumanMethylation450 v1.2 Manifest (Illumina Inc., San Diego, USA), and
(iii) enhancer elements as those expressed in whole blood, peripheral blood mononuclear cells, natural killer cells, CD4+ T cells, CD8+ T cells, monocytes, neutrophils, eosinophils or B cells[91],

Altogether, these encompassed 14 enrichment analyses for 8 variables. Enrichment was tested using χ^2-tests of unequal proportions. The enrichment *P*-value threshold was Bonferroni-corrected for multiple tests (i.e. $p < 4.46x10^{-04}$ for 8x14 = 122 tests). Second, we tested enrichment of Gene Ontology (GO) categories for genes linked to CpG sites surviving adjustment for multiple tests ($p < 1x10^{-07}$) for each of the main variables of interest. This analysis was adjusted for gene size and pruned for near-identical terms (see elsewhere for a full description[92]). For completeness, terms with nominal $p < 0.05$ were reported. Last, we tested enrichment of age-related DNA methylation trajectories (11 different age-related patterns: M1 decreasing, increasing, and inter-individual variation in linear change, M2 Positive-Neutral, Negative-Neutral, other nonlinear, inter-individual variation in change from birth, in slope change at 6 years, and in slope change at 9 years, and M3 stable sex differences and sex differences in DNA methylation change) in EWASs on age, prenatal smoking, smoking, cardiovascular-associated traits, C-reactive protein, allergies, educational attainment, and cellular heterogeneity. EWAS summary statistics were retrieved from the EWAS Catalog (http://www.ewascatalog.org/) and studies were included when performed with the 450K array in peripheral or cord blood, resulting in 33 EWASs. Enrichment was tested with Fisher's exact tests, the enrichment *P*-value threshold was Bonferroni-corrected for multiple tests (i.e. $p < 1.38x10^{-04}$ for 11x33 = 363 tests).

References

1. Khavari DA, Sen GL, Rinn JL. DNA methylation and epigenetic control of cellular differentiation. *Cell Cycle.* 2010;9:3880-3.
2. Meissner A, Mikkelsen TS, Gu H, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature.* 2008;454:766-70.
3. Fraser A, Macdonald-Wallis C, Tilling K, et al. Cohort profile: the Avon Longitudinal Study of Parents and Children: ALSPAC mothers cohort. *International Journal of Epidemiology.* 2012;42:97-110.
4. Chow J, Heard E. X inactivation and the complexities of silencing a sex chromosome. *Current Opinion in Cell Biology.* 2009;21:359-66.
5. Koukoura O, Sifakis S, Spandidos DA. DNA methylation in the human placenta and fetal growth. *Molecular Medicine Reports.* 2012;5:883-9.
6. Eden A, Gaudet F, Waghmare A, Jaenisch R. Chromosomal instability and tumors promoted by DNA hypomethylation. *Science.* 2003;300:455-0.
7. Comet I, Riising EM, Leblanc B, Helin K. Maintaining cell identity: PRC2-mediated regulation of transcription and cancer. *Nature Reviews Cancer.* 2016;16:803.
8. Natoli G. Maintaining cell identity through global control of genomic organization. *Immunity.* 2010;33:12-24.
9. Bojesen SE, Timpson N, Relton C, Smith GD, Nordestgaard BG. AHRR (cg05575921) hypomethylation marks smoking behaviour, morbidity and mortality. *Thorax.* 2017;72:646-53.
10. Cao-Lei L, Massart R, Suderman MJ, et al. DNA methylation signatures triggered by prenatal maternal stress exposure to a natural disaster: Project ice storm. *PLoS ONE.* 2014;9.
11. Baccarelli A, Wright RO, Bollati V, et al. Rapid DNA methylation changes after exposure to traffic particles. *American Journal of Respiratory and Critical Care Medicine.* 2009;179:572-8.
12. Suarez-Alvarez B, Rodriguez RM, Fraga MF, López-Larrea C. DNA methylation: A promising landscape for immune system-related diseases. *Trends in Genetics.* 2012;28:506-14.
13. Szulwach KE, Li X, Li Y, et al. 5-hmC–mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nature Neuroscience.* 2011;14:1607.
14. Gaunt TR, Shihab HA, Hemani G, et al. Systematic identification of genetic influences on methylation across the human life course. *Genome Biology.* 2016;17:61.
15. Feil R, Fraga MF. Epigenetics and the environment: emerging patterns and implications. *Nature Reviews Genetics.* 2012;13:97.
16. Bell JT, Tsai P-C, Yang T-P, et al. Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS Genetics.* 2012;8:e1002629.
17. Hernandez DG, Nalls MA, Gibbs JR, et al. Distinct DNA methylation changes highly correlated with chronological age in the human brain. *Human Molecular Genetics.* 2011;20:1164-72.
18. Alisch RS, Barwick BG, Chopra P, et al. Age-associated DNA methylation in pediatric populations. *Genome Research.* 2012;22:623-32.
19. Wang Y, Karlsson R, Lampa E, et al. Epigenetic influences on aging: a longitudinal genome-wide methylation study in old Swedish twins. *Epigenetics.* 2018;13:975-87.

20. Wang Y, Pedersen NL, Hägg S. Implementing a method for studying longitudinal DNA methylation variability in association with age. *Epigenetics.* 2018;13:866-74.
21. Lévesque ML, Casey KF, Szyf M, et al. Genome-wide DNA methylation variability in adolescent monozygotic twins followed since birth. *Epigenetics.* 2014;9:1410-22.
22. Martino DJ, Loke YJ, Gordon L, et al. Longitudinal, genome-scale analysis of DNA methylation in twins from birth to 18 months of age reveals rapid epigenetic change in early life and pair-specific effects of discordance. *Genome Biology.* 2013;14:R42.
23. Martino DJ, Tulic MK, Gordon L, et al. Evidence for age-related and individual-specific changes in DNA methylation profile of mononuclear cells during early immune development in humans. *Epigenetics.* 2011;6:1085-94.
24. Fitzmaurice GM, Laird NM, Ware JH. *Applied longitudinal analysis*. John Wiley & Sons, 2012.
25. Ng JWY, Barrett LM, Wong A, Kuh D, Smith GD, Relton CL. The role of longitudinal cohort studies in epigenetic epidemiology: challenges and opportunities. *Genome Biology.* 2012;13:246.
26. Naumova EN, Must A, Laird NM. Tutorial in biostatistics: evaluating the impact of 'critical periods' in longitudinal studies of growth using piecewise mixed effects models. *International Journal of Epidemiology.* 2001;30:1332-41.
27. Cudeck R, Harring JR. Analysis of nonlinear patterns of change with random coefficient models. *Annual Reviews in Psychology.* 2007;58:615-37.
28. Xu C-J, Bonder MJ, Söderhäll C, et al. The emerging landscape of dynamic DNA methylation in early childhood. *BMC Genomics.* 2017;18:25.
29. Richmond RC, Simpkin AJ, Woodward G, et al. Prenatal exposure to maternal smoking and offspring DNA methylation across the lifecourse: findings from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Human Molecular Genetics.* 2014;24:2201-17.
30. Simpkin AJ, Suderman M, Gaunt TR, et al. Longitudinal analysis of DNA methylation associated with birth weight and gestational age. *Human Molecular Genetics.* 2015;24:3752-63.
31. Sharp GC, Lawlor DA, Richmond RC, et al. Maternal pre-pregnancy BMI and gestational weight gain, offspring DNA methylation and later offspring adiposity: findings from the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology.* 2015;44:1288-304.
32. Ball MP, Li JB, Gao Y, et al. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nature Biotechnology.* 2009;27:361.
33. Wagner JR, Busche S, Ge B, Kwan T, Pastinen T, Blanchette M. The relationship between DNA methylation, genetic and expression inter-individual variation in untransformed human fibroblasts. *Genome Biology.* 2014;15:R37.
34. Illingworth RS, Bird AP. CpG islands–'a rough guide'. *FEBS Letters.* 2009;583:1713-20.
35. Ong C-T, Corces VG. Enhancer function: new insights into the regulation of tissue-specific gene expression. *Nature Reviews Genetics.* 2011;12:283.
36. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. *Genome Biology.* 2014;15:R31.
37. Bohlin J, Håberg SE, Magnus P, et al. Prediction of gestational age based on genome-wide differentially methylated regions. *Genome Biology.* 2016;17:207.

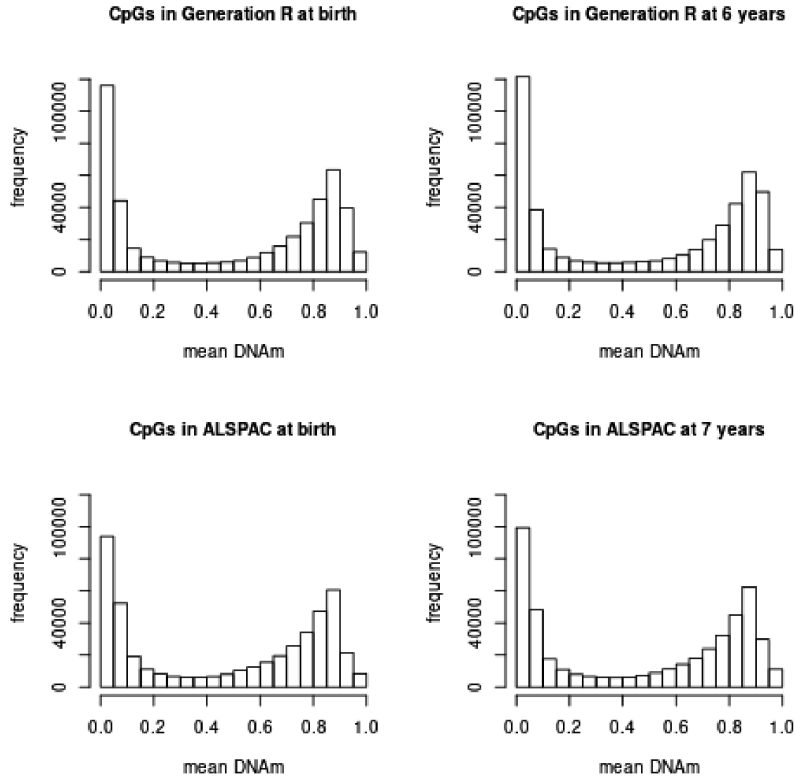
38. Hedman ÅK, Mendelson MM, Marioni RE, et al. Epigenetic patterns in blood associated with lipid traits predict incident coronary heart disease events and are enriched for results from genome-wide association studies. *Circulation: Cardiovascular Genetics.* 2017;10:e001487.
39. Ligthart S, Marzi C, Aslibekyan S, et al. DNA methylation signatures of chronic low-grade inflammation are associated with complex diseases. *Genome Biology.* 2016;17:255.
40. Mendelson MM, Marioni RE, Joehanes R, et al. Association of body mass index with DNA methylation and gene expression in blood cells and relations to cardiometabolic disease: a Mendelian randomization approach. *PLoS Medicine.* 2017;14.
41. Demerath EW, Guan W, Grove ML, et al. Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. *Human Molecular Genetics.* 2015;24:4464-79.
42. Joehanes R, Just AC, Marioni RE, et al. Epigenetic signatures of cigarette smoking. *Circulation: Cardiovascular Genetics.* 2016;9:436-47.
43. Sayols-Baixeras S, Lluís-Ganella C, Subirana I, et al. Identification of a new locus and validation of previously reported loci showing differential methylation associated with smoking. The REGICOR study. *Epigenetics.* 2015;10:1156-65.
44. Sayols-Baixeras S, Subirana I, Lluis-Ganella C, et al. Identification and validation of seven new loci showing differential DNA methylation related to serum lipid profile: an epigenome-wide approach. The REGICOR study. *Human Molecular Genetics.* 2016;25:4556-65.
45. Singmann P, Shem-Tov D, Wahl S, et al. Characterization of whole-genome autosomal differences of DNA methylation between men and women. *Epigenetics & Chromatin.* 2015;8:43.
46. Joubert BR, Felix JF, Yousefi P, et al. DNA methylation in newborns and maternal smoking in pregnancy: genome-wide consortium meta-analysis. *The American Journal of Human Genetics.* 2016;98:680-96.
47. Zeilinger S, Kühnel B, Klopp N, et al. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS ONE.* 2013;8:e63812.
48. Guida F, Sandanger TM, Castagné R, et al. Dynamics of smoking-induced genome-wide methylation changes with time since smoking cessation. *Human Molecular Genetics.* 2015;24:2349-59.
49. Ambatipudi S, Cuenin C, Hernandez-Vargas H, et al. Tobacco smoking-associated genome-wide DNA methylation changes in the EPIC study. *Epigenomics.* 2016;8:599-618.
50. Chatziioannou A, Georgiadis P, Hebels DG, et al. Blood-based omic profiling supports female susceptibility to tobacco smoke-induced cardiovascular diseases. *Scientific Reports.* 2017;7:42870.
51. Linnér RK, Marioni RE, Rietveld CA, et al. An epigenome-wide association study meta-analysis of educational attainment. *Molecular Psychiatry.* 2017;22:1680-90.
52. Zhu X, Li J, Deng S, et al. Genome-wide analysis of DNA methylation and cigarette smoking in a Chinese population. *Environmental Health Perspectives.* 2016;124:966-73.
53. Chen W, Wang T, Pino-Yanes M, et al. An epigenome-wide association study of total serum IgE in Hispanic children. *Journal of Allergy and Clinical Immunology.* 2017;140:571-7.
54. Yousefi P, Huen K, Davé V, Barcellos L, Eskenazi B, Holland N. Sex differences in DNA methylation assessed by 450 K BeadChip in newborns. *BMC Genomics.* 2015;16:911.

55. Shah S, McRae AF, Marioni RE, et al. Genetic and environmental exposures constrain epigenetic drift over the human life course. *Genome Research.* 2014;24:1725-33.
56. Markunas CA, Xu Z, Harlid S, et al. Identification of DNA methylation changes in newborns related to maternal smoking during pregnancy. *Environmental Health Perspectives.* 2014;122:1147-53.
57. Joubert BR, Herman T, Felix JF, et al. Maternal plasma folate impacts differential DNA methylation in an epigenome-wide meta-analysis of newborns. *Nature Communications.* 2016;7:10577.
58. Marttila S, Kananen L, Häyrynen S, et al. Ageing-associated changes in the human DNA methylome: genomic locations and effects on gene expression. *BMC Genomics.* 2015;16:179.
59. Reynolds LM, Taylor JR, Ding J, et al. Age-related variations in the methylome associated with gene expression in human monocytes and T cells. *Nature Communications.* 2014;5:1-8.
60. Florath I, Butterbach K, Müller H, Bewerunge-Hudler M, Brenner H. Cross-sectional and longitudinal changes in DNA methylation with age: an epigenome-wide analysis revealing over 60 novel age-associated CpG sites. *Human Molecular Genetics.* 2014;23:1186-201.
61. Li S, Christiansen L, Christensen K, et al. Identification, replication and characterization of epigenetic remodelling in the aging genome: a cross population analysis. *Scientific Reports.* 2017;7:1-8.
62. Johansson Å, Enroth S, Gyllensten U. Continuous aging of the human DNA methylome throughout the human lifespan. *PLoS ONE.* 2013;8:e67378.
63. Felix JF, Joubert BR, Baccarelli AA, et al. Cohort profile: pregnancy and childhood epigenetics (PACE) Consortium. *International Journal of Epidemiology.* 2017;47:22-3u.
64. Irizarry RA, Ladd-Acosta C, Wen B, et al. The human colon cancer methylome shows similar hypo-and hypermethylation at conserved tissue-specific CpG island shores. *Nature Genetics.* 2009;41:178.
65. Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nature Reviews Genetics.* 2012;13:484.
66. Voigt M, Hermanussen M, Wittwer-Backofen U, Fusch C, Hesse V. Sex-specific differences in birth weight due to maternal smoking during pregnancy. *European Journal of Pediatrics.* 2006;165:757-61.
67. Paus T, Nawazkhan I, Leonard G, et al. Corpus callosum in adolescent offspring exposed prenatally to maternal cigarette smoking. *Neuroimage.* 2008;40:435-41.
68. Toro R, Leonard G, Lerner JV, et al. Prenatal exposure to maternal cigarette smoking and the adolescent cerebral cortex. *Neuropsychopharmacology.* 2008;33:1019-27.
69. Jacobsen LK, Slotkin TA, Mencl WE, Frost SJ, Pugh KR. Gender-specific effects of prenatal and adolescent exposure to tobacco smoke on auditory and visual attention. *Neuropsychopharmacology.* 2007;32:2453-64.
70. Richmond RC, Suderman M, Langdon R, Relton CL, Davey Smith G. DNA methylation as a marker for prenatal smoke exposure in adults. *International Journal of Epidemiology.* 2018;47:1120-30.
71. Kupers LK, Xu X, Jankipersadsing SA, et al. DNA methylation mediates the effect of maternal smoking during pregnancy on birthweight of the offspring. *International Journal of Epidemiology.* 2015.
72. Nielsen CH, Larsen A, Nielsen AL. DNA methylation alterations in response to prenatal exposure of maternal cigarette smoking: A persistent epigenetic impact on health from maternal lifestyle? *Archives of Toxicology.* 2014.

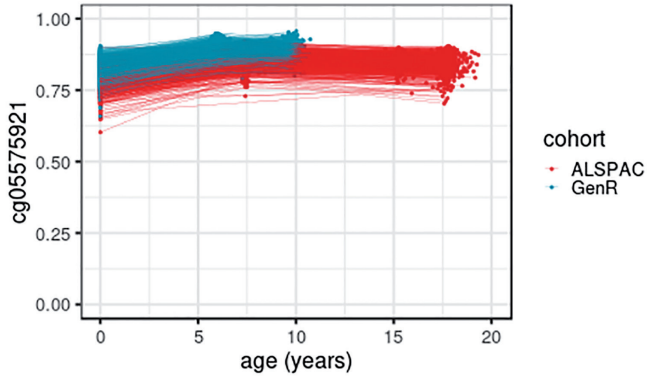
73. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nature Reviews Genetics.* 2018;19:371.
74. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biology.* 2013;14:3156.
75. Hannum G, Guinney J, Zhao L, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Molecular Cell.* 2013;49:359-67.
76. Heyn H, Moran S, Hernando-Herraez I, et al. DNA methylation contributes to natural human variation. *Genome Research.* 2013;23:1363-72.
77. Barfield RT, Almli LM, Kilaru V, et al. Accounting for population stratification in DNA methylation studies. *Genetic Epidemiology.* 2014;38:231-41.
78. Fraser HB, Lam LL, Neumann SM, Kobor MS. Population-specificity of human DNA methylation. *Genome Biology.* 2012;13:R8.
79. Gopalan S, Carja O, Fagny M, et al. Trends in DNA methylation with age replicate across diverse human populations. *Genetics.* 2017;206:1659-74.
80. Kooijman MN, Kruithof CJ, van Duijn CM, et al. The Generation R Study: design and cohort update 2017. *European Journal of Epidemiology.* 2016;31:1243-64.
81. Boyd A, Golding J, Macleod J, et al. Cohort profile: the 'children of the 90s'—the index offspring of the Avon Longitudinal Study of Parents and Children. *International Journal of Epidemiology.* 2013;42:111-27.
82. Medina-Gomez C, Felix JF, Estrada K, et al. Challenges in conducting genome-wide association studies in highly admixed multi-ethnic populations: the Generation R Study. *European Journal of Epidemiology.* 2015;30:317-30.
83. Relton CL, Gaunt T, McArdle W, et al. Data resource profile: accessible resource for integrated epigenomic studies (ARIES). *International Journal of Epidemiology.* 2015;44:1181-90.
84. Lehne B, Drong AW, Loh M, et al. A coherent approach for analysis of the Illumina HumanMethylation450 BeadChip improves data quality and performance in epigenome-wide association studies. *Genome Biology.* 2015;16:37.
85. Min JL, Hemani G, Davey Smith G, Relton C, Suderman M, Hancock J. Meffil: efficient normalization and analysis of very large DNA methylation datasets. *Bioinformatics.* 2018.
86. R Core Team. *R: A language and environment for statistical computing.* 2013.
87. Bakulski KM, Feinberg JI, Andrews SV, et al. DNA methylation of cord blood cell types: applications for mixed cell birth studies. *Epigenetics.* 2016;11:354-62.
88. Houseman EA, Accomando WP, Koestler DC, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics.* 2012;13:1.
89. Hothorn T, Bretz F, Westfall P, et al. Package 'multcomp'. Simultaneous inference in general parametric models. *Project for Statistical Computing*, Vienna, Austria 2016.
90. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *arXiv preprint* arXiv:1406.5823 2014.
91. Andersson R, Gebhard C, Miguel-Escalada I, et al. An atlas of active enhancers across human cell types and tissues. *Nature.* 2014;507:455-61.

92. Cecil CAM, Smith RG, Walton E, Mill J, McCrory EJ, Viding E. Epigenetic signatures of childhood abuse and neglect: implications for psychiatric vulnerability. *Journal of Psychiatric Research.* 2016;83:184-94.

Supplemental Figures



Supplemental Figure 1. Distribution of CpG mean methylation levels in (a) Generation R at birth; (b) ALSPAC at birth; (c) Generation R 6 years; (d) ALSPAC at 7 years



Supplemental Figure 2. Longitudinal DNA methylation of *AHRR* cg05575921 as predicted by Model 2

Supplemental Tables

Supplemental Table 1. Cohort characteristics

	Generation R (*N* = 1399)	**ALSPAC (*N* = 949)**
Child characteristics		
Sex (% boys)	50.8	48.7
Gestational age (weeks)	40.1 (1.5)	39.6 (1.5)
Age (years)		
@6 or 7 years, n = 488 / 970	6.0 (0.4)	7.5 (0.1)
@10 years, n = 463	9.8 (0.3)	
@17 years, n = 70		17.1 (1.0)
Birth weight (grams)	3546 (510)	3489 (490)
BMI		
@6 or 7 years, n = 488 / 970	15.9 (1.3)	16.2 (2.0)
@10 years, n = 463	17.1 (2.0)	
@17 years, n = 970		22.5 (3.8)
Mother characteristics		
Maternal age at birth (years)	32.2 (4.2)	30.0 (4.4)
BMI early pregnancy	24.2 (4.0)	22.8 (3.7)
Education level (%)		
low	11.2	8.6
medium	23.5	41.1
high	65.3	50.3
Prenatal smoking (% sustained)	13.3	10.2

Supplemental Tables 3A-H. CpG characteristics per gene region compared to CpGs at all other regions

Supplemental Table 3A. M1 linear DNA methylation change

	CpGs in area	other CpGs			
	% null/negative/ positive	% null/negative/ positive	χ^2	df	*p-value*
TSS1500	48.63/31.78/19.59	48.41/36.30/15.29	1164.52	2	$1.34x10^{-253}$
TSS200	52.79/28.19/19.02	47.82/36.59/15.60	1694.99	2	$<9.88x10^{-324}$
5'UTR	50.50/30.06/19.44	48.13/36.36/15.51	1200.84	2	$1.7 x10^{-261}$
1st exon	51.63/25.35/23.02	48.17/36.4/15.43	2527.50	2	$<9.88x10^{-324}$
Gene body	48.31/38.75/12.94	48.53/33.67/17.80	2420.25	2	$<9.88x10^{-324}$
3'UTR	48.03/42.77/9.20	48.47/35.21/16.32	880.58	2	$6.08x10^{-192}$
Intergenic	45.71/38.73/15.56	49.35/34.46/16.19	718.15	2	$1.14x10^{-156}$

Supplemental Table 3B. M1 inter-individual variation in DNA methylation change

	CpGs in area	other CpGs			
	% no/yes	% no/yes	χ^2	df	*p-value*
TSS1500	77.10/22.90	71.71/28.29	984.90	1	$3.43x10^{-216}$
TSS200	77.76/22.24	71.89/28.11	908.37	1	$1.49x10^{-119}$
5'UTR	76.11/23.89	72.10/27.9	442.63	1	$2.90x10^{-98}$
1st exon	76.33/23.67	72.31/27.69	281.83	1	$3.00x10^{-63}$
Gene body	71.14/28.86	73.49/26.51	305.25	1	$2.36x10^{-68}$
3'UTR	72.79/27.21	72.63/27.37	0.24	1	$6.24x10^{-01}$
Intergenic	69.42/30.58	73.70/26.3	814.75	1	$3.35x10^{-179}$

Supplemental Table 3C. M2 nonlinear DNA methylation change

	CpGs in area	other CpGs			
	% linear or null/ Positive-Neutral/ Negative-Neutral/ other nonlinear	% linear or null/ Positive-Neutral/ Negative-Neutral/ other nonlinear	χ^2	df	*p-value*
TSS1500	88.55/4.70/3.77/2.98	89.11/4.84/3.00/3.05	132.58	3	$1.50x10^{-28}$
TSS200	92.85/3.14/1.38/2.63	88.46/5.06/3.39/3.09	1227.78	3	$6.90x10^{-266}$
5'UTR	89.73/4.13/3.06/3.08	88.91/4.92/3.15/3.03	75.95	3	$2.27x10^{-16}$
1st exon	92.11/3.21/1.56/3.12	88.75/4.95/3.27/3.03	589.94	3	$1.53x10^{-127}$
Gene body	88.40/4.84/3.66/3.10	89.37/4.80/2.84/2.99	252.42	3	$1.96x10^{-54}$
3'UTR	88.91/4.64/3.40/3.05	89.02/4.82/3.12/3.03	5.89	3	$1.17x10^{-01}$
Intergenic	88.12/5.82/2.92/3.14	89.31/4.48/3.21/3.00	371.72	3	$2.96x10^{-80}$

Supplemental Table 3D. M2 inter-individual variation in DNA methylation change from birth

	CpGs in area	other CpGs			
	% no/yes	% no/yes	χ^2	df	*p-value*
TSS1500	96.99/3.01	96.58/3.42	35.70	3	$2.30x10^{-09}$
TSS200	95.98/4.02	96.75/3.25	95.82	3	$1.26x10^{-22}$
5'UTR	96.20/3.80	96.72/3.28	45.97	3	$1.20x10^{-11}$
1st exon	95.09/4.91	96.79/3.21	308.77	3	$4.05x10^{-69}$
Gene body	97.04/2.96	96.43/3.57	128.26	3	$9.82x10^{-30}$
3'UTR	97.89/2.11	96.60/3.40	94.92	3	$1.98x10^{-22}$
Intergenic	96.54/3.46	96.69/3.31	6.45	3	$1.11x10^{-02}$

Supplemental Table 3E. M2 inter-individual variation in DNA methylation change from 6 years

	CpGs in area	other CpGs			
	% no/yes	% no/yes	χ^2	df	*p-value*
TSS1500	99.83/0.17	99.83/0.17	0.21	2	$6.49x10^{-01}$
TSS200	99.64/0.36	99.86/0.14	149.48	2	$2.25x10^{-34}$
5'UTR	99.72/0.28	99.85/0.15	56.08	2	$6.96x10^{-14}$
1st exon	99.47/0.53	99.86/0.14	310.00	2	$2.18x10^{-69}$
Gene body	99.90/0.10	99.80/0.20	63.49	2	$1.61x10^{-15}$
3'UTR	99.94/0.06	99.83/0.17	13.93	2	$1.90x10^{-04}$
Intergenic	99.9/0.10	99.81/0.19	43.23	2	$4.88x10^{-11}$

Supplemental Table 3F. M2 inter-individual variation in DNA methylation change from 9 years

	CpGs in area	other CpGs			
	% no/yes	% no/yes	χ^2	df	*p-value*
TSS1500	93.34/6.66	91.52/8.48	296.39	2	$2.01x10^{-66}$
TSS200	91.47/8.53	91.89/8.11	12.44	2	$4.21x10^{-04}$
5'UTR	92.08/7.92	91.80/8.20	5.76	2	$1.64x10^{-02}$
1st exon	90.79/9.21	91.92/8.08	59.47	2	$1.24x10^{-14}$
Gene body	91.18/8.82	92.21/7.79	156.20	2	$7.64x10^{-36}$
3'UTR	92.86/7.14	91.79/8.21	28.01	2	$1.20x10^{-07}$
Intergenic	92.33/7.67	91.67/8.33	51.46	2	$7.30x10^{-13}$

Supplemental Table 3G. M3 stable sex differences in DNA methylation

	CpGs in area	other CpGs			
	% no/yes	% no/yes	χ^2	df	*p-value*
TSS1500	94.49/5.51	95.33/4.67	544.43	2	$5.99x10^{-119}$
TSS200	96.66/3.34	94.97/5.03	400.83	2	$9.14x10^{-88}$
5'UTR	96.41/3.59	95.00/5.00	237.15	2	$3.19x10^{-52}$
1st exon	96.23/3.77	95.09/4.91	100.04	2	$1.89x10^{-22}$
Gene body	96.12/3.88	94.65/5.35	755.26	2	$9.93x10^{-165}$
3'UTR	96.45/3.55	95.13/4.87	103.99	2	$2.63x10^{-23}$
Intergenic	93.18/6.82	95.85/4.15	1377.23	2	$8.68x10^{-300}$

Supplemental Table 3H. M3 sex differences in DNA methylation change

	CpGs in area	other CpGs			
	% no/yes	% no/yes	χ^2	df	*p-value*
TSS1500	99.64/0.36	99.62/0.38	7.67	2	2.16×10^{-02}
TSS200	99.73/0.27	99.61/0.39	25.52	2	2.88×10^{-06}
5'UTR	99.68/0.32	99.62/0.38	6.16	2	4.60×10^{-02}
1st exon	99.76/0.24	99.62/0.38	28.74	2	5.74×10^{-07}
Gene body	99.65/0.35	99.62/0.38	7.59	2	2.25×10^{-02}
3'UTR	99.70/0.30	99.62/0.38	2.96	2	2.28×10^{-01}
Intergenic	99.51/0.49	99.66/0.34	54.36	2	1.57×10^{-12}

Supplemental Tables 4A-H. Chi-square of CpG characteristics for each CpG island region compared to CpG characteristics in all other regions

Supplemental Table 4A. M1 linear DNA methylation change

	CpGs in area	other CpGs			
	% null/negative/ positive	% null/negative/ positive	χ^2	df	*p-value*
N shelf	50.09/42.08/7.83	48.41/36.3/15.29	1395.43	2	$9.68x10^{-304}$
N shore	49.11/31.25/19.64	47.82/36.59/15.60	940.13	2	$7.13x10^{-205}$
Island	51.82/23.00/25.18	48.13/36.36/15.51	20798.58	2	$<9.88x10^{-324}$
S shore	49.26/31.66/19.08	48.17/36.40/15.43	540.32	2	$4.68x10^{-118}$
S shelf	50.09/42.08/7.83	48.53/33.67/17.80	1344.60	2	$1.05x10^{-292}$
Open sea	44.67/46.91/8.42	48.47/35.21/16.32	20603.72	2	$<9.88x10^{-324}$

Supplemental Table 4B. M1 inter-individual variation in DNA methylation change

	CpGs in area	other CpGs			
	% no/yes	% no/yes	χ^2	df	*p-value*
N shelf	75.05/24.95	71.71/28.29	74.74	1	$5.38x10^{-18}$
N shore	73.98/26.02	71.89/28.11	63.72	1	$1.44x10^{-15}$
Island	71.40/28.60	72.10/27.90	160.31	1	$9.69x10^{-37}$
S shore	74.26/25.74	72.31/27.69	70.67	1	$4.23x10^{-17}$
S shelf	75.05/24.95	73.49/26.51	87.17	1	$9.97x10^{-21}$
Open sea	72.06/27.94	72.63/27.37	44.92	1	$2.05x10^{-11}$

Supplemental Table 4C. M2 nonlinear DNA methylation change

	CpGs in area	other CpGs			
	% linear or null/ Positive-Neutral/ Negative-Neutral/ other nonlinear	% linear or null/ Positive-Neutral/ Negative-Neutral/ other nonlinear	χ^2	df	*p-value*
N shelf	88.51/5.04/3.89/2.56	89.04/4.80/3.10/3.06	68.65	3	$8.31x10^{-15}$
N shore	86.10/6.79/4.50/2.62	89.45/4.52/2.93/3.10	1090.63	3	$3.94x10^{-236}$
Island	93.10/3.25/0.79/2.87	87.20/5.51/4.18/3.11	5182.32	3	$<9.88x10^{-324}$
S shore	85.75/6.86/4.73/2.66	89.38/4.58/2.96/3.08	984.09	3	$5.08x10^{-213}$
S shelf	88.51/5.04/3.89/2.56	89.04/4.80/3.10/3.06	49.85	3	$8.58x10^{-11}$
Open sea	87.56/4.85/4.02/3.57	89.85/4.79/2.63/2.73	992.66	3	$7.03x10^{-215}$

Supplemental Table 4D. M2 inter-individual variation in DNA methylation change from birth

	CpGs in area	other CpGs			
	% no/yes	% no/yes	χ^2	df	*p-value*
N shelf	98.09/1.91	96.58/3.42	163.36	3	$2.09x10^{-37}$
N shore	96.90/3.10	96.75/3.25	13.13	3	$2.91x10^{-04}$
Island	94.67/5.33	96.72/3.28	2544.09	3	$<9.88x10^{-324}$
S shore	96.79/3.21	96.79/3.21	3.25	3	$7.15x10^{-02}$
S shelf	98.09/1.91	96.43/3.57	197.61	3	$6.94x10^{-45}$
Open sea	97.78/2.22	96.60/3.40	1070.38	3	$9.06x10^{-235}$

Supplemental Table 4E. M2 inter-individual variation in DNA methylation change from 6 years

	CpGs in area	other CpGs			
	% no/yes	% no/yes	χ^2	df	*p-value*
N shelf	99.97/0.03	99.83/0.17	26.83	2	$2.22x10^{-07}$
N shore	99.94/0.06	99.86/0.14	44.75	2	$2.24x10^{-11}$
Island	99.55/0.45	99.85/0.15	1027.50	2	$1.89x10^{-225}$
S shore	99.93/0.07	99.86/0.14	27.85	2	$1.31x10^{-07}$
S shelf	99.97/0.03	99.80/0.20	24.15	2	$8.93x10^{-07}$
Open sea	99.97/0.03	99.83/0.17	328.18	2	$2.40x10^{-73}$

Supplemental Table 4F. M2 inter-individual variation in DNA methylation change from 9 years

	CpGs in area	other CpGs			
	% no/yes	% no/yes	χ^2	df	*p-value*
N shelf	94.58/5.42	91.52/8.48	255.42	2	$1.71x10^{-57}$
N shore	92.96/7.04	91.89/8.11	118.35	2	$1.45x10^{-27}$
Island	88.40/11.60	91.80/8.20	3305.19	2	$<9.88x10^{-324}$
S shore	92.97/7.03	91.92/8.08	91.09	2	$1.37x10^{-21}$
S shelf	94.58/5.42	92.21/7.79	244.06	2	$5.13x10^{-55}$
Open sea	93.27/6.73	91.79/8.21	753.18	2	$8.18x10^{-166}$

Supplemental Table 4G. M3 stable sex differences in DNA methylation

	CpGs in area	other CpGs			
	% no/yes	% no/yes	χ^2	df	*p-value*
N shelf	96.54/3.46	95.33/4.67	104.49	2	$2.04x10^{-23}$
N shore	92.65/7.35	94.97/5.03	1214.08	2	$2.32x10^{-264}$
Island	95.46/4.54	95.00/5.00	99.69	2	$2.25x10^{-22}$
S shore	92.24/7.76	95.09/4.91	1379.57	2	$2.69x10^{-300}$
S shelf	96.54/3.46	94.65/5.35	94.48	2	$3.05x10^{-21}$
Open sea	96.31/3.69	95.13/4.87	1553.00	2	$<9.88x10^{-324}$

Supplemental Table 4H. M3 sex differences in DNA methylation change

	CpGs in area	other CpGs			
	% no/yes	% no/yes	χ^2	df	*p-value*
N shelf	99.62/0.38	99.62/0.38	0.37	2	$8.29x10^{-01}$
N shore	99.45/0.55	99.61/0.39	66.45	2	$3.72x10^{-15}$
Island	99.74/0.26	99.62/0.38	89.21	2	$4.26x10^{-20}$
S shore	99.36/0.64	99.62/0.38	115.80	2	$7.17x10^{-26}$
S shelf	99.62/0.38	99.62/0.38	4.16	2	$1.25x10^{-01}$
Open sea	99.68/0.32	99.62/0.38	21.98	2	$1.68x10^{-05}$

Supplemental Tables 5. CpGs inside and outside enhancer regions

	% enhancer CpGs	% other CpGs	χ^2	df	*p-value*
M1 linear DNAm change null/negative/positive	47.54/28.58/23.88	48.47/35.66/15.87	508.87	2	$3.16x10^{-111}$
M1 inter-individual variation in DNAm change no/yes	72.55/27.45	72.64/27.36	0.03	1	$8.70x10^{-01}$
M2 nonlinear DNAm change linear or null/Positive-Neutral/Negative-Neutral/other nonlinear	78.65/9.85/8.21/3.29	89.23/4.71/3.03/3.03	1434.20	3	$1.11x10^{-310}$
M2 inter-individual variation in DNAm change from birth no/yes	94.41/5.59	96.7/3.30	150.83	1	$1.13x10^{-34}$
M2 inter-individual variation in DNAm change from 6 years no/yes	99.83/0.17	99.83/0.17	0.00	1	1.00
M2 inter-individual variation in DNAm change from 9 years no/yes	92.28/7.72	91.82/8.18	2.55	1	$1.10x10^{-01}$
M3 stable DNAm sex differences no/yes	94.93/5.07	95.19/4.81	1.30	1	$2.55x10^{-01}$
M3 sex differences in DNAm change no/yes	99.74/0.26	99.62/0.38	2.95	1	$8.58x10^{-02}$

DNAm: DNA methylation

Supplemental Table 9. CpGs part of Horvath age estimator

	% Horvath CpGs	% other CpGs	χ^2	df	*p-value*
M1 linear DNAm change null/negative/positive	37.11/29.75/33.14	48.46/35.52/16.02	77.22	2	$1.71x10^{-17}$
M1 inter-individual variation in DNAm change no/yes	74.22/25.78	72.63/27.37	0.37	1	$5.42x10^{-01}$
M2 nonlinear DNAm change linear or null/Positive-Neutral/ Negative-Neutral/other nonlinear	73.65/5.38/8.22/12.75	89.03/4.81/3.13/3.03	148.54	3	$5.43x10^{-32}$
M2 inter-individual variation in DNAm change from birth no/yes	95.75/4.25	96.65/3.35	0.63	1	$4.28x10^{-01}$
M2 inter-individual variation in DNAm change from 6 years no/yes	100/0	99.83/0.17	0.01	1	$9.04x10^{-01}$
M2 inter-individual variation in DNAm change from 9 years no/yes	93.2/6.8	91.83/8.17	0.71	1	$4.00x10^{-01}$
M3 stable DNAm sex differences no/yes	86.97/13.03	95.19/4.81	50.23	1	$1.37x10^{-12}$
M3 sex differences in DNAm change no/yes	98.58/1.42	99.63/0.37	7.73	1	$5.44x10^{-03}$

DNAm: DNA methylation

Supplemental Table 10. CpGs part of Hannum age estimator

	% Hannum CpGs	% other CpGs	χ^2	df	*p-value*
M1 linear DNAm change null/negative/positive	8.45/46.48/45.07	48.45/35.52/16.03	63.2	2	$1.89x10^{-14}$
M1 inter-individual variation in DNAm change no/yes	46.48/53.52	72.64/27.36	23.14	1	$1.51x10^{-06}$
M2 nonlinear DNAm change linear or null/Positive-Neutral/Negative-Neutral/other nonlinear	54.93/7.04/7.04/30.99	89.02/4.81/3.14/3.03	196.34	3	$2.61x10^{-42}$
M2 inter-individual variation in DNAm change from birth no/yes	83.10/16.90	96.65/3.35	36.22	1	$1.76x10^{-09}$
M2 inter-individual variation in DNAm change from 6 years no/yes	100/0	99.83/0.17	0	1	1.00
M2 inter-individual variation in DNAm change from 9 years no/yes	71.83/28.17	91.84/8.16	35.27	1	$2.87x10^{-09}$
M3 stable DNAm sex differences no/yes	90.14/9.86	95.18/4.82	2.92	1	$8.76x10^{-01}$
M3 sex differences in DNAm change no/yes	100/0	99.63/0.37	0	1	1.00

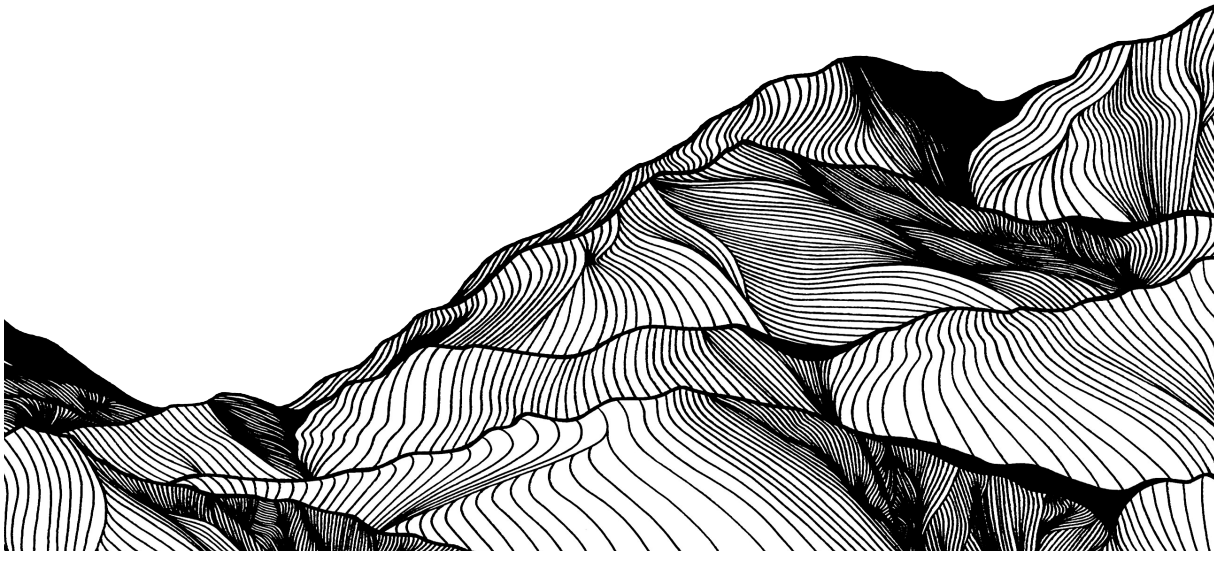
DNAm: DNA methylation

Supplemental Table 11. Estimated white blood cell proportions

	Generation R **mean (*SD*)**	***ALSPAC*** **mean (*SD*)**
0 years – Bakulski method		
CD8T	0.13 (0.05)	0.09 (0.05)
CD4T	0.16 (0.05)	0.18 (0.06)
NK	0.03 (0.03)	0.01 (0.02)
Bcell	0.10 (0.03)	0.17 (0.04)
Mono	0.09 (0.02)	0.01 (0.02)
Gran	0.41 (0.11)	0.35 (0.10)
nRBC	0.12 (0.07)	0.20 (0.09)
6 / 7 years – Houseman method		
CD8T	0.12 (0.04)	0.04 (0.03)
CD4T	0.17 (0.05)	0.21 (0.05)
NK	0.02 (0.02)	0.19 (0.04)
Bcell	0.13 (0.03)	0.14 (0.03)
Mono	0.06 (0.02)	0.06 (0.03)
Gran	0.53 (0.08)	0.44 (0.08)
10 years – Houseman method		
CD8T	0.12 (0.04)	
CD4T	0.18 (0.05)	
NK	0.03 (0.03)	
Bcell	0.11 (0.03)	
Mono	0.06 (0.02)	
Gran	0.52 (0.08)	
17 years – Houseman method		
CD8T		0.03 (0.03)
CD4T		0.18 (0.05)
NK		0.20 (0.05)
Bcell		0.11 (0.03)
Mono		0.07 (0.03)
Gran		0.48 (0.09)

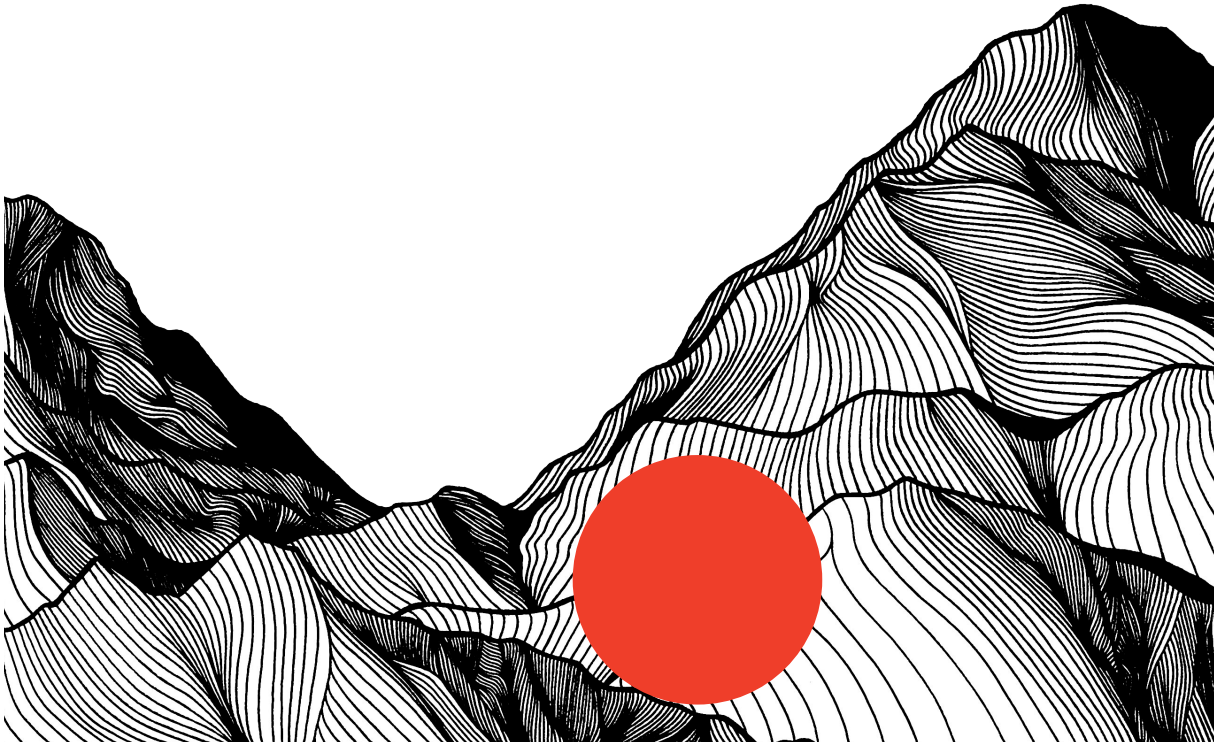
DNAm: DNA methylation

Supplemental Table 2 and 5-8 are available online in the published preprint–https://www.biorxiv.org/content/10.1101/2020.06.09.142620v2



CHAPTER IX

Discussion



The putative role of DNA methylation in the effects of stress on the developing child are much alluded to in the literature on genetic and environment risk factors, and psychological outcomes. In the current thesis, we set out to study associations between stress and DNA methylation in child development to explain how stress can get 'under the skin'. We reviewed existing literature on DNA methylation and parenting stress and examined associations between DNA methylation and family stress as well as interpersonal stress in social situations with peers throughout childhood in population-based samples in the Netherlands (Generation R and the United Kingdom (ALSPAC. Furthermore, we used these datasets to create an epigenome-wide characterization of change in DNA methylation from birth to late adolescence.

Stress and DNA methylation

Throughout this thesis, we tried to link DNA methylation levels to the occurrence of stress. In our review of the literature on DNA methylation and stress resulting from parenting behavior in Chapter II, we saw that there are many studies associated DNA methylation with both parenting stress and child psychological outcomes, but also noted important methodological drawbacks in the set-up of these studies. Mainly, issues included those of (i reliability and validity of the DNA methylation measurement, (ii reproducibility of findings, and (iii causality. We will discuss how we addressed each of these methodological issues in our own studies in the following paragraphs.

In our own studies on DNA methylation on stress, using a population-based approach, we identified several small associations between stress and DNA methylation. In Chapter III, we found that methylation of *FKBP5*, a gene involved in the binding of cortisol to glucocorticoid receptors[1], *together* with *FKBP5* genotype and resistant attachment of the child to the mother, was associated with cortisol reactivity. Since the attachment behavior in the child is seen as response to the parenting behavior of the attachment figure, we conclude that for *FKBP5*, genotype, DNA methylation, and environment *together* are associated with stress regulation. In Chapter IV, we found that maternal sensitivity at 3 and 4 years was associated with DNA methylation in the child at 6 years at several genomic regions. The top region included *DOCK4*, a gene involved in neuronal growth and migration previously linked to the stress response in mice. Hence, in this study, we were able to relate an index of early maternal caretaking to child DNA methylation. In Chapter V, we studied sleep in association with genome-wide DNA methylation and found associations at chromosome 17, spanning *MAPT,* a gene for Tau protein and linked to Alzheimer's disease, as well as *CRHR1*, a gene coding for corticotropin-releasing hormone receptor 1 and pivotal for HPA-axis functioning. Thus, sleep was associated with methylation of, again, genes involved in memory and stress. Last, in Chapter VII, we studied the association of exposure to bullying victimization in children of Generation R and ALSPAC and change in DNA methylation in a longitudinal epigenome-wide association study. We found that DNA methylation at a site on

RAB14 changed differently for children who had been bullied than for those who had not been bullied. This gene is important in Golgi apparatus functioning and has previously been associated with stress exposure in rats. The effect was small but consistent between cohorts and indicates that there might be an association between social stress exposure and change in DNA methylation. Together, we conclude that there are indeed associations between stress and DNA methylation that were small but measurable in the children from these population-based samples. Methodological considerations will be discussed.

Candidate-gene methylation versus epigenome-wide studies–and the issue of small effect sizes

Since the study of DNA methylation in humans is relatively new, methodological considerations remain a topic of ongoing debate. First of all, when studying DNA methylation one can use a candidate-gene or epigenome-wide approach. A candidate-gene methylation study can be a great place to start; generally there is already an understanding of the gene's variants in relation to its functioning and a study into DNA methylation can add to the prediction of variation in gene function or related outcomes. *FKBP5* is an example of such a gene; its key role in HPA functioning is well-known[1]. The study presented in in Chapter III thereby worked as a proof-of-principle, showing how gene, environment, and DNA methylation might interact to affect HPA functioning. On the other hand, adopting an hypothesis free epigenome-wide opens up the possibility of finding new correlates and processes. This is the approach taken in Chapter IV, V, VII, and VIII and as a result, we see that the hits that come up are not the ones that would be selected for a candidate gene study. Moreover, in candidate-gene look-ups within an epigenome-wide approach in Chapter IV and VII, we noted that *p*-values for the sites on these genes do not reach gene-wide significance (as opposed to genome-wide significance). Similarly, other studies on childhood stress do not report hits associated with popular candidate-genes[2-11] and neither did one such study that also did a specific candidate-genes look-up [12].

A reason for this contrast might be that effect sizes in this field of research are generally small. In Chapter VIII we saw that changes in DNA methylation were often as small as a cumulative change of less than 2% from birth to late adolescence; similarly, in Chapters III, IV, V, and VII we saw that associations between DNA methylation and stress were small. This is in line with results from other population-based studies on DNA methylation and stress[2, 12] as well as with results from population-based studies on DNA methylation and exposure to, e.g., air pollution[13] or prenatal exposure to alcohol[14], or arsenic[15]. When having to control for a false positive rate otherwise inflated by half a million tests, the adjusted significance level is greatly reduced and true, small, associations may fly under the radar. One way of increasing the power to detect an association is through the study of large samples. Generation R has one of the largest datasets of childhood DNA methylation in the world, yet larger is always better. In Chapter VII and VIII therefore, we studied DNA methylation of the children of

Generation R, as well as ALSPAC in a combined dataset. In Chapter VII however, we only found one association between DNA methylation and bullying victimization. This indicates that even larger sample sizes might be necessary to study such associations, something that could be achieved in consortia[16]. In the epigenome-wide studies performed in Chapter VII and VIII, we used a Bonferroni-correction to adjust for multiple testing, thereby treating each test as an independent added risk of obtaining a false positive. DNA methylation at different sites is however correlated. Research techniques that take into account this correlational structure of the data and reduce the multidimensionality based on that structure may therefore be worthwhile from a methodological perspective, but also from a theoretical, systems biology perspective, and perhaps converging with the idea of epistasis, that a phenotype is build up from the interaction of multiple genes[17]. We indeed applied such reductionist methods in Chapter IV where correlated sites were clustered at the regional level, and in Chapter V where a network approach was applied to form different modules of sites. In Chapter III, moreover, we used a factor analysis to reduce the burden of multiple testing at the gene-level of *FKBP5*. These approaches were successful in that we showed significant associations between DNA methylation and maternal sensitivity, sleep, and cortisol reactivity, respectively. The power in these studies was increased via the reduced number of tests performed and in Chapter IV, we additionally saw that effect sizes increased by examining related methylation sites in concert as opposed to individually. This favors the interpretation that single sites are unlikely to act alone and is in agreement with the systems biology perspective.

Multiple testing could also be reduced by simply not testing all measured methylation sites. Results from our study into DNA methylation in development in Chapter VIII indicated that there are many sites that do not change and do not show inter-individual variation in change. Such sites are unlikely to be affected by environmental stressors in childhood and excluding such sites pre-emptively from an epigenome-wide association study can reduce the burden of multiple testing, which can easily be done using the website we developed (http://epidelta. mrcieu.ac.uk/).

Another way to examine the candidate gene methylation versus the epigenome-wide dilemma is through the lens of the developments in the field of Genetics and Psychiatry. The field started with candidate studies and evolved towards larger and larger genome-wide associations studies (GWASs), and retrospectively one can see that the earlier candidate genes do not come up as 'frontrunners' in GWAS studies on psychiatric diseases[18, 19]. It is now thought that many such candidate studies were underpowered and showed false positive results[20]. However, a recent epigenome-wide association study of physical aggression did observe a genome-wide corrected associations within a gene often

used in candidate-gene studies; in dopamine-receptor coding gene *DRD4*[21]. This offers hope that ultimately, the gene-candidate and epigenome-wide approach might produce converging or complementary results. To this end, it is important keep power enhancing methods in mind, not by increasing sample size and reducing the number of tests, but also by improving the quality of the measurements and thus reducing measurement error. Such approaches are discussed in the paragraphs below.

Temporal change, reliability, and validity of DNA methylation measurement

The first methodological issue noted in our literature review in Chapter II was that essential characteristics of the DNA methylation measurement were unknown, the most important ones being lack of clarity to which extent DNA methylation changes over time, or to which extent intra-individual DNA methylation patterns between different tissues are the same.

It is important to understand whether DNA methylation changes, and furthermore, to what extent the rate of DNA methylation change differs between people, in order to identify sites that might be susceptible to environmental influences, or might be responsible for a changing phenotype. In Chapter VIII, we were able to track almost half a million genomic sites and characterize their DNA methylation levels from birth to late adolescence. Here we saw a large proportion of sites that changed, and a considerable proportion of sites with a different rate of DNA methylation change for different individuals, or for the two sexes. These findings could aid in narrowing down the search for sites that are affected by stressful environments and in interpreting findings from such studies by placing them in a developmental context. For example, inter-individual variation in change was more often measured between 9 and 18 years of age than between birth and 9 years, indicating that the former period may be of more interest for researchers interested in developmentally formative periods for DNA methylation. With our freely accessible website that visualizes the trajectory of change and contains downloadable results at the level of the individual site, we provide a valuable tool for the research community. Last, it was encouraging for other studies using similar array-type DNA methylation data that the epigenome-wide estimates of DNA methylation stability and change between Generation R and ALSPAC in the study of Chapter VIII were remarkably similar. This adds to the reliability of these data in our and other cohorts and supports the practice of pooling effect estimates from different epidemiological cohorts in epigenome-wide meta-analyses in consortia.

Another challenge to the validity of the DNA methylation measurement is the extent to which DNA methylation levels in one tissue are informative of those in another tissue. In our sample, as is not uncommon for population-based samples, we had access to cord blood and peripheral blood. Not much research existed however on the extent to which cord blood and peripheral blood provide information on the same biological processes even if measured

cross-sectionally. In Chapter VIII we found many sites at which DNA methylation changed consistently across individuals from birth (cord blood to late adolescence (peripheral blood, indicating that relative levels of DNA methylation in one tissue were indicative of relative levels in another. This is encouraging for developmental population-based epigenomic studies interested in longitudinal study designs.

On a related note, in the study of DNA methylation in stress and psychosocial functioning, brain tissue would be a tissue of high interest. In humans, DNA methylation studies in stress have been performed in post-mortem brain tissues[22], but such studies do not have the advantages of a prospective research design; information on psychopathology is generally obtained retrospectively, often via family members. Further, the post-mortem state of the brain itself may affect the epigenetic configuration of the DNA[23]. A study of epigenome-wide concordance of DNA methylation in peripheral blood and brain tissue showed that only around 6% of DNA methylation in blood at least moderately predicts DNA methylation in the brain, with the exact percentages depending on the location in the brain[24]. In addition, DNA methylation levels at sites with such cross-tissue covariance were more determined by genetic variation, indicating that sites with high brain-blood concordance may be influenced relatively less by the environment. Online tools (http://epigenetics.essex.ac.uk/bloodbrain/; https://redgar598.shinyapps.io/BECon/ stemming from such studies[24, 25] are useful in determining the extent to which DNA methylation for sites of interest in blood mirrors that in the brain. In Chapter IV and V we performed such a look-up of hits of regions associated with maternal sensitivity and sleep, respectively, and in both cases meaningful blood-brain associations were found for top regions, indicating that in these cases, blood proved an informative tissue on DNA methylation in the brain. Genetic associations with DNA methylation were apparent for the region associated with sleep in Chapter V, but not for the top region associated with maternal sensitivity in Chapter IV. It should be noted, however, that even when DNA methylation in peripheral or cord blood is not associated with DNA methylation in the brain, blood DNA methylation in patterns can still be informative for research on psychosocial functioning, since there are many peripheral bodily functions associated with the functions of the mind. For example, when cortisol is produced by the adrenal cortices, the hormone travels through the blood and leukocytes themselves express glucocorticoid receptors, which have been associated with stress-induced inflammatory reactions[26, 27]. Relatedly, glucocorticoid sensitivity in leukocytes has been associated with *FKBP5* expression[28], indicating that blood can be a tissue of interest in the study of DNA methylation and stress.

Reproducibility

A second issue we discussed in Chapter II, is that previous research on DNA methylation and parenting stress was often based on small samples, thereby increasing the chances of reporting falsely positive results. We noted a need for large samples and built-in replication or meta-analytic effort. Results from meta-analyses, further, are more generalizable since

they are produced by converging results from different populations, with variables measured at different time-points and often with different methods. Indeed, we made use of the large datasets of Generation R in all our studies and combined our data with those of the ALSPAC cohort in Chapter VII on bullying and VIII on change in DNA methylation. Despite our best efforts to maximize sample sizes, as mentioned, effect sizes were small and confirmation of findings by replication would be optimal. Consortia of cohorts with DNA methylation data offer great opportunities for replication or meta-analysis, however the number of cohorts that have longitudinal DNA methylation is currently small, as are the number of cohorts that have used observational measures such as those on attachment (Chapter III) or maternal sensitivity (Chapter III, IV). Observational measures are costly but valuable, for example when children are too young to report how they feel, and, as discussed in Chapter VI with regard to facial expressions, may not be influenced by reporter biases that affect questionnaire data. In the current thesis therefore, we took a meta-analytic and longitudinal approach where possible, while also aiming to maintain a high quality of data by using observational measures where possible.

Causality

The last issue raised in our literature review in Chapter II concerned directionality of causality, an matter not limited to DNA methylation research, but relevant to research in general, especially epidemiological research. Within epidemiology, longitudinal modelling is a method to study temporality and approach causality. In Chapter VII, we studied if being bullied was associated with *change* in DNA methylation from pre- to post bullying. This does not preclude the possibility that some children were already bullied before the first measurement of DNA methylation, or that children who are at risk of being bullied also, for a reason unrelated to the bullying itself, show certain changes in DNA methylation. Yet, longitudinal modelling within epidemiological DNA methylation research is a large step forward from the commonly used cross-sectional research designs. Another methodological tool is mediation analysis, which can help understand if one variable affects another variable *via* a third variable. Mediation analysis requires the putative predictor, mediator, and outcome to be measured in chronological order, which unfortunately made it impossible for us to perform a mediational analysis. In future research, such an analysis could be applied to study if stress affects DNA methylation, which in turn affects mental health. Another issue of causality, for example discussed in Chapters IV and V, is that associations between DNA methylation and a phenotype, may be caused by genetic variants associated with DNA methylation, and that phenotype. Online tools[29-31] are helpful in finding such associations between DNA methylation and genetics, but direct measurement of both genetic and epigenetic influences within the same dataset are necessary to disentangle the in(ter)dependence of effects. We took this approach in Chapter III by measuring both *FKBP5* genotype and *FKBP5* methylation, but in genome-wide analyses the combination of both aspects is more complex and further increases the burden of multiple testing. Genetics can also be used as a tool in finding causal pathways; in Mendelian

randomization, genetic variants are used as instrumental variables[32, 33]. With genome-wide analyses becoming more ubiquitous, and polygenic risk scores thereby more readily available, these methods are gaining popularity, although limitations should be taken into account[34]. The ultimate approach to tackle issues of causality are intervention studies, preferably within the setting of randomized controlled trials. Such studies are pivotal in understanding the causal associations between stress and DNA methylation.

Future research

In order to fully understand how stress and DNA methylation are related during development, we will need to further address issues of power, tissue, and causality. Within our own field of epidemiology, more collaboration between different cohorts will help increase sample size and add to the reproducibility of our results. Carefully conducted data reduction approaches, that are mindful of increased type I error, can be applied to minimize the burden of multiple testing. As the number of sites measured at profiling platforms increases, most recently from the Illumina 450K to 850K array (Illumina Infinium HumanMethylation450 BeadChip to Infinium MethylationEPIC Beadchip; Illumina Inc., San Diego, USA), such approaches may become indispensable. Moreover, estimates of change and inter-individual variation might change with this new platform. The new platform includes more sites located at enhancer elements[35], sites that in Chapter VIII more often had increasing DNA methylation over time with more inter-individual variation in change, thus this pattern may more often occur on the new platform. Longitudinal patterns will need to be readdressed, however, to establish whether this is the case.

In addition, more research is needed to learn to map out blood-brain associations of DNA methylation. Current online tools are valuable but based on small samples with data stemming from older individuals. Tissues from children would be necessary to understand how patterns of covariation change throughout development. The measurement of buccal cells may be a useful and accessible addition to that of blood[36], since buccal and brain cells both originate from the mesodermal germ layer that develops during gastrulation, whereas blood cells originate from the ectodermal germ layer, thus variation in DNA methylation may be more coherent between buccal and brain cells, than between blood and brain cells.

Longitudinal measurements and mediation models will help to separate cause and effect in the study of DNA methylation and stress. Natural experiments, such as the current COVID-19 crisis, may offer epidemiological studies an even more effective approach to study directionality of effects. Assuming that the crisis increases stress, a study on perinatal stress and postnatal DNA methylation can be envisioned not unlike the one applied in the Dutch Hunger Winter studies[37]. For example, DNA methylation of children conceived at the beginning of the crisis could be compared to that of children conceived just after the crisis. However, to truly model stress as an independent factor, intervention studies with random group assignment are pivotal[38].

Lastly, in the current thesis, we have treated DNA methylation as an isolated player in the epigenetic machinery. Whereas data on DNA methylation is relatively easily obtainable and therefore often studied, DNA methylation interacts with other forms of epigenetics[39-41]. On a related note, the research fields of Genomics, Epigenomics, and Transcriptomics are by and large separate entities. Integration of different types of – omics data may come with statistical and computational challenges but is ultimately a necessary step in understanding the full picture of how genetics and the environment together are associated with stress in development.

Conclusions

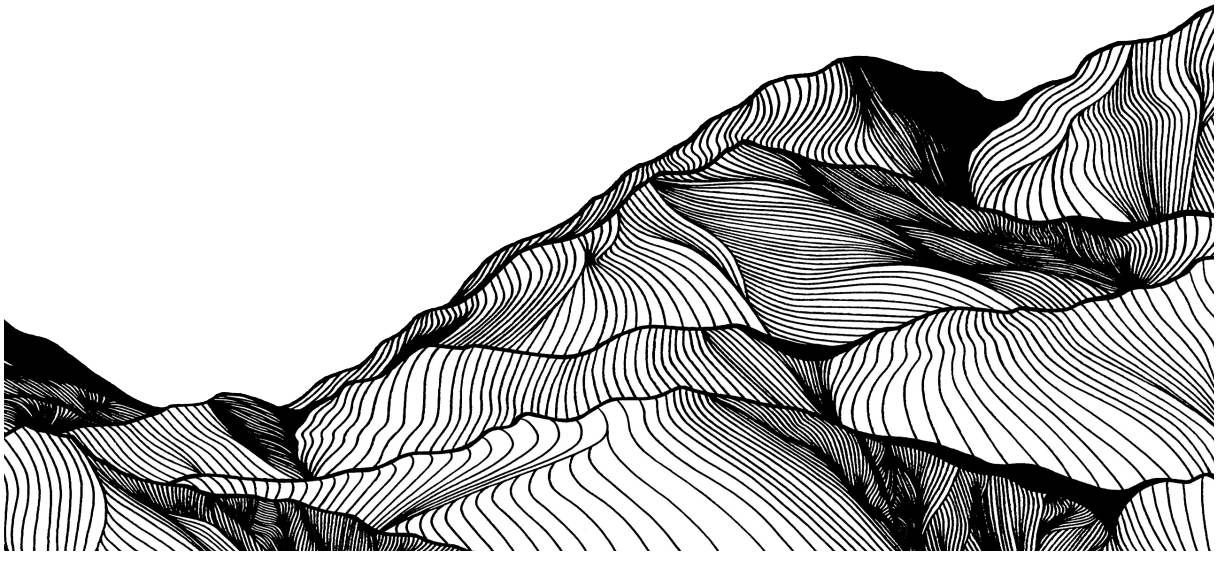
In this thesis, we studied associations between stress and DNA methylation in the developing child. In the process of doing so, we developed a micro-coding system for facial expressions within an experiment on interpersonal stress and we detailed an epigenetic landscape of change by mapping out half a million longitudinal patterns of DNA methylation. Throughout our studies, we found associations between stress and DNA methylation, and indications that effect sizes are small. These results call for a continuation of research in the field of stress and DNA methylation, but we advise to do so thoughtfully. Large sample sizes are necessary, preferably within consortial efforts. Longitudinal studies will help the researcher understanding the pathways of effects, as will mediation analyses. We recommend an integration of –omics data on the one hand, and reductionist approaches on the other. Ultimately, this might lead us to better understand how we are all just rolling down the ridges and valleys of the epigenetic landscape.

References

1. Binder EB. The role of FKBP5, a co-chaperone of the glucocorticoid receptor in the pathogenesis and therapy of affective and anxiety disorders. *Psychoneuroendocrinology.* 2009;34:S186-S95.
2. Houtepen LC, Hardy R, Maddock J, et al. Childhood adversity and DNA methylation in two population-based cohorts. *Translational Psychiatry.* 2018;8:266.
3. Houtepen LC, Vinkers CH, Carrillo-Roa T, et al. Genome-wide DNA methylation levels and altered cortisol stress reactivity following childhood trauma in humans. *Nature Communications.* 2016;7.
4. Cecil CAM, Smith RG, Walton E, Mill J, McCrory EJ, Viding E. Epigenetic signatures of childhood abuse and neglect: implications for psychiatric vulnerability. *Journal of Psychiatric Research.* 2016;83:184-94.
5. Mehta D, Klengel T, Conneely KN, et al. Childhood maltreatment is associated with distinct genomic and epigenetic profiles in posttraumatic stress disorder. *Proceedings of the National Academy of Sciences. 2013*;110:8302-7.
6. Prados J, Stenz L, Courtet P, et al. Borderline personality disorder and childhood maltreatment: A genome-wide methylation analysis. *Genes, Brain and Behavior.* 2015;14:177-88.
7. Suderman M, Pappas JJ, Borghol N, et al. Lymphoblastoid cell lines reveal associations of adult DNA methylation with childhood and current adversity that are distinct from whole blood associations. *International Journal of Epidemiology*. 2015;44:1331-40.
8. Yang BZ, Zhang H, Ge W, et al. Child abuse and epigenetic mechanisms of disease risk. *American Journal of Preventive Medicine.* 2013;44:101-7.
9. Kumsta R, Kreppner J, Kennedy M, Knights N, Rutter M, Sonuga-Barke E. Psychological Consequences of Early Global Deprivation An Overview of Findings From the English & Romanian Adoptees Study. *European Psychologist.* 2015;20:138-51.
10. Marinova Z, Maercker A, Küffer A, et al. DNA methylation profiles of elderly individuals subjected to indentured childhood labor and trauma. *BMC Medical Genetics.* 2017;18:21.
11. Dunn EC, Soare TW, Zhu Y, et al. Sensitive periods for the effect of childhood adversity on DNA methylation: Results from a prospective, longitudinal study. *Biological Psychiatry.* 2019;85:838-49.
12. Marzi SJ, Sugden K, Arseneault L, et al. Analysis of DNA methylation in young people: limited evidence for an association between victimization stress and epigenetic variation in blood. *American Journal of Psychiatry.* 2018;175:517-29.
13. Panni T, Mehta AJ, Schwartz JD, et al. Genome-wide analysis of DNA methylation and fine particulate matter air pollution in three study populations: KORA F3, KORA F4, and the normative aging study. *Environmental Health Perspectives.* 2016;124:983-90.
14. Sharp GC, Arathimos R, Reese SE, et al. Maternal alcohol consumption and offspring DNA methylation: findings from six general population-based birth cohorts. *Epigenomics.* 2018;10:27-42.
15. Green BB, Karagas MR, Punshon T, et al. Epigenome-wide assessment of DNA methylation in the placenta and arsenic exposure in the New Hampshire Birth Cohort Study (USA). *Environmental Health Perspectives.* 2016;124:1253-60.

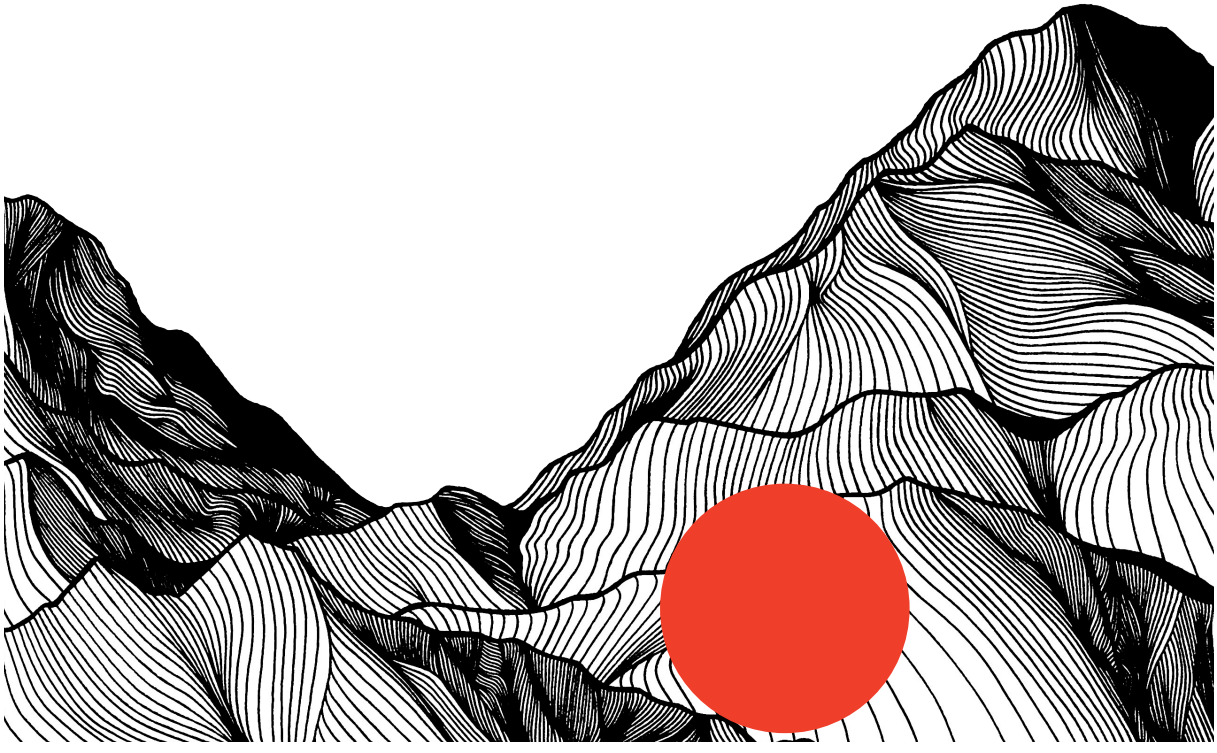
16. Felix JF, Joubert BR, Baccarelli AA, et al. Cohort profile: pregnancy and childhood epigenetics (PACE) Consortium. *International Journal of Epidemiology.* 2017;47:22-3u.
17. Carlborg Ö, Haley CS. Epistasis: too often neglected in complex trait studies? *Nature Reviews Genetics.* 2004;5:618-25.
18. Johnson EC, Border R, Melroy-Greif WE, de Leeuw CA, Ehringer MA, Keller MC. No evidence that schizophrenia candidate genes are more associated with schizophrenia than noncandidate genes. *Biological Psychiatry.* 2017;82:702-8.
19. Border R, Johnson EC, Evans LM, et al. No support for historical candidate gene or candidate gene-by-interaction hypotheses for major depression across multiple large samples. *American Journal of Psychiatry.* 2019;176:376-87.
20. Duncan LE, Ostacher M, Ballon J. How genome-wide association studies (GWAS) made traditional candidate gene studies obsolete. *Neuropsychopharmacology.* 2019;44:1518-23.
21. Cecil CAM, Walton E, Pingault JB, et al. DRD4 methylation as a potential biomarker for physical aggression: An epigenome-wide, cross-tissue investigation. American Journal of Medical Genetics Part B: *Neuropsychiatric Genetics.* 2018;177:746-64.
22. McGowan PO, Sasaki A, D'Alessio AC, et al. Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nature Neuroscience.* 2009;12:342-8.
23. Sjöholm LK, Ransome Y, Ekström TJ, Karlsson O. Evaluation of post-mortem effects on global brain DNA methylation and hydroxymethylation. *Basic & Clinical Pharmacology & Toxicology.* 2018;122:208-13.
24. Hannon E, Lunnon K, Schalkwyk L, Mill J. Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. *Epigenetics.* 2015;10:1024-32.
25. Edgar RD, Jones MJ, Meaney MJ, Turecki G, Kobor MS. BECon: a tool for interpreting DNA methylation findings from blood in the context of brain. *Translational Psychiatry.* 2017;7:e1187-e.
26. Yehuda R, Golier JA, Yang R-K, Tischler L. Enhanced sensitivity to glucocorticoids in peripheral mononuclear leukocytes in posttraumatic stress disorder. *Biological Psychiatry.* 2004;55:1110-6.
27. De Kloet CS, Vermetten E, Bikker A, et al. Leukocyte glucocorticoid receptor expression and immunoregulation in veterans with and without post-traumatic stress disorder. *Molecular Psychiatry.* 2007;12:443-53.
28. Vermeer H, Hendriks-Stegeman BI, van der Burg B, van Buul-Offers SC, Jansen M. Glucocorticoid-induced increase in lymphocytic FKBP51 messenger ribonucleic acid expression: a potential marker for glucocorticoid sensitivity, potency, and bioavailability. *The Journal of Clinical Endocrinology & Metabolism.* 2003;88:277-84.
29. Hannon E, Knox O, Sugden K, et al. Characterizing genetic and environmental influences on variable DNA methylation using monozygotic and dizygotic twins. *PLoS Genetics.* 2018;14:e1007544.
30. Gaunt TR, Shihab HA, Hemani G, et al. Systematic identification of genetic influences on methylation across the human life course. *Genome Biology.* 2016;17:1-14.
31. Bonder MJ, Luijk R, Zhernakova DV, et al. Disease variants alter transcription factor levels and methylation of their binding sites. *Nature Genetics.* 2017;49:131-8.

32. Caramaschi D, Hatcher C, Mulder RH, et al. Epigenome-wide association study of seizures in childhood and adolescence. *Clinical Epigenetics.* 2020;12:1-13.
33. Pingault J-B, Cecil C, Murray J, Munafo M, Viding E. Causal inference in psychopathology: Using Mendelian randomisation to identify environmental risk factors for psychopathology. *Psychopathology Review.* 2016;4:4-25.
34. Burgess S, Labrecque JA. Mendelian randomization with a binary exposure variable: interpretation and presentation of causal estimates. *European Journal of Epidemiology.* 2018;33:947-52.
35. Moran S, Arribas C, Esteller M. Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. *Epigenomics.* 2016;8:389-99.
36. Lowe R, Gemma C, Beyan H, et al. Buccals are likely to be a more informative surrogate tissue than blood for epigenome-wide association studies. *Epigenetics.* 2013;8:445-54.
37. Tobi EW, Goeman JJ, Monajemi R, et al. DNA methylation signatures link prenatal famine exposure to growth and metabolism. *Nature Communications.* 2014;5.
38. Moffitt TE. Childhood exposure to violence and lifelong health: Clinical intervention science and stress biology research join forces. *Development and Psychopathology.* 2013;25.
39. Li X, Wang X, He K, et al. High-resolution mapping of epigenetic modifications of the rice genome uncovers interplay between DNA methylation, histone methylation, and gene expression. *The Plant Cell.* 2008;20:259-76.
40. Ball MP, Li JB, Gao Y, et al. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nature Biotechnology.* 2009;27:361.
41. Fuks F. DNA methylation and histone modifications: teaming up to silence genes. *Current Opinion in Genetics & Development.* 2005;15:490-5.



CHAPTER X

Summary / Samenvatting



Summary

In research on the interaction between genes and the environment in psychosocial development of children, attention has turned towards epigenetics. 'Epigenetics' are the molecular structures on and around the DNA that regulate the expression of the DNA. Because it is believed that these epigenetic structures adapt to the environment in which one grows up, epigenetics has been lauded as the putative mechanism through which 'nature affects nurture'. Such a mechanism could explain how early stressful experiences can have major emotional psychological consequences later in life. An often-studied form of epigenetics is DNA methylation – a methyl group bound to the DNA sequence itself–since it is relatively easy to measure. However, research on DNA methylation is oftentimes limited by small sample sizes as well as cross-sectional study designs, which make it difficult to interpret research results within a developmental framework. In the current thesis, we therefore studied DNA methylation throughout development, in a large prospective population-based study of the Generation R Study in Rotterdam, in some cases supplemented by data from the prospective population-based Avon Longitudinal Study of Parent and Children (ALSPAC) in the United Kingdom. We studied associations between DNA methylation and stress in child development.

In **Chapter II**, we performed a literature study on DNA methylation and stress within the family environment to obtain an understanding of existing studies. We described that associations were found between parenting stress and DNA methylation, as well as between DNA methylation and child psychosocial outcomes in multiple studies. The designs of these studies were however often limited by small sample sizes and cross-sectional designs. We further noted that little was known about stability and change in DNA methylation in development. In designing our own studies, we therefore set out to address these issues.

We studied early parent-child relations and DNA methylation in Generation R in Chapter III and IV, firstly focusing on a single gene, and secondly on a genome-wide level. In **Chapter III**, we selected *FKBP5* as a candidate gene. This gene codes for a co-chaperone of the glucocorticoid receptor, which makes it important for stress regulation. We studied attachment behavior of the child towards the mother as well as reactivity of the stress hormone cortisol in the child, during the Strange Situation Procedure. We found that DNA methylation of *FKBP5*, together with *FKBP5* genotype and attachment, was associated to cortisol reactivity in the child. In **Chapter IV**, we studied associations between observed maternal sensitivity towards the child and genome-wide DNA methylation. DNA methylation was studied at the site-level, as well as at the level of co-localized correlating sites, i.e. the region level. No associations were found at site-level, but DNA methylation was associated to maternal sensitivity at 13 different regions, 4 of

which remained after adjustment for DNA methylation at birth. One of the top-regions was located at the *DOCK4* gene, a gene involved neuronal growth and migration and previously related to stress reactivity in mice.

We report on the association between sleep and DNA methylation in **Chapter V**. Sleep is related to diurnal cortisol fluctuations. We applied a network analysis to cluster genome-wide DNA methylation and found a cluster encompassing *MAPT* and *CRH1*. *MAPT* is a gene coding for Tau protein, a protein often linked to Alzheimer's disease; *CRHR1* codes for corticotropin-releasing hormone, which is a key player in stress regulation.

Social stress was examined in studies presented in both Chapter VI and Chapter VII. In **Chapter VI**, we tested a new method for the micro-coding of facial expressions of emotions within a computerized social exclusion paradigm. We found that facial expressions of sadness and anger were associated to self-report on feelings during the game. We conclude that we have created a new measure of emotions during stress with good utility. In **Chapter VII**, we examined bullying exposure and epigenome-wide changes of DNA methylation in Generation R and ALSPAC. In this longitudinal study, we found an association between being bullied and changes in DNA methylation at a site located on *RAB14.* This gene is involved in Golgi apparatus functioning and has previously been associated to stress exposure in rats.

We joined forces again with ALSPAC in our last study in **Chapter VIII**, to address an important issue of validity within research on DNA methylation, which is to find out to what extent DNA methylation changes over time. We combined the longitudinal sets of DNA methylation of the two cohorts into one dataset with measurements from birth until late adolescence. We studied epigenome-wide change in DNA methylation, within and between individuals. We found that a little over half of all the methylation sites change, and that the rate of change was different for a little over a quarter of all sites. The website we developed to present our results can function as a stepping stone for other researchers, by placing findings on DNA methylation within a longitudinal perspective.

From this thesis, we conclude that associations between stress and DNA methylation in the developing child can be found. Such associations are small, which leads us to recommend performing studies of DNA methylation in large samples and to explore data reduction methods.

Samenvatting

In onderzoek naar de interactie tussen genen en omgeving in de psychosociale ontwikkeling van het kind zijn de ogen vaak gericht op epigenetica. Met 'epigenetica' worden de moleculaire structuren op en rondom het DNA bedoeld, die de expressie van het DNA beïnvloeden. Omdat wordt gedacht dat deze epigenetische structuren zich aanpassen aan de omgeving waarin men opgroeit, wordt epigenetica gezien als een mogelijk mechanisme waardoor 'nurture' 'nature' kan beïnvloeden. Zo'n mechanisme zou kunnen verklaren hoe vroege stressvolle ervaringen grote emotionele gevolgen kunnen hebben later in het leven. In onderzoek naar epigenetica wordt vooral gekeken naar DNA-methylering, een vorm van epigenetica waarbij een methylgroep is gebonden aan de DNA sequentie en die relatief makkelijk te meten is in het lab. Echter, onderzoek naar DNA-methylering werd in het verleden vaak gekenmerkt door kleine streekproeven en wordt vaak cross-sectioneel uitgevoerd, waardoor het moeilijk is om de resultaten in een ontwikkelingsperspectief te plaatsen. In de huidige studiereeks keken we daarom naar epigenetica tijdens de ontwikkeling van het kind, in de grote prospectieve populatiestudie van Generation R Study in Rotterdam, waar mogelijk aangevuld door data uit de prospectieve populatiestudie van de Avon Longitudinal Study of Parent and Children (ALSPAC) in het Verenigd Koninkrijk. We onderzochten associaties tussen DNA-methylering en stress in de ontwikkeling van het kind.

In **Hoofdstuk II** hebben we een literatuurstudie gedaan om een beeld te krijgen van bestaande literatuur aangaande DNA-methylering en stress in de familiesfeer. We beschreven dat verschillende studies associaties hebben laten zien tussen een stressvolle opvoeding en DNA-methylering en tussen DNA-methylering en psychosociale uitkomsten in het kind. Deze studies zijn echter vaak beperkt door kleine streekproeven en vaak alleen cross-sectioneel in opzet. We benoemden daarnaast dat er nog te weinig bekend was over de stabiliteit en verandering van DNA-methylering tijdens de ontwikkeling van het kind. In de opzet van onze eigen onderzoeken hebben we daarom geprobeerd deze kwesties te adresseren.

In Hoofdstuk III en IV hebben we onderzoek gedaan naar de vroege ouder-kindrelatie en DNA-methylering in Generation R, in eerste instantie gericht op één enkel gen en in tweede instantie op het niveau van het hele genoom. In de studie beschreven in **Hoofdstuk III** richtten we ons op het *FKBP5* gen, welke codeert voor een co-chaperone van de glucocorticoid receptor en daarmee belangrijk is voor de stressregulatie. We observeerden de gehechtheidsrelatie van het kind met de moeder en de reactiviteit van het stresshormoon cortisol in het kind tijdens de *Strange Situation Procedure*. We vonden dat DNA-methylering van *FKBP5*, samen met *FKBP5* genotype en gehechtheidsrelatie, geassocieerd is met cortisolreactiviteit in het kind. In **Hoofdstuk IV** onderzochten we associaties tussen geobserveerde sensitiviteit van de moeder in de interactie met haar kind en genoomwijde DNA-methylering. Associaties werden zowel

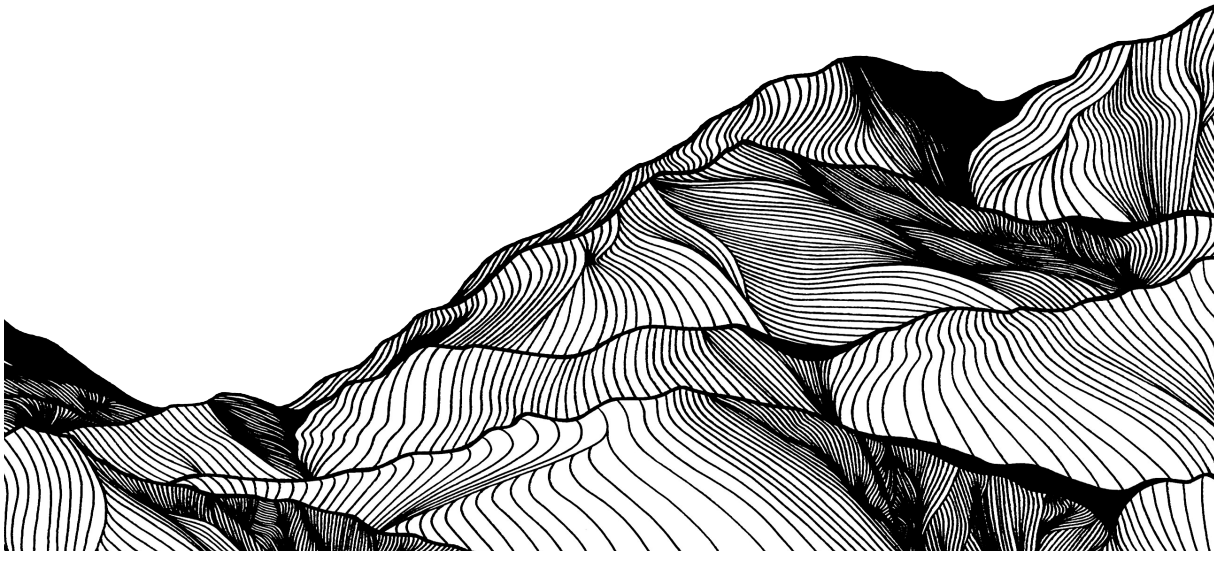
per locatie (site) bekeken als per regio, dat is een gebied met geassocieerde sites. We vonden geen associaties op site-niveau, maar wel 13 regio's waar DNA-methylering geassocieerd is met sensitiviteit van de moeder, waarvan er vier overbleven na adjustering voor DNA-methylering vlak na de geboorte. Eén van deze regio's lag op het *DOCK4* gen, een gen dat betrokken is bij neuronale groei en migratie en welke eerder geassocieerd is met stressreactiviteit in muizen.

In **Hoofdstuk V** onderzochten we de associatie tussen DNA-methylering en slaap. Slaap is gerelateerd aan cortisolfluctuaties over de dag heen. We pasten een netwerkanalyse toe om genoomwijde DNA-methylering te clusteren en vonden een cluster met onder andere *MAPT* en *CRHR1. MAPT* is een gen dat codeert voor het Tau eiwit en dat vaak geassocieerd wordt met Alzheimer; *CRHR1* codeert voor corticotropine-releasing hormoon, een belangrijk hormoon voor de stressregulatie.

In de studies beschreven in Hoofdstuk VI en VII hebben we gekeken naar sociale stress. In **Hoofdstuk VI** testten we een nieuwe methode om emotionele expressies in het gezicht tijdens een gecomputeriseerd sociaal-exclusie paradigma minutieus te coderen. We vonden dat gezichtsuitdrukkingen van boosheid en verdriet samenhangen met de gevoelens die het kind zelf rapporteerde te hebben gehad tijdens de sociale exclusie. Hieruit concluderen we dat deze nieuwe codeermethode goed bruikbaar is om emoties tijdens sociale stress te meten. In **Hoofdstuk VII** hebben we op epigenoomwijd niveau gekeken of gepest worden samenhangt met verandering in DNA-methylering in Generation R en ALSPAC. In deze longitudinale studie konden we 'gepest worden' relateren aan verandering in DNA-methylering voor één site op *RAB14*. Dit gen is belangrijk voor het functioneren van het Golgi apparaat en is eerder in verband gebracht met blootstelling aan stress in knaagdieronderzoek.

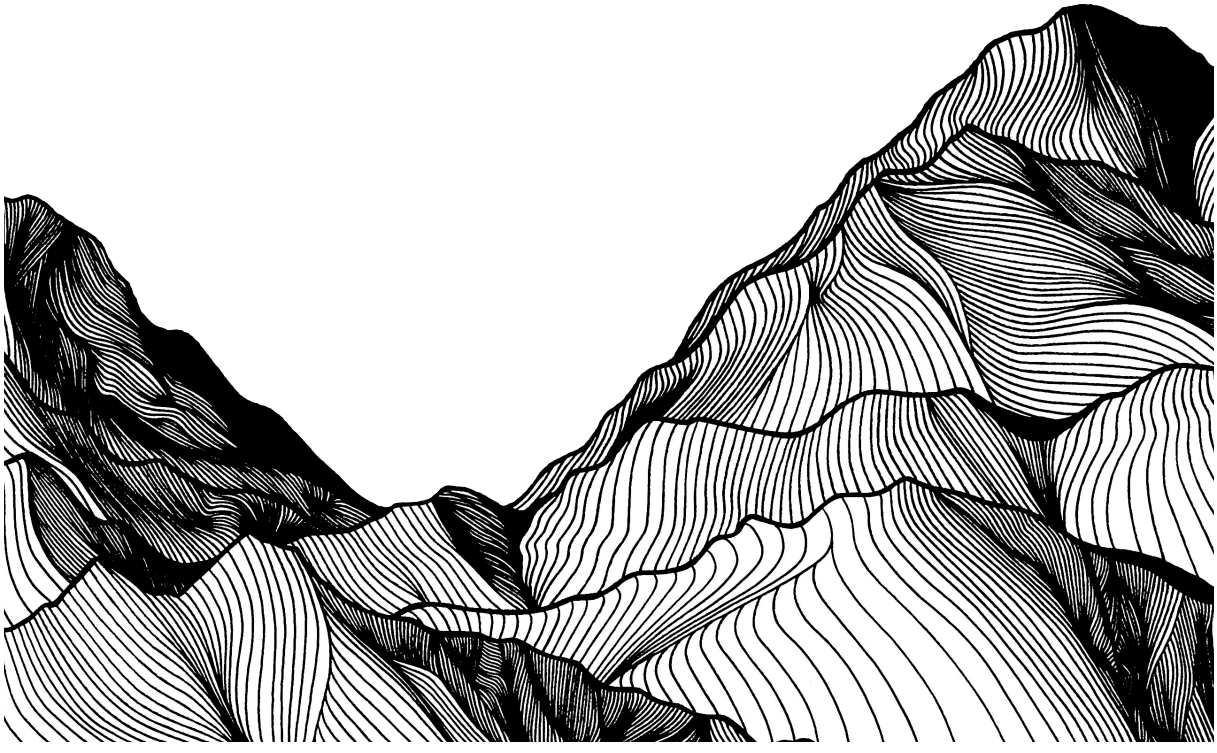
In onze laatste studie, beschreven in **Hoofdstuk VIII**, hebben wederom de krachten van Generation R en ALSPAC gebundeld om een belangrijke valideitskwestie binnen het onderzoek naar DNA-methylering aan te kaarten, namelijk in hoeverre genoomwijde DNA-methylering verandert over de tijd. We hebben de longitudinale DNA-methylering datasets van beide cohorten samengevoegd tot één set met metingen van de geboorte tot de late adolescentie. We onderzochten de epigenoom-wijde verandering in DNA-methylering, waarbij we zowel verschillen binnen als tussen personen in beeld brachten. We vonden dat iets meer dan de helft van de methylatie sites veranderen over de tijd en dat voor iets meer dan een kwart van de sites de richting van verandering verschillend is voor verschillende personen. De website die we hebben ontwikkeld om deze resultaten te presenteren kan als springplank fungeren voor andere onderzoekers door bevindingen van DNA-methylering in een longitudinaal perspectief te plaatsen.

Uit deze studiereeks concluderen we dat er associaties zijn tussen stress en DNA-methylering in de ontwikkeling van het kind. Deze associaties zijn klein. Daarom raden we aan om onderzoek naar DNA-methylering uit te voeren in grote groepen en waar mogelijk methoden voor datareductie te benutten.



ADDENDUM

Author affiliations
List of publications
About the author
PhD portfolio
Dankwoord (words of gratitude)
Acknowledgements



Authors and affiliations

Department of Child and Adolescent Psychiatry/Psychology, Erasmus MC, Erasmus University Medical Center - Sophia Children's Hospital, Rotterdam, the Netherlands
Rosa H. Mulder, Alexander Neumann, M. Elisabeth Koopman-Verhoeff, Jolien Rijlaarsdam, Maartje P. C. M. Luijk, Charlotte A. M. Cecil, Frank C. Verhulst, Henning Tiemeier

The Generation R Study Group, Erasmus MC, Erasmus University Medical Center, Rotterdam, the Netherlands
Rosa H. Mulder, Alexander Neumann, M. Elisabeth Koopman-Verhoeff, Jolien Rijlaarsdam, Janine F. Felix, Irwin Reiss, Vincent W. V. Jaddoe

Institute of Education and Child Studies, Leiden University, Leiden, The Netherlands
Rosa H. Mulder

Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, Qc, Canada
Alexander Neumann

EP Bradley Hospital Sleep Laboratory, Alpert Medical School of Brown University, Providence, RI, United States of America
M. Elisabeth Koopman-Verhoeff, Jared M. Saletin, Mary A. Carskadon

Alpert Medical School of Brown University, Department of Psychiatry and Human Behavior, Providence, RI
M. Elisabeth Koopman-Verhoeff, Mary A. Carskadon, Jared M. Saletin

Department of Molecular Genetics, Erasmus University Medical Center, Rotterdam, the Netherlands
Gijsbertus T.J. van der Horst

School of Mathematics, Statistics and Applied Mathematics, National University of Ireland, Galway, Ireland
Andrew J. Simpkin

Molecular Epidemiology, Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, The Netherlands
Bastiaan T. Heijmans

Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
Charlotte A. M. Cecil, Vincent W. V. Jaddoe

Department of Psychology, Institute of Psychology, Psychiatry & Neuroscience, King's College London, London, UK
Charlotte A. M. Cecil

MRC Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
Esther Walton, Lotte C. Houtepen, Andrew J. Simpkin, Tom R. Gaunt, Matthew Suderman, Caroline L. Relton

Department of Psychology, University of Bath, Bath, UK
Esther Walton

Department of Pediatrics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
Janine F. Felix, Irwin Reiss, Vincent W. V. Jaddoe

Child and Adolescent Mental Health Centre, Mental Health Services Capital Region, Research Unit, Copenhagen University Hospital, Copenhagen, Denmark
Frank C. Verhulst

Clinical Child & Family Studies, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
Marian J. Bakermans-Kranenburg

Department of Social and Behavioral Science, Harvard TH Chan School of Public Health, Boston USA
Henning Tiemeier

School of Clinical Medicine, University of Cambridge, Cambridge, UK
Marinus H. van IJzendoorn

Department of Psychology, Education and Child Studies, Erasmus University Rotterdam, Rotterdam, the Netherlands
Maartje P. C. M. Luijk, Rianne Kok, Marinus H. van IJzendoorn

Publications

This thesis

Mulder, R. H., Rijlaarsdam, J., & van IJzendoorn, M. H. (2017). DNA Methylation: a mediator between parenting stress and adverse child development? *Parental Stress and Early Child Development* (pp. 157-180): Springer.

Mulder, R. H., Rijlaarsdam, J., Luijk, M. P. C. M., Verhulst, F. C., Felix, J. F., Tiemeier, H., . . . van IJzendoorn, M. H. (2017). Methylation matters: FKBP506 binding protein 51 (FKBP5) methylation moderates the associations of FKBP5 genotype and resistant attachment with stress regulation. *Developmental Psychopathology*, 29(2), 491.

Dall'Aglio, L., Rijlaarsdam, J.* Mulder, R. H.*, Neumann, A., Felix, J. F., Kok, R., Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., Tiemeier, H., Cecil, C. A. M. (in press). Epigenome-wide associations between observed maternal sensitivity and offspring DNA methylation: A population-based prospective study in children. *Psychological Medicine*.

Koopman-Verhoeff, M. E., Mulder, R. H., Saletin, J. M., Reiss, I., van der Horst, G. T. J., Felix, J. F., . . . Cecil, C. A. M. (2020). Genome-wide DNA methylation patterns associated with sleep and mental health in children: a population-based study. *Journal of Child Psychology and Psychiatry*.

Mulder, R. H., Walton, E., Neumann, A., Houtepen, L. C., Felix, J. F., Bakermans-Kranenburg, M. J., . . . Relton, C. L. (2020). Epigenomics of being bullied: changes in DNA methylation following bullying exposure. *Epigenetics*, 15(6-7), 750-764.

Mulder, R. H., Neumann, A. H.*, Cecil, C. A. M.*, Walton, E., Houtepen, L. C., Simpkin, A. J., . . . Suderman, M. (2020). Epigenome-wide change and variation in DNA methylation from birth to late adolescence. *bioRxiv*.

Not part of this thesis

Caramaschi, D., Hatcher, C., Mulder, R. H., Felix, J. F., Cecil, C. A. M., Relton, C. L., & Walton, E. (2020). Epigenome-wide association study of seizures in childhood and adolescence. *Clinical Epigenetics*, 12(1), 1-13.

van Dongen, J., Hagenbeek, F. A., Suderman, M., Roetman, P. J., Sugden, K., Chiocchetti, A. G., Ismail, K., Mulder, R. H. . . . Boomsma, D. I. (in press). DNA methylation signatures of aggression and closely related constructs: A meta-analysis of epigenome-wide studies across the lifespan. *Molecular Psychiatry*.

Jansen, P. R., Dremmen, M., Van Den Berg, A., Dekkers, I. A., Blanken, L. M. E., Muetzel, R. L., Bolhuis, K., Mulder, R. H. . . . White, T. J. H. (2017). Incidental Findings on Brain Imaging in the General Pediatric Population. *The New England Journal of Medicine*, 377(16), 1593.

Muetzel, R. L., Mulder, R. H., Lamballais, S., Cortes, A. P., Jansen, P., Güroğlu, B., . . . El Marroun, H. (2019). Frequent bullying involvement and brain morphology in children. *Frontiers in Psychiatry*, 10, 696.

Sammallahti, S.*, Cortes Hidalgo, A. P.*, Tuominen, S., Malmberg, A., Mulder, R. H., Brunst, K. J., Alemany, S., . . . Lahti, J. (in press). Maternal anxiety during pregnancy and newborn epigenome-wide DNA methylation. *Molecular Psychiatry*.

White, T., Muetzel, R. L., El Marroun, H., Blanken, L. M. E., Jansen, P., Bolhuis, K., Kocevska, D., Mous, S. E., Mulder, R. H., . . . Tiemeier, H. (2018). Paediatric population neuroimaging and the Generation R Study: the second wave. *European Journal of Epidemiology*, 33(1), 99-125.

*authors contributed equally.

About the author

Rosa Henriëtte Mulder was born on the 24th of May 1988 in Amsterdam, the Netherlands. She grew up in Abcoude and enjoyed her pre-university secondary education at St. Ignatiusgymnasium in Amsterdam, where she graduated with a Nature and Health (Natuur en Gezondheid) profile in 2006. She studied Psychology at the University of Amsterdam (UvA) and obtained her bachelor's degree with merit in 2009. She followed up with another bachelor program, Neurobiology (Psychobiologie), at the University of Amsterdam UvA, and obtained her degree in 2011. After backpacking through South America for half a year, she continued her university studies with a selective Research Master in Psychology at the UvA in 2012, with Clinical Psychology and Brain and Cognition as focus areas. She obtained her master's degree cum laude in 2014. During her studies, Rosa developed an interest in research and worked as a research assistant at the Social Psychology department and as a teaching assistant at the laboratory practicals of the Biology department of the UvA. In November of 2014, she started working as a PhD candidate at the Center for Child and Family Studies of Leiden University in Leiden and at the Department of Child and Adolescent Psychiatry/ Psychology and Generation R Study Group at the Erasmus MC-Sophia in Rotterdam, resulting in the work presented in this thesis. During this candidacy, amongst others, she was responsible for two major projects. The first pertains the development of an observational measurement method for micro-coding of facial expressions and supervision of the subsequent coding of more than 6000 videos. For the second project, she studied the longitudinal change of epigenome-wide DNA methylation from birth to late adolescence in samples from both the Generation R and the Avon Longitudinal Study of Parents and Children (ALSPAC) in the UK, and built an accompanying interactive website. In 2017 she was awarded a Van der Gaag Grant of the Royal Netherlands Academy of Arts and Sciences (Koninklijke Nederlandse Akademie van Wetenschappen, KNAW), which gave her the opportunity to work as a visiting fellow at the MRC Integrative Epidemiological Unit – an expertise centrum for genetic and epigenetic epidemiology – of Bristol University in the United Kingdom. From late 2018 onwards, Rosa started a postdoc position studying bullying and brain development at Generation R within the larger consortium of NeurolabNL – a Dutch nationwide collaboration for brain, cognitive, and behavioral research. From 2020 onwards, she has combined this position with a postdoc position at Generation R within the EarlyCause consortium – a European collaboration studying early life stress and psychological and cardiometabolic outcomes, which allows her to follow her passion for research of stress and development.

PhD portfolio

Name PhD student:	Rosa Mulder
Erasmus MC department:	The Generation R Study
	Child and Adolescent Psychiatry/Psychology
PhD period:	November 2014 - January 2021
Promotors:	Prof. dr. M. H. van IJzendoorn
	Prof. dr. H. Tiemeier
Copromotor:	Prof. dr. M. J. Bakermans-Kranenburg

PhD training			
	Institute	**Year**	**ECTS***
General courses			
Applied Multivariate Data Analysis	Leiden University	2015	5.0
Basic Course on R	Erasmus MC	2017	1.8
Writing a Grant Proposal	Leiden University	2017	0.5
Scientific Integrity	Erasmus MC	2020	0.3
Epidemiology courses			
Principles of Genetic Epidemiology	Erasmus MC	2015	0.7
Epigenetic Epidemiology	Bristol University	2015	0.6
Multilevel Modelling in R	Leiden University	2017	1.0
Epigenetic Clock Symposium	Bristol University	2018	0.3
GWAS Blitz Course	Erasmus MC	2019	1.0
Other specialized courses			
E-prime Programming	Leiden University	2015	0.5
Linux for Scientists	Erasmus MC	2015	0.6
Current Issues in Clinical Neuroscience	Utrecht University	2016	1.5
Attachment and Developmental Psychopathology	Leiden University	2016	5.0
Training			
Emotional Availability Scales observational coding	Leiden University	2015	2.0
MRI incidental findings	Erasmus MC	2015	1.5
MRI safety	Erasmus MC	2016	1.0
Wechsler Intelligence Scale for Children testing	Erasmus MC	2016	1.0
Seminars			
Colloquia Family and Child Studies	Leiden University	2014-2018	2.0
Research meetings Generation R	Erasmus MC	2014-2020	3.0

PhD training–*continued*			
	Location	**Year**	**ECTs**
National meetings			
Aggression in Children: unraveling gene-environment interplay to inform Treatment and Intervention strategies (ACTION) meeting	Amsterdam	2016	0.3
Pregnancy And Childhood Epigenetics (PACE) consortium meeting	Rotterdam	2017	0.3
PACE consortium meeting	Rotterdam	2019	0.3
International conferences and meetings			
Society for Research in Child Development (SRCD) biennal meeting (two poster presentations)	Austin, Texas, USA	2017	2.0
2-day research visit IEU–MRC Integrative Epidemiological Unit of Bristol University (oral presentation)	Bristol, UK	2017	1.0
2-week research visit IEU–MRC Integrative Epidemiological Unit of Bristol University (oral presentation)	Bristol, UK	2017	1.0
6th Haruv International PhD Workshop on Child Maltreatment (oral presentation)	Jerusalem, Israel	2017	1.2
ACTION meeting (oral presentation)	Sardinia, Italy	2019	1.2
Epigenomics of Common Diseases (poster presentation)	Cambridge, UK	2019	1.0

Teaching activities			
	Institute	**Year**	**ECTS**
Supervision bachelor- and mastertheses			
Luna Vesseur, master student Psychology: *Stress and ostracism: A diffusion tensor imaging-study*	Leiden University	2015-2016	3.0
Jiske van Zeijl, master student Psychology: *The relationship between symptoms of autism spectrum disorder and facial expression during social exclusio*n	Leiden University	2015-2016	3.0
Johan Veenstra, master student Psychology: *The relationship between problem behavior and facial expression during ostracism: A new facial expression scoring system*	Leiden University	2015-2016	3.0
Lisette van de Graaf, bachelor student Pedagogy: *Associatie tussen hechting en emoties: De invloed van ouder-kind interacties in de eerste levensjaren op emotionele uitingen van kinderen op negenjarige leeftijd tijdens Cyberball*	Erasmus University	2016	3.0
Manouk Stam, master student Psychology: *De relatie tussen hechting, psychopathologie van de ouders en emotionele reactie op sociale exclusie*	Erasmus University	2017	3.0
Eleni Marantzani, master student Pedagogy: *Family functioning an childend's emotion regulation after social exclusion: The moderating role of child gender*	Erasmus University	2017	3.0

Teaching activities–*continued*			
	Institute	**Year**	**ECTS**
Burcu Özaydin, master student Pedagogy: *De relatie tussen pesten, sensitiviteit van de moeder en emotionele reactie op sociale exclusie*	Erasmus University	2017	3.0
Lorenza Dall'Aglio, master student Psychology: *Maternal sensitivity and DNA methylation in the child*	Erasmus University		1.0
Group-supervision theses (scriptie-atelier)	Leiden University	2017–2018	1.0
Training of facial micro-expression coding			
Training students in coding of facial micro-expressions	Erasmus MC	2015-2017	6.0
Lecture			
Stress and epigenetics for 2nd year bachelor students	Erasmus University	2018	0.3
Reviewing assignments and exams			
Reviewing assignments and exams within the Pedagogy bachelor and master studies	Leiden University	2015-2017	4.0

Data collection		
	Year	**ECTS**
General data collection		
Generation R data collection Focus@9 and Focus@13, including intelligence testing children (WISC), stressful life events interviews parents, MRI testing, phone and personal contact with parents	2014-2017	50.0
Social exclusion task and facial micro-expressions (Cyberball)		
Coordination social exclusion task research center	2014-2016	1.0
Development of coding system and accompanying software for facial micro-expressions	2014-2016	3.0
Coded over 600 videos for facial micro-expressions	2016-2017	3.0
Other data collection		
Checked 500 MRI scans for incidental findings	2015-2017	3.0
Programming and coordination of moral decision task research center	2016-2017	4.0

Reviews		
	Year	**ECTS**
PLOS ONE	2018	0.5
Development and Psychopathology	2020	0.5

Scientific outreach		
	Year	**ECTS**
Built freely available website presenting interactive graphs regarding longitudinal DNA methylation (http://epidelta.mrcieu.ac.uk/)	2019-2020	5.0
Blog on longitudinal DNA methylation (https://ieureka.blogs.bristol.ac.uk/2020/08/11/epigenetics-regulate-our-genes-but-how-do-they-change-as-we-grow-up/)	2020	0.3

Fellowship		
	Institute	**Year**
Four-month fellowship at MRC Integrative Epidemiological Unit–expertise centrum for genetic and epigenetic epidemiology–at Bristol University, UK, working bullying and epigenetic development of children of the ALSPAC cohort	Bristol University	2018

Grants	
	Year
Van der Gaag Grant of the Royal Netherlands Academy of Arts and Sciences (Koninklijke Nederlandse Akademie van Wetenschappen, KNAW), for a fellowship at Bristol University, UK (€6,600)	2017
Travel grant for 6th Haruv International PhD Workshop on Child Maltreatment in Jerusalem, Israel	2017
Travel grant *Epigenomics for Common Diseases* conference in Cambridge, UK	2019

* 1 ECTS (European Credit Transfer System) is equal to a workload of 28 hours

Dankwoord
(words of gratitude)

Dit boekje had nooit tot stand kunnen komen zonder de helpende handen, waardevolle bijdragen en steunbetuigingen van velen. Allereerst wil ik de kinderen en ouders van Generation R bedanken. Jullie trouwe toewijding en openheid in eindeloze interviews en vragenlijsten maakt Generation R tot wat het is. Dan alle focusmedewerkers; de bezige bijen op het onderzoekscentrum die de boel binnen strakke tijdsschema's draaiende houden, onder de geduldige supervisie van Ronald. En speciaal de dames van het gedragsblok–Anneke, Rukiye, Sabah, Ineke en later ook Yvette – veel dank voor jullie inzet, vooral ook met betrekking tot de afname van Cyberball. Michael, op het lab, dank je voor je werk aan de DNA methylatiedata en je uitleg. Datamanagement natuurlijk, Claudia, Marjolein, Annemiek en Eline, dank voor jullie hulp. Op de verschillende secretariaten in Rotterdam: Patricia, Erica, Laureen, Mireille, dank, dank, dank. In Leiden: Gea en Esther, dank voor de kletspraatjes en de flexibiliteit die af en toe nodig was door het schipperen tussen twee universiteiten. Ook de vele stagiaires die hebben geholpen in het ontwikkelen van een codeermethode van Cyberball en het coderen van de vele video's wil ik graag bedanken: Luna, Jiske, Johan, Lisette, Burcu, Manouk, Eleni. Johan en Etiënne, jullie hebben een prachtig programma ontwikkeld waarmee we een immense hoeveelheid data snel konden coderen. Eleni, you pushed through the last videos and played a big part in finalizing the dataset.

Rien en Marian, terwijl jullie je in een onvoorzien moeilijke en slepende situatie bevonden heb ik me altijd gesteund gevoeld door jullie. Veel dank voor jullie toewijding. Henning, ook na jouw emigratie kon ik altijd rekenen op kritisch commentaar. Je daagt me uit en bent een leermeester waar menig onderzoeker jaloers op zou zijn. Dank voor je begeleiding.

Vanuit de afdeling Kinder- en Jeugdpsychiatrie en Psychologie wil ik graag ok Frank en Manon bedanken voor hun steun. Frank, ook bedankt voor je co-auteurschap.

De commissieleden, Pauline, wat fijn dat je secretaris wil zijn van mijn leescommissie, het is een eer. Professor Relton, Caroline, I learned so much at the IEU in Bristol and am honored that you agreed to be part of the reading committee. Professor Slagboom, hartelijk dank voor de tijd en moeite die u neemt om deel uit te maken van de leescommissie. Janine, de DNA-methylering data in Generation R staat of valt bij jouw inzet, ik ben daarom dankbaar dat je onderdeel wilt zijn van de oppositiecommissie. Daarnaast ben ik ontzettend blij dat ik mijn werk aan deze dataset mag voortzetten onder jouw supervisie. Professor Williams, the Cyberball paradigm has been an integral part of my PhD; I am honored that you agreed to be

a part of my opposition committee. Professor Franken, hartelijk dank dat u wilt plaatsnemen in de oppositiecommissie.

Naar werk gaan was altijd leuk, dankzij de gezellige collega's van de Gedragsgroep. Koen, mijn Roffa-buddy, Pamela wil graag weer witte-chocolade-oreos eten met Sjoerd! Elize, ik vind het nog steeds een raar idee dat je niet meer op de afdeling werkt. Gelukkig hebben wij nog wat slaapprojecten lopen! Ivonne, met jou door Rotterdam racen in een klein Generation R autootje was de beste vervulling van de algemene taken. Drinken uit waterflesjes terwijl jij rijdt heb ik snel afgeleerd. Yllza, you are such a lovely person and obviously, I've been missing your chocolates a lot lately! Michiel, wat fijn om de promotieperikelen met jou te kunnen delen! Binnenkort zijn we allebei klaar! Clair, ons adoptiekind binnen de Gedragsgroep, jouw vrolijkheid en enthousiasme waren altijd een welkome afleiding. Gosia, die trip naar Barcelona komt eraan! Alex, I feel lucky to have been able to work with you. Your statistical insight (I know you think otherwise) has definitely elevated my work. Fadila, never have I met such a wonderfully chaotic person. I am happy our paths collided. Desi, I remember being placed next to you in our office when I just started, and eyeballing you to get an idea if we could be friends. That turned out pretty well. You made my time in Rotterdam so much more worthwhile, I cannot imagine how it would have been without you. Lorenza, starting as a master student and now fulfilling your own PhD-candidacy, thank you for your contribution, it really ties this thesis together. Ryan, is there a more reasonable person than you? I feel so lucky I get to learn all about neuroimaging from you. Jolien, beide zijn we in de afgelopen periode heen-en-weer gebounced tussen het EMC en Universiteit Leiden. Dank je voor je begeleiding. Charlotte, your work is inspiring and I was so happy that you came to Rotterdam, am thankful for your supervision, and delighted we get to continue to work together. A 'thank you' to many others of the Behavioral group that make for such a wonderful working environment – I pray I didn't miss anyone: Hanan, Andrea, Fernando, Monica (looking at brains has never been so fun), Tessa, Lisa, Elisabeth (it's all about the pineal gland), Runyu, Deborah, Sander, Sara (the meta-analysis-queen), Louk, Annemarie, Mannan (and the Scottish egg with Watson and Crick), Cees, Novika, Lisanne, Lizzy, Tessa, Nienke (convergeren kun je leren), Isabel, Serena, Patricia, as well as all those that I've worked with before but have gone on to work in different places.

Twee daarvan toch even direct: Dafna en Andrea, nu alweer zo lang geleden, maar ontzettend fijn dat ik bij jullie terecht kon aan het begin van mijn PhD voor alle Leiden-Generation R gerelateerde vragen! Rens, ook jij bedankt voor je hulp bij het programmeren van het Cyberball paradigma. En niet te vergeten de vele andere collega's in Leiden die mijn werkdagen opleukten. Katharina, met jou heb ik vele problemen beklommen, of het nou ging om Leidse perikelen, of om Italiaans rotsformaties. Renate, de kletspraatjes met jou waren altijd zeer welkom! Fenne, kamergenoot wiens bureau altijd geordender was dan dat van mij, dank voor je gezelschap. Chloë, je droge humor gebracht met zacht fluwelen Vlaamse tongval

is ongeëvenaard. Esther, met jou valt er altijd wat te lachen, op de afdeling, op congres in Austin, én tijdens het crossen over Texaanse achterwegen.

Then many thanks to the amazing team at ALSPAC: Esther, Lotte, Matt, Caroline, thank you for giving me the opportunity to learn so much. Matt, thank you for introducing me to the wonderful world of longitudinal modelling. Thank you for your supervision, and for all your help in setting up the EpiDelta website. Professor Gaunt, Tom, thank you as well for morphing something that was just a piece of code, into an actual website people can actually visit. Esther, your analytical eye has improved many of my analyses, and besides, you're just a wonderful person. I feel lucky to be able to continue to work with you. Lotte, een klein stukje Nederland in Bristol, dank je voor al je hulp in de 'Epi-' projecten en wat leuk dat ik nu af en toe een klein stukje Bristol tegen het lijf loop op consortium meetings in Nederland! Nancy, Charlie, Matt, Steph, Paul, Nour, and of course Jeroen, thank you for making me feel so welcome at the department. Nancy, Steph, Paul, can we go swimming in the woolly whatever again please? Tasha, partner in climb, thank you for showing me the best of Bristol. Giordano, the best housemate I could've wished for, can you play the guitar yet?

Dan alle co-auteurs die ik nog niet genoemd heb: Maartje, Vincent, Bas, dank voor de samenwerking en jullie waardevolle bijdragen aan de stukken in dit boekje. Andrew, thank you for your collaboration, thank you for your help in the nonlinear models.

Stefanie, Sabrina, Jentien, jullie vriendschap is me goud waard. Steef, jij wist gedurende deze PhD periode altijd haarfijn aan te voelen wanneer het weer eens tijd werd mij naar Amsterdam toe te halen. Sab, fantastisch mens dat je er bent, ook een afstand(je) voel ik me gesteund door jou. Jeanny, je bent een lieverd en ik ben blij dat ik af en toe even met je kon nerden over onderzoek. Roel, ook met jou is het heerlijk nerden. Indra, ons middagje scripten was super. Mickey, wat fijn dat jij en Juul weer in het land zijn, met jou is het altijd een feestje. Div, talking with you about everything and nothing is the best pastime there is. Especially when it is in a tree house in between the coconut-palms in Chitrapur. Benny-boy, you're mostly far away but somehow you manage to keep me close. Counting seven cities in five countries. Where to next? Diana, I love how we went through the same PhD process from different places in the world. Un apapacho desde Holanda. Nellie, van onbevangen psychologiestudentjes naar verstandige volwassenen (of zoiets), dank voor al je steun en gezelligheid! Megan en Marg, vriendinnen van het eerste uur, dank voor jullie vriendschap, voor, tijdens en na deze PhD.

Papa en mama, dankzij jullie heb ik altijd geloofd in mijn eigen kunnen. Papa, jouw mogelijkheid om iets vanuit vele verschillende perspectieven te benaderen heeft mij altijd geïntrigeerd en inspireert me om altijd kritisch blijven. Mama, jouw nuchterheid en inzicht hebben mij juist geholpen om dingen vooral niet ingewikkelder te maken dan ze hoeven zijn. Lieve Nienke,

als er iemand is die me heeft laten zien dat je met een duidelijk doel en een goede dosis doorzettingsvermogen precies kunt bereiken wat je wilt, ben jij het wel.

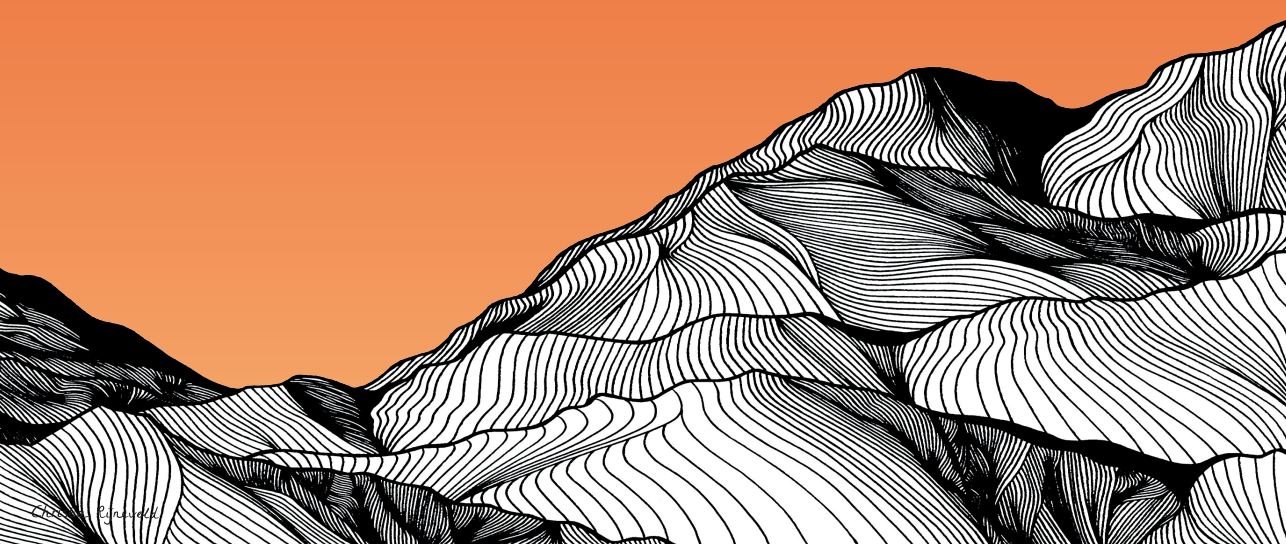
Some things, like this thesis, take effort and patience. Other ‘things’, like you, Michalis, just sort of seem to appear. Thank you for your support through the last part of my PhD, thank you for wandering into my life. Who knew a heart could be so full?

Acknowledgements

The Generation R Study is conducted by the Erasmus Medical Center in close collaboration with the Faculty of Social Sciences of the Erasmus University Rotterdam, the Municipal Health Service Rotterdam area, Rotterdam, the Rotterdam Homecare Foundation, Rotterdam and the Stichting Trombosedienst & Artsenlaboratorium Rijnmond (STAR-MDC), Rotterdam. We gratefully acknowledge the contribution of children and parents, general practitioners, hospitals, midwives and pharmacies in Rotterdam. The study protocol was approved by the Medical Ethical Committee of the Erasmus Medical Centre, Rotterdam. Written informed consent was obtained for all participants. The generation and management of the Illumina 450K methylation array data (EWAS data) for the Generation R Study was executed by the Human Genotyping Facility of the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, the Netherlands. We thank Mr Michael Verbiest, Ms Mila Jhamai, Ms Sarah Higgins, Mr Marijn Verkerk and Dr. Lisette Stolk for their help in creating the EWAS database. We thank Dr A. Teumer for his work on the quality control and normalization scripts. The general design of the Generation R Study is made possible by financial support from the Erasmus Medical Center, Rotterdam, the Erasmus University Rotterdam, the Netherlands Organization for Health Research and Development and the Ministry of Health, Welfare and Sport. The EWAS data was funded by a grant from the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) Netherlands Consortium for Healthy Aging (NCHA; project nr. 050-060-810), by funds from the Genetic Laboratory of the Department of Internal Medicine, Erasmus MC, and by a grant from the National Institute of Child and Human Development (R01HD068437).

We are extremely grateful to all the families who took part in the ALSPAC study, the midwives for their help in recruiting them, and the whole ALSPAC team, which includes interviewers, computer and laboratory technicians, clerical workers, research scientists, volunteers, managers, receptionists and nurses. The UK Medical Research Council and Wellcome (grant number 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. A comprehensive list of grants funding is available on the ALSPAC website (http://www.bristol.ac.uk/alspac/external/documents/grant-acknowledgements.pdf).

Analysis of the ALSPAC data was funded by UK Economic & Social Research Council grant (grant number ES/N000498/1). ARIES was funded by the BBSRC (BBI025751/1 and BB/I025263/1). Supplementary funding to generate DNA methylation data which are (or will be) included in ARIES has been obtained from the MRC, ESRC, NIH and other sources. ARIES is maintained under the auspices of the MRC Integrative Epidemiology Unit at the University of Bristol (grant numbers MC_UU_00011/4 and MC_UU_00011/5).

Rosa Mulder was additionally supported by the Van der Gaag Grant of the Royal Netherlands Academy of Arts and Sciences and The Netherlands Organization for Scientific Research (Koninklijke Nederlandse Akademie van Wetenschappen).

EPIGENETICS HAS BEEN LAUDED AS THE PUTATIVE MECHANISM THROUGH WHICH 'NATURE AFFECTS NURTURE'. SUCH A MECHANISM COULD EXPLAIN HOW EARLY STRESSFUL EXPERIENCES CAN HAVE MAJOR PSYCHOLOGICAL CONSEQUENCES LATER IN LIFE. AN OFTEN-STUDIED FORM OF EPIGENETICS IS DNA METHYLATION. HOWEVER, RESEARCH ON DNA METHYLATION IS TYPICALLY LIMITED BY SMALL SAMPLE SIZES AS WELL AS CROSS-SECTIONAL STUDY DESIGNS, WHICH MAKE IT DIFFICULT TO INTERPRET RESEARCH RESULTS WITHIN A DEVELOPMENTAL FRAMEWORK. IN THE CURRENT THESIS, WE THEREFORE STUDIED DNA METHYLATION AND STRESS THROUGHOUT CHILD DEVELOPMENT IN A POPULATION-BASED APPROACH.