POTENTIAL BIOLOGICAL EMBEDDING OF ADVERSE AND ENRICHING ENVIRONMENTS THROUGH EPIGENETIC MEDIATORS

BY

JOHN RYAN PFEIFFER

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Psychology
in the Graduate College of the
University of Illinois Urbana-Champaign, 2021

Urbana, Illinois

Doctoral Committee:

Professor Josh Gulley, Chair
Professor Monica Uddin, Director of Research, University of South Florida
Professor Janice Juraska
Associate Professor David Zhao
Associate Professor Jamie Derringer

ABSTRACT

In humans, adverse social exposures (ASEs) such as stress or trauma can exert both acute and long-lasting biological and psychological effects. This is especially true during critical periods of development such as childhood, adolescence, and young adulthood. Numerous biological systems are affected in response to ASEs, but variability in neuroimaging structural, functional, and connectivity measures within the fronto-limbic pathway are repeatedly observed in association with ASEs. The hippocampus, amygdala, and frontal cortex, key members of the fronto-limbic pathway, support critical cognitive functions that are notably impaired in numerous mood- and stress-related psychiatric illnesses. Replicable variability of the aforementioned neuroimaging measures are also noted in associated with psychiatric illnesses. Together, variability in these fronto-limbic structure, function, and connectivity measures are regarded as neural correlates of ASEs and neural endophenotypes of psychiatric illness.

In rodents, stressor paradigms are commonly employed to investigate the molecular effects of ASEs and can be a significant source of human-relevant insight. With these rodent models, researchers have observed cognitive and neurobiological effects that mirror ASE-associated effects in humans. In contrast, environmental enrichment (EE) can exert *positive* physiological and cognitive effects and can rescue the negative effects of ASEs (as observed in both humans and rodents). It appears then, that the effects of ASEs and EE potentially operate on similar underlying neurobiological substrates in the fronto-limbic pathway.

To this end, the molecular mechanisms supporting the biological embedding of our external environments have come under investigation. Epigenetic mechanisms, such as 5-methyl-cytosine (5mC) marks at cytosine-phosphate-guanine (CpG) base pair sequences, are now

regarded as potential mediators of gene by environment interactions. However, current technological limitations make it impossible to measure epigenetics in the living human brain--the primary etiological tissue of interest regarding mental health outcomes.

To overcome these barriers, researchers have forged alternative paths towards elucidating the relationships between our external environments and the epigenetic mechanisms that purportedly potentiate experience-dependent changes across the CNS. Two such methods are utilized in the current research, after an initial overview of relevant literature in Chapter 1. In Chapter 2 and 3, I use genome-scale bioinformatics approaches in tandem with structural neuroimaging measures from the fronto-limbic pathway to investigate whether variability in peripheral tissue-derived 5mC clusters may represent pathways through which ASEs experienced in childhood and adulthood may become biologically embedded. In Chapter 4, I leverage secondary data analysis methods on a mouse model of EE to investigate the effects of EE on locus-specific epigenomic and transcriptomic measures from dorsal and ventral dentate gyrus in the hippocampus. Chapter 5 discusses the implications of these findings and their contributions to the existing literature surrounding neurobiological and epigenetic variability associated with ASEs and EE.

Overall, this work uses novel genome-scale bioinformatics approaches to investigate underlying epigenetic mechanisms associated with, and/or *potentially* responsible for the biological embedding of external environments. It provides novel insights to the fields of stress, trauma, and psychiatric illness by providing network-level and locus-specific peripheral biomarkers of ASEs, peripheral epigenetic profiles associated with neural endophenotypes of psychiatric illness, in addition to potentially mechanistic epigenetic mediators of EE-related biological embedding in the hippocampus.

ACKNOWLEDGEMENTS

I look back at the whole of my PhD journey and view it as a truly transformative period in my life. I see strength and weakness, laughter and tears, confidence and self-doubt; I wouldn't trade it for the world. I would first like to thank my advisor, Monica Uddin, in addition to my committee members: Josh Gulley, Dave Zhao, Janice Juraska, and Jamie Derringer. I would also like to thank my Aunt Cindy, Alaina Kanfer, and my teammates at SimBioSys: Joe Peterson, John Cole, Tyler Earnest, Mike Hallock, and Tushar Pandey.

When people look at a freshly minted PhD, I believe they seldom consider the immense network of support that the new doctor required throughout the process. I believe my journey exemplifies this. When I first came on to Monica's team at the University of Illinois, I was the canonical example of a bright-eyed first year eager to please and absorb all the knowledge that I could. With the guidance and attention of now Drs. Angie Bustamante and Grace Kim, I was able to get a grip on the molecular and computational basics that would come to be the foundation of my knowledge. Don Armstrong, quiet and reserved, yet vastly knowledgeable beyond the comprehension of a first year (and a sixth year), offered advice and guidance where-ever he could. Monica was the ever-present captain, providing the highest quality perspective, wisdom, and guidance that a student could ask for. This team of people set me up for success.

The years progressed, and I became intimately aware that the adage, "change is the only constant in life", was much more than just an adage; it was ground-truth. Only with the graduation of Angie and Grace did I fully realize what they were to me: mentors, friends, confidants, peers. Good people. Strong people. I am forever grateful for the things they taught

me and the moments we shared. The wheel of change kept on rolling, and Monica accepted a highly coveted position at the University of South Florida (as successful members of academia will do). I was faced with, what was at that point, one of the most momentous decisions of my life. I could move to USF, I could find a new lab at the University of Illinois, or I could maintain the mentor/mentee relationship with Monica (remotely). I took my time and consulted with several U of I Professors and Psychology Department administrators. To those people, I say thank you. Even if you couldn't offer me advice or guidance on a clear path forward, it meant a great deal that I could talk openly with you.

Ultimately, I decided to maintain my relationship with Monica remotely, and to focus my efforts fully on computational genomic biology. The whole process was, candidly, very difficult at first. Then, it was difficult in the middle. Finally, it was difficult at the end. But it's important to conceptualize that the best things in life are worth working *hard* for. Throughout the entirety of this experience, my significant other Valeria Sanabria Guillen was my ever-shining North Star. My parents, John and Diana, were always willing and eager to lend an ear, or a shoulder to cry on. My brother and sister, Michael and Ashton, were ever present in my life, regardless of what was going on in theirs. My "grad" friends Dominic, Gaby, Ben, Ami, Stef, and Alex were a tremendous source of wonderment, shenanigans, debauchery, and fundamental support that I couldn't go without. I would also like to thank Jimmy, Devin, and Mike. You guys surprise me every single day. To all these people, and to the many more who I crossed paths with during my time here, I give my sincerest thanks and gratitude. I couldn't have done it without you.

I would also like to thank my collaborators and co-authors, the researchers and participants from the Duke Neurogenetics Study (Chapter 2), and the researchers and participants from the Grady Trauma Project (Chapter 3).

Dedication

This work is dedicated to my parents, Diana and John Pfeiffer, to my siblings, Ashton and Michael Pfeiffer, to my partner, Valeria Sanabria Guillen, and to my pup, Murph. Without your unwavering support, I would not be the man I am today.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

1.1 Adverse social exposures and their prevalence

Humans can experience a wide range of adverse social exposures (ASEs) throughout their lifetimes. These could include childhood maltreatment, trauma exposure (TE) across childhood and adulthood, stressful life events (SLEs), measures of low socioeconomic status (SES) such as low income or low education attainment, or even living with a caregiver diagnosed with psychopathology, among others.

These exposures are both prevalent and impactful in the United States (US). Regarding TE, past year physical or psychological abuse in US children ranges from 4% to 16%[1], meaning that roughly up to one in six children has been subjected to abuse in the past year. Projecting forward in development, it is estimated that between 50% and 90% of adults will have experienced as least one occurrence of a TE in their lifetime[2–5] – the majority of adults in our social networks have likely been exposed to trauma. In addition, poverty rates in the US range from 12% to 15%[6], and over one million students per year leave high school without a diploma or GED[7]. Likely due, in part, to the high prevalence of these exposures, children often grow up in homes in which a parent is diagnosed with a psychiatric illness. To this end, it is estimated that ~12.8 million parents (18.2%) suffer yearly from some form of psychiatric illness, and that ~2.7 million parents (3.8%) suffer yearly from a *serious* psychiatric illness[8].

1.2 Effects of adverse social exposures in humans

In addition to their wide-reaching prevalence in the US, it is apparent that these exposures can exert both acute and long-lasting cognitive and psychological effects. For

example, investigations into the effects of early life stress and trauma reveal decreases in executive function as measured in adolescents[9] and adults[10]. Low SES is associated with similar effects in children[11] and adults[12], and poverty-related stress is associated with increased internalizing/externalizing symptoms, in addition to involuntary engagement/disengagement stress responses in those same age groups[13]. In a similar vein, researchers have also observed that the children of caregivers with major depressive disorder (MDD), are more likely to develop aberrant internalizing/externalizing symptoms as well as social, cognitive, and academic difficulties[14, 15].

Given the notable cognitive effects associated with these exposure, it is not surprising that ASEs are also frequently associated with an increased risk of psychopathology development[16–23]. Indeed, the effects of childhood maltreatment or trauma on long term mental health prognoses are well-documented. In a recent meta-analysis of over 190 studies and 68,000 participants, researchers observed that higher measures of childhood maltreatment (including traumas such as emotional abuse, physical abuse, sexual abuse) were positively associated with MDD diagnosis later in life and with depression symptom severity[24]. This built upon previous work in the field showing that exposure to *any* form of maltreatment during childhood was associated with a two-fold risk of MDD development, and nearly a three-fold risk of generalized anxiety disorder (GAD) development in adulthood[25]. TE in adulthood is also associated with an increased risk of MDD, which in turn is highly comorbid with post-traumatic stress disorder (PTSD) development[22, 23]. Other ASEs such as low SES and SLEs are also associated with the development of psychopathology, specifically including MDD[16–19], schizophrenia[20], and neuroticism[21].

The societal impact of these psychiatric illnesses cannot be understated. In fact, in calculating and assessing the years lived with disability (YLD) across 289 diseases from 1990 to 2010, researchers found that the overall top contributors to global YLD were mental and behavioral disorders[26]. More specifically, MDD was second only to lower back injury in their ranking, and anxiety disorders occupied the sixth highest position on the list[26]. A similar study conducted in the US specifically found that one of the top three leading causes of disability was depression[27]. Branching out, researchers noted that risk of suicide is particularly high in those diagnosed with PTSD, that a majority of potential PTSD sufferers do not receive treatment, and that the overall consequences of PTSD are comparable to those of other seriously impairing mental health disorders[28].

1.3 Potential substrates of biological embedding

Although individual mood- or stress-related psychiatric illnesses can present with heterogenous symptom clusters, across illnesses a number of symptoms are shared[29]. Therefore, it stands to reason that ASEs act on common physiological or neurobiological substrates to portend such common symptoms. To this end, major physiological systems affected by these exposures include the hypothalamic-pituitary-adrenal (HPA)-axis[30, 31], immune system[32–34], and central nervous system (CNS)[35–52]. The HPA-axis is the key mediator of the body's response to perceived stressors. In "healthy" individuals, in response to such stressors, fronto-limbic brain regions associated with cognitive function, emotional reactivity, and memory are activated[53]. These brain regions, which are critical in mediating adaptive behavioral responses, include the hippocampus, amygdala, and frontal cortices[53]. Through bi-directional axonal projections, these regions activate the HPA-axis' neural hub, the paraventricular nucleus

of the hypothalamus. From there, an endocrine cascade stimulates the eventual release of glucocorticoids (GCs) from the adrenal cortices of the kidneys; GCs are critical to immunological, metabolic, cardiac, and homeostatic functions throughout the body[54]. GCs act mechanistically in and around the CNS and peripheral nervous system, including serving as a long-range negative feedback mechanism to inhibit further HPA-axis activity[54]. Additionally, the ligands and receptors of cytokines and neurotransmitters are shared throughout the CNS, HPA-axis, and immune systems[55, 56]. Aberrant responses to this signaling can result in neural growth, inflammation, metabolism, and stress-related pathway disruption through molecular mechanisms[57].

Integrating some of these proposed mechanisms more comprehensively, the neuroimmune network hypothesis is one framework used to explain the physiological mechanisms via which ASEs are biologically embedded. The neuroimmune network hypothesis focuses on the integrated, bi-directional network of the CNS and the immune system[58], and posits that exposure to ASEs impacts communication between peripheral inflammatory signals and fronto-limbic brain regions. Researchers posit that this relationship is especially salient during childhood, a notably plastic window of physiological and psychological development[59]. These inflammatory signals disrupt the inter-dependent functions of the fronto-limbic pathway, leading to altered behavioral states, and the pre-disposition to develop aberrant stress responses later in life[60].

Given the complex interactions between the body's stress-response, immune, and central nervous systems, and the range of cognitive and psychopathology effects seen in association with ASEs, one could justly posit that the effects of ASEs could also observed in the CNS. To this end, variability in structural, functional, and connectivity measures within the fronto-limbic

pathway are repeatedly observed in association with ASEs[35–52]. These included volumetric, cortical thickness (CT), and surface area (SA) measures taken from the hippocampus, amygdala, frontal cortex, and more specifically from frontal cortex subregions (frontal pole: FP, superior frontal gyrus: SFG, rostral middle frontal gyrus: RMFG, anterior cingulate cortex: ACC, and the orbito-frontal cortex: OFC), among others. Exposures of interest from these studies in children/adolescents included low maternal support[45], hostile parenting[40], childhood neglect[36] and physical abuse[36], interpersonal trauma[49], low familial income-to-needs ratio[40], and low SES[36]. In a strong majority of these studies, researchers identified lower hippocampus, amygdala, or frontal cortex volumes in association with ASE exposure. Studies investigating ASEs in adults rendered similar results as investigations in children[47–52]. It appears then, that ASEs experienced across childhood and adulthood are generally associated with lower hippocampal, amygdala, and frontal cortex morphometry measures, and that these signatures should be considered neural correlates of such exposures. Importantly though, depending on the exposure of interest, the hemisphere-specificity, and the brain region of interest, these associations go unobserved in some studies which also report positive results[39], highlighting the heterogenous nature of the potential neurobiological effects of ASEs.

Importantly, the critical cognitive functions of the fronto-limbic pathway brain regions are notably impaired in numerous psychiatric illnesses, including but not limited to PTSD[61], MDD[62], GAD[63], bipolar disorder[64], and schizophrenia[64]. Together, variability in these fronto-limbic structure, function, and connectivity measures are regarded as *neural correlates* of ASEs and as *endophenotypes* of psychiatric illness development[65–67].

1.4 Rodent models for human-relevant insights

In rodents, stressor paradigms are commonly employed to investigate the molecular effects of ASEs and can be a significant source of human-relevant insight. These models are especially important, considering that the CNS appears to be a major substrate in the biological embedding of the external environment. In rodents, these paradigms have well-documented morphological effects on brain regions belonging to the fronto-limbic pathway[68–70], in addition to cognitive systems in rodents[71–75] that mirror effects of ASEs observed in humans.

On the other hand, environmental enrichment (EE), typically regarded as housing methods in which the potential for social interactions and novel object exploration are increased compared to standard housing[76], compensates for or rescues many effects of stress exposure in rodents[77–81]. Researchers have also observed that EE in rodents is associated with notable improvement in cognitive and behavioral domains[82–86], and neurobiological changes to the hippocampus and associated cortical regions. Increased neurogenesis[87–90], synaptogenesis[91], and gliogenesis[90], as well as increased dendrite outgrowth[83], cortical neurotransmission[92], and long term potentiation[93] have all been observed in rodent models of EE. EE has also been shown to increase adult hippocampal volume[94] and granule cell neurogenesis in the dentate gyrus[95] (a hippocampal sub-region which is responsible for neurogenesis throughout the life cycle).

Importantly, the effects of ASEs and EE are especially salient during periods in physiological development where the CNS is particularly plastic, namely childhood, adolescence, and young adulthood. Indeed, a number of the neurobiological changes associated with EE, as noted above, have been observed in these periods of developmental plasticity[90, 91, 93]. This concept was further highlighted in a recent study that applied early life stress in a post-

natal rodent model, and then rescued the behavioral (fear response to conditioned stimulus) and neurobiological (survival and activation of newborn hippocampal neurons) effects through EE[96].

It appears then, that the effects of stressor and EE paradigms in rodents operate on similar underlying neurobiological substrates: the fronto-limbic pathway, including the hippocampus and dentate gyrus. This implies that bi-directional alterations to underlying molecular regulatory networks may be at play in response to these divergent exposures. This concept is not specific to rodents, as noted by the hypothetical neglect-enrichment continuum discussed in the context of human children/adolescents[97].This hypothetical framework states that the "healthy" development of children and adolescents is dependent on environmental/caregiver facilitation of either deprivation or enrichment. As reviewed in King et al 2019[97], evidence suggests that exposure to both ends of this common spectrum act on the fronto-limbic pathways of the brain. Therefore, identifying molecular alterations in the rodent brain associated with EE could provide valuable insight regarding molecular resilience to ASEs in humans.

1.5 Studying potential mechanisms of biological embedding

To this end, the molecular mechanisms playing a role in the biological embedding of external exposures, such as ASEs and EE, have come under investigation in both humans and rodents. Numerous epigenetic mechanisms have been identified that serve as potential substrates for gene by environment interaction, including post-translational histone tail modifications, small non-coding RNA alterations, and long non-coding RNA molecule alterations. The epigenetic process of DNA methylation, which commonly refers to the addition of a 5-methyl-cytosine (5mC) residue to a cytosine-phosphate-guanine (CpG) base-pair, is another epigenetic

mechanism that is understood to enhance or reduce mRNA transcription of genes in an experience-dependent, and temporally-stable manner[98–100]. In addition to CpG methylation (mCG), neuronal cell types have specifically unique patterns of *non-CpG* methylation (mCH) that can also dictate transcriptional regulation[101]. These epigenetic mechanisms are known mediators of gene by environment interactions[102–104], but current technological limitations make it impossible to measure epigenetics in the living human brain--the primary etiological tissue of interest in regards to mental health outcomes. To overcome these innate barriers, researchers have forged alternative paths towards elucidating the relationships between our external environments and the epigenetic mechanisms that potentiate experience-dependent changes across the CNS.

Two such approaches are utilized in the current research. The first approach is utilized in Chapters 2 and 3, where human peripheral tissue measurements of mCG serve as proxies for those taken from etiological tissue. It is well-documented that these peripheral epigenetic measures *can* index changes to the HPA-axis[105, 106], immune system[107, 108], and CNS[109–111], although the direction of effect may be discordant across tissues[110, 112]. Peripheral epigenetic measures can also index CNS-relevant endophenotypes of psychiatric illness development, as observed in studies using neuroimaging and peripheral epigenetic measures *in tandem.* However, these studied have primarily utilized candidate gene approaches, for example, investigating mCG in the *SLC6A4*[113–115], *NR3C1*[116, 117], *FKBP5*[118], and *SKA2*[119, 120] genes and neuroimaging measures of structure and function from the frontal cortex, hippocampus, and amygdala. Findings show that peripheral mCG can index CNS structural variability[113–120] associated with psychiatric illness development, and that locus-specific peripheral mCG *can* mediate ASE-associated CNS structure variability[120]. In these

chapters, I use genome-scale bioinformatics approaches in tandem with structural neuroimaging measures from the fronto-limbic pathway to investigate whether variability in peripheral tissue-derived mCG clusters represent pathways through which ASEs experienced in childhood and adulthood may become biologically embedded.

The second approach, as utilized in Chapter 4, uses secondary data analysis methods on a rodent model of EE to investigate epigenetic and transcriptomic mechanisms of biological embedding in the dentate gyrus of the hippocampus. In the original publication, researchers used high-resolution in vivo structural MRI to observe distinct increases in hippocampal volume in EE-exposed mice, in both dorsal (+8.5%) and ventral poles (+6.1%)[121]. In addition, EE was associated with a +60% increase in newborn neurons labeled with 5′-bromo-2′-dexoyuridine across the dorsal and ventral subregions of the dentate gyrus. In a separate cohort of animals, they used next-generation sequencing methods and identified altered mCG and mCH profiles in dentate gyrus subregions, which correlated with mRNA expression across discrete regions of the genome. Although the aforementioned research effectively investigated the effects of EE on CNS methylation and transcription, their methods focused on *genomic-region* level effects, i.e. differentially methylated regions characterized by shared methylation patterns across large stretches of DNA sequence. They were also focused on investigating the molecular differences between the dorsal and ventral subregions of the dentate gyrus. However, their research did not clearly investigate the *locus-specific* effects of EE on DNA methylation, or the potential mediating effects of DNA methylation on the relationship *between* EE and differential mRNA transcription, as done in Chapter 4.

1.6 Summary

Overall, this work uses novel genome-scale bioinformatics approaches to investigate potential underlying epigenetic mechanisms associated with, and/or potentially responsible for, the biological embedding of external environments. It provides novel insights to the fields of stress, trauma, and psychiatric illness by providing network-level and locus-specific peripheral biomarkers of ASEs, peripheral epigenetic profiles associated with neural endophenotypes of psychiatric illness, in addition to potentially mechanistic epigenetic mediators of EE-related biological embedding in the hippocampal dentate gyrus in mouse.

CHAPTER 2: ASSOCIATIONS BETWEEN CHILDHOOD FAMILY EMOTIONAL HEALTH, FRONTO-LIMBIC GREY MATTER VOLUME, AND SALIVA 5MC IN YOUNG ADULTHOOD[1]

2.1 Abstract

Poor family emotional health (FEH) during childhood is prevalent and impactful, and likely confers similar neurodevelopmental risks as other adverse social environments (ASEs). Pointed FEH study efforts are underdeveloped, and the mechanisms by which poor FEH are biologically embedded are unclear. The current exploratory study examined whether variability in 5-methyl-cytosine (5mC) and fronto-limbic grey matter volume may represent pathways through which FEH may become biologically embedded.

In 98 university students aged 18-22 years, retrospective self-reported childhood FEH was associated with right hemisphere hippocampus (b=10.4, p=0.005), left hemisphere amygdala (b=5.3, p=0.009), and right hemisphere amygdala (b=5.8, p=0.016) volumes. After pre-processing and filtering to 5mC probes correlated between saliva and brain, analyses showed that childhood FEH was associated with 49 5mC principal components (module eigengenes; MEs) (prange=3x10-6 to 0.047). Saliva-derived 5mC MEs partially mediated the association between FEH and right hippocampal volume (Burlywood ME indirect effect b=-111, p=0.014), and fully mediated the FEH and right amygdala volume relationship (Pink4 ME indirect effect b=-48, p=0.026). Modules were enriched with probes falling in genes with immune, central nervous system (CNS), cellular development/differentiation, and metabolic functions.

[1] This chapter is available at BMC Clinical Epigenetics (10.1186/s13148-021-01056-y).

Findings extend work highlighting neurodevelopmental variability associated with ASE exposure during childhood by specifically implicating poor FEH, while informing a mechanism of biological embedding. FEH-associated epigenetic signatures could function as proxies of altered fronto-limbic grey matter volume associated with poor childhood FEH and inform further investigation into primarily affected tissues such as endocrine, immune, and CNS cell types.

2.2 Introduction

Children can be exposed to an array of adverse social environments (ASEs) throughout their development, such as stressful life events (SLEs) and trauma. In addition, growing up in a low socioeconomic status (SES) household, or a household in which a caregiver is diagnosed with psychopathology, conveys risk towards potentially adverse exposures. Caregiver psychopathology is of particular interest to the current research due to its wide-reaching effects throughout the family unit, and its prevalence in the United States; it is estimated that ~12.8 million parents suffer yearly from some form of mental illness (18.2%), and that ~2.7 million parents suffer yearly from a *serious* mental illness (3.8%)[8]. The psychological effects of living with caregivers with a mental illness can be notably deleterious. Children of caregivers with major depressive disorder (MDD), for example, are subject to elevated risk of more hostile, negative, and withdrawn parenting[122]. Estimates range from two to 13 times increased risk for children to develop either their caregiver's mental illness or a mental illness different from their caregiver's[123]. Children growing up in these conditions are also more likely to develop internalizing or externalizing behavioral problems, as well as social, cognitive, and academic difficulties[14, 15]. An extension of caregiver psychopathology exposure is family emotional

health (FEH). Importantly, the mechanisms by which poor FEH are biologically embedded and produce these adverse outcomes are unclear.

The neuroimmune network hypothesis is one framework used to explain the physiological mechanisms via which ASEs and caregiver mental illness affect the mental health of offspring. The neuroimmune network hypothesis focuses on the integrated, bi-directional network of the central nervous system (CNS) and the immune system[58]. It posits that exposure to ASEs during childhood, an especially plastic window of development[59], impacts communication between peripheral inflammatory signals and fronto-limbic brain regions (i.e. prefrontal cortex, hippocampus, and amygdala). These brain regions support threat, reward, executive control, memory, and adaptive behavioral/emotional responses, among others[53]. Importantly, these functions are impaired in numerous affective and stress-related mental illnesses[61–64]. These inflammatory signals disrupt the inter-dependent functions of the fronto-limbic pathways, leading to altered behavioral states, and the pre-disposition to develop aberrant stress responses later in life[60]. These concepts are supported by a significant body of research that has shown immune system[32–34], hypothalamic-pituitary-adrenal (HPA)-axis[30, 35], and fronto-limbic pathway[35–37, 41, 42] associations with ASE exposure. More specifically, researchers have shown that childhood exposure to factors similar to poor FEH, such as maternal support and supportive/hostile parenting, are associated with lower hippocampus and amygdala grey matter volume later in life[40, 45]. Observed in association with ASE exposures, the signatures of morphometric variability within the fronto-limbic pathway are regarded as neural correlates of these exposures[35–37, 40–42, 45], and as neural endophenotypes of psychiatric illness[65–67].

The molecular mechanisms by which ASEs, including caregiver mental illness, become biologically embedded in the CNS are currently under investigation[124], and research has pointed to the importance of epigenetics, particularly cytosine-phosphate-guanine (CpG) 5-methyl-cytosine (5mC) levels, in this process[31, 125]. 5mC serves as a mediator of gene by environment interaction[102, 104, 126, 127], but it remains challenging to measure epigenetics in the living human brain -- the primary etiologic tissue of interest in regards to mental health-related outcomes. This limitation has prompted investigation into epigenetic measures collected from peripheral tissue, such as saliva, which may serve as proxies for etiological tissue. Previous studies have provided a framework for the use of peripheral tissues in epigenome-wide association studies (EWAS) and support the potential use of peripheral 5mC as a proxy for etiological tissue 5mC[128]. Further bolstering the notion that peripheral 5mC is an efficacious proxy for etiological tissue 5mC, is research showing that peripheral epigenetic measures can index changes in the HPA-axis[105, 106], immune system[107, 108], and the CNS[109–111]. However, these relationships do not directly indicate association between peripheral epigenetic measures and CNS-relevant endophenotypes of psychopathology. On this note, studies have used human structural and functional neuroimaging data in tandem with epigenetic measures but have primarily utilized candidate gene approaches. Measuring peripheral 5mC of the *SLC6A4*[113–115], *NR3C1*[116, 117], *FKBP5*[118], and *SKA2*[119, 120] genes, these studies have investigated associations between peripheral 5mC and variability in the structure and function of the frontal cortex, hippocampus, and amygdala. Findings suggest that locus-specific peripheral 5mC can index CNS structural alterations[113–120], and may statistically mediate ASE-induced CNS structural alterations[120].

Despite the evidence that peripheral 5mC can index CNS-related phenotypes, to date few studies, to our knowledge, have examined these relations in a hemisphere-specific manner within the brain. Importantly, numerous aspects of human behavior and biology are subject to hemisphere-specific brain lateralization[129–131]. This, coupled with evidence of hemisphere-specific fronto-limbic variability in association with ASEs in humans[35–37, 40–42, 45], provides a solid framework to address the potential associations of poor FEH with *hemisphere-specific* volume measurements. Beyond the aforementioned reports, studies of poor FEH or caregiver mental illness on CNS structure are sparse and limited to biological offspring of parents with genetically heritable psychopathology, although some investigate associations of exposure with outcome on a hemisphere-specific basis[132, 133]. These types of ASEs are also associated with changes in cell type-specific and tissue-specific 5mC[134], further highlighting the potential biologic embedding of these adverse exposures.

However, to our knowledge, investigations into the role of poor FEH in association with neural endophenotypes of psychopathology development have yet to be reported, and therefore, the magnitude of risk associated with poor childhood FEH has not been elucidated. In addition, investigations into the potential epigenetic mechanisms explaining the biological embedding of poor FEH have yet to be carried out.

To address these gaps in the field, the current exploratory study applied genome-scale approaches to assess whether saliva-derived 5mC measurements might index CNS endophenotypes of psychopathology in a sample of 98 young adult volunteers. We were specifically interested whether saliva-derived 5mC principal components (module eigengenes; MEs) might statistically mediate the relationship between poor FEH and hemisphere-specific fronto-limbic grey matter volume, while controlling for age, sex, cellular heterogeneity, genomic

ancestry, past year SLEs, and total brain volume (TBV). Such a result may serve as a peripheral proxy of such CNS variability, while informing a *potential* biological mechanism of physiological embedding. Based on previous work, we hypothesized that identified 5mC modules would be enriched with 5mC probes falling in genes with HPA-axis, immune system, and CNS-relevant gene ontology (GO) functions.

2.3 Materials and Methods

Participants

The current study draws on data from 98 university students (19.8±1.2 years old; 69% women; 49% white) (mean±SD) who successfully completed the Duke Neurogenetics Study (DNS). The DNS aims to assess the associations among a wide range of behavioral, neural, and genetic variables in a large sample of young adults, with one of the core goals being to establish a link between these various phenotypes and psychopathology**[37, 135–137]**. Data from a subset of participants with overlapping demographic, psychosocial, epigenetic, and neuroimaging data were included in the current cohort. This study was approved by the Duke University Medical Center Institutional Review Board, and all experiments were performed in accordance to its guidelines. Prior to the study, all participants provided informed consent. To be eligible for the DNS, all participants were free of: 1) medical diagnoses of cancer, stroke, head injury with loss of consciousness, untreated migraine headaches, diabetes requiring insulin treatment, chronic kidney or liver disease, or lifetime history of psychotic symptoms; 2) use of psychotropic, glucocorticoid, or hypolipidemic medication; and 3) conditions affecting cerebral blood flow and metabolism (e.g., hypertension)[136]. Neither past nor present diagnosis of Diagnostic and Statistical Manual for Mental Disorders, Fourth Edition (DSM-IV) Axis I or select Axis II

disorders (borderline and antisocial personality disorders) were exclusion criterion because the DNS seeks to establish broad variability in multiple behavioral phenotypes related to psychopathology[138]. Categorical diagnoses were assessed by trained staff using the electronic Mini International Neuropsychiatric Interview[139] and Structured Clinical Interview for the DSM-IV subtests[140]. Of the total sample reported here, 20 participants (~20%) met criteria for at least one lifetime DSM-IV diagnosis (Supplementary Table A.1).

Family emotional health (FEH)

Participants were asked to complete the Family History Questionnaire (FHQ), which produced the current study's measure of FEH. The FHQ is composed *fully* of questions from previously validated inventories[141–147]; of 70 total FHQ questions, 55 were included from the Family History Screen (FHS)[141, 142]. Researchers assessing psychometric properties of the FHS observed specificity in the range of 76.0% (MDD) to 97.1% (suicide attempt), and sensitivity in the range of 31.7% (alcohol dependence) to 60.0% (conduct disorder)[142]. The FHQ and FHS both capture family-wide psychiatric illness, but the FHQ is more encompassing of other ASEs, including 15 questions pertaining to cognitive decline of family members[143], externalizing behaviors[144], exposure to smoking[145], and drug/alcohol abuse treatment[146, 147]. Example questions from the FHQ include, "Has anyone in your family ever felt sad, blue, or depressed for most of the time for two weeks or more?", "Has anyone had several attacks of extreme fear or panic, even though there was nothing to be afraid of?", and "Has anyone in your immediate family ever tried kill to himself or herself?". The summed responses from 70 "yes/no" questions based on the aforementioned topics from the FHQ represent the current study's

measure of FEH (Supplementary Table A.2). Each "no" response corresponded to an additional score of one, with lower values representing poor FEH.

Cumulative perceived impact of past year stressful life events (past year SLEs)

Participants were administered an inventory measuring the cumulative perceived impact of SLEs from the past year ("past year SLEs"). Prior research reported associations between stress exposure and significant variability in fronto-limbic brain region volumes (BRVs) [35, 36, 41]. Therefore, throughout the current study, we controlled for the effect of past years SLEs using a summation of 45 negatively valenced items[135, 148] from the Life Events Scale for Students (LESS)[149].

Neuroimaging

ASEs and exposures similar to poor FEH are known to impact fronto-limbic pathways in the human CNS[35–37, 40–42, 45]. In addition, a meta-analysis has shown that both the hippocampus and amygdala have hemisphere-specific volume differences in healthy adults[150], and ASEs are known to have hemisphere-specific effects on fronto-limbic brain regions[40, 151]. Due to the dearth of research surrounding the potential effects of 5mC and ASEs (independently or in causal pathway models) on fronto-limbic brain region volume variability, and in order to avoid omission of hemisphere-specific effects by taking the mean of hemisphere volumes, hemisphere-specific amygdala, hippocampus, dorso-lateral pre-frontal cortex (dlPFC), and medial pre-frontal cortex (mPFC) volume measures were analyzed. Volume measurements of dlPFC and mPFC were chosen as outcome variables from the frontal cortex due to the opposing nature of their afferent and efferent projections to hippocampus and amygdala, and

their functional relationships with each region[53]. To compare differences between hemisphere volumes of mPFC, dlPFC, amygdala, and hippocampus, we performed paired sample t-tests of each brain region.

Data were collected at the Duke-UNC Brain Imaging and Analysis Center using one of two identical research-dedicated GE MR750 3T scanners (General Electric Healthcare, Little Chalfont, United Kingdom) equipped with high-power high-duty cycle 50-mT/m gradients at 200 T/m/s slew rate, and an eight-channel head coil for parallel imaging at high bandwidth up to 1 MHz. T1-weighted images were obtained using a 3D Ax FSPGR BRAVO with the following parameters: TR = 8.148 ms; TE = 3.22 ms; 162 axial slices; flip angle, 12°; FOV, 240 mm; matrix = 256 × 256; slice thickness = 1 mm with no gap; and total scan time = 4 min and 13 seconds. To generate regional measures of brain morphometry, anatomical images for each subject were first skull-stripped[152], then submitted to Freesurfer's (Version 5.3) "recon-all" with the "noskullstrip" option[153, 154], using an x86_64 linux cluster. CT and SA for 31 regions in each hemisphere, as defined by the Desikan-Killiany-Tourville atlas[155], were extracted using Freesurfer. Additionally, gray matter volumes from eight subcortical regions (including hippocampus and amygdala) were extracted with Freesurfer's subcortical segmentation ("aseg") pipeline[156], along with estimated TBV. The gray and white matter boundaries determined by recon-all were visually inspected using FreeSurfer QA Tools and determined to be sufficiently accurate for all subjects.

Molecular

Saliva was collected from participants using the Oragene-DNA OG-500 kit (Oragene; Ottawa, Canada). DNA was extracted and cleaned using the DNA Genotek prepIT PT-L2P kit

(DNA Genotek Inc; Ottawa, Canada) using manufacturer recommended methods. Purity of extracted DNA samples was assessed by absorbance using Nanodrop 1000 spectrophotometer (Thermo Fisher Scientific Inc; Waltham, Massachusetts). The quantity of double-stranded DNA was assessed using Quant-iT PicoGreen dsDNA kits with manufacturer recommended protocols (Invitrogen; Carlsbad, California). A total of 500 ng of genomic DNA was bisulfite-converted (BSC) using manufacturer-recommended EZ DNAm kits (Zymo Research; Irvine, California). After conversion, BSC DNA was applied to the Infinium MethylationEPIC BeadChip (Illumina; San Diego, California) (850k) to measure 5mC at ~850k loci.

5mC pre-processing

Beta-values measured from the 850k platform were background corrected in GenomeStudio, quality controlled (QCed), and filtered according to previously published methods[157]. All quality control and pre-processing was performed in R, version 3.6.1[158]. These steps removed 112,307 low quality and potentially cross-hybridizing probes, quantile-normalized probe beta-values, and removed potentially confounding technical and batch effects. 5mC beta-values were variance stabilized and logit-transformed into M-values[159]. 15,063 X-chromosome, Y-chromosome, and *rs*-mapped probes were removed to focus the analysis on genomic loci common between both biological sexes. The remaining ~739k probes were then subset to include only those with observed significant Pearson correlation ($p<0.05$) between saliva and brain tissue from the ImageCpG data repository[160]. This was done to focus the analysis on loci with greater prospect for proxy or surrogate status with etiologically relevant CNS tissue. Probes removed during QC, pre-processing, and subsetting were not analyzed. Afterwards, 62,422 probes remained for following analyses.

Cellular heterogeneity

Cell heterogeneity was estimated using a reference-free deconvolution method[161, 162]. Briefly, the top 15k most variable CpG sites were selected from the pre-processed/quality controlled 850k data and used to estimate the number of cell types and generate a matrix containing the proportions. Based on these methods, the number of cell types was set at five. Estimated proportions were used as covariates in relevant analyses to account for cellular heterogeneity.

Genomic ancestry

To avoid potential inaccuracies and confounding effects of self-reported race/ethnicity, genetic ancestry was modeled using multi-dimensional scaling (MDS) measures extracted from participant genomic data using PLINK[163]. Using previously collected genome-wide association study (GWAS) data from the DNS, the first four MDS genetic ancestry measures were calculated and used as covariates across pertinent models based on visual inspection of scree plots. This methodology is in line with previous publications[136].

Probe clustering

To remove non-desired effects, we fit linear models with age, validated biological sex, cellular heterogeneity, and genomic ancestry as predictors of probe-wise 5mC M-value. For each probe, residual values ("residualized M-values") were extracted for clustering. Taking the 62,422 residualized M-values, the "WGCNA" R package was used to build a co-methylation network[164]. First, scale-free topology model fit was analyzed. As recommended, a soft-

threshold value of four was chosen based on the lowest power at which adjusted R^2>0.90. Adjacency and dissimilarity matrices were generated, and unsupervised hierarchical clustering was used to generate a clustered, residual M-value network. Setting a minimum cluster size of ten generated 194 modules, each identified by a unique color. The first principal component (ME) of each module was then calculated. Compared to EWAS, which assess differential methylation on the level of individual 5mC loci, network-based methods, as used in the current research, utilize dimension-reduction techniques to create a much smaller network of related 5mC probe clusters. This reduces the burden of multiple hypothesis testing, and provides the potential for increased statistical power in circumstances with a small number of biological replicates[165].

Statistical analyses

In order to understand the relationships between variables, we computed Pearson correlations and mapped their correlation coefficients. Based on these correlations, we conducted a set of analyses, as shown in Figure 2.1. In Arm A analyses, FEH was used as a predictor of hemisphere-specific BRVs, while including age, biological sex, four genomic ancestry MDS measures, past year SLEs, and TBV as covariates. In Arm B analyses, FEH was used as a predictor of ME values, while including past year SLEs as a covariate. Age, sex, and genomic ancestry effects were accounted for previously by using residualized M-values as input for clustering. In Arm C analyses, ME values were used as individual predictors of BRV, while including the same covariates as in Arm A. Throughout the current research, past year SLEs were included as a covariate because our FEH measure only captures SLEs from childhood, and recent stress exposure is associated with variability in our outcome variables[35, 36, 40, 166].

TBV was included as a covariate but, where pertinent, non-TBV controlled model results are also reported. Within each phase of the analyses, non-standardized continuous measures were used resulting in non-standardized effect estimates. Due to the exploratory nature of the current study, dependent variables were graduated to subsequent study arms if $p<0.05$; sequential Benjamini-Hochberg (BH)-significant results at false discovery rate (FDR) = 0.10 are also reported where applicable [167], and all results with $p<0.05$ were considered for interpretation. Briefly, for each p-value, a BH critical value was calculated where the p-value's assigned rank over the number of tests was multiplied by the accepted FDR. P-values less than this threshold were deemed BH-significant.

Mediation analyses

To investigate whether the effect of poor FEH on hemisphere-specific BRV is *statistically* mediated via peripheral 5mC signatures, MEs were tested for mediating status between FEH and hemisphere-specific BRVs using the "mediation" package in R[168] (Figure 2.1). Importantly, only hemisphere-specific BRVs associated with FEH (Figure 2.1, Arm A) were considered. Similarly, MEs tested for mediation included *only those* associated with both FEH (Figure 2.1, Arm B) and hemisphere-specific BRV (Figure 2.1, Arm C). Mediation model inputs were assembled per recommended "mediation" package protocol. Therefore, Arm A (plus ME as a covariate) and Arm B models were used as inputs. For each ME, indirect effects (IDE), direct effects (DE), and total effects (TE) were calculated as a result of 10,000 non-parametric bootstrap simulations. Consistent with published methods[40], we considered an ME a full mediator if the DE=0 while the IDE and TE$\neq$ 0, or a partial mediator if the DE, IDE, and TE$\neq$ 0. Individual probes from full mediator modules were assessed for mediation status as well.

Gene set enrichment analysis

To assess the underlying methylomic network enrichment of the ~62,000 brain-saliva correlated probes, individual residualized probe M-values were used as predictors of FEH in Bayesian regression models. Age, sex, genomic ancestry measures, cell heterogeneity measures, and past year SLEs were included as covariates. From this analysis, BH-significant probe p-values were extracted and used as input to gene set enrichment analysis (GSEA) in the "methylGSA" package[169]. GO sets composed of 50 to 1,000 genes were allowed, which eliminated high-level GO-terms such as "biological process" and facilitated testing of 3,186 GO sets. To produce a condensed summary of non-redundant GO-terms, the web-based tool "Revigo" was used[170].

2.4 Results

Study participants

Descriptive statistics for demographic, psychosocial, and neuroimaging variables in study participants are shown in Table 2.1. FEH ranged from 34 to 70; the mean in the study sample was 60 (+/-8.5). To compare differences between hemisphere volumes of mPFC, dlPFC, amygdala, and hippocampus we performed paired-sample t-tests of each brain region. We found significant differences between left and right hemisphere volume in dlPFC ($mean_{Left}$ = 11980 mm3, $mean_{Right}$ = 10639 mm3, t = 9.1, p = 1x10-14), amygdala ($mean_{Left}$ = 1655 mm3, $mean_{Right}$ = 1859 mm3, t = -12.5, p = 0.005), and hippocampus ($mean_{Left}$ = 4650 mm3, $mean_{Right}$ = 4741 mm3, t = -2.9, p = 2x10-16), but no difference in mPFC (p > 0.05).

Correlation analyses

Pearson correlations between variables used in the current study were mapped (Figure 2.2). Of note, a moderate negative association was observed between FEH and past year SLEs (Pearson's correlation: r=-0.44, p=7x10^{-6}).

FEH predicts hemisphere-specific BRV

FEH was positively associated with right hippocampus (b=10.4, SE=3.6, t=2.9, p=0.005), left amygdala (b=5.3, SE=2.0, t=2.7, p=0.009), and right amygdala volumes (b=5.8, SE=2.3, t=2.4, p=0.016). These significant relationships were also observed in models without controlling for the covarying effect of TBV (right hippocampus p=0.015; left amygdala p=0.018; right amygdala p=0.023). FEH was not associated with left hippocampus (p=0.62), left dlPFC (p=0.10), right dlPFC (p=0.62), left mPFC (p=0.98), or right mPFC volume (p=0.09). In controlling for seventy-two tests at FDR=0.10, all three brain regions with p<0.05 were BH-significant (Table 2.2). Regions associated with FEH were carried into following analyses.

FEH predicts ME values

FEH was associated with 49 MEs (b_{min}=-0.006, b_{max}=0.006, p_{min}=3x10^{-6}, p_{max}=0.047). Twenty-nine out of 49 MEs achieved BH-significance, including the Burlywood and Pink4 MEs, taking 194 tests into account at FDR=0.10 (Supplementary Table A.3).

ME values predict hemisphere-specific BRV

Forty-nine MEs associated with FEH were tested for association with right hippocampus, left amygdala, and right amygdala volumes (Supplementary Table A.4). Seven MEs were

associated with right hippocampus volume, four of which were BH-significant: Burlywood (b=874.2, SE=252.1, t=3.5, $p=8x10^{-4}$) (Figure 2.3a), Darkolivegreen1 (b=770.0, SE=258.6, t=3.0, p=0.004), Thistle2 (b=728.1, SE=261.8, t=2.8, p=0.007), and Chocolate2 (b=-713.3, SE=259.3, t=-2.8, p=0.007). The Darkgray ME (b=-374.2, SE=140.1, t=-2.7, p=0.009) (Figure 2.3b) was negatively associated with left amygdala volume, in addition to the Darkolivegreen ME (b=-300.6, SE=142.2, t=-2.1, p=0.037). The Lavenderblush2 ME was positively associated with left amygdala volume (b=295.1, SE=144.5, t=2.0, p=0.044). Finally, the Pink4 ME (b=467.5, SE=165.8, t=2.8, p=0.006) (Figure 2.3c) was positively associated with right amygdala volume. In controlling for 49 tests within each of the three BRVs at FDR=0.10, only the aforementioned MEs associated with right hippocampus volume were BH-significant.

ME mediation

Eleven MEs were tested for mediation between FEH and BRVs**.** The Burlywood ME was a partial *statistical* mediator between FEH and right hippocampus volume (b_{TE}=-366, $p=8x10^{-4}$; b_{IDE}=-111, p=0.014; b_{DE}=-254, p=0.037). The TE indicated that right hippocampal volume was 366 mm^3 less under poor FEH conditions compared to high FEH conditions, while the IDE of the Burlywood ME was accountable for 111 mm^3 (30%) of that effect. Without controlling for the covarying effect of TBV, the Burlywood ME was a full mediator (b_{TE}=-376, p=0.006; b_{IDE}=-114, p=0.031; b_{DE}=-261, p=0.071). The Darkolivegreen1 (b_{TE}=-369, p=0.001; b_{IDE}=-66, p=0.042; b_{DE}=-303, p=0.008) and Thistle2 (b_{TE}=-373, p=0.002; b_{IDE}=-64, p=0.025; b_{DE}=-309, p=0.010) MEs were also partial statistical mediators of the FEH and right hippocampus volume relationship. The Thistle ME was also a partial mediator in analyses without controlling for TBV (b_{TE}=-382, p=0.007; b_{IDE}=-85, p=0.017; b_{DE}=-297, p=0.042). On the other hand, the Chocolate2,

Cornflowerblue, Aliceblue, and Yellow MEs were neither partial nor full mediators of the relationship ($p_{TE}<0.05$; $p_{IDE} > 0.05$; $p_{DE}<0.05$). Six of the 11 Burlywood probes are mapped to known genes (*MSH2, ATXN7L1, ODF2, SLC22A6, TGFB3,* and *DYX1C1*) with GO-terms including GO:0005245 voltage-gated calcium channel activity, GO:0002700 regulation of production of molecular mediator of immune response, and GO:0043524 negative regulation of neuron apoptotic process. GO-terms associated with probes from the Darkolivegreen1 and Thistle2 modules include GO:0001829 trophectodermal cell differentiation, GO:0045087 innate immune response, GO:0042552 myelination, and GO:0010506 regulation of autophagy, among others[171, 172].

None of the Darkgray (b_{TE}=-183, p=0.014; b_{IDE}=-47, p=0.095; b_{DE}=-135, p=0.094), Darkolivegreen (b_{TE}=-185, p=0.013; b_{IDE}=-32, p=0.205; b_{DE}=-153, p=0.057), or Lavenderblush2 (b_{TE}=-187, p=0.011; b_{IDE}=-30, p=0.181; b_{DE}=-156, p=0.044) MEs were mediators of the relationship between FEH and left amygdala volume. However, the significant TE values indicated ~185 mm^3 lower left amygdala volume in poor FEH conditions. Regarding FEH and right amygdala volume, Pink4 ME value was a full statistical mediator of the relationship (b_{TE}=-204, p=0.017; b_{IDE}=-48, p=0.026; b_{DE}=-156, p=0.069), indicating that right amygdala volume was 204 mm^3 less in poor FEH conditions than in high FEH conditions. Results additionally indicate that Pink4 ME value accounted for 48 mm^3 (24%) of the aforementioned effect. Pink4 was also a full mediator in the non-TBV controlled model (b_{TE}=-208, p=0.017; b_{IDE}=-52, p=0.025; b_{DE}=-157, p=0.087). Pink4 is composed of 21 probes mostly mapped to known genes (*SNORD123, TBCD, FN3K, NRXN3, GLB1L2, SBF2, PSMB1, SYT1, BEST2, TBATA,* and *GNA12*). GO-terms associated with mapped genes include GO:0048487 beta-tubulin binding, GO:0007158 neuron cell-cell adhesion, GO:0007212 dopamine receptor signaling pathway,

GO:0010762 regulation of fibroblast migration, GO:0019905 syntaxin binding, and GO:0031683 G-protein beta/gamma-subunit complex binding, among others[171, 172]. Probe-specific genomic biology annotation for partial and full mediating modules can be found in Supplementary Table A.5. In controlling for 33 tests at FDR=0.10, all significant ME IDEs, DEs, and TEs were BH-significant (Table 2.3). Mediation analyses were then performed on individual probes from the Pink4 module in order to assess locus-specific effects.

Probe-wise mediation

Three out of 21 probes from the Pink4 module were full mediators between FEH and right amygdala volume: cg22325292 (b_{TE}=-204, p=0.013; b_{IDE}=-53, p=0.018; b_{DE}=-151, p=0.087), cg02398342 (b_{TE}=-204, p=0.014; b_{IDE}=-44, p=0.038; b_{DE}=-161, p=0.060), and cg00809820 (b_{TE}=-205, p=0.013; b_{IDE}=-48, p=0.049; b_{DE}=-157, p=0.064). These three probes also had extremely high Pearson correlation values with the Pink4 ME (r>0.93, p<$2x10^{-44}$), indicating that they are strong representatives of the Pink4 ME. In controlling for 63 tests at FDR=0.10, all significant probe IDEs, DEs, and TEs were BH-significant (Supplementary Table A.6).

Gene set enrichment analysis

We performed GSEA using probe M-values as predictors of FEH and used resultant p-values to facilitate the testing of 3,186 GO-terms. After redundancy reduction, 45 BH-significant GO-terms remained for interpretation. CNS-related GO-terms included: beta-amyloid clearance (GO:0097242, p=$8x10^{-11}$, rank=2), filopodia assembly (GO:0046847, p=$2x10^{-8}$, rank=5), catecholamine metabolic process (GO:0006584, p=$4x10^{-5}$, rank=11), and positive regulation of

neuron apoptotic process (GO:0043525, p=0.013, rank=25) among others. Although immune-related terms were limited, one was present in the top three: cytokine receptor activity (GO:0004896, $p=8x10^{-11}$, rank=3). Numerous metabolic functions were identified: negative regulation of stress-activated MAPK cascade (GO:0032873, $p=4x10^{-12}$, rank=1), NAD metabolic process (GO:0019674, $p=4x10^{-9}$, rank=4), and TOR signaling (GO:0031929, $p=3x10^{-4}$, rank=12) among others. A complete list of BH-significant GO-terms can be found in Supplementary Table A.7.

2.5 Discussion

The current exploratory study examined whether variability in 5mC and fronto-limbic grey matter volume represent pathways through which FEH becomes biologically embedded. Based on previous work, we hypothesized that 5mC modules would be enriched for immune system[32–34], HPA-axis[30, 35], and CNS-relevant[35–37, 41, 42] GO-terms. Our study findings indicated that exposure to poor FEH during childhood was associated with CNS endophenotypes of psychiatric illness, and that a subset of saliva-derived 5mC measurements *statistically* mediated this relationship. Additionally, we found the mediating 5mC modules were enriched with probes associated with GO-terms relating to the CNS and immune system, as well as cellular differentiation, regulation, specialization, and organization (mostly relating to neuronal cell populations). Finally, we found that the underlying FEH-associated methylomic network was enriched with CNS-related, immune system, and metabolic GO-terms. Overall, we posit that the FEH-associated epigenetic signatures could function as proxies of altered fronto-limbic grey matter volume associated with poor childhood FEH; peripheral epigenetic signatures indexing our relationships of interest may be explained by peripheral inflammation related to

development of stress-related psychopathology, thereby supporting the neuroimmune network hypothesis[58].

The relationships observed between poor childhood FEH and left/right amygdala volume in the current study mirrored relationships observed throughout the literature regarding direction of effect and magnitude, but not hemisphere-specificity[35, 36]. Studies show hemisphere-specific effects of ASEs on amygdala volume, with stressors exerting notable statistical effects on left but not right amygdala volume. In one such prospective longitudinal study, SLEs negatively predicted left, but not right, amygdala volume in children with low to average polygenic risk scores. They showed that children exposed to the highest level of SLEs had ~9% less left amygdala volume than those exposed to the lowest levels of SLEs[35]. A more recent study showed lower left amygdala volumes in children who had experienced early neglect, low SES, or physical abuse compared to non-exposed controls[36]. Although we observed bilateral amygdala grey matter volume associations with poor childhood FEH exposure, our study did show a similar magnitude of effect; poor childhood FEH exposure was explanatory (DE) of -8.9% difference in left and -8.4% difference in right amygdala volume. The peripheral 5mC signature (Pink4 ME) mediating right amygdala volume and FEH accounted for -2.5% of additional volumetric difference (IDE).

Similar to our amygdala-related findings, the reported relationship between poor childhood FEH and low hippocampus volume supports previous findings from the field regarding direction and estimated magnitude of effect, but not hemisphere-specificity[40, 45]. In a prospective longitudinal study, researchers focused on childhood “maternal support” as their exposure of interest, finding that maternal support of children, three to five years old, was associated with increased hippocampal volume in both hemispheres later in childhood (seven to

thirteen years old). Specifically, they found that children exposed to low maternal support during that time span had a difference in hippocampal volume of -7.1%[45]. This magnitude closely mirrors the findings of the current study, which show poor childhood FEH has a DE that explains -6.1% difference in right hippocampal volume, and peripheral 5mC signatures have an IDE responsible for an additional -1.7% of difference. A more recent study from the same group found that the positive association between SES and hippocampal volume was mediated by "supportive/hostile parenting" in both hemispheres, but only by SLEs in left hippocampus[40]. These studies identified significant associations of maternal support and supportive/hostile parenting in *both* hippocampal hemispheres, whereas the current study identified a significant association only in right hippocampus.

No salient effects of FEH were observed in dlPFC or mPFC, in either hemisphere. This finding does not support research showing deleterious effects of ASEs on frontal cortex morphometry[48–50]. Our findings across fronto-limbic brain regions imply that poor childhood FEH has specific morphometric associations with variability in subcortical structures responsible for memory, avoidance, fear, stress, and negative valence, but not with variability in cortical structures managing those functions.

Beyond the observed associations between poor childhood FEH and fronto-limbic brain morphometry, we were interested in the peripheral epigenetic signatures that index the relationships, and that provide a potential mechanism of biological embedding of ASEs. The Pink4 module, which fully mediated the relationship between poor childhood FEH and right amygdala volume in both TBV-controlled and non TBV-controlled models, is composed of 21 probes, three of which were full mediators of the FEH and right amygdala volume relationship: cg22325292, cg02398342, and cg00809820. Probes cg22325292 and cg02398342 exist in the

sixth of six exons of the *FN3K* gene and fall in a putative CpG island and DNaseI hypersensitive region ~1,000 base pairs upstream of the *TBCD* transcription start site (TSS)[171, 172]. The main *TBCD* protein isomer plays a major role in the assembly of microtubules[173], the cell-cycle progression to mitosis[174], and neuronal morphogenesis[175]. Hypermethylation of the *TBCD* gene in CD4+ T-cells is also associated with rheumatoid arthritis[176], an autoimmune disorder associated with stress exposure[177]. Additionally, cg02398342 falls in the transcription factor binding site of the *EGR1* protein, which has integral, dynamic interactions with genes responsible for vesicular release and endocytosis, neurotransmitter metabolism and receptors, and actin cytoskeleton organization[178]. These interactions facilitate *EGR1*'s significant impact on synaptic and neuronal activation. Module-wide and probe-specific results suggest that the association of poor childhood FEH with right amygdala volume is indexed and *statistically* mediated by peripheral epigenetic signatures relevant to synapse development, neurotransmitter signal transduction, and cytoskeleton organization.

Three modules were partial mediators of the relationship between right hippocampus volume and poor childhood FEH: Burlywood, Darkolivegreen1, and Thistle2. However, the Burlywood module was a full mediator in the non TBV-controlled mediation model, implying that this peripheral epigenetic signature exerts more statistical effect on absolute right hippocampus volume, agnostic of TBV, through poor childhood FEH. GO-terms associated with the hippocampus mediating modules include GO:0002700 regulation of production of molecular mediator of immune response, GO:0043524 negative regulation of neuron apoptotic process, GO:0042552 myelination, and GO:0010506 regulation of autophagy, among others[171, 172]. Module-wide and probe-specific results imply that the associations of poor childhood FEH with

right hippocampal volume are indexed by peripheral epigenetics signatures related to immune response and CNS cell development/lifecycle.

The top three GO-terms from our methylome network analysis were: 1. negative regulation of stress-activated MAPK cascade 2. beta-amyloid clearance 3. cytokine receptor activity. The MAPK cascade has long been established as a key driver of eukaryotic signal transduction, but more recently as an integral contributor to cell proliferation, differentiation, and inflammatory processes[179]. There is also a building body of evidence suggesting a significant role of the MAPK cascade in mental health outcomes. In a mouse model, modulation of the MAPK cascade in the forebrain is associated with both anxiety-like and depressive-like behaviors[180]. When p38 MAPK protein is selectively knocked out (KO) of the dorsal raphe nucleus, rodents subjected to social defeat stress show significantly reduced social avoidance compared to wild-type animals[181]. Additionally, pro-inflammatory cytokine administration induces a state of increased serotonergic CNS activity (canonically thought to be depleted in MDD), and induction towards that state is blocked with p38 MAPK inhibition[182]. In humans, MDD is a common co-morbidity of rheumatoid arthritis (RA)[183]; peripheral inflammation is a hallmark of RA and is also observed in MDD patients[184]. Therefore, it is hypothesized that within the context of psychopathology development, environmental stressors induce peripheral cytokine signaling that communicates with fronto-limbic brain regions including the amygdala, hippocampus, and frontal cortex through mechanisms including the MAPK cascade[185]. To this end, numerous RA and anti-depressant drugs are observed to reduce canonical disease symptoms, while also reducing clinical inflammation markers and MAPK signaling[185].

Based on the current study results, it appears that variability in peripheral 5mC and fronto-limbic grey matter volume *potentially* represent pathways through which FEH becomes

biologically embedded, with 5mC signatures that mediate the relation between FEH and grey matter volume being especially enriched with GO-terms related to the peripheral inflammatory sequela of stress-related psychopathology development. To our knowledge, the degree to which peripheral 5mC serves as a statistical mediator between poor childhood FEH (or ASEs in general) and variable fronto-limbic brain morphometry had not been previously elucidated. In addition, the observed GO-terms support potential mechanisms of biological embedding that are actively being considered in the field[180–182, 185, 186].

Dimension reduction techniques used throughout our research represent the foremost strengths of this study. These methods focus the analysis onto loci with greater prospect for proxy or surrogate status with etiologically relevant CNS tissue, and reduce the burden of multiple hypothesis testing. Clustering similarly methylated probes creates a relatively small number of modules which potentially contain probes from functionally related genes. On the other hand, limitations of the current study that may reduce the generalizability of the results include retrospective self-report of the main exposure of interest, a relatively small sample size, lack of replication in an independent cohort, the balance of biological sex within the cohort, the cohort enrichment of higher SES and FEH participants, the inability to correct for smoking-related effects, and the interpretation of nominally significant ($p<0.05$) results instead of more rigorous multiple hypothesis testing corrected results. Additionally, in analyzing grey matter volume of fronto-limbic brain regions as outcomes of interest, we have omitted surface area- or cortical thickness-specific effects. The current study also falls short in establishing whether the mediation by peripheral 5mC modules is causal in nature. Longitudinal data could provide more precise insight into whether such relationships exist. Future studies on this topic should capture

longitudinal data from a diverse, increased sample size and could investigate genetic factors or tissues of etiological interest.

2.6 Conclusion

The current study showed that, in support of prior literature, exposure to poor childhood FEH is associated with low fronto-limbic BRV as measured in young adulthood. Newly reported here is the finding that saliva-derived 5mC modules mediate the FEH and BRV relationship and are enriched for immune system, CNS-related, and CNS cell development/specialization functions; with additional validation in independent cohorts, these 5mC modules could potentially be used as peripheral biomarkers of poor FEH exposure during childhood. Overall, the findings of the current study support the neuroimmune network hypothesis[58], extend the body of work highlighting neurodevelopmental variability associated with childhood ASE exposure, and inform a potential molecular mechanism of biologic embedding. Future research on these peripheral signatures could validate their use as proxies/biomarkers of perturbed underlying neurobiology in response to poor FEH exposure and could inform further investigation into primarily effected tissue such as endocrine, immune, and CNS cell types.

2.7 Figures and Tables

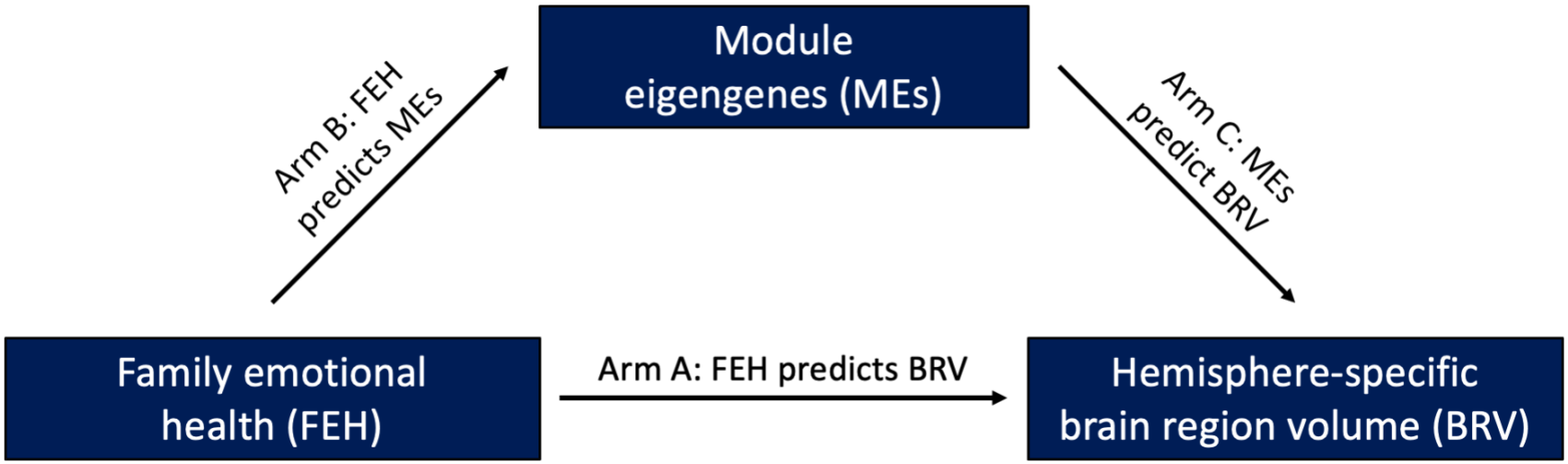


Figure 2.1 Conceptual model testing module eigengenes (MEs) as mediators of the hypothesized association between family emotional health (FEH) and variability in hemisphere-specific brain region volume (BRV). Arm A. FEH was used as a predictor of hemisphere-specific BRV, while including age, biological sex, four genomic ancestry MDS measures, past year SLEs, and total brain volume as covariates. Arm B. FEH was used as a predictor of ME value, while including past year SLEs as a covariate. The age, sex, and genomic ancestry effects on ME components were previously removed. Arm C. ME values were used as individual predictors of BRV, while including age, biological sex, four genomic ancestry MDS measures, past year SLEs, and total brain volume as covariates.

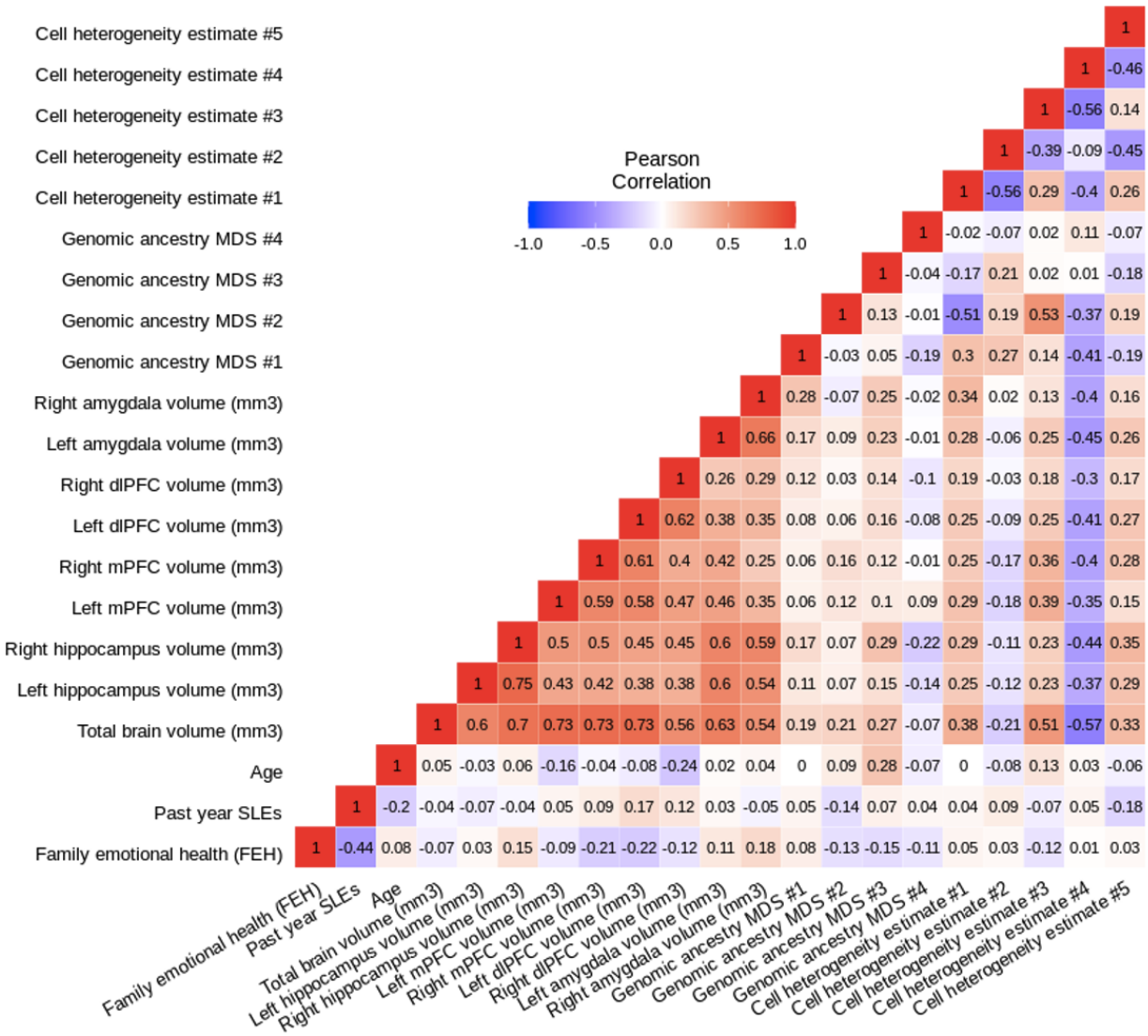


Figure 2.2 Pearson correlation heat map of variables used throughout the current analyses. A strong negative relationship was observed between FEH and past year SLEs (Pearson's correlation: r=-0.44, p=7x10-6). Strong positive relationships are also observed between hemisphere-specific brain regions (Pearson's correlation r range: 0.26 – 0.75).

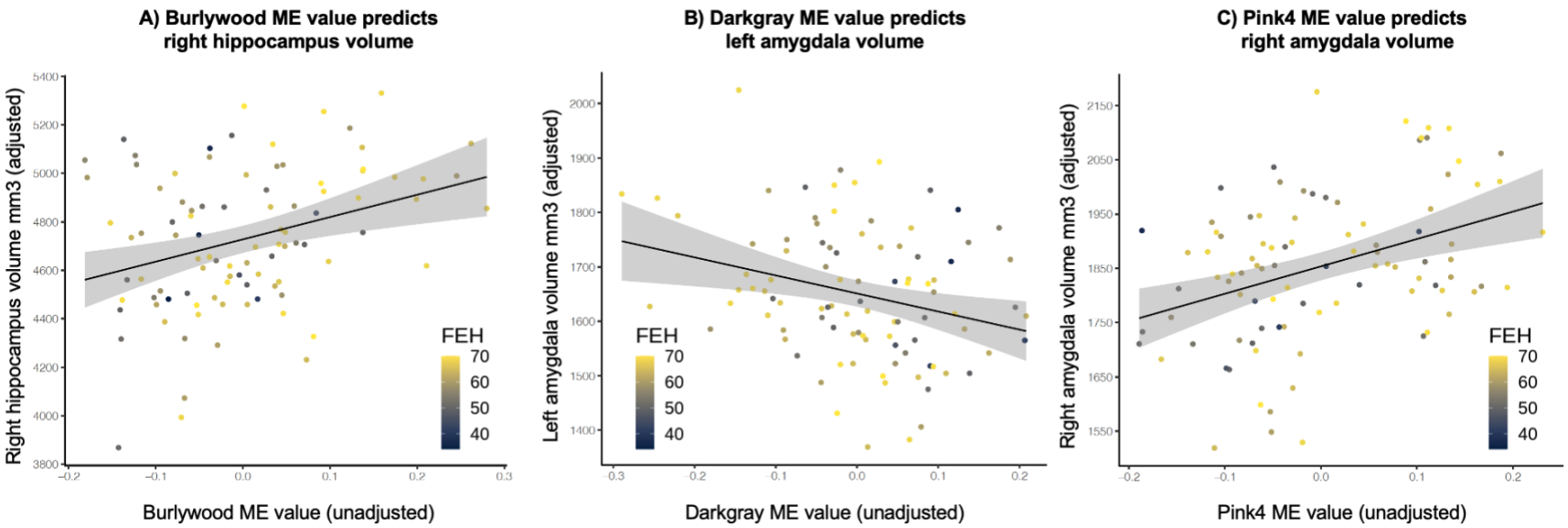


Figure 2.3 ME values are associated with high right hippocampal, low left amygdala, and high right amygdala volume. BRVs values shown are adjusted by covariates. Covariates across all models: age, sex, four genomic ancestry measures, past year SLEs, and total brain volume. The line of best fit (via least squares) is shown with a grey 95% SE confidence range. A. High Burlywood ME value is associated with high right hippocampal volume (b=874.2, SE=252.1, t=3.5, p=0.0008). B. High Darkgray ME value is associated with low left amygdala volume (b=- 374.2, SE=140.1, t=-2.7, p=0.009). C. High Pink4 ME value is associated with high right amygdala volume (b=467.5, SE=165.8, t=2.8, p=0.006).

Table 2.1. Demographic, psychosocial, and neuroimaging summary stats for the current sample (n = 98)

Characteristic	Description	Value
Age	Mean [SD] (Range)	19.8 [1.2] (18 - 22)
Sex	Male	31%
	Female	69%
Self-reported race/ethnicity	Caucasian/White	49%
	African American/Black	6%
	Asian	22%
	American Indian	13%
	Bi- or multiracial	0%
	Other	9%
Cumulative perceived impact of past year SLEs	Mean [SD] (Range)	11 [7.6] (1 - 42)
Family emotional health (FEH)	Mean [SD] (Range)	60 [8.5] (34 - 70)
Left hippocampus volume (mm3)	Mean [SD] (Range)	4650 [446] (3267 - 5679)
Right hippocampus volume (mm3)	Mean [SD] (Range)	4741 [392] (3825 - 6031)
Left dlPFC volume (mm3)	Mean [SD] (Range)	11980 [1785] (7308 - 16007)
Right dlPFC volume (mm3)	Mean [SD] (Range)	10639 [1585] (7426 - 14819)
Left mPFC volume (mm3)	Mean [SD] (Range)	5745 [853] (3871 - 8686)
Right mPFC volume (mm3)	Mean [SD] (Range)	5790 [717] (4238 - 7442)
Left amygdala volume (mm3)	Mean [SD] (Range)	1655 [186] (1292 - 2157)
Right amygdala volume (mm3)	Mean [SD] (Range)	1859 [210] (1371 - 2417)
Total brain volume (mm3)	Mean [SD] (Range)	$1.2x10^6$ [$1.2x10^5$] ($9.0x10^5$-$1.5x10^{6)}$

SLE: stressful life event, FEH: family emotional health, dlPFC: dorso-lateral pre-frontal cortex, mPFC (medial PFC)

Table 2.2 Family emotional health (FEH) predicts hemisphere-specific brain region volume (BRV) (n = 98)

L. hipp vol. (mm3)	**adj. R2**	**RSE**	**b**	**SE**	**R. hipp vol. (mm3)**	**adj. R2**	**RSE**	**b**	**SE**
Model	0.325***	369.2			Model	0.532***	255.9		
FEH			2.6	5.2	FEH			**10.4****	**3.6**
Sex (female)			-60.6	114.9	Sex (female)			-88.5	79.7
Age			-32.1	33.2	Age			-8.4	23.0
Cumulative impact of past year SLE			-2.5	5.9	Cumulative impact of past year SLE			4.1	4.1
Total brain volume (mm3)			**0.002*****	**0.0005**	Total brain volume (mm3)			**0.002*****	**0.0003**
L. dlPFC vol. (mm3)	**adj. R2**	**RSE**	**b**	**SE**	**R. dlPFC vol. (mm3)**	**adj. R2**	**RSE**	**b**	**SE**
Model	0.564***	1182.0			Model	0.344**	1254.0		
FEH			-27.9	16.7	FEH			-8.8	17.7
Sex (female)			366.8	367.9	Sex (female)			219.2	390.3
Age			-82.5	106.3	Age			**-323.2****	**112.8**
Cumulative impact of past year SLE			28.3	18.9	Cumulative impact of past year SLE			10.9	20.0
Total brain volume (mm3)			**0.013*****	**0.0015**	Total brain volume (mm3)			**0.008*****	**0.0016**
L. mPFC vol. (mm3)	**adj. R2**	**RSE**	**b**	**SE**	**R. mPFC vol. (mm3)**	**adj. R2**	**RSE**	**b**	**SE**
Model	0.559***	557.6			Model	0.539***	479.1		
FEH			0.2	7.9	FEH			-11.6	6.8
Sex (female)			159.9	173.6	Sex (female)			53.1	149.1
Age			-102.9*	50.2	Age			-11.6	43.1
Cumulative impact of past year SLE			4.8	8.9	Cumulative impact of past year SLE			5.7	7.7
Total brain volume (mm3)			**0.006*****	**0.0007**	Total brain volume (mm3)			**0.005*****	**0.0006**
L. amygdala vol. (mm3)	**adj. R2**	**RSE**	**b**	**SE**	**R. amygdala vol. (mm3)**	**adj. R2**	**RSE**	**b**	**SE**
Model	0.404***	139.2			Model	0.356***	166.2		
FEH			**5.2****	**2.0**	FEH			**5.8***	**2.3**
Sex (female)			-34.5	43.3	Sex (female)			-34.3	51.7
Age			-3.6	12.5	Age			-5.1	15.0
Cumulative impact of past year SLE			3.5	2.2	Cumulative impact of past year SLE			0.7	2.7
Total brain volume (mm3)			**0.001*****	**0.0002**	Total brain volume (mm3)			**0.001*****	**0.0002**

*** P < 0.001, ** P < 0.01, * P < 0.05, **P: BH-significant (bolded)**

Four genomic ancestry multi-dimensional scaling measures were included in each model, but are not shown here.

RSE: relative standard error, adj: adjusted, b: estimate, SE: standard error, L: left, R: right, hipp: hippocampus, vol: volume

Table 2.3 Module eigengenes (MEs) mediating observed family emotional health (FEH) and brain region volume (BRV) relationships (n = 98)

Burlywood: Right hippocampus	b	95% CI Lower	95% CI Upper
Average indirect effect (IDE)	**-111.2***	-255.5	-17.9
Average direct effect (DE)	**-254.4***	-528.2	-15.8
Average total effect (TE)	**-365.7*****	-641.0	-151.0
Darkolivegreen1: Right hippocampus	b	95% CI Lower	95% CI Upper
Average indirect effect (IDE)	**-65.6***	-161.3	-1.6
Average direct effect (DE)	**-302.9****	-553.6	-84.3
Average total effect (TE)	**-368.5****	-623.6	-152.7
Thistle2: Right hippocampus	b	95% CI Lower	95% CI Upper
Average indirect effect (IDE)	**-64.4***	-156.6	-5.6
Average direct effect (DE)	**-308.5***	-567.7	-79.2
Average total effect (TE)	**-372.9****	-640.1	-150.5
Chocolate2: Right hippocampus	b	95% CI Lower	95% CI Upper
Average indirect effect (IDE)	-82.8	-204.7	4.9
Average direct effect (DE)	**-281.5***	-585.5	-24.2
Average total effect (TE)	**-364.3****	-635.0	-144.5
Cornflowerblue: Right hippocampus	b	95% CI Lower	95% CI Upper
Average indirect effect (IDE)	-48.3	-132.5	14.3
Average direct effect (DE)	**-321.5****	-606.7	-96.5
Average total effect (TE)	**-369.8****	-641.6	-151.8
Aliceblue: Right hippocampus	b	95% CI Lower	95% CI Upper
Average indirect effect (IDE)	-53.4	-146.6	6.6
Average direct effect (DE)	**-317.6****	-591.1	-98.8
Average total effect (TE)	**-371.1*****	-642.8	-162.5
Yellow: Right hippocampus	b	95% CI Lower	95% CI Upper
Average indirect effect (IDE)	-39.7	-115.2	14.7
Average direct effect (DE)	**-332.2****	-610.1	-113.6
Average total effect (TE)	**-371.9****	-642.0	-158.7
Darkgray: Left amygdala	b	95% CI Lower	95% CI Upper
Average indirect effect (IDE)	-47.4	-115.8	8.9
Average direct effect (DE)	-135.2	-293.1	22.0
Average total effect (TE)	**-182.5***	-324.2	-38.6
Darkolivegreen: Left amygdala	b	95% CI Lower	95% CI Upper
Average indirect effect (IDE)	-32.1	-89.7	21.8
Average direct effect (DE)	-152.5	-310.8	4.6
Average total effect (TE)	**-184.6***	-323.9	-39.9
Lavenderblush2: Left amygdala	b	95% CI Lower	95% CI Upper
Average indirect effect (IDE)	-30.4	-93.4	13.7
Average direct effect (DE)	**-156.4***	-301.0	-3.0
Average total effect (TE)	**-186.9***	-324.4	-44.0
Pink4: Right amygdala	b	95% CI Lower	95% CI Upper
Average indirect effect (IDE)	**-47.9***	-117.4	-3.7
Average direct effect (DE)	-156.1	-338.9	12.9
Average total effect (TE)	**-204.1***	-380.1	-44.2

*** P < 0.001, ** P < 0.01, * P < 0.05, **P: BH-significant (bolded)**

ME: module eigenvalue, b: estimate, CI: confidence interval

CHAPTER 3: BLOOD 5MC CLUSTERS ASSOCIATED WITH ADVERSE SOCIAL EXPOSURES AND ENDOPHENOTYPES OF PTSD-RELATED PSYCHIATRIC ILLNESS IN A TRAUMA-EXPOSED COHORT OF WOMEN

3.1 Abstract

Adverse social exposures (ASEs) such as low income, low educational attainment, and childhood/adult trauma exposure are associated with variability in fronto-limbic grey matter volume, surface area (SA), and cortical thickness (CT) measurements. These central nervous system (CNS) morphometries, in turn, are associated with mood-related psychiatric illnesses and represent endophenotypes of mood-related psychiatric illness development. In African American (AA) women, ASEs and the development of mood-related psychiatric illness are prevalent. Epigenetic mechanisms, such as 5-methyl-cytosine (5mC), may contribute to the biological embedding of the environment in these populations but are understudied and not well understood. How these epigenetic mechanisms relate to CNS endophenotypes of psychiatric illness is also unclear.

In 97 mostly AA, all female, trauma-exposed participants from the Grady Trauma Project, we examined the associations of childhood trauma burden (CTQ), adult trauma burden, low income, and low education with blood-derived 5mC clusters and fronto-limbic grey matter measurements. Specifically, we assessed whether 5mC clusters statistically mediated the relationships between ASEs and fronto-limbic grey matter measurements, in order to elucidate whether peripheral 5mC would index CNS endophenotypes of psychopathology.

CTQ was negatively associated with rostral middle frontal gyrus (RMFG) SA (β = -0.231, p = 0.041). Low income and low education were also associated with a number of fronto-

limbic grey matter measurements. Seven 5mC clusters were associated with CTQ (pmin = 0.002), two with low education (pmin = 0.010), and three with low income (pmin = 0.007). Two clusters were full mediators of the relation between CTQ and RMFG SA, accounting for 47% and 35% of variability respectively. These clusters were enriched for probes falling in DNA regulatory regions, as well as signal transduction and immune signaling gene ontology functions.

Our results support the body of work highlighting fronto-limbic variability associated with ASEs, while informing a peripheral inflammation-based epigenetic mechanism of biological embedding of such exposures. These signatures may be proxies of perturbed underlying neurobiology in response to childhood trauma burden and inform further investigation into primarily affected tissue such as endocrine, immune, and CNS cell types. These findings could also serve to potentiate increased investigation of understudied populations at significant risk for PTSD-related psychiatric illness development, such as AA women in low income, high trauma environments.

3.2 Introduction

Adverse social exposures (ASEs) such as low income, low educational attainment, childhood trauma, and adult trauma exposure (TE), are common in the United States (US). Over the past five years, poverty rates in the US have ranged from 12 to 15%[6], and are likely to increase due to the impact of the ongoing COVID-19 pandemic[187]. The percentage of US adults aged 18 to 24 that finish high school with a diploma or GED is roughly 88%, meaning that over one million students per year leave high school without completion credentials[7]. Past year exposure to physical abuse in US children range from 4% to 16%, whereas past year psychological abuse is reported in 10% of children in the US; sexual abuse is experienced at

some point during childhood in 15% to 30% of children[1]. Furthermore, in a US population-based study of 34,653 participants, researchers found that depending on self-reported race/ethnicity, respondents' prevalence of lifetime TE varied between 66% and 84%[4]. Other researcher has observed lifetime TE prevalence ranging from 50% to 90%[2, 3, 5].

To this end, ASEs carry with them both acute and chronic risks across biological and psychological domains; one salient risk is the development of PTSD-related psychiatric illness or symptoms of such illnesses. To this end, in *both* adults and children, poverty-related stress is associated with increased internalizing symptoms (e.g., sadness or anxiety), involuntary engagement stress responses (e.g., emotional arousal or intrusive thoughts), externalizing symptoms (e.g., acts of aggression or theft), and involuntary disengagement stress responses (e.g., emotional numbing or cognitive interference)[13]. In addition to these associations between poverty-related stress and symptoms of potential psychiatric illness, rates of anxiety, mood, and substance-use disorders are all positively correlated with measures of socioeconomic disadvantage[188]. In the context of this international World Health Organization study, socioeconomic disadvantage included *both* low income and low educational attainment[188]. Moving a step beyond correlation, a longitudinal study found that higher educational attainment significantly attenuated both the overall risk of adult onset major depressive disorder (MDD), particularly in women, and the conditional risk of MDD as well[189].

The effects of childhood maltreatment or trauma on long term mental health prognoses are also well-documented. In a recent meta-analysis of over 190 studies and 68,000 participants, researchers observed that higher childhood maltreatment (including potential traumas such as emotional abuse, physical abuse, sexual abuse) scores were associated with MDD diagnosis later in life and with depression symptom severity[24]. This built upon previous work in the field

showing that exposure to any form of maltreatment during childhood was associated with a two-fold risk of MDD development, and nearly a three-fold risk of generalized anxiety disorder (GAD) development in adulthood[25]. Although childhood and adolescence are regarded as particularly vulnerable periods of development, the effects of trauma on psychiatric illness/symptom development are also notable when trauma is experienced in adulthood. A particularly salient example of this is evident in the potential development of post-traumatic stress disorder (PTSD) after a trauma. Interestingly, recent population-based work highlights race/ethnicity-based differences in psychiatric illness development and PTSD in particular. Researchers found that African Americans (AAs) are *generally* more resilient to developing psychiatric illness after a TE, compared to other races/ethnicities[190]; however, this resilience does not apply to PTSD, where AAs experience elevated conditional PTSD development rates relative to other races/ethnicities[190]. This, in tandem with research showing that females are subject to almost double the risk of PTSD/MDD development throughout their lifetime[2, 191–193], highlights that AA women in particular may be subject to elevated risk of PTSD-related psychiatric illness development.

A great deal of work has been placed in identifying the neural correlates of poverty, low education, and childhood/adult TEs. Researchers have identified specific signatures of fronto-limbic central nervous system (CNS) variability in association with ASEs using methods such as structural magnetic resonance imaging (MRI) and functional MRI (fMRI). In rodent[194, 195] and human[35–39] studies, observed neural correlates of ASEs include variability in volumetric, cortical thickness (CT), surface area (SA), excitability, and connectivity measures. Fronto-limbic brain regions, including frontal cortex, hippocampus, and amygdala, contribute to adaptive behavioral and emotional responses, and are relevant to psychiatric illness development due to

their contributions to the generation and regulation of major decision-making, memory formation, and stress-reactivity processes[53]. To this end, a meta-analytic study showed that frontal pole (FP), superior frontal gyrus (SFG), and rostral middle frontal gyrus (RMFG) frontal cortex subregions are heavily involved in attention, working memory, cognitive flexibility, and executive reasoning[196], with added functionality of the frontal pole to introspection and longitudinal thinking[197]. The anterior cingulate cortex (ACC) is regarded as a mediator of the interactions between higher level cognitive processes and emotion/stress reactivity[198], and the orbito-frontal cortex (OFC) contributes to decision making and emotional processing functions[199]. These functions are impaired in numerous mental illnesses, including but not limited to PTSD[61], MDD[62], GAD[63], bipolar disorder[64], and schizophrenia[64]. These neural correlates of ASEs are associated with the development of PTSD-related psychiatric illnesses, and the replicated associations of fronto-limbic variability with psychiatric illness represent neural endophenotypes of said disorders[65, 67, 200].

Importantly though, the effects of ASEs do not travel directly from environment to psychiatric illness. The sequalae of effect, rather, travels initially through the neuroendocrine system as a response to stressful external stimuli[201], whereby cortisol signaling occurs throughout the body, engaging immune (and other) networks[202]. Aberrant immune network activation and responses to cortisol signaling result in neural growth, inflammation, metabolism, and stress-related pathway disruption through molecular and epigenetic mechanisms[57, 203]. The process of DNA methylation, which commonly refers to the addition of a 5-methyl-cytosine (5mC) residue to a cytosine-phosphate-guanine (CpG) base-pair, is one such epigenetic mechanism, and is understood to enhance or reduce mRNA transcription of genes in an experience-dependent, and temporally-stable manner[204]; continued research has pointed to the

importance of 5mC in this biological embedding of external stimuli[31, 125]. Because of challenges in taking epigenetic measurements from a living human brain, the primary etiologic tissue of interest in regard to psychiatric illness development, peripheral tissue such as blood or saliva serve as proxies for etiological tissue. It is well-documented that these peripheral epigenetic measures can index changes to the hypothalamic-pituitary-adrenal (HPA)-axis[105, 106], immune system[107, 108], and CNS[109–111], although the direction of effect may be discordant across tissues[110, 112]. The aforementioned aberrant physiological responses and their accompanying epigenetic alterations are thought to confer the biological embedding of ASEs[205, 206], and can ultimately increase the risk of PTSD-related psychiatric illness[203, 207]. Peripheral epigenetic measures can also index CNS-relevant endophenotypes of psychiatric illness development, as observed in studies using neuroimaging and peripheral epigenetic measures *in tandem.* However, these studied have primarily utilized candidate gene approaches, investigating 5mC in the *SLC6A4*[113–115], *NR3C1*[116, 117], *FKBP5*[118], and *SKA2*[119, 120] genes, and neuroimaging measures of structure and function from the frontal cortex, hippocampus, and amygdala. Findings show that peripheral 5mC can index CNS structural variability[113–120] associated with psychiatric illness development, and that locus-specific peripheral 5mC can mediate ASE-associated CNS structure variability[120].

To date, however, thorough investigation into the relationships between ASEs, peripheral 5mC, and CNS endophenotypes of psychiatric illness have not been carried out at the genome-scale. To this end, the current exploratory study applied genome-scale approaches in an all-female, majority AA cohort from the Grady Trauma Project (GTP) to assess whether blood-derived 5mC measurements might index CNS endophenotypes of PTSD-related psychiatric illness. We focus specifically on the GTP because AA women may be subject to elevated risk of

PTSD-related psychiatric illness development, but the mechanisms underlying that elevated risk are still unclear. In this study, we were specifically interested whether clusters of peripheral 5mC measurements *statistically* mediate the relation between ASEs and fronto-limbic grey matter volume, SA, and CT measures. Based on previous work, we hypothesized that identified 5mC modules would be enriched with 5mC probes falling in genes with HPA-axis[105, 106], immune system[107, 108], and CNS-relevant [109–111] gene ontology (GO) functions.

3.3 Materials and Methods

Participants

The current research draws on 97 female, trauma-exposed participants drawn from the GTP. Participants were part of a larger investigation of genetic and environmental factors that predict the response to stressful life events (SLEs) in a predominantly AA, urban population of low socio-economic status (SES)[191]. Research participants were approached in the waiting rooms of primary care clinics of a large, public hospital. After the subjects provided written informed consent, they participated in a verbal interview and gave a blood sample for genetic and epigenetic analyses. Participants in this cohort also had available demographic, psychosocial (Table 3.1), and neuroimaging data. Exclusion criteria included mental retardation or active psychosis. All GTP procedures were approved by the Institutional Review Boards of Emory University School of Medicine and Grady Memorial Hospital.

Exposures of interest

Income, education, adult trauma burden, and childhood trauma burden were considered exposures of interest in current analyses. *Income:* Self-reported approximate household monthly

income (“household income”) was measured as part of a demographic-focused inventory. Participants indicated total household income within one of the following ranges: $0 to $499, $500 to $999, or greater than $1000 per month. *Education:* Participants also reported educational attainment history (“education”) and were binned into one of the following categories: did not complete high school, completed high school or GED, or completed more schooling than high school. *Traumatic events inventory:* Adult trauma burden was assessed via a semi-structured interview using the Traumatic Events Inventory (TEI), a scale developed by the GTP researchers during their prior work in the Grady hospital primary care population[208]. TEI responses pertaining to childhood were excluded in favor of the Childhood Trauma Questionnaire (CTQ). In the current study, the continuous TEI score was used (“TEI”); the higher the TEI score, the more traumatic life events one encountered. *Childhood trauma questionnaire:* Childhood trauma burden was assessed using the self-report 28 question CTQ[209]. This inventory consists of five subscales, including sexual abuse, physical abuse, emotional abuse, as well as physical and emotional neglect. In the current study, the continuous CTQ score was used (“CTQ total”). Information regarding previous usage of exposure of interest variables in GTP studies can be found elsewhere[31, 191, 208, 210–212].

Neuroimaging

In the parent study/project, hemisphere-specific frontal cortex CT and SA measurements were taken along with amygdala and hippocampal gray matter volume to assess brain morphometry of the fronto-limbic pathway. For the purpose of this study, hippocampus and amygdala measures, we used an average of the volumes taken across hemispheres (e.g., “hemisphere-mean hippocampal volume”, “hemisphere-mean amygdala volume”). CT and SA

develop along discordant temporal trajectories[213], and are effected in different manners based on underlying genetic sequence[214]. Based on this rationale, we examined CT *and* SA measures of frontal cortex subregions, instead of gray matter volume. Frontal cortex subregions included in the current study map to Brodmann's areas 8, 9, 10, 11, 24, 25, 32, and 46[53], and include FP, medial OFC (mOFC), lateral OFC (lOFC), SFG, RMFG, rostral ACC (rACC), and caudal ACC (cACC). For frontal cortex SA and CT, we used an average of values taken across hemispheres (e.g., "hemisphere-mean RMFG SA", "hemisphere-mean RMFG CT"). Structural MRI study procedures followed methods as previously described in the GTP cohort [215, 216].

Molecular

Whole-blood was collected in EDTA vacuum tubes prior to DNA extraction. Genomic DNA was extracted and purified using the AllPrep DNA/RNA Mini Kit (Qiagen, Valencia, CA) using manufacturer recommended methods. DNA was denatured and bisulfite converted (BSC) using the EZ DNA Methylation-GoldTM Kit (Zymo Research, Irvine, CA). After conversion, BSC DNA was applied to the Infinium MethylationEPIC BeadChip (Illumina; San Diego, California) (850k) using manufacturer recommended protocol to measure DNA 5mC at ~850,000 loci.

5mC pre-processing

Beta-values measured from the 850k platform were background corrected with GenomeStudio, quality controlled (QCed), and filtered according to previously published methods[157]. All quality control and pre-processing was performed in R version 3.6.1[158]. These steps removed low quality and potentially cross-hybridizing probes, quantile-normalized

probe Beta-values, and removed technical and batch effects[217–220]. 5mC beta-values were variance stabilized and logit-transformed into M-values[159]. X- and Y-chromosome-mapped probes were removed, along with *rs*-mapped probes. The remaining ~827k probes were then subset to include only those with observed nominally significant Pearson correlation ($p < 0.05$) between blood and brain tissue from the ImageCpG data repository[160], which focused analysis onto loci with greater prospect for proxy or surrogate status with etiologically-relevant CNS tissue. Following QC and filtering steps, 92,208 probes remained.

Covariates

Age, sex, cellular heterogeneity estimates, multi-dimensional scaling (MDS) of genomic ancestry, an epigenetic proxy of smoking[221], and employment/disability status were included in models as covariates, where pertinent. *Cellular heterogeneity:* Genome-wide DNAm patterns were used to estimate the proportions of granulocytes, monocytes, B cells, natural killer cells, CD4T and CD8T cells within each sample. Calculations were done according to methodology described in Houseman et al [222]. Estimated proportions of monocytes, B cells, natural killer cells, CD4T and CD8T were included as covariates in multiple regression models, where pertinent. *Genomic ancestry:* To avoid potential inaccuracies/confounding effects of self-reported race/ethnicity, genetic ancestry was modeled using MDS measures extracted from participant genomic data using PLINK[163]. The first four MDS genetic ancestry measures were chosen to be used as covariates based on visual inspection of scree plots. This methodology is in line with previous publications[223]. *Smoking:* Cigarette smoking is known to exert significant effects across the methylome[224]. Using a method discussed in detail elsewhere[225], the current study corrected for the methylomic effects of smoking by calculating effect size

estimates of the top twenty-six probes from a recent smoking epigenome-wide association study (EWAS) performed in participants with AA genomic ancestry[221]. *Employment and disability:* As applied in previous GTP publications[210], employment/disability status was considered as: unemployed; unemployed receiving disability support; or employed, with or without disability support.

Probe clustering

In order to remove non-desired effects from probe M-values, we composed linear models in R using age, sex, cellular heterogeneity, genomic ancestry, and the smoking proxy as predictors of probe-wise 5mC M-values. For each probe, residual values ("residualized M-values") were extracted for clustering. Taking the 92,208 residualized M-values, the "WGCNA" package was then used in R to build a co-methylation network[164]. First, scale-free topology model fit was analyzed. As recommended, a soft-threshold value of ten was chosen based on the lowest power at which adjusted $R^2 > 0.90$. Adjacency and dissimilarity matrices were generated, and unsupervised hierarchical clustering was used to generate a clustered residual M-value network. Setting a minimum cluster size of ten generated 73 clusters of 5mC probes, where each cluster was identified by a unique color. After clustering, we extracted the first principal component of each cluster, referred to from here on as a "module eigengene" (ME). Compared to EWAS-type analyses, which assess differential methylation on the level of individual 5mC loci, network-based methods, as used in the current research, utilize dimension-reduction techniques to create a much smaller network of related 5mC probe clusters. This reduces the burden of multiple hypothesis testing and provides the potential for increased statistical power in circumstances with a small number of biological replicates[165].

Statistical analyses

To understand the relationships between variables used throughout the current analyses, we carried out Pearson correlations and mapped their correlation coefficients. Based on these correlations, we conducted a set of analyses, as shown in Figure 3.1. Figure 3.1 Arm A analyses included hemisphere-mean hippocampus and amygdala volumes as dependent variables (in separate models). Income, education, CTQ total, and TEI were included as independent variables of interest, whereas genomic ancestry, age, employment/disability and intra-cranial volume (ICV) were included as covariates. The same predictors and covariates were included, minus ICV, to predict frontal cortex subregions SA and CT measures. Overall CT and/or SA were not included as covariates based on prior work with this cohort [226]. Figure 3.1 Arm B analyses included 5mC MEs as dependent variables (in separate models). Income, education, CTQ total, and TEI were included as independent variables of interest, and employment/disability was included as a covariate. Covariates were limited in this stage of analysis due to the removal of confounding effects through residual extraction in prior pre-processing steps. Figure 3.1 Arm C analyses included individual fronto-limbic morphometry measures as dependent variables (in separate models). Individual 5mC MEs were included as independent variables of interest, whereas age, genomic ancestry, employment, and ICV (ICV was only included in hippocampus and amygdala models) were included as covariates. Within each phase of the analyses, including mediation analyses, continuous dependent and independent variables were standardized, resulting in standardized effect estimates. Nominal p-values were corrected for multiple hypothesis testing by controlling the false discovery rate (FDR=0.10) using the Benjamini Hochberg (BH) procedure(85). Briefly, for each nominal p-value, a BH critical value was calculated where nominal p-value's assigned rank over the number of tests was multiplied by the

accepted FDR. Nominal p-values less than this threshold were deemed BH-significant. Due to the exploratory nature of the current work, both nominal and BH-significant terms were considered for interpretation; and nominal p-values are presented within the results section of the current text. We also highlight results from the same analyses as seen above, where we excluded the single non-AA participant.

Mediation analyses

MEs were tested for *statistical* mediating status between independent variables of interest (ASEs) and dependent variables (hemisphere-mean fronto-limbic brain morphometry) using the "mediation" package in R[168] (Figure 3.1). Importantly, only fronto-limbic brain morphometry measures nominally associated with income, education, CTQ total or TEI (Figure 3.1, Arm A) were considered. Similarly, MEs tested for mediation included *only those* nominally associated with an exposure (Figure 3.1, Arm B) and a fronto-limbic brain morphometry measure (Figure 3.1, Arm C). For each ME, average indirect effect (IDE), direct effect (DE), and total effect estimates and confidence intervals were calculated as a result of 10,000 quasi-Bayesian Monte Carlo approximations, applied with the mediation package. Consistent with methodology employed in the field[40, 227], we considered an ME a full mediator if the DE=0 while the IDE and total effect $\neq 0$, or a partial mediator if the DE, IDE, and total effect $\neq 0$. Individual probes from full mediator modules were assessed for mediation status, in order to gain insight into potential locus-specific mediation effects. Additionally, we mention results from the same analyses as seen above, where we excluded the single non-AA participant.

Gene set enrichment analyses

To assess underlying methylomic network enrichment, all four ASEs of interest were included in linear models as either independent variables or covariates, where individual residualized probe M-values were included as dependent variables, while controlling for employment/disability status. Importantly, only exposures with a fully-mediating 5mC cluster were considered independent variables of interest. Other exposures which did not exhibit a fully-mediating 5mC cluster were considered as covariates. For probes/ASEs with nominally significant relationships ($p < 0.05$), we extracted probe p-values and probe names and used them as input to gene set enrichment analyses (GSEA) with the "methylGSA" package[169]. Importantly, ASE-associated probe p-values were only captured for ASEs in which a full mediator was previously observed, which limited analyses to CTQ only. GO sets of 25 to 1,000 genes were allowed, eliminating high-level GO-terms such as "biological process", which facilitated testing of 5,478 gene sets. To produce a condensed summary of non-redundant GO-terms, the web-based tool "Revigo" was used[170]. Again, we mention results from GSEA where we excluded the single non-AA participant.

3.4 Results

Study participants

Descriptive statistics for demographic and psychosocial measures in study participants are shown in Table 3.1. Mean age of participants was 40 (+/-12.5), with 100% of the sample female and 99% AA. Thirty percent of the sample reported household income in the range of $0 to $499 per month, while 34% of the sample reported household income above $1000 per month. Fifteen percent of the sample reported not having finished high school, whereas 55% of the

sample reported having more schooling than high school. Mean CTQ total value for participants was 40.5 (+/- 15.4), whereas mean TEI value for participants was 4.2 (+/- 2.3). Pearson correlations between variables used in the current study are shown in Figure 3.2.

ASEs predict fronto-limbic brain morphometry

At nominal significance, CTQ total was negatively associated with RMFG SA (β = - 0.231, SE = 0.111, t = - 2.079, p = 0.041), but was not associated with other outcomes of interest ($p > 0.05$). Relative to high educational attainment (completed more schooling than high school), low educational attainment (no high school certificate) was negatively associated with cACC SA (β = - 0.855, SE = 0.280, t = - 3.049, p = 0.003), and positively associated with FP CT (β = 0.684, SE = 0.302, t = 2.267, p = 0.026) and SFG CT (β = 0.553, SE = 0.256, t = 2.162, p = 0.034) at the nominal significance threshold, but no other brain morphometry outcomes of interest ($p > 0.05$). On the other hand, relative to high household income ($1000+/month), low household income ($0-$499/month) was nominally negatively associated with amygdala grey matter volume (β = - 0.649, SE = 0.283, t = - 2.299, p = 0.024), but no other neuroimaging outcome of interest ($p > 0.05$). TEI was not associated with any neuroimaging measures ($p > 0.05$). Nominally significant model outcomes can be found in Table 3.2. In controlling for 64 total tests at FDR = 0.10, no outcomes were BH-significant. All of the nominally significant relationships mentioned above were maintained in post-hoc analyses including *only* participants (n = 96) that self-reported race as AA (data not shown). Brain morphometry outcomes nominally associated with ASEs were carried into downstream analyses.

ASEs predict 5mC ME

CTQ total was nominally associated with seven MEs, the strongest of which were Darkolivegreen (Darkolivegreen $\beta = 0.354$, SE = 0.110, $t = 3.214$, $p = 0.002$) and Steelblue (Steelblue $\beta = -0.306$, SE = 0.111, $t = -2.748$, $p = 0.007$). Compared to high educational attainment, low educational attainment was positively associated with one ME (Coral1 $\beta = 0.750$, SE = 0.286, $t = 2.618$, $p = 0.010$), and negatively associated with one ME (Lightcoral $\beta = -0.615$, SE = 0.296, $t = -2.079$, $p = 0.040$) at nominal levels. Relative to high household income ($1000+/month), low household income ($0-$499/month) was positively associated with three MEs (Antiquewhite4 $\beta = 0.770$, SE = 0.277, $t = 2.781$, $p = 0.007$; Indianred4 $\beta = 0.717$, SE = 0.276, $t = 2.602$, $p = 0.011$; Orange $\beta = 0.571$, SE = 0.274, $t = 2.084$, $p = 0.040$) at nominal levels. Because TEI was not associated with any fronto-limbic brain morphometric measures (Figure 3.1, Arm A), we do not report any results for this exposure. Results from all Arm B analyses can be found in Supplementary Table B.1. In controlling for 73 tests at FDR = 0.10, no outcomes were BH-significant. All nominally significant relationships reported above maintained their significance in analyses excluding the single non-AA participant. 5mC MEs nominally associated with ASEs were carried into downstream analyses.

5mC MEs predict fronto-limbic brain morphometry

Four out of seven CTQ-associated MEs showed a nominally significant relationship with hemisphere-mean RMFG SA (Maroon, White, Tan, Yellow). The three strongest relationships were the Maroon ME ($\beta = -0.451$, SE = 0.091, $t = -4.966$, $p = 3 \times 10^{-6}$) (Figure 3.3a), White ME ($\beta = -0.392$, SE = 0.093, $t = -4.231$, $p = 6 \times 10^{-5}$) (Figure 3.3b), and Tan ME ($\beta = 0.233$, SE = 0.100, $t = 2.322$, $p = 0.023$) (Figure 3.3c). Two low educational attainment-associated MEs were

tested independently for associations with hemisphere-mean cACC SA, frontal pole CT, and SFG CT. Of these, the Lightcoral ME was nominally associated with cACC SA ($\beta = 0.215$, SE = 0.099, $t = 2.182$, $p = 0.032$) (Figure 3.3d). Finally, of the three low income-associated MEs, none showed a nominally significant relationship with hemisphere-mean amygdala volume (Supplementary Table B.2). In controlling for 16 tests at FDR = 0.10, the Maroon and White ME relationships with hemisphere-mean RMFG SA were BH-significant. In assessing these same relationships after removing the non-AA participant, we found no differences from results reported above (data not shown); all nominally significant and BH-significant relationships remained. 5mC MEs nominally associated with fronto-limbic brain morphometry measures were carried into downstream analyses.

ME mediation

Four CTQ-associated MEs that also showed a nominally significant relationship with hemisphere-mean RMFG SA were tested for their mediating status between CTQ total and hemisphere-mean RMFG SA (Maroon, White, Tan, Yellow). Both the Maroon ME (TE $\beta = -0.983$, $p = 0.047$; IDE $\beta = -0.466$, $p = 0.008$; DE $\beta = -0.517$, $p = 0.277$) and the White ME (TE $\beta = -1.014$, $p = 0.034$; IDE $\beta = -0.348$, $p = 0.032$; DE $\beta = -0.667$, $p = 0.164$) were full mediators of this relationship. TE results show that RMFG SA was an estimated ~589 mm^2 lower at the highest levels of CTQ exposure compared to the lowest levels of CTQ exposure. Independently, the Maroon ME accounted for 275 mm^2 (47%) of that effect, whereas the White ME accounted for 204 mm^2 (34%). Importantly, the Maroon and White MEs were both full mediators in analyses excluding the non-AA participant (data not shown). Neither the Tan nor Yellow MEs were partial or full mediators of the relationship between CTQ and hemisphere-mean RMFG SA

(TE $p < 0.05$; IDE $p > 0.05$; DE $p > 0.05$), although the direction of TEs in these analyses closely mirrored those in the Maroon and White ME mediation analyses. The Lightcoral ME, which we previously observed to be associated with low education attainment and cACC SA, was neither a partial nor full mediator of the relationship between low education attainment and hemisphere-mean cACC SA (TE $\beta = -0.844$, $p = 0.003$; IDE $\beta = -0.107$, $p = 0.100$; DE $\beta = -0.737$, $p = 0.010$). However, the DE indicates that low educational attainment status accounted for cACC SA that is 84 mm^2 smaller than high educational attainment status. In controlling for ten tests of the IDE and DE p-values for the five relationships above at FDR = 0.10, all nominally significant ME IDEs were BH-significant and all DEs remained non-significant (Table 3.3). Follow-up mediation analyses were then performed on individual probes from the Maroon and White modules.

Probe-wise mediation: Maroon module

Six out of the 11 Maroon ME probes were full mediators of the relationship between CTQ total and RMFG SA (Table 3.4). The top three strongest mediators based on IDE β coefficients were: cg21622733 (TE $\beta = -0.983$, $p = 0.044$; IDE $\beta = -0.512$, $p = 0.003$; DE $\beta = -0.471$, $p = 0.320$), cg19805668 (TE $\beta = -1.003$, $p = 0.042$; IDE $\beta = -0.411$, $p = 0.009$; DE $\beta = -0.592$, $p = 0.217$), and cg03710029 (TE $\beta = -0.990$, $p = 0.048$; IDE $\beta = -0.321$, $p = 0.019$; DE $\beta = -0.670$, $p = 0.182$). Respectively, these probes mapped to the long intergenic non-coding (LINC) RNA gene *LINC01531*, the placental growth factor gene *PGF*, and the solute carrier family 38 member 10 gene *SLC38A10*. The other three full mediating probes mapped to myosin heavy chain 14 (*MYH14*), DNA methyltransferase 3 alpha (*DNMT3A*), and neuralized E3 ubiquitin protein ligase 4 (*NEURL4).* In controlling for 22 tests of Maroon probe IDE and DE p-values,

only two nominally significant probes were also BH-significant (cg21622733: *LINC01531* and cg19805668: *PGF*). Comprehensive probe-wise mediation results are found in Supplementary Table B.3.

In all, probes from the Maroon module map to eight known protein coding genes (*MYH14, HGSNAT, SFTPA1, PGF, NEURL4, DNMT3A, SLC38A10,* and *PEBP4*) and two non-coding RNAs (*LINC01531 and MIR3659HG*). GO-terms associated with these loci include: *MYH14* (neuronal action potential GO:0019228, actomyosin structure organization GO:0031032), *HGSNAT* (lysosomal transport GO:0007041, neutrophil degranulation GO:0043312), *SFTPA1* (toll-like receptor signaling pathway GO:0002224), *PGF* (chemoattractant activity GO:0042056, vascular endothelial growth factor receptor binding GO:0005172), *NEURL4* (ubiquitin protein ligase activity GO:0061630), *DNMT3A* (DNA (cytosine-5-methyltransferase activity GO:0003886), *SLC38A10* (amino acid transmembrane transporter activity GO:0015171), and *PEBP4* (protein binding GO:0005515), among others.

Probe-wise mediation: White module

On the other hand, three out of 19 White ME probes were full mediators of the relationship between CTQ total and RMFG SA (Table 3.4): cg01544227 (TE β = -0.994, p = 0.044; IDE β = -0.305, p = 0.036; DE β = -0.689, p = 0.175), cg00686169 (TE β = -1.017, p = 0.037; IDE β = -0.277, p = 0.043; DE β = -0.740, p = 0.126), and cg08002107 (TE β = -1.006, p = 0.040; IDE β = -0.268, p = 0.047; DE β = -0.738, p = 0.122). These three probes mapped to *AC092745.5*, *FAM90A24P*, and *USP17L7* loci, respectively. In controlling for 38 tests of White probe IDE and DE p-values, no nominally significant probes were BH-significant. In performing the same analyses as above without inclusion of the non-AA participant, we found that in the

Maroon module, cg15150970 did not maintain a nominally significant full mediator status (data not shown). In the White module, we found that zero out of the three nominally significant full mediators maintained this same status, although they each approached nominal significance ($p < 0.08$). Comprehensive probe-wise mediation results are found in Supplementary Table B.3.

In all, only one probe from the White module maps to a known protein coding gene, *USP17L7*. The GO term associated with *USP17L7* is thiol-dependent deubiquitinase (GO:0004843*).* One other probe maps to a LINC RNA, *FAM66D*. The remaining probes map to a pseudogene family, named family with sequence similarity 90. Members include: *FAM90A10*, *FAM90A11P*, *FAM90A20*, *FAM90A24P*, *FAM90A25P*, *FAM90A5*, *FAM90A6P*, and *FAM90A7*. The only GO term associated with these pseudogenes is nucleic acid binding (GO:0003676).

Gene set enrichment analysis

CTQ total (childhood trauma burden) was included as the independent variable of interest in linear models with probe residualized M-values as dependent variables. Income, education, and adult trauma burden were included and considered as covariates, as no 5mC clusters mediated the relationships between our exposures and outcomes of interest. We used the resultant 6,335 nominally significant p-values and probe names as GSEA input, facilitating the testing of 5,478 GO-terms. After Revigo redundancy reduction, 30 BH-significant GO-terms remained for interpretation. These included immune-related GO-terms: macrophage migration (GO:1905517, $p = 9x10^{-8}$, rank = 10), T cell receptor complex (GO:0042101, $p = 6x10^{-6}$, rank = 13), and chemokine-mediated signaling pathway (GO:0070098, $p = 7x10^{-4}$, rank = 28), among four others. Two CNS-related GO-terms were present: negative regulation of smoothened signaling pathway (GO:0045879, $p = 2x10^{-6}$, rank = 12), and glutamine family amino acid

biosynthetic process (GO:0009084, $p = 1x10^{-4}$, rank = 19), along with one HPA-axis related term: thyroid hormone metabolic process (GO:0042403, $p = 2x10^{-15}$, rank = 5). Additionally, numerous signal transduction and membrane transport GO-terms were present: positive regulation of phospholipase activity (GO:0010518, $p = 5x10^{-25}$, rank = 1), intrinsic component of peroxisomal membrane (GO:0031231, $p = 7x10^{-18}$, rank = 2), and RNA polymerase II carboxy-terminal domain kinase activity (GO:0008353, $p = 8x10^{-15}$, rank = 6), among others (Supplementary Table B.4). Analysis performed while excluding the single non-AA participant rendered highly similar results to those reported above (data not shown).

3.5 Discussion

As discussed, previous work has shown that women are subject to elevated risk of PTSD-related psychiatric illness development, and that AAs are at higher conditional risk of PTSD following trauma. Therefore, we applied genome-scale approaches to a cohort of majority AA women from a low SES, trauma-enriched environment, representing an understudied population, where our main goal was to assess whether blood-derived 5mC cluster MEs might index CNS endophenotypes of PTSD-related psychiatric illness. We were specifically interested in exploring whether blood-derived 5mC MEs *statistically* mediated the relations between ASEs (low income, low education, childhood trauma burden, adult trauma burden) and fronto-limbic brain morphometry measures (volume, SA, CT), thus serving as a peripheral proxy of such CNS effects. Based on previous work, we hypothesized that identified 5mC modules would be enriched with 5mC probes falling in genes with HPA-axis[105, 106], immune system[107, 108], and CNS-relevant [109–111] GO-functions.

Our study findings indicate that high childhood trauma burden, low educational attainment, and low income are each associated with CNS endophenotypes of PTSD-related psychiatric illness, and that a subset of blood-derived 5mC MEs *statistically* mediate some of these relationships. We further found that the individual probes from mediating 5mC MEs fell in genes with CNS-relevant and immune system GO-functions. Finally, we found that the underlying childhood trauma burden-associated methylomic network was enriched with HPA-axis, immune, and CNS-related gene sets, in addition to signal transduction and membrane transport functions. Overall, we posit that the peripheral epigenetic signatures indexing and/or mediating our relationships of interest are consistent with previously observed patterns in which PTSD-related psychiatric illness is accompanied by profiles of peripheral inflammation[33, 58, 60, 62].

In the current study, childhood trauma burden was negatively associated with hemisphere-mean RMFG SA, which mirrors findings from cohorts enriched for stress, trauma, and psychiatric illness [51, 52]. We also observed higher CT of both the FP and SFG in participants with low educational attainment (compared to high educational attainment), as well as higher cACC SA. Although studies on FP CT and ASEs (besides trauma) are scarce, our findings support past research investigating whether high CT is representative of a predisposition to psychiatric illness development, or whether high CT emerges after psychiatric illness onset. Specifically, prior longitudinal work showed higher cortical thickness in high risk (of mood disorder) participants compared to healthy controls at baseline. The high-risk participants that went on to develop MDD then showed increased FP CT over the two-year time period following their diagnosis[46]. This research implied that high FP CT *could* represent both a predisposition to MDD development risk, and a post-onset emergent phenotype of MDD.

Our study also shows that low household income is negatively associated with hemisphere-mean amygdala grey matter volume. This finding aligns with research in children/adolescents showing that low familial income-to-needs ratio[40] is associated with decreased amygdala grey matter volume, and with studies in adults showing that past year financial hardship[47] is associated with decreased amygdala grey matter volume. However, the relationships between SES-like measures and amygdala grey matter volume are well-studied, and variable conclusions between studies are common[228]. Hemisphere-mean hippocampal volume was not associated with any ASE; as in the case of amygdala grey matter volume, study conclusions are mixed[228], and the well-documented phenomenon of hemisphere-specific laterality may be implicated in the current negative findings[129–131].

In all, four blood-derived MEs were tested for mediating status between childhood trauma burden and RMFG SA. Both the Maroon and the White ME fully mediated the relationship and accounted for significant proportions of variability (47% and 35% respectively).

The Maroon module is composed of 11 probes, six of which are also full mediators of the aforementioned relationship. The six probes fully mediating the childhood trauma burden and RMFG SA relationship are located in the *MYH14, PGF, NEURL4, DNMT3A, SLC38A10,* and *LINC01531*. GO-terms associated with these loci include: *MYH14* (neuronal action potential GO:0019228, actomyosin structure organization GO:0031032), *PGF* (chemoattractant activity GO:0042056, vascular endothelial growth factor receptor binding GO:0005172), *NEURL4* (ubiquitin protein ligase activity GO:0061630), *DNMT3A* (DNA (cytosine-5-methyltransferase activity GO:0003886), and *SLC38A10* (amino acid transmembrane transporter activity GO:0015171). Of particular interest is the probe falling in the *DNMT3A* gene. Past longitudinal study in a predominantly female AA sample showed differential blood *DNMT3A* 5mC in

response to traumatic event exposure[229], and emerging work focused specifically on African ancestry individuals has identified a genome-wide significant association in *DNMT3A* in relation to MDD[230]. What is more, in a rodent model of fear conditioning, *Dnmt3a* mRNA expression is altered in forebrain neurons, leading to differential DNA 5mC of synaptic plasticity and memory formation genes[231]. *Dnmt3A* mRNA expression in rodent nucleus accumbens is also altered in response to chronic restraint stress, and targeted inhibition of *Dnmt3A* 5mC potentiates anti-depressant-like effects[232]. Based on the current and prior findings, it appears that the *DNMT3A* gene in humans is a *potential* locus of stress- or trauma-related biological embedding, and that measurements of *DNMT3A* 5mC taken from peripheral tissues *could be* indexing fronto-limbic variability associated with these exposures. Overall, our findings in the Maroon module imply that immune signaling, cellular proliferation, neuronal development, and epigenetic regulatory pathways could be *potential* substrates for the biological embedding of childhood trauma. However, whether these profiles, as measured in peripheral tissue, are indicative of neurobiological variability resulting from childhood trauma is beyond the scope of the current work. What is more, although our current results are derived from an all-female, majority AA cohort of participants, we cannot express with confidence whether these signatures are sex- or ancestry-specific.

On the other hand, the White module is comprised of 19 probes, three of which are full mediators of the relation between CTQ total and RMFG SA. Thirteen probes from this module are mapped to the "Family with sequence similarity 90" superfamily of pseudogenes, and all but two probes exist in a one megabase stretch of chromosome eight. The three full mediator probes (cg00686169, cg01544227, and cg08002107) fall in the *FAM90A24P, LINC00937, and FAM66D* loci, respectively. Little is known of the *FAM90A24P* locus, but the *LINC00937 locus*

encodes a LINC RNA species that is strongly associated with cutaneous melanoma prognosis[233], as well as endocervical cancer progression[234], suggesting that it plays a role in cellular proliferation. On the other hand, the *FAM66D* locus encodes a LINC RNA species whose expression is significantly upregulated in both Crohn's disease (CD) and ulcerative colitis (UC), the two most common types of inflammatory bowel disease (IBD)[235]. Efforts made towards elucidating the function of the broader chromosome eight locus highlight its structurally dynamic nature, in addition to its harboring of the defensin genes, the protein products of which play significant roles in the innate immune response, as well as antitumor response[236]. Defensin proteins are also recognized as key contributors to innate immune defense against UC, CD, and IBD in general[237].

With this in mind, it appears that 5mC of CpG probes mediating the childhood trauma burden and RMFG SA relationship interact with endothelial growth and innate immune response regulatory pathways. This notion is supported by repeated observation of increased blood-based inflammation profiles, as measured in adults that had previously experienced childhood TE[238–240]. More CNS-focused research has also shown significant dysregulation of stress-response and inflammation-related mRNA/protein levels in post-mortem frontal cortex of patients with neurodevelopmental[241] and psychiatric illnesses[242, 243]. What is more, decreased CT of medial pre-frontal cortex is observed in MDD patients, a relationship that was partially mediated by inflammatory factors known to influence neuroplasticity[244]. Unfortunately, the relevant studies informing our current results are sparse, and could be considered underpowered. This precludes a concrete ability to postulate on the relationships between peripheral inflammation and neural endophenotypes of PTSD-related psychiatric illness.

In our GSEA of childhood trauma burden -associated probes, the top enriched term was "positive regulation of phospholipase activity", meaning that multiple probes fall in genes that increase the frequency, rate, or extent of phospholipase activity. Of close relation is the 11th ranked GO-term, "inositol phosphate phosphatase activity". Together, these terms appear to represent one of the major mechanisms of neuronal and hormonal signal transduction in mammals, the activation of phospholipase C (PLC) and subsequent activation of inositol signaling pathways[245]. In brief, extracellular stimuli receptor binding activates PLC. PLC then functions to convert phosphatidylinositol-4-5 bisphosphate (P45P2) into inositol-1-4-5 trisphosphate (I145P3). From this, the kinase-activating secondary messenger diacylglycerol (DAG) is activated[245]. There are many isoforms of the PLC enzyme family, but PLC-δ1 is strongly implicated in regulating cell-motility & cytoskeletal organization[246], and in inhibiting inflammatory immune responses[247]. Quantities of this PLC isoform, PLC-δ1, are highly correlated with nuclear levels of P45P2[248]; interestingly, P45P2 regulates transcription through binding interaction with the RNA polymerase II C-terminal tail[249]. The aforementioned processes are specifically represented within our GSEA results (Supplementary Table B.4): RNA polymerase II carboxy-terminal domain kinase activity, regulation of mitotic spindle organization, and numerous immune signaling and motility terms. Given childhood trauma burden is the ASE of interest, it stands to reason that in our cohort of AA women, peripheral inflammatory signals could be heightened in those with high childhood trauma burden, and phospholipase activity may be upregulated in an attempt to restore homeostatic balance. To this end, genomic, transcriptomic, and metabolic investigations into phospholipase involvement in psychiatric illness is ongoing, supporting the feasibility of this concept[250–252], but studies focused ancestry- or race-specific effects are currently non-existent.

Overall, we posit that the blood-based epigenetic signatures indexing and/or mediating our relationships of interest are potentially explained by ASE-associated peripheral inflammation. In response to perceived stress or trauma, fronto-limbic brain regions associated with cognitive function, emotional reactivity, and memory are activated, signaling to the HPA-axis' neural hub, and eventually potentiating glucocorticoid (GC) release. Critical to immunological, metabolic, cardiac, and homeostatic functions, GCs act mechanistically in and around the CNS and peripheral nervous system (PNS), including serving as a long-range negative feedback mechanism to inhibit further HPA-axis activity. Additionally, the ligands and receptors of cytokines and neurotransmitters are shared throughout the CNS, HPA-axis, and immune systems[55, 56]. Aberrant responses to this signaling can result in neural growth, inflammation, metabolism, and stress-related pathway disruption through molecular and epigenetic mechanisms[57], which are consistent with 5mC-derived patterns observed in peripheral tissue in the current report.

The foremost strength of the current research is the inclusion of multi-modal data types (epigenetic/neuroimaging) to help explain the biological embedding of ASEs. Furthermore, the dimension reduction techniques used in the current work, clustering and principal component extraction, were valuable in reducing the burden of multiple hypothesis testing and in describing potentially related networks of 5mC loci. It also focused the analysis onto loci with greater prospect for proxy or surrogate status with etiologically-relevant CNS tissue. The composition of the cohort is also a strength of the current work, being all-female and majority AA. This sample is reflective of an understudied population at elevated risk for PTSD and for persistence of PTSD-related psychiatric illness[253]. However, because of the homogenous nature of the cohort, we are unable to make confident conclusions about ancestry-specific or sex-specific

effects. Limitations of the current study include relatively small sample size and the inability to infer causality from cross-sectional samples. In addition, the current research used peripheral tissue in lieu of etiological CNS tissue for obvious technical reasons. Functional MRI measures such as neurotransmitter specific PET imaging could also provide deeper endophenotypic insights. Future studies on this topic should capture longitudinal data from a larger sample and could investigate genetic factors or tissues of etiological interest.

The current study showed that exposure to high childhood trauma burden, low educational attainment, and low income are each associated with neuroimaging endophenotypes of psychiatric illness, and that a subset of blood-derived 5mC measurements statistically mediate these relationships. In addition, we found that the mediating 5mC MEs were enriched with probes falling in genomic regulatory regions, and in genes with CNS-relevant and immune system GO-functions. Finally, we found that the underlying childhood trauma burden-associated methylomic network was enriched with HPA-axis, immune, and CNS-related gene sets. To this end, the current results provide a feasible epigenetic mechanism through which ASEs are embedded in human physiological systems, and through which ASEs contribute to the development of PTSD-related psychiatric illness. Although these concepts are broadly applicable across race, ethnicity, and biological sex, the current results highlight mechanisms of biological embedding that *may* be specific to AA women. These epigenetic signatures could be taken as peripheral biomarkers or proxies of perturbed underlying neurobiology associated with ASEs, and could further research efforts into etiological tissues of interest such as endocrine, immune, and CNS cell types. These findings could also serve to potentiate increased investigation of understudied populations at significant risk for onset or persistence of certain PTSD-related psychiatric illnesses, such as AA women in low income, high trauma environments.

3.6 Figures and Tables

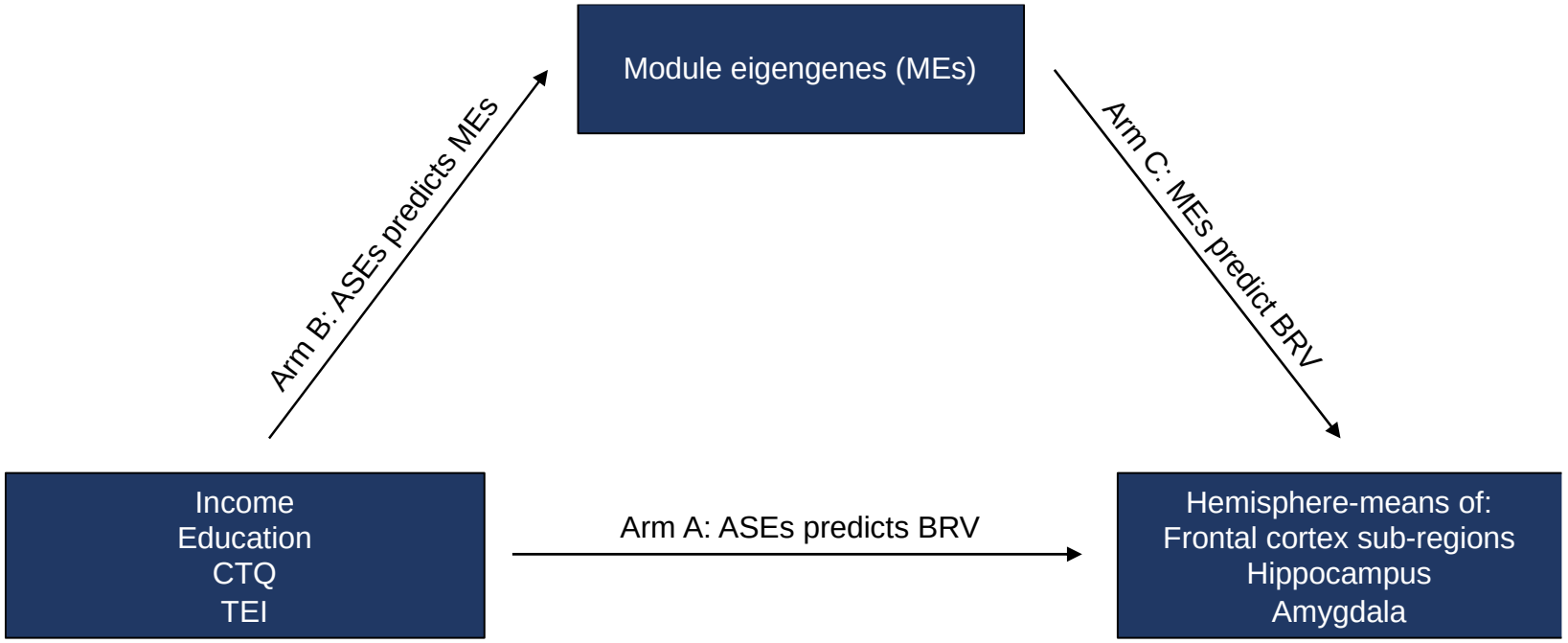


Figure 3.1 Conceptual model testing individual module eigengenes (MEs) as mediators of the hypothesized associations between adverse social exposures (ASEs) and variability in fronto-limbic brain morphometries. Arm A. Analyses included hemisphere-mean hippocampus and amygdala volumes as dependent variables (in separate models). Income, education, CTQ total, and TEI were included as independent variables of interest, whereas genomic ancestry, age, employment/disability and ICV were included as covariates. The same predictors and covariates were included, minus ICV, to predict frontal cortex subregions SA and CT measures. Arm B. Analyses included 5mC MEs as dependent variables (in separate models). Income, education, CTQ total, and TEI were included as independent variables of interest, and employment/disability was included as a covariate. Arm C. Analyses included individual fronto-limbic morphometry measures as dependent variables (in separate models). Individual 5mC MEs were included as independent variables of interest, whereas age, genomic ancestry, employment, and ICV (ICV was only included in hippocampus and amygdala models) were included as covariates.

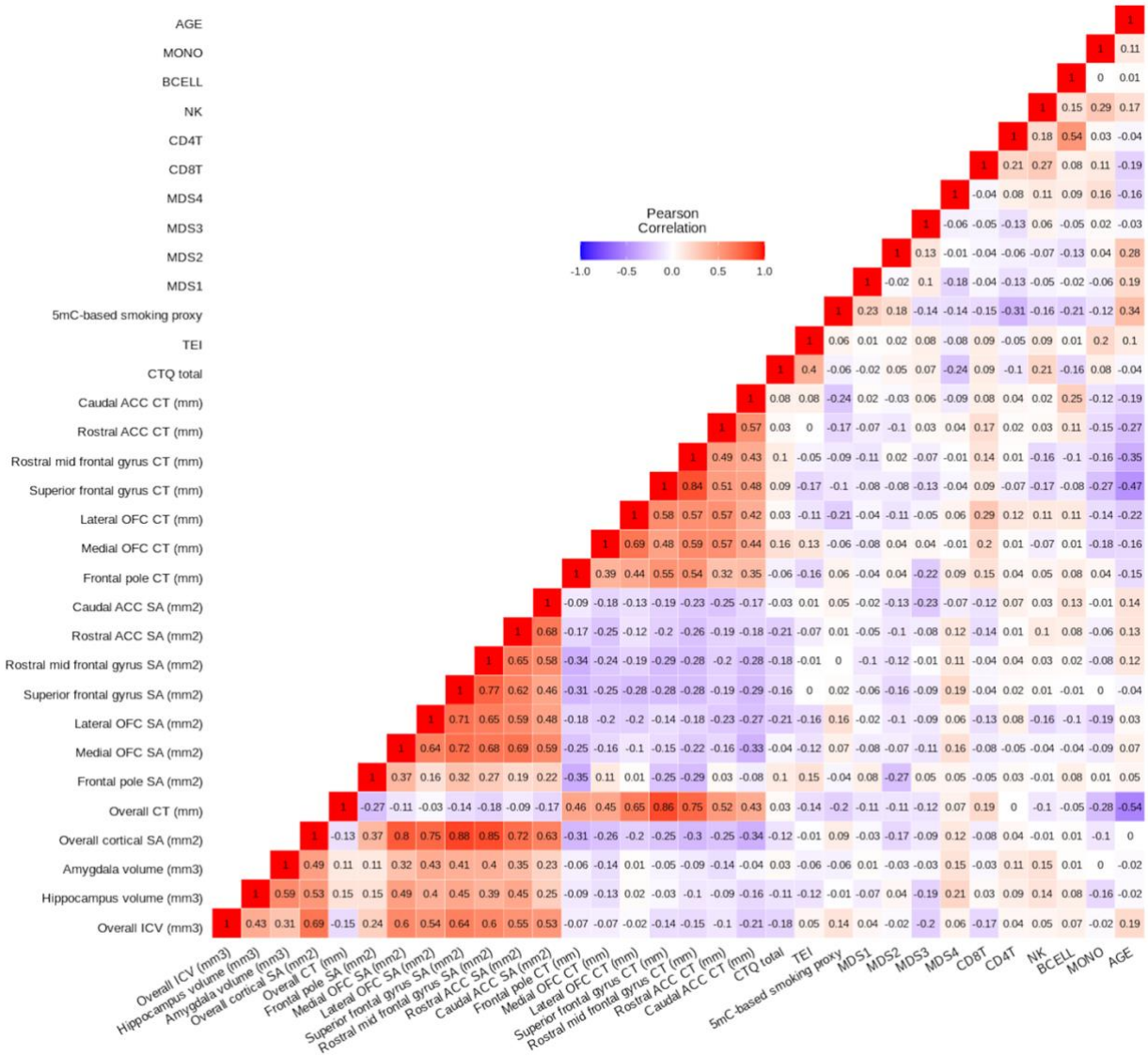


Figure 3.2 Pearson's correlation heatmap of variables used throughout the current study. Age is negatively associated with multiple frontal cortex cortical thickness (CT) measures (Pearson's correlation r_{range} = -0.15 to -0.47, p_{range} = 0.092 to $6x10^{-7}$). Within frontal cortex CT measures, multiple subregions show positive correlations with one another (Pearson's correlation r_{range} = 0.32 to 0.84, p_{range} = 0.004 to < $2x10^{-16}$). The same phenomenon is observed within frontal cortex surface area (SA) measures (Pearson's correlation r_{range} = 0.16 to 0.77, p_{range} = 0.210 to <$2x10^{-16}$). Overall CT is positively correlated with each of the frontal cortex CT measures (Pearson's correlation r_{range} = 0.43 to 0.86, p_{range} = $1x10^{-7}$ to <$2x10^{-16}$), and overall SA is positively correlated with each of the frontal cortex SA measures (Pearson's correlation r_{range} = 0.37 to 0.88, p_{range} = $5x10^{-4}$ to < $2x10^{-16}$).

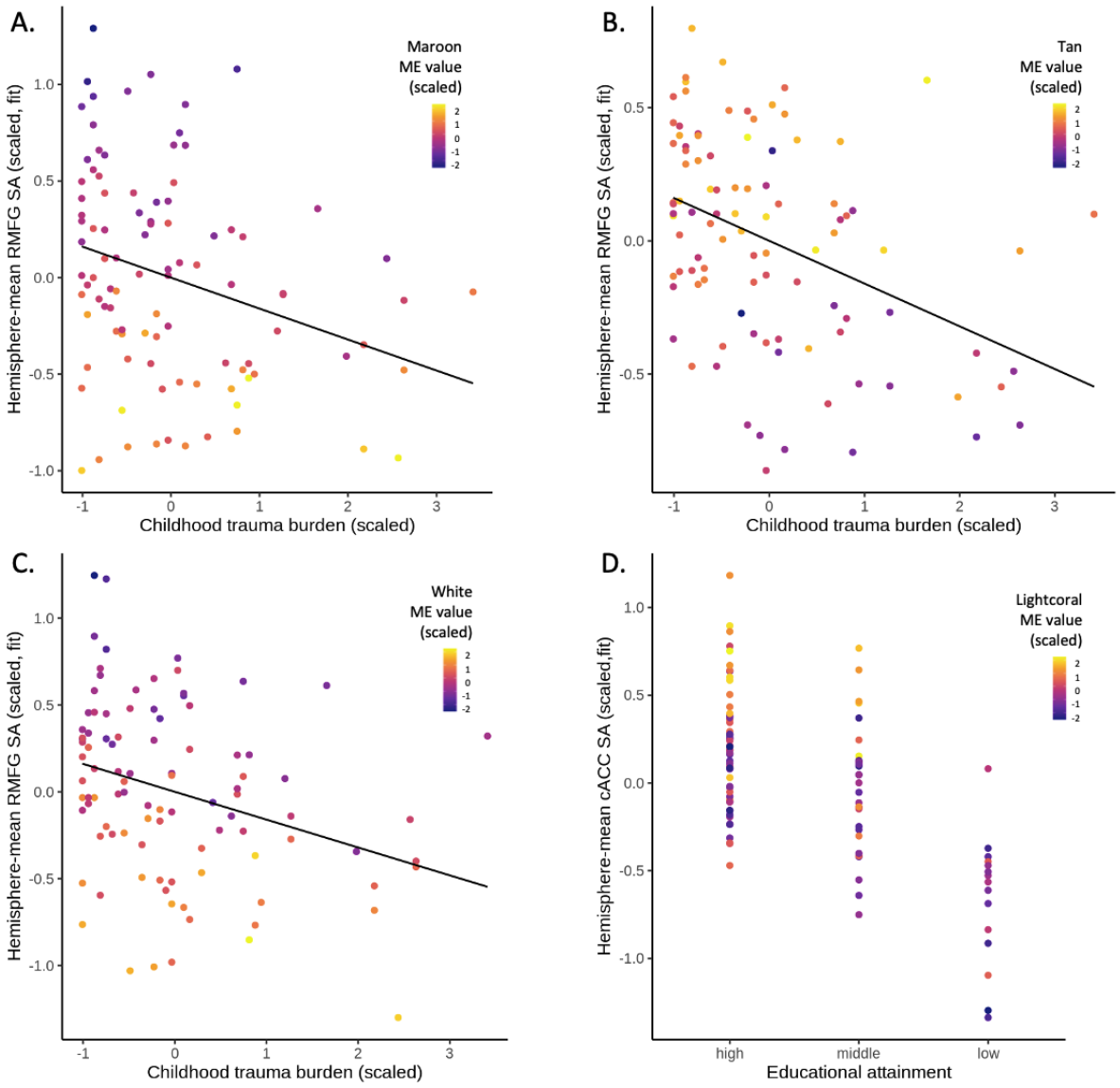


Figure 3.3 CTQ total is negatively associated with rostral middle frontal gyrus (RMFG) surface area (SA). Low educational attainment status is associated with low caudal anterior cingulate cortex (cACC) SA. Out of seven CTQ-associated module eigengenes (MEs), the three strongest relationships with hemisphere-mean RMFG SA were the Maroon (A), White (B), and Tan (C) MEs. Out of two low education attainment-associated MEs, the strongest relationship with hemisphere-mean cACC SA was the Lightcoral ME (D). As shown, neuroimaging measures are adjusted by age, ME value, genomic ancestry, employment/disability status, and an adverse social exposure (ASE). Dependent and independent variables were scaled. A. CTQ total was positively associated with the Maroon ME (β = 0.289, SE = 0.109, t = 2.639, p = 9.8×10^{-3}). In turn, the Maroon ME was negatively associated with RMFG SA (β = - 0.451, SE = 0.091, t = - 4.966, p = 3×10^{-6}). B. CTQ total was negatively associated with the Tan ME (β = - 0.248, SE = 0.112, t = - 2.220, p = 0.029). The Tan ME was, in turn, positively associated with RMFG SA (β = 0.233, SE = 0.100, t = 2.322, p = 0.023). C. CTQ total was positively associated with the White ME (β = 0.236, SE = 0.110, t = 2.151, p = 0.034). In turn, the White ME was negatively associated with RMFG SA (β = - 0.392, SE = 0.093, t = - 4.231, p = 6×10^{-5}). D. Low educational attainment (no high school certificate) was associated with the Lightcoral ME (β = - 0.615, SE = 0.296, t = - 2.079, p = 0.040). The Lightcoral ME, in turn, was positively associated with cACC SA (β = 0.215, SE = 0.099, t = 2.182, p = 0.032).

Table 3.1 Demographic & psychosocial measure summary statistics (non-scaled) for the current sample (n = 97)

Measure	Description	Value
Age	Mean [SD] (Range)	40 [12.5] (19 - 62)
Sex	Female	100%
Self-reported race/ethnicity	African American	99%
Household income (income)	$0-499	30%
	$499-999	36%
	$1000+	34%
Educational attainment (education)	Some high school	15%
	High school grad or GED	30%
	≥ Undergraduate degree	55%
Employment/disability	Employed	33%
	Disabled (not employed)	7%
	Not employed or disabled	60%
Childhood trauma questionnaire (CTQ total)	Mean [SD] (Range)	40.5 [15.4] (25 - 93)
Traumatic events inventory (TEI)	Mean [SD] (Range)	4.2 [2.3] (0 - 10.7)

Table 3.2 Adverse social exposures (ASEs) predict hemisphere-mean fronto-limbic brain morphometry

	Low household income				Low educational attainment				Childhood trauma burden				Adult trauma burden			
	β	**SE**	**T**	**P**	**β**	**SE**	**T**	**P**	**β**	**SE**	**T**	**P**	**β**	**SE**	**T**	**P**
Hipp. Volume (mm3)	-0.425	0.267	-1.590	0.116	0.005	0.279	0.019	0.985	0.030	0.115	0.260	0.796	-0.167	0.101	-1.644	0.104
Amyg. Volume (mm3)	-0.649	0.283	-2.299	0.024*	0.091	0.294	0.310	0.757	0.145	0.118	1.228	0.223	-0.146	0.106	-1.380	0.171
FP SA (mm2)	0.556	0.285	1.950	0.055	-0.065	0.294	-0.221	0.826	-0.004	0.114	-0.037	0.971	0.109	0.107	1.017	0.312
Medial OFC SA (mm2)	0.271	0.293	0.926	0.357	-0.374	0.302	-1.237	0.220	-0.052	0.117	-0.440	0.661	-0.112	0.110	-1.021	0.310
Lateral OFC SA (mm2)	-0.033	0.298	-0.111	0.912	-0.244	0.307	-0.796	0.428	-0.211	0.119	-1.769	0.081	-0.093	0.111	-0.834	0.407
SFG SA (mm2)	0.386	0.289	1.336	0.185	-0.435	0.298	-1.460	0.148	-0.223	0.116	-1.926	0.058	0.039	0.108	0.358	0.721
RMFG SA (mm2)	-0.050	0.278	-0.181	0.857	-0.543	0.286	-1.897	0.061	-0.231	0.111	-2.079	0.041*	0.036	0.104	0.346	0.730
Rostral ACC SA (mm2)	-0.176	0.291	-0.602	0.549	-0.511	0.301	-1.700	0.093	-0.229	0.117	-1.961	0.053	0.029	0.109	0.268	0.790
Caudal ACC SA (mm2)	0.121	0.272	0.443	0.659	-0.855	0.280	-3.049	0.003**	-0.110	0.109	-1.006	0.317	0.044	0.102	0.436	0.664
FP CT (mm)	-0.137	0.293	-0.468	0.641	0.684	0.302	2.267	0.026*	-0.079	0.117	-0.671	0.504	-0.120	0.110	-1.098	0.275
Medial OFC CT (mm)	-0.002	0.293	-0.008	0.993	0.346	0.315	1.099	0.275	0.138	0.118	1.172	0.244	0.022	0.111	0.202	0.840
Lateral OFC CT (mm)	-0.023	0.304	-0.074	0.941	0.241	0.314	0.769	0.444	0.055	0.122	0.451	0.653	-0.112	0.114	-0.983	0.328
SFG CT (mm)	-0.088	0.248	-0.356	0.723	0.553	0.256	2.162	0.034*	0.070	0.100	0.699	0.487	-0.170	0.093	-1.831	0.071
RMFG CT (mm)	-0.078	0.281	-0.276	0.783	0.443	0.290	1.530	0.130	0.072	0.113	0.637	0.526	-0.112	0.105	-1.062	0.291
Rostral ACC CT (mm)	-0.063	0.297	-0.211	0.833	0.060	0.318	0.188	0.851	0.019	0.118	0.163	0.871	-0.034	0.112	-0.305	0.761
Caudal ACC CT (mm)	-0.124	0.296	-0.419	0.676	0.187	0.318	0.589	0.558	0.046	0.119	0.388	0.699	0.004	0.112	0.039	0.969

*** P < 0.001, **P < 0.01, * P < 0.05, **P: BH significant (bold)**

Both dependent and independent variables were scaled prior to analysis. Childhood trauma burden was measured with the Childhood Trauma Questionnaire, and Adult trauma burden was measured with the Traumatic Events Inventory. In separate linear models including hippocampus and amygdala volumes as dependent variables, income, education, CTQ total, and TEI were included as independent variables of interest, whereas genomic ancestry, age, employment/disability and ICV were included as covariates. The same independent variables and covariates were included, minus ICV, to predict frontal cortex subregions SA and CT measures.

Abbreviations. Hipp: hippocampus, Amyg: amygdala, FP: frontal pole, mOFC: medial orbito-frontal cortex, lOFC: lateral orbito-frontal cortex, SFG: superior frontal gyrus, RMFG: rostral medial frontal gyrus, rACC: rostral anterior cingulate cortex, cACC: caudal anterior cingulate cortex, SA: surface area, CT: cortical thickness, β: scaled beta value, SE: standard error, T: t statistic value.

Table 3.3 Module eigengenes (MEs) mediating observed adverse social exposure (ASE), fronto-limbic brain morphometry relationships			
Maroon: Rostral Mid Frontal Gyrus SA	β	95% CI Lower	95% CI Upper
Average indirect effect (IDE)	**-0.466****	-0.952	-0.098
Average direct effect (DE)	-0.517	-1.464	0.407
Average total effect (TE)	-0.983*	-1.959	-0.016
Tan: Rostral Mid Frontal Gyrus SA	β	95% CI Lower	95% CI Upper
Average indirect effect (IDE)	-0.124	-0.448	0.100
Average direct effect (DE)	-0.884	-1.869	0.085
Average total effect (TE)	-1.001*	-1.955	-0.040
White: Rostral Mid Frontal Gyrus SA	β	95% CI Lower	95% CI Upper
Average indirect effect (IDE)	**-0.348***	-0.781	-0.026
Average direct effect (DE)	-0.667	-1.590	0.251
Average total effect (TE)	-1.014*	-1.983	-0.069
Yellow: Rostral Mid Frontal Gyrus SA	β	95% CI Lower	95% CI Upper
Average indirect effect (IDE)	-0.128	-0.474	0.142
Average direct effect (DE)	-0.888	-1.908	0.148
Average total effect (TE)	-1.016*	-2.003	-0.019
Lightcoral: Caudal ACC SA	β	95% CI Lower	95% CI Upper
Average indirect effect (IDE)	-0.107	-0.307	0.015
Average direct effect (DE)	**-0.737***	-1.296	-0.176
Average total effect (TE)	-0.844**	-1.394	-0.283

*** P < 0.001, ** P < 0.01, * P < 0.05, **P: BH-significant (bolded)**

Both dependent and independent variables were scaled prior to analysis. SA: surface area (mm2), CT: cortical thickness (mm), β: coefficient estimate, CI: confidence interval

Table 3.4 Probe-wise mediation analysis of the relationship between childhood trauma burden (CTQ total) & rostral middle frontal gyrus (RMFG) surface area (SA)

Probe name	Module name	Chr:pos (hg38)	Gene name	IDE β	IDE P	DE β	DE P	TE β	TE P
cg21622733	**Maroon**	Chr19:35408551	*LINC01531*	**-0.512**	**0.003****	-0.471	0.320	-0.983	0.044*
cg19805668	**Maroon**	Chr14:74951972	*PGF*	**-0.411**	**0.009****	-0.592	0.217	-1.003	0.042*
cg10592766	Maroon	Chr19:50226846	*MYH14*	-0.321	0.019*	-0.670	0.182	-0.990	0.048*
cg03710029	Maroon	Chr17:81291801	*SLC38A10*	-0.383	0.025*	-0.613	0.188	-0.996	0.043*
cg15150970	Maroon	Chr2:25250660	*DNMT3A*	-0.253	0.049*	-0.753	0.125	-1.006	0.040*
cg12219789	Maroon	Chr17:7324180	*NEURL4*	-0.352	0.049*	-0.622	0.183	-0.974	0.049*
cg01544227	White	Chr12:8323720	*AC092745.5*	-0.305	0.036*	-0.689	0.175	-0.994	0.044*
cg00686169	White	Chr8:8024026	*FAM90A24P*	-0.277	0.043*	-0.740	0.126	-1.017	0.037*
cg08002107	White	Chr8:12135255	*USP17L7*	-0.268	0.047*	-0.738	0.122	-1.006	0.040*

*** P < 0.001, **P < 0.01, * P < 0.05, **P: BH-significant (bold)**

Both dependent and independent variables were scaled prior to analysis. IDE: indirect effect, DE: direct effect, TE: total effect, β: coefficient estimate, P: nominal p-value, BH: Benjamini Hochberg

CHAPTER 4: LOCUS-SPECIFIC CPG METHYLATION MEDIATES ENVIRONMENTAL ENRICHMENT AND MRNA EXPRESSION RELATIONSHIPS IN MOUSE DENTATE GYRUS

4.1 Abstract

Environmental enrichment (EE) and adverse social exposures/environments exert effects on similar physiological, neurological, and cognitive substrates, especially in neurodevelopmentally plastic windows of time such as childhood, adolescence, and young adulthood. In humans and rodents, the effects of these exposures appear to converge on the hippocampus, but the molecular mechanisms by which EE is biologically embedded in the hippocampus are not well characterized. Research has pointed towards the importance of CpG and non-CpG methylation (mCG; mCH) in this process. Past research has not clearly investigated the *locus-specific* effects of EE on mCG/mCG, or the potential causal mediation effects of locus-specific mCG/mCH on the relationship *between* EE and differential mRNA transcription in the hippocampus.

We applied secondary data analysis methods on a next-generation sequencing dataset in mouse (GSE95740) to investigate the effects of EE on locus-specific mCG, mCH, and mRNA expression in dorsal and ventral dentate gyrus (dDG; vDG) in the hippocampus. We set out to assess the putative causal mediation status of mCG and mCH on the relationship between EE and differential mRNA transcription in both dentate gyrus sub-sections, in order to elucidate underlying epigenetic mechanisms responsible for the biological embedding of EE.

At nominal significance thresholds, 404 mCG loci were full mediators of the relationship between their cis-located mRNA transcript and EE in dDG, while 211 were full mediators in

vDG. Both the MAPK and TNF-B signaling pathways were implicated in the biological embedding of EE, and numerous mediating loci were identified in genes with distinct roles in hippocampal development and mental/neurodevelopmental disorders, including *Dab1*, *Rora*, *Card14*, *Kalrn*, and *Cdh13*. Epigenetic regulation of gene expression through differential mCG was observed at a higher degree in dDG than in vDG, whereas the opposite was true of mCH in vDG versus dDG.

The findings of the current study inform divergent molecular mechanisms of the biologic embedding of EE on a locus-specific, and brain region-specific basis. These findings could be used in the future to drive pointed locus-specific investigations of hippocampal development or mental health disorder etiology in humans.

4.2 Introduction

In rodents, stressor paradigms are commonly employed to investigate the molecular effects of adverse social exposures (ASEs) and can be a significant source of human-relevant insight. These paradigms have well-documented effects on biological and cognitive systems in rodents that mirror effects of ASEs in humans. In rodents, biological effects of ASEs include alterations to the hypothalamic-adrenal-pituitary-axis (HPA-axis)[71, 72], immune system[73, 74], and central nervous system (CNS)[68, 72]. Of particular importance, though, are structural and functional alterations to the evolutionarily conserved fronto-limbic brain regions[68–70]: hippocampus, amygdala, and frontal cortices. These brain regions support threat, reward, executive control, memory, and adaptive behavioral/emotional responses among others[53], and are impaired in numerous affective and stress-related mental illnesses in humans[61–64]. To this end, the cognitive effects of stressor paradigms in rodents also mirror those of mental health

disorders in humans; researchers have observed learning and memory deficits[75], and increases in anxiety-like[72] and depression-like behaviors[72] in rodents exposed to stressor paradigms.

On the other hand, environmental enrichment (EE), typically regarded as housing methods in which the potential for social interactions and novel object exploration are increased compared to standard housing[76], compensates for or rescues many effects of stress exposure in rodents[77–81]. Researchers have also observed that EE in rodents leads to notable improvement in many cognitive and behavioral domains[83–86]. Accompanying these cognitive and behavioral alterations in response to EE are a host of neurobiological changes to the hippocampus and associated cortical regions, including increased neurogenesis[87–90], synaptogenesis[91], and gliogenesis[90], as well as increased dendrite outgrowth[83], cortical neurotransmission[92], and long term potentiation[93]. EE also increases adult hippocampal volume[94] and granule cell neurogenesis in the dentate gyrus[95] (a hippocampal sub-region which is responsible for neurogenesis throughout the life cycle). Importantly, the effects of ASEs and EE are especially salient during periods in physiological development where the CNS is particularly plastic, namely childhood, adolescence, and young adulthood. Indeed, a number of the neurobiological changes associated with EE, as noted above, are observed in these periods of developmental plasticity[90, 91, 93]. This concept is further highlighted in a recent study that applied early life stress in a post-natal rodent model, and then rescued the behavioral (fear response to conditioned stimulus) and neurobiological (survival and activation of newborn hippocampal neurons) effects through EE[96].

It appears then, that the effects of stressor and EE paradigms in rodents operate on similar underlying neurobiological substrates, namely the hippocampus and dentate gyrus. This implies that bi-directional alterations to underlying molecular regulatory networks may be at play in

response to these divergent exposures. This concept is not specific to rodents, as noted by the hypothetical neglect-enrichment continuum discussed in the context of human children/adolescents[97].This hypothetical framework states that the "healthy" development of children and adolescents is dependent on environmental/caregiver facilitation of either deprivation or enrichment. As reviewed in King et al 2019[97], evidence suggests that exposure to both ends of this common spectrum act on the fronto-limbic pathways of the brain. Therefore, identifying molecular alterations in the rodent brain associated with EE could provide valuable insight regarding molecular resilience to ASEs in humans.

To this end, the molecular mechanisms playing a role in the biological embedding of external exposures, such as stress and EE, have come under investigation, and recent research in the field of epigenetics has implicated the importance of 5-methyl-cytosine (5mC; methylation) marks at cytosine-phosphate-guanine (CpG) base pair sequences in this process[98–100]. In addition to CpG methylation (mCG), neuronal cell types have unique patterns of non-CpG methylation (mCH) that dictate transcriptional regulation[101]. These epigenetic mechanisms are known mediators of gene by environment interactions[102–104], but current technological limitations make it impossible to measure epigenetics in the living human brain--the primary etiological tissue of interest in regards to mental health outcomes. In lieu of these innate barriers, researchers have forged alternative paths forward towards elucidating the relationship between EE and epigenetic mechanisms that potentiate EE-related changes across the CNS. One such alternative path is the use of EE animal models that allow direct access to CNS tissue in tandem with genome-scale molecular data.

Recently, researchers using high-resolution in vivo structural MRI observed distinct increases in hippocampal volume in EE-exposed mice, in both dorsal (+8.5%) and ventral poles

(+6.1%)[121]. In addition, EE was associated with a +60% increase in newborn neurons labeled with 5′-bromo-2′-dexoyuridine (BrdU) across the dorsal and ventral regions of the dentate gyrus. These results aligned with previous findings in the same brain regions[94, 95], further bolstering the notion that EE has significant effects on the hippocampus and dentate gyrus during post-natal development. Through next-generation sequencing methods, researchers also identified altered mCG and mCH profiles in these same brain regions, which correlated with mRNA expression across discrete regions of the genome. Interestingly, immature neuron and neural stem cell marker mRNA was upregulated in the dorsal dentate gyrus (dDG), whereas mature neuron marker mRNA was upregulated in the ventral dentate gyrus (vDG). Researchers posit that their findings are indicative of region-specific transcriptional regulation of neurogenesis mediated by epigenetic mechanisms in response to EE[121].

Although the aforementioned research effectively investigated the effects of EE on CNS methylation and transcription, their methods focused on *genomic-region* level effects, i.e. differentially methylated regions characterized by shared methylation patterns across large stretches of DNA sequence. They were also focused on investigating the molecular differences between the dDG and vDG. However, their research did not clearly investigate the *locus-specific* effects of EE on DNA methylation, or the potential mediating effects of DNA methylation on the relationship *between* EE and differential mRNA transcription. Importantly, researchers have observed locus-specific effects of psychosocial exposures on DNA methylation throughout the genome in both humans[105, 113, 117, 120, 254, 255] and rodents[31, 256–258]. This research has also identified associations between locus-specific variability in DNA methylation and physiological/behavioral measures[42, 117, 118, 120], further informing the hypothesis of biological embedding of the psychosocial environment through epigenetic mechanisms.

With these limitations in mind, the current study applied secondary data analysis methods on the Zhang et al 2018 next-generation sequencing data to investigate the locus-specific effects of EE on dDG/vDG mCG and mCH. We set out to assess the potential mediating status of mCG and mCH on the relationship between EE and differential mRNA transcription in both dentate gyrus sub-sections, in order to elucidate underlying epigenetic mechanisms responsible for the biological embedding of EE.

We hypothesized that methylation at a number of mCG and mCH sites would mediate the EE/mRNA transcription relationship on a locus-specific basis; and, due to their known roles in regulating gene expression[102–104, 259], that transcription start-sites (TSS) and transcription factor binding sites (TFBS) underlying differentially expressed genes (DEGs) would be enriched with these mediating loci. Based on previous work in humans[260, 261], we also hypothesized that the network of differentially methylated loci (DMLs) that mediate mRNA/EE relationships would be enriched with gene ontology functions relating to: synapse development, neurotransmitter signal transduction, neuronal development, inflammation, immune cell signaling and migration, cytoskeleton organization, MAPK cascade, and regulation of phospholipase activity.

Finally, we compared results from our previous studies on associations between human peripheral tissue mCG, fronto-limbic neuroimaging measures, and ASEs[260, 261] to the locus-specific DML results from the current study. We hypothesized that DMLs from human peripheral tissues associated with ASE exposure would also be differentially methylated in rodent CNS tissue in association with EE exposure, and that the observed direction of effect on mCG between EE and ASEs would be opposite one another.

4.3 Methods and Materials

Zhang et al 2018: research paradigm

The current research draws on existing data from a study investigating the molecular effects of EE in mouse dentate gyrus[121]. All procedures carried out by the original research team were performed in accordance with the guidelines established by the Canadian Council on Animal Care (CCAC) with protocols approved by the McGill University Facility Animal Care Committee (FACC). In brief, male C57/Bl/6 mice were weaned on postnatal day 22. Afterwards, siblings were assigned to either the enriched housing or standard housing conditions. Standard housed animals were raised in groups of three male mice from different mothers in a 30 × 18 cm cage. The enriched group contained 12 male mice, housed in a larger rectangular plexiglass cage (78 × 86 cm) with a plexiglass top, which contained toys such as running wheels, a bridge, and novel objects (changed weekly). All animals remained in their respective housing conditions for eight weeks, and animals were sacrificed on age day 80 (post sexual maturation). More detailed information on the primary study paradigm can be found in the original publication[121]. Visual representation of the study design is observed in Figure 4.1.

Zhang et al 2018: molecular methods

Comprehensive next-generation sequencing data from the aforementioned study of EE[121] was used as the basis for the following analyses. Researchers took a 300 μm diameter punch of each dDG and vDG region separately for tissue collection. Five biological replicates (samples) per condition were used for sequencing assays; each sample was composed of tissue from ten to 12 mice (technical replicates) ($n_{biological}$ = 5 per group; $n_{technical}$ = 50 to 60 per group). Resulting gene expression profiles from bulk tissue showed strong correlation with profiles

characteristic of neurons[262]. Regarding RNA sequencing (RNA-seq), as detailed in the original publication, data were collected via a HiSeq 2000 at a sequencing depth of 30 million using paired-end 100 bp reads. Regarding DNA methylation (MethylC-seq), one μg of DNA from each tissue type was bisulfite converted, and methylated/unmethylated spike in controls were added to assess conversion efficacy. Whole genome bisulfite sequencing (WGBS) libraries were prepared, and library amplification was performed. Across all samples, WGBS average genome coverage was 96.3%, average cytosine coverage was 89.0%, and average bisulfite conversion efficacy was 99.6%. More detailed information on the collection of tissues, and processing of molecular data can be found elsewhere[121].

Molecular data pre-processing and quality control: Current study

For each dDG and vDG, we used genome-wide RNA seq and MethylC-seq sequencing data, paired from each biological replicate.

RNA-seq: Mouse genome mm10[263] RNA-seq count and transcript per million (TPM) data were subset and assessed for quality thresholds after download from public repository GSE95740[264]. Reads mapping to the X-chromosome were removed, to limit sex-specific effects in downstream analyses. Following this, in similarity with the original authors' methodology, reads passed this phase of QC if we observed read count of greater than ten in two or more samples, *and* TPM of one or greater in three or more samples. Across both tissues, from an original 46,480 reads, ~17k (16,913 dDG; 16,952 vDG) reads survived this phase of QC. Finally, we removed reads with zero variance between SH and EE groups. From both tissues, this removed an additional 376 reads, leaving 16,537 reads in dDG and 16,576 in vDG out of the

original 46,480. These remaining reads were used for downstream analyses. A visual representation of pre-processing and QC steps taken in the current study is shown in Figure 4.2a.

CG methylation: After download from public repository GSE95740[264], MethylC-seq data were read into R by sample. Reads were only considered if they were represented across all ten samples and if they were mapped to a CG site (one of: CCG, CGA, CGC, CGG, CGN, or CGT). Remaining reads were then subset and QC'ed. Loci mapping to CG sites were considered if read count across all samples was greater than five. X-chromosome mapped loci were removed, as well as loci with zero variance between groups. For dDG and vDG mCG sites, 5,444,204 and 5,520,791 reads survived this QC, respectively. After QC, loci-specific mCG read count was divided by total count, resulting in a proportion ranging from zero to one. These proportions were then variance stabilized and logit-transformed[159] towards a normal distribution. A visual flowchart of QC methods is seen in Figure 4.2b.

CH methylation: MethylC-seq data was read into R by sample, and reads were kept where represented across all ten biological replicates. Reads were then subset to include only those mapping to CH sites (one of: CAA, CAC, CAG, CAN, CAT, CCA, CCC, CCN, CCT, CTA, CTC, CTG, or CTT). Remaining reads were then subset and QC'ed using the same methods as mCG loci. For dDG and vDG mCH sites, 704,180 and 1,023,564 reads survived QC, respectively. For each remaining locus, we captured a Boolean "methylated" call for use in downstream analyses. Calls were produced by the "call-methylation-state" function in the "methylpy" Python package[265]. This function performs a binomial test on each cytosine to delineate cytosines that are significantly methylated compared to noise. A visual flowchart of QC methods is seen in Figure 4.2c.

Mediation: mCG mediates EE/mRNA relationships

To investigate the molecular mechanisms underlying the biological embedding of EE on mRNA transcription, we tested the mediating effects of locus-specific mCG on the relationship between EE and mRNA expression using the "mediation" package in R[168] (Figure 4.3, Mediation). Importantly, only a targeted subset of mCG loci were assessed in this fashion due to our testing framework requirements.

Described with more detail in the individual sections below, our mediation testing framework was composed of four arms (A, B, C, Mediation). We first tested the effects of EE on mRNA transcription, and advanced transcripts with nominally significant ($p < 0.05$) relationships with EE to analysis Arm C. We then carried out false discovery rate (FDR) and log2 fold-chance (FC) thresholding to identify transcripts with the strongest relationships to EE. In Arm B analyses, we tested the effect of EE on individual mCG loci. Again, mCG loci with nominally significant relationships to EE were advanced to study Arm C. We carried out FDR and percent difference (PD) thresholding to identify mCG loci with the strongest relationships to EE. Before carrying out Arm C analyses, we assessed the genes to which the nominally significant transcripts and mCG loci from Arms A and B were mapped. We only kept DEGs that had a DML mapped to it. We only kept DMLs that belonged to a DEG. This reduced the Arm C testing pool to a focused number of mRNA/mCG read pairs, where a putative cis-regulatory relationship between mRNA/mCG was present. We then tested the individual effects of mCG loci on their mapped mRNA transcripts. mRNA/mCG read pairs exhibiting a nominally significant relationship were advanced to the mediation phase of testing. We carried out FDR thresholding on the nominally significant Arm C results to assess which pairs showed the strongest relationships with one another. Taking nominally significant mRNA/mCG read pairs from Arm

C, we tested the mediating effect of mCG on EE/mRNA relationships. For each mCG locus, average indirect effect (IDE), direct effect (DE), and total effect (TE) estimates and confidence intervals were calculated as a result of 10,000 quasi-Bayesian Monte Carlo approximations. Consistent with methodology employed in the field[40, 227], we considered an mCG locus a full mediator if the DE=0 while the IDE and TE≠ 0, or a partial mediator if the DE, IDE, and TE≠ 0. Nominally significant, full-mediators were used in downstream transcription factor binding site, transcription start site, and GO pathway enrichment analyses.

As these mediation analyses required input data to be in linear rather than logistic form, similar tests of mCH were not possible; instead, investigation of mCH marks focused on pathway-level analyses and enrichment of differential gene body mCH in DEGs. The overall workflow of the current study is visualized in Figure 4.4.

Arm A: Differential mRNA expression by EE

Within each tissue type, and for each of the ~16,500 QC'ed reads, we first assessed the effect of EE on mRNA expression (Figure 4.3, Arm A). In fitting negative-binomial distribution models, read count was normalized by total library size for each sample, and a metric of global dispersion (variance) was used to further adjust the distribution of errors[266]. Importantly, this method mirrors the underlying statistical strategy used by "edgeR"[266], the method applied by the original authors, and an approach commonly used in the field[267–269]. This custom methodology was utilized to enable downstream mediation testing through the creation of individual model objects in R, and concordance of current RNA-seq results with the original author's RNA-seq results is strong (Log Z-value vs. Log FC) Pearson's correlation: dDG = 0.76, vDG = 0.77) (Supplemental Figure C.1). Genes differentially expressed at a nominal level ($p <$

0.05) were considered DEGs due to the exploratory nature of the current work, and were advanced to the subsequent phase of the analysis (Arm C). We also carried out FDR thresholding ($q < 0.05$) and log2 FC thresholding (log2 FC > |0.25|) to identify the DEGs with the strongest relationship with EE. To conceptualize the network of altered biological processes in response to EE, we performed gene set enrichment analysis (GSEA) for nominally significant RNA-seq results in each tissue type, using the "topGO" package[270] in R. We used default settings, with the "classic" algorithm, and "fisher" statistical test. We considered a gene ontology pathway enriched at $p < 0.05$ (exploratory), or $q < 0.05$ (stringent).

Arm B: Differential mCG by EE

Within each tissue type, for each of the ~5.5 million QC'ed reads, we assessed the effect of EE on mCG by fitting linear models (LMs) of the logit-transformed proportion methylated values (Figure 4.3, Arm B). An mCG locus was considered differentially methylated by EE and nominally significant at $p < 0.05$. The loci with nominally significant relationships to EE were advanced to Arm C of the analysis. Similar to Arm A, we carried out FDR thresholding ($q < 0.05$) and percent difference (PD) thresholding (PD > |10%| between SH and EE groups) to identify the loci with the strongest associations to EE.

Arm B: Differential mCH by EE

Within each tissue type, we assessed the effect of EE on mCH by fitting generalized linear models (GLMs) of the Boolean "methylated" call of each locus (Figure 4.3, Arm B). GLMs were fit with using the binomial distribution and logit linkage. mCH loci were considered differentially methylated by EE at $p < 0.05$ and advanced to the subsequent study arm (Arm C).

We then applied FDR thresholding ($q < 0.05$) to identify mCH loci with the strongest relationships to EE. Because the metric of interest in this stage of analysis was Boolean (mCH), representative values (methylated/unmethylated) had neither variability in magnitude or spread, and therefore, we did not apply additional thresholding steps as in the case of mRNA (log2FC) or mCG (PD).

Arm C: Differential mRNA expression by DNA methylation

We compared the nominally significant DEGs from Arm A with the nominally significant DMLs from Arm B, and *only* included the intersecting subset of mCG/mCH DMLs that mapped to a DEG. This focused the analysis on DMLs with a putative cis-regulatory function on their DEG. We then tested the locus-specific effect of DNA methylation (mCG/mCH) on mRNA expression (Figure 4.3, Arm C) (e.g., we tested the effect of the differentially methylated chromosome 19, position 50,394,496 mCG locus on the differentially expressed cis-located mRNA transcript mapped to the *Sorcs1* gene). Again, we fit negative-binominal models where read count was normalized by total library size, and a metric of global dispersion was used to adjust the distribution of errors for each read. We also included EE as a covariate. Transcripts were considered differentially expressed by DNA methylation at $p < 0.05$ and these nominally significant relationships were advanced to the mediation phase of testing. In order to identify mRNA transcripts with the strongest relationships to their mapped DMLs, we applied FDR thresholding ($q < 0.05$).

Transcription factor binding site enrichment

To assess EE's effect on genomic regulatory regions, we investigated the enrichment of TFBSs for DMLs that fully mediated an EE/mRNA relationship (at nominal significance, $p < 0.05$). We performed the same analysis using mCH loci that correlated with gene expression (at nominal significance, $p < 0.05$) (Figure 4.4). Using the "chipenrich" package in R[271], we identified TFBSs for known mouse genes, and tested whether DMLs were located closer to these TFBSs than by chance. This package was originally designed to be used with chromatin-immunoprecipitation sequencing data but was adapted for use in the context of mCG/mCH loci by treating each fully mediating DML as a peak. In brief, this method calculates the distance between DMLs and TFBSs across the genome while adjusting for gene length. A Wilcoxon rank-sum test is then applied to test for differences in the adjusted distances between loci of interest and the nearest TFBSs. We corrected for FDR at $q < 0.05$ to identify the TFBSs most strongly enriched for DMLs.

Importantly, this TFBS enrichment analysis differs in regard to methodology and application compared to the TFBS analysis performed by the original authors. In their analysis, Zhang et al first identified differentially methylated regions (DMRs) within the mouse genome. Then, they assessed the +/- 200 bp range surrounding the DMR center for enrichment of specific TFBS DNA sequences, including that of *NeuroD1*. In contrast, our method identified TFBSs enriched with nominally significant fully-mediating mCG loci, or mCH loci that were nominally associated with mRNA expression of their underlying mapped gene.

Transcription start-site/promoter & gene body enrichment

To assess the effects of EE on methylation at cis-located regulatory regions, we investigated the enrichment of DEG TSSs/promoter regions compared to their gene body, with regard to nominally significant fully mediating mCG sites (Figure 4.4). Enrichment was assessed by first identifying the total number of mCG loci (whether fully mediating or not) mapped to a given gene's TSS/promoter regions, and the number of mediators within those regions. This was accomplished using the "TxDb.Mmusculus.UCSC.mm10.knownGene"[272], and "BSgenome.Mmusculus.UCSC.mm10"[273] packages in R. Then, we identified the total number of mCG loci present across the gene as a whole (exons and introns, but not including the TSS/promoter regions), as well as the number of mediators mapped to the gene. On a gene-by-gene basis, these values were then used as input to Fisher's Exact Test, where our null hypothesis was that mediating mCG loci would be evenly distributed across the TSS/promoters and gene body. Our alternative hypothesis stated that across the entirety of the gene (TSS/promoter/gene body) the probability of a mediating mCG locus being located in the TSS/promoter would be higher than in the gene body. Genes were considered enriched for TSS/promoter mediators if $p < 0.05$, and number of mediators in the TSS/promoter regions > 1. To identify the most enriched examples, we applied FDR at $q < 0.05$.

In the pool of correlated mCH/mRNA read pairs from Arm C analyses, we assessed the enrichment of individual gene bodies for differential mCH loci relative to the total pool (Figure 4.4), testing the canonical concept that gene body mCH can be associated with transcriptional repression[274]. From this analysis pool, we took the following values: total number of differential mCH gene body sites and total number of tested mCH gene body sites. Then, for each gene, we collected the number of different mCH loci in its gene body, and the overall

number of gene body mCH sites tested in Arm C. This allowed us to test enrichment of gene body mCH, for each gene using Fisher's Exact Test, against the rest of the analysis pool. Once again, to identify the most strongly enriched examples, we applied FDR at $q < 0.05$.

Biological and molecular function enrichment

We assessed the network-based gene ontology biological process (GOBP) and molecular function (GOMF) enrichment of nominally significant fully mediating mCG loci, again using the "chipenrich" package in R (Figure 4.4). The chipenrich package was utilized in this context because, due to its design for ChIPseq peak analyses, it was better suited for locus-specific DML testing. We also tested GOBP and GOMF enrichment of in mCH loci from correlated mCH/mRNA read pairs Arm C analyses (Figure 4.4). We corrected for FDR at $q < 0.05$.

Cross-species & cross-tissue mCG

In previous studies[260, 261], we investigated the relationships between ASEs, peripheral mCG clusters, and neuroimaging endophenotypes of mental health disorders. From these analyses we identified six peripheral mCG clusters comprised of 118 total mCG loci, that mediated our target relationship. These loci mapped to genes including *MYH14, DNMT3A, PGF, NRXN3, and SYT1* among others, and are involved in biological and molecular function pathways implicated in neutrophil degranulation, toll-like receptor binding, vascular endothelial growth factor binding, vesicular trafficking, and CNS cell adhesion binding. Assuming that ASE exposure and EE exist on the same spectrum with opposing directions of effect[97], we sought to investigate whether these mediating ASE-associated human loci were affected by EE in DG

tissue. We also wanted to determine whether these loci were subject to an opposing direction of effect as in the human ASE context (Figure 4.4).

The following steps were taken to create a concise mapping of human loci to mouse loci. First, we used the "biomaRt"[275] package in R to map CpG loci from the human genome (hg38)[276] to their mouse genome (mm10) orthologous counterparts, only keeping mappings with greater than 75% sequence similarity between human and mouse. Then, we considered the +/- 250 bp sequence surrounding each human CpG probe. Using default parameters of the UCSC "liftOver"[277] package in R, we generated mouse genome sequences of high similarity (+75%) to the human sequences in question. Finally, we compared the UCSC "liftOver" mappings to the "biomaRt" ortholog mappings, and kept only those with concordant chromosome, strand, gene start position, and gene end position information between the two sets of results. This created our final and singular list of mouse genome locations that exhibited high sequence similarity with, and that were orthologous with the human CpG locus locations. We then qualitatively evaluated, on a locus-specific basis, whether the direction of effect between our previous human studies and the current mouse study were in opposition to one another.

4.4 Results

Arm A: Differential mRNA expression by EE

dDG: We tested the effect of EE on mRNA expression of 16,537 dDG transcripts (Figure 4.5, Figure 4.6) and found 2,021 nominally significant DEGs (p_{range} = 0.05 to $9.7x10^{-55}$), which were carried through to Arm C analysis. We identified 161 differentially expressed when thresholding for FDR and log2 FC. The top three DEGs in response to EE included: *Igfbp6* ($log2FC_{EE}$ = 1.59, q = $1.6x10^{-50}$, rank = 1), *Penk* ($log2FC_{EE}$ = 1.39, q = $8.4x10^{-38}$, rank = 2), and

Ogn ($log2FC_{EE}$ = - 0.77, q = $1.3x10^{-24}$, rank = 3) (Figure 4.7a). Top five results are documented in Table 4.1, and comprehensive results are provided in Supplementary Table C.1. Through GSEA, we observed significant ($p < 0.05$) enrichment of 153 biological process networks composed of more than ten DEGs, including many large networks related to cellular proliferation/differentiation, sexual system development, ion transport/metabolic processes, and neuropeptide/neurotransmitter signaling (Supplementary Table C.2). Of note were 78 smaller enriched networks (between five and ten DEGs, $p < 0.05$), that largely focused on leukocyte differentiation, tumor necrosis factor signaling, and interferon gamma signaling (GO:0032720, GO:1903556, GO:2000106, GO:0001776, GO:0032609), in addition to chondrocyte development/differentiation (GO:0003413, GO:0003418, GO:0003433) (Supplementary Table C.2). No pathways were enriched at our FDR significant threshold ($q < 0.05$).

vDG: We performed the same test of 16,576 transcripts in mouse vDG tissue (Figure 4.5, Figure 4.6). We found 1,672 nominally significant DEGs (p_{range} = 0.05 to $2.5x10^{-49}$), which were carried through to Arm C analyses. We observed 131 DEGs when thresholding for FDR and log2 FC. The top three DEGs were up-regulated in response to EE and included: *Fos* ($log2FC_{EE}$ = 1.22, q = $4.2x10^{-45}$, rank = 1), *Nptx2* ($log2FC_{EE}$ = 0.81, q = $4.2x10^{-22}$, rank = 2), and *Npas4* ($log2FC_{EE}$ = 1.45, q = $2.8x10^{-19}$, rank = 3) (Figure 4.7b). The top five corrected/thresholded results are documented in Table 4.1, and comprehensive results are provided in Supplementary Table C.3. We observed significant ($p < 0.05$) enrichment of 129 biological process networks composed of more than ten DEGs, including large networks related to cellular localization and secretion, vesicular transport, neurotransmitter signaling, and synapse development (Supplementary Table C.4). Of note were the top five enriched networks, which all reached FDR significance. These included GO:0042326 (negative regulation of phosphorylation, p = $3.5x10^{-5}$,

$q = 2.4x10^{-2}$) as well as GO:0043409 (negative regulation of MAPK cascade, $p = 4.6x10^{-5}$, $q = 2.4x10^{-2}$). Of further interest were 68 smaller networks ($p < 0.05$) that focused on more specific pathways within the MAPK cascade (GO:0043407, GO:0070373), in addition to synapse/dendrite development (GO:0050805, GO:0048169, GO:0008088), and leukocyte homeostasis/apoptotic regulation (GO:0070232, GO:0070227, GO:0070228) (Supplementary Table C.4). These smaller networks were not enriched at FDR significant levels.

Arm B: Differential mCG by EE

dDG: We tested the effect of EE on locus-specific mCG at 5,445,098 positions across the genome in mouse dDG tissue (Figure 4.5). Out of 5,445,098 loci, 267,826 were differentially methylated ($p_{range} = 0.05$ to $6.1x10^{-14}$); 66,598 were mapped to known genes, 8,849 were mapped to gene promoters or TSSs, and these nominally significant/mapped DMLs were carried through to Arm C analyses. The top 1,078 survived FDR correction and PD threshold (990 gene body, 88 TSS/promoter). The top three hits mapped to the genes: *Plxna4* (Chr6:32340658, $PD_{EE} = -14.0$, $q = 3.3x10^{-7}$, rank = 1), *B3glct* (Chr5:149753329, $PD_{EE} = 16.2$, $q = 4.4x10^{-7}$, rank = 2), and *Gstt3* (Chr10:75776018, $PD_{EE} = -11.3$, $q = 7.8x10^{-6}$, rank = 3). Top five results are documented in Table 4.2, and comprehensive results are provided in Supplementary Table C.5.

vDG: We performed the same analysis on 5,520,790 positions across the genome in vDG tissue (Figure 4.5). Out of 5,520,790 loci, 225,311 were differentially methylated at a nominal level ($p_{range} = 0.05$ to $1.6x10^{-14}$). Of these, 54,985 loci mapped to known gene bodies and 6,707 mapped to gene promoters or TSSs. These nominally significant/mapped DMLs were carried through to Arm C analyses. In all, the top 784 loci survived FDR correction and PD thresholding; the top two hits were observed to have lower methylation in response to EE, and

the top three mapped to genes: *Mettl15* (Chr2:109243084, PD_{EE} = -10.2, q = $8.7x10^{-8}$, rank = 1), *Nubp1* (Chr16:10411612, PD_{EE} = -10.7, q = $1.7x10^{-5}$, rank = 2), and *Prss48* (Chr3:86001744, PD_{EE} = 16.5, q = $2.4x10^{-5}$, rank = 3). Top five results are documented in Table 4.2, and comprehensive results are provided in Supplementary Table C.6.

In post-hoc analyses, by comparing DML location between dDG and vDG using Fisher's Exact test, we found that the dDG is significantly enriched with TSS/promoter DMLs compared to vDG (odds ratio = 1.09, 95% CI = 1.05 to 1.13, p = $6.3x10^{-7}$).

Arm B: Differential mCH by EE

dDG: We tested the effect of EE on mCH in mouse dDG. We performed tests at 740,180 loci (Figure 4.6) and found 114,742 to be differentially methylated (p_{range} = 0.05 to $1.9x10^{-4}$). Of these, 25,880 were mapped to known gene bodies, whereas 1,706 were mapped to gene promoter regions or TSSs. These mapped/nominally significant DMLs were carried through to the following analysis arm. No loci survived FDR/PD thresholding due to the nature of the Boolean "methylated" calls used in our analysis (e.g., *Phactr2:* chr10:13373248, $SH_{n\ methylated\ calls}$ = 0, $EE_{n\ methylated\ calls}$ = 5, PD_{EE} = 100, p = $1.96x10^{-4}$, q = 0.64). Comprehensive results are reported in Supplementary Table C.7.

vDG: We performed the same tests in vDG on 1,023,564 mCH positions across the genome (Figure 4.6). Of these 1,023,564 loci, 174,711 were differentially methylated (p_{range} = 0.05 to $1.9x10^{-4}$). Within these ~175k loci, 37,486 were mapped to known genes, 2,386 were mapped to gene promoters or TSSs and all nominally significant and mapped mCH loci were advanced to Arm C analysis. None survived FDR/PD thresholding for reasons stated above. Comprehensive results are reported in Supplementary Table C.8.

In post-hoc analyses, by comparing mCH DML location between dDG and vDG using Fishers' Exact test, we found no difference when assessing the proportion of DMLs located in gene bodies (odds ratio = 1.04, 95% CI = 0.97 to 1.10, p = 0.29).

Arm C: Differential mRNA expression by mCG

dDG: Using only RNA-seq and mCG reads deemed differentially expressed and methylated by EE at the nominal threshold, and taking only DMLs which mapped to a DEG, we carried out 10,425 hypothesis tests of mCG's effect on mRNA expression, while including EE as a covariate (Figure 4.5). In all, a total of 10,425 DMLs across 1,367 DEGs were tested. 987 loci were nominally significant predictors of mRNA transcription as a function of mCG (p_{range} = 0.05 to $1.0x10^{-12}$) (Figure 4.8a), and were carried through to the mediation phase of our analysis. In correcting for multiple hypothesis testing bias at FDR q < 0.05, we found 86 DMLs that were significantly predictive of their mapped transcript's expression. The top five FDR significant results are documented in Table 4.3, and comprehensive results are provided in Supplementary Table C.9.

vDG: Carrying out the same strategy in mouse vDG, we tested 6,006 hypotheses of mCG's effect on mRNA expression (Figure 4.5). A total of 1,005 DEGs were represented in this phase of the analysis. We found 827 mCG loci to be nominally significant predictors of mRNA expression at their cis-located transcript (p_{range} = 0.05 to $8.9x10^{-12}$) (Figure 4.8b). These nominally significant mRNA/mCG read pairs were carried through to the subsequent mediation testing phase. We corrected for multiple hypothesis testing bias at FDR q < 0.05 and found that 127 mCG loci from vDG were FDR significant predictors of their mapped transcript's

expression. The top five FDR significant results are documented in Table 4.3, and comprehensive results are provided in Supplementary Table C.10.

Arm C: Differential mRNA expression by mCH

dDG: Using only RNA-seq and mCH reads deemed differentially expressed and methylated by EE at the nominal threshold ($p < 0.05$), and taking only DMLs which mapped to a DEG, we carried out 4,440 hypothesis tests of mCH's effect on mRNA expression of 779 unique transcripts (Figure 4.6). EE was again included as a covariate in all tested models. 472 mCH loci were significant predictors of their cis-located transcripts expression ($p_{range} = 0.05$ to $2.4x10^{-9}$) (Figure 4.8c) and were utilized in downstream enrichment testing analyses. In applying an FDR threshold at $q < 0.05$, we identified that ten mCH DMLs were FDR significant predictors of their mapped transcript mRNA expression. Top five FDR significant results are documented in Table 4.4, and comprehensive results are provided in Supplementary Table C.11.

vDG: Using the same strategy in vDG, we tested the effect of 4,058 individual mCH loci on their 640 cis-located mRNA transcripts (Figure 4.6). 628 mCH loci were nominally significant predictors ($p_{range} = 0.05$ to $6.8x10^{-16}$) (Figure 4.8d) and were utilized in downstream enrichment testing analyses. One hundred and four mCH loci were FDR significant predictors of their cis-located transcripts expression. Top five FDR significant results are documented in Table 4.4, and comprehensive results are provided in Supplementary Table C.12.

Mediation: mCG mediates EE/mRNA relationships

dDG: We performed 10,000 Monte Carlo simulations per mCG locus to test the mediating effects of mCG between EE and differential mRNA expression of cis-located

transcripts (Figure 4.5). Of 987 mCG loci (mapping to 533 individual genes) that satisfied nominal Arm C requirements, 404 were nominally significant full mediators of their EE/mRNA relationships, and 502 were partial mediators. These 404 nominally significant full mediators were used in downstream enrichment analyses for mCG. In applying FDR at $q < 0.05$ to IDE pvalues, we found that 17 mCG loci were FDR significant, full mediators of an EE/mRNA relationship. Twenty mCG loci were FDR significant, partial mediators. Top five results are presented in Table 4.5; comprehensive results are provided in Supplementary Table C.13. Of note are two fully mediating mCG probes located in the *Smtnl2* (smoothelin-like 2: Chr11:72394459, Chr11:72403429) gene, as well as full mediators within the *Myo5b* (myosin VB: Chr18:74659297) and *Tnfrsf19* (tumor necrosis factor receptor superfamily, member 19: Chr14:60967710) genes.

vDG: Performing the same tests of mediation on 827 mCG loci across 396 mRNA reads whose relationships satisfied nominal Arm C requirements (Figure 4.5), we found 211 were nominally significant full mediators of an EE/mRNA relationship, while 255 were partial mediators. These 211 nominally significant full mediators were used in downstream enrichment analyses specific to vDG. Out of this pool, we also identified two mCG loci that were FDR significant, full mediators, and four mCG loci that were FDR significant, partial mediators. Top five results are documented in Table 4.5, and comprehensive results are provided in Supplementary Table C.14. Of note was a fully mediating probe in the *Ror1* (receptor tyrosine kinase-like orphan receptor 1: Chr4:100390763) gene locus, along with two partially mediating probes in the *Myo5b* gene (myosin VB: Chr18:74556187, Chr18:74576505).

In comparing the proportion of fully mediating mCG loci between dDG and vDG, we found that dDG was significantly enriched (odds ratio = 1.60, 95% CI = 1.32 to 1.95, $p = 9.1x10^{-7}$) for full mediators compared to vDG.

Transcription factor binding site enrichment

dDG: To assess EE's effect on genomic regulatory regions, we investigated the enrichment of TFBSs across the entirety of the mouse genome for nominally significant fully mediating mCG loci in dDG tissue (Figure 4.5). We found 16 TFBSs to be positively enriched, including OCT_C (p = 0.001, rank = 1), TST1_01 (p = 0.003, rank = 2), and NKX25_02 (p = 0.005, rank = 3). In correcting these TFBS results for multiple hypothesis testing bias, none surpassed the FDR threshold. Because we could not assess the mediating status of mCH loci, we tested TFBS enrichment for mCH DMLs which correlated with gene expression from Arm C analyses (Figure 4.6). We found 17 enriched TFBSs, including ARNT_02 ($p = 4.6x10^{-5}$, $q = 1.0x10^{-2}$, rank = 1), TST1_01 ($p = 2.8 x10^{-4}$, $q = 3.1x10^{-2}$, rank = 2), and CREBP1_Q2 ($p = 5.8x10^{-4}$, $q = 4.3x10^{-2}$, rank = 3), all of which were significant at FDR levels. Top enriched dDG TFBSs from Arm C mCH and from the mCG full mediation pools are presented in Table 4.6.

vDG: In fully mediating mCG vDG loci (Figure 4.5), we found 16 positively enriched TFBSs. The top three TFBSs were FDR significant, and included POU3F2_02 ($p = 2.2x10^{-5}$, $q = 4.6x10^{-3}$, rank = 1), STAT5A_02 ($p = 4.1x10^{-4}$, $q = 3.7x10^{-2}$, rank = 2), and LHX3_01 ($p = 5.5x10^{-4}$, $q = 3.7x10^{-2}$, rank = 3). In mCH DMLs (Figure 4.6), we found 18 enriched TFBSs, including FOXO3_01 (p = 0.002, rank = 1), POU3F2_02 (p = 0.002, rank = 2), and LMO2COM_02 (p = 0.005, rank = 3). None reached FDR significance. Top enriched TFBSs are presented in Table 4.6.

Transcription start-site/promoter & gene body enrichment

dDG: In assessing genes with an mRNA expression/EE relationship fully mediated by an mCG locus (Figure 4.5), we found only one out of 242 that were nominally enriched for TSS/promoter region DMLs: *Npas2* (Neuronal PAS Domain Protein 2: odds ratio = 15.9, 95% CI = 1.6 to Inf, p = 0.021). *Npas2* promoter enrichment did not reach FDR significance. No genes were significantly enriched for correlated mCH loci in their gene bodies compared to the rest of the analysis pool (Figure 4.6).

vDG: Out of the 136 DEGs with a fully mediating mCG locus (Figure 4.5), none had enriched TSS/promoter regions, by our current analysis. Again, no genes were significantly enriched for correlated mCH loci in their gene bodies compared to the rest of the analysis pool (Figure 4.6).

Biological and molecular function enrichment

dDG: We assessed the enrichment of gene ontology biological process and molecular function pathways for 404 fully mediating DMLs in dDG tissue (Figure 4.5). We found numerous nominally enriched biological process pathways including myeloid leukocyte migration (GO:0097529, p = 0.003, rank = 1), divalent inorganic cation homeostasis (GO:0072507, p = 0.004, rank = 2), and molecular mediator of immune response (GO:0002440, p = 0.004, rank = 3), although none were FDR significant. Enriched molecular function pathways included steroid binding (GO:0009416, p = 0.009, rank = 16), sterol binding (GO:0048565, p = 0.009, rank = 17), and guanyl nucleotide binding (GO:0055123, p = 0.009, rank = 19). No pathways reached FDR significance. Biological process and molecular function pathways enriched in dDG for correlated mCH loci (Figure 4.6) included: osteoblast

differentiation (GO:0001649, $p = 8.8x10^{-6}$, rank = 1), steroid hormone receptor activity (GO:0003707, $p = 1.8x10^{-3}$, rank = 31), and RNA polymerase II transcription factor activity (GO:0004879, $p = 1.8x10^{-3}$, rank = 32), among others (Table 4.7). Of these, only osteoblast differentiation (GO:0001649) was significant at FDR < 0.05.

vDG: Biological process pathways enriched for the 211 fully mediating DMLs from vDG tissue (Figure 4.5) included regulation of peptide secretion (GO:0002791, $p = 2.8x10^{-5}$, rank = 1), protein localization to plasma membrane (GO:0072659, $p = 1.3x10^{-4}$, rank = 2), and regulation of cell morphogenesis (GO:0022604, $p = 2.4x10^{-4}$, rank = 3), which all reached FDR significance. Molecular function pathways enriched for mediating DMLs included protein serine/threonine kinase activity (GO:0004674, $p = 2.6x10^{-4}$, rank = 11), Rac guanyl-nucleotide exchange factor activity (GO:0030676, $p = 6.6x10^{-4}$, rank = 25), and calmodulin binding (GO:0005516, p = 0.015 , rank = 95) (Table 4.7). In vDG, the top pathways enriched for correlated mCH loci (Figure 4.6) included many related to regulation of potassium ion transport (GO:0043266, $p = 9.1x10^{-5}$, rank = 1). The top two enriched biological processes reached FDR significance. In addition, nine terms related to microtubule/cytoskeletal organization were observed (e.g., GO:0000226, p = 0.015, rank = 66) (Table 4.7) at nominally significant levels.

Cross-species & cross-tissue mCG

We sought to investigate whether ASE-associated human CpG loci were affected by EE in mouse DG tissue, and whether EE and ASEs differed in their direction of effect on these loci. Out of 118 CpGs previously associated with ASEs[260, 261], 46 human CpG loci mapped to 27 orthologous mouse genes (Supplementary Table C.15).

Mouse dDG: In dDG, we identified 101 mouse mCG loci mapped to 19 human CG loci (Supplementary Table C.16). Of these, we identified a nominally significant DML in the promoter/CpG island of the *Rbm46* gene (Chr3:82876701, PD_{EE} = - 13.3, p_{EE} = 0.030, q_{EE} = 0.896, rank = 124,244). This single mouse locus mapped to six human loci which also reside in the promoter/CpG island of the *RMB46* gene. One other locus, in a *Serinc5* intron (Chr13:92693210: PD_{EE} = - 21.1, p_{EE} = 0.015, q_{EE} = 0.896, rank = 79,210), was differentially methylated (Table 4.8). As noted by the negative PD values In both loci, we observed that EE-group samples had lower mCG than SH-group samples.

Human saliva: In looking to our prior research on peripheral epigenetic signatures that mediate the relationship between childhood family emotional health (FEH) and neural endophenotypes of mental health disorders[260], we found that high FEH was positively associated with high mCG of the Thistle2 cluster (b = 0.003, SE = 0.001, t = 2.2, p = 0.027). Put another way, participants with poor FEH had lower Thistle2 mCG than participants with high FEH. In addition, Thistle2 mCG was positively associated with right hippocampal volume (b = 728.1, SE = 261.8, t = 2.8, p = 0.007). The same positive direction of effect was present when observing mCG values of the individual probes that made up the Thistle2 cluster. A prime example was the *RBM46* Thistle2 locus cg01289218 (b = 0.012, SE = 0.005, t = 2.7, p = 0.007), which was also positively associated with childhood FEH.

Mouse dDG vs. human saliva: Overall, we found two EE-associated DMLs in mouse dDG, and we observed that for both loci, EE-group samples had lower mCG than SH-group samples. In the corresponding probes in human saliva, we found that participants with poor FEH had lower mCG than those with high FEH. It appears, then, that the direction of effect between

mouse mCG and human mCG loci between EE and low FEH is congruent, which fails to support our hypothesis.

Mouse vDG: We performed the same analyses in mouse vDG tissue and identified 101 mouse mCG loci mapped to 18 of the 118 human CpG loci of interest (Supplementary Table C.17). Of these mouse vDG mCG loci, none were differentially methylated ($p>0.05$) (Table 4.8).

4.5 Discussion

The current study examined the effects of EE on mRNA transcription, mCG, and mCH in dDG and vDG from mouse. We sought to determine whether mCG or mCH, on a *locus-specific* basis, putatively causally mediated the relationship between EE and mRNA transcription, in order to elucidate underlying epigenetic mechanisms responsible for the biological embedding of EE. We hypothesized that methylation at a number of mCG and mCH sites would mediate their EE/mRNA transcription relationships, and that the TSS/promoter regions and TFBSs underlying DEGs would be enriched with these mediating loci. Based on previous work[260, 261], we also hypothesized that the network of fully mediating mCG/mCH loci would be enriched with gene ontology relating to: synapse development, neurotransmitter signal transduction, neuronal development, inflammation, immune cell signaling and migration, cytoskeleton organization, MAPK cascade, and regulation of phospholipase activity. Finally, we hypothesized that ASE-associated DMLs from previous human peripheral tissue studies would *also* be differentially methylated in rodent CNS tissue in association with EE exposure, but with opposite direction of effect.

In regard to our Arm A analysis of differential mRNA expression in dDG and vDG tissue, we find that our results largely correspond to those of the original authors in regard to

direction and magnitude of effect (Supplemental Figure C.1). Furthering the continuity between studies is widespread differential gene expression across both subregions of the dentate gyrus, but with more differential gene expression in dDG relative to vDG (+21%). Novel findings not previously discussed in past analyses on this dataset include observations that the MAPK cascade is likely subject to targeted downregulation in vDG. This position is evidenced by enrichment of GO pathways including: negative (down) regulation of phosphorylation, negative regulation of kinase activity, negative regulation of ERK1 and ERK2 cascade, and negative regulation of MAPK cascade. Of note, these terms were not observed in the dDG, where immune signaling and cellular differentiation/proliferation pathways were enriched. A well-established driver of eukaryotic signal transduction, the MAPK cascade has been implicated in mental health outcomes and the development of synaptic plasticity in mature neurons. To this end, modulation of the MAPK cascade in mouse forebrain is associated with both anxiety-like and depressive-like behaviors[180], and inhibition of the MAPK cascade reduces hippocampal long-term potentiation[278]. Therefore, the negative (down) regulation of these pathways as observed in this work stand in opposition to the expected upregulation of the MAPK cascade in response to EE. However, previous research shows that the MAPK cascade is subject to temporally specific, multi-modal negative feedback loops in response to an initial upregulation[279]. It appears, then, that the aforementioned behavioral and neurobiological effects of EE could be driven strongly through transcriptional regulation of the MAPK cascade[279], and that observations of the molecular mechanisms in vDG made at p80 capture negative feedback signals operating on the MAPK cascade after a previous, unobserved increase *potentially* due to EE. This position was supported by post-hoc analyses showing significant upregulation of the dual-specificity MAPK phosphatases *Dusp1* and *Dusp5* (both of which survived FDR correction and FC thresholding),

which serve as nuclear proteins that dephosphorylate ERK1/2, p38, and JNK, and downregulate the MAPK cascade[280]. Furthering the contrast between dorsal and ventral regions of the dentate gyrus are the enriched transcriptional pathways relating to neuronal maturation (synapse development, dendrite development, vesicular transport) in the vDG. Taken together with signals from the dDG showing cellular differentiation/proliferation pathway enrichment it appears that differential expression in the dDG *could be* representative of a proliferating, immature neuron-like phenotype, whereas expression in the vDG is representative of a more mature neuron-like phenotype which has shifted focus towards dendrite and synapse development mediated through MAPK the signaling cascade.

Regarding mCG, a similar number of loci were tested for differential methylation between dDG and vDG (~5.5 million). However, nineteen percent more DMLs were observed in dDG compared to vDG, and dDG genes were specifically enriched for DMLs in their promoter regions relative to vDG. This distinct difference in EE-associated transcriptional regulatory mechanism alterations aligns with the enrichment of DEGs in dDG tissue, and with previously reported results from this dataset. However, when comparing differential mCH between dDG and vDG, we found the opposite pattern as in mCG-- more differential mCH occurs in the vDG compared to the dDG. Previously published next-generation sequencing data from human and rodent cortex shows that mCH increases in maturing neuron populations through the young adult stages of development, and that mCH is the dominant form of methylation in human neurons[274]. This body of work, in tandem with the observed synapse development pathway enrichment and heightened mCH methylation in vDG compared to dDG in this study, again bolster the notion that neuronal maturation could occur preferentially in the vDG throughout the p22 to p80 stage of development in mouse, whereas neurogenesis occurs throughout this time

period in the dDG. Interestingly, we found evidence that the transforming growth factor beta (TGF-B) pathway, canonically thought to oppose the MAPK cascade in regard to cellular differentiation, was transcriptionally downregulated in vDG (*Tgfb2*, *Tgfbr1*, *Tgfa*). We posit this transcriptional downregulation is possibly achieved via mCH, due to the observed mRNA/mCH relationships from the current study (*Tgfb2* chr1:186640735/chr1:186672797, *Tgfa* chr6:86201987, *Tgfbr1* chr4:47415330). Therefore, we posit that in the vDG, TGF-B signaling and MAPK signaling are regulated through distinct epigenetic mechanisms.

We found that dDG is significantly enriched for fully mediating mCG loci compared to vDG, indicating that the ultimate transcriptional effects of EE are potentially exerted more so through mCG in dDG than in vDG. This implies that other epigenetic mechanisms may be playing a greater role in vDG compared to dDG. Our results show ten times more FDR significant correlated mRNA/mCH read pairs in vDG (104 pairs at $q < 0.05$) compared to dDG (ten pairs at $q < 0.05$) and imply that mCH could be operating in this role. However, due to statistical barriers, we were unable to specifically investigate mCH mediation. Other epigenetic mechanisms such as hydroxy-methylation, and post-translational histone modifications could also be considered. Out of the top ten mCG mediators in each of dDG and vDG, only one locus is located in a TSS/promoter region, *Card14* (Caspase Recruitment Domain Family Member 14). *Card14* encodes a protein that plays a role in cellular adhesion and NF-kappa-B activation; and mutations in this gene are a known mechanistic contributor, when assessed in epithelial cells, of psoriasis and psoriatic arthritis[281]. Interestingly, *Card14* is upregulated in post-mortem hippocampus from humans with Alzheimer's disease[282], relative to "healthy" age-matched controls. Paradoxically, *Card14* promoter locus mCG is downregulated in EE-group samples, in tandem with transcriptional upregulation, as observed in dDG.

In assessing TFBS enrichment of fully mediating loci, we identified a family of 8 genes that were enriched with 21 fully-mediating mCG loci located within *Nkx2-5* TFBSs. Canonically understood to play a major role in heart and spleen formation/development[283], the *Nkx2-5* gene is a homeobox-domain containing transcription factor that also functions to regulate heart progenitor cell differentiation and proliferation[284]. What is more, targeted over-expression of Nkx2-5 transforms in vitro human fetal myoblasts into neuron-like cells with noted expression of neuronal markers[285]. Researchers posit that when *Nkx2-5* acts as a transcription factor in CNS tissue, it interacts with *Adora1*, the adenosine A1 receptor[286], which is co-expressed in heart tissue. However, little else is known regarding *Nkx2-5* function in the CNS. In finding that *Nkx2-5* TFBFs were enriched with fully-mediating mCG loci, our results indicate that *Nkx2-5* could be involved the epigenetically-mediated biological embedding of EE in dDG. Equally implicated in this *Nkx2*-5-related pathway of biological embedding are the genes to which the fully mediating mCG loci belong, including *Dab1* (DAB Adaptor Protein 1; five mediating loci), and *Rora* (RAR Related Orphan Receptor A; six mediating loci). *Dab1*, and the previously mentioned *Card14,* are targets of the nuclear steroid hormone receptor and transcriptional activator *Rora*, which also plays a key role in the differentiation of neuronal cell types in the brain[287]. Just as the case with *Card14, Dab1* and *Rora* transcription are markedly upregulated in the hippocampus of deceased Alzheimer's patients[282]. Single nucleotide polymorphisms identified within the human *RORA* gene are associated with post-traumatic stress disorder (PTSD) [225], depressive symptoms[288, 289], depression vulnerability[290], aberrant internalizing of distress behaviors[291], and sex-specific anxiety sensitivity[292]. *RORA* has also been identified as a sex-specific candidate gene in Autism spectrum disorder (ASD)[293], with mCG playing a notable role in *RORA* mRNA expression in the brain of deceased ASD patients[294]. One

mediating *Rora* mCG locus stands out (Chr9:69304549), as it is located in the vicinity of the *Rora* TSS, in a region particularly enriched for cis-regulatory elements and an mCG island. In our analyses, *Rora* gene expression is significantly lower in EE-condition mice compared to SH-condition mice, and increased mCG of this locus is responsible for an estimated 86% of that effect. The identification of this largely mediator-enriched *Nkx2-5* TFBS family could shed light on the potentially related mechanistic underpinnings of these disorders in humans by highlighting locus-specific targets of epigenetic regulation in response to EE in the brain.

We also found that genes with *Pou3f2* TFBSs were enriched with both fully mediating mCG loci and mRNA-correlated mCH loci in vDG. *Pou3f2* (POU class 3 homeobox 2) is a protein coding gene that produces a neural transcription factor that is important in neural differentiation[295]. Our findings of EE-induced epigenetic regulation at *Pou3f2* target sites lend support to previous research showing that *Pou3f2* is critical for hippocampal-dependent cognitive function and adult neurogenesis in the dentate gyrus[296]. Of the fully mediating mCG loci located within *Pou3f2* TFBSs, five map near (but > 1000 BP from) the *Kalrn* (Kalirin RhoGEF Kinase) gene promoter region, with three of these mapping to a cis-regulatory element with a distal-enhancer-like signature (ENCODE Accession: EM10E0622110). The *Kalrn* gene encodes a protein playing critical roles in CNS excitatory synapse development[297, 298]. Out of the mCH pool, on the other hand, an additional ten loci map to a *Pou3f2* TFBS and to the aforementioned *Cdh13* (Cadherin 13) gene, which encodes a protein that acts as a negative regulator of neurite outgrowth during cellular differentiation[299]. The majority of these probes map to the gene body of *Cdh13,* are positively associated with EE, and are negatively related to expression of the *Cdh13* transcript. This aligns with the canonical understanding that gene body mCH is associated with transcriptional repression[274].

One of the major findings from analyses performed on this data by the original authors was the noted enrichment of dorsal mCH DMR's for *NeuroD1* TFBSs. *NeuroD1* is a transcription factor known to contribute to the early differentiation and maturation processes of granule neurons in the DG[300]. They also noted that *NeuroD1* transcription was not altered in response to EE in either dDG or vDG. In contrast to Zhang et al 2018, the current study did not assess the effects of EE on potential DMRs, or sequence-specific enrichment of TFBS within said DMRs. However, in contrast to the results from the original authors, we did observe nominal downregulation of the *NeuroD1* transcript in vDG in tandem with nominal differential methylation of four *NeuroD1* mCG loci. One of these, Chr2:79,459,144, is methylated at a higher level in EE than in SH samples and is located near the *NeuroD1* TSS and a *Pax6* TFBS. Our results suggest that differential expression of NeuroD1 in response to EE may be regulated by locus-specific mCG but not mCH in the vDG.

In previous human studies, we focused on investigating epigenetically-mediated relationships between ASEs and neural endophenotypes of mental health disorders. Through these earlier investigations, we found epigenetic signatures in saliva and blood that were associated with ASEs and that did mediate the relationships in question. These pathways were largely related to: synapse development, neurotransmitter signal transduction, neuronal development, inflammation, immune cell signaling and migration, cytoskeleton organization, the MAPK cascade, and regulation of phospholipase activity. These findings were intriguing, in light of their relation to CNS function, but the direct relationship with CNS molecular measures were unknown. Leveraging the EE paradigm in the current study, we qualitatively assessed overlap between mediator enriched pathways in the mouse dentate gyrus, and ASE-associated pathways in human peripheral tissues. In assessing the biological process and molecular function pathways

enriched for mediating mCG loci and mRNA-correlated mCH loci in mouse dentate, we find signatures of epigenetic regulation that largely mirror results from our RNA-seq analysis in mouse, and our previous work with human ASEs. These results lend support to the notion that differential methylation at these loci could be altering downstream, transcription-related processes. The dDG is enriched for fully-mediating mCG loci in a number of pathways focused on the migration or differentiation of myeloid precursor cells, as well as stem cell maintenance. Additionally, we observed numerous microtubule organization and osteoblast differentiation pathways enriched with mRNA-correlated mCH loci. We posit that our observation of osteoblast differentiation in this context is actually representative of signaling through the TGF-B pathway *towards* hippocampal neuron maturation. In brief, certain members of the TGF-B signaling ligand family, such as bone morphogenetic proteins, play marked roles in osteoblast differentiation[301] and specialized roles in hippocampal neurons critical to dendrite formation[302], axon guidance[303], and synaptic stability[304]. These processes require dynamic remodeling of both actin and microtubule assemblies[305], and here we provide putative evidence that mCH could regulate them, through TGF-B signaling, in dDG. Results from dDG stand in contrast to the epigenetic signatures identified in vDG mCG. Some of the stronger signals from the vDG included peptide/hormone secretion and ion transport, but there was also notable enrichment of pathways dedicated to synaptogenesis, synaptic plasticity, dendritic spine development, glutamatergic signaling, and protein serine/threonine kinase activity. The latter plays a major role in the MAPK cascade[306], further bolstering our position that mature neurons in the vDG are under influence by the MAPK cascade, which in turn is likely under epigenetic regulation by mCG.

Overall, our results showed CNS-derived mRNA, mCG, and mCH signatures correspond closely with enriched pathways from human peripheral tissue, with pointed focus towards the MAPK and TGF-B signaling pathways. The concordance of enriched pathways between species/tissue types lends support to the notion that a peripheral epigenetic signature could be considered a proxy of effects exerted within CNS tissue. However, when assessing the locus-specific differential methylation and expression of human ASE-associated loci in mouse, the picture is less clear. A relatively small number of CpG loci mapped between the human and mouse genome with high enough sequence specificity to consider for analyses. Of those that did accurately map in our dDG analysis, we observed only two nominally significant DMLs in the *Serinc5* and *Rbm46* genes, and they showed the same direction of effect in relation to EE and poor FEH. We posit that the observed lack of corresponding DMLs between human and mouse could be due to a number of reasons, including differences in human and mouse physiology, the well-documented tissue-specific and cell type-specific nature of mCG[307, 308], or divergent function of genes between the two species. Although clear limitations exist for cross-species, cross-tissue analyses, it is our position that continued assessment of these relationships are integral to future forward movement of the field.

The foremost strengths of the current work were our ability to test enrichment of genomic regulatory regions for loci that putatively causally mediate the EE/mRNA relationship, and our conceptualization of results at both the locus-specific and network-based level; our causal mediation framework enabled us to target analyses to loci with the greatest potential for functional downstream repercussions. Additionally, we performed the analysis in the etiological tissue of interest, a clear advantage over most human epigenetic studies that are typically limited to peripheral tissues such as saliva or blood. This approach also allowed us to explore CNS-

based measures in loci implicated in previous human work[260, 261], to evaluate whether measures of adverse social context indirectly linked to CNS measures in humans would show expected (i.e. opposite) directions of effect in an animal model that directly measured CNS tissue in relation to positive social context (i.e. environmental enrichment). Limitations of the current work include the inability to clearly delineate the proportion of cell types included in samples. Gene expression profiles correlated most strongly with neuron, but other cell type transcriptional patterns were also correlated. Additional limitations include the absence of other epigenetic measures, such as post-translational histone modifications, the inability to compare transcription/epigenetic data to the original authors' mouse hippocampal volume and immunohistochemistry measurements, and the omission of RNA-seq and MethylC-seq data without full representation across all ten biological replicates.

4.6 Conclusion

The current study showed that, in support of prior literature and analyses performed on this dataset, EE affects distinct transcriptional networks when comparing dorsal and ventral dentate subregions. Additionally, in congruence with previously published analyses of the current dataset, the dDG is subject to a greater degree of locus-specific mCG-based epigenetic regulation, and the higher overall level of locus-specific differential mCG corresponds with a heightened number of mediating mCG loci. On the other hand, vDG transcription appears subject to greater epigenetic regulation through mCH. Our findings support the notion that during adolescence/young adulthood in mice (p22 to p80), the dorsal subregion is enriched for heightened neurogenesis compared to the ventral subregion, and that a major regulatory force in this process is locus-specific mCG. In contrast, it appears that the vDG is more finely tuned

toward the maturation of neurons and then strengthening of synaptic connections, with epigenetic regulation provided through relatively higher levels of differential mCH. Additionally, it appears that the MAPK and TGF-B signaling cascades could be operating in response to EE in spatio-temporally specific patterns within the dentate gyrus, with epigenetic regulation of these pathways observed through divergent epigenetic mechanisms (mCG and mCH respectively). We identified mediating loci existing in TFBSs of genes critical for progenitor cell differentiation, hippocampal-dependent cognitive function, and adult neurogenesis in the dentate gyrus. The mediator-enriched TFBS gene families provided strong supporting evidence for the roles of *Dab1*, *Rora*, *Card14*, *Kalrn*, and *Cdh13* in hippocampal developmental and mental health disorder etiology, and our research shed light on the potentially related epigenetic underpinnings of these disorders in humans. Finally, we identified that a number of loci that mediate the relationship between poor FEH and neural endophenotypes of mental health disorders, were differentially expressed and differentially methylated at nominal levels in response to EE, although the direction of effect in the observed DMLs was unexpectedly congruent between EE and poor FEH. Overall, the findings of the current study extend the body of work highlighting neurodevelopmental variability putatively caused by EE and inform divergent spatio-temporal molecular mechanisms of biologic embedding in the dDG vs. vDG on a locus-specific basis. These findings could be used in the future to drive pointed locus-specific investigations of hippocampal development or mental health disorder etiology in humans.

4.7 Figures and Tables

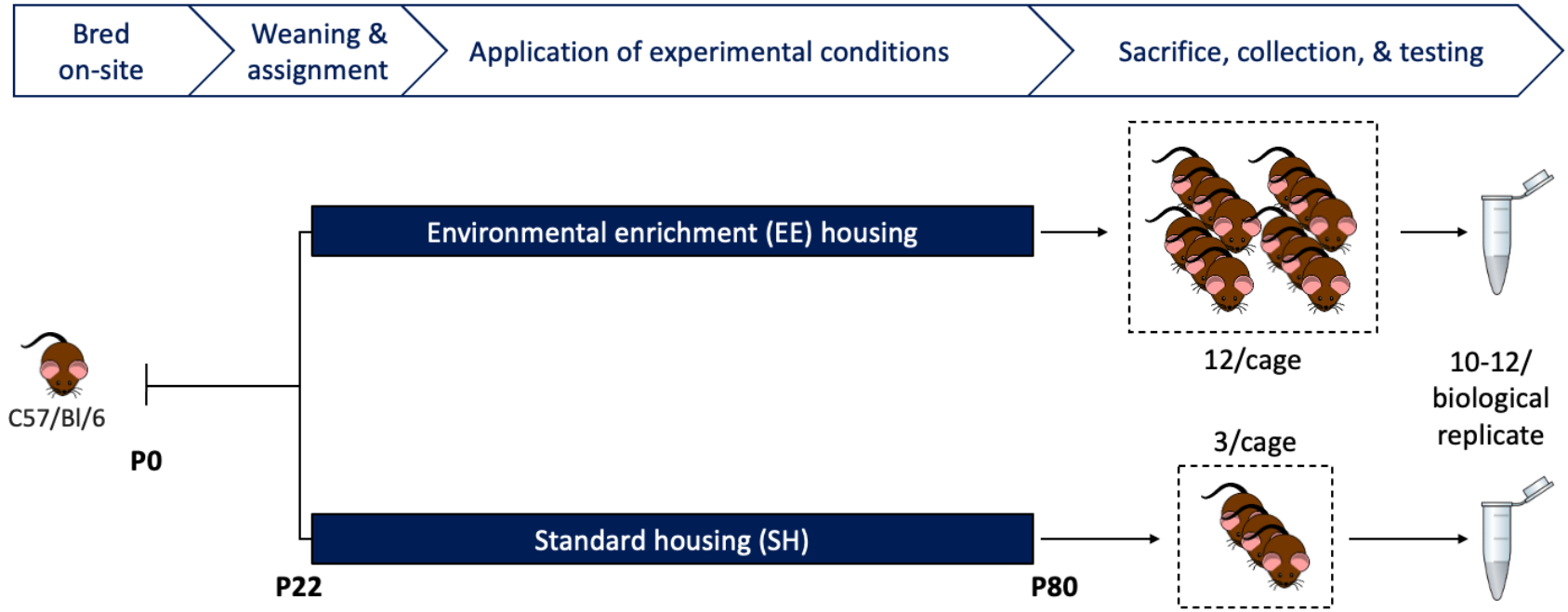


Figure 4.1 Zhang et al 2018 overall study design. Animals were bred in-house and weaned at P22. Siblings were separated and assigned to either the EE or SH conditions, in groups of ten to 12 or three respectively. Animals stayed in their respective housing conditions until P80, whereupon they were sacrificed for tissue collection. Importantly, ten to 12 mice (technical replicates), were used per sample (biological replicates) generated for sequencing. This schema facilitated greatly increased sequencing depth, compared to other experimental designs.

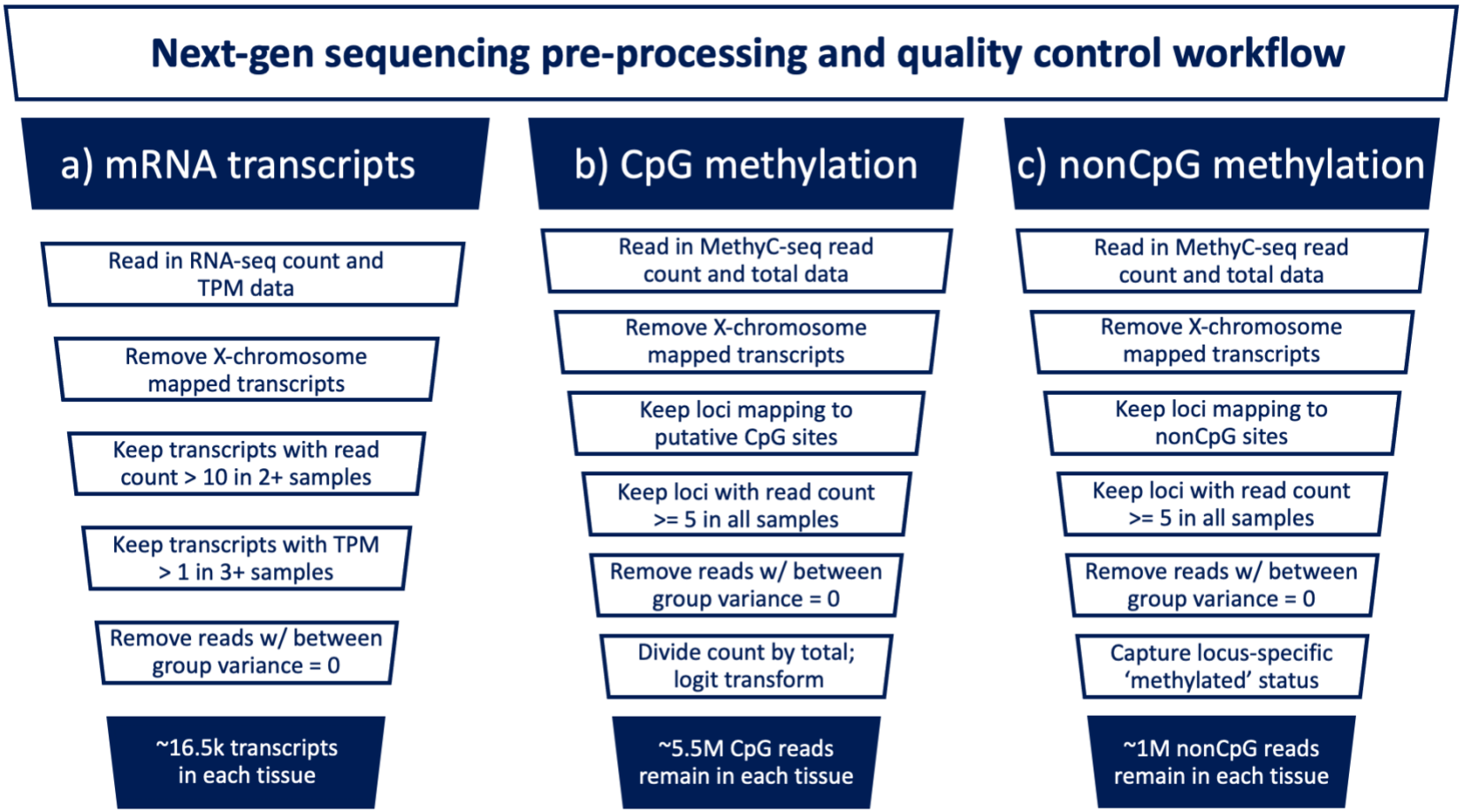


Figure 4.2 Visual representation of major pre-processing and quality control steps taken across next-generation sequencing molecular measures. Across molecular data types, we removed sex-chromosome specific reads, removed reads without sufficient sequencing depth, and removed reads with 0 variance between SH and EE groups. A. RNA-seq pre-processing and quality control. B. MethylC-seq pre-processing and quality control of CpG methylation sites. C. MethylC-seq pre-processing and quality control of non CpG methylation sites.

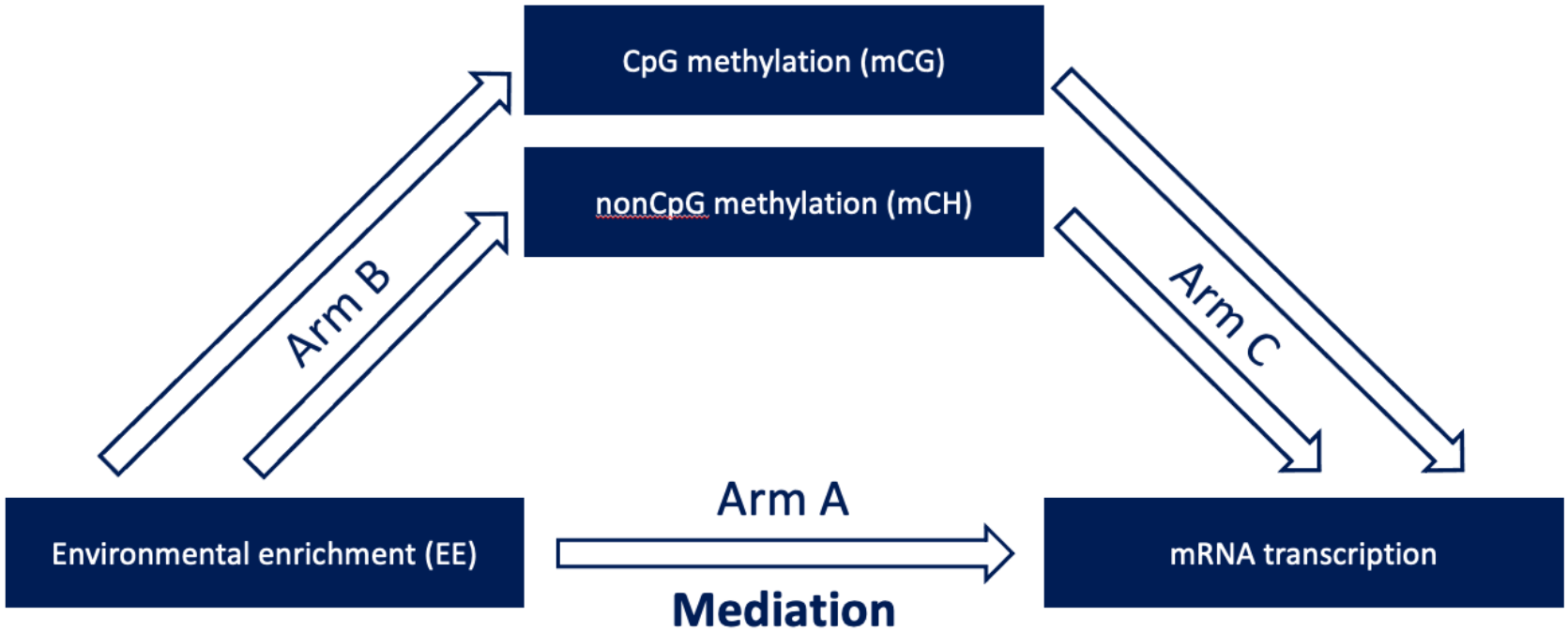


Figure 4.3 Locus-specific mediation testing framework. Arm A. Environmental enrichment is associated with differential mRNA expression. We tested the effect of environmental enrichment on mRNA expression of individual transcripts using negative binomial models, adjusting for total library size and global dispersion. Arm B. Environmental enrichment is associated with differential CpG and non CpG methylation. We tested the effects of environmental enrichment on CpG and non CpG methylation separately, using linear models (gaussian) and generalized linear models (binomial) respectively. Arm C. CpG and non CpG methylation are associated with differential mRNA expression. We tested the effects of CpG methylation and non CpG methylation on mRNA expression separately. Importantly, we only tested the effect of methylation on mRNA transcripts mapped to the same underlying gene as the methylation read. Additionally, we only tested the relationship between mRNA and methylation reads that were each differentially expressed, as observed in Arms A and B respectively. We used negative binomial models, adjusting for total library size and global dispersion. Environmental enrichment was included as a covariate. Mediation. mCG loci statistically mediate the relationships between environmental enrichment and mRNA expression. Taking only significantly correlated read pairs from Arm C analyses, we tested the mediating effect of mCG on EE/mRNA transcription relationships.

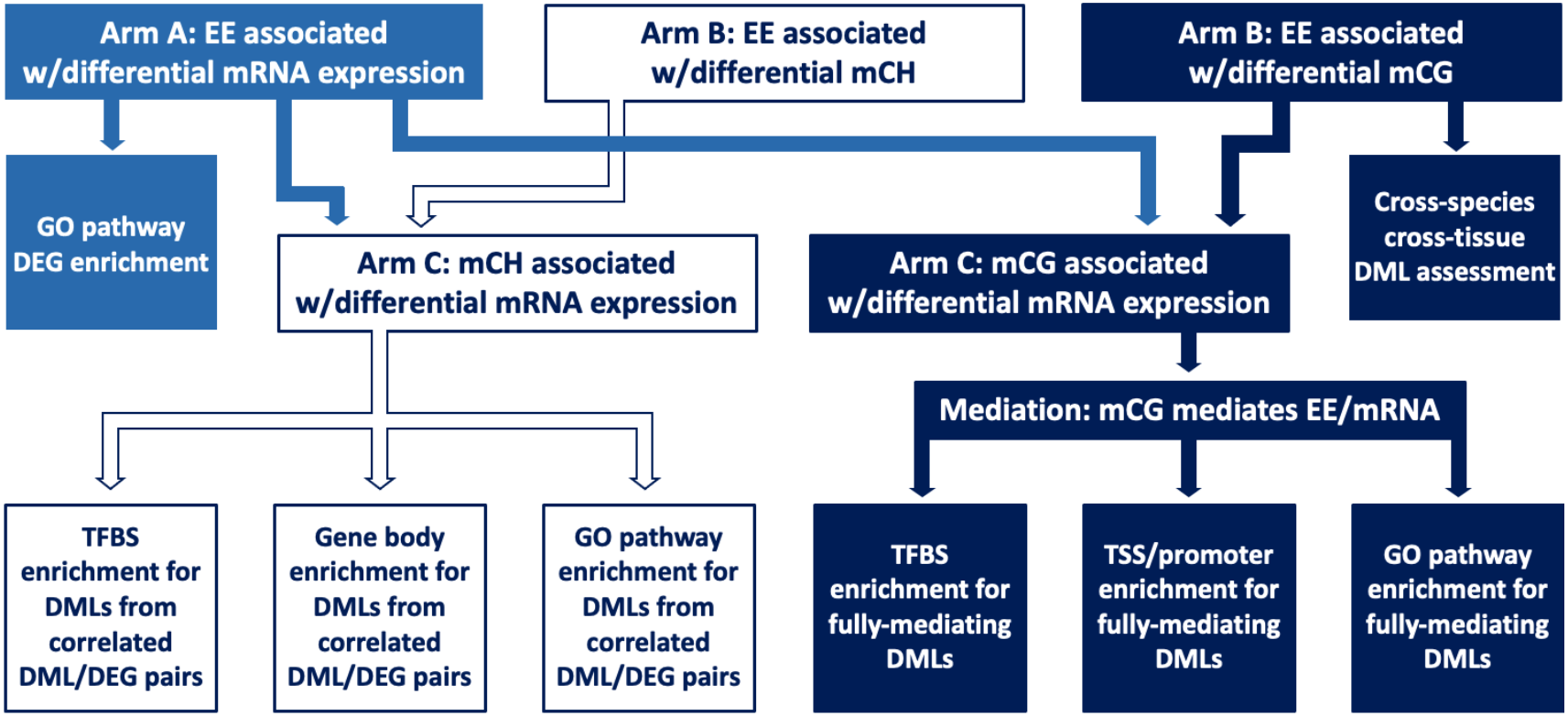


Figure 4.4 Next-generation sequencing mRNA, mCG, and mCH analytics pipeline. Light blue boxes denote mRNA-specific workflows, dark blue boxes denote mCG-specific workflows, and white boxes denote mCH-specific workflows.

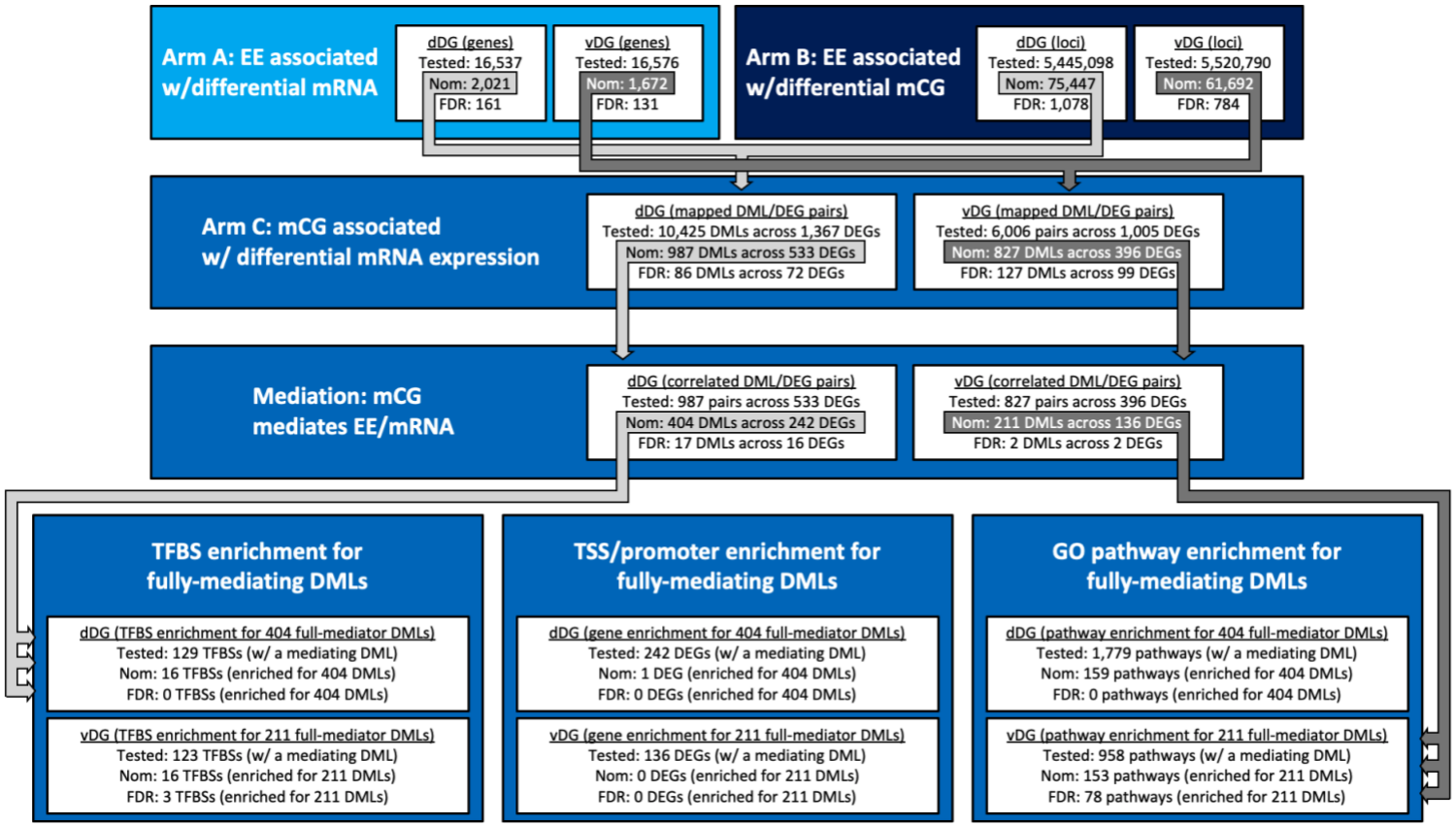


Figure 4.5 Next-generation sequencing mRNA and mCG mediation analytic pipeline with results. In Arm A analyses, we tested the effect of EE on ~16.5k mRNA transcripts in both dDG and vDG (separately). Transcripts with nominally significant ("nom") relationships to EE were advanced to Arm C. We carried out FDR and log2 FC thresholding to identify transcripts with the strongest relationships to EE. In Arm B analyses, we tested the effect of EE on ~5.5 million mCG loci in both dDG and vDG (separately). Again, mCG loci with nominally significant relationships to EE were advanced to study Arm C. We carried out FDR and percent difference (PD) thresholding to identify mCG loci with the strongest relationships to EE. Before carrying out Arm C analyses, we assessed the gene to which the nominally significant transcripts and mCG loci from Arms A and B were mapped. We only kept differentially expressed genes (DEGs) that had a differentially methylated loci (DML) mapped to it. We only kept DMLs that belonged to a DEG. This reduced the Arm C testing pool to ~10k mRNA/mCG read pairs in dDG, and ~6k mRNA/mCG read pairs in vDG. We then tested the individual effects of mCG loci on their mapped mRNA transcripts. mRNA/mCG read pairs exhibiting a nominally significant relationship were advanced to the mediation phase of testing. We carried out FDR thresholding to assess which pairs showed the strongest relationships to one another. Taking nominally significant mRNA/mCG read pairs from Arm C, we tested the mediating effect of mCG on EE/mRNA relationships. Nominally significant, full-mediators were used in downstream transcription factor binding site, transcription start site, and GO pathway enrichment analyses.

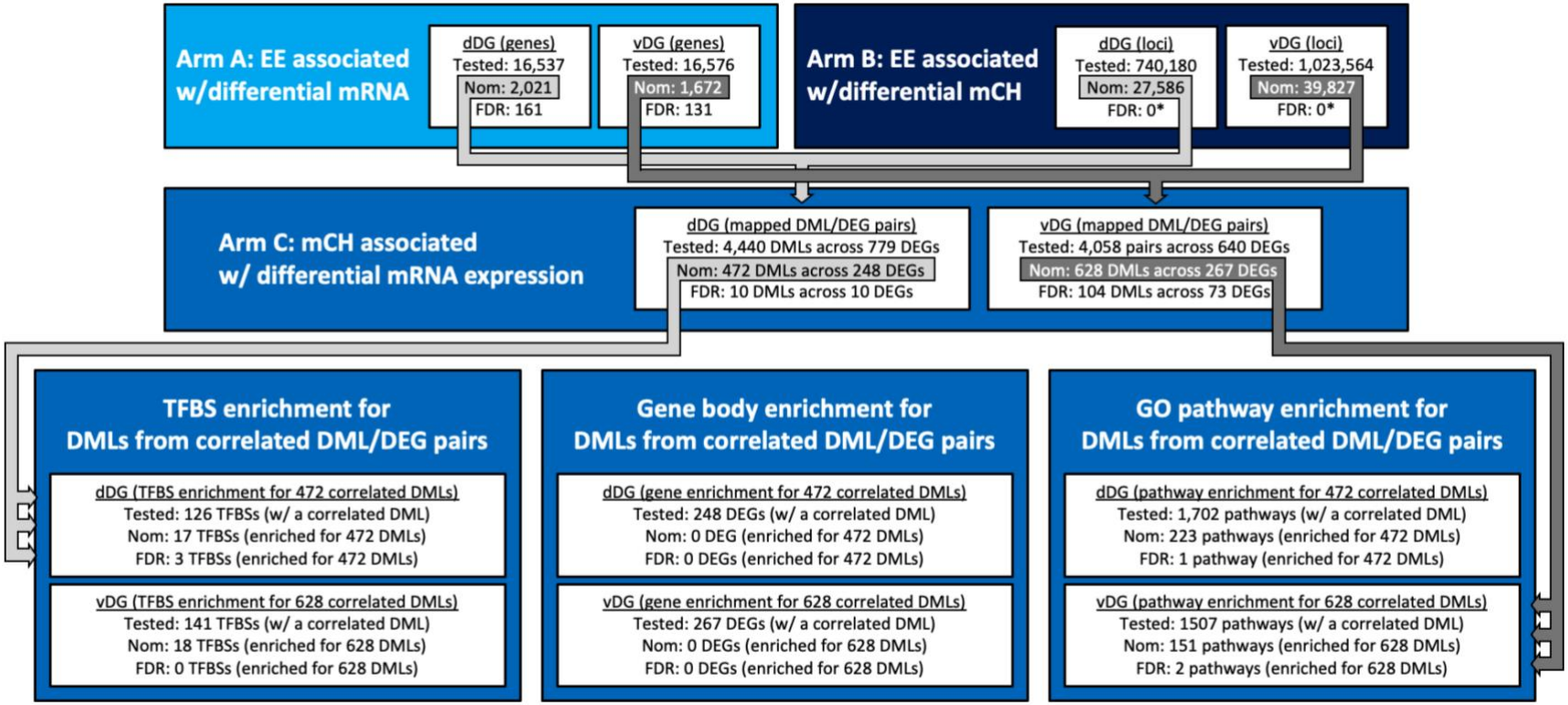


Figure 4.6 Next-generation sequencing mRNA and mCH mediation analytic pipeline with results. As explained in Figure 4.5, in Arm A analyses, we tested the effect of EE on ~16.5k mRNA transcripts in both dDG and vDG (separately). Transcripts with nominally significant ("nom") relationships to EE were advanced to Arm C. We carried out FDR and log2 FC thresholding to identify transcripts with the strongest relationships to EE. In Arm B analyses, we tested the effect of EE on ~740k and ~1 million mCH loci in both dDG and vDG (respectively). mCH loci with nominally significant relationships to EE were advanced to study Arm C. Before carrying out Arm C analyses, we assessed the gene to which the nominally significant transcripts and mCH loci from Arms A and B were mapped. We only kept differentially expressed genes (DEGs) that had a differentially methylated loci (DML) mapped to it. We only kept DMLs that belonged to a DEG. This reduced the Arm C testing pool to ~4.4k mRNA/mCH read pairs in dDG, and ~4.0k mRNA/mCH read pairs in vDG. We then tested the individual effects of mCH loci on their mapped mRNA transcripts. mRNA/mCH read pairs exhibiting a nominally significant relationship were advanced to downstream transcription factor binding site, transcription start site, and GO pathway enrichment analyses.

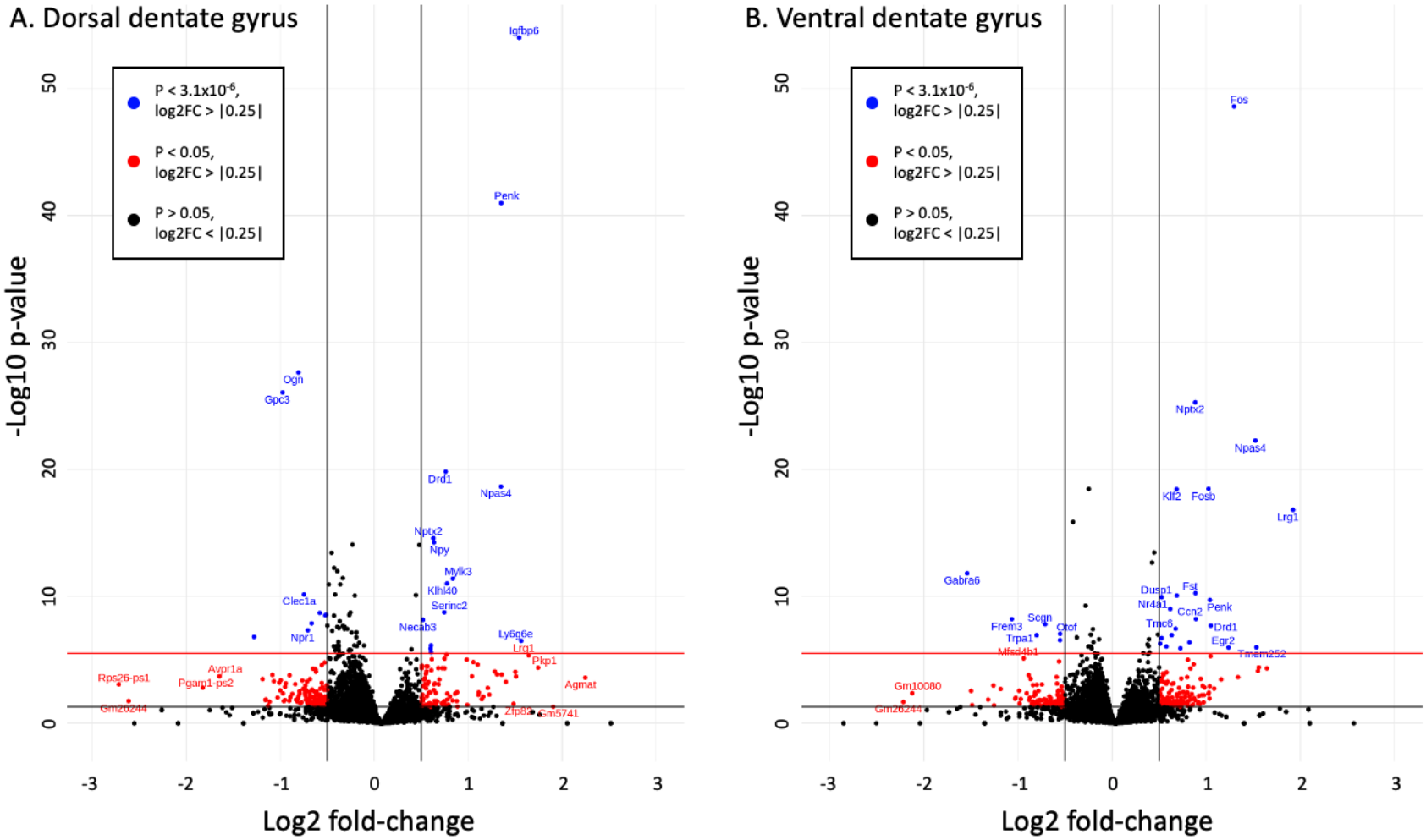


Figure 4.7 Volcano plots of dorsal and ventral dentate gyrus Arm A RNA-seq analysis in mouse. Blue points represent transcripts with log2 FC greater than |0.25| and $p < 3.1x10^{-6}$. Red points represent transcripts with log2 FC greater than |0.25| and $p > 3.1x10^{-6}$ but < 0.05. Black points represent transcripts with log2 FC < |0.25| or p > 0.05. A. Dorsal dentate gyrus RNA-seq log2 FC versus p-value. Twenty-three transcripts had $p < 3.1x10^{-6}$ and log2 FC greater than |0.25|. 228 transcripts had p < 0.05 and log2 FC greater than |0.25|. 2021 transcripts had p < 0.05. B. Ventral dentate gyrus RNA-seq log2 FC versus p-value. Twenty-eight transcripts had $p < 3.1x10^{-6}$ and log2 FC greater than |0.25|. 218 transcripts had p < 0.05 and log2 FC greater than |0.25|. 1669 transcripts had p < 0.05.

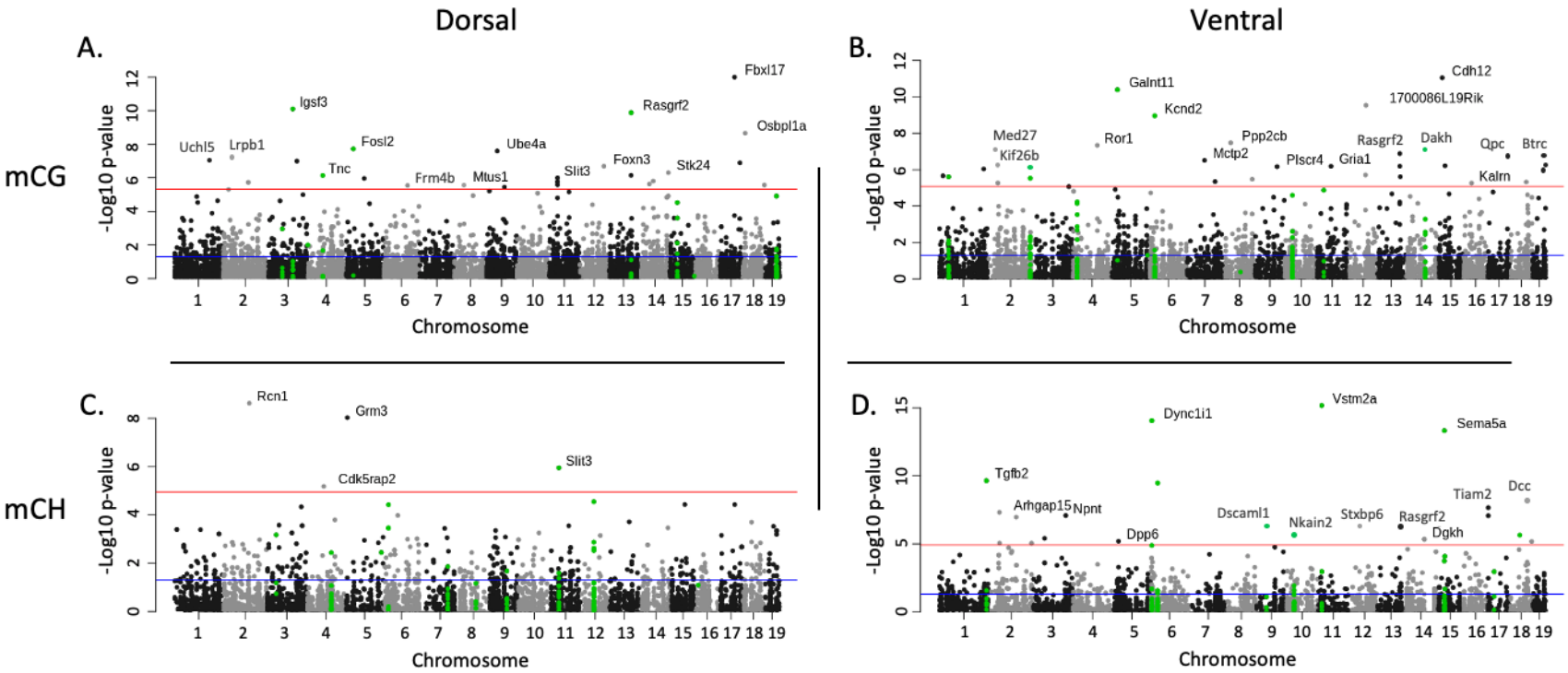


Figure 4.8 Manhattan plots of dorsal and ventral dentate gyrus Arm C analyses in mouse, across mCG and mCH. Figures contain p-values representative of Arm C analyses, where methylation in DMLs was tested for association with mRNA expression in mapped DEGs. The higher the plotted dot on the y-axis, the more predictive that locus' methylation was of mRNA expression. The blue horizontal lines represent the nominally significant p value cutoff < 0.05. The red horizontal lines represent the Bonferroni significant p value cutoff < 0.05/# of relationships tested in Arm C. Green dots belong to genes with the top ten level of differential mRNA expression. Visible gene labels are applied to loci with p < Bonferroni threshold within Arm C analyses. A. CpG methylation relationship with mRNA expression of underlying mapped transcripts in dorsal dentate gyrus, controlling for the effect of environmental enrichment. In dorsal dentate gyrus, CpG methylation was significantly associated (at Bonferroni threshold) with 85 DEGs, including *Rasgrf2*, *Fosl2*, *Tnc*, and *Igsf3*. B. CpG methylation relationship with mRNA expression of underlying mapped transcripts in ventral dentate gyrus, controlling for the effect of environmental enrichment. In ventral dentate gyrus, CpG methylation was significantly associated (at Bonferroni threshold) with 127 DEGs, including those of *Dakh*, *Galnt11*, and *Kcnd2*, among others. C. Non-CpG methylation relationship with mRNA expression of underlying mapped transcripts in dorsal dentate gyrus, controlling for the effect of environmental enrichment. In dorsal dentate gyrus, non CpG methylation was significantly associated (at Bonferroni threshold) with differential expression of only four genes. One of these, *Slit3*, was in the top ten of DEGs in this analysis pool. D. Non-CpG methylation relationship with mRNA expression of underlying mapped transcripts in ventral dentate gyrus, controlling for the effect of environmental enrichment. In ventral dentate gyrus, non CpG methylation was significantly associated (at Bonferroni threshold) with 104 DEGs. These included, among others, *Sema5a*, *Vstm2a*, and *Dync1i1*.

Table 4.1 Top 5 differentially expressed transcripts (by P-value) in response to environmental enrichment in dorsal and ventral mouse dentate gyrus.

Gene name	Tissue	Rank	EE log2 FC	EE P-value	EE FDR	Gene name description	Feature type	MGI Gene/Marker ID
Igfbp6	dDG	1	1.59	9.66E-55	1.58E-50	insulin-like growth factor binding protein 6	protein coding gene	MGI:96441
Penk	dDG	2	1.39	1.03E-41	8.40E-38	preproenkephalin	protein coding gene	MGI:104629
Ogn	dDG	3	-0.77	2.29E-28	1.25E-24	osteoglycin	protein coding gene	MGI:109278
Gpc3	dDG	4	-0.94	8.48E-27	3.47E-23	glypican 3	protein coding gene	MGI:104903
Drd1	dDG	5	0.80	1.49E-20	4.88E-17	dopamine receptor D1	protein coding gene	MGI:99578
Fos	vDG	1	1.21	2.53E-49	4.15E-45	FBJ osteosarcoma oncogene	protein coding gene	MGI:95574
Nptx2	vDG	2	0.81	5.10E-26	4.18E-22	neuronal pentraxin 2	protein coding gene	MGI:1858209
Npas4	vDG	3	1.44	5.17E-23	2.83E-19	neuronal PAS domain protein 4	protein coding gene	MGI:2664186
Fosb	vDG	4	0.94	3.35E-19	9.87E-16	FBJ osteosarcoma oncogene B	protein coding gene	MGI:95575
Igfbp5	vDG	5	-0.32	3.46E-19	9.87E-16	insulin-like growth factor binding protein 5	protein coding gene	MGI:96440

Legend: EE: environmental enrichment; FC: fold-change; FDR: false discovery rate, dDG: dorsal dentate gyrus; vDG: ventral dentate gyrus

Table 4.2 Top 5 differentially methylated CG loci (by P-value) in response to environmental enrichment in dorsal and ventral mouse dentate gyrus.

Gene name	Chr:position	Tissue	Rank	EE PD	EE P-value	EE FDR	Gene name description	Feature type	MGI Gene/Marker ID
Plxna4	Chr6:32340658	dDG	1	-14.0	6.06E-14	3.30E-07	plexin A4	protein coding gene	MGI:2179061
B3glct	Chr5:149753329	dDG	2	16.2	8.04E-14	4.38E-07	beta-3-glucosyltransferase	protein coding gene	MGI:2685903
Gstt3	Chr10:75776018	dDG	3	-11.3	2.88E-12	7.84E-06	glutathione S-transferase, theta 3	protein coding gene	MGI:2143526
Adarb2	Chr13:8569865	dDG	4	19.4	2.89E-12	7.88E-06	adenosine deaminase, RNA-specific, B2	protein coding gene	MGI:2151118
Lrp1b	Chr2:41419042	dDG	5	18.7	8.81E-12	1.56E-05	low density lipoprotein-related protein 1B	protein coding gene	MGI:2151136
Mettl15	Chr2:109243084	vDG	1	-10.2	1.58E-14	8.71E-08	methyltransferase like 15	protein coding gene	MGI:1924144
Nubp1	Chr16:10411612	vDG	2	-10.7	3.14E-12	1.73E-05	nucleotide binding protein 1	protein coding gene	MGI:1347073
Prss48	Chr3:86001744	vDG	3	16.5	3.35E-12	2.39E-05	protease, serine 48	protein coding gene	MGI:2685865
Nom1	Chr5:29434466	vDG	4	17.1	4.95E-12	2.39E-05	nucleolar protein with MIF4G domain 1	protein coding gene	MGI:1861749
Alx1	Chr10:103026849	vDG	5	-18.5	1.15E-11	5.32E-05	ALX homeobox 1	protein coding gene	MGI:104621

Legend: EE: environmental enrichment; PD: percent difference, FDR: false discovery rate, Chr: chromosome; dDG: dorsal dentate gyrus; vDG: ventral dentate gyrus

Table 4.3 Top 5 differentially expressed transcripts (by P-value) in association with mCG in dorsal and ventral mouse dentate gyrus.

Gene name	Chr:position	Tissue	Rank	mCG estimate	mCG P-value	mCG FDR	Gene name description	Feature type	MGI Gene/Marker ID
Fbxl17	Chr17:63340223	dDG	1	-0.089	9.98E-13	1.04E-08	F-box and leucine-rich repeat protein 17	protein coding gene	MGI:1354704
Igsf3	Chr3:101402606	dDG	2	-0.03	7.84E-11	4.09E-07	immunoglobulin superfamily, member 3	protein coding gene	MGI:1926158
Rasgrf2	Chr13:91972970	dDG	3	-0.043	1.32E-10	4.58E-07	RAS protein-specific guanine nucleotide-releasing factor 2	protein coding gene	MGI:109137
Osbpl1a	Chr18:12933637	dDG	4	0.026	2.19E-09	5.72E-06	oxysterol binding protein-like 1A	protein coding gene	MGI:1927551
Fosl2	Chr5:32157447	dDG	5	-0.035	1.89E-08	3.95E-05	fos-like antigen 2	protein coding gene	MGI:102858
Cdh12	Chr15:21290339	vDG	1	0.11	8.91E-12	5.35E-08	cadherin 12	protein coding gene	MGI:109503
Galnt11	Chr5:25257599	vDG	2	-0.027	4.03E-11	1.21E-07	polypeptide N-acetylgalactosaminyltransferase 11	protein coding gene	MGI:2444392
1700086L19Rik	Chr12:74285044	vDG	3	-0.023	2.89E-10	5.78E-07	RIKEN cDNA 1700086L19 gene	protein coding gene	MGI:1921534
Kcnd2	Chr6:21700384	vDG	4	0.022	1.10E-09	1.65E-06	potassium voltage-gated channel, Shal-related family, member 2	protein coding gene	MGI:102663
Ppp2cb	Chr8:33603280	vDG	5	-0.064	3.32E-08	3.98E-05	protein phosphatase 2 (formerly 2A), catalytic subunit, b isoform	protein coding gene	MGI:1321161

Legend: mCG: CpG methylation; Chr: chromosome; FDR: false discovery rate, dDG: dorsal dentate gyrus; vDG: ventral dentate gyrus

Table 4.4 Top 5 differentially expressed transcripts (by P-value) in association with mCH in dorsal and ventral mouse dentate gyrus.

Gene name	Chr:position	Tissue	Rank	mCH estimate	mCH P-value	mCG FDR	Gene name description	Feature type	MGI Gene/Marker ID
Rcn1	chr2:105390536	dDG	1	0.355	2.40E-09	1.07E-05	reticulocalbin 1	protein coding gene	MGI:104559
Grm3	chr5:9709659	dDG	2	-0.159	9.47E-09	2.10E-05	glutamate receptor, metabotropic 3	protein coding gene	MGI:1351340
Slit3	chr11:35277380	dDG	3	-0.186	1.13E-06	1.67E-03	slit guidance ligand 3	protein coding gene	MGI:1315202
Cdk5rap2	chr4:70361841	dDG	4	-0.231	6.64E-06	7.37E-03	CDK5 regulatory subunit associated protein 2	protein coding gene	MGI:2384875
Npas3	chr12:53363315	dDG	5	0.217	2.84E-05	2.11E-02	neuronal PAS domain protein 3	protein coding gene	MGI:1351610
Vstm2a	chr11:16285206	vDG	1	0.19	6.82E-16	2.77E-12	V-set and transmembrane domain containing 2A	protein coding gene	MGI:2384826
Dync1i1	chr6:5861525	vDG	2	0.194	8.92E-15	1.21E-11	dynein cytoplasmic 1 intermediate chain 1	protein coding gene	MGI:107743
Dync1i1	chr6:5961368	vDG	3	0.194	8.92E-15	1.21E-11	dynein cytoplasmic 1 intermediate chain 1	protein coding gene	MGI:107743
Sema5a	chr15:32643391	vDG	4	0.195	4.80E-14	4.87E-11	semaphorin 5A	protein coding gene	MGI:107556
Tgfb2	chr1:186640735	vDG	5	0.243	2.30E-10	1.86E-07	transforming growth factor, beta 2	protein coding gene	MGI:98726

Legend: mCH: CH methylation; Chr: chromosome; EE: environmental enrichment; dDG: dorsal dentate gyrus; vDG: ventral dentate gyrus; SE: standard error

Table 4.5 Top 5 mRNA/EE relationships mediated by mCG in dorsal and ventral mouse dentate gyrus.

Gene name	Chr:position	Tissue	IDE b	IDE FDR	DE b	DE P-value	TE b	TE P-value	Status	Gene name description	Feature type	MGI Gene/Marker ID
Pappa	Chr4:65276112	dDG	11.3	4.60E-02	5.0	2.45E-01	17.3	< 1E-16	Full mediator	pregnancy-associated plasma protein A	PC	MGI:97479
Myo5b	Chr18:74659297	dDG	-168.0	4.60E-02	-96.2	1.71E-01	-253.6	< 1E-16	Full mediator	myosin VB	PC	MGI:106598
Tnfrsf19	Chr14:60967710	dDG	-109.4	4.60E-02	-16.1	6.40E-01	-124.4	1.00E-03	Full mediator	TNF receptor superfamily, member 19	PC	MGI:1352474
Ercc4	Chr16:13127072	dDG	-28.7	< 1E-16	-3.6	6.98E-01	-32.1	1.60E-03	Full mediator	excision repair cross-complementing deficiency	PC	MGI:1354163
Smtnl2	Chr11:72403429	dDG	-14.1	< 1E-16	-1.2	8.10E-01	-15.3	4.00E-03	Full mediator	smoothelin-like 2	PC	MGI:2442764
Frzb	Chr2:80428126	vDG	-97.8	< 1E-16	57.9	1.21E-01	-54.1	2.88E-02	Full mediator	frizzled-related protein	PC	MGI:892032
Ror1	Chr4:100390763	vDG	-18.6	< 1E-16	3.2	7.06E-01	-16.5	1.92E-02	Full mediator	receptor tyrosine kinase-like orphan receptor 1	PC	MGI:1347520
Nkain3	Chr4:20553490	vDG	234.5	< 1E-16	-178.2	< 1E-16	-61.6	< 1E-16	Partial mediator	Na+/K+ transporting ATPase interacting 3	PC	MGI:2444830
Myo5b	Chr18:74556187	vDG	637.4	< 1E-16	-813.5	< 1E-16	-390.7	1.36E-02	Partial mediator	myosin VB	PC	MGI:106598
Myo5b	Chr18:74576505	vDG	553.0	< 1E-16	-767.7	< 1E-16	-389.3	1.52E-02	Partial mediator	myosin VB	PC	MGI:106598

Legend: EE: environmental enrichment; mCG: CpG methylation; Chr: chromosome; FDR: false discovery rate, IDE: indirect effect; DE: direct effect; TE: total effect; dDG: dorsal dentate gyrus; vDG: ventral dentate gyrus; PC: protein coding gene

Table 4.6 Top TFBS's enriched for DML's across dorsal and ventral mouse dentate gyrus and analysis pools.

TFBS	Tissue	Analysis pool	Locus type	Estimate	P-value	FDR	Rank	N Gene Set Genes	N Probes	N Genes	Represented genes
ARNT_02	dDG	Arm C	mCH	4.075	4.59E-05	1.02E-02	1	303	2	2	*Rfx4, Catsperd*
TST1_01	dDG	Arm C	mCH	3.632	2.82E-04	3.13E-02	2	247	11	8	*Gpm6A, Pax6, Plcg2, Ptprk, Zfhx4, Btbd3, Sema5B, Csmd3*
CREBP1_Q2	dDG	Arm C	mCH	3.443	5.75E-04	4.27E-02	3	332	2	1	*Grm3*
OCT_C	dDG	Mediation	mCG	3.209	1.33E-03	2.48E-01	1	261	10	10	*Mtus1, Gcnt2, C2Cd5, Ebf1, Lpl, Cadm1, Nfia, Rab26, Map2K6, Slc6A15*
TST1_01	dDG	Mediation	mCG	2.923	3.47E-03	2.48E-01	2	247	6	4	*Smarca2, Ptprk, Btbd3, Arl4A*
NKX25_02	dDG	Mediation	mCG	2.828	4.68E-03	2.48E-01	3	166	21	8	*Foxn3, Zdhhc21, Dab1, Rora, Ntrk2, Lrp1B, Tet1, Slc6A15*
FOXO3_01	vDG	Arm C	mCH	3.085	2.03E-03	1.57E-01	1	207	13	5	*Cacna2D1, Lhx9, Rbfox1, Stard13, Pdzrn4*
POU3F2_02	vDG	Arm C	mCH	3.026	2.48E-03	1.57E-01	2	189	29	13	*Bmp5, Lyst, Zpld1, Rbpj, Add3, Kalrn, Cadm2, Cdh13, Smarca2, Col25A1, Lmo4, Pdzrn4, Slc6A15*
LMO2COM_02	vDG	Arm C	mCH	2.829	4.67E-03	1.57E-01	3	209	14	5	*Rbfox1, Pkia, Stard13, Nr3C1, Vstm2A*
POU3F2_02	vDG	Mediation	mCG	4.238	2.25E-05	4.55E-03	1	189	11	5	*Add3, Kalrn, Cadm2, Cdh13, Col25A1*
STAT5A_02	vDG	Mediation	mCG	3.537	4.05E-04	3.74E-02	2	251	7	3	*Arhgap44, Kalrn, Rbfox1*
LHX3_01	vDG	Mediation	mCG	3.452	5.56E-04	3.74E-02	3	72	7	3	*Arhgap44, Kalrn, Rbfox1*

Legend: TFBS: transcription factor binding site; DML: differentially methylated loci; mCG: CpG methylation; mCH: CH methylation; dDG: dorsal dentate gyrus; vDG: ventral dentate gyrus

Table 4.7 Top gene ontology biological process and molecular function pathways enriched for mediating mCG loci.

Geneset ID	Tissue	Analysis pool	Locus type	Geneset type	Description	Estimate	P-value	FDR	N Geneset genes	N Geneset peak genes	Rank
GO:0001649	dDG	Arm C	mCH	BP	osteoblast differentiation	4.4	8.80E-06	2.29E-02	220	11	1
GO:0045667	dDG	Arm C	mCH	BP	regulation of osteoblast differentiation	3.9	1.03E-04	5.85E-02	134	9	2
GO:0001754	dDG	Arm C	mCH	BP	eye photoreceptor cell differentiation	3.9	1.12E-04	5.85E-02	53	8	3
GO:0003707	dDG	Arm C	mCH	MF	steroid hormone receptor activity	3.1	1.78E-03	1.76E-01	58	13	31
GO:0004879	dDG	Arm C	mCH	MF	RNA pol II transcription factor activity, ligand-activated sequence-specific DNA binding	3.1	1.78E-03	1.76E-01	46	13	32
GO:0098531	dDG	Arm C	mCH	MF	transcription factor activity, direct ligand regulated sequence-specific DNA binding	3.1	1.78E-03	1.76E-01	46	13	33
GO:0097529	dDG	Mediation	mCG	BP	myeloid leukocyte migration	2.9	3.30E-03	5.35E-01	164	3	1
GO:0072507	dDG	Mediation	mCG	BP	divalent inorganic cation homeostasis	2.9	3.67E-03	5.35E-01	454	6	2
GO:0002440	dDG	Mediation	mCG	BP	production of molecular mediator of immune response	2.9	4.33E-03	5.35E-01	173	3	3
GO:0009416	dDG	Mediation	mCG	MF	steroid binding	2.6	9.23E-03	3.46E-01	98	8	16
GO:0048565	dDG	Mediation	mCG	MF	sterol binding	2.6	9.23E-03	3.46E-01	41	8	17
GO:0055123	dDG	Mediation	mCG	MF	guanyl nucleotide binding	2.6	9.48E-03	3.46E-01	363	7	19
GO:0043266	vDG	Arm C	mCH	BP	regulation of potassium ion transport	3.9	9.10E-05	1.11E-01	93	14	3
GO:1901379	vDG	Arm C	mCH	BP	regulation of potassium ion transmembrane transport	3.9	9.10E-05	1.11E-01	66	14	4
GO:0034764	vDG	Arm C	mCH	BP	positive regulation of transmembrane transport	3.7	2.14E-04	1.11E-01	134	22	5
GO:0015459	vDG	Arm C	mCH	MF	potassium channel regulator activity	3.9	8.92E-05	2.04E-02	39	14	1
GO:0099106	vDG	Arm C	mCH	MF	ion channel regulator activity	3.9	8.92E-05	2.04E-02	78	14	2
GO:0016247	vDG	Arm C	mCH	MF	channel regulator activity	3.4	6.48E-04	5.85E-02	121	16	11
GO:0002791	vDG	Mediation	mCG	BP	regulation of peptide secretion	4.2	2.81E-05	2.63E-02	453	14	1
GO:0072659	vDG	Mediation	mCG	BP	protein localization to plasma membrane	3.8	1.26E-04	2.63E-02	224	7	2
GO:0022604	vDG	Mediation	mCG	BP	regulation of cell morphogenesis	3.7	2.36E-04	2.63E-02	451	14	3
GO:0004674	vDG	Mediation	mCG	MF	protein serine/threonine kinase activity	3.7	2.58E-04	7.55E-02	438	16	11
GO:0030676	vDG	Mediation	mCG	MF	Rac guanyl-nucleotide exchange factor activity	3.4	6.55E-04	9.59E-02	15	5	25
GO:0005516	vDG	Mediation	mCG	MF	calmodulin binding	2.4	1.52E-02	3.88E-01	182	12	95

Legend: mCG: CpG methylation; mCH: non-CpG methylation; BP: biological process; MF: molecular function; FDR: false discovery rate

Table 4.8 Cross-species mCG mapping between human hg38 and mouse mm10 genomes with differential mCG results (Arm B).

Human Illumina probe ID	Human module	Human HGCN symbol	Tissue	Mouse Chr:position	Mouse MGI symbol	EE estimate	EE P-value	DML rank	Gene body	TSS/promoter
cg10189441	Thistle2	*SERINC5*	dDG	chr13:92693210	*Serinc5*	-4.557	1.53E-02	79210	TRUE	FALSE
cg01289218	Thistle2	*RBM46*	dDG	chr3:82876701	*Rbm46*	-1.129	3.02E-02	124244	FALSE	TRUE
cg01827861	Thistle2	*RBM46*	dDG	chr3:82876701	*Rbm46*	-1.129	3.02E-02	124244	FALSE	TRUE
cg08282428	Thistle2	*RBM46*	dDG	chr3:82876701	*Rbm46*	-1.129	3.02E-02	124244	FALSE	TRUE
cg22103164	Thistle2	*RBM46*	dDG	chr3:82876701	*Rbm46*	-1.129	3.02E-02	124244	FALSE	TRUE
cg22496683	Thistle2	*RBM46*	dDG	chr3:82876701	*Rbm46*	-1.129	3.02E-02	124244	FALSE	TRUE
cg23483495	Thistle2	*RBM46*	dDG	chr3:82876701	*Rbm46*	-1.129	3.02E-02	124244	FALSE	TRUE
cg01583753	Thistle2	*CDKL4*	dDG	chr17:80577026	*Cdkl4*	4.909	6.76E-02	350123	FALSE	TRUE
cg09834706	Thistle2	*CDKL4*	dDG	chr17:80577026	*Cdkl4*	4.909	6.76E-02	350123	FALSE	TRUE
cg12302982	Thistle2	*CDKL4*	dDG	chr17:80577026	*Cdkl4*	4.909	6.76E-02	350123	FALSE	TRUE
cg00722631	Pink4	*SEMA5A*	vDG	chr15:32242559	*Sema5a*	3.185	7.41E-02	321375	FALSE	TRUE
cg00722631	Pink4	*SEMA5A*	vDG	chr15:32242493	*Sema5a*	1.145	8.02E-02	343989	FALSE	TRUE
cg01289218	Thistle2	*RBM46*	vDG	chr3:82876684	*Rbm46*	4.064	1.08E-01	441721	FALSE	TRUE
cg01827861	Thistle2	*RBM46*	vDG	chr3:82876684	*Rbm46*	4.064	1.08E-01	441721	FALSE	TRUE
cg08282428	Thistle2	*RBM46*	vDG	chr3:82876684	*Rbm46*	4.064	1.08E-01	441721	FALSE	TRUE
cg22103164	Thistle2	*RBM46*	vDG	chr3:82876684	*Rbm46*	4.064	1.08E-01	441721	FALSE	TRUE
cg22496683	Thistle2	*RBM46*	vDG	chr3:82876684	*Rbm46*	4.064	1.08E-01	441721	FALSE	TRUE
cg23483495	Thistle2	*RBM46*	vDG	chr3:82876684	*Rbm46*	4.064	1.08E-01	441721	FALSE	TRUE
cg15150970	Maroon	*DNMT3A*	vDG	chr12:3893206	*Dnmt3a*	-0.868	1.22E-01	494814	FALSE	TRUE
cg10189441	Thistle2	*SERINC5*	vDG	chr13:92693210	*Serinc5*	-2.237	1.27E-01	510307	TRUE	FALSE

Legend: mCG: CpG methylation; hg38: human genome build 38; mm10: mouse genome build 10; Chr: chromosome; DML: differentially methylated loci; TSS: transcription start site

CHAPTER 5: CONCLUSION

5.1 General Discussion

Overall, this body of work used novel genome-scale bioinformatics approaches to investigate underlying epigenetic profiles associated with, and/or *potentially* responsible for, the biological embedding of both adverse and enriching external environments. It provided novel insights to the fields of stress, trauma, and psychiatric illness by providing network-level and locus-specific peripheral biomarkers of adverse social exposures (ASEs) (such as family emotional health (FEH) and childhood trauma), and by identifying peripheral epigenetic correlates of neural endophenotypes of psychiatric illness. The current body of work also identified potentially mechanistic locus-specific epigenetic mediators of the biological embedding of environmental enrichment (EE) in the dorsal and ventral subregions of the mouse dentate gyrus (dDG; vDG).

One of the major overarching themes of this work is the potential for peripheral epigenetic measures to index ASE-associated variability in genes/loci associated with the hypothalamic-pituitary-adrenal (HPA)-axis, immune system, and central nervous system (CNS). While previous studies have focused on peripheral epigenetic measures in relation to adverse exposures or neural endophenotypes of psychiatric illnesses, few studies to date have incorporated these three disparate data types in pointed investigations into peripheral epigenetic signatures indexing the relationships between ASEs, epigenetic measures, and neural endophenotypes of psychiatric illness, or the *potential* epigenetic mechanisms explaining the biological embedding of ASEs.

To address these gaps, our first study applied genome-scale bioinformatics approaches to assess whether saliva-derived CpG 5mC (CpG methylation; mCG) clusters might statistically mediate the relationship between poor FEH, as experienced during childhood, and hemisphere-specific fronto-limbic grey matter volume (neural endophenotypes of psychiatric illness).

To this end, we found that exposure to poor FEH during childhood was associated with lower hippocampus and amygdala grey matter volume (in a hemisphere-specific manner), and that a subset of saliva-derived mCG clusters statistically mediated these relationships. Cluster-wide and probe-specific results implied that the association of poor childhood FEH with right hippocampal volume was indexed by peripheral epigenetic signatures related to mediation of the immune response, CNS cell migration, TGF-B signaling, and cytoskeletal organization. On the other hand, the association of poor childhood FEH with right amygdala volume was indexed and statistically mediated by signatures relevant to synapse development, neurotransmitter signal transduction, and cytoskeleton organization. Our methylome analysis highlighted the enrichment of stress-activated MAPK, beta-amyloid clearance, and cytokine receptor activity pathways in association with poor FEH. Overall, we posit that these FEH-associated epigenetic signatures could function as proxies of altered fronto-limbic grey matter volume associated with poor childhood FEH, and that peripheral epigenetic signatures indexing our relationships of interest may be explained by TGF-B- and/or MAPK-driven peripheral inflammation related to the development of psychiatric illness.

Our second study took a similar approach as in Chapter 2, while focusing on ASEs in a trauma-exposed, all female, majority African American (AA) cohort (Grady Trauma Project; GTP) with blood-derived mCG measurements. We focused specifically on the GTP because AA women are subject to elevated risk of PTSD-related psychiatric illness development and

persistence, but the mechanisms underlying that elevated risk are unclear. As in Chapter 2, we investigated whether clusters of peripheral mCG measurements *statistically* mediated the relation between ASEs and fronto-limbic grey matter volume. Besides cohort composition, a key difference between Chapters 2 and 3, however, was the inclusion of surface area (SA) and cortical thickness (CT) measures across discrete sub-sections of the frontal cortex.

We found that participants reporting low household income had lower amygdala grey matter volume than those reporting higher household incomes. Mirroring findings from cohorts enriched for stress, trauma, and psychiatric illness, we also found that childhood trauma burden (CTQ) was negatively associated with rostral medial frontal gyrus (RMFG) SA, and that this relationship was mediated and indexed by two clusters of mCG probes with gene functions related to immune signaling, cytoskeletal organization, methyltransferase activity, and endothelial growth factor pathways. Our methylome analysis highlighted the enrichment of phospholipase C (PLC) and inositol activation signaling pathways, which are major mechanisms of neuronal and hormonal signal transduction, in association with CTQ. Once again, we posit that the peripheral epigenetic signatures indexing our relationship of interest may be explained by peripheral inflammation, but that epigenetic regulation of the PLC and inositol signaling pathways may be dysregulated in association with CTQ.

Another major theme of the current work is the potential for mCG and non-CpG 5mC (mCH) to be mechanisms by which the external environment biologically embeds itself into the physiology of an organism, thereby leaving a temporally stable footprint on the genome through which downstream biological processes may be altered. In the case of Zhang et al 2018, researchers observed increased volume of both hippocampal poles, and increased newborn neuron count in the dorsal and ventral dentate gyrus (dDG; vDG) in response to EE[121].

Through next-generation sequencing *of a separate cohort of mice* exposed to the same EE paradigm, researchers also identified differentially methylation regions (DMRs) in these same hippocampal sub-regions, which correlated with mRNA expression across a limited number of genomic regions. This cohort of animals with molecular data formed the basis of the secondary data analysis performed in Chapter 4. Their research did not, however, clearly investigate the locus-specific effects of EE on mCG/mCH, or the potential mediating effects of mCG/mCH on the relationship between EE and differential mRNA transcription.

Taking an approach which posited that the effects of ASEs and EE operate on similar underlying neurobiological substrates, and that bi-directional alterations to underlying molecular regulatory networks may be at play in response to these divergent exposures, we sought to determine whether mCG or mCH causally mediated the relationship between EE and mRNA transcription. We hypothesized that the transcription start sites (mCG), promoter regions (mCG), gene bodies (mCH), and transcription factor binding sites (mCG/mCH) underlying differentially expressed genes (DEGs) would be enriched with these mediating loci. We posited that enrichment of these features *could be* related to the increased hippocampal volume and newborn neuron count observed in the separate cohort of animals, but were unable to assess this position.

We found 404 fully-mediating mCG loci in dDG, and 211 in vDG. Both the MAPK and TNF-B signaling pathways were implicated in the biological embedding of EE, albeit to differing degrees between dDG and vDG. Furthermore, it appeared that the MAPK and TGF-B signaling cascades could have been operating in response to EE in spatio-temporally specific patterns within the dentate gyrus, with epigenetic regulation of these pathways observed through divergent epigenetic mechanisms (mCG and mCH respectively). Numerous mediating loci were identified in genes with distinct roles in hippocampal development, psychiatric illness, and

neurodevelopmental disorders including *Dab1*, *Rora*, *Card14*, *Kalrn*, and *Cdh13*. Epigenetic regulation of gene expression through differential mCG was observed at a higher degree in dDG than in vDG, whereas the opposite was true of mCH in vDG versus dDG. In assessing the biological process and molecular function pathways enriched for mediating mCG loci and mRNA-correlated mCH loci in mouse dentate gyrus, we found signatures of epigenetic regulation that largely mirror results from our work in Chapters 2 with human ASEs. The findings of the current study inform divergent molecular mechanisms of the biologic embedding of EE on a locus-specific, epigenetic mechanism-specific, and hippocampal sub-region-specific basis. These findings could be used in the future to drive pointed locus-specific investigations of hippocampal development psychiatric illness etiology in humans.

5.2 Childhood versus Adult Exposures

The ASEs of interest investigated across Chapters 2 and 3 could be grossly classified into two groups, dependent upon the developmental period in which the ASEs were experienced. In this manner, FEH and CTQ are distinct from the other exposure of interest, in that they capture experiences throughout earlier stages of development, as opposed to adulthood (as in the case of low income, and adult trauma burden). Given the sensitivity of CNS development to stress and trauma during early developmental stages and the temporally stable nature of mCG, it follows that the relationships between FEH/CTQ and neural endophenotypes of psychiatric illness were statistically mediated by mCG clusters. In other words, adverse environments or trauma, as experienced during childhood, interact with neural substrates that are already primed towards plasticity, partially through epigenetic mechanisms. To this end, pre-frontal cortex vulnerability is specifically heightened due to its protracted maturation into early adulthood[309]. On the other

hand, the hippocampus exhibits developmental malleability in part due to the generation of adult-born neurons, a process which is influenced by experiences throughout the organisms life[310]. The malleable neural substrates and their accompanying epigenetic regulators are therefore highly susceptible to external inputs, which may leave a lasting impression.

On the other hand, we did observe relationships between adult exposures and CNS morphometry measures, such as the relationship between low income and low amygdala grey matter volume. We posit that the lack of mediating mCG clusters for these types of relationships could be potentially attributed to indirect pathways of effect between the environmental exposures and outcomes of interest. More specifically, it is well-known that socioeconomic disadvantage experienced during childhood portends a range of negative outcomes in adulthood relating to educational achievement and economic circumstances[311]. It is feasible, therefore, that latent childhood ASEs could represent the ultimate source of the association between low household income in adulthood and low amygdala grey matter volume, and that our testing framework failed to facilitate the identification of mCG clusters relating to this. On the other side of the putative causal pathway, it is equally feasible that low household income in adulthood facilitates exposure to other potentially detrimental factors or a lack of exposure to supporting factors. These factors could more directly mediate the relationship between low income and amygdala grey matter volume, and could potentially be biological embedded at a greater magnitude that low income itself. Such factors could include lack of access to quality health care or mental health care, lack of access to other avenues of institutional support, or lack of social capital, among others.

Understanding the degree to which a study participants' broader environment and household provide adversity versus support could be instrumental in defining the dynamics of

the biological embedding of such environments. In addition, collecting longitudinal ASE-related and molecular data from childhood through adulthood could help to address these issues, although the financial and logistical burdens associated with such as study are readily apparent.

5.3 Cytokine Signaling and the Hippocampus

Of additional interest are the similarities in the epigenetic signatures observed when comparing results from Chapters 2 (FEH) and 4 (EE), specifically in the context of the hippocampus. The FEH/right hippocampal volume relationship was indexed by peripheral epigenetic signatures related to CNS cell migration, TGF-B signaling, and cytoskeletal organization, while FEH-associated enriched methylomic pathways included terms relating to cytokine receptor activity and regulation of osteoblast differentiation.

We posit that our observation of enriched TGF-B signaling and cytokine receptor activity pathways, in the context of Chapter 2, could be representative of macrophage or microglia-mediated release of TGF-B and other anti-inflammatory cytokines. Recently, it has become known that the blood brain barrier (BBB) and other areas of PNS versus CNS interaction can actively transport immune cells, immune molecules (such as pro-inflammatory or anti-inflammatory cytokines), or immune signals from the periphery to the CNS[312]. These newly discovered mechanisms may help to explain why we are able to observe these signatures in peripheral tissue. Alternatively, observations of these signatures in peripheral tissue may be due in part to microglia activation throughout the peripheral nervous system. We further posit that these processes could potentially be indicative of a temporally stable compensatory neuroprotective process associated with poor childhood FEH induced neuroinflammation, or a

dysregulation of the epigenetic regulatory mechanisms of the microglia-mediated pro-inflammatory or anti-inflammatory processes.

In the context of Chapter 4, we posit that our observation of osteoblast differentiation is potentially representative of signaling through the TGF-B pathway *towards* hippocampal neuron maturation. We raise that certain members of the TGF-B signaling ligand family, such as bone morphogenetic proteins, play marked roles in osteoblast differentiation[301] and specialized roles in hippocampal neurons critical to dendrite formation[302], axon guidance[303], and synaptic stability[304]; these processes require dynamic remodeling of both actin and microtubule assemblies[305]. Our observation that pathways in the DG were enriched for migration/differentiation of myeloid precursor cell activity is another indicator of the potential involvement of microglia, as microglia are derived from the myeloid lineage[313].

It appears then, that FEH and EE, although diametrically opposed in regard to their physiological and cognitive effects, could potentially operate on similar underlying neurobiological substrates, through similar cellular and molecular mechanisms, regulated in part through similar epigenetic mediators. However, the diverging point in the above-outlined hypothetical model at which FEH and EE begin to exert their opposing downstream physiological and cognitive effects is unclear.

5.4 Moving Forward

The current research presents novel findings regarding peripheral epigenetic correlates of psychiatric illness neural endophenotypes, and *potentially* mechanistic locus-specific epigenetic mediators of the biological embedding of EE in the mouse dentate gyrus, while highlighting common underlying molecular pathways through which ASEs and EE *potentially* bi-

directionally exert their effects. However, additional steps should be taken in the future that could help to inform psychiatric illness risk/resilience factors, and mechanisms of psychiatric illness onset, prognosis, and therapy response.

In the past, the field of oncology served as a "proving ground" for epigenetic research[314]. Eventually, the measures and methodologies used to study cancer were transferred to the study of psychiatric illness, stress, and trauma as researchers focused on how it was possible for our psychosocial environment to "get under the skin". Currently within the field of oncology, researchers in both the academic-sphere and in the private sector are pushing towards the inclusion of multi-omic and multi-modal data sets to further drive our understanding of cancer and the identification of novel therapeutic agents[315]. Moving forward, researchers investigating epigenetic contributions to psychiatric illness should once again look to the field of oncology and follow suit in incorporating both multi-omic and multi-modal arrays of data into their analyses. The eventual incorporation of genomic, epigenomic, and transcriptomic molecular data in tandem with structural and functional neuroimaging measures could eventually push research towards a more multi-system-level representation of psychiatric illness. These models could serve as a basis to investigate interactions between the different biological domains affected by ASEs or EE.

Super-charging the potential of these multi-modal datasets, would be the longitudinal collection of such data. Depending on the specific questions at hand, researchers could utilize a number of longitudinal approaches. One such approach could be the identification and subsequent enrollment of "high-risk" individuals through prior familial diagnoses. Other methods could include the enrollment of participants with newly onset psychiatric illnesses, or

participants with a newly started therapy regimen. These strategies would help to inform mechanisms of disease onset, prognosis, and therapy response.

An important question worth posing, though, is whether researchers should continue to group study participants based on diagnoses alone. In recent years, considerable effort has been put forth towards the continued development of the Diagnostic and Statistical Manual of Mental Disorders, Version Five (DSM 5)[29]. This manual, commonly used by mental health care practitioners to inform their assessment, diagnosis, and treatment of psychiatric illnesses, categorizes these illnesses based largely on the magnitude and range of symptom constellations exhibited by a patient. To this end, many psychiatric illnesses share symptoms and are commonly comorbid with one another[22, 316]. These relationships strongly hint towards common etiologies between illnesses. On this notion, identifying discrete symptom clusters of psychiatric illness for use as new study endpoints, as opposed to categorical diagnoses, could facilitate additional insights into the biological networks underlying such diseases.

In addition, the continued development of endophenotypes, or intermediate phenotypes that fill the gap between underlying genomic sequence and psychiatric illness diagnoses (or ideally symptom clusters), is paramount to identifying the etiology of such illnesses. These could take the form of structural or functional neuroimaging measures, collected through methods such as functional MRI (fMRI), positron emission tomography (PET), or diffusion tensor imaging (DTI). They could also include measures taken from other physiological systems, such as the HPA-axis or immune system (cortisol[49], C-reactive protein[317], blood leukocyte composition[318], cytokines[62], etc).

The aforementioned, and idealized, paths forward are largely the focus of the National Institute of Mental Health's (NIMH) Research Domain Criteria (RDoC) initiative[319, 320]. In

brief, RDoC is a novel research framework meant to facilitate the investigation of the biological and psychosocial bases of psychiatric illness. It implores the eventual dismissal of categorical psychiatric illness diagnoses as the use of study endpoints, in favor of using deconstructed elements or core features of such categories. One example of this could include focusing a study on identifying physiological features associated with heightened startle reactivity in participants across a range of psychiatric diagnoses.

Through the application of the RDoC framework to multi-modal/omic, longitudinal, and symptom-oriented studies that focus inquisition on the biological and psychosocial bases of psychiatric illness, future researchers can begin to gather insights on causal mechanisms of illness onset versus resultant physiological or psychological effects of such illnesses. This path could also prove fruitful regarding the development of efficacious and replicable predictive biomarkers of disease onset, prognosis, or therapy response, hopefully cutting across stratifying factors such as genomic ancestry and race, or biological sex and gender. Such as path would hopefully inform further education and legislation regarding psychiatric illness prevention and intervention, while reducing the immense disease burden of psychiatric illness across the United States, and the world as a whole.

REFERENCES

1. Gilbert R, Widom CS, Browne K, Fergusson D, Webb E, Janson S. Burden and consequences of child maltreatment in high-income countries. The Lancet. 2009;373:68–81.

2. Kessler RC, Sonnega A, Bromet E, Hughes M, Nelson CB. Posttraumatic stress disorder in the National Comorbidity Survey. Arch Gen Psychiatry. 1995;52:1048–60.

3. Breslau N. The epidemiology of trauma, PTSD, and other posttrauma disorders. Trauma Violence Abuse. 2009;10:198–210.

4. Roberts AL, Gilman SE, Breslau J, Breslau N, Koenen KC. Race/ethnic differences in exposure to traumatic events, development of post-traumatic stress disorder, and treatment-seeking for post-traumatic stress disorder in the United States. Psychol Med. 2011;41:71–83.

5. Kilpatrick DG, Resnick HS, Milanak ME, Miller MW, Keyes KM, Friedman MJ. National Estimates of Exposure to Traumatic Events and PTSD Prevalence Using DSM-IV and DSM-5 Criteria. J Trauma Stress. 2013;26:537–47.

6. Semega J, Kollar M, Creamer J, Mohanty A. Income and poverty in the United States: 2018 (Current Population Reports No. P60-266). Retrieved from US Census Bureau website: https://www. census. gov/library/publications/2019/demo/p60-266. html. 2019.

7. Tyler JH, Lofstrom M. Finishing High School: Alternative Pathways and Dropout Recovery. Future Child. 2009;19:77–103.

8. Stambaugh LF, Forman-Hoffman V, Williams J, Pemberton MR, Ringeisen H, Hedden SL, et al. Prevalence of serious mental illness among parents in the United States: results from the National Survey of Drug Use and Health, 2008-2014. Ann Epidemiol. 2017;27:222–4.

9. Mueller SC, Maheu FS, Dozier M, Peloso E, Mandell D, Leibenluft E, et al. Early-life stress is associated with impairment in cognitive control in adolescence: an fMRI study. Neuropsychologia. 2010;48:3037–44.

10. Navalta CP, Polcari A, Webster DM, Boghossian A, Teicher MH. Effects of childhood sexual abuse on neuropsychological and cognitive function in college women. J Neuropsychiatry Clin Neurosci. 2006;18:45–53.

11. Mezzacappa E. Alerting, orienting, and executive attention: developmental properties and sociodemographic correlates in an epidemiological sample of young, urban children. Child Dev. 2004;75:1373–86.

12. Turrell G, Lynch JW, Kaplan GA, Everson SA, Helkala E-L, Kauhanen J, et al. Socioeconomic position across the lifecourse and cognitive function in late middle age. J Gerontol B Psychol Sci Soc Sci. 2002;57:S43-51.

13. Wadsworth ME, Raviv T, Santiago CD, Etter EM. Testing the Adaptation to Poverty-Related Stress Model: Predicting Psychopathology Symptoms in Families Facing Economic Hardship. J Clin Child Adolesc Psychol. 2011;40:646–57.

14. Connell AM, Goodman SH. The association between psychopathology in fathers versus mothers and children's internalizing and externalizing behavior problems: a meta-analysis. Psychol Bull. 2002;128:746–73.

15. Suveg C, Shaffer A, Morelen D, Thomassin K. Links Between Maternal and Child Psychopathology Symptoms: Mediation Through Child Emotion Regulation and Moderation Through Maternal Behavior. Child Psychiatry Hum Dev. 2011;42:507.

16. Kendler KS, Karkowski LM, Prescott CA. Stressful Life Events and Major Depression: Risk Period, Long-Term Contextual Threat, and Diagnostic Specificity. J Nerv Ment Dis. 1998;186:661–9.

17. Everson SA, Maty SC, Lynch JW, Kaplan GA. Epidemiologic evidence for the relation between socioeconomic status and depression, obesity, and diabetes. J Psychosom Res. 2002;53:891–5.

18. Mistry RS, Benner AD, Tan CS, Kim SY. Family Economic Stress and Academic Well-Being Among Chinese-American Youth: The Influence of Adolescents' Perceptions of Economic Strain. J Fam Psychol JFP J Div Fam Psychol Am Psychol Assoc Div 43. 2009;23:279–90.

19. Kendler KS, Myers J, Halberstadt LJ. SHOULD THE DIAGNOSIS OF MAJOR DEPRESSION BE MADE INDEPENDENT OF OR DEPENDENT UPON THE PSYCHOSOCIAL CONTEXT? Psychol Med. 2010;40:771–80.

20. Eaton W, Harrison G. Life Chances, Life Planning, and Schizophrenia: A Review and Interpretation of Research on Social Deprivation. Int J Ment Health. 2001;30:58–81.

21. Lewis G, Bebbington P, Brugha T, Farrell M, Gill B, Jenkins R, et al. Socioeconomic status, standard of living, and neurotic disorder. The Lancet. 1998;352:605–9.

22. Breslau N, Davis GC, Peterson EL, Schultz LR. A second look at comorbidity in victims of trauma: the posttraumatic stress disorder–major depression connection. Biol Psychiatry. 2000;48:902–9.

23. Breslau N, Peterson EL, Schultz LR. A second look at prior trauma and the posttraumatic stress disorder effects of subsequent trauma: a prospective epidemiological study. Arch Gen Psychiatry. 2008;65:431–7.

24. Humphreys KL, LeMoult J, Wear JG, Piersiak HA, Lee A, Gotlib IH. Child maltreatment and depression: A meta-analysis of studies using the Childhood Trauma Questionnaire. Child Abuse Negl. 2020;102:104361.

25. Li M, D'Arcy C, Meng X. Maltreatment in childhood substantially increases the risk of adult depression and anxiety in prospective cohort studies: systematic review, meta-analysis, and proportional attributable fractions. Psychol Med. 2016;46:717–30.

26. Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, Ezzati M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet Lond Engl. 2012;380:2163–96.

27. Begg S, Tomijima N, Majmudar M, Bulzacchelli MT, Ebrahim S, Gaber Kreiser J, et al. The Burden of Disease and Injury in the United States 1996. Popul Health Metr. 2006.

28. Kessler RC. Posttraumatic stress disorder: the burden to the individual and to society. J Clin Psychiatry. 2000;61:4–14.

29. Association AP. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). American Psychiatric Pub; 2013.

30. Lupien SJ, King S, Meaney MJ, McEwen BS. Child's stress hormone levels correlate with mother's socioeconomic status and depressive state. Biol Psychiatry. 2000;48:976–80.

31. Klengel T, Mehta D, Anacker C, Rex-Haffner M, Pruessner JC, Pariante CM, et al. Allele-specific FKBP5 DNA demethylation mediates gene–childhood trauma interactions. Nat Neurosci. 2013;16:33–41.

32. Cohen S, Tyrrell DA, Smith AP. Negative life events, perceived stress, negative affect, and susceptibility to the common cold. J Pers Soc Psychol. 1993;64:131–40.

33. Segerstrom SC, Miller GE. Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry. Psychol Bull. 2004;130:601–30.

34. Salleh MohdR. Life Event, Stress and Illness. Malays J Med Sci MJMS. 2008;15:9–18.

35. Pagliaccio D, Luby JL, Bogdan R, Agrawal A, Gaffrey MS, Belden AC, et al. Stress-System Genes and Life Stress Predict Cortisol Levels and Amygdala and Hippocampal Volumes in Children. Neuropsychopharmacology. 2014;39:1245–53.

36. Hanson JL, Nacewicz BM, Sutterer MJ, Cayo AA, Schaefer SM, Rudolph KD, et al. Behavior Problems After Early Life Stress: Contributions of the Hippocampus and Amygdala. Biol Psychiatry. 2015;77:314–23.

37. Swartz JR, Knodt AR, Radtke SR, Hariri AR. A Neural Biomarker of Psychological Vulnerability to Future Life Stress. Neuron. 2015;85:505–11.

38. Teicher MH, Samson JA, Anderson CM, Ohashi K. The effects of childhood maltreatment on brain structure, function and connectivity. Nat Rev Neurosci. 2016;17:652–66.

39. Cassiers LLM, Sabbe BGC, Schmaal L, Veltman DJ, Penninx BWJH, Van Den Eede F. Structural and Functional Brain Abnormalities Associated With Exposure to Different Childhood

Trauma Subtypes: A Systematic Review of Neuroimaging Findings. Front Psychiatry. 2018;9. doi:10.3389/fpsyt.2018.00329.

40. Luby J, Belden A, Botteron K, Marrus N, Harms MP, Babb C, et al. The effects of poverty on childhood brain development: the mediating effect of caregiving and stressful life events. JAMA Pediatr. 2013;167:1135–42.

41. Kim P, Evans GW, Angstadt M, Ho SS, Sripada CS, Swain JE, et al. Effects of childhood poverty and chronic stress on emotion regulatory brain function in adulthood. Proc Natl Acad Sci U S A. 2013;110:18442–7.

42. Swartz JR, Hariri AR, Williamson DE. An epigenetic mechanism links socioeconomic status to changes in depression-related brain function in high-risk adolescents. Mol Psychiatry. 2017;22:209–14.

43. Pagliaccio D, Luby JL, Bogdan R, Agrawal A, Gaffrey MS, Belden AC, et al. Amygdala functional connectivity, HPA axis genetic variation, and life stress in children and relations to anxiety and emotion regulation. J Abnorm Psychol. 2015;124:817–33.

44. Park AT, Leonard JA, Saxler PK, Cyr AB, Gabrieli JDE, Mackey AP. Amygdala–medial prefrontal cortex connectivity relates to stress and mental health in early childhood. Soc Cogn Affect Neurosci. 2018;13:430–9.

45. Luby JL, Barch DM, Belden A, Gaffrey MS, Tillman R, Babb C, et al. Maternal support in early childhood predicts larger hippocampal volumes at school age. Proc Natl Acad Sci. 2012;109:2854–9.

46. Papmeyer M, Giles S, Sussmann JE, Kielty S, Stewart T, Lawrie SM, et al. Cortical Thickness in Individuals at High Familial Risk of Mood Disorders as They Develop Major Depressive Disorder. Biol Psychiatry. 2015;78:58–66.

47. Butterworth P, Cherbuin N, Sachdev P, Anstey KJ. The association between financial hardship and amygdala and hippocampal volumes: results from the PATH through life project. Soc Cogn Affect Neurosci. 2012;7:548–56.

48. Tomoda A, Suzuki H, Rabi K, Sheu Y-S, Polcari A, Teicher MH. Reduced Prefrontal Cortical Gray Matter Volume in Young Adults Exposed to Harsh Corporal Punishment. NeuroImage. 2009;47 Suppl 2:T66–71.

49. Carrion VG, Weems CF, Richert K, Hoffman BC, Reiss AL. Decreased Prefrontal Cortical Volume Associated With Increased Bedtime Cortisol in Traumatized Youth. Biol Psychiatry. 2010;68:491–3.

50. Corbo V, Salat DH, Amick MM, Leritz EC, Milberg WP, McGlinchey RE. Reduced cortical thickness in veterans exposed to early life trauma. Psychiatry Res. 2014;223:53–60.

51. Eckart C, Stoppel C, Kaufmann J, Tempelmann C, Hinrichs H, Elbert T, et al. Structural alterations in lateral prefrontal, parietal and posterior midline regions of men with chronic posttraumatic stress disorder. J Psychiatry Neurosci JPN. 2011;36:176–86.

52. Legge RM, Sendi S, Cole JH, Cohen-Woods S, Costafreda SG, Simmons A, et al. Modulatory effects of brain-derived neurotrophic factor Val66Met polymorphism on prefrontal regions in major depressive disorder. Br J Psychiatry. 2015;206:379–84.

53. Kovner R, Oler JA, Kalin NH. Cortico-Limbic Interactions Mediate Adaptive and Maladaptive Responses Relevant to Psychopathology. Am J Psychiatry. 2019;176:987–99.

54. Bellavance M-A, Rivest S. The HPA – Immune Axis and the Immunomodulatory Actions of Glucocorticoids in the Brain. Front Immunol. 2014;5. doi:10.3389/fimmu.2014.00136.

55. Haddad JJ, Saadé NE, Safieh-Garabedian B. Cytokines and neuro-immune-endocrine interactions: a role for the hypothalamic-pituitary-adrenal revolving axis. J Neuroimmunol. 2002;133:1–19.

56. Sternberg EM. Neural regulation of innate immunity: a coordinated nonspecific host response to pathogens. Nat Rev Immunol. 2006;6:318–28.

57. Jensen SKG, Berens AE, Nelson CA. Effects of poverty on interacting biological systems underlying child development. Lancet Child Adolesc Health. 2017;1:225–39.

58. Nusslock R, Miller GE. Early-Life Adversity and Physical and Emotional Health Across the Lifespan: A Neuroimmune Network Hypothesis. Biol Psychiatry. 2016;80:23–32.

59. Dunn EC, Soare TW, Zhu Y, Simpkin AJ, Suderman MJ, Klengel T, et al. Sensitive Periods for the Effect of Childhood Adversity on DNA Methylation: Results From a Prospective, Longitudinal Study. Biol Psychiatry. 2019;85:838–49.

60. Anisman H. Cascading effects of stressors and inflammatory immune system activation: implications for major depressive disorder. J Psychiatry Neurosci JPN. 2009;34:4–20.

61. Walter KH, Palmieri PA, Gunstad J. More than symptom reduction: Changes in executive function over the course of PTSD treatment. J Trauma Stress. 2010;23:292–5.

62. Misiak B, Beszłej JA, Kotowicz K, Szewczuk-Bogusławska M, Samochowiec J, Kucharska-Mazur J, et al. Cytokine alterations and cognitive impairment in major depressive disorder: From putative mechanisms to novel treatment targets. Prog Neuropsychopharmacol Biol Psychiatry. 2018;80:177–88.

63. Marganska A, Gallagher M, Miranda R. Adult attachment, emotion dysregulation, and symptoms of depression and generalized anxiety disorder. Am J Orthopsychiatry. 2013;83:131–41.

64. Green MF. Cognitive impairment and functional outcome in schizophrenia and bipolar disorder. J Clin Psychiatry. 2006;67:e12.

65. Hasler G, Northoff G. Discovering imaging endophenotypes for major depression. Mol Psychiatry. 2011;16:604–19.

66. Manoach DS, Agam Y. Neural markers of errors as endophenotypes in neuropsychiatric disorders. Front Hum Neurosci. 2013;7. doi:10.3389/fnhum.2013.00350.

67. Matsubara T, Matsuo K, Harada K, Nakano M, Nakashima M, Watanuki T, et al. Distinct and Shared Endophenotypes of Neural Substrates in Bipolar and Major Depressive Disorders. PLOS ONE. 2016;11:e0168493.

68. Joels M, Karst H, Alfarez D, Heine VM, Qin Y, Riel E van. Effects of Chronic Stress on Structure and Cell Function in Rat Hippocampus and Hypothalamus: Stress: Vol 7, No 4. https://www.tandfonline.com/doi/abs/10.1080/10253890500070005. Accessed 20 Mar 2021.

69. Sapolsky RM. A mechanism for glucocorticoid toxicity in the hippocampus: increased neuronal vulnerability to metabolic insults. J Neurosci. 1985;5:1228–32.

70. Vyas A, Mitra R, Shankaranarayana Rao BS, Chattarji S. Chronic stress induces contrasting patterns of dendritic remodeling in hippocampal and amygdaloid neurons. J Neurosci Off J Soc Neurosci. 2002;22:6810–8.

71. Nishi M, Horii-Hayashi N, Sasagawa T, Matsunaga W. Effects of early life stress on brain activity: Implications from maternal separation model in rodents. Gen Comp Endocrinol. 2013;181:306–9.

72. Chiba S, Numakawa T, Ninomiya M, Richards MC, Wakabayashi C, Kunugi H. Chronic restraint stress causes anxiety- and depression-like behaviors, downregulates glucocorticoid receptor expression, and attenuates glutamate release induced by brain-derived neurotrophic factor in the prefrontal cortex. Prog Neuropsychopharmacol Biol Psychiatry. 2012;39:112–9.

73. O'Mahony SM, Marchesi JR, Scully P, Codling C, Ceolho A-M, Quigley EMM, et al. Early Life Stress Alters Behavior, Immunity, and Microbiota in Rats: Implications for Irritable Bowel Syndrome and Psychiatric Illnesses. Biol Psychiatry. 2009;65:263–7.

74. Kubera M, Basta-Kaim A, Holan V, Simbirtsev A, Roman A, Pigareva N, et al. Effect of mild chronic stress, as a model of depression, on the immunoreactivity of C57BL \ 6 mice. Int J Immunopharmacol. 1998;20:781–9.

75. Luine V. Sex differences in chronic stress effects on memory in rats. Stress Amst Neth. 2002;5:205–16.

76. Toth LA, Kregel K, Leon L, Musch TI. Environmental enrichment of laboratory rodents: the answer depends on the question. Comp Med. 2011;61:314–21.

77. Francis DD, Diorio J, Plotsky PM, Meaney MJ. Environmental enrichment reverses the effects of maternal separation on stress reactivity. J Neurosci Off J Soc Neurosci. 2002;22:7840–3.

78. Artola A, von Frijtag JC, Fermont PCJ, Gispen WH, Schrama LH, Kamal A, et al. Long-lasting modulation of the induction of LTD and LTP in rat hippocampal CA1 by behavioural stress and environmental enrichment. Eur J Neurosci. 2006;23:261–72.

79. Morley-Fletcher S, Rea M, Maccari S, Laviola G. Environmental enrichment during adolescence reverses the effects of prenatal stress on play behaviour and HPA axis reactivity in rats. Eur J Neurosci. 2003;18:3367–74.

80. Schloesser RJ, Lehmann M, Martinowich K, Manji HK, Herkenham M. Environmental enrichment requires adult neurogenesis to facilitate the recovery from psychosocial stress. Mol Psychiatry. 2010;15:1152–63.

81. Chourbaji S, Hörtnagl H, Molteni R, Riva MA, Gass P, Hellweg R. The impact of environmental enrichment on sex-specific neurochemical circuitries - effects on brain-derived neurotrophic factor and the serotonergic system. Neuroscience. 2012;220:267–76.

82. Greenough WT, Madden TC, Fleischmann TB. Effects of isolation, daily handling, and enriched rearing on maze learning. Psychon Sci. 1972;27:279–80.

83. Leggio MG, Mandolesi L, Federico F, Spirito F, Ricci B, Gelfo F, et al. Environmental enrichment promotes improved spatial abilities and enhanced dendritic growth in the rat. Behav Brain Res. 2005;163:78–90.

84. Foti F, Laricchiuta D, Cutuli D, De Bartolo P, Gelfo F, Angelucci F, et al. Exposure to an Enriched Environment Accelerates Recovery from Cerebellar Lesion. Cerebellum Lond Engl. 2011;10:104–19.

85. Hannan AJ. Environmental enrichment and brain repair: harnessing the therapeutic effects of cognitive stimulation and physical activity to enhance experience-dependent plasticity. Neuropathol Appl Neurobiol. 2014;40:13–25.

86. Mandolesi L, Gelfo F, Serra L, Montuori S, Polverino A, Curcio G, et al. Environmental Factors Promoting Neural Plasticity: Insights from Animal and Human Studies. Neural Plast. 2017;2017:e7219461.

87. Kempermann G, Kuhn HG, Gage FH. Experience-Induced Neurogenesis in the Senescent Dentate Gyrus. J Neurosci. 1998;18:3206–12.

88. Bergami M. Experience-dependent plasticity of adult-born neuron connectivity. Commun Integr Biol. 2015;8:e1038444.

89. Sakalem ME, Seidenbecher T, Zhang M, Saffari R, Kravchenko M, Wördemann S, et al. Environmental enrichment and physical exercise revert behavioral and electrophysiological impairments caused by reduced adult neurogenesis. Hippocampus. 2017;27:36–51.

90. Chakrabarti L, Scafidi J, Gallo V, Haydar TF. Environmental Enrichment Rescues Postnatal Neurogenesis Defect in the Male and Female Ts65Dn Mouse Model of Down Syndrome. Dev Neurosci. 2011;33:428–41.

91. Gelfo F, Florenzano F, Foti F, Burello L, Petrosini L, De Bartolo P. Lesion-induced and activity-dependent structural plasticity of Purkinje cell dendritic spines in cerebellar vermis and hemisphere. Brain Struct Funct. 2016;221:3405–26.

92. Green EJ, Greenough WT. Altered synaptic transmission in dentate gyrus of rats reared in complex environments: evidence from hippocampal slices maintained in vitro. J Neurophysiol. 1986;55:739–50.

93. Duffy SN, Craddock KJ, Abel T, Nguyen PV. Environmental Enrichment Modifies the PKA-Dependence of Hippocampal LTP and Improves Hippocampus-Dependent Memory. Learn Mem. 2001;8:26–34.

94. Scholz J, Allemang-Grand R, Dazai J, Lerch JP. Environmental enrichment is associated with rapid volumetric brain changes in adult mice. NeuroImage. 2015;109:190–8.

95. Rizzi S, Bianchi P, Guidi S, Ciani E, Bartesaghi R. Impact of environmental enrichment on neurogenesis in the dentate gyrus during the early postnatal period. Brain Res. 2011;1415:23–33.

96. Rule L, Yang J, Watkin H, Hall J, Brydges NM. Environmental enrichment rescues survival and function of adult-born neurons following early life stress. Mol Psychiatry. 2020;:1–11.

97. King LS, Humphreys KL, Gotlib IH. The neglect–enrichment continuum: Characterizing variation in early caregiving environments. Dev Rev. 2019;51:109–22.

98. Rossetti MF, Varayoud J, Moreno-Piovano GS, Luque EH, Ramos JG. Environmental enrichment attenuates the age-related decline in the mRNA expression of steroidogenic enzymes and reduces the methylation state of the steroid 5α-reductase type 1 gene in the rat hippocampus. Mol Cell Endocrinol. 2015;412:330–8.

99. Griñan-Ferré C, Puigoriol-Illamola D, Palomera-Ávalos V, Pérez-Cáceres D, Companys-Alemany J, Camins A, et al. Environmental Enrichment Modified Epigenetic Mechanisms in SAMP8 Mouse Hippocampus by Reducing Oxidative Stress and Inflammaging and Achieving Neuroprotection. Front Aging Neurosci. 2016;8. doi:10.3389/fnagi.2016.00241.

100. Griñán-Ferré C, Izquierdo V, Otero E, Puigoriol-Illamola D, Corpas R, Sanfeliu C, et al. Environmental Enrichment Improves Cognitive Deficits, AD Hallmarks and Epigenetic Alterations Presented in 5xFAD Mouse Model. Front Cell Neurosci. 2018;12. doi:10.3389/fncel.2018.00224.

101. Guo JU, Su Y, Shin JH, Shin J, Li H, Xie B, et al. Distribution, recognition and regulation of non-CpG methylation in the adult mammalian brain. Nat Neurosci. 2014;17:215–22.

102. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet. 2003;33 Suppl:245–54.

103. Zhang T-Y, Meaney MJ. Epigenetics and the Environmental Regulation of the Genome and Its Function. Annu Rev Psychol. 2010;61:439–66.

104. Talens RP, Boomsma DI, Tobi EW, Kremer D, Jukema JW, Willemsen G, et al. Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. FASEB J Off Publ Fed Am Soc Exp Biol. 2010;24:3135–44.

105. Perroud N, Paoloni-Giacobino A, Prada P, Olié E, Salzmann A, Nicastro R, et al. Increased methylation of glucocorticoid receptor gene (NR3C1) in adults with a history of childhood maltreatment: a link with the severity and type of trauma. Transl Psychiatry. 2011;1:e59.

106. Tyrka AR, Price LH, Marsit C, Walters OC, Carpenter LL. Childhood adversity and epigenetic modulation of the leukocyte glucocorticoid receptor: preliminary findings in healthy adults. PloS One. 2012;7:e30148.

107. Uddin M, Aiello AE, Wildman DE, Koenen KC, Pawelec G, de los Santos R, et al. Epigenetic and immune function profiles associated with posttraumatic stress disorder. Proc Natl Acad Sci U S A. 2010;107:9470–5.

108. Smith AK, Conneely KN, Kilaru V, Mercer KB, Weiss TE, Bradley B, et al. Differential immune system DNA methylation and cytokine regulation in post-traumatic stress disorder. Am J Med Genet Part B Neuropsychiatr Genet Off Publ Int Soc Psychiatr Genet. 2011;156B:700–8.

109. Davies MN, Volta M, Pidsley R, Lunnon K, Dixit A, Lovestone S, et al. Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. Genome Biol. 2012;13:R43.

110. Hannon E, Lunnon K, Schalkwyk L, Mill J. Interindividual methylomic variation across blood, cortex, and cerebellum: implications for epigenetic studies of neurological and neuropsychiatric phenotypes. Epigenetics. 2015;10:1024–32.

111. Yamamoto T, Toki S, Siegle GJ, Takamura M, Takaishi Y, Yoshimura S, et al. Increased amygdala reactivity following early life stress: a potential resilience enhancer role. BMC Psychiatry. 2017;17. doi:10.1186/s12888-017-1201-x.

112. Illingworth RS, Gruenewald-Schneider U, De Sousa D, Webb S, Merusi C, Kerr ARW, et al. Inter-individual variability contrasts with regional homogeneity in the human brain DNA methylome. Nucleic Acids Res. 2015;43:732–44.

113. Booij L, Szyf M, Carballedo A, Frey E-M, Morris D, Dymov S, et al. DNA methylation of the serotonin transporter gene in peripheral cells and stress-related changes in hippocampal volume: a study in depressed patients and healthy controls. PloS One. 2015;10:e0119061.

114. Frodl T, Szyf M, Carballedo A, Ly V, Dymov S, Vaisheva F, et al. DNA methylation of the serotonin transporter gene (SLC6A4) is associated with brain function involved in processing emotional stimuli. J Psychiatry Neurosci JPN. 2015;40:296–305.

115. Ismaylova E, Lévesque ML, Pomares FB, Szyf M, Nemoda Z, Fahim C, et al. Serotonin transporter promoter methylation in peripheral cells and neural responses to negative stimuli: A study of adolescent monozygotic twins. Transl Psychiatry. 2018;8:1–9.

116. Vukojevic V, Kolassa I-T, Fastenrath M, Gschwind L, Spalek K, Milnik A, et al. Epigenetic modification of the glucocorticoid receptor gene is linked to traumatic memory and post-traumatic stress disorder risk in genocide survivors. J Neurosci Off J Soc Neurosci. 2014;34:10274–84.

117. Schechter DS, Moser DA, Paoloni-Giacobino A, Stenz L, Gex-Fabry M, Aue T, et al. Methylation of NR3C1 is related to maternal PTSD, parenting stress and maternal medial prefrontal cortical activity in response to child separation among mothers with histories of violence exposure. Front Psychol. 2015;6:690.

118. Tozzi L, Farrell C, Booij L, Doolin K, Nemoda Z, Szyf M, et al. Epigenetic Changes of FKBP5 as a Link Connecting Genetic and Environmental Risk Factors with Structural and Functional Brain Changes in Major Depression. Neuropsychopharmacol Off Publ Am Coll Neuropsychopharmacol. 2018;43:1138–45.

119. Sadeh N, Wolf EJ, Logue MW, Hayes JP, Stone A, Griffin LM, et al. EPIGENETIC VARIATION AT SKA2 PREDICTS SUICIDE PHENOTYPES AND INTERNALIZING PSYCHOPATHOLOGY. Depress Anxiety. 2016;33:308–15.

120. Sadeh N, Spielberg JM, Logue MW, Wolf EJ, Smith AK, Lusk J, et al. SKA2 methylation is associated with decreased prefrontal cortical thickness and greater PTSD severity among trauma-exposed veterans. Mol Psychiatry. 2016;21:357–63.

121. Zhang T-Y, Keown CL, Wen X, Li J, Vousden DA, Anacker C, et al. Environmental enrichment increases transcriptional and epigenetic differentiation between mouse dorsal and ventral dentate gyrus. Nat Commun. 2018;9:298.

122. National Research Council (US) and Institute of Medicine (US) Committee on Depression PP, England MJ, Sim LJ. Associations Between Depression in Parents and Parenting, Child Health, and Child Psychological Functioning. National Academies Press (US); 2009. https://www.ncbi.nlm.nih.gov/books/NBK215128/. Accessed 13 May 2020.

123. Dean K, Stevens H, Mortensen PB, Murray RM, Walsh E, Pedersen CB. Full Spectrum of Psychiatric Outcomes Among Offspring With Parental History of Mental Disorder. Arch Gen Psychiatry. 2010;67:822–9.

124. Toyokawa S, Uddin M, Koenen KC, Galea S. How does the social environment 'get into the mind'? Epigenetics at the intersection of social and psychiatric epidemiology. Soc Sci Med 1982. 2012;74:67–74.

125. Weaver ICG. Integrating Early Life Experience, Gene Expression, Brain Development, and Emergent Phenotypes. In: Advances in Genetics. Elsevier; 2014. p. 277–307. doi:10.1016/B978-0-12-800222-3.00011-5.

126. Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. Nat Genet. 2007;39:61–9.

127. Michaud EJ, van Vugt MJ, Bultman SJ, Sweet HO, Davisson MT, Woychik RP. Differential expression of a new dominant agouti allele (Aiapy) is correlated with methylation state and is influenced by parental lineage. Genes Dev. 1994;8:1463–72.

128. Smith AK, Kilaru V, Klengel T, Mercer KB, Bradley B, Conneely KN, et al. DNA extracted from saliva for methylation studies of psychiatric traits: evidence tissue specificity and relatedness to brain. Am J Med Genet Part B Neuropsychiatr Genet Off Publ Int Soc Psychiatr Genet. 2015;168B:36–44.

129. Schaafsma S m, Riedstra B j, Pfannkuche K a, Bouma A, Groothuis T g. g. Epigenesis of behavioural lateralization in humans and other animals. Philos Trans R Soc B Biol Sci. 2009;364:915–27.

130. Gotts SJ, Jo HJ, Wallace GL, Saad ZS, Cox RW, Martin A. Two distinct forms of functional lateralization in the human brain. Proc Natl Acad Sci. 2013;110:E3435–44.

131. Francks C. Exploring human brain lateralization with molecular genetics and genomics. Ann N Y Acad Sci. 2015;1359:1–13.

132. Kim E, Garrett A, Boucher S, Park M-H, Howe M, Sanders E, et al. Inhibited Temperament and Hippocampal Volume in Offspring of Parents with Bipolar Disorder. J Child Adolesc Psychopharmacol. 2016;27:258–65.

133. Haren NEM van, Setiaman N, Koevoets MGJC, Baalbergen H, Kahn RS, Hillegers MHJ. Brain structure, IQ, and psychopathology in young offspring of patients with schizophrenia or bipolar disorder. Eur Psychiatry. 2020;63. doi:10.1192/j.eurpsy.2019.19.

134. Klengel T, Pape J, Binder EB, Mehta D. The role of DNA methylation in stress-related psychiatric disorders. Neuropharmacology. 2014;80:115–32.

135. Nikolova YS, Bogdan R, Brigidi BD, Hariri AR. Ventral Striatum Reactivity to Reward and Recent Life Stress Interact to Predict Positive Affect. Biol Psychiatry. 2012;72:157–63.

136. Kim MJ, Avinun R, Knodt AR, Radtke SR, Hariri AR. Neurogenetic plasticity and sex influence the link between corticolimbic structural connectivity and trait anxiety. Sci Rep. 2017;7:10959.

137. Mareckova K, Hawco C, Dos Santos FC, Bakht A, Calarco N, Miles AE, et al. Novel polygenic risk score as a translational tool linking depression-related changes in the corticolimbic transcriptome with neural face processing and anhedonic symptoms. Transl Psychiatry. 2020;10. doi:10.1038/s41398-020-01093-w.

138. Kim MJ, Farber MJ, Knodt AR, Hariri AR. Corticolimbic circuit structure moderates an association between early life stress and later trait anxiety. NeuroImage Clin. 2019;24:102050.

139. Lecrubier Y, Sheehan D, Weiller E, Amorim P, Bonora I, Sheehan KH, et al. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. Eur Psychiatry. 1997;12:224–31.

140. First MB, Spitzer RL, Gibbon M, Williams JB. Structured Clinical Interview for DSM-IV Axis I Disorders, Research Version, Nonpatient Edition. N Y State Psychiatr Inst Biom Res Dep. 1996.

141. Weissman MM, Wickramaratne P, Adams P, Wolk S, Verdeli H, Olfson M. Brief screening for family psychiatric history: the family history screen. Arch Gen Psychiatry. 2000;57:675–82.

142. Milne BJ, Caspi A, Crump R, Poulton R, Rutter M, Sears MR, et al. The Validity of the Family History Screen for Assessing Family History of Mental Disorders. Am J Med Genet Part B Neuropsychiatr Genet Off Publ Int Soc Psychiatr Genet. 2009;0:41–9.

143. Devi G, Marder K, Schofield PW, Tang MX, Stern Y, Mayeux R. Validity of family history for the diagnosis of dementia among siblings of patients with late-onset Alzheimer's disease. Genet Epidemiol. 1998;15:215–23.

144. Bidaut-Russell M, Reich W, Cottler LB, Robins LN, Compton WM, Mattison RE. The Diagnostic Interview Schedule for Children (PC-DISC v.3.0): parents and adolescents suggest reasons for expecting discrepant answers. J Abnorm Child Psychol. 1995;23:641–59.

145. Williams RR, Hunt SC, Barlow GK, Chamberlain RM, Weinberg AD, Cooper HP, et al. Health family trees: a tool for finding and helping young family members of coronary and cancer prone pedigrees in Texas and Utah. Am J Public Health. 1988;78:1283–6.

146. Skinner HA. The drug abuse screening test. Addict Behav. 1982;7:363–71.

147. Selzer ML, Vinokur A, van Rooijen L. A self-administered Short Michigan Alcoholism Screening Test (SMAST). J Stud Alcohol. 1975;36:117–26.

148. Avinun R, Nevo A, Knodt AR, Elliott ML, Radtke SR, Brigidi BD, et al. Reward-Related Ventral Striatum Activity Buffers against the Experience of Depressive Symptoms Associated with Sleep Disturbances. J Neurosci Off J Soc Neurosci. 2017;37:9724–9.

149. Clements K, Turpin G. The Life Events Scale for Students: Validation for use with British samples. Personal Individ Differ. 1996;20:747–51.

150. Pedraza O, Bowers D, Gilmore R. Asymmetry of the hippocampus and amygdala in MRI volumetric measurements of normal adults. J Int Neuropsychol Soc. 2004;10:664–78.

151. Lindauer RJL, Vlieger E-J, Jalink M, Olff M, Carlier IVE, Majoie CBLM, et al. Smaller hippocampal volume in Dutch police officers with posttraumatic stress disorder. Biol Psychiatry. 2004;56:356–63.

152. Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, Chiang M-C, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. NeuroImage. 2009;46:786–802.

153. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. NeuroImage. 1999;9:179–94.

154. Fischl B, Sereno MI, Dale AM. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. NeuroImage. 1999;9:195–207.

155. Klein A, Tourville J. 101 Labeled Brain Images and a Consistent Human Cortical Labeling Protocol. Front Neurosci. 2012;6. doi:10.3389/fnins.2012.00171.

156. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron. 2002;33:341–55.

157. Ratanatharathorn A, Boks MP, Maihofer AX, Aiello AE, Amstadter AB, Ashley-Koch AE, et al. Epigenome-wide association of PTSD from heterogeneous cohorts with a common multi-site analysis pipeline. Am J Med Genet Part B Neuropsychiatr Genet Off Publ Int Soc Psychiatr Genet. 2017;174:619–30.

158. R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

159. Du P, Zhang X, Huang C-C, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics. 2010;11:587.

160. Braun PR, Han S, Hing B, Nagahama Y, Gaul LN, Heinzman JT, et al. Genome-wide DNA methylation comparison between live human brain and peripheral tissues within individuals. Transl Psychiatry. 2019;9:1–10.

161. Houseman EA, Molitor J, Marsit CJ. Reference-free cell mixture adjustments in analysis of DNA methylation data. Bioinforma Oxf Engl. 2014;30:1431–9.

162. Houseman EA, Kile ML, Christiani DC, Ince TA, Kelsey KT, Marsit CJ. Reference-free deconvolution of DNA methylation data and mediation by cell composition effects. BMC Bioinformatics. 2016;17:259.

163. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–75.

164. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.

165. Lin X, Barton S, Holbrook JD. How to make DNA methylome wide association studies more powerful. Epigenomics. 2016;8:1117–29.

166. Unternaehrer E, Luers P, Mill J, Dempster E, Meyer AH, Staehli S, et al. Dynamic changes in DNA methylation of stress-associated genes (OXTR, BDNF) after acute psychosocial stress. Transl Psychiatry. 2012;2:e150.

167. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B Methodol. 1995;57:289–300.

168. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. mediation: R Package for Causal Mediation Analysis. J Stat Softw. 2014;59. doi:10.18637/jss.v059.i05.

169. Ren X, Kuan PF. methylGSA: a Bioconductor package and Shiny app for DNA methylation data length bias adjustment in gene set testing. Bioinforma Oxf Engl. 2019;35:1958–9.

170. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. PloS One. 2011;6:e21800.

171. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res. 2002;12:996–1006.

172. Rosenbloom KR, Sloan CA, Malladi VS, Dreszer TR, Learned K, Kirkup VM, et al. ENCODE Data in the UCSC Genome Browser: year 5 update. Nucleic Acids Res. 2013;41:D56–63.

173. Tian G, Thomas S, Cowan NJ. Effect of TBCD and its regulatory interactor Arl2 on tubulin and microtubule integrity. Cytoskelet Hoboken NJ. 2010;67:706–14.

174. Flex E, Niceta M, Cecchetti S, Thiffault I, Au MG, Capuano A, et al. Biallelic Mutations in TBCD, Encoding the Tubulin Folding Cofactor D, Perturb Microtubule Dynamics and Cause Early-Onset Encephalopathy. Am J Hum Genet. 2016;99:962–73.

175. Miyake N, Fukai R, Ohba C, Chihara T, Miura M, Shimizu H, et al. Biallelic TBCD Mutations Cause Early-Onset Neurodegenerative Encephalopathy. Am J Hum Genet. 2016;99:950–61.

176. Guo S, Zhu Q, Jiang T, Wang R, Shen Y, Zhu X, et al. Genome-wide DNA methylation patterns in CD4+ T cells from Chinese Han patients with rheumatoid arthritis. Mod Rheumatol. 2017;27:441–7.

177. Song H, Fang F, Tomasson G, Arnberg FK, Mataix-Cols D, Cruz LF de la, et al. Association of Stress-Related Disorders With Subsequent Autoimmune Disease. JAMA. 2018;319:2388–400.

178. Duclot F, Kabbaj M. The Role of Early Growth Response 1 (EGR1) in Brain Plasticity and Neuropsychiatric Disorders. Front Behav Neurosci. 2017;11. doi:10.3389/fnbeh.2017.00035.

179. Zhang W, Liu HT. MAPK signal pathways in the regulation of cell proliferation in mammalian cells. Cell Res. 2002;12:9–18.

180. Wefers B, Hitz C, Hölter SM, Trümbach D, Hansen J, Weber P, et al. MAPK Signaling Determines Anxiety in the Juvenile Mouse Brain but Depression-Like Behavior in Adults. PLOS ONE. 2012;7:e35035.

181. Bruchas MR, Schindler AG, Shankar H, Messinger DI, Miyatake M, Land BB, et al. Selective p38α MAPK Deletion in Serotonergic Neurons Produces Stress Resilience in Models of Depression and Addiction. Neuron. 2011;71:498–511.

182. Zhu C-B, Lindler KM, Owens AW, Daws LC, Blakely RD, Hewlett WA. Interleukin-1 Receptor Activation by Systemic Lipopolysaccharide Induces Behavioral Despair Linked to MAPK Regulation of CNS Serotonin Transporters. Neuropsychopharmacology. 2010;35:2510–20.

183. Firestein GS. Immunologic mechanisms in the pathogenesis of rheumatoid arthritis. J Clin Rheumatol Pract Rep Rheum Musculoskelet Dis. 2005;11 3 Suppl:S39-44.

184. Fifield J, Tennen H, Reisine S, McQuillan J. Depression and the long-term risk of pain, fatigue, and disability in patients with rheumatoid arthritis. Arthritis Rheum. 1998;41:1851–7.

185. Malemud CJ, Miller AH. Pro-inflammatory cytokine-induced SAPK/MAPK and JAK/STAT in rheumatoid arthritis and the new anti-depression drugs. Expert Opin Ther Targets. 2008;12:171–83.

186. Chan KL, Cathomas F, Russo SJ. Central and Peripheral Inflammation Link Metabolic Syndrome and Major Depressive Disorder. Physiology. 2019;34:123–33.

187. Martin A, Markhvida M, Hallegatte S, Walsh B. Socio-Economic Impacts of COVID-19 on Household Consumption and Poverty. Econ Disasters Clim Change. 2020;4:453–79.

188. World Health Organization. Cross-national comparisons of the prevalences and correlates of mental disorders. WHO International Consortium in Psychiatric Epidemiology. Bull World Health Organ. 2000;78:413–26.

189. Chevalier A, Feinstein L. Sheepskin or Prozac: The Causal Effect of Education on Mental Health. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network; 2006. https://papers.ssrn.com/abstract=923530. Accessed 17 Apr 2021.

190. McLaughlin KA, Alvarez K, Fillbrunn M, Green JG, Jackson JS, Kessler RC, et al. Racial/ethnic variation in trauma-related psychopathology in the United States: a population-based study. Psychol Med. 2019;49:2215–26.

191. Gillespie CF, Bradley B, Mercer K, Smith AK, Conneely K, Gapen M, et al. Trauma exposure and stress-related disorders in inner city primary care patients. Gen Hosp Psychiatry. 2009;31:505–14.

192. Holbrook TL, Hoyt DB, Stein MB, Sieber WJ. Gender differences in long-term posttraumatic stress disorder outcomes after major trauma: women are at higher risk of adverse outcomes than men. J Trauma. 2002;53:882–8.

193. Seedat S, Scott KM, Angermeyer MC, Berglund P, Bromet EJ, Brugha TS, et al. Cross-national associations between gender and mental disorders in the World Health Organization World Mental Health Surveys. Arch Gen Psychiatry. 2009;66:785–95.

194. Golub Y, Kaltwasser SF, Mauch CP, Herrmann L, Schmidt U, Holsboer F, et al. Reduced hippocampus volume in the mouse model of Posttraumatic Stress Disorder. J Psychiatr Res. 2011;45:650–9.

195. Rosenkranz JA, Venheim ER, Padival M. Chronic Stress Causes Amygdala Hyperexcitability in Rodents. Biol Psychiatry. 2010;67:1128–36.

196. Gilbert SJ, Spengler S, Simons JS, Steele JD, Lawrie SM, Frith CD, et al. Functional specialization within rostral prefrontal cortex (area 10): a meta-analysis. J Cogn Neurosci. 2006;18:932–48.

197. Tsujimoto S, Genovesio A, Wise SP. Frontal pole cortex: encoding ends at the end of the endbrain. Trends Cogn Sci. 2011;15:169–76.

198. Stevens FL, Hurley RA, Taber KH, Hurley RA, Hayman LA, Taber KH. Anterior Cingulate Cortex: Unique Role in Cognition and Emotion. J Neuropsychiatry Clin Neurosci. 2011;23:121–5.

199. Bechara A, Damasio H, Damasio AR. Emotion, Decision Making and the Orbitofrontal Cortex. Cereb Cortex. 2000;10:295–307.

200. Glahn DC, Thompson PM, Blangero J. Neuroimaging endophenotypes: Strategies for finding genes influencing brain structure and function. Hum Brain Mapp. 2007;28:488–501.

201. Doom JR, Gunnar MR. Stress physiology and developmental psychopathology: Past, present and future. Dev Psychopathol. 2013;25 4 0 2. doi:10.1017/S0954579413000667.

202. O'Connor TG, Moynihan JA, Caserta MT. Annual Research Review: The neuroinflammation hypothesis for stress and psychopathology in children – developmental psychoneuroimmunology. J Child Psychol Psychiatry. 2014;55:615–31.

203. Dirven BCJ, Homberg JR, Kozicz T, Henckens MJ a. G. Epigenetic programming of the neuroendocrine stress response by adult life stress. J Mol Endocrinol. 2017;59:R11–31.

204. Clayton DF, Anreiter I, Aristizabal M, Frankland PW, Binder EB, Citri A. The role of the genome in experience-dependent plasticity: Extending the analogy of the genomic action potential. Proc Natl Acad Sci. 2020;117:23252–60.

205. Klengel T, Binder EB. Epigenetics of Stress-Related Psychiatric Disorders and Gene × Environment Interactions. Neuron. 2015;86:1343–57.

206. Heim CM, Entringer S, Buss C. Translating basic research knowledge on the biological embedding of early-life stress into novel approaches for the developmental programming of lifelong health. Psychoneuroendocrinology. 2019;105:123–37.

207. Miller GE, Chen E, Parker KJ. Psychological Stress in Childhood and Susceptibility to the Chronic Diseases of Aging: Moving Towards a Model of Behavioral and Biological Mechanisms. Psychol Bull. 2011;137:959–97.

208. Kilaru V, Iyer SV, Almli LM, Stevens JS, Lori A, Jovanovic T, et al. Genome-wide gene-based analysis suggests an association between Neuroligin 1 (NLGN1) and post-traumatic stress disorder. Transl Psychiatry. 2016;6:e820–e820.

209. Bernstein DP, Stein JA, Newcomb MD, Walker E, Pogge D, Ahluvalia T, et al. Development and validation of a brief screening version of the Childhood Trauma Questionnaire. Child Abuse Negl. 2003;27:169–90.

210. Dunn EC, Nishimi K, Powers A, Bradley B. Is developmental timing of trauma exposure associated with depressive and post-traumatic stress disorder symptoms in adulthood? J Psychiatr Res. 2017;84:119–27.

211. Binder EB, Bradley RG, Liu W, Epstein MP, Deveau TC, Mercer KB, et al. Association of FKBP5 Polymorphisms and Childhood Abuse With Risk of Posttraumatic Stress Disorder Symptoms in Adults. JAMA. 2008;299:1291–305.

212. Powers A, Ressler KJ, Bradley RG. The protective role of friendship on the effects of childhood abuse and depression. Depress Anxiety. 2009;26:46–53.

213. Wierenga LM, Langen M, Oranje B, Durston S. Unique developmental trajectories of cortical thickness and surface area. NeuroImage. 2014;87:120–6.

214. Panizzon MS, Fennema-Notestine C, Eyler LT, Jernigan TL, Prom-Wormley E, Neale M, et al. Distinct genetic influences on cortical surface area and cortical thickness. Cereb Cortex N Y N 1991. 2009;19:2728–35.

215. Stevens JS, Jovanovic T, Fani N, Ely TD, Glover EM, Bradley B, et al. Disrupted amygdala-prefrontal functional connectivity in civilian women with posttraumatic stress disorder. J Psychiatr Res. 2013;47:1469–78.

216. Stevens JS, Kim YJ, Galatzer-Levy IR, Reddy R, Ely TD, Nemeroff CB, et al. Amygdala Reactivity and Anterior Cingulate Habituation Predict Posttraumatic Stress Disorder Symptom Maintenance After Acute Civilian Trauma. Biol Psychiatry. 2017;81:1023–9.

217. Barfield RT, Kilaru V, Smith AK, Conneely KN. CpGassoc: an R function for analysis of DNA methylation microarray data. Bioinforma Oxf Engl. 2012;28:1280–1.

218. McCartney DL, Walker RM, Morris SW, McIntosh AM, Porteous DJ, Evans KL. Identification of polymorphic and off-target probe binding sites on the Illumina Infinium MethylationEPIC BeadChip. Genomics Data. 2016;9:22–4.

219. Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, et al. A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinforma Oxf Engl. 2013;29:189–96.

220. Pidsley R, Y Wong CC, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. BMC Genomics. 2013;14:293.

221. Barcelona V, Huang Y, Brown K, Liu J, Zhao W, Yu M, et al. Novel DNA methylation sites associated with cigarette smoking among African Americans. Epigenetics. 2019;14:383–91.

222. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics. 2012;13:86.

223. Meyers JL, Salling MC, Almli LM, Ratanatharathorn A, Uddin M, Galea S, et al. Frequency of alcohol consumption in humans; the role of metabotropic glutamate receptors and downstream signaling pathways. Transl Psychiatry. 2015;5:e586–e586.

224. Joehanes Roby, Just Allan C., Marioni Riccardo E., Pilling Luke C., Reynolds Lindsay M., Mandaviya Pooja R., et al. Epigenetic Signatures of Cigarette Smoking. Circ Cardiovasc Genet. 2016;9:436–47.

225. Logue MW, Miller MW, Wolf EJ, Huber BR, Morrison FG, Zhou Z, et al. An epigenome-wide association study of posttraumatic stress disorder in US veterans implicates several new DNA methylation loci. Clin Epigenetics. 2020;12:46.

226. Katrinli S, Stevens J, Wani AH, Lori A, Kilaru V, van Rooij SJH, et al. Evaluating the impact of trauma and PTSD on epigenetic prediction of lifespan and neural integrity. Neuropsychopharmacology. 2020;:1–8.

227. Barfield R, Shen J, Just AC, Vokonas PS, Schwartz J, Baccarelli AA, et al. Testing for the indirect effect under the null for genome-wide mediation analyses. Genet Epidemiol. 2017;41:824–33.

228. Brito NH, Noble KG. Socioeconomic status and structural brain development. Front Neurosci. 2014;8. doi:10.3389/fnins.2014.00276.

229. Sipahi L, Wildman DE, Aiello AE, Koenen KC, Galea S, Abbas A, et al. Longitudinal epigenetic variation of DNA methyltransferase genes associated with vulnerability to post-traumatic stress disorder (PTSD). Psychol Med. 2014;44:3165–79.

230. Lewis C. Major Depressive Disorder Working Group of the Psychiatric Genomics Consortium. 2020.

231. Feng J, Zhou Y, Campbell SL, Le T, Li E, Sweatt JD, et al. Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. Nat Neurosci. 2010;13:423–30.

232. LaPlant Q, Vialou V, Covington HE, Dumitriu D, Feng J, Warren BL, et al. Dnmt3a regulates emotional behavior and spine plasticity in the nucleus accumbens. Nat Neurosci. 2010;13:1137–43.

233. Xu S, Sui J, Yang S, Liu Y, Wang Y, Liang G. Integrative analysis of competing endogenous RNA network focusing on long noncoding RNA associated with progression of cutaneous melanoma. Cancer Med. 2018;7:1019–29.

234. Li H, Jia Y, Cheng J, Liu G, Song F. LncRNA NCK1-AS1 promotes proliferation and induces cell cycle progression by crosstalk NCK1-AS1/miR-6857/CDK1 pathway. Cell Death Dis. 2018;9:1–15.

235. Mirza AH, Berthelsen CH, Seemann SE, Pan X, Frederiksen KS, Vilien M, et al. Transcriptomic landscape of lncRNAs in inflammatory bowel disease. Genome Med. 2015;7:39.

236. Taudien S, Galgoczy P, Huse K, Reichwald K, Schilhabel M, Szafranski K, et al. Polymorphic segmental duplications at 8p23.1 challenge the determination of individual defensin gene repertoires and the assembly of a contiguous human reference sequence. BMC Genomics. 2004;5:92.

237. Huang Y, Chen Z. Inflammatory bowel disease related innate immunity and adaptive immunity. Am J Transl Res. 2016;8:2490–7.

238. Danese A, Caspi A, Williams B, Ambler A, Sugden K, Mika J, et al. Biological embedding of stress through inflammation processes in childhood. Mol Psychiatry. 2011;16:244–6.

239. Baumeister D, Akhtar R, Ciufolini S, Pariante CM, Mondelli V. Childhood trauma and adulthood inflammation: a meta-analysis of peripheral C-reactive protein, interleukin-6 and tumour necrosis factor-α. Mol Psychiatry. 2016;21:642–9.

240. Danese A, Pariante CM, Caspi A, Taylor A, Poulton R. Childhood maltreatment predicts adult inflammation in a life-course study. Proc Natl Acad Sci U S A. 2007;104:1319–24.

241. Patel N, Crider A, Pandya CD, Ahmed AO, Pillai A. Altered mRNA Levels of Glucocorticoid Receptor, Mineralocorticoid Receptor, and Co-Chaperones (FKBP5 and PTGES3) in the Middle Frontal Gyrus of Autism Spectrum Disorder Subjects. Mol Neurobiol. 2016;53:2090–9.

242. Sinclair D, Tsai SY, Woon HG, Weickert CS. Abnormal Glucocorticoid Receptor mRNA and Protein Isoform Expression in the Prefrontal Cortex in Psychiatric Illness. Neuropsychopharmacology. 2011;36:2698–709.

243. Sinclair D, Fillman SG, Webster MJ, Weickert CS. Dysregulation of glucocorticoid receptor co-factors FKBP5, BAG1 and PTGES3 in prefrontal cortex in psychotic illness. Sci Rep. 2013;3. doi:10.1038/srep03539.

244. Meier TB, Drevets WC, Wurfel BE, Ford BN, Morris HM, Victor TA, et al. Relationship between neurotoxic kynurenine metabolites and reductions in right medial prefrontal cortical thickness in major depressive disorder. Brain Behav Immun. 2016;53:39–48.

245. Gresset A, Sondek J, Harden TK. The Phospholipase C Isozymes and Their Regulation. Subcell Biochem. 2012;58:61–94.

246. Hu X-T, Zhang F-B, Fan Y-C, Shu X-S, Wong AHY, Zhou W, et al. Phospholipase C delta 1 is a novel 3p22.3 tumor suppressor involved in cytoskeleton organization, with its epigenetic silencing correlated with high-stage gastric cancer. Oncogene. 2009;28:2466–75.

247. Ichinohe M, Nakamura Y, Sai K, Nakahara M, Yamaguchi H, Fukami K. Lack of phospholipase C-δ1 induces skin inflammation. Biochem Biophys Res Commun. 2007;356:912–8.

248. Stallings JD, Tall EG, Pentyala S, Rebecchi MJ. Nuclear Translocation of Phospholipase C-δ1 Is Linked to the Cell Cycle and Nuclear Phosphatidylinositol 4,5-Bisphosphate. J Biol Chem. 2005;280:22060–9.

249. Yu H, Fukami K, Watanabe Y, Ozaki C, Takenawa T. Phosphatidylinositol 4,5-bisphosphate reverses the inhibition of RNA transcription caused by histone H1. Eur J Biochem. 1998;251:281–7.

250. Udawela M, Scarr E, Boer S, Um JY, Hannan AJ, McOmish C, et al. Isoform specific differences in phospholipase C beta 1 expression in the prefrontal cortex in schizophrenia and suicide. Npj Schizophr. 2017;3:1–9.

251. Su K-P, Yang H-T, Chang JP-C, Shih Y-H, Guu T-W, Kumaran SS, et al. Eicosapentaenoic and docosahexaenoic acids have different effects on peripheral phospholipase A2 gene expressions in acute depressed patients. Prog Neuropsychopharmacol Biol Psychiatry. 2018;80:227–33.

252. Yamawaki Y, Shirawachi S, Mizokami A, Nozaki K, Ito H, Asano S, et al. Phospholipase C-related catalytically inactive protein regulates lipopolysaccharide-induced hypothalamic inflammation-mediated anorexia in mice. Neurochem Int. 2019;131:104563.

253. Vilsaint CL, NeMoyer A, Fillbrunn M, Sadikova E, Kessler RC, Sampson NA, et al. Racial/Ethnic Differences in 12-Month Prevalence and Persistence of Mood, Anxiety, and Substance Use Disorders: Variation by Nativity and Socioeconomic Status. Compr Psychiatry. 2019;89:52–60.

254. Koenen KC, Uddin M, Chang S-C, Aiello AE, Wildman DE, Goldmann E, et al. SLC6A4 methylation modifies the effect of the number of traumatic events on risk for posttraumatic stress disorder. Depress Anxiety. 2011;28:639–47.

255. Matosin N, Cruceanu C, Binder EB. Preclinical and Clinical Evidence of DNA Methylation Changes in Response to Trauma and Chronic Stress. Chronic Stress. 2017;1:2470547017710764.

256. Furukawa-Hibi Y, Nagai T, Yun J, Yamada K. Stress increases DNA methylation of the neuronal PAS domain 4 (Npas4) gene. Neuroreport. 2015;26:827–32.

257. Mifsud KR, Saunderson EA, Spiers H, Carter SD, Trollope AF, Mill J, et al. Rapid Down-Regulation of Glucocorticoid Receptor Gene Expression in the Dentate Gyrus after Acute Stress in vivo: Role of DNA Methylation and MicroRNA Activity. Neuroendocrinology. 2017;104:157–69.

258. Weaver ICG, Cervoni N, Champagne FA, D'Alessio AC, Sharma S, Seckl JR, et al. Epigenetic programming by maternal behavior. Nat Neurosci. 2004;7:847–54.

259. Tate PH, Bird AP. Effects of DNA methylation on DNA-binding proteins and gene expression. Curr Opin Genet Dev. 1993;3:226–31.

260. Pfeiffer JR, Bustamante AC, Kim GS, Armstrong D, Knodt AR, Koenen KC, et al. Associations between childhood family emotional health, fronto-limbic grey matter volume, and saliva 5mC in young adulthood. Clin Epigenetics. 2021;13:68.

261. Pfeiffer J, Stevens J, Smith A, Mekawi Y, Uddin M. Blood DNA methylation clusters associated with adverse social exposures and endophenotypes of PTSD-related psychiatric illness in a female, trauma-exposed cohort. Prep. 2021.

262. Zhang Y, Chen K, Sloan SA, Bennett ML, Scholze AR, O'Keeffe S, et al. An RNA-Sequencing Transcriptome and Splicing Database of Glia, Neurons, and Vascular Cells of the Cerebral Cortex. J Neurosci. 2014;34:11929–47.

263. GRCm38 - mm10 - Genome - Assembly - NCBI. https://www.ncbi.nlm.nih.gov/assembly/GCF_000001635.20/. Accessed 19 Mar 2021.

264. GEO Accession viewer. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE95740. Accessed 19 Mar 2021.

265. Schultz MD, He Y, Whitaker JW, Hariharan M, Mukamel EA, Leung D, et al. Human Body Epigenome Maps Reveal Noncanonical DNA Methylation Variation. Nature. 2015;523:212–6.

266. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26:139–40.

267. Hwang Y, Kim J, Shin J-Y, Kim J-I, Seo J-S, Webster MJ, et al. Gene expression profiling by mRNA sequencing reveals increased expression of immune/inflammation-related genes in the hippocampus of individuals with schizophrenia. Transl Psychiatry. 2013;3:e321–e321.

268. Breen MS, Maihofer AX, Glatt SJ, Tylee DS, Chandler SD, Tsuang MT, et al. Gene networks specific for innate immunity define post-traumatic stress disorder. Mol Psychiatry. 2015;20:1538–45.

269. Miao Z, Fengbiao M, Jialong L, Szyf M, Wang Y, Sun Z. Anxiety-Related Behaviours Associated with microRNA-206-3p and BDNF Expression in Pregnant Female Mice Following Psychological Social Stress. Mol Neurobiol. 2018;55.

270. Alexa A, Rahnenfuhrer J. topGO: Enrichment Analysis for Gene Ontology. Bioconductor version: Release (3.12); 2021. doi:10.18129/B9.bioc.topGO.

271. Welch RP, Lee C, Imbriano PM, Patil S, Weymouth TE, Smith RA, et al. ChIP-Enrich: gene set enrichment testing for ChIP-seq data. Nucleic Acids Res. 2014;42:e105–e105.

272. TxDb.Mmusculus.UCSC.mm10.knownGene. Bioconductor. http://bioconductor.org/packages/TxDb.Mmusculus.UCSC.mm10.knownGene/. Accessed 19 Mar 2021.

273. BSgenome.Mmusculus.UCSC.mm10. Bioconductor. http://bioconductor.org/packages/BSgenome.Mmusculus.UCSC.mm10/. Accessed 19 Mar 2021.

274. Lister R, Mukamel EA, Nery JR, Urich M, Puddifoot CA, Johnson ND, et al. Global epigenomic reconfiguration during mammalian brain development. Science. 2013;341:1237905.

275. Durinck S, Spellman PT, Birney E, Huber W. Mapping Identifiers for the Integration of Genomic Datasets with the R/Bioconductor package biomaRt. Nat Protoc. 2009;4:1184–91.

276. GRCh38 - hg38 - Genome - Assembly - NCBI. https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.26/. Accessed 19 Mar 2021.

277. liftOver. Bioconductor. http://bioconductor.org/packages/liftOver/. Accessed 19 Mar 2021.

278. Thomas GM, Huganir RL. MAPK cascade signalling and synaptic plasticity. Nat Rev Neurosci. 2004;5:173–83.

279. Lake D, Corrêa SAL, Müller J. Negative feedback regulation of the ERK1/2 MAPK pathway. Cell Mol Life Sci. 2016;73:4397–413.

280. Caunt CJ, Keyse SM. Dual-specificity MAP kinase phosphatases (MKPs): shaping the outcome of MAP kinase signalling. FEBS J. 2013;280:489–504.

281. Jordan CT, Cao L, Roberson EDO, Pierson KC, Yang C-F, Joyce CE, et al. PSORS2 is due to mutations in CARD14. Am J Hum Genet. 2012;90:784–95.

282. Acquaah-Mensah GK, Agu N, Khan T, Gardner A. A Regulatory Role for the Insulin- and BDNF-Linked RORA in the Hippocampus: Implications for Alzheimer's Disease. J Alzheimers Dis. 2015;44:827–38.

283. Akazawa H, Komuro I. Cardiac transcription factor Csx/Nkx2-5: Its role in cardiac development and diseases. Pharmacol Ther. 2005;107:252–68.

284. Prall OWJ, Menon MK, Solloway MJ, Watanabe Y, Zaffran S, Bajolle F, et al. An Nkx2-5/Bmp2/Smad1 Negative Feedback Loop Controls Heart Progenitor Specification and Proliferation. Cell. 2007;128:947–59.

285. Riazi AM, Lee H, Hsu C, Van Arsdell G. CSX/Nkx2.5 Modulates Differentiation of Skeletal Myoblasts and Promotes Differentiation into Neuronal Cells in Vitro*. J Biol Chem. 2005;280:10716–20.

286. Rivkees SA, Chen M, Kulkarni J, Browne J, Zhao Z. Characterization of the Murine A1 Adenosine Receptor Promoter, Potent Regulation by GATA-4 and Nkx2.5*. J Biol Chem. 1999;274:14204–9.

287. Boukhtouche F, Doulazmi M, Frederic F, Dusart I, Brugg B, Mariani J. RORalpha, a pivotal nuclear receptor for Purkinje neuron survival and differentiation: from development to ageing. Cerebellum Lond Engl. 2006;5:97–104.

288. Terracciano A, Tanaka T, Sutin AR, Sanna S, Deiana B, Lai S, et al. Genome-wide association scan of trait depression. Biol Psychiatry. 2010;68:811–7.

289. Maglione JE, Nievergelt CM, Parimi N, Evans DS, Ancoli-Israel S, Stone KL, et al. Associations of PER3 and RORA Circadian Gene Polymorphisms and Depressive Symptoms in Older Adults. Am J Geriatr Psychiatry Off J Am Assoc Geriatr Psychiatry. 2015;23:1075–87.

290. Lavebratt C, Sjöholm LK, Partonen T, Schalling M, Forsell Y. PER2 variantion is associated with depression vulnerability. Am J Med Genet Part B Neuropsychiatr Genet Off Publ Int Soc Psychiatr Genet. 2010;153B:570–81.

291. Miller MW, Wolf EJ, Logue MW, Baldwin CT. The retinoid-related orphan receptor alpha (RORA) gene and fear-related psychopathology. J Affect Disord. 2013;151:702–8.

292. Min J-A, Lee H-J, Lee S-H, Park Y-M, Kang S-G, Park Y-G, et al. RORA Polymorphism Interacts with Childhood Maltreatment in Determining Anxiety Sensitivity by Sex: A Preliminary Study in Healthy Young Adults. Clin Psychopharmacol Neurosci Off Sci J Korean Coll Neuropsychopharmacol. 2017;15:402–6.

293. Sarachana T, Xu M, Wu R-C, Hu VW. Sex Hormones in Autism: Androgens and Estrogens Differentially and Reciprocally Regulate RORA, a Novel Candidate Gene for Autism. PLoS ONE. 2011;6:e17116.

294. Nguyen A, Rauch TA, Pfeifer GP, Hu VW. Global methylation profiling of lymphoblastoid cell lines reveals epigenetic contributions to autism spectrum disorders and a novel autism candidate gene, RORA, whose protein product is reduced in autistic brain. FASEB J Off Publ Fed Am Soc Exp Biol. 2010;24:3036–51.

295. Lin Y-MJ, Hsin I-L, Sun HS, Lin S, Lai Y-L, Chen H-Y, et al. NTF3 Is a Novel Target Gene of the Transcription Factor POU3F2 and Is Required for Neuronal Differentiation. Mol Neurobiol. 2018;55:8403–13.

296. Hashizume K, Yamanaka M, Ueda S. POU3F2 participates in cognitive function and adult hippocampal neurogenesis via mammalian-characteristic amino acid repeats. Genes Brain Behav. 2018;17:118–25.

297. Mandela P, Ma X-M. Kalirin, a key player in synapse formation, is implicated in human diseases. Neural Plast. 2012;2012:728161.

298. Parnell E, Shapiro LP, Voorn RA, Forrest MP, Jalloul HA, Loizzo DD, et al. KALRN: A central regulator of synaptic function and synaptopathies. Gene. 2021;768:145306.

299. Takeuchi T, Misaki A, Liang SB, Tachibana A, Hayashi N, Sonobe H, et al. Expression of T-cadherin (CDH13, H-Cadherin) in human brain and its characteristics as a negative growth regulator of epidermal growth factor in neuroblastoma cells. J Neurochem. 2000;74:1489–97.

300. Beckervordersandforth R, Zhang C-L, Lie DC. Transcription-Factor-Dependent Control of Adult Hippocampal Neurogenesis. Cold Spring Harb Perspect Biol. 2015;7. doi:10.1101/cshperspect.a018879.

301. Wu M, Chen G, Li Y-P. TGF-β and BMP signaling in osteoblast, skeletal development, and bone formation, homeostasis and disease. Bone Res. 2016;4:1–21.

302. Lee-Hoeflich ST, Causing CG, Podkowa M, Zhao X, Wrana JL, Attisano L. Activation of LIMK1 by binding to the BMP receptor, BMPRII, regulates BMP-dependent dendritogenesis. EMBO J. 2004;23:4792–801.

303. Wen Z, Han L, Bamburg JR, Shim S, Ming G, Zheng JQ. BMP gradients steer nerve growth cones by a balancing act of LIM kinase and Slingshot phosphatase on ADF/cofilin. J Cell Biol. 2007;178:107–19.

304. Eaton BA, Davis GW. LIM Kinase1 controls synaptic stability downstream of the type II BMP receptor. Neuron. 2005;47:695–708.

305. Moustakas A, Heldin C-H. Dynamic control of TGF-β signaling and its links to the cytoskeleton. FEBS Lett. 2008;582:2051–65.

306. Kostenko S, Shiryaev A, Dumitriu G, Gerits N, Moens U. Cross-talk between protein kinase A and the MAPK-activated protein kinases RSK1 and MK5. J Recept Signal Transduct. 2011;31:1–9.

307. Montaño CM, Irizarry RA, Kaufmann WE, Talbot K, Gur RE, Feinberg AP, et al. Measuring cell-type specific differential methylation in human brain tissue. Genome Biol. 2013;14:R94.

308. Jaffe AE, Irizarry RA. Accounting for cellular heterogeneity is critical in epigenome-wide association studies. Genome Biol. 2014;15:R31.

309. Gogtay N, Giedd JN, Lusk L, Hayashi KM, Greenstein D, Vaituzis AC, et al. Dynamic mapping of human cortical development during childhood through early adulthood. Proc Natl Acad Sci. 2004;101:8174–9.

310. Deng W, Aimone JB, Gage FH. New neurons and new memories: how does adult hippocampal neurogenesis affect learning and memory? Nat Rev Neurosci. 2010;11:339–50.

311. Gibb SJ, Fergusson DM, Horwood LJ. Childhood family income and life outcomes in adulthood: Findings from a 30-year longitudinal study in New Zealand. Soc Sci Med. 2012;74:1979–86.

312. Fumagalli S, Perego C, Pischiutta F, Zanier ER, De Simoni M-G. The Ischemic Environment Drives Microglia and Macrophage Function. Front Neurol. 2015;6. doi:10.3389/fneur.2015.00081.

313. Kabba JA, Xu Y, Christian H, Ruan W, Chenai K, Xiang Y, et al. Microglia: Housekeeper of the Central Nervous System. Cell Mol Neurobiol. 2018;38:53–71.

314. Baylin SB, Jones PA. A decade of exploring the cancer epigenome — biological and translational implications. Nat Rev Cancer. 2011;11:726–34.

315. NIH/NCI. NIH/NCI 428 - Cloud-Based Multi-Omic and Imaging Software for the Cancer Research Data Commons | NCI: SBIR & STTR. https://sbir.cancer.gov/funding/contracts/428. Accessed 27 Apr 2021.

316. Rapaport MH. Prevalence, recognition, and treatment of comorbid depression and anxiety. J Clin Psychiatry. 2001;62 Suppl 24:6–10.

317. Wium-Andersen MK, Ørsted DD, Nielsen SF, Nordestgaard BG. Elevated C-Reactive Protein Levels, Psychological Distress, and Depression in 73 131 Individuals. JAMA Psychiatry. 2013;70:176–84.

318. Kim GS, Smith AK, Xue F, Michopoulos V, Lori A, Armstrong DL, et al. Methylomic profiles reveal sex-specific differences in leukocyte composition associated with post-traumatic stress disorder. Brain Behav Immun. 2019;81:280–91.

319. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. Am J Psychiatry. 2010;167:748–51.

320. Insel TR. The NIMH Research Domain Criteria (RDoC) Project: precision medicine for psychiatry. Am J Psychiatry. 2014;171:395–7.

APPENDIX A: SUPPLEMENTARY MATERIALS – CHAPTER 2

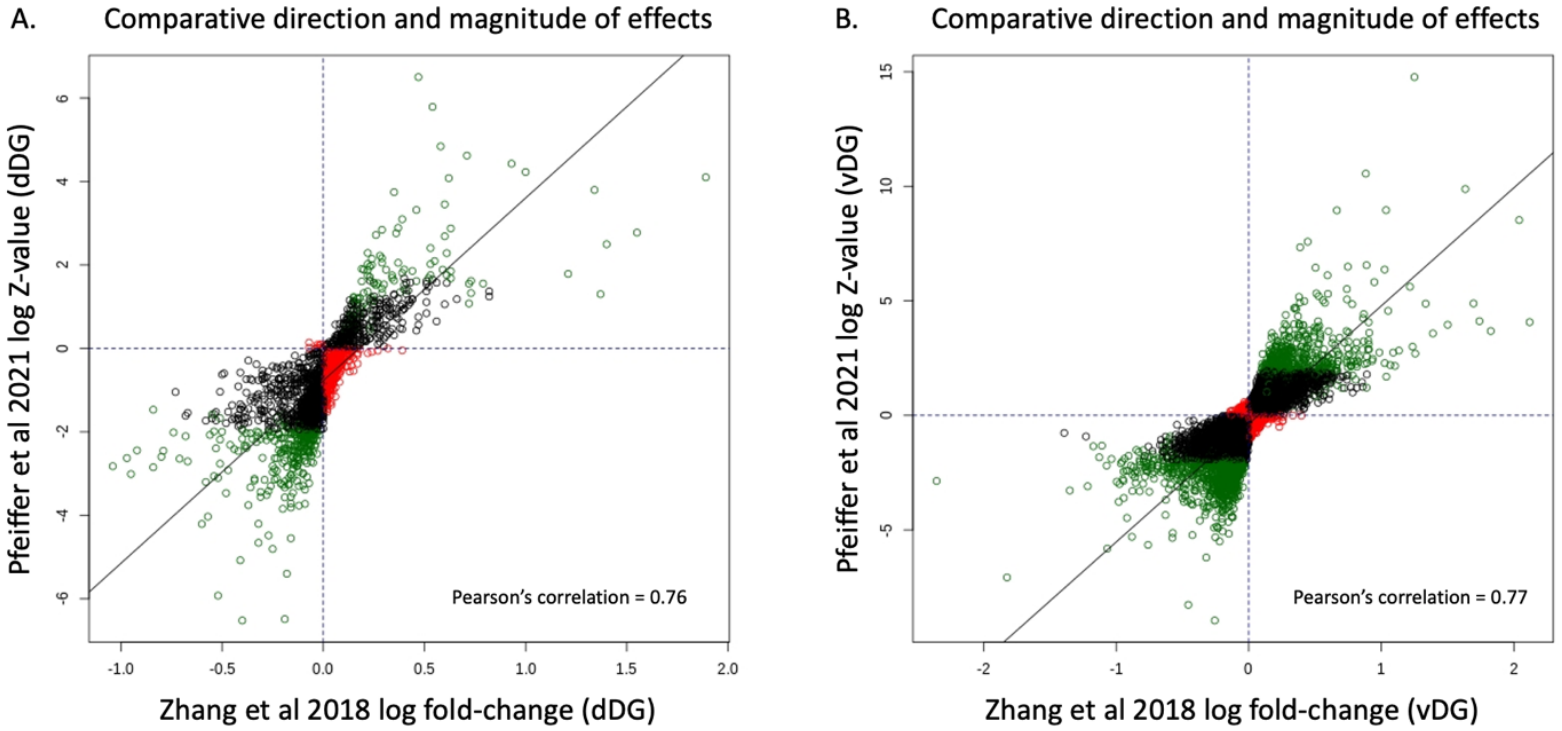
The supplementary excel file ("Chapter2_SupplementalTables_JP.xlsx") includes results from numerous analyses performed throughout Chapter 2. The "SuppTabA1.MentalHealthDX" tab (Supplementary Table A.1) contains a cohort-level summary of lifetime mental health disorder diagnoses. The "SuppTabA2.FHQquestionnaire" tab (Supplementary Table A.2) contains the questions used to generate our retrospective measure of FEH. The "SuppTabA3.FEHpredictsMEs" tab contains Supplementary Table A.3 and includes results from Arm B analyses where family emotional health (FEH) was used as an independent variable to predict module eigengene (ME) value. The "SuppTabA4.MEpredBRV" tab (Supplementary Table A.4) contains results from Arm C analyses where MEs were used as independent variables to predict brain region volume (BRV) of fronto-limbic brain regions. The "SuppTabA5.ProbeSpecificAnnot" tab (Supplementary Table A.5) contains annotation information for probes from partially and fully mediating modules. The "SuppTabA6.ProbeMediation" tab (Supplementary Table A.6) contains results from probe-wise mediation analyses. The "SuppTabA7.CondensedGSEA" tab (Supplementary Table A.7) contains Revigo-condensed GSEA results of the family emotional health-associated methylome.

APPENDIX B: SUPPLEMENTARY MATERIALS – CHAPTER 3

The supplementary excel file ("Chapter3_SupplementalTables_JP.xlsx") includes results from numerous analyses performed throughout Chapter 3. The "SuppTabB1.ASEpredictME" tab contains Supplementary Table B.1 and includes results from Arm B analyses where adverse social exposures (ASEs) were used as independent variables to predict module eigengene (ME) value. The "SuppTabB2.MEpredictFLBM" tab (Supplementary Table B.2) contains results from Arm C analyses where MEs were used as independent variables to fronto-limbic brain morphology measures. The "SuppTabB3.ProbeMediationFull" tab (Supplementary Table B.3) results from probe-wise mediation analyses from modules that were full mediators. The "SuppTabB4.CondensedGSEA" tab (Supplementary Table B.4) contains Revigo-condensed GSEA results of the childhood trauma burden-associated methylome.

APPENDIX C: SUPPLEMENTARY MATERIALS – CHAPTER 4

C.1 Supplementary Figures and Tables



Supplementary Figure C.1 Comparison of Zhang et al 2018 log fold-change difference between EE and SH groups in dDG (A), and vDG (B) versus log Z-value derived from hypothesis tests of environmental enrichment effect on mRNA expression. Plots show strong concordance in direction of effect and magnitude of effect between the two methodologies (Zhang vs. Pfeiffer).

The supplementary excel file ("Chapter4_SupplementalTables_JP.xlsx") includes results from numerous analyses performed throughout Chapter 4.

The "SuppTabC1.dDG_RNAseq" tab contains Supplementary Table C.1 and includes nominally significant results from RNAseq analyses performed in dDG (Arm A). The "SuppTabC2.dDG_GSEA" tab contains Supplementary Table C.2 and includes nominally significant results from GSEA of DEGs from Arm A in dDG. The "SuppTabC3.vDG_RNAseq" tab contains Supplementary Table C.3 and includes nominally significant results from RNAseq analyses performed in vDG (Arm A). The "SuppTabC4.vDG_GSEA" tab contains Supplementary Table C.4 and includes nominally significant results from GSEA of DEGs from Arm A in vDG.

The "SuppTabC5.dDG_mCG" tab contains Supplementary Table C.5 and includes nominally significant differential methylation (mCG) results as performed in dDG (Arm B). The "SuppTabC6.vDG_mCG" tab contains Supplementary Table C.6 and includes nominally significant differential methylation (mCG) results as performed in vDG (Arm B). The "SuppTabC7.dDG_mCH" tab contains Supplementary Table C.7 and includes nominally significant differential methylation (mCH) results as performed in dDG (Arm B). The "SuppTabC8.vDG_mCH" tab contains Supplementary Table C.8 and includes nominally significant differential methylation (mCH) results as performed in vDG (Arm B).

The "SuppTabC9.dDG_mCG_ArmC" tab contains Supplementary Table C.9 and includes all results from Arm C analyses assessing the effect of mCG on dDG mRNA transcription. The "SuppTabC10.vDG_mCG_ArmC" tab contains Supplementary Table C.10 and includes all results from Arm C analyses assessing the effect of mCG on vDG mRNA

transcription. The “SuppTabC11.dDG_mCG_ArmC” tab contains Supplementary Table C.11 and includes all results from Arm C analyses assessing the effect of mCH on dDG mRNA transcription. The “SuppTabC12.vDG_mCG_ArmC” tab contains Supplementary Table C.12 and includes all results from Arm C analyses assessing the effect of mCH on vDG mRNA transcription.

The “SuppTabC13.dDG_mCG_mediation” tab contains Supplementary Table C.13 and includes all results from mediation analyses testing the mediating status of locus-specific mCG sites between environmental enrichment and mRNA transcription in dDG. The “SuppTabC14.vDG_mCG_mediation” tab contains Supplementary Table C.14 and includes all results from mediation analyses testing the mediating status of locus-specific mCG sites between environmental enrichment and mRNA transcription in vDG.

The “SuppTabC15.ortho_mapping” tab contains Supplementary Table C.15 and includes results from analyses mapping human loci to mouse orthologue genes. The “SuppTabC16.dDG_ortho_mapping” tab contains Supplementary Table C.16 and includes results from analyses where we first identified sequence specific orthologue mappings between human and mouse loci, and then used those mappings to assess whether the loci were differentially methylated in mouse dDG. The “SuppTabC17.vDG_ortho_mapping” tab contains Supplementary Table C.17 and includes results from analyses where we first identified sequence specific orthologue mappings between human and mouse loci, and then used those mappings to assess whether the loci were differentially methylated in mouse vDG.