



King's Research Portal

DOI:

[10.1111/jcpp.12782](https://doi.org/10.1111/jcpp.12782)

Document Version

Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Barker, E. D., Walton, E., & Cecil, C. A. M. (2017). 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*. <https://doi.org/10.1111/jcpp.12782>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Annual Research Review:

DNA methylation as a mediator in the association between risk exposure and child and adolescent psychopathology

Edward D. Barker¹, Esther Walton² & Charlotte A.M. Cecil¹

SELF-ARCHIVING VERSION

¹Department of Psychology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, UK.

²Medical Research Council Integrative Epidemiology Unit, University of Bristol, UK

Short title: DNAm and child psychopathology

Conflict of interest: For all authors no conflict of interest is declared.

Corresponding Authors: Edward D. Barker and Charlotte AM Cecil

Department of Psychology, King's College London. Institute of Psychiatry, Psychology and Neuroscience, De Crespigny Park, London, SE5 8AF. Email: ted.barker@kcl.ac.uk (Phone: +44 (0)207 848 0992) or charlotte.cecil@kcl.ac.uk.

Citation: Barker ED, Walton E & Cecil CAM (2017). 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*

Key Points

- DNA methylation is a potential mechanism for propagating the effects of environmental risk exposures on child and adolescent mental health;
- There is recent growth in publications assessing DNA methylation from peripheral samples (blood, buccal cells) in children and adolescents;
- Very few published studies exist that can test key mediational DNAm hypotheses – ones that measure risk exposures, DNAm, and psychopathology in the context of a prospective longitudinal design;
- We provide a strategic review of existing studies on living children – with risk exposure and/or psychopathology measured in childhood or adolescence
- We highlight current challenges that, if overcome, can help assess DNA methylation as a causal mechanism, or biomarker, in the link between environmental risk exposure and child and adolescent psychopathology

Abstract

Background: DNA methylation (DNAm) is a potential mechanism for propagating the effects of environmental exposures on child and adolescent mental health. In recent years, this field has experienced steady growth. **Methods:** We provide a strategic review of the current child and adolescent literature in order to evaluate evidence for a mediating role of DNAm in the link between environmental risks and psychopathological outcomes, with a focus on internalising and externalising difficulties. **Results:** Based on the studies presented, we conclude that there is preliminary evidence to support that (i) environmental factors, such as diet, neurotoxic exposures and stress, influence offspring DNAm, and that (ii) variability in DNAm, in turn, is associated with child and adolescent psychopathology. Overall, very few studies have examined DNAm in relation to *both* exposures and outcomes, and almost all analyses have been correlational in nature. **Conclusions:** DNAm holds potential as a biomarker indexing both environmental risk exposure and vulnerability for child psychopathology. However, the extent to which it may represent a causal mediator is not clear. In future, collection of prospective risk exposure, DNAm and outcomes – as well as functional characterization of epigenetic findings – will assist in determining the role of DNAm in the link between risk exposure and psychopathology.

Keywords: DNA methylation; Epigenetics; Developmental Psychopathology; Childhood; Adolescence; Environmental risk; Internalising problems; Externalising problems

Abbreviations: 5hmC, 5-hydroxymethylcytosine; ADHD, attention deficit hyperactivity disorder; CP, conduct problems; CpG, cytosine-guanine pair; CU, callous-unemotional; DNAm, DNA methylation; EWAS, epigenome-wide association studies; HPA, hypothalamic–pituitary axis; mQTL, methylation quantitative trait locus; MR, Mendelian randomization; ODD, oppositional defiant disorder; WGBS, whole-genome bisulfite sequencing.

Introduction

Data indicate the importance of early adversity in the development of child and adolescent psychopathology (Barker & Maughan, 2009; O'Connor, Heron, Golding, Beveridge, & Glover, 2002). Indeed, both prenatal (e.g. maternal stress, psychopathology) and postnatal (e.g. childhood maltreatment) adversities have been found to increase risk for a wide range of negative mental health outcomes, including internalising and externalising difficulties. Importantly, the effects of early adversity can persist well into the adult years, long after the exposure itself has ceased (e.g. Yehuda, McFarlane, & Shalev, 1998). Yet, the exact mechanisms mediating this enduring vulnerability remain unclear. Consequently, a key challenge for research is to understand how adverse experiences impact biological development (and function) in a manner that engenders long-term risk for psychopathology.

In recent years DNA methylation (DNAm) – an epigenetic process that regulates gene expression – has emerged as a potential mechanism through which the genome can ‘capture’ the effects of environmental exposures and propagate their influence (McKay et al., 2012). DNAm refers to the addition of a methyl group, primarily in the context of cytosine-guanine (CpG) dinucleotides (Jaenisch & Bird, 2003). In the human genome, CpG sites often cluster in CpG-islands, which themselves tend to be embedded in promoter regions of genes. Methylated CpG islands impede transcription factors from accessing the DNA sequence. Increased methylation in these regions is typically associated with inhibition of gene transcription (i.e. gene silencing) and chromatin compaction, and hence can provide a mechanism by which the DNAm can trigger long-term alterations in phenotypes. Of note, DNAm is also influenced by genetic factors (McRae et al., 2014). Indeed, twin research has shown that DNAm is highly heritable near the promoter regions of genes (Kaminsky et al., 2009) and molecular studies report that genetic influence on DNAm levels in peripheral samples (blood) can be identified as early infancy (cord samples at birth; Teh et al., 2014) and across the life course (Gaunt et al., 2016). The interest in DNAm within the field of developmental psychology stems from studies that suggest DNAm can be both sensitive to a range of environmental exposures across the lifespan (Kofink, Boks, Timmers, & Kas, 2013) and – at the same time – associate with altered biological processes underlying the emergence of disease states (Szyf, 2015). Of note,

Over the past decade, a growing number of studies have tested whether DNAm is associated with environmental risks or with different forms of psychopathology. As shown in **Figure 1**, DNAm studies have been steadily rising since 2010, with many more studies relating DNAm to risk exposures compared to psychopathological outcomes. Of these

published studies, the majority have focused on adult populations, with a much smaller number focusing on childhood. In fact, the least common studies are those that would make it possible to test key mediational DNAm hypotheses – ones that measure risk exposures, DNAm, and psychopathology in the context of a prospective longitudinal design. Thus, while the field is growing fast, there is still a relative dearth of studies featuring the necessary data to test DNAm as a mediator in the link between prior environmental risk exposure(s) and subsequent psychopathological outcomes during childhood and adolescence.

Delineating the mechanistic role of DNAm is crucial in order to clarify its potential utility in the prevention, detection, and treatment of psychopathology. The ‘promise’ of a mediational framework is that – if DNAm is identified as a causal link in the etiology of a disease – then reversing epigenetic marks (e.g. through multimodal interventions including epigenetic therapy; Szyf, 2015) might help alleviate the burdens of disease. It is also equally possible, however, that DNAm functions as a non-causal biomarker of environmental risk exposure and/or stress-related disorders. Here, differences in DNAm may be a consequence of disease etiology (e.g. risk exposure and/or psychopathology) rather than a causal mechanism within the disease process. In aetiologic epidemiology, this is termed “reverse causality” (Richmond, Al-Amin, Smith, & Relton, 2014). Yet even in this situation DNAm can still serve as an important biomarker of disease and have clinical utility. For example, epigenetic patterns have already been shown to be useful in cancer detection, prognosis and even predicting response to treatment (Ladd-Acosta & Fallin, 2015).

In this review, we discuss DNAm as a potential mechanism linking environmental risk exposure and the development of child and adolescent psychopathology. Figure 2 (panel A) contains the conceptual mediation model that guides this strategic review. Specifically, we provide an overview of studies (published by 15-Oct-2016) that have examined DNAm based on peripheral tissues (e.g. blood, saliva/buccal cells) in living children, with risk exposure and/or psychopathology measured in childhood or adolescence. This is therefore a strategic (non-systematic) review of the literature. First, we begin with a brief introduction to epigenetics in general, and DNAm in particular. Second, we review studies that have examined the *a path* of our proposed mediation model (i.e. exposure \rightarrow DNAm). We include environmental risks that have been previously associated with child and adolescent psychopathology, including dietary, neurotoxic and stress exposures. Third, we describe studies that have examined the *b path* of the model (i.e. DNAm \rightarrow psychopathology), focusing on internalising and externalising problems. We then highlight the small number of integrative studies that have examined DNAm as a mediator in the link between risk exposure

and psychopathology (i.e. the $a*b$ path). Finally, we close with a discussion of current challenges and propose five key ways to move the field forward in future.

Overview of the epigenome

Epigenetic mechanisms influence dynamic changes in transcription independent of the genomic DNA sequence, primarily via modifications to DNA, histone proteins, and chromatin structure (Jaenisch & Bird, 2003). Epigenetic processes are essential for normal cellular development and differentiation, and allow the long-term regulation of gene function through non-mutagenic mechanisms (Henikoff & Matzke, 1997). The term epigenetics was first introduced by Waddington, who tried to understand how cells differentiate into diverse tissue types, despite containing the same genome (Waddington, 1957). Nowadays, it is clear that epigenetic modifications are not only fundamental for the establishment and maintenance of cellular identity, but also co-ordinate a much wider range of biological processes, including genomic imprinting and X chromosome inactivation, as well as stress-response, immune function and neurodevelopment. Importantly, because epigenetic processes have been shown to respond to both genetic and environmental factors, they represent a potential mechanism that can help explain the biology of gene-environmental interplay and disease susceptibility across the lifespan (Meaney, 2010).

One of the most extensively researched epigenetic mechanisms is DNAm, which – as stated earlier – refers to the addition of a methyl group, primarily in the context of cytosine-guanine (CpG). This process affects about 4 % of CpG pairs, which in turn impacts the binding potential of transcription factors. Many (but not all) genes demonstrate an inverse correlation between the degree of methylation and the level of expression. While 5-methylcytosine (5mC) is the most common form of DNAm, particularly in peripheral tissues (blood, buccal cells), other DNAm processes have recently been discovered, such as 5-hydroxymethylcytosine (5hmC). As this review is focused on 5mC DNAm, we refer the reader to a high quality review that covers 5hmC as well as a range of other epigenetic mechanisms (Szyf, 2015).

DNAm can be studied through a variety of different approaches. Global DNAm is used as proxy for the overall degree of methylation in the genome and can be derived using methods such as high performance liquid chromatography or mass spectrophotometry. Candidate gene approaches focus on preselected genes based on a priori hypotheses (e.g. via pyrosequencing and related PCR or sequencing applications). Despite their utility, candidate gene approaches may offer a limited insight into the pathophysiology of complex diseases,

which are likely to reflect the influence of multiple genes of potentially unknown function. In contrast, the use of epigenome-wide, hypothesis-free approaches can enable the identification of epigenetic marks in previously unconsidered biological systems. Methods to assess DNAm across the genome include whole-genome bisulfite sequencing (WGBS) or array, bead-type hybridization. While WGBS has in-depth coverage of millions of CpGs, costs can be high. As an alternative, bead-type hybridization has become very popular, particularly the use of the Illumina microarrays. Of these, the most often-used platform has been the Illumina 450k array, which interrogates >450,000 CpGs across the genome. Although the array is human-specific, making cross-species comparisons unfeasible, it has become an array-of-choice to many researchers interested in studying the associations between DNAm and psychopathological traits or their risk factors. However, overall coverage of CpG sites is still low (only around 2% of all sites in the human genome) with some features such as enhancer regions being severely underrepresented. Also, as the platform was originally designed to study cancer, findings may be somewhat biased toward annotated genes related to oncological function, omitting probes that are currently unannotated or might be more relevant in other processes such as epigenetic inheritance or imprinting. In terms of technical artefacts, batch effects and the presence of cross-reactive probes can be a significant source of noise. The introduction of the newer Illumina EPIC BeadChip shows promise in addressing this bias by measuring DNAm at over 850,000 CpG sites, with increased genomic coverage, high reproducibility and reliability (Pidsley et al., 2016). However, increased coverage also implies statistical difficulties related to multiple testing and false positives. For a more in-depth discussion on these methods and arrays, see Non and Thayer (2015).

Much of what is known about DNAm has come from experimental animal models, which enable the manipulation of environmental exposures, the investigation of time- and tissue-specific effects on DNAm, and the characterization of downstream consequences on gene expression and behaviour. For example, ground-breaking rodent studies by Weaver et al. (2004) showed for the first time that differences in the *social* environment, such as maternal licking and grooming behaviour, could cause long-term epigenetic alterations of the glucocorticoid receptor (*NR3C1*) gene within the hippocampus of offspring, in turn driving inter-individual differences in stress response. Deriving similar evidence in human populations has been difficult, as experimental manipulation and access to relevant tissues is often unfeasible. Post-mortem brain tissues represent one important tool for testing the translational potential of animal findings to humans. For example, Weaver et al.'s *NR3C1* findings have been replicated in humans, based on hippocampal tissue from suicide

completers retrospectively identified as abused during childhood (e.g. McGowan et al., 2009). Among the drawbacks of post-mortem studies, however, is that factors such as pH, postmortem interval, and preservation methods might have confounding effects on DNA methylation and that these studies are usually retrospective, precluding the possibility of testing mediational hypotheses. In living humans, studies have to rely on DNAm from peripheral tissues (e.g. blood, saliva). Because DNAm patterns can differ across tissues, the extent to which we may infer epigenetic patterns in the living brain (likely the most relevant organ for the study of psychopathology) based on peripheral samples is still unclear, although there is evidence that a subset of peripheral DNAm markers may proxy methylation status of brain tissue (Walton et al., 2015).

Environmental risk exposure and DNAm: the *a* path

In this section, we review evidence for an association between environmental effects on DNAm, spanning gestation through to childhood. We focus on risks that have been previously associated with child internalising and externalising problems, including dietary, neurotoxic and stress exposures.

Nutrition. Pre- and postnatal nutrition is critical for healthy neurological development. DNAm is also highly responsive to diet as nutrients and bioactive compounds can alter the expression of genes at the transcriptional level and result in long-term phenotypic changes (Choi & Friso, 2010). For example, using Mendelian randomization (MR) analysis to strengthen causal inference, Binder and Michels (2013) found that maternal folate levels affect epigenome-wide DNAm patterns in infant cord blood and that genes in close proximity to several CpG loci were involved in metabolic processing. Different types of prenatal diet have been examined in relation to offspring DNAm.

On the one end of the spectrum is *under-nutrition*. In perhaps the first study to contribute empirical support for the hypothesis that early dietary conditions can cause persistent epigenetic changes in humans, Heijmans et al. (2008) examined DNAm (whole blood) in survivors of the Dutch hunger winter, a severe famine at the end of World War II. Even 60 years after the event, siblings who had been prenatally exposed to the famine showed lower DNAm of the *IGF2* gene (implicated in foetal development) compared with their unexposed, same-sex sibling. As an extension to this study, Tobi, Heijman and colleagues (2014) examined a subset of the same siblings using a hypothesis-free, epigenome-wide approach. The researchers identified additional loci associated with prenatal famine exposure, including *INSR* (involved in prenatal growth and insulin signalling) and

CPT1A (involved in fatty acid oxidation). In turn, these changes associated with relevant developmental outcomes, including birth weight and low-density lipoprotein cholesterol levels. Using data from the multi-generational Barbados Nutrition Study, Peter et al. (2016) examined whole blood from adults in their fifth decade of life ($n = 94$), including a group ($n = 44$) who had been hospitalised in early infancy due to severe protein-energy malnutrition. Epigenome-wide DNAm analyses identified 134 nutrition-sensitive, differentially methylated genomic regions, including loci annotated to a number of neuropsychiatric risk genes (e.g. *COMT*, *IFNG*, *SYNGAP1*, and *VIPR2*), which in turn associated with cognitive outcomes (i.e. IQ and attention). Of note, less than 3% of loci in the adults showed DNAm in their children. This result could reflect a low rate of intergenerational transmission of DNAm patterns associated with early nutrition. It could also be due to the Illumina 450k technology not being sufficient to detect such changes (e.g. not annotating key intragenic regions); or due to factors experienced (e.g. improved nutrition) by the second generation that could normalise DNAm levels (see Szutorisz & Hurd, 2016). Related to this last point, both the Dutch Hunger Winter and the Barbados studies were retrospective in terms of later measures of DNAm (≥ 50 years) being associated with earlier measures of the nutritional environment. Drawbacks to retrospective studies include that it is difficult to establish the temporal relationship between events and to control for potential confounding factors (e.g. cohort effects; see Talati, Keyes, & Hasin, 2016) that can affect DNAm levels.

On the other end of the spectrum is *over-nutrition*, which can be indexed by obesity. Greater adiposity during pregnancy (either as a result of being more overweight at the start, or by gaining weight during pregnancy) can deliver greater concentrations of glucose, fatty acids and inflammatory markers to the developing fetus (Lawlor et al., 2007; Poston, 2012). Obesity-focused studies during pregnancy have typically utilised prospective research designs. Godfrey et al. (2011) examined DNAm in healthy neonates, focusing on 68 CpGs across five candidate genes for obesity (measured in cord blood). Differential DNAm in the vicinity of two genes – *RXRA* and *eNOS* – associated with child fat mass at age 9. DNAm in *RXRA* also associated with maternal prenatal diet, a result that was replicated in an independent sample. Based on 1018 mother-child pairs in the Accessible Resource for Integrated Epigenomic Studies (ARIES; Relton et al., 2015) – a subsample of the Avon Longitudinal Study of Parents and Children (ALSPAC) – Sharp et al. (2015) examined associations between prenatal maternal weight and neonatal epigenome-wide DNAm patterns (measured in cord blood). Compared to ‘normal’ weight mothers, offspring of overweight mothers and underweight mothers showed many more differentially epigenome-wide

methylated sites (underweight $n=28$; overweight $n = 1621$). Of interest, offspring of overweight and underweight mothers showed no overlap in loci at the epigenome-wide level. Using a negative control design to strengthen causal inference, the authors additionally found that epigenetic changes associated with maternal obesity were stronger than for paternal obesity, supporting an intrauterine environmental effect of nutritional status on DNAm. Studies have also investigated the effect of postnatal dietary patterns on children's DNAm (Zheng, Xiao, Zhang, & Yu, 2014). One candidate gene study by Obermann-Borst et al. (2013), based on whole blood draws from offspring at 17 months of age, found that longer breastfeeding duration associated with lower levels of infant *LEP* methylation. As discussed by the authors, higher leptin levels can associate with deficient appetite regulation and risk for obesity.

Exposure to toxins. In general, prenatal exposure to bioactive compounds found in cigarettes, such as arsenic and lead, have been shown to affect DNAm patterns in neonates (Cardenas et al., 2015; Sen et al., 2015). In the field of epigenetics, the association between tobacco smoke exposure (whether directly or via prenatal exposure) and DNAm is perhaps the most extensively replicated finding. Similar to studies on obesity, toxin-focused epigenetic studies have primarily utilised prospective research designs. One of the earliest reports came from Breton et al. (2009), who found lower levels of global (LINE-1 and AluYb8) and CpG-specific methylation (Illumina GoldenGate Cancer methylation panel I measuring 1,505 CpGs) in buccal cells from children aged 5-7. Joubert et al. (2012) supported and extended these findings in the Norwegian Mother and Child Cohort Study (MoBA; $n = 1,062$ newborns, cord blood), by examining maternal plasma cotinine (a biomarker of smoking) in relation to a higher number of CpG sites (i.e. 473,844 via the Illumina 450k). The authors identified differential DNAm for 26 CpGs mapped to 10 genes, primarily involved in the detoxification of chemicals found in tobacco smoke. Of note, similar DNAm levels of the *AHRR*, *CYP1A1* and *GFII* genes were found when using plasma cotinine vs maternal self-reports, and results were also replicated in the Newborn Epigenetics Study (Hoyo et al., 2011). In 800 mother-child pairs from ALSPAC, Richmond et al. (2015), replicated many of these loci, as well as extending findings by examining how these methylation patterns change over time (cord blood at birth; whole blood at age 7 and 17). The authors reported that between birth and age 17, sites annotated to a number of genes showed reversible methylation (e.g. *GFLI*), whereas other showed persistently perturbed patterns (e.g. *CYP1A1*, *AHRR*). In summary, findings by Richmond et al. (2015), Breton et al. (2009) and Joubert et al. (2012) all converge in showing that prenatal exposure to teratogens can

have a prospective long-term impact on the methylome.

Early life stress and adversity. Stress and adversity are frequently assessed risk factors both during pregnancy and after. In animal models, prenatal and postnatal stress can cause long-term elevations in hypothalamic–pituitary axis (HPA) reactivity and anxiety-like behaviours, which can be explained by in part altered *NR3C1* (glucocorticoid receptor) gene expression (Weaver et al., 2004). These findings have been validated by associational studies in living humans based on peripheral tissues. For example, *NR3C1* DNAm in the offspring has been associated with prenatal factors including maternal depression and anxiety based on cord blood at birth (Hompeš et al., 2013; Oberlander et al., 2008), as well as postnatal factors, such as frequency of mothers' stroking of the child based on saliva in infancy (i.e. 14 months; Murgatroyd, Quinn, Sharp, Pickles, & Hill, 2015). As Murgatroyd et al. (2015) points out, candidate gene studies on *NR3C1* methylation have generated similar results in brain, saliva, and blood samples, supporting the idea that early life adversities can influence DNAm patterns across a variety of tissues.

Epigenome-wide changes in DNAm have also been investigated in relation to prenatal maternal stress. Of interest, these studies typically do not identify significant *NR3C1* changes after multiple correction. Cao-Lei et al. (2014) examined DNAm patterns (via whole blood draws) in children ($n = 34$) of women who, 13 years earlier, were pregnant during the 1998 Quebec ice storm. Here, prenatal maternal objective stress (e.g. injury, days without electricity/in shelter) was associated with offspring DNAm at 1,675 CpG sites (from the Illumina 450k), spanning 957 genes primarily related to immune function. Subjective stress (i.e. perceived impact of events) was uncorrelated with DNAm. Results were validated (10 of 12 loci investigated) within the same youth using pyrosequencing. In collaboration with Rijlaarsdam et al. (2016), we examined the prospective association between prenatal exposure to a general index of maternal stress and offspring epigenome-wide cord blood methylation in Generation R cohort ($n = 912$) and in ALSPAC ($n = 828$). A meta-analysis of results between the two studies indicated an over-representation of the methyltransferase activity pathway, although none of the individual DNAm loci identified survived genome-wide correction. Of note, the prenatal measure of maternal stress employed by Generation R and ALSPAC contain both objective (e.g. housing adequacy) and subjective (e.g. emotional support) measures of stress, which may (in part) explain the discrepancies in results between Cao-Lei et al. (2014) and Rijlaarsdam et al. (2016).

With regard to postnatal influences, early exposure to poverty and adversity in community samples has been associated with altered DNAm in large-scale hypothesis free

analytic approaches. For example, retrospective studies, based on whole blood draws in adolescence and adulthood, have implicated genes involved in basic cell processes, such as the protocadherin superfamily of genes encodes proteins involved in cell–cell adhesion and communication, as well as *EUROGI*, a gene that is involved in neuronal differentiation and cell-type specification in the developing nervous system (Borghol et al., 2012; Essex et al., 2013). In addition, ‘natural experiments’ have been used as a powerful method for examining epigenetic changes following exposure to severe environmental conditions (Mill & Heijmans, 2013). For example, research suggests that children raised in institutions (orphanages), who are deprived of normative rearing experiences, show changes in DNAm patterns that may be important for development and wellbeing (see Kumsta et al., 2016). In a Russian sample, Naumova et al. (2012) found that children raised in an institution since birth ($n = 14$) showed higher epigenome-wide DNAm patterns compared to high-poverty children living with their families ($n = 14$), particularly in genes related to immune regulation and cellular signalling. Blood samples were taken between 7-10 years of age and DNAm was assessed with the Illumina 27k. In another study, Esposito et al (2016) examined epigenome-wide DNAm (whole blood; Illumina 450k) in a sample of Russian and Eastern European adopted children ($n = 50$) compared with matched non-adopted children ($n = 33$). The authors identified 30 differentially methylated sites spanning 19 genes enriched for neural development and developmental biology. Of interest, one of the genes, *CYP1A1* (part of the cytochrome P450 family), had previously been associated with prenatal smoking (e.g. Richmond et al. 2015), perhaps reflecting early life exposure to cigarette smoke. A third epigenome-wide study by Kumsta et al. (2016) compared Romanian adoptees exposed to either extended (6–43 months; $n = 16$) or limited duration (<6 months; $n = 17$) early-life deprivation, in addition to a matched sample of UK adoptees ($n=16$) not exposed to deprivation. Although no probes were significant at the genome-wide level, an exposure-associated differentially methylated region was identified that spanned 9 sequential CpG sites in the promoter-regulatory region of the *CYP2E1* gene.

While there are some similarities in findings across these three ‘institutional’ studies (e.g. methylation markers involved in (neuro-)development and cell communication), differences in results might be attributable to a range of factors including the use of diverse cell types (buccal, whole blood or white blood cells), methylation assays (MeDIP, Illumina 27k or 450k), measures of postnatal adversity (socio-economic position as measured via father’s occupation and household amenities; parental adversity measured via depressive symptoms, anger, role overload and financial stress; institutional deprivation; adoption),

timing of exposure (at age 7; during infancy and preschool; before 6 months of age) or sample size (ranging between $n = 14$ per group to $n=109$).

Another severe form of adversity, childhood maltreatment (e.g. abuse, neglect), has also been found to associate with DNAm alterations in genes important for stress-response, immune function and neurodevelopment (see Lutz & Turecki, 2014, for a review). With few exceptions, most published research to date on maltreatment has focused on candidate genes (e.g. *NR3C1*, *SLC6A4*, *FKBP5*, *BDNF* and *5-HTT*) and been retrospective in nature (see Beach, Brody, Todorov, Gunter, & Philibert, 2010; Turecki & Meaney, 2016). Yang et al., (2013) examined DNAm differences (saliva; Illumina 450k) in 96 children who were removed from parental care due to reports of abuse or neglect vs 96 children with no history of abuse (age range 5-14 years). After controlling for multiple comparisons, maltreated and control children showed significantly different methylation levels across 2,868 CpG sites, which contained numerous markers of physical and psychiatric morbidity (e.g. *CCDC85*, *PTPRN*). More recently, based on a cross-sectional sample of high-risk youth (buccal cells; $n = 124$; age range = 16-24), we sought to characterise the DNAm ‘signatures’ of different forms of maltreatment, using an epigenome-wide approach (Cecil, Smith, et al., 2016). We found that physical maltreatment showed the strongest associations with DNAm, implicating multiple genes previously associated with psychiatric and physical disorders (e.g. *GABBR1*, *GRIN2D*, *CACNA2D4*, *PSEN2*). Based on gene ontology analyses, we also found that different types of maltreatment showed *unique* methylation patterns enriched for specific biological processes (e.g. physical abuse and cardiovascular function vs physical neglect and nutrient metabolism), but also shared a ‘common’ epigenetic signature enriched for biological processes related to regulation of nervous system development and organismal growth. Although the experience of victimization outside of the family is also important, it is a considerably under-researched area. One candidate gene study in 28 monozygotic twin pairs discordant for bullying victimization found that the bullied twins showed increased prospective *5-HTTPLR* DNAm between ages 5 and 10 (Ouellet-Morin et al., 2013). It is noteworthy that most of these studies overlap little with respect to the risk measure (abuse, neglect, bullying), sample design (population vs twin) and genetic coverage (candidate vs epigenome-wide), providing limited opportunity to derive integrative conclusions. Of interest, the genome-wide loci identified by Yang et al. and Cecil et al. did not show strong overlap, which may be due to differences in the study samples (i.e. children with an ‘official’ history of maltreatment vs. self-reports from high risk youth), and the different age ranges of the study participants (i.e. 5-14 vs. 16-24 years of age). In future, efforts to maximise

comparability between studies will hopefully allow a deeper insight into the epigenetic specifics of maltreatment characteristics.

Summary. In support of the *a* path of the mediation model (Figure 2 panel A; risk exposure → DNAm), DNAm has been associated with a range of pre- and postnatal adversities. Whereas prenatal studies have generally been prospective, postnatal studies, especially those examining early nutrition and childhood maltreatment, have been retrospective and/or cross-sectional. In reminder, to test DNAm as a mediator, the risk exposure should come before DNAm (e.g., prenatal risk → DNAm at birth). In addition, only a handful of studies have replicated findings in an independent sample (e.g. Godfrey et al., 2011), integrated genetic information or used methods to strengthen causal inference (e.g. *MR*: Binder & Michels, 2013; *negative control*: Sharp et al., 2015; *twin discordance*: Oullet-Morin et al., 2013).

DNAm and child psychopathology: the *b* path

In this section, we review studies that have examined DNAm in relation to psychopathology from childhood through to adolescence (i.e. the *b* path), including internalising and externalising difficulties.

Internalising difficulties. Research on childhood internalising problems has primarily focussed on DNAm of *NR3C1*. In a cross-sectional analysis of high-risk pre-schoolers, Parade et al. (2016) found that higher DNAm levels at exons 1D and 1F (extracted from saliva) associated with higher internalising – but not externalising – behaviour problems. Consistent with this, Dadds, Moul, Hawes, Mendoza Diaz, and Brennan (2015) reported that increased DNAm in the 1F region across whole blood and saliva associated with morning cortisol levels and higher risk of co-occurring internalising problems among clinically referred conduct disordered children. Furthermore, *NR3C1* methylation (extracted from whole blood) has been found to positively associate with risk of a lifetime diagnosis of internalizing disorders in an adolescent population sample (van der Knaap, Oldehinkel, Verhulst, van Oort, & Riese, 2015). Taking an epigenome-wide approach, Weder et al. (2014) reported that saliva-derived DNAm in three genes involved in stress response and neural plasticity (*ID3*, *TPPP*, and *GRINI*) associated with child depression levels. Nominal associations were also identified in a number of traditional candidate genes, including *NR3C1*, *FKBP5* and *BDNF* – particularly among children exposed to maltreatment. Finally, capitalizing on a genetically-sensitive design, Dempster et al. (2014) examined epigenome-wide patterns in 18 monozygotic twin pairs discordant for adolescent depression, based on

buccal cell DNA. The most differentially methylated site, annotated to *STK32C* (a serine/threonine kinase gene), was also found to be independently associated with major depression in post-mortem cerebellum samples, suggesting that, despite tissue specificity in DNAm, some disease epimutations may be detectable across multiple tissues – consistent with prior findings (e.g. Davies et al., 2012). Likewise, the consistency of some findings (such as *NR3C1*) across studies is remarkable considering the heterogeneity of study characteristics (e.g. blood vs saliva; population vs twin samples; preschool age vs adolescence; children with or without a history of maltreatment or co-occurring externalizing symptoms). The moderate degree of stability in results related to *NR3C1* might originate from a slight increase in sample size ranging from $n = 171$ to $n = 361$ (compared to – for instance – a range of $n = 14$ – $n = 109$ in studies on postnatal adversity, as discussed above).

Overall, the candidate gene studies described above support a link between higher *NR3C1* methylation and internalising difficulties, with associations observed across different peripheral tissues and sample characteristics. However, *NR3C1* methylation does not emerge as a significant predictor of internalising difficulties based on existing epigenome-wide studies, suggesting that – if associations are indeed true – they are likely to be of small effect size. As commonly observed in epigenome-wide association studies (EWAS), the two studies that employed an EWAS approach did not identify overlapping DNAm sites. While both studies extracted DNA from similar sources (saliva/buccal cells), used the same array (Illumina 450k) and examined depression specifically, the studies differed in sample characteristics (children differing in maltreatment exposure vs monozygotic twins discordant for depression) and analytical strategies (linear mixed effects model vs in-house method combining significance level and effect size), which may have contributed to the divergent findings. Of note, because EWAS studies do not typically publish their full results as supplementary material, it is not possible to establish whether DNAm sites identified in one study may show sub-threshold associations in other studies.

Externalising difficulties. A larger number of studies have investigated DNAm patterns associated with externalising difficulties, including physical aggression and conduct problems (CP), callous-unemotional (CU) traits, attention deficit hyperactivity disorder (ADHD), oppositional defiant disorder (ODD) and substance use.

Physical aggression and conduct problems. In contrast to *higher* DNAm reported for internalising problems, Heinrich et al. (2015) found that youth with a lifetime diagnosis of an externalising disorder showed *lower* levels of *NR3C1* methylation at exon 1F (whole blood), compared to both youth with a depressive disorder and healthy controls. Based on data

extracted from white blood cells, one research group also found that, compared to controls, adult males with a history of chronic childhood aggression differed in (i) *SLC6A4* promoter methylation and in vivo levels of brain serotonin synthesis (Wang et al., 2012); (ii) DNAm levels in a set of genes involved in cytokine function and inflammation (Provençal, Suderman, Vitaro, Szyf, & Tremblay, 2013), and (iii) DNAm patterns across a large number of gene promoter regions in an epigenome-wide scan (Provençal et al., 2014) – a finding that was later extended to a small sample of adult females as well (Guillemín et al., 2014). In the ALSPAC sample, we found that methylomic variation at seven loci across the genome at birth (cord blood) differentiated children who go on to develop early-onset ($n = 174$) vs low ($n = 86$) conduct problems, including sites in the vicinity of *MGLL* – a gene involved in endocannabinoid signalling and pain perception. (Cecil, Jaffee, et al., 2016). In addition, sub-threshold associations with DNAm sites across three candidate genes (*MAOA*, *BDNF*, *FKBP5*) were identified, supporting the idea that DNAm patterns in candidate genes may reliably associate with psychopathological outcomes, but only exert small-to-moderate effects that do not survive genome-wide correction. Finally, we examined whether the identified loci associated with genetic and environmental risks. As our sample was underpowered to examine genetic polymorphisms directly, we consulted an online catalogue of known methylation quantitative trait loci (mQTL) based on the Illumina 450k resource in ARIES (Gaunt et al., 2016). None of the identified loci were linked to known mQTLs, although nominal associations with prenatal exposures were observed (e.g. maternal smoking and alcohol use).

Callous-unemotional traits. In a clinical sample of boys diagnosed with conduct and oppositional defiant disorders, those with more severe CU traits (e.g. low capacity for empathy, lack of guilt, shallow affect) have been reported to show lower levels of *HTR1B* methylation in saliva (Moul, Dobson-Stone, Brennan, Hawes, & Dadds, 2015), as well as higher *OXTR* methylation in blood, which in turn correlated with lower circulating oxytocin levels (Dadds et al., 2014), compared to boys with less severe CU traits. Using a genome-wide approach in ARIES, we examined DNAm patterns associated with a correlate of CU traits – low prosocial behaviour (e.g. inconsiderate of other's feelings, not helpful if someone is hurt; Meehan et al., 2017). We found that, at birth (cord blood), two loci located in the vicinity of *NDUFS8* and *SGCE/PEG10* differentiated chronic-low prosocial youth from typical comparisons. In turn, higher DNAm in *SGCE/PEG10* associated with lower empathy, higher social-cognitive difficulties and greater victimization during childhood.

Attention deficit hyperactivity. On the basis of findings from animal and molecular genetic work, there has been growing interest in the involvement of dopamine signalling genes in ADHD development. In a sample of Chinese Han children, Xu et al. (2015) found that promoter DNAm and expression levels of the dopamine genes *DAT1*, *DRD4* and *DRD5* related to ADHD risk. Consistent with this, Dadds, Schollar-Root, Lenroot, Moul, and Hawes (2016) reported that higher *DRD4* methylation related to more severe ADHD symptomatology in high-risk children across peripheral tissues (saliva and blood). Interestingly, these associations were found to be specific to the *attentional* features of ADHD, and were independent of genetic variation, environmental adversity and comorbid conduct problems. However, based on another sample of Chinese Han children with a diagnosis of ADHD, Ding et al. (2016) found no evidence for a significant association between *DAT1* or *DRD4* promoter methylation and baseline ADHD symptoms, although *DAT1* (but not *DRD4*) methylation did correlate with response to methylphenidate treatment at follow-up. Focusing on a different candidate gene in 6-15 year olds with ADHD, Park et al. (2015) found that increased *SLC6A4* promoter methylation specifically associated with more severe *hyperactive-impulsive* symptoms, as well as lower cortical thickness in occipital-temporal regions. Using the prospective Generation R cohort, van Mil et al. (2014) examined associations between DNAm in 7 candidate genes (measured in cord blood at birth) with ADHD symptoms (at age 6). DNAm levels were inversely related to ADHD symptoms, with the *DRD4* and *SLC6A4* regions largely explaining this association. Supporting evidence for a role of altered neurotransmitter function has also emerged from an epigenome-wide study of ADHD in 7-12 year old children. After filtering genes based on statistical and biological criteria, Wilmot et al (2016) found that DNAm of the neuropeptide receptor gene *VIPR2* in saliva was lower in the ADHD group vs controls (all boys), and this association was confirmed in an independent sample of age-matched children. At a broader level, genes that were nominally associated with ADHD were found to be enriched for biological processes related to inflammation and monoamine neurotransmission. Finally, using ARIES, we performed an epigenome-wide study of DNAm (cord blood at birth; whole blood at age 7) and trajectories of ADHD (ages 7-15) (Walton et al., 2016). DNAm at birth differentially associated with a high vs low ADHD trajectory across 13 probes, including those annotated to *ZNF544* (previously implicated in ADHD) and *PEX2* (related to fatty acid metabolism of n-3 PUFAs and axon myelination). None of the 13 probes maintained an association with ADHD at age 7, which suggests that DNAm associations can be time- specific, perhaps reflecting critical – or sensitive – periods of development. Together, tentative evidence from

epigenetic studies on ADHD points to the potential relevance of DNAm alterations in genes implicated in neurotransmission, particularly dopaminergic signalling.

Oppositional defiant disorder and ADHD. Twin studies suggest that externalising disorders, such as ODD and ADHD, share considerable common genetic variance (Barker, Cecil, Walton, & Meehan, 2017); however, the potential role of DNAm in this common biologic variability is unclear. Based on ARIES, we (Barker, Cecil, Walton, Jaffee, et al., 2017) investigated (i) prospective associations between neonatal DNAm (cord blood) and trajectories (ages 7-13) of ODD symptoms, as well as the ODD sub-dimensions of Irritable and Headstrong; and (ii) biological overlap with the ADHD-associated loci identified by Walton et al. (2016). We identified a larger number of loci in relation to ODD as a whole ($n = 30$) compared to either the sub-dimension of Headstrong ($n = 11$) or Irritable (no loci) alone. Overlap analysis also indicated shared biological influences between ODD and ADHD, including glutamate signalling and the protocadherin superfamily of genes involved in cell–cell adhesion and communication, a result similar to a study that examined the impact of early poverty on DNAm (Borghol et al., 2012). We also examined potential genetic influences on DNAm through the use of the mQTLdb database. Three of the 30 loci for ODD and two of the 11 loci for Headstrong associated with known mQTLs (i.e. *DNAm* loci that are likely to be under genetic control).

Substance use. Finally, very little work has sought to characterize the relationship between DNAm and substance abuse during adolescence (Cecil, Walton & Viding, 2015). Researching *COMT* methylation in whole blood – an important gene for neurotransmitter catalysis – Van der Knaap et al. (2014) found no main effect on adolescent cannabis use. However, a significant methylation by genotype interaction was identified, where Met/Met carriers with higher DNAm were least likely to be frequent cannabis users. (Philibert, Gunter, Beach, Brody, & Madan, 2008) reported that DNAm in *MAOA* (derived from lymphoblast lines) significantly associated with alcohol and nicotine dependence for females, but not males – also showing a trend-level effect of genotype on DNAm in the female group. A study by Ruggeri et al. (2015) examined epigenome-wide DNAm differences (whole blood) in monozygotic twin pairs discordant for alcohol use disorders in young adulthood. The most differentially methylated region was located in *PPM1G*, a gene previously linked to alcohol dependence. This association was verified technically using mass spectrometry and replicated in an independent sample of 499 adolescents from the IMAGEN cohort, whereby higher DNAm associated with decreased gene expression (independently of genotype), early escalation of alcohol use and greater impulsiveness, as well as increased blood-oxygen-level-

dependent response in the right subthalamic nucleus during an impulsiveness task. While it is not possible to compare these three studies given their different objectives, level of coverage (candidate gene vs EWAS) and substance type examined (cannabis, nicotine, alcohol use), it is noteworthy that all sought to account for genetic influences, with Ruggeri et al (2015) also featuring an independent replication and the integration of DNAm with additional biological data (i.e. gene expression and brain imaging).

Summary. Overall, the majority of studies on the *b* path of the mediation model (Figure 2 panel A; DNAm → psychopathology) were retrospective and focussed on candidate genes (particularly those implicated in stress response and neurotransmitter function), although an increasing number of longitudinal, epigenome-wide investigations are emerging. As with the *a* path, only a few studies, such as Ruggeri et al., 2015, have replicated findings in an independent cohort (e.g.), accounted for genetic influences and/or used advanced methods for causal inference (e.g. *twin discordance*: Dempster et al., 2014). However, a handful of studies have taken steps to biologically characterise the identified DNAm changes by testing whether effects replicated across peripheral tissues, correlated with gene product levels, or were associated with measures of brain activity (e.g. *multiple tissues*: Dadds et al., 2016; *gene product levels*: Dadds et al., 2014; *brain activity*: Ruggeri et al., 2015). Similarly to the *a* path section reviewed above, genes investigated by candidate studies did not converge with those identified by studies that used a hypothesis-free, epigenome-wide analyses. Furthermore, little or no overlap was evident in top DNAm sites identified across epigenome-wide analyses of psychopathology. Together, these discrepancies are likely due to wide methodological differences across studies, including sample characteristics (e.g. age, risk level), phenotype operationalisation (e.g. symptom-based vs clinical threshold), and analytical strategy (e.g. choice of covariates, pruning of DNAm sites and statistical models performed). The discordance between candidate gene studies and EWAS studies may also reflect differences in effect sizes of the identified sites.

Integrative models:

Combining risk exposure, DNAm and child psychopathology

In this section, we highlight four studies that have explicitly tested whether DNAm mediates associations between risk exposures and psychopathology. Specifically, these are studies that have integrated environmental, epigenetic and phenotypic measures, using a prospective design that captures the temporal ordering of the variables, with DNAm as the intermediary variable (i.e. risk → DNAm → psychopathology).

First, Monk et al. (2016) examined whether mothers ($n = 61$) with higher levels of stress or depressed mood in the 2nd trimester of pregnancy showed lower levels of foetal coupling – an index of foetal nervous system development – in the third trimester, due to DNAm in three placental glucocorticoid pathway genes – *HSD11B2*, *NR3C1* and *FKBP5*. Maternal DNAm was measured by saliva and results were validated in placental tissue. The authors found that mothers who reported higher rates of depression had higher *HSD11B2* methylation, which in turn, associated with lower foetal coupling.

Second, in a sample of CP youth from ALSPAC, we examined whether different developmental pathways to callous-unemotional traits exist, based on the presence ($n = 45$) or absence ($n = 39$) of co-occurring internalising symptoms (Cecil et al., 2014). Specifically, we used path analysis to trace the prospective inter-relationships between environmental risk (prenatal-to-age 9) and *OXTR* DNAm (birth, age 7 and 9) in the prediction of CU traits (age 13), for CP youth who either showed high or low levels of internalising problems. For those with *low* levels of internalising problems, higher prenatal risks (maternal psychopathology, criminal behaviours, substance use) associated with higher *OXTR* methylation at birth, which, in turn, associated with higher CU traits in early adolescence. Intriguingly, we also identified in this group an “evocative epigenetic-environment correlation”, whereby higher *OXTR* methylation at birth (but not at any other time point) was prospectively associated with lower subsequent experience of victimization during childhood (birth-age 7). In contrast, no associations between *OXTR* methylation and CU traits was identified in youth with high internalising problems.

Third, applying a similar analytical model in ALSPAC, we examined the degree to which a high-fat, high-sugar diet (*maternal*: prenatal; *child*: age 3) might associate with higher ADHD symptoms for children with early-onset ($n = 83$) vs low CP ($n = 81$) (Rijlaarsdam et al., 2017). Results indicated that, across both groups, prenatal diet associated with higher *IGF2* methylation at birth. However, only in the early-onset CP group, higher *IGF2* methylation also associated higher ADHD symptoms. Similar to the Cecil et al., (2014) study, epigenetic associations were localised at birth (i.e. no associations with DNAm at age 7).

Finally, we again used the ALSPAC sample ($n = 244$) to investigate epigenome-wide, prospective associations between DNAm (birth, age 7) and substance use in adolescence (tobacco, alcohol and cannabis use; $n = 244$; age 14–18; Cecil, Walton, et al., 2016). We found that at birth (but not at age 7), epigenetic variation across a tightly interconnected genetic network ($n = 65$ epigenome-wide corrected loci) associated with greater levels of

substance use during adolescence, as well as an earlier age of onset amongst users. Key annotated genes included *PACSINI*, *NEUROD4* and *NTRK2*, implicated in neurodevelopmental processes. Several of the identified loci were associated with known methylation quantitative trait loci, and consequently likely to be under significant genetic control. In addition, we found evidence for a prenatal environmental effect, whereby these 65 loci collectively mediated the influence of prenatal maternal tobacco smoking on adolescent substance use (see Figure 2, Panel B). Although the above studies are promising, in that they show that DNAm can act as a mediator between risk exposure and child psychopathology, it is important to note that they are nonetheless associational in nature (i.e. non-causal).

Discussion

In this review, we summarised evidence linking environmental risk exposure, DNAm and childhood psychopathology. We focused on studies that have examined living children and adolescents (i.e. the measure of risk and/or psychopathology is in childhood or adolescence) and that have used DNAm extracted from peripheral tissues. Based on the studies presented, there is preliminary evidence to support that (i) pre- and postnatal environmental factors, such as diet, neurotoxic exposures and stress, may influence offspring DNAm (i.e. *a* path), and that (ii) variability in DNAm may, in turn, associate with internalising and externalising difficulties during childhood and adolescence (i.e. *b* path). Despite these promising findings, very few studies have examined DNAm in relation to *both* exposures and outcomes, only a handful have accounted for genetic confounding and most all analyses have been correlational in nature – to the best of our knowledge, models examining causal mediational pathways in living children and adolescents have yet to be published. Consequently, the extent to which DNAm may truly mediate environmental influences on the development and course of psychopathology is not clear. Below, we discuss five key areas that can help move the field forward.

Key challenges, recommendations and future directions

1. Characterizing environmental effects with greater precision (i.e. the *a* path)

First, we need to reach a more complete and fine-tuned understanding of environmental effects on DNAm (see top third of Figure 3, the *a* path section). As seen earlier in our review, the number of reported associations between environmental influences and DNAm is exponentially growing, implicating a wide range of exposures. These exposures, however, are often correlated – for example, stress levels in pregnancy have been associated with quality of diet (Barker, Kirkham, Ng, & Jensen, 2013), and childhood

maltreatment has been found to cluster in geographical areas characterized by increased poverty and violence exposure (Cecil, Viding, Barker, Guiney, & McCrory, 2014). Yet, epigenetic studies to date have typically examined single exposures in isolation, potentially resulting in the overestimation of effects observed – in some cases, the risk under investigation may even be a proxy for unmeasured exposures (Monk, Georgieff, & Osterholm, 2013). In future, modelling multiple exposures and/or outcomes simultaneously (e.g. via path analysis) may help to isolate specific risk pathways with greater accuracy, particularly in cases where an exposure has been found to associate with multiple negative outcomes (i.e. 'multifinality'), or conversely, where an outcome has been linked to different exposures (i.e. 'equifinality'; Cicchetti & Rogosch, 1996).

Furthermore, studies will need to begin testing the potential moderating role of exposure characteristics, such as type, severity, duration and timing (Barker, 2013). Specifically, studies that record this information should aim to address a number of outstanding questions, including (i) the extent to which epigenetic patterns may reliably differentiate between types of exposures (e.g. childhood abuse vs neglect); (ii) whether the effects of acute exposures differ from chronic ones; and (iii) whether environmental effects may be developmentally-dependent (e.g. pre- vs post-natal), for example, based on the degree of maturation/plasticity of relevant biological systems.

Another factor that may contribute to the overestimation of effects is genetic influence. Research suggests that (i) genetic effects on DNAm can be highly stable across the life course, and (ii) the genetic component of DNAm may have an important role in the development of complex traits (Gaunt et al., 2016; Hannon et al., 2016). However, very few studies reviewed here have modelled genetic influences when examining the association between exposures and DNAm. In order to address this, future studies should seek to use genetically-informative designs that make it possible to isolate environmental influences (e.g. monozygotic twin difference design; Oullet-Morin et al., 2013) or incorporate genetic sequencing data in order to control for genetic variants known to influence DNAm levels. These mQTLs (Gaunt et al., 2016) have been found to associate with gene expression and may serve as markers for genetic influences on gene regulation (see Hannon et al., 2016). There are now online mQTL databases based on published large-scale analyses that can be used when genetic information is unavailable, or when studies are underpowered to perform genetic analyses (e.g. ARIES mQTL database; www.mqtladb.org; as done for example in Cecil et al., 2016). It should be noted that the heritability of DNAm patterns via twin studies (e.g. between 20-97% across different genes; Heijmans, Kremer, Tobi, Boomsma, &

Slagboom, 2007) is greater than what can currently be explained using known mQTLs (Gaunt et al., 2016). Therefore, the small amount of genetic influence identified in our research (e.g. Barker et al., 2017; Cecil et al., 2016, 2017) may underestimate genetic variance that could be due to polygenic effects, involving many mQTLs, each of explains a small amount of variance (as detectable in the ARIES resource given the phenotypes we examined). Another current limitation is that existing array platforms (e.g. 450k) may not provide representative data for methylomic variation and its underlying genetic architecture (Taudt, Colomé-Tatché, & Johannes, 2016). However, the introduction of newer array-based platforms, such as the EPIC array could help to more comprehensively assess genetic influences on DNAm (Taudt et al., 2016).

2. Improving understanding of the methylome (i.e. the mediator)

Another barrier to establishing causal mediation relates to our limited knowledge of DNAm itself (see Figure 3, the *mediator* section). We highlight here three aspects of DNAm that warrant further investigation. First, *variability*. Unlike the genome – which remains mostly stable across the lifespan – DNAm is dynamic over time and varies across multiple factors, including sex, age, tissue, and cell type (Liang & Cookson, 2014). Consequently, it has been difficult to establish what a ‘normative’ profile of DNAm is (as this may depend on when and where the sample is collected), and how far such profile must deviate in order to confer risk for psychopathology. To address this, it will be important to collect data across multiple tissues and time points. Specifically, the availability of cross-tissue data will make it possible to quantify peripheral-CNS variability (Walton et al., 2015) and locate peripheral biomarkers that most closely resemble DNAm patterns in neural networks underlying psychopathology. Establishing whether exposures of interest exert tissue-specific vs global effects will also be important, as this will bear on the selection of appropriate samples for testing mediation hypotheses. A notable example of this is *NR3C1* methylation, which so far has been linked to early adversity across multiple peripheral (e.g. blood, saliva, placenta) and neural (e.g. hippocampus) tissues (see Turecki & Meaney, 2016). At the same time, collection of repeated measures of DNAm – so far a rare occurrence (with notable exceptions, e.g. ARIES, Generation R) – will be crucial for establishing patterns of stability vs change in DNAm across development. Particularly, studies will need to investigate whether certain genomic regions are more dynamic than others (perhaps reflecting greater responsivity to the environment, or even gene-environment interactions on DNAm levels),

and whether longitudinal DNAm trajectories may be informative in predicting the onset and course of psychopathology.

The second aspect is *scale*. At present, we are only able to access a small part of a much wider system that needs to be fully mapped out. Indeed, commonly used platforms such as the Illumina 450k only capture around 2% of methylomic variation across the genome. In turn, the methylome as a whole is only one of multiple epigenetic mechanisms that work in concert, including non-coding RNAs, histone modifications and other types of chromatin remodelling. As a result, many epigenetic patterns of potential relevance to psychopathology remain largely inaccessible (Non & Thayer, 2015). In the near future, rapid technological advances such as whole-genome bisulfite sequencing, will make it increasingly possible to obtain a more complete picture of the methylome. The compilation of large-scale reference epigenome datasets (e.g. BLUEPRINT; <http://www.blueprint-epigenome.eu/>) will also be important for establishing how DNAm fits within the broader machinery of epigenetic regulation (Shakya, O'Connell, & Ruskin, 2012).

The third aspect is *transmission*. Although it is generally assumed that in the zygote patterns in DNAm are erased after fertilization – allowing for a complete epigenetic resetting in the offspring – recent studies in animals have shown that in some instances DNAm patterns are copied, providing a potential mechanism for transgenerational inheritance. For example, rodent studies have shown that environmental exposure to endocrine disruptors can lead to reduced spermatogenic capacities in the following four generations via altered DNAm patterns in the germ line (Anway, Cupp, Uzumcu, & Skinner, 2005). In humans, the possibility that exposure-related DNAm patterns may be passed on across generations has received considerable theoretical interest, and recent research suggests this possibility (Yehuda et al., 2016). However, at present, little empirical data exists. An intriguing question is whether an exposure can potentially affect DNAm patterns in three generations directly – without the DNAm patterns being inherited; e.g., a prenatal DNAm stress effect on the mother, foetus and, in turn, DNAm in the foetus' germ cells (see Figure 3 on the left-hand side; Bowers & Yehuda, 2016). Testing hypotheses about transmission and epigenetic inheritance will require the collection of data spanning several generations, presenting perhaps one of our greatest challenges.

3. Establishing the functional significance of identified loci (i.e. the *b* path)

The next area for future research concerns the pathway between DNAm and psychopathology (see Figure 3: the *b path* section). As we have seen in this review, DNAm

patterns have been associated with a wide range of psychopathological outcomes in children and adolescents. However, the presence of an association alone does not necessarily imply a functional effect. Indeed, it is unclear to what extent *statistical* significance overlaps with *functional* significance. For example, reported effects are sometimes highly significant but involve only a small percent change in methylation, with unknown biological consequence. Furthermore, genes selected for candidate gene studies have not typically converged with those identified using hypothesis-free, epigenome-wide approaches. This is well illustrated in the case of *NR3C1*, which, despite being the most commonly investigated candidate gene in the field (with associations having been reported across tissues and species), is not usually identified in EWAS studies that employ epigenome-wide thresholds of significance.

In order to move beyond statistical associations, researchers should, whenever possible, assess the functional significance of psychopathology-related DNAm changes at multiple biological levels (e.g. transcriptome, proteome, metabolome). Relevant strategies include (i) profiling gene expression and additional *omics* data from the same samples as the DNAm (e.g. Ruggeri et al., 2015); (ii) using animal models (e.g. which permit chemical alteration of DNAm patterns); and (iii) carrying out in vitro experiments (e.g. cell cultures). In cases where this is not feasible, researchers may use online resources to check whether the identified DNAm loci overlap with established key regulatory elements, such as transcription factor binding sites (e.g. using ENCODE data; <http://genome.ucsc.edu/ENCODE/>; e.g. as done in Cecil et al., 2017; Provencal et al., 2013). In-vivo neuroimaging data could also be used in future to examine whether peripheral DNAm markers associate with structural and functional correlates of psychopathology (e.g. amygdala volume; Walton et al., 2017). Finally, in light of studies pointing to the temporal specificity of DNAm effects on psychopathology (e.g. our work pointing to localized effects at birth; Cecil et al., 2014; Rijdsdaam et al., 2016), more research is needed to determine whether DNAm can trigger developmental cascades, without the mark itself being sustained over time. More broadly, research will need to trace exactly how the downstream effects of DNAm engender biological vulnerability to psychopathology.

4. Maximising comparability across studies and opportunities for replication

DNAm studies to date have been highly heterogeneous, differing widely in the choice of design, sample characteristics, genomic coverage, quality control procedures, variables of interest and analytical routines. Consequently, comparability across findings has been low, and opportunities for replication scarce. Going forward, the establishment of best practice

guidelines will considerably help in this respect. Indeed, areas of consensus are already emerging (e.g. need for appropriate covariates), and standardized pipelines for data normalization and quality control are also increasingly available (Morris & Beck, 2015). Furthermore, the integration of discovery and replicate samples will become more important (e.g. as done in Godfrey et al., 2011; Ruggeri et al., 2015), as was the case for genomic studies.

One main challenge to maximising comparability is that DNAm data are multifactorial, high dimensional, and inter-correlated, raising questions as to how best they should be analysed in the first place (Almouzni et al., 2014). In Figure 4 (left-hand side), we provide a flowchart of analytical approaches that may be undertaken, depending on study characteristics such as genomic coverage (i.e. epigenome-wide vs candidate-focussed), preferred level of analysis (i.e. single-site analysis vs application of data reduction strategies), availability of repeated DNAm measures, and access to *both* exposure and phenotype data. These include commonly-used approaches in the field, as well as methods that are likely to become increasingly applied in future, such as polyepigenetic scores and structural equation modelling.

5. Strengthening causal inference

Finally, we need to move towards stronger causal inference. It is important to note that a non-causal DNAm biomarker can satisfy steps for statistical significance in mediation (e.g. the $a*b$ path). However, in order for DNAm to truly be a causal mediator, it must be shown to be causally affected by an exposure (a path), and to causally affect a psychopathological outcome (b path; Figure 2A). These pathways will be particularly difficult to isolate as (i) environmental influences are confounded by genetic factors; (ii) epigenetic effects are complex and inter-correlated; and (iii) psychopathological outcomes are heterogeneous and multi-determined. One key issue here is *reverse causation*, whereby the observed DNAm changes may be a consequence – as opposed to a risk factor for – psychopathology. This can be partially addressed by using a prospective design that accurately captures temporal order, especially in the case where the availability of repeated measures data makes it possible to trace longitudinal interrelationships between risk exposures, DNAm and psychopathology.

Even so, correlational studies in general are vulnerable to confounding, and it will become increasingly important to establish the robustness of findings using advanced inference methods. Ideally, studies should aim to draw on multiple approaches to test causal pathways (i.e. triangulation; Lawlor, Tilling, & Smith, 2017), as each has its own strengths

and limitations (see right-hand side of Figure 4 for an overview). For example, studies may use a mixture of natural experiments (e.g. negative control, Sharp et al., 2015; mendelian randomization; Relton & Davey Smith, 2012) and methodological approaches (e.g. cross-setting and cross-species comparisons). For a comprehensive review of methodological approaches for drawing causal inferences from epidemiological birth cohorts, we refer the reader to Richmond et al (2014). It is important to note that peripheral DNAm marks may still be a useful biomarker, even in the absence of causal effects (see Ladd-Acosta & Fallin, 2015). As stated, in cancer research, epigenetic biomarkers have even been used to predict treatment response. However, without a causal relationship, any type of ‘epigenetic therapy’ (e.g. Syzf, 2015) will be unlikely to have an effect on the psychopathological outcome.

Implications and translational potential

DNA methylation is a promising molecular mechanism by which the environment can increase risk for psychopathology. Yet, as we gain an appreciation of the challenges characterizing epigenetic research, we must be mindful to manage expectations. In future, the use of strategies such as careful selection of research design (e.g. appropriate tissue, sample size, time points), collection of prospective and integrative data, functional characterization of epigenetic findings, transparency in reporting and replication, and the application of causal inference methods, will mark an important step for overcoming current hurdles and moving the field forward. Bearing this in mind, there are a number of ways in which epigenetic research may contribute to the understanding, prevention, and treatment of psychopathology.

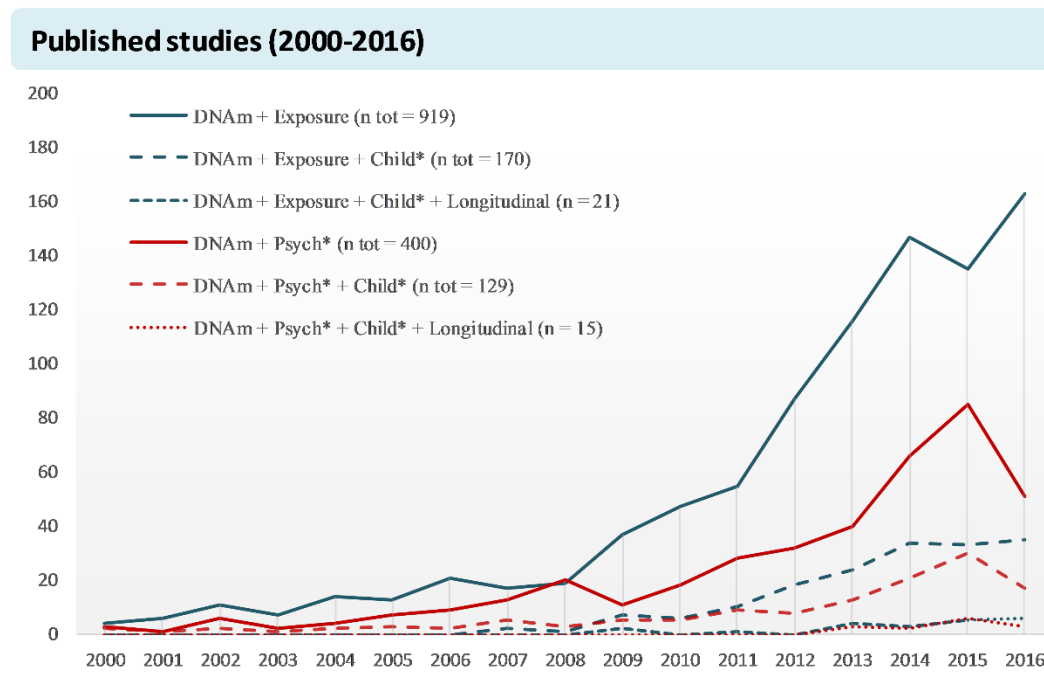
In the short-term, findings may be used to refine existing models of how environmental influences become biologically embedded, shifting developmental trajectories and engendering latent vulnerability for psychopathology. Longitudinal models may also be used to investigate the timing of environmental effects and pinpoint specific windows of biological vulnerability (e.g. prenatal period, infancy) that could benefit most from preventive action. In the medium-term, as the number of replications grow and robust associations are identified, epigenetic variation in specific genes may be used across clinical and research settings as biomarkers for environmental exposures, psychopathology risk, and response to treatment. Furthermore, the comparison of DNAm pre- vs post-intervention (e.g. via environmental enrichment, psychological therapy, medication), could lend insights into the potential reversibility of psychopathology-related patterns and how best to promote resilience. Indeed, tentative evidence that psychopathology-related DNAm patterns respond to psychological and pharmaceutical intervention is already beginning to emerge (Ding et al., 2016; Roberts et

al., 2014). In the long-term, establishing causal pathways between DNAm and psychopathology knowledge may provide greater knowledge about the aetiology of child and adolescent psychopathology, and perhaps even lead to the development of novel strategies for treating mental health problems.

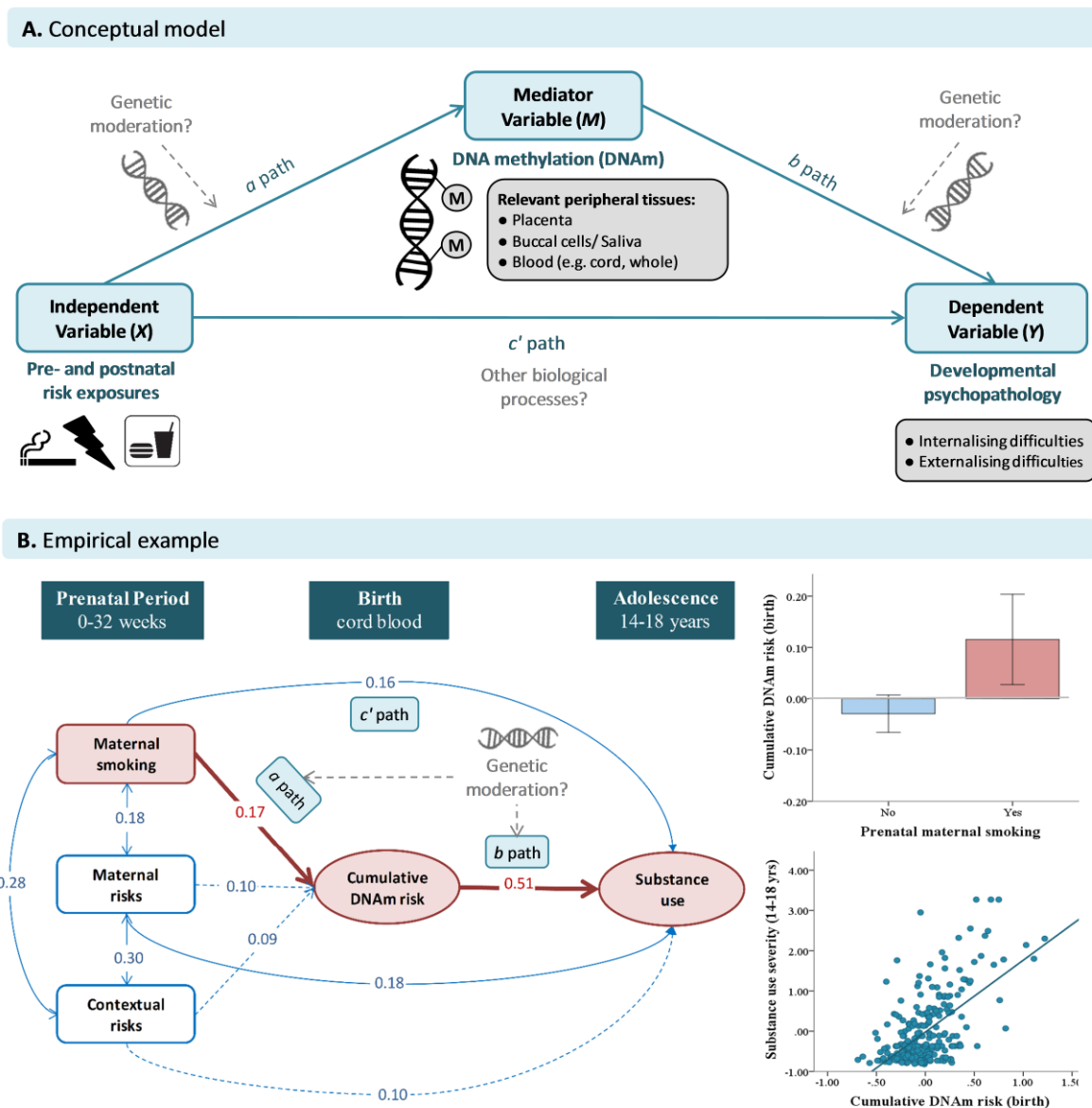
Acknowledgements

We are extremely grateful to the Avon Longitudinal Study of Parents and children. Specifically, we thank Caroline L. Relton, Tom R Gaunt, Sue M Ring, Wendy McArdle and George Davey Smith, and all the laboratory scientists and bioinformaticians who contributed considerable time and expertise to the ARIES DNA methylation resource. More generally we thank all the families who took part in ALSPAC, 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. Of note, ALSPAC is funded by the UK Medical Research Council and the Wellcome Trust (Grant ref: 102215/2/13/2) and the University of Bristol provide core support for ALSPAC. ARIES was funded by the BBSRC (BBI025751/1 and BB/I025263/1). This publication is the work of the authors who will serve as guarantors for the contents of this paper. This research was specifically supported by the National Institute of Child and Human Development grant to Edward D. Barker (R01HD068437). Charlotte AM Cecil was supported by the Economic and Social Research Council (ESRC, grant no. ES/N001273/1). We also thank Jonathan Mill, Sara R. Jaffee, Thomas G. O'Connor, Barbara Maughan and Isabelle Ouellet-Morin who were key collaborators on our DNA methylation project. We thank Sara R. Jaffee and Bonamy Oliver, for comments on a previous version of this manuscript.

Figure 1. Published human studies (2000-2016) relating DNAm to environmental exposures (blue) or psychopathology-related phenotypes (red)

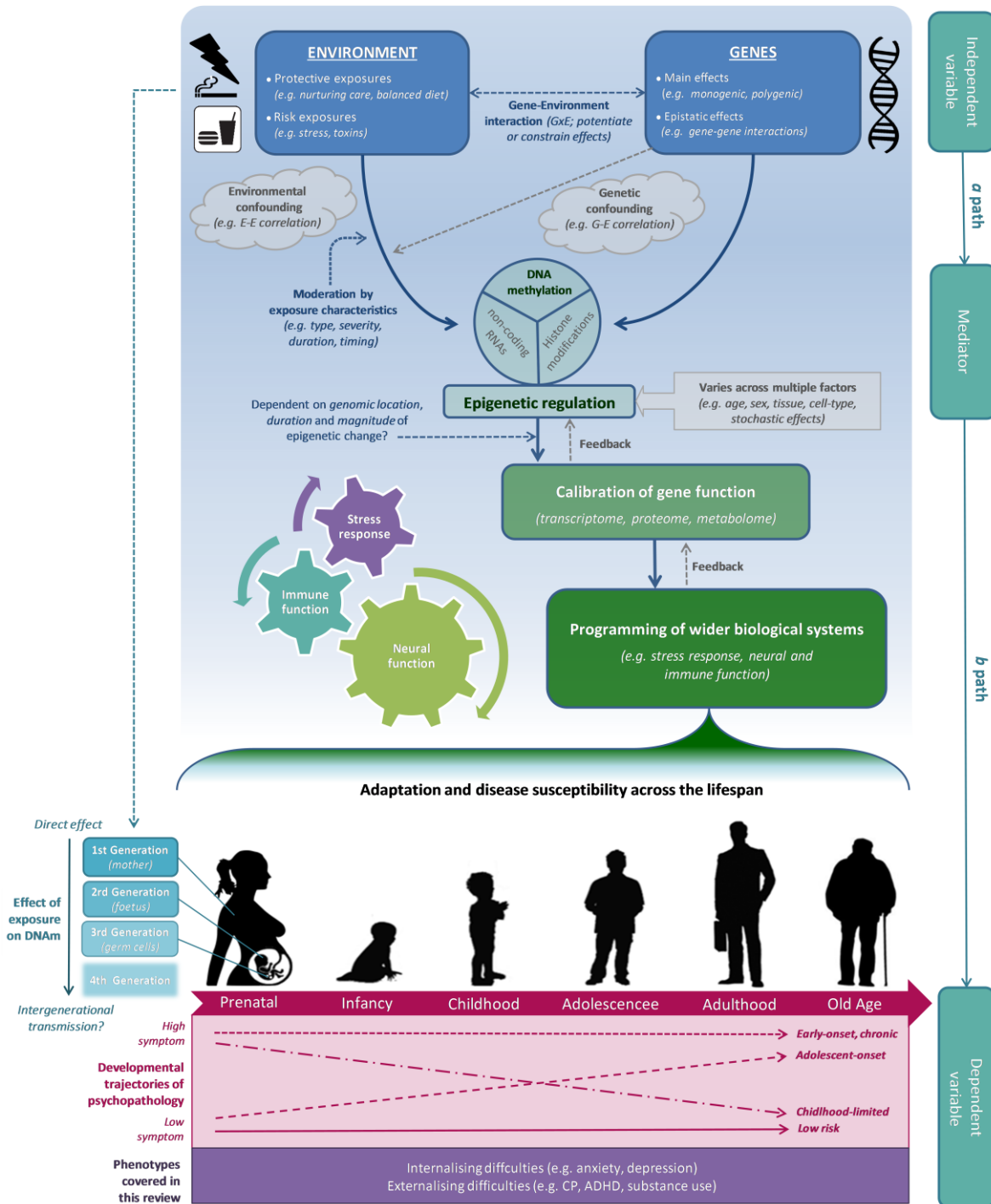


Note. Lines represent search terms on SCOPUS. Blue lines = *a* path (exposure). Red Lines = *b* path (psychopathology). Solid lines = general terms including any study across age and research design (children and adults; cross sectional and longitudinal). Wide dots = additionally specify child studies across designs. Small dots = narrow search specifying childhood and longitudinal.

Figure 2. Mediation model reviewed in this article

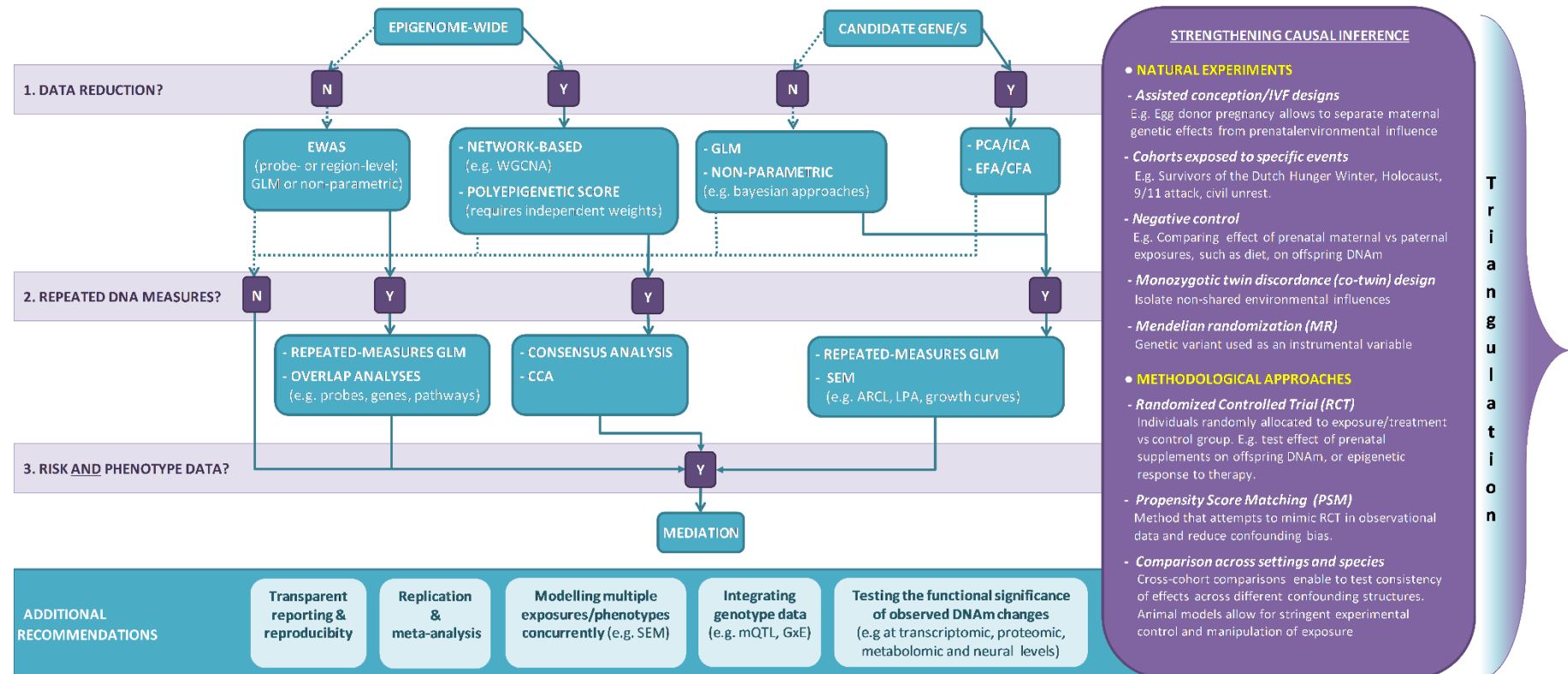
A. The conceptual mediation model reviewed in this article, whereby the independent variable (X) is the pre- and post-natal environment (e.g. toxic exposures, stress, diet); the mediator variable (M) is DNAm extracted from peripheral tissues in living subjects; and the dependent variable (Y) is developmental psychopathology, including internalising and externalising difficulties in childhood and adolescence. Of note, both the a path and the b path may be moderated by genetic effects. **B.** An empirical example of the mediation model, adapted from Cecil et al. (2016) with permission from *Translational Psychiatry*. Specifically, the model shows that, over and above other prenatal risk factors, maternal smoking prospectively associates with variation in DNAm at birth (measured as a cumulative risk index encompassing 65 genome-wide corrected loci), which in turn associates with greater substance use in adolescence (tobacco, alcohol and cannabis use). The effect of maternal smoking on substance use is partially mediated by DNAm at birth. The sample in this study was underpowered to test for genetic effects, which may have moderated the a and/or b path.

Figure 3. A simplified, integrative model of the relationship between environmental risk exposure, DNAm and psychopathology



The figure illustrates a simplified, integrative model of the relationship between environmental risks, DNAm and developmental psychopathology, and how this model relates to the mediation framework (right-hand side). Briefly, the top of the figure shows environmental effects on DNAm, how these may interact with - and be confounded by - genetic factors, and how they may be moderated by exposure characteristics. The dotted arrow on the left-hand side also demonstrates how environmental exposures may influence DNAm across multiple

generations, beginning with (empirically-supported) direct effects to hypothesised (but not empirically verified) indirect intergenerational effects. The middle of the figure depicts DNAm as the mediator, how it is embedded into the wider machinery of epigenetic regulation, and how it varies across multiple factors. Together, these epigenetic changes are shown to influence gene activity and function, which in turn contribute to the programming of wider, interconnected biological systems (e.g. immune, neural and stress function). The bottom of the figure shows how this biological cascade modulates adaptation to the environment and disease susceptibility over the lifespan, shifting developmental trajectories (typical to atypical) and contributing to the phenotypic manifestations reviewed in this article, including internalising and externalising difficulties in childhood and adolescence. Of note, the trajectories depicted are only examples from a wide range of possible trajectories, and are modelled around findings regarding the development of conduct problems (Barker & Maughan, 2009). *Abbreviations:* E-E, environment-environment; G-E, gene-environment; CP, conduct problems.

Figure 4. Analysis flowchart and methodological recommendations for future research

The flowchart on the left-hand side of the figure provides an overview of analytical approaches that may be undertaken by future DNAm studies in the field. Analyses are subdivided based on genomic coverage (i.e. epigenome-wide vs candidate gene/s), whether data is analysed at the probe-level or reduced (e.g. via network-based strategies or factor analyses), whether repeated measures of DNAm are available, and whether analyses seek to integrate both risk exposures and phenotypic outcomes. General methodological recommendations are also provided at the bottom of the figure. On the right-hand side of the figure, we present a summary of the main strategies for strengthening causal inference and highlight the need for triangulation. *Abbreviations:* N, no; Y, yes; EWAS, epigenome-wide association study; GLM, general linear model; WGCNA, weighted gene co-expression network analysis; PCA, principal component analysis; ICA, independent component analysis; EFA, exploratory factor analysis; CFA, confirmatory factor analysis; CCA, canonical correlation analysis; SEM, structural equation modelling; ARCL, auto-regressive cross lag; LPA, latent profile analysis; mQTL, methylation quantitative trait loci; GxE, gene-environment interaction.

References

- Almouzni, G., Altucci, L., Amati, B., Ashley, N., Baulcombe, D., Beaujean, N., . . . Braun, S. (2014). Relationship between genome and epigenome-challenges and requirements for future research. *BMC genomics*, *15*(1), 487.
- Anway, M. D., Cupp, A. S., Uzumcu, M., & Skinner, M. K. (2005). Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science*, *308*(5727), 1466-1469.
- Barker, E. D. (2013). The duration and timing of maternal depression as a moderator of the relationship between dependent interpersonal stress, contextual risk and early child dysregulation. *Psychological Medicine*, *43*(08), 1587-1596.
- Barker, E. D., Cecil, C. A. M., Walton, E., Jaffee, S. R., Rowe, R., Maughan, B., . . . Gaunt, T. R. (2017). A methylome-wide association study of trajectories of oppositional defiant behaviors and biological overlap with attention deficit hyperactivity disorder. *under review*.
- Barker, E. D., Cecil, C. A. M., Walton, E., & Meehan, A. J. (2017). Genetic influences on childhood disruptive behaviors *The Wiley Handbook of Disruptive and Impulse-Control Disorders* (pp. in press). USA: Wiley.
- Barker, E. D., Kirkham, K., Ng, J., & Jensen, S. M. K. (2013). Prenatal maternal depression, prenatal nutrition and child cognitive function. *British Journal of Psychiatry*, *203*, 417-421.
- Barker, E. D., & Maughan, B. (2009). Differentiating Early-Onset Persistent Versus Childhood-Limited Conduct Problem Youth. *American Journal of Psychiatry*, *166*, 900-908. doi:10.1176/appi.ajp.2009.08121770
- Beach, S. R., Brody, G. H., Todorov, A. A., Gunter, T. D., & Philibert, R. A. (2010). Methylation at SLC6A4 is linked to family history of child abuse: an examination of the Iowa Adoptee sample. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, *153*(2), 710-713.
- Binder, A. M., & Michels, K. B. (2013). The causal effect of red blood cell folate on genome-wide methylation in cord blood: a Mendelian randomization approach. *BMC Bioinformatics*, *14*(1), 1.
- Borghol, N., Suderman, M., McArdle, W., Racine, A., Hallett, M., Pembrey, M., . . . Szyf, M. (2012). Associations with early-life socio-economic position in adult DNA

- methylation. *International journal of epidemiology*, *41*(1), 62-74.
doi:10.1093/ije/dyr147
- Bowers, M. E., & Yehuda, R. (2016). Intergenerational transmission of stress in humans. *Neuropsychopharmacology*, *41*(1), 232-244.
- Breton, C. V., Byun, H.-M., Wenten, M., Pan, F., Yang, A., & Gilliland, F. D. (2009). Prenatal tobacco smoke exposure affects global and gene-specific DNA methylation. *American journal of respiratory and critical care medicine*, *180*(5), 462-467.
- Cao-Lei, L., Massart, R., Suderman, M. J., Machnes, Z., Elgbeili, G., Laplante, D. P., . . . King, S. (2014). DNA methylation signatures triggered by prenatal maternal stress exposure to a natural disaster: Project Ice Storm. *PLoS ONE*, *9*(9), e107653.
- Cardenas, A., Koestler, D. C., Houseman, E. A., Jackson, B. P., Kile, M. L., Karagas, M. R., & Marsit, C. J. (2015). Differential DNA methylation in umbilical cord blood of infants exposed to mercury and arsenic in utero. *Epigenetics*, *10*(6), 508-515.
- Cecil, C. A., Viding, E., Barker, E. D., Guiney, J., & McCrory, E. J. (2014). Double disadvantage: The influence of childhood maltreatment and community violence exposure on adolescent mental health. *Journal of Child Psychology and Psychiatry*, *55*(7), 839-848.
- Cecil, C. A. M., Jaffee, S. R., O'Connor, T. G., Oullet-Morin, I., Smith, R., Relton, C. L., . . . Barker, E. D. (2016). Neonatal DNA methylation and early-onset conduct problems: A genome-wide, prospective study. *under review*.
- Cecil, C. A. M., Lysenko, L., jaffee, S. R., Relton, C. L., Mill, J., & Barker, E. D. (2014). Environmental risk, Oxytocin Receptor Gene (OXTR) methylation and youth callous-unemotional traits: a 13-year longitudinal study. *Molecular Psychiatry*, *19*, 1071-1077. doi:doi:10.1038/mp.2014.95
- Cecil, C. A. M., Smith, R. G., Walton, E., Mill, J., McCrory, E. J., & Viding, E. (2016). Epigenetic signatures of childhood abuse and neglect: Implications for psychiatric vulnerability Assessment. *Journal of Psychiatric Research*, *in press*.
- Cecil, C. A. M., Walton, E., Smith, R. G., Viding, E., McCrory, E. J., Relton, C. L., . . . Barker, E. D. (2016). DNA methylation and substance use risk: A prospective, genome-wide study spanning gestation to adolescence. *Translational Psychiatry*, *in press*(e), e976. doi:doi:10.1038/tp.2016.247
- Choi, S.-W., & Friso, S. (2010). Epigenetics: a new bridge between nutrition and health. *Advances in Nutrition: An International Review Journal*, *1*(1), 8-16.

- Cicchetti, D., & Rogosch, F. (1996). Equifinality and multifinality in developmental psychopathology. *Development and Psychopathology*, *8*, 597-600.
- Dadds, M. R., Moul, C., Cauchi, A., Dobson-Stone, C., Hawes, D. J., Brennan, J., & Ebstein, R. E. (2014). Methylation of the oxytocin receptor gene and oxytocin blood levels in the development of psychopathy. *Development and Psychopathology*, *26*(01), 33-40.
- Dadds, M. R., Moul, C., Hawes, D. J., Mendoza Diaz, A., & Brennan, J. (2015). Individual differences in childhood behavior disorders associated with epigenetic modulation of the cortisol receptor gene. *Child Development*, *86*(5), 1311-1320.
- Dadds, M. R., Schollar-Root, O., Lenroot, R., Moul, C., & Hawes, D. J. (2016). Epigenetic regulation of the DRD4 gene and dimensions of attention-deficit/hyperactivity disorder in children. *European Child & Adolescent Psychiatry*, *25*(10), 1081-1089.
- Davies, M., Volta, M., Pidsley, R., Lunnon, K., Dixit, A., Lovestone, S., . . . Mill, J. (2012). Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome Biology*, *13*(6), R43.
- Dempster, E. L., Wong, C. C., Lester, K. J., Burrage, J., Gregory, A. M., Mill, J., & Eley, T. C. (2014). Genome-wide methylomic analysis of monozygotic twins discordant for adolescent depression. *Biological Psychiatry*, *76*(12), 977-983.
- Ding, K., Yang, J., Reynolds, G. P., Chen, B., Shao, J., Liu, R., . . . Wen, J. (2016). DAT1 methylation is associated with methylphenidate response on oppositional and hyperactive-impulsive symptoms in children and adolescents with ADHD. *The World Journal of Biological Psychiatry*, 1-9.
- Essex, M. J., Thomas Boyce, W., Hertzman, C., Lam, L. L., Armstrong, J. M., Neumann, S. M. A., & Kobor, M. S. (2013). Epigenetic Vestiges of Early Developmental Adversity: Childhood Stress Exposure and DNA Methylation in Adolescence. *Child Development*, *84*(1), 58-75. doi:10.1111/j.1467-8624.2011.01641.x
- Gaunt, T. R., Shihab, H. A., Hemani, G., Min, J. L., Woodward, G., Lyttleton, O., . . . Ho, K. (2016). Systematic identification of genetic influences on methylation across the human life course. *Genome Biology*, *17*(1), 1.
- Godfrey, K. M., Sheppard, A., Gluckman, P. D., Lillycrop, K. A., Burdge, G. C., McLean, C., . . . Crozier, S. R. (2011). Epigenetic gene promoter methylation at birth is associated with child's later adiposity. *Diabetes*, *60*(5), 1528-1534.
- Guillemin, C., Provençal, N., Suderman, M., Côté, S. M., Vitaro, F., Hallett, M., . . . Szyf, M. (2014). DNA methylation signature of childhood chronic physical aggression in T cells of both men and women. *PLoS ONE*, *9*(1), e86822.

- Hannon, E., Spiers, H., Viana, J., Pidsley, R., Burrage, J., Murphy, T. M., . . . Schalkwyk, L. C. (2016). Methylation QTLs in the developing brain and their enrichment in schizophrenia risk loci. *Nature Neuroscience*, *19*(1), 48-54.
- Heijmans, B. T., Kremer, D., Tobi, E. W., Boomsma, D. I., & Slagboom, P. E. (2007). Heritable rather than age-related environmental and stochastic factors dominate variation in DNA methylation of the human IGF2/H19 locus. *Human Molecular Genetics*, *16*(5), 547-554.
- Heijmans, B. T., Tobi, E. W., Stein, A. D., Putter, H., Blauw, G. J., Susser, E. S., . . . Lumey, L. (2008). Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proceedings of the National Academy of Sciences*, *105*(44), 17046-17049.
- Henikoff, S., & Matzke, M. A. (1997). Exploring and explaining epigenetic effects. *Trends in Genetics*, *13*, 293-295. doi:10.1016/s0168-9525(97)01219-5
- Hompes, T., Izzi, B., Gellens, E., Morreels, M., Fieuws, S., Pexsters, A., . . . Freson, K. (2013). 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. *Journal of Psychiatric Research*, *47*(7), 880-891.
- Hoyo, C., Murtha, A. P., Schildkraut, J. M., Forman, M. R., Calingaert, B., Demark-Wahnefried, W., . . . Murphy, S. K. (2011). Folic acid supplementation before and during pregnancy in the Newborn Epigenetics Study (NEST). *BMC Public Health*, *11*(1), 46.
- Jaenisch, R., & Bird, A. (2003). Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nature Genetics*, *33*, 245-254.
- Joubert, B. R., Håberg, S. E., Nilsen, R. M., Wang, X., Vollset, S. E., Murphy, S. K., . . . Cupul-Uicab, L. A. (2012). 450K epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy.
- Kaminsky, Z. A., Tang, T., Wang, S.-C., Ptak, C., Oh, G. H., Wong, A. H., . . . Tysk, C. (2009). DNA methylation profiles in monozygotic and dizygotic twins. *Nature Genetics*, *41*(2), 240-245.
- Kofink, D., Boks, M. P., Timmers, H. M., & Kas, M. J. (2013). Epigenetic dynamics in psychiatric disorders: environmental programming of neurodevelopmental processes. *Neuroscience & Biobehavioral Reviews*, *37*(5), 831-845.

- Kumsta, R., Marzi, S., Viana, J., Dempster, E., Crawford, B., Rutter, M., . . . Sonuga-Barke, E. (2016). Severe psychosocial deprivation in early childhood is associated with increased DNA methylation across a region spanning the transcription start site of CYP2E1. *Translational Psychiatry*, 6(6), e830.
- Ladd-Acosta, C., & Fallin, M. D. (2015). The role of epigenetics in genetic and environmental epidemiology.
- Lawlor, D. A., Smith, G. D., O'Callaghan, M., Alati, R., Mamun, A. A., Williams, G. M., & Najman, J. M. (2007). Epidemiologic evidence for the fetal overnutrition hypothesis: findings from the mater-university study of pregnancy and its outcomes. *American Journal of Epidemiology*, 165(4), 418-424.
- Lawlor, D. A., Tilling, K., & Smith, G. D. (2017). Triangulation in aetiological epidemiology. *International Journal of Epidemiology*, dyw314.
- Liang, L., & Cookson, W. O. (2014). Grasping nettles: cellular heterogeneity and other confounders in epigenome-wide association studies. *Human Molecular Genetics*, ddu284.
- McGowan, P. O., Sasaki, A., D'Alessio, A. C., Dymov, S., Labonté, B., Szyf, M., . . . Meaney, M. J. (2009). Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nature Neuroscience*, 12(3), 342-348.
- McKay, J. A., Groom, A., Potter, C., Coneyworth, L. J., Ford, D., Mathers, J. C., & Relton, C. L. (2012). Genetic and non-genetic influences during pregnancy on infant global and site specific DNA methylation: role for folate gene variants and vitamin B 12. *PLoS ONE*, 7(3), e33290.
- McRae, A. F., Powell, J. E., Henders, A. K., Bowdler, L., Hemani, G., Shah, S., . . . Montgomery, G. W. (2014). Contribution of genetic variation to transgenerational inheritance of DNA methylation. *Genome Biology*, 15(5), R73.
- Meaney, M. J. (2010). Epigenetics and the Biological Definition of Gene & Environment Interactions. *Child Development*, 81, 41-79.
- Meehan, A. J., Cecil, C. A. M., Maughan, B., Relton, C. L., Gaunt, T. R., McArdle, W., & Barker, E. D. (2017). Epigenetic Correlates of Low Prosocial Behavior: A Prospective, Genome-Wide Study. *under review*.
- Mill, J., & Heijmans, B. T. (2013). From promises to practical strategies in epigenetic epidemiology. *Nat Rev Genet*, 14(8), 585-594. doi:10.1038/nrg3405

- Monk, C., Feng, T., Lee, S., Krupka, I., Champagne, F. A., & Tycko, B. (2016). Distress During Pregnancy: Epigenetic Regulation of Placenta Glucocorticoid-Related Genes and Fetal Neurobehavior. *American Journal of Psychiatry*.
- Monk, C., Georgieff, M. K., & Osterholm, E. A. (2013). Research Review: Maternal prenatal distress and poor nutrition – mutually influencing risk factors affecting infant neurocognitive development. *Journal of Child Psychology and Psychiatry*, *54*(2), 115-130. doi:10.1111/jcpp.12000
- Morris, T. J., & Beck, S. (2015). Analysis pipelines and packages for Infinium HumanMethylation450 BeadChip (450k) data. *Methods*, *72*, 3-8.
- Moul, C., Dobson-Stone, C., Brennan, J., Hawes, D. J., & Dadds, M. R. (2015). Serotonin 1B Receptor Gene (HTR1B) Methylation as a Risk Factor for Callous-Unemotional Traits in Antisocial Boys. *PLoS ONE*, *10*(5), e0126903.
- Murgatroyd, C., Quinn, J., Sharp, H., Pickles, A., & Hill, J. (2015). Effects of prenatal and postnatal depression, and maternal stroking, at the glucocorticoid receptor gene. *Translational Psychiatry*, *5*(5), e560.
- Naumova, O. Y., Lee, M., Kuposov, R., Szyf, M., Dozier, M., & Grigorenko, E. L. (2012). Differential patterns of whole-genome DNA methylation in institutionalized children and children raised by their biological parents. *Development and Psychopathology*, *24*(01), 143-155.
- Non, A. L., & Thayer, Z. M. (2015). Epigenetics for anthropologists: An introduction to methods. *American Journal of Human Biology*, *27*(3), 295-303.
- O'Connor, T. G., Heron, J., Golding, J., Beveridge, M., & Glover, V. (2002). Maternal antenatal anxiety and children's behavioural/emotional problems at 4 years. *British Journal of Psychiatry*, *180*, 502-508.
- Oberlander, T. F., Weinberg, J., Papsdorf, M., Grunau, R., Misri, S., & Devlin, A. M. (2008). Prenatal exposure to maternal depression, neonatal methylation of human glucocorticoid receptor gene (NR3C1) and infant cortisol stress responses. *Epigenetics*, *3*(2), 97-106.
- Obermann-Borst, S. A., Eilers, P. H., Tobi, E. W., de Jong, F. H., Slagboom, P. E., Heijmans, B. T., & Steegers-Theunissen, R. P. (2013). Duration of breastfeeding and gender are associated with methylation of the LEPTIN gene in very young children. *Pediatric Research*, *74*(3), 344-349.
- Ouellet-Morin, I., Wong, C. C. Y., Danese, A., Pariante, C. M., Papadopoulos, A. S., Mill, J., & Arseneault, L. (2013). 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, 43(09), 1813-1823. doi:doi:10.1017/S0033291712002784

Parade, S. H., Ridout, K. K., Seifer, R., Armstrong, D. A., Marsit, C. J., McWilliams, M. A., & Tyrka, A. R. (2016). Methylation of the glucocorticoid receptor gene promoter in preschoolers: Links with internalizing behavior problems. *Child Development*, 87(1), 86-97.

Park, S., Lee, J.-M., Kim, J.-W., Cho, D.-Y., Yun, H., Han, D., . . . Kim, B.-N. (2015).

Associations between serotonin transporter gene (SLC6A4) methylation and clinical characteristics and cortical thickness in children with ADHD. *Psychological Medicine*, 45(14), 3009-3017.

Peter, C. J., Fischer, L. K., Kundakovic, M., Garg, P., Jakovcevski, M., Dincer, A., . . . Bryce, C. P. (2016). DNA methylation signatures of early childhood malnutrition associated with impairments in attention and cognition. *Biological Psychiatry*.

Philibert, R. A., Gunter, T. D., Beach, S. R., Brody, G. H., & Madan, A. (2008). MAOA methylation is associated with nicotine and alcohol dependence in women. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 147(5), 565-570.

Pidsley, R., Zotenko, E., Peters, T. J., Lawrence, M. G., Risbridger, G. P., Molloy, P., . . . Clark, S. J. (2016). Critical evaluation of the Illumina MethylationEPIC BeadChip microarray for whole-genome DNA methylation profiling. *Genome Biology*, 17(1), 208.

Poston, L. (2012). Maternal obesity, gestational weight gain and diet as determinants of offspring long term health. *Best Practice & Research Clinical Endocrinology & Metabolism*, 26(5), 627-639.

Provençal, N., Suderman, M. J., Guillemin, C., Vitaro, F., Côté, S. M., Hallett, M., . . . Szyf, M. (2014). Association of childhood chronic physical aggression with a DNA methylation signature in adult human T cells. *PLoS ONE*, 9(4), e89839.

Provençal, N., Suderman, M. J., Vitaro, F., Szyf, M., & Tremblay, R. E. (2013). Childhood Chronic Physical Aggression Associates with Adult Cytokine Levels in Plasma. *PLoS ONE*, 8(7), e69481. doi:10.1371/journal.pone.0069481

Relton, C. L., & Davey Smith, G. (2012). Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. *International Journal of Epidemiology*, 41(1), 161-176.
doi:10.1093/ije/dyr233

- Relton, C. L., Gaunt, T., McArdle, W., Ho, K., Duggirala, A., Shihab, H., . . . Reik, W. (2015). Data resource profile: accessible resource for integrated epigenomic studies (aries). *International Journal of Epidemiology*, dyv072.
- Richmond, R. C., Al-Amin, A., Smith, G. D., & Relton, C. L. (2014). Approaches for drawing causal inferences from epidemiological birth cohorts: a review. *Early Human Development*, 90(11), 769-780.
- Richmond, R. C., Simpkin, A. J., Woodward, G., Gaunt, T. R., Lyttleton, O., McArdle, W. L., . . . Tilling, K. (2015). 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*, 24, 2201-2217.
- Rijlaarsdam, J., Cecil, C. A., Walton, E., Mesirov, M. S., Relton, C. L., Gaunt, T. R., . . . Barker, E. D. (2017). Prenatal unhealthy diet, insulin-like growth factor 2 gene (IGF2) methylation, and attention deficit hyperactivity disorder symptoms in youth with early-onset conduct problems. *Journal of Child Psychology and Psychiatry*, 58, 19–27.
- Rijlaarsdam, J., Pappa, I., Walton, E., Bakermans-Kranenburg, M. J., Mileva-Seitz, V. R., Rippe, R. C., . . . Felix, J. F. (2016). An epigenome-wide association meta-analysis of prenatal maternal stress in neonates: A model approach for replication. *Epigenetics*, 11(2), 140-149.
- Roberts, S., Lester, K., Hudson, J., Rapee, R., Creswell, C., Cooper, P. J., . . . Wong, C. (2014). Serotonin transporter methylation and response to cognitive behaviour therapy in children with anxiety disorders. *Translational Psychiatry*, 4(9), e444.
- Ruggeri, B., Nymberg, C., Vuoksima, E., Lourdasamy, A., Wong, C. P., Carvalho, F. M., . . . Banaschewski, T. (2015). Association of protein phosphatase PPM1G with alcohol use disorder and brain activity during behavioral control in a genome-wide methylation analysis. *American Journal of Psychiatry*, 172(6), 543-552.
- Sen, A., Heredia, N., Senut, M.-C., Hess, M., Land, S., Qu, W., . . . Ruden, D. M. (2015). Early life lead exposure causes gender-specific changes in the DNA methylation profile of DNA extracted from dried blood spots.
- Shakya, K., O'Connell, M. J., & Ruskin, H. J. (2012). The landscape for epigenetic/epigenomic biomedical resources. *Epigenetics*, 7(9), 982-986.
- Sharp, G. C., Lawlor, D. A., Richmond, R. C., Fraser, A., Simpkin, A., Suderman, M., . . . Ring, S. M. (2015). 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*, dyv042.
- Szutorisz, H., & Hurd, Y. L. (2016). Feeding the Developing Brain: The Persistent Epigenetic Effects of Early Life Malnutrition. *Biological Psychiatry*, 80(10), 730-732.
- Szyf, M. (2015). Prospects for the development of epigenetic drugs for CNS conditions. *Nature Reviews Drug Discovery*, 14(7), 461-474.
- Talati, A., Keyes, K., & Hasin, D. (2016). Changing relationships between smoking and psychiatric disorders across twentieth century birth cohorts: clinical and research implications. *Molecular Psychiatry*.
- Taudt, A., Colomé-Tatché, M., & Johannes, F. (2016). Genetic sources of population epigenomic variation. *Nature Reviews Genetics*.
- Teh, A. L., Pan, H., Chen, L., Ong, M.-L., Dogra, S., Wong, J., . . . Saw, S.-M. (2014). The effect of genotype and in utero environment on interindividual variation in neonate DNA methylomes. *Genome Research*, 24(7), 1064-1074.
- Tobi, E. W., Goeman, J. J., Monajemi, R., Gu, H., Putter, H., Zhang, Y., . . . Müller, F. (2014). DNA methylation signatures link prenatal famine exposure to growth and metabolism. *Nature communications*, 5.
- Turecki, G., & Meaney, M. J. (2016). Effects of the social environment and stress on glucocorticoid receptor gene methylation: a systematic review. *Biological Psychiatry*, 79(2), 87-96.
- Van der Knaap, L., Schaefer, J., Franken, I., Verhulst, F., van Oort, F., & Riese, H. (2014). Catechol-O-methyltransferase gene methylation and substance use in adolescents: the TRAILS study. *Genes, Brain and Behavior*, 13(7), 618-625.
- van der Knaap, L. J., Oldehinkel, A. J., Verhulst, F. C., van Oort, F. V., & Riese, H. (2015). Glucocorticoid receptor gene methylation and HPA-axis regulation in adolescents. The TRAILS study. *Psychoneuroendocrinology*, 58, 46-50.
- van Mil, N. H., Steegers-Theunissen, R. P., Bouwland-Both, M. I., Verbiest, M. M., Rijlaarsdam, J., Hofman, A., . . . Verhulst, F. C. (2014). DNA methylation profiles at birth and child ADHD symptoms. *Journal of Psychiatric Research*, 49, 51-59.
- Waddington, C. H. (1957). The strategy of the genes. A discussion of some aspects of theoretical biology. With an appendix by H. Kacser. *The strategy of the genes. A discussion of some aspects of theoretical biology. With an appendix by H. Kacser.*

- Walton, E., Cecil, C. A. M., Suderman, M., Liu, J., Turner, J. A., Calhoun, V. D., . . . Barker, E. D. (2017). Longitudinal epigenetic predictors of amygdala:hippocampus volume ratio *under review*.
- Walton, E., Hass, J., Liu, J., Roffman, J. L., Bernardoni, F., Roessner, V., . . . Ehrlich, S. (2015). Correspondence of DNA methylation between blood and brain tissue and its application to schizophrenia research. *Schizophrenia Bulletin*, sbv074.
- Walton, E., Pingault, J.-B., Cecil, C., Gaunt, T., Relton, C., Mill, J., & Barker, E. (2016). Epigenetic profiling of ADHD symptoms trajectories: a prospective, methylome-wide study. *Molecular Psychiatry*.
- Wang, D., Szyf, M., Benkelfat, C., Provençal, N., Turecki, G., Caramaschi, D., . . . Booij, L. (2012). Peripheral SLC6A4 DNA methylation is associated with in vivo measures of human brain serotonin synthesis and childhood physical aggression. *PLoS ONE*, 7(6), e39501.
- Weaver, I. C. G., Cervoni, N., Champagne, F. A., D'Alessio, A. C., Sharma, S., Seckl, J. R., . . . Meaney, M. J. (2004). Epigenetic programming by maternal behavior. *Nature Neuroscience*, 7(8), 847-854.
- Weder, N., Zhang, H., Jensen, K., Yang, B. Z., Simen, A., Jackowski, A., . . . Perepletchikova, F. (2014). Child abuse, depression, and methylation in genes involved with stress, neural plasticity, and brain circuitry. *Journal of the American Academy of Child & Adolescent Psychiatry*, 53(4), 417-424. e415.
- Xu, Y., Chen, X.-T., Luo, M., Tang, Y., Zhang, G., Wu, D., . . . Wang, H.-L. (2015). Multiple epigenetic factors predict the attention deficit/hyperactivity disorder among the Chinese Han children. *Journal of Psychiatric Research*, 64, 40-50.
- Yehuda, R., Daskalakis, N. P., Bierer, L. M., Bader, H. N., Klengel, T., Holsboer, F., & Binder, E. B. (2016). Holocaust exposure induced intergenerational effects on FKBP5 methylation. *Biological Psychiatry*, 80(5), 372-380.
- Yehuda, R., McFarlane, A., & Shalev, A. (1998). Predicting the development of posttraumatic stress disorder from the acute response to a traumatic event. *Biological Psychiatry*, 44(12), 1305-1313.
- Zheng, J., Xiao, X., Zhang, Q., & Yu, M. (2014). DNA methylation: the pivotal interaction between early-life nutrition and glucose metabolism in later life. *British Journal of Nutrition*, 112(11), 1850-1857.