



Gene by environment interaction mouse model reveals a functional role for 5-hydroxymethylcytosine in neurodevelopmental disorders

Ligia A Papale, Andy Madrid, Qi Zhang, et al.

Genome Res. published online December 23, 2021

Access the most recent version at doi:[10.1101/gr.276137.121](https://doi.org/10.1101/gr.276137.121)

P<P	Published online December 23, 2021 in advance of the print journal.
Accepted Manuscript	Peer-reviewed and accepted for publication but not copyedited or typeset; accepted manuscript is likely to differ from the final, published version.
Creative Commons License	This article is distributed exclusively by Cold Spring Harbor Laboratory Press for the first six months after the full-issue publication date (see https://genome.cshlp.org/site/misc/terms.xhtml). After six months, it is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at http://creativecommons.org/licenses/by-nc/4.0/ .
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here .



To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Published by Cold Spring Harbor Laboratory Press

Gene by environment interaction mouse model reveals a functional role for 5-hydroxymethylcytosine in neurodevelopmental disorders

Ligia A. Papale^{1a}, Andy Madrid^{1,3a}, Qi Zhang^{4a}, Kailei Chen², Lara Sak¹, Sündüz Keleş², and Reid S. Alisch^{1*}

Departments of ¹Neurological Surgery, ²Statistics, Biostatistics, and Medical Informatics, and

³Neuroscience training program, University of Wisconsin, Madison, Wisconsin 53719, USA.

⁴Department Mathematics and Statistics, University of New Hampshire, Durham, NH 03824, USA.

*Correspondence to:

Reid S. Alisch

E-mail: alisch@wisc.edu

Phone: (608) 263-3793

Running Head: 5hmC role in neurodevelopmental disorders

^aThese authors contributed equally to this work.

Abstract

Mouse knockouts of *Cntnap2* exhibit altered neurodevelopmental behavior, deficits in striatal GABAergic signaling and a genome-wide disruption of an environmentally sensitive DNA methylation modification (5-hydroxymethylcytosine, 5hmC) in the orthologs of a significant number of genes implicated in human neurodevelopmental disorders. We tested adult *Cntnap2* heterozygous mice (*Cntnap2*^{+/-}, lacking behavioral or neuropathological abnormalities) subjected to a prenatal stress and found that prenatally stressed *Cntnap2*^{+/-} female mice showed repetitive behaviors and altered sociability, similar to the homozygote phenotype. Genomic profiling revealed disruptions in hippocampal and striatal 5hmC levels that were correlated to altered transcript levels of genes linked to these phenotypes (*e.g.*, *Reln*, *Dst*, *Trio* and *Epha5*). Chromatin-immunoprecipitation coupled with high-throughput sequencing and hippocampal nuclear lysate pull-down data indicated that 5hmC abundance alters the binding of the transcription factor CLOCK near the promoters of these genes (*e.g.*, *Palld*, *Gigyf1*, and *Fry*), providing a mechanistic role for 5hmC in gene regulation. Together, these data support gene by environment hypotheses for the origins of mental illness and provide a means to identify the elusive factors contributing to complex human diseases.

Keywords: Developmental brain disorders; Gene by Environment; 5hmC

Introduction

Developmental brain disorders usually present years after birth; however, the molecular pathogenesis is thought to arise earlier, either during pregnancy or just after birth. Genetic studies have revealed hundreds of rare risk genes that cumulatively account for less than 25% of neurodevelopmental disorder cases (Sebat et al. 2007; Glessner et al. 2009; Weiss et al. 2009). The search for environmental factors contributing to the origins of neurodevelopmental disorders has been motivated by their multifactorial genetic heredity and incomplete concordance in monozygotic twins (Tordjman et al. 2014). There is clear evidence that individual differences in the developmental-timing and severity of early-life stressors, including *in utero*, can predict emotional dispositions later in life and influence the risk for the development of psychopathology (O'Connor et al. 2003; Heim et al. 2008; Ceschi et al. 2014). This process of “fetal programming” may be mediated, in part, by the impact of the early-life experience on the developing hypothalamic-pituitary-adrenal (HPA) axis (Weinstock 1997; Van den Hove et al. 2006). The HPA axis responds to stressors in early embryonic development, and is highly sensitive to early life adversities (Meaney 2001).

5-hydroxymethylcytosine (5hmC) is an environmentally sensitive epigenetic mark that is highly enriched in neuronal cells (Kriaucionis and Heintz 2009; Song et al. 2011; Szulwach et al. 2011a; Szulwach et al. 2011b) and is associated with the regulation of neuronal activity, suggesting that 5hmC plays an important role in coordinating transcriptional activity in the brain, where it has been linked to developmental brain disorders (*e.g.*, fragile X syndrome, Rett syndrome, and autism) (Mellen et al. 2012; Wang et al. 2012; Yao and Jin 2014; Zhubi et al. 2014). Recently we reported that the *Cntnap2* knockout mouse model of autism exhibits a genome-wide disruption of 5hmC in a significant number of orthologous genes ($N = 57/232$) and

pathways common to human neurodevelopmental disorders (Papale et al. 2015a), supporting an interaction between *Cntnap2* and 5hmC in developmental brain disorders.

The *Cntnap2* (or *Caspr2*) gene encodes a neuronal transmembrane protein member of the neurexin superfamily involved in neuron–glia interactions and clustering of potassium channels in myelinated axons (Poliak et al. 1999; Poliak et al. 2003). Humans harboring a homozygous loss of *CNTNAP2* function exhibit severe developmental delays and intellectual disabilities, contributing to a variety of neurodevelopmental disorders such as Gilles de la Tourette syndrome, obsessive-compulsive disorder, cortical dysplasia-focal epilepsy syndrome, autism, schizophrenia, Pitt-Hopkins syndrome, and attention deficit hyperactivity disorder (Poot 2015). Mice harboring a homozygous loss of *Cntnap2* function exhibit parallels to the major neuropathological features in autism spectrum disorders, including abnormal social behavior, deficits in communication, stereotypic repetitive behaviors, cortical dysplasia, and focal epilepsy (Penagarikano et al. 2011). In addition, these mice exhibit defects in neuronal migration of cortical projection neurons and a reduction in the number of striatal GABAergic interneurons. In contrast, heterozygous *Cntnap2* mutant mice lack any of the behavioral or neuropathological abnormalities found in the homozygous mutant (Penagarikano et al. 2011).

Recent evidence in humans does not support a relationship between heterozygous *CNTNAP2* deletions and neurodevelopmental disorders (Toma et al. 2018); however, it remains possible that the combination of heterozygous *CNTNAP2* deletions in a genomic background of increased risk may lead to mental illness (Toma et al. 2018). With the onset of *Cntnap2* expression in mice beginning on embryonic day 14, its expression may be sensitive to prenatal stressors that alter the neuronal epigenome. Thus, we hypothesized that *Cntnap2* heterozygote

mutants exposed to prenatal stress would have genome-wide disruptions in 5hmC levels and behavioral deficits, similar to the homozygote mutant.

Results

*Prenatal stressed *Cntnap2* heterozygous mice show altered social & repetitive behaviors*

To test the long-lasting effects of a prenatal environmental stress on the development of altered behaviors, we subjected WT pregnant mice (carrying WT and *Cntnap2*^{+/-} pups) to seven-days of variable prenatal stress from embryonic day (E) 12 to E18, which overlaps the onset of *Cntnap2* expression (E14; Methods) (Mueller and Bale 2008; Penagarikano et al. 2011). These mild stressors were previously published and were selected because they do not include pain or directly influence maternal food intake or weight gain (Mueller and Bale 2006; Mueller and Bale 2007). Both WT and HET mice were taken from the same litter, providing an ideal internal control for our findings. Offspring were monitored for consistent maternal/offspring interactions and left undisturbed until weaning day (postnatal day 18; Supplemental Methods). At three months of age, prenatal stressed (PS) heterozygous *Cntnap2* (PS-HET) and PS wild-type (PS-WT), as well as non-PS HET (control-HET) and non-PS WT (control-WT) offspring, were subjected to behavioral testing or sacrificed for molecular analysis. All behavioral groups (PS and controls for both genotypes and sex, $N = 9-12$ per genotype and for each sex) were assessed using a variety of behavioral tests for general locomotor activity, anxiety and depression-like levels and social aptitude (Supplemental Methods). Only female PS-HET mice showed altered social and repetitive behaviors in two independent social behavioral tests, the 3-chambersocial test and a 10-minute reciprocal social interaction test, respectively. As expected for the highly social strain C57/B6J, female controls (control-WT, PS-WT, and control-HET) showed a

significant preference to be in the chamber and interacting with the cup containing an unfamiliar mouse in the 3-chamber social test. In contrast, female PS-HET mice mirror the homozygous *Cntnap2* mutants by not showing a significant preference (differential in mean time spent in each chamber was 65.3 seconds (s) for control-WT, 80.3s for PS-WT, 102.4s for control-HET, and 27.8s for PS-HET; Two-way Anova, P -value < 0.05 for controls only; **Figure 1A-B; Supplemental Fig S1**) (Penagarikano et al. 2011). Similarly, the time spent interacting with the unfamiliar mouse instead of the empty cup during the 3-chamber test was significant for the controls but not for the PS-HET female mice (differential in mean time spent interacting in each chamber was 50.2s for control-WT, 53.6s for PS-WT, 65.4s for control-HET, 16.1s for PS-HET; Two-way Anova, P -value < 0.05 for controls only; **Figure 1C-D; Supplemental Fig S1F**). The mice were then subjected to an additional social interaction test that confirmed the female PS-HET mice spent significantly more time grooming/digging than interacting with an unfamiliar mouse compared to all controls (total time spent grooming/digging among pairs of mice matched for treatment, genotype, and sex are 96.4s for control-WT, 74.4s for PS-WT, 101.2s for control-HET, and 160.8s for PS-HET; One-way Anova, P -value < 0.05 , **Figure 1E**). This repetitive grooming/digging behavior phenotype was only present in a social setting, as it was not observed in mice allowed to explore a cage alone for 10 min (total time spent grooming/digging during a 10 minute cage exploratory test are 64.5s for control-WT, 52.2s for PS-WT, 59.3s for control-HET, and 44.2s for PS-HET; Two-way Anova, P -value > 0.05 , **Table 1**). Deficits in these 3 social-related measures were reproducible in an independent cohort of female mice ($N = 8-10$ per treatment, sex, and genotype, **Supplemental Fig S1**). Male mice (control-WT, control-HET, PS-WT, and PS-HET) did not show any altered social or repetitive behavior phenotypes (**Supplemental Fig S1**). In addition, both sexes and all experimental groups showed no

significant differences in general locomotor activity and anxiety and depression-like levels (P -value > 0.05 ; **Table 1 and Supplemental Table S1**). Together, these data support altered social and repetitive behaviors in female PS-HET mice, revealing a sex-specific gene by environment (GxE) interaction effect on behavior. Since homozygous mutants show similar altered social and repetitive-like behaviors (Penagarikano et al. 2011) and a genome-wide disruptions in 5hmC (Papale et al. 2015a), we next sought to examine the 5hmC profile in this GxE interaction model.

Long-lasting disruption of 5hmC in the female GxE interaction model

To determine the effect of this GxE interaction on the genome-wide distribution of 5hmC, we employed an established chemical labeling and affinity purification method, coupled with high-throughput sequencing technology (Szulwach et al. 2011b; Papale et al. 2015a; Li et al. 2016; Papale et al. 2016). Age-matched PS-HET and control HET, as well as PS-WT and control-WT female offspring ($N = 3/\text{group}$) were sacrificed at 3 months of age as independent biological replicates. 5hmC containing DNA sequences were enriched from pools ($N = 2$) of hippocampal tissues, a brain region involved in the pathology of social behavior with high expression of the *Cntnap2* gene since embryonic day 14 and previously shown to be correlated with altered sociability in the *Cntnap2*^{-/-} mice (Penagarikano et al. 2015). High-throughput sequencing resulted in a range of ~25-43 million uniquely mapped reads from each pool of biological replicates (**Supplemental Table S2**, Supplemental Methods). These data showed no visible differences among the Chromosomes, except depletion on Chromosomes X, which is consistent with previous observations (Szulwach et al. 2011b; Papale et al. 2015a; Papale et al.

2016), and an overall equal distribution of 5hmC levels in both genotypes and in both conditions on defined genomic structures and repetitive elements. (**Supplemental Fig S2**).

Identification and characterization of differentially hydroxymethylated regions (DhMRs) in the female GxE interaction model

To identify distinct 5hmC distribution patterns following prenatal stress throughout the genome of both *Cntnap2* heterozygous and WT female mice, we characterized differentially hydroxymethylated regions (DhMRs) between: 1) PS-HET and control-HET groups, to investigate the long lasting effects of the GxE interaction on 5hmC levels (HET-DhMRs); as well as 2) PS-WT and control-WT groups, to determine the long-lasting effects of prenatal stress on 5hmC levels (WT-DhMRs). A total of 1,381 PS-HET increases in hydroxymethylation (PS-HET-specific 5hmC) and 1,231 PS-HET decreases in hydroxymethylation (control-HET-specific 5hmC) were found and these loci were distributed across all Chromosomes (**Figure 2A**; Dataset 1; Supplemental Methods). In contrast, only a total of 735 significant PS-WT increases in hydroxymethylation (PS-WT-specific 5hmC) and 710 significant PS-WT decreases in hydroxymethylation (control-WT-specific 5hmC) were found and these loci were distributed across all Chromosomes (**Supplemental Fig S3A**; Dataset 1; Supplemental Methods). These data indicate that the GxE interaction results in nearly twice the number of changes in 5hmC levels, but also suggest that prenatal stress alone is sufficient to cause stable changes in 5hmC levels. Since specific regions of the genome are differentially methylated based on the biological functions of the genes contained within each region, we next determined if GxE- and PS-induced DhMRs are enriched or depleted on certain Chromosomes using a binomial test of all detected 5hmC peaks as the background (Supplemental Methods). This analysis revealed that DhMRs

were not significantly enriched or depleted on any Chromosomes (binomial P -value < 0.05 ; **Figure 2A**; **Supplemental Fig S3A**).

We proceeded to define the genomic features associated with the GxE interaction by annotating HET-DhMRs to overlapping gene structures or to intergenic regions if more than 10 kilobases (kb) away from any gene structure. The overall distribution of these data indicated that the largest fractions of HET-DhMRs were intronic (~34%) or in distal intergenic regions (~30%), which is *not* a significant change in distribution compared to distribution of the 5hmC data, suggesting that DhMR are not targeted to specific gene structures. (**Figure 2B**). Binomial testing revealed that the HET-DhMRs were significantly depleted in exons and promoter regions (**Supplemental Fig S4A**; binomial P -value < 0.05). While the distribution of WT-DhMRs also revealed structures with significant fluctuations, the genomic location of the WT-DhMR fluctuations were unique to those found for the HET-DhMRs (**Supplemental Figs S3B and S4B**). Together, these data indicate that the GxE interaction- and stress-alone-induced DhMRs are unique and not randomly distributed throughout the genome.

Finally, DNA hypermethylation on repetitive elements is believed to play a critical role in maintaining genomic stability (Coufal et al. 2009). To investigate the genome-wide disruptions of 5hmC on repetitive elements, we aligned the total 5hmC reads and DhMRs to RepeatMasker and the segmental duplication tracks of NCBI37v1/mm9 and found that the largest fraction of HET-DhMRs (~70%) was in the SINE and simple repeat regions of the genome; however, they were only enriched in LINES (**Supplemental Fig S4C**). In contrast, the WT-DhMRs were enriched in LTR repeats and were specifically enriched in DNA repeats and SINEs (PS-WT- and Control-WT-specific, respectively; **Supplemental Fig S4D**; binomial P -value < 0.05). These

data suggest a long-term stability of 5hmC at the majority of repeat sequences in female mice following prenatal stress.

Annotation of GxE interaction DhMRs to genes reveals known and potentially novel stress and neurodevelopmental-related genes

To determine the genes and pathways affected by the long-lasting molecular effects of a GxE interaction on female mice, we annotated the HET-DhMRs to genes and found 895 and 770 genes that contained PS-HET- and control-HET-specific 5hmC levels, respectively. One-hundred and ninety-five genes were found to have more than one DhMR, leading to a total of 1,600 unique genes with disruptions in 5hmC levels (Dataset 1; **Supplemental Fig S5**). While sixty-five genes were identified to have both PS-HET- and control-HET-specific 5hmC, DhMRs residing on the same gene were at distinct genomic loci (average >50kb away from each other), suggesting they have unique molecular roles (*e.g.*, gene regulation) following a GxE interaction. A significant number of HET-DhMR-associated genes were found in well-known stress genes such as *Homer1* (Wagner et al. 2015) and *Ncam1* (Bisaz et al. 2009) (524/1600 χ^2 test P -value < 0.0001). Moreover, several of the HET-DhMR-associated genes were previously linked with developmental brain disorders including genes that encode synaptic proteins/receptors (*e.g.*, glutamate, GABAergic, calcium channels, potassium channels, sodium channels, as well as neurexin and neuroligin transmembrane proteins). To examine whether the DhMR-associated genes were enriched for neurodevelopmental-related genes, we compared them to a list of developmental brain disorder genes ($N = 232$ genes; Supplemental Methods) (Gonzalez-Mantilla et al. 2016). This comparison revealed that a significant number of orthologs ($N = 56$ of 232; hypergeometric test P -value < 0.01) harbor HET-DhMRs, suggesting that 5hmC has a molecular

role in the disruption of neurodevelopment following a GxE interaction. Indeed, this finding is similar to the DhMR-associated genes found in the *Cntnap2* homozygous mouse (Papale et al. 2015a). The HET-DhMR-associated genes also included several genes known to function in epigenetic pathways, including *Dnmt3a*, *Tet2*, and *Hdac4*. To further examine the biological significance of the HET DhMRs-associated genes, we performed separate gene ontological (GO) analyses of the PS- and control-specific HET DhMR-associated genes ($N = 895$ and 770 , respectively) and found a significant enrichment of neuronal ontological terms only among the PS-specific HET DhMR-associated genes, which included regulation of synaptic organization, axon development, and neuron projection morphogenesis (P -value < 0.05 & $FC > 1.5$ for GO enrichment tests; **Figure 2C**; Dataset 2). In contrast, the control-specific HET DhMR-associated genes were related to organ-related developmental and morphological processes (**Figure 2D**; Dataset 2). Together, increased 5hmC in the GxE model results in disruptions of genes that encode synaptic proteins/receptors and signaling pathways, supporting a link to brain-related disorders.

Genomic annotation of the WT-DhMRs to genes revealed that 510 and 450 genes contained PS-WT- and control-WT-specific 5hmC, respectively. Sixty genes were found to have more than one DhMR, leading to a total of 931 unique genes with disruptions in 5hmC levels (Dataset 1). While twenty-eight genes were identified to have both PS-WT- and control-WT-specific 5hmC, WT-DhMRs residing on the same gene were at distinct genomic loci (average >50 kb away from each other), suggesting that when found proximal to the same gene WT-DhMRs likely have unique molecular roles (*e.g.*, gene regulation) following prenatal stress. Nearly 80% of WT-DhMR-associated genes were unique from the HET-DhMR-associated genes (**Supplemental Fig S3C**), indicating that the majority of hydroxymethylation changes were

specific to genotype. Moreover, to gain insight into 5hmC changes that may predispose heterozygous mice to exhibit repetitive behaviors and altered sociability when exposed to prenatal stress, we compared control-HET and control-WT groups and found 1,430 DhMR-associated genes. Overlap of these genes with PS-HET and PS-WT DhMR-associated genes revealed a significant overlap among all three groups ($N = 50$; **Supplemental Fig S3D**; Dataset 1). The overlapping DhMR-associated genes may indicate top candidates contributing to the gene by environment behavioral outcomes.

A significant number of WT DhMR-associated genes were previously linked to stress, including *Dnmt1*, *Pten*, and *Fkbp5* (Unterberger et al. 2006; Bassi et al. 2013; Zannas et al. 2016) (312/931; hypergeometric test P -value < 0.01). Moreover, a significant number of orthologs, best known for their role in human neurodevelopmental disorders, have disruptions in 5hmC levels (35/232; hypergeometric P -value < 0.01 ; *e.g.*, *Aut2*, *Nckap5*, *Setbp1*) (Papale et al. 2015a; Gonzalez-Mantilla et al. 2016), further implicating stress-alone in the pathogenesis of neurodevelopmental disorders. Finally, WT DhMR-associated genes included *Dnmt1* from the epigenetic pathway. Annotation of the PS- and control-specific WT DhMR-associated genes ($N = 510$ and 450 genes, respectively) to GO terms found a significant overrepresentation of terms connected with metabolism/catabolism and limb development/morphogenesis (**Supplemental Figs S3E and S3F**; Dataset 2). Together, these data indicate that prenatal stress alone disrupts 5hmC on biologically relevant genes and pathways.

Confirmation of long-lasting disruptions of 5hmC in the GxE interaction model

As an initial confirmation of genome-wide disruption of 5hmC following prenatal stress, we profiled 5hmC levels in an independent cohort of mice and another brain region associated

with social and stress-related behaviors (striatum) that may contribute to pathways of risk in female autism spectrum disorders (Jack et al. 2021). Age-matched PS-HET and control-HET, as well as PS-WT and control-WT female mice ($N = 3/\text{group}$) were sacrificed at 3 months of age as independent biological replicates. 5hmC containing DNA sequences were enriched from striatal total DNA and high-throughput sequencing resulted in a range of ~20-70 million uniquely mapped reads from each biological replicate (**Supplemental Table S3**, Supplemental Methods). Similar to the hippocampal 5hmC data, sequence read density mapping showed no visible differences among the Chromosomes, except depletion on Chromosomes X, and an overall equal distribution of 5hmC levels in both genotypes and in both conditions on defined genomic structures and repetitive elements (**Supplemental Fig S6**). To compare the long-lasting molecular effects of a gene by environment interaction on female mice in the striatum and hippocampus, we annotated the striatal HET-DhMRs to genes ($N = 2,224$) and found a significant overlap of differentially hydroxymethylated genes between both tissues ($N = 395$ genes; $P\text{-value} < 0.05$), including genes well known to play a role in neurodevelopmental disorders (*e.g.*, *Nrxn1*, *Nlgn1*, and *Grip1/Ncoa2*; **Supplemental Fig S7A**). Again, similar to the hippocampal findings, the gene ontologies of the striatal DhMR-associated gene contained a significant neuronal enrichment of terms (*e.g.*, synapse organization/transmission (glutamatergic) and negative regulation of axonogenesis ($P\text{-value} < 0.05$ & $\text{FC} > 1.5$ for GO enrichment tests; Dataset 2). The DNA hydroxymethylation levels of several regions also were validated using an alternative method (**Supplemental Fig S7B**; Supplemental Methods). Together, these striatal 5hmC data provide a validation that 5hmC levels are stably disrupted following a GxE interaction in brain regions linked to stress response and social behavior.

Since the prenatal stressed *Cntnap2*^{+/-} female mice showed repetitive behaviors and altered sociability, similar to the homozygote phenotype, we compared the DhMR-associated genes between these two models (*e.g.*, PS-HET and female homozygotes) and found a significant overlap of genes in the striatum ($N = 443$; P -value < 0.05 ; **Supplemental Fig S7C**). In addition, the DhMR-associated genes from the homozygote striatum also had a significant overlap of genes when compared to the PS-HET genes from the hippocampus ($N = 346$; P -value < 0.05 ; **Supplemental Fig S7C**). Together, these data further validate the finding that 5hmC disruptions are linked to neurodevelopment-related disorders and support gene by environment hypotheses for complex disorders.

Candidate functional DhMRs in the female GxE interaction model

To gain insight into the potential molecular mechanism(s) for long-lasting GxE interaction-induced DhMRs, we used RNA sequencing (RNA-seq) to profile gene expression in the hippocampus of the same mice surveyed for 5hmC (see Supplemental Methods). Comparison of transcript levels in PS-HET and control-HET mice revealed 524 differentially expressed genes, 162 upregulated and 363 downregulated (FDR P -value < 0.1 ; **Figure 3A**; Dataset 3). Separate gene ontologies of these 162 upregulated and 363 downregulated genes found terms related to metabolic-related processes (upregulated genes; **Figure 3B**; Dataset 4) and axonal development and histone modifications (downregulated genes; **Figure 3C**; Dataset 4; Supplemental Methods). DhMRs associated with differentially expressed genes represent candidate functional DhMRs that may have a direct role in gene regulation. An overlay of the HET-DhMR data with these HET-RNA-seq data revealed that 72 differentially expressed genes harbor a HET-DhMR, which included neurodevelopmental disorder genes (*e.g.*, *Reln*, *Dst*, and

Trio) (Gonzalez-Mantilla et al. 2016), known stress-related genes (*e.g.*, *Ncoa2*) (Barko et al. 2019) and epigenetic genes (*Dnmt3a* and *Hdac4*; **Figure 3D**; Dataset 3). The Pearson's correlation between DNA methylation and gene expression across all genes was similar to the correlations found in previous studies ($r = 0.267$) (Zeng et al. 2012; Alisch et al. 2014). While general relationships between 5hmC abundance and gene expression levels were not observed in these genes, there was a significant enrichment of DhMRs located in the introns of differentially expressed genes (P -value < 0.05). These results indicate that the GxE interaction induces long-lasting effects in the regulation of genes functioning in neuronal development and synaptic plasticity and supports a role for 5hmC in this dysregulation that is associated with altered social behaviors in adult GxE mice.

Comparison of transcript levels in PS-WT and control-WT mice revealed only 5 genes that were altered by prenatal stress, all downregulated (FDR P -value < 0.1 ; **Supplemental Fig S8**; Dataset 3). An overlay of the WT-DhMR data with the WT-RNA-seq data did not reveal any potentially functional DhMRs. These results suggest that WT-DhMRs represent stable molecular alterations that may have affected gene expression at earlier developmental time-points, suggesting that the PS-WT mice may have earlier phenotypes related to the prenatal stress.

The GxE interaction alters transcription factor binding in DhMRs

Syndromic forms of neurodevelopmental disorders can involve the disruption of transcription factor function (Gharani et al. 2004; Bienvenu and Chelly 2006); thus, we investigated the potential for the HET-DhMRs to alter DNA-binding of transcription factors by testing for enrichments of known transcription factor sequence motifs among the DhMRs (Supplemental Methods). Several transcription factor sequence motifs were significantly

enriched in the hippocampal HET-DhMRs. Many of the transcription factors that bind these motifs have links to neurodevelopmental behaviors and disorders, such as CLOCK, NPAS2, HIF-1A, ARNT (HIF-1B), and USF1 (**Table 2**; Dataset 5) (Schmidt-Kastner et al. 2006; Nicholas et al. 2007; Shi et al. 2016; Hamilton et al. 2018). Together, these findings suggest that 5hmC influences transcription factor binding, which may explain the observed correlated disruptions in gene expression. An initial test of this hypothesis examined transcription factor sequence motif enrichments among only the potentially functional DhMRs (*i.e.*, genes with correlated disruptions in 5hmC and expression; $N = 72$; Supplemental Methods). Five transcription factors sequence motifs were significantly enriched among the potentially functional HET-DhMRs (BHLHE40, MYC, CLOCK, ARNT, and USF1). The CLOCK transcription factor binds E-box enhancer elements, its consensus binding sequence contains a CpG dinucleotide (CACGTG) that can be hydroxymethylated, and its expression/dysregulation is associated with autism (Hu et al. 2013; Schuch et al. 2018). RNA sequence data and Western blot analysis revealed that *Clock* is not differentially expressed at the transcript and protein levels due genotype and/or condition (*i.e.*, PS-HET, control-HET, PS-WT, and control-WT; (**Figure 4A**; Dataset 3). To determine the effect of this GxE interaction on the genome-wide distribution of CLOCK binding, we employed chromatin affinity purification coupled with high-throughput sequencing technology (ChIP-seq). Age-matched PS-HET and control HET, as well as PS-WT and control-WT female offspring ($N = 3/\text{group}$) were sacrificed at 3 months of age as independent biological replicates. DNA sequences containing CLOCK binding sites were enriched from the hippocampus total DNA and high-throughput sequencing resulted in ~10 million uniquely mapped, non-duplicated, reads from each biological replicate (Methods; Dataset 6). A differential analysis of these data filtered to only DhMRs containing a putative CLOCK

binding site ($N = 577$; CACGTG) identified 19 DhMRs with differential CLOCK binding (**Figure 4B**; Dataset 6). Three of these DhMRs were associated with genes that also were differentially expressed as a result of the GxE interaction (*Gigyf1* ~1.8x and both *Palld* and *Fry* ~1.5x; Dataset 3). The DhMR containing CLOCK binding sequences reside near the promoters of each of these genes (**Figure 4C**) and all three genes have links to autism (Garbett et al. 2008; Voineagu and Eapen 2013; Paulraj et al. 2019; Satterstrom et al. 2020). Together, these data suggest that differential hydroxymethylation levels may mediate CLOCK binding, supporting a mechanism for 5hmC in the regulation of gene expression of genes associated with the observed phenotype.

5hmC alters CLOCK binding

To formally test whether 5hmC can alter CLOCK binding, hippocampal nuclear lysates were incubated with oligonucleotide probes containing sequences from the differentially bound CLOCK sites in the DhMRs associated with *Gigyf1*, *Palld*, and *Fry* genes. The CpG dinucleotide in the E-box motif of CLOCK located in each probe contained either cytosine, 5-methylcytosine, or 5-hydroxymethylcytosine (Methods). Electromobility shift assays using unmodified cytosine probes from each of these genes revealed successful binding of CLOCK from hippocampal nuclear lysates (**Figure 4D-G**; **Supplemental Fig S9**). These DNA::protein complexes were abrogated in cold competition assays and super-shifted with the addition of antibodies against CLOCK (**Figure 4D**, left panel; **Supplemental Fig S9**). In contrast, hydroxymethylated probes significantly reduced the binding of CLOCK (**Figure 4D-G**; **Supplemental Fig S9**), indicating that 5hmC prevents CLOCK from binding to the E-box motifs located near the promoters of the differentially expressed genes *Gigyf1*, *Palld*, and *Fry*. Positive controls for both 5mC (*Kcnk4*)

and 5hmC (*Camk2a*) confirm that these DNA modifications exhibit unique binding efficiencies for specific transcription factors (**Supplemental Fig S10**). Together, these findings suggest a working model for 5hmC regulation of developmentally important genes that when altered can result in social-related disorders. In the example presented here, the GxE interaction results in a loss of 5hmC near the promoters of *Gigyf1*, *Palld*, and *Fry* that subsequently improves the binding of CLOCK to act as a repressor of *Gigyf1*, *Palld*, and *Fry* expression (**Figure 5**). Thus, *Gigyf1*, *Palld*, and *Fry* are top candidate genes contributing to the adult sex-specific phenotype found in female *Cntnap2* heterozygous mice exposed to prenatal stress.

Discussion

Here we provide evidence of a functional role for 5hmC in a gene (partial loss of *Cntnap2*) by environment (prenatal stress; GxE) interaction model that results in sex-specific alterations of neurodevelopmental behaviors. These 5hmC disruptions were found in the orthologs of a significant number of genes implicated in human developmental brain disorders: significantly overlapping with those found in both male and female *Cntnap2* homozygous mouse mutants (Papale et al. 2015b). Integration of epigenetic and gene expression data from the same mice revealed relations with genes (*e.g.*, *Nrxn1*, *Reln*, *Grip1*, *Dlg2* and *Sdk1*) known to contribute to the behavioral deficits of neurodevelopmental disorders. Finally, we focused on an enrichment of transcription factors sequence motifs among the potentially functional PS-HET DhMRs and found that the loss of 5hmC in these regions resulted in the loss of CLOCK binding. Together these data demonstrate a molecular mechanism for 5hmC in the regulation of developmentally important genes that when altered can result in social-related disorders. A GxE model and 5hmC disruptions may represent a common etiology for developmental brain disorders in humans.

Having data from both striatal and hippocampal tissues provide a unique opportunity to examine the combined and differential contributions of these tissues to neurodevelopment-related behaviors, which both supported that the epigenome plays a role in the long lasting effects of prenatal environmental exposures on brain and behavior. For example, both tissues carried GxE DhMRs in orthologous genes common to human neurodevelopmental disorders, such as nuclear receptor genes, glucocorticoid receptor-interaction protein 1 (*Grip1* also called *Ncoa2*) (Barko et al. 2019), as well as several genes associated with neurotransmission and synaptic plasticity (*Nrxn1*, *Nlgn1*, *Dlg2*, and *Negr1*) (Gonzalez-Mantilla et al. 2016; Szczurkowska et al. 2018). The hippocampus-specific GxE DhMRs also included genes related to human neurodevelopmental disorders, including *Homer1*, which forms high-order complexes with *Shank1* that are necessary for the structural and functional integrity of dendritic spines (Orlowski et al. 2012). In addition, the hippocampus-specific GxE DhMRs contained links to circadian function (e.g., *Clock*, *Homer1*, and *Shank2*); the prenatal stress administered here is sleep disruptive, supporting a hypothesis that circadian rhythms contribute to the pathogenesis of neurodevelopmental disorders (Geoffroy et al. 2016). Finally, investigation of the hippocampus-specific GxE DhMRs with correlated disruptions in gene expression revealed links to the observed adult behavioral deficits, including specific links to the limbic system and synaptic plasticity (e.g., *Dst*, *Reln*, *Lpp*, *Trio*, *Notch1*, *Col4a1*, and *Utrn*) (Sibbe et al. 2009; Bhandare et al. 2010; Wong et al. 2014; Wang et al. 2015). Together, these data suggest insights into the molecular pathogenesis of GxE interactions that result in neurobehavioral alterations. While prenatal stress alone disruptions in 5hmC (WT-DhMRs) were largely unique from the GxE DhMRs (HET-DhMRs), WT-DhMR-associated genes also included previous links to neurodevelopmental disorders (e.g., *Auts2*, *Cacna1c* and *Setd5*) (Gonzalez-Mantilla et al. 2016).

Although 5hmC disruption on these genes was not correlated with altered transcript levels and were not sufficient to cause a behavioral phenotype, many of them have been implicated in disorders with social deficits (Gonzalez-Mantilla et al. 2016), suggesting that prenatal stressed WT mice have uncharacterized deficits and may be susceptible to additional environmental stressors that trigger later life brain-related disorders.

Mechanistically, it is intriguing that the binding motifs of several transcription factors with known roles in neuronal development were found in the GxE DhMRs. These findings included binding motifs for transcription factors best known for their roles in circadian rhythm, immune response, learning and memory, and hypoxia. For example, CLOCK, ARNTL, BHLHE40, and BHLHE41 all regulate the circadian clock; in addition, they are also linked to neurodevelopmental disorders, such as major depressive disorders, bipolar disorders, and autism spectrum disorders (Nicholas et al. 2007; O'Roak et al. 2012; Shi et al. 2016; Hamilton et al. 2018); again, consistent with the prenatal stress paradigm used here being sleep disruptive. CLOCK functions as a pioneer-like transcription factor (Menet et al. 2014), while also displaying histone acetyltransferase activity (Doi et al. 2006), corroborating CLOCK as a regulator of chromatin accessibility and contributor to the epigenetic framework in cells. While CLOCK generally acts as a transcriptional activator, it also exhibits repressive activity that is dependent on the assembly of larger protein complexes (Michael et al. 2017; Yang et al. 2020; Cao et al. 2021). Paired box protein 5 (*Pax5*) and aryl hydrocarbon receptor (*Ahr*) have defined roles in immune response and also are connected to autism spectrum disorders and major depressive disorders (Wong et al. 2008; O'Roak et al. 2012; Neavin et al. 2018). The stress paradigm used in this study increases placental inflammation (Bronson and Bale 2014); indeed, immune activation history in the mother is associated to increased symptom severity in children

with mental illness (Patel et al. 2018). It is notable that the GxE interaction resulted in distinct differences between PS-HET and PS-WT enriched transcription factor sequence motifs, perhaps suggesting an interaction between CNTNAP2 and 5hmC in transcription factor binding and function.

Focusing more specifically on significant enrichments of transcription factor sequence motifs in potentially functional DhMRs, especially enriched motifs located in the promoter regions, revealed known and novel genes contributing to the adult phenotype. *Gigyf1*, *Palld*, and *Fry* were the examples reported here with differential CLOCK binding and links to autism (Garbett et al. 2008; Voineagu and Eapen 2013; Paulraj et al. 2019; Satterstrom et al. 2020). In addition, catenin (cadherin associated protein), alpha 2 (*CTNNA2*) is an example of a gene that is differentially expressed and contains a DhMR located in its promoter harboring a different enriched transcription factor sequence motif (USF1). Rodent *Ctnna2* encodes a protein that regulates morphological plasticity of synapses and its loss in neurons leads to defects in neurite stability and migration, and cortical dysplasia (Schaffer et al. 2018). Indeed, cortical dysplasia is a feature in humans and mice lacking expression of CNTNAP2, and mice exhibit defects in neuronal migration of cortical projection neurons and a reduction in the number of striatal GABAergic interneurons (Penagarikano et al. 2011; Poot 2015). It will be important to characterize the role of 5hmC on the binding of other transcription factors (*e.g.*, USF1) in promoters and enhancers of genes containing potentially functional DhMRs. While only the 5hmC probes containing the transcription factor binding sites had statistically significant different binding efficiencies compared to the 5C probes, it will be important for future studies to characterize the potential role of 5mC in altered transcription factor binding associated with gene by environment interactions resulting in neurodevelopmental disorders. Despite that enriched

motifs in DhMRs located in promoter regions of differentially expressed genes represent top candidate genes directly contributing to the adult phenotype, enriched motifs in DhMR-associated genes that are not differentially expressed in the adult may identify genes contributing to the origins of the adult behavioral deficits. Further studies at earlier developmental time-points are needed to determine the role of these DhMRs in the development of the observed adult behavioral deficits.

Exposure to stress during different gestational periods has unique effects on epigenetic programming of the developing embryo. During early gestation, prenatal stress-induced epigenetic changes in neuronal precursor cells are predominantly mediated by hormonal mechanisms, which in part can be transmitted via placental pathways (Bock et al. 2015). Stress during this time period predominately affects behavioral development in males (Li et al. 2008; Mueller and Bale 2008). However, in late gestational stages, the mechanism underlying prenatal stress-induced epigenetic changes become more complex involving the activation of the excitatory and inhibitory endocrine modulatory systems, which may interfere with stress-induced synaptic reorganization (Bock et al. 2015). Stress during this time period predominately affects behavioral development in females (Li et al. 2008; Mueller and Bale 2008); thus, there is a sex-specific effect on behavior depending on the gestational timing of the stress exposure. Consistent with these reports, late gestational stress in this present study results in long-lasting behavioral and molecular alterations in females on genes known to be associated with synaptic plasticity and mental disorders. The repetitive behaviors and social deficits render the PS-HET mice an attractive model for other psychiatric disorders in humans. Since both insults (*i.e.*, genetic and environment) are required for the expression of an altered behavioral phenotype, this model (*i.e.*, a two-hit model) directs studies to consider environmentally sensitive molecular mechanisms

contributing to stress-induced behavioral outcomes. Moreover, it supports the study of other “multi-hit” hypotheses in the origins of mental illness to identify elusive factors that could contribute to the complexity of psychiatric disorders in humans. While there is precedence for multi-hit models disrupting brain development (van den Hove et al. 2011; Schraut et al. 2014; Bell et al. 2016; Schaafsma et al. 2017), it will be important to examine the generalizability of these findings using other developmentally important genes in the brain that also may contribute to the vulnerability/resilience toward mental illness.

CNTNAP2 has been implicated in numerous neurodevelopmental disorders and heterozygous disruptions have been found in humans with intellectual disability (ID), seizures, and signs of autism spectrum disorders (*e.g.*, repetitive behaviors, and social deficits). However, heterozygous losses of *CNTNAP2* also have been inherited from healthy parents, suggesting that heterozygous disruptions of *CNTNAP2* by themselves are not sufficient to elicit cellular or organismal phenotypes. Aside from striatum and hippocampus, high levels of *CNTNAP2* expression have been found in the olfactory bulb, ventricular zone, and thalamus (Abrahams et al. 2007; Penagarikano et al. 2011), warranting future examinations of 5hmC in other brain-related behaviors that are linked to these additional brain regions. *CNTNAP2* also has an organizing function to assemble neurons into neural circuits (Anderson et al. 2012), suggesting future studies should examine the electrophysiology of the excitatory and inhibitory synaptic transmissions in the PS-HET mice. Finally, prenatal stress may have revealed a molecular connection between 5hmC and *Cntnap2* involving the binding of a transcription-enhancing transcription factor *TCF4*, which can transactivate the *CNTNAP2* and *NRXN1* promoters. Heterozygous deletions of *TCF4* causes Pitt-Hopkins syndrome: a syndrome that also is linked to the loss of either *CNTNAP2* or *NRXN1* function (Amiel et al. 2007; Brockschmidt et al. 2007;

Zweier et al. 2007). Thus, *TCF4* may modulate the expression of *CNTNAP2* and *NRXN1* in the regulatory network involved in Pitt-Hopkins syndrome (Forrest et al. 2012). We find several DhMRs associated with *Tcf4* and *Nrxn1* following GxE interaction, suggesting that the regulatory network involved in Pitt-Hopkins syndrome may be similarly disrupted in the PS-HET mice. Together, these findings potentially provide novel insight into the molecular connection between 5hmC, *Cntnap2*, *Tcf4*, and *Nrxn1* in neurodevelopmental disorders, such as Pitt-Hopkins syndrome.

Molecular factors influencing vulnerability and resilience to environmental stress clearly involve environmentally sensitive epigenetic mechanisms such as DNA methylation. Thus, perturbations of gene by environment interactions may underlie common pathways for many mental illness-related disorders in humans. The identification and characterization of a potentially modifiable substrate (*e.g.*, 5hmC) contributing to gene by environment interaction-induced sex-specific neurologic behaviors is significant and comes at a time when there is great interest to harness the diagnostic and therapeutic power of these substrates toward healthy outcomes.

Methods

Prenatal stress paradigm

All experiments were approved by the University of Wisconsin – Madison Institutional Animal Care and Use Committee (M02529). To minimize the stress of animal handling, all of the following were conducted by a single researcher: animal colony maintenance; breeding, prenatal stress; and behavioral tests. Pregnant mice assigned to the variable stress group experienced a different daily stressor on each of the seven days during late pregnancy (E12 to

E18; the timing of which was chosen because it overlaps with the onset of *Cntnap2* expression (E14)). These variable stressors included: 36 hours of constant light, 15 min of fox odor (Cat# W332518) beginning 2 hours after lights on, overnight exposure to novel objects (8 marbles), 5 minutes of restraint stress (beginning 2 hours after lights on), overnight novel noise (white noise, nature sound-sleep machine®, Brookstone), 12 cage changes during light period, and overnight saturated bedding (700 mL, 23°C water) (Mueller and Bale 2006; Mueller and Bale 2007). These mild stressors were previously published and were selected because they do not include pain or directly influence maternal food intake or weight gain. Both WT and HET sibling pairs from multiple litters were used for both behavior and molecular studies, providing an ideal internal control for our findings and negates the need for cross-fostering of the prenatally stressed pups to non-stressed moms to find effects of prenatal exposure to stress.

Statistics for behavioral tests

Data is reported as mean \pm S.E.M. Homogeneity of variance was assessed using the Brown-Forsythe and the normal distribution of the data was tested by the Shapiro-Wilk test. Two-way ANOVA also was used to detect differences between genotype (WT x HET) and treatment (Control x PS) for time spent in the inner circle and number of entries into that circle (Open field test); time spent in closed arm, time spent in open arm, and time spent in center (Elevated plus maze); time spent in light side (Light/Dark box test) and time spent floating (Forced swim test). Two-way ANOVA was used to detect differences between treatment (Control x PS) for time spent in the chamber (3-chamber social test) and interacting with the novel mouse and time spent in the chamber with the empty cup. One-way ANOVA was used to

detect differences between groups for time spent engaged in repetitive behaviors (grooming and digging). Post hoc analysis was performed using a Bonferroni correction. Statistics were performed using SigmaPlot version 14.

5hmC enrichment of genomic DNA, library preparation, and high-throughput sequencing

Chemical labeling-based 5hmC enrichment was described previously (Song et al. 2011; Papale et al. 2015a). 5hmC-enriched libraries were generated using the NEBNext ChIP-seq Library Prep Reagent Set for Illumina sequencing, according to the manufacturer's protocol. 50-cycle single-end (striatum) or paired-end (hippocampus) sequencing was performed by Beckman Coulter Genomics or the University of Wisconsin Biotechnology Center, respectively. Paired-end sequencing for hippocampal tissue was chosen for reasons unrelated to this project. Image processing and sequence extraction were done using the standard Illumina Pipeline.

Identification of differentially hydroxymethylated regions (DhMRs)

We mapped the sequence reads to mouse NCBI37v1/mm9 reference genome using Bowtie 0.12.7 (Langmead et al. 2009), called peaks using the MACS2 algorithm v2.1.2 (Zhang et al. 2008). Newer versions of the genome always add more sequence data, but since we achieved >90% alignment of the reads to mm9, realigning to mm10 or beyond would only nominally increase the alignment efficiency, not altering the main findings of these analyses. Summits from both striatal and hippocampal data were extracted for each peak for each sample and extended +/-500bp for downstream analysis. For each genotype, the stress and control groups were pooled and merged to form the candidate differential regions within each genotype. The Bioconductor package edgeR was used to test each candidate region (P -value<0.05) (Robinson et al. 2010).

Analysis of RNA-sequence data

Read alignment and calculation of transcript expression levels: we used the mm9 assembly as our reference genome and the GENCODE M1 (NCBIM37) as our gene annotation library (the Y Chromosome was deleted for the alignment of female samples). We ran RSEM to calculate the expression at both gene-level and isoform-level (Li and Dewey 2011). EBSeq was used to detect the DE genes and transcripts in each of the pairwise comparisons of PS-WT vs control WT, and PS-HET vs control HET, with an FDR < 0.1.

ChIP-sequencing and analysis

High-throughput sequence libraries were generated from immunoprecipitated and IgG eluted DNA using the NEBNext ChIP-seq library prep reagent set for Illumina sequencing, according to the manufacturer's instructions. 50-cycle paired-end sequencing was performed by the University of Wisconsin – Madison Biotechnology Center. Image processing and sequence extraction were done using the standard Illumina Pipeline. Raw paired-end sequencing files were assessed for quality using FastQC (<https://www.bioinformatics.babraham.ac.uk/>), trimmed, using the Trim Galore! package in R (<https://github.com/FelixKrueger/TrimGalore>), and aligned to the *mm9* genome using Bowtie 2 v2.3.4 (Langmead et al. 2009), local alignment. Following alignment, SAMtools to filter multi-mapping reads (Li et al. 2009) and PCR duplicates. Uniquely-mapped and filtered sequence reads were mapped to DhMRs found to contain a putative CLOCK binding site (CACGTG) by motif enrichment analysis ($N = 577$). Only data from these sites were used for differential binding analyses (R package *edgeR* (Robinson et al. 2010)), after controlling for background through normalization to sequenced IgG reads. A *P*-

value threshold of 0.05 was used to determine significant differential binding of CLOCK between groups.

Electrophoretic mobility shift assays

Complementary 5'-biotinylated (0.5 μ M) and unlabeled oligonucleotides (Integrated DNA Technologies; Supplemental Material) were annealed by heating equal concentrations of sense and anti-sense oligonucleotides to 95°C and lowering the temperature by 2°C in 3-minute intervals until reaching 23°C. Mouse hippocampal nuclear extracts were isolated from adult female mice (P90) using a nuclear and cytoplasmic extraction kit per the manufacturer's instructions (Thermo Fisher Scientific 78833). Shift assays were performed using the LightShift® Chemiluminescent EMSA Kit (Thermo Fisher Scientific 20148). Reaction products were separated by electrophoresis and protein::DNA complexes were transferred onto a positively-charged nylon membrane (Thermo Fisher Scientific 77016). Protein::DNA complexes were detected using the Nucleic Acid Detection Module Kit (Thermo Fisher Scientific 89880) per the manufacturer's instructions, and imaged using an Odyssey® Fc imaging system (Li-Cor). A two-sided *t*-test (R environment) was used to determine significant alterations in binding affinity to methylated- and hydroxymethylated-oligonucleotides, using three independent EMSA replicates.

Data Access

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE87032.

Competing interest statement

The coauthors of this manuscript do not have any conflicts of interest regarding the contents of this manuscript. The authors declare no competing financial interests.

Acknowledgements

The authors would like to thank Dr. Sergio Tufik for his endless support, the UW biotechnology center, and Dr. Sisi Li for her assistance with brain dissections. This work was supported in part by the University of Wisconsin-Madison departments of Neurosurgery and Psychiatry, University of Wisconsin Vilas Cycle Professorships #133AAA2989 and University of Wisconsin Graduate School #MSN184352 (all to RSA), NARSAD Young Investigator Grant from the Brain & Behavioral Research Foundation #22669 & 25212 (LP), Ruth L. Kirschstein National Research Service Award MH113351-02 (AM), NIH grants HG003747 and U54AI117924 (all to SK), HG007019 (SK and QZ) and the University of Nebraska Lincoln department of Statistics (QZ). Part of QZ's research was performed when he was a postdoc in the statistical genomics group at UW Madison led by SK. *Authors' contributions:*

Conceptualization: LAP, AM, RSA; Methodology: LAP, AM, QZ, KC, LS, SK, RSA;

Validation: LAP, AM, LS, RSA; Software: AM, QZ, KC, SK, RSA; Formal analysis: LAP, AM, QZ, KC, RSA; Writing - original draft: LAP, AM, RSA; Writing - review/editing: LAP, AM, QZ, KC, LS, SK, RSA.

References

- Abrahams BS, Tentler D, Perederiy JV, Oldham MC, Coppola G, Geschwind DH. 2007. Genome-wide analyses of human perisylvian cerebral cortical patterning. *Proc Natl Acad Sci U S A* **104**: 17849-17854.
- Alisch RS, Chopra P, Fox AS, Chen K, White AT, Roseboom PH, Keles S, Kalin NH. 2014. Differentially methylated plasticity genes in the amygdala of young primates are linked to anxious temperament, an at risk phenotype for anxiety and depressive disorders. *J Neurosci* **34**: 15548-15556.
- Amiel J, Rio M, de Pontual L, Redon R, Malan V, Boddaert N, Plouin P, Carter NP, Lyonnet S, Munnich A et al. 2007. Mutations in TCF4, encoding a class I basic helix-loop-helix transcription factor, are responsible for Pitt-Hopkins syndrome, a severe epileptic encephalopathy associated with autonomic dysfunction. *Am J Hum Genet* **80**: 988-993.
- Anderson GR, Galfin T, Xu W, Aoto J, Malenka RC, Sudhof TC. 2012. Candidate autism gene screen identifies critical role for cell-adhesion molecule CASPR2 in dendritic arborization and spine development. *Proc Natl Acad Sci U S A* **109**: 18120-18125.
- Barko K, Paden W, Cahill KM, Seney ML, Logan RW. 2019. Sex-Specific Effects of Stress on Mood-Related Gene Expression. *Mol Neuropsychiatry* **5**: 162-175.
- Bassi C, Ho J, Srikumar T, Dowling RJ, Gorrini C, Miller SJ, Mak TW, Neel BG, Raught B, Stambolic V. 2013. Nuclear PTEN controls DNA repair and sensitivity to genotoxic stress. *Science* **341**: 395-399.
- Bell MR, Hart BG, Gore AC. 2016. Two-hit exposure to polychlorinated biphenyls at gestational and juvenile life stages: 2. Sex-specific neuromolecular effects in the brain. *Mol Cell Endocrinol* **420**: 125-137.
- Bhandare R, Damera G, Banerjee A, Flammer JR, Keslacy S, Rogatsky I, Panettieri RA, Amrani Y, Tliba O. 2010. Glucocorticoid receptor interacting protein-1 restores glucocorticoid responsiveness in steroid-resistant airway structural cells. *Am J Respir Cell Mol Biol* **42**: 9-15.
- Bienvenu T, Chelly J. 2006. Molecular genetics of Rett syndrome: when DNA methylation goes unrecognized. *Nat Rev Genet* **7**: 415-426.
- Bisaz R, Conboy L, Sandi C. 2009. Learning under stress: a role for the neural cell adhesion molecule NCAM. *Neurobiology of learning and memory* **91**: 333-342.
- Bock J, Wainstock T, Braun K, Segal M. 2015. Stress In Utero: Prenatal Programming of Brain Plasticity and Cognition. *Biol Psychiatry* **78**: 315-326.
- Brockschmidt A, Todt U, Ryu S, Hoischen A, Landwehr C, Birnbaum S, Frenck W, Radlwimmer B, Lichter P, Engels H et al. 2007. Severe mental retardation with breathing abnormalities (Pitt-Hopkins syndrome) is caused by haploinsufficiency of the neuronal bHLH transcription factor TCF4. *Hum Mol Genet* **16**: 1488-1494.
- Bronson SL, Bale TL. 2014. Prenatal stress-induced increases in placental inflammation and offspring hyperactivity are male-specific and ameliorated by maternal antiinflammatory treatment. *Endocrinology* **155**: 2635-2646.
- Cao X, Yang Y, Selby CP, Liu Z, Sancar A. 2021. Molecular mechanism of the repressive phase of the mammalian circadian clock. *Proceedings of the National Academy of Sciences of the United States of America* **118**.

- Ceschi G, Billieux J, Hearn M, Furst G, Van der Linden M. 2014. Trauma exposure interacts with impulsivity in predicting emotion regulation and depressive mood. *European journal of psychotraumatology* **5**.
- Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, Lovci MT, Morell M, O'Shea KS, Moran JV, Gage FH. 2009. L1 retrotransposition in human neural progenitor cells. *Nature* **460**: 1127-1131.
- Doi M, Hirayama J, Sassone-Corsi P. 2006. Circadian regulator CLOCK is a histone acetyltransferase. *Cell* **125**: 497-508.
- Forrest M, Chapman RM, Doyle AM, Tinsley CL, Waite A, Blake DJ. 2012. Functional analysis of TCF4 missense mutations that cause Pitt-Hopkins syndrome. *Hum Mutat* **33**: 1676-1686.
- Garbett K, Ebert PJ, Mitchell A, Lintas C, Manzi B, Mirnics K, Persico AM. 2008. Immune transcriptome alterations in the temporal cortex of subjects with autism. *Neurobiology of disease* **30**: 303-311.
- Geoffroy MM, Nicolas A, Speranza M, Georgieff N. 2016. Are circadian rhythms new pathways to understand Autism Spectrum Disorder? *J Physiol Paris* **110**: 434-438.
- Gharani N, Benayed R, Mancuso V, Brzustowicz LM, Millonig JH. 2004. Association of the homeobox transcription factor, ENGRAILED 2, 3, with autism spectrum disorder. *Mol Psychiatry* **9**: 474-484.
- Glessner JT, Wang K, Cai G, Korvatska O, Kim CE, Wood S, Zhang H, Estes A, Brune CW, Bradfield JP et al. 2009. Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* **459**: 569-573.
- Gonzalez-Mantilla AJ, Moreno-De-Luca A, Ledbetter DH, Martin CL. 2016. A Cross-Disorder Method to Identify Novel Candidate Genes for Developmental Brain Disorders. *JAMA psychiatry* **73**: 275-283.
- Hamilton KA, Wang Y, Raefsky SM, Berkowitz S, Spangler R, Suire CN, Camandola S, Lipsky RH, Mattson MP. 2018. Mice lacking the transcriptional regulator Bhlhe40 have enhanced neuronal excitability and impaired synaptic plasticity in the hippocampus. *PLoS One* **13**: e0196223.
- Heim C, Newport DJ, Mletzko T, Miller AH, Nemeroff CB. 2008. The link between childhood trauma and depression: insights from HPA axis studies in humans. *Psychoneuroendocrinology* **33**: 693-710.
- Hu S, Wan J, Su Y, Song Q, Zeng Y, Nguyen HN, Shin J, Cox E, Rho HS, Woodard C et al. 2013. DNA methylation presents distinct binding sites for human transcription factors. *eLife* **2**: e00726.
- Jack A, Sullivan CAW, Aylward E, Bookheimer SY, Dapretto M, Gaab N, Van Horn JD, Eilbott J, Jacokes Z, Torgerson CM et al. 2021. A neurogenetic analysis of female autism. *Brain : a journal of neurology* doi:10.1093/brain/awab064.
- Kriaucionis S, Heintz N. 2009. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324**: 929-930.
- Langmead B, Trapnell C, Pop M, Salzberg SL. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome biology* **10**: R25.
- Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**: 323.

- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)* **25**: 2078-2079.
- Li H, Li X, Jia N, Cai Q, Bai Z, Chen R, Song T, Zhu Z, Liu J. 2008. NF-kappaB regulates prenatal stress-induced cognitive impairment in offspring rats. *Behav Neurosci* **122**: 331-339.
- Li S, Papale LA, Zhang Q, Madrid A, Chen L, Chopra P, Keles S, Jin P, Alisch RS. 2016. Genome-wide alterations in hippocampal 5-hydroxymethylcytosine links plasticity genes to acute stress. *Neurobiology of disease* **86**: 99-108.
- Meaney MJ. 2001. Maternal care, gene expression, and the transmission of individual differences in stress reactivity across generations. *Annu Rev Neurosci* **24**: 1161-1192.
- Mellen M, Ayata P, Dewell S, Kriaucionis S, Heintz N. 2012. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell* **151**: 1417-1430.
- Menet JS, Pescatore S, Rosbash M. 2014. CLOCK:BMAL1 is a pioneer-like transcription factor. *Genes & development* **28**: 8-13.
- Michael AK, Fribourgh JL, Chelliah Y, Sandate CR, Hura GL, Schneidman-Duhovny D, Tripathi SM, Takahashi JS, Partch CL. 2017. Formation of a repressive complex in the mammalian circadian clock is mediated by the secondary pocket of CRY1. *Proceedings of the National Academy of Sciences of the United States of America* **114**: 1560-1565.
- Mueller BR, Bale TL. 2006. Impact of prenatal stress on long term body weight is dependent on timing and maternal sensitivity. *Physiology & behavior* **88**: 605-614.
- Mueller BR, Bale TL. 2007. Early prenatal stress impact on coping strategies and learning performance is sex dependent. *Physiol Behav* **91**: 55-65.
- Mueller BR, Bale TL. 2008. Sex-specific programming of offspring emotionality after stress early in pregnancy. *J Neurosci* **28**: 9055-9065.
- Neavin DR, Liu D, Ray B, Weinshilboum RM. 2018. The Role of the Aryl Hydrocarbon Receptor (AHR) in Immune and Inflammatory Diseases. *Int J Mol Sci* **19**.
- Nicholas B, Rudrasingham V, Nash S, Kirov G, Owen MJ, Wimpory DC. 2007. Association of Per1 and Npas2 with autistic disorder: support for the clock genes/social timing hypothesis. *Mol Psychiatry* **12**: 581-592.
- O'Connor TG, Heron J, Golding J, Glover V. 2003. Maternal antenatal anxiety and behavioural/emotional problems in children: a test of a programming hypothesis. *J Child Psychol Psychiatry* **44**: 1025-1036.
- O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, Levy R, Ko A, Lee C, Smith JD et al. 2012. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* **485**: 246-250.
- Orlowski D, Elfving B, Muller HK, Wegener G, Bjarkam CR. 2012. Wistar rats subjected to chronic restraint stress display increased hippocampal spine density paralleled by increased expression levels of synaptic scaffolding proteins. *Stress* **15**: 514-523.
- Papale LA, Li S, Madrid A, Zhang Q, Chen L, Chopra P, Jin P, Keles S, Alisch RS. 2016. Sex-specific hippocampal 5-hydroxymethylcytosine is disrupted in response to acute stress. *Neurobiol Dis* **96**: 54-66.
- Papale LA, Zhang Q, Li S, Chen K, Keles S, Alisch RS. 2015a. Genome-wide disruption of 5-hydroxymethylcytosine in a mouse model of autism. *Hum Mol Genet* **24**: 7121-7131.

- Papale LA, Zhang Q, Li S, Chen K, Keles S, Alisch RS. 2015b. Genome-wide disruption of 5-hydroxymethylcytosine in a mouse model of autism. *Hum Mol Genet* doi:10.1093/hmg/ddv411.
- Patel S, Masi A, Dale RC, Whitehouse AJO, Pokorski I, Alvares GA, Hickie IB, Breen E, Guastella AJ. 2018. Social impairments in autism spectrum disorder are related to maternal immune history profile. *Mol Psychiatry* **23**: 1794-1797.
- Paulraj P, Bosworth M, Longhurst M, Hornbuckle C, Gotway G, Lamb AN, Andersen EF. 2019. A Novel Homozygous Deletion within the FRY Gene Associated with Nonsyndromic Developmental Delay. *Cytogenet Genome Res* **159**: 19-25.
- Penagarikano O, Abrahams BS, Herman EI, Winden KD, Gdalyahu A, Dong H, Sonnenblick LI, Gruver R, Almajano J, Bragin A et al. 2011. Absence of CNTNAP2 leads to epilepsy, neuronal migration abnormalities, and core autism-related deficits. *Cell* **147**: 235-246.
- Penagarikano O, Lazaro MT, Lu XH, Gordon A, Dong H, Lam HA, Peles E, Maidment NT, Murphy NP, Yang XW et al. 2015. Exogenous and evoked oxytocin restores social behavior in the Cntnap2 mouse model of autism. *Science translational medicine* **7**: 271ra278.
- Poliak S, Gollan L, Martinez R, Custer A, Einheber S, Salzer JL, Trimmer JS, Shrager P, Peles E. 1999. Caspr2, a new member of the neurexin superfamily, is localized at the juxtaparanodes of myelinated axons and associates with K⁺ channels. *Neuron* **24**: 1037-1047.
- Poliak S, Salomon D, Elhanany H, Sabanay H, Kiernan B, Pevny L, Stewart CL, Xu X, Chiu SY, Shrager P et al. 2003. Juxtaparanodal clustering of Shaker-like K⁺ channels in myelinated axons depends on Caspr2 and TAG-1. *The Journal of cell biology* **162**: 1149-1160.
- Poot M. 2015. Connecting the CNTNAP2 Networks with Neurodevelopmental Disorders. *Mol Syndromol* **6**: 7-22.
- Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)* **26**: 139-140.
- Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An JY, Peng M, Collins R, Grove J, Klei L et al. 2020. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell* **180**: 568-584 e523.
- Schaafsma SM, Gagnidze K, Reyes A, Norstedt N, Mansson K, Francis K, Pfaff DW. 2017. Sex-specific gene-environment interactions underlying ASD-like behaviors. *Proc Natl Acad Sci U S A* **114**: 1383-1388.
- Schaffer AE, Breuss MW, Caglayan AO, Al-Sanaa N, Al-Abdulwahed HY, Kaymakcalan H, Yilmaz C, Zaki MS, Rosti RO, Copeland B et al. 2018. Biallelic loss of human CTNNA2, encoding alphaN-catenin, leads to ARP2/3 complex overactivity and disordered cortical neuronal migration. *Nat Genet* **50**: 1093-1101.
- Schmidt-Kastner R, van Os J, H WMS, Schmitz C. 2006. Gene regulation by hypoxia and the neurodevelopmental origin of schizophrenia. *Schizophr Res* **84**: 253-271.
- Schraut KG, Jakob SB, Weidner MT, Schmitt AG, Scholz CJ, Strekalova T, El Hajj N, Eijssen LM, Domschke K, Reif A et al. 2014. Prenatal stress-induced programming of genome-wide promoter DNA methylation in 5-HTT-deficient mice. *Translational psychiatry* **4**: e473.

- Schuch JB, Genro JP, Bastos CR, Ghisleni G, Tovo-Rodrigues L. 2018. The role of CLOCK gene in psychiatric disorders: Evidence from human and animal research. *American journal of medical genetics Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics* **177**: 181-198.
- Sebat J, Lakshmi B, Malhotra D, Troge J, Lese-Martin C, Walsh T, Yamrom B, Yoon S, Krasnitz A, Kendall J et al. 2007. Strong association of de novo copy number mutations with autism. *Science* **316**: 445-449.
- Shi SQ, White MJ, Borsetti HM, Pendergast JS, Hida A, Ciarleglio CM, de Verteuil PA, Cadar AG, Cala C, McMahon DG et al. 2016. Molecular analyses of circadian gene variants reveal sex-dependent links between depression and clocks. *Translational psychiatry* **6**: e748.
- Sibbe M, Forster E, Basak O, Taylor V, Frotscher M. 2009. Reelin and Notch1 cooperate in the development of the dentate gyrus. *J Neurosci* **29**: 8578-8585.
- Song CX, Szulwach KE, Fu Y, Dai Q, Yi C, Li X, Li Y, Chen CH, Zhang W, Jian X et al. 2011. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nature biotechnology* **29**: 68-72.
- Szczurkowska J, Pischedda F, Pinto B, Manago F, Haas CA, Summa M, Bertorelli R, Papaleo F, Schafer MK, Piccoli G et al. 2018. NEGR1 and FGFR2 cooperatively regulate cortical development and core behaviours related to autism disorders in mice. *Brain : a journal of neurology* **141**: 2772-2794.
- Szulwach KE, Li X, Li Y, Song CX, Han JW, Kim S, Namburi S, Hermetz K, Kim JJ, Rudd MK et al. 2011a. Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells. *PLoS Genet* **7**: e1002154.
- Szulwach KE, Li X, Li Y, Song CX, Wu H, Dai Q, Irier H, Upadhyay AK, Gearing M, Levey AI et al. 2011b. 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nature neuroscience* **14**: 1607-1616.
- Toma C, Pierce KD, Shaw AD, Heath A, Mitchell PB, Schofield PR, Fullerton JM. 2018. Comprehensive cross-disorder analyses of CNTNAP2 suggest it is unlikely to be a primary risk gene for psychiatric disorders. *PLoS Genet* **14**: e1007535.
- Tordjman S, Somogyi E, Coulon N, Kermarrec S, Cohen D, Bronsard G, Bonnot O, Weismann-Arcache C, Botbol M, Lauth B et al. 2014. Gene x Environment interactions in autism spectrum disorders: role of epigenetic mechanisms. *Frontiers in psychiatry* **5**: 53.
- Unterberger A, Andrews SD, Weaver IC, Szyf M. 2006. DNA methyltransferase 1 knockdown activates a replication stress checkpoint. *Mol Cell Biol* **26**: 7575-7586.
- van den Hove DL, Jakob SB, Schraut KG, Kenis G, Schmitt AG, Kneitz S, Scholz CJ, Wiescholleck V, Ortega G, Prickaerts J et al. 2011. Differential effects of prenatal stress in 5-Htt deficient mice: towards molecular mechanisms of gene x environment interactions. *PLoS One* **6**: e22715.
- Van den Hove DL, Steinbusch HW, Scheepens A, Van de Berg WD, Kooiman LA, Boosten BJ, Prickaerts J, Blanco CE. 2006. Prenatal stress and neonatal rat brain development. *Neuroscience* **137**: 145-155.
- Voineagu I, Eapen V. 2013. Converging Pathways in Autism Spectrum Disorders: Interplay between Synaptic Dysfunction and Immune Responses. *Frontiers in human neuroscience* **7**: 738.
- Wagner KV, Hartmann J, Labermaier C, Hausl AS, Zhao G, Harbich D, Schmid B, Wang XD, Santarelli S, Kohl C et al. 2015. Homer1/mGluR5 activity moderates vulnerability to

- chronic social stress. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* **40**: 1222-1233.
- Wang HY, Hsieh PF, Huang DF, Chin PS, Chou CH, Tung CC, Chen SY, Lee LJ, Gau SS, Huang HS. 2015. RBFOX3/NeuN is Required for Hippocampal Circuit Balance and Function. *Scientific reports* **5**: 17383.
- Wang T, Pan Q, Lin L, Szulwach KE, Song CX, He C, Wu H, Warren ST, Jin P, Duan R et al. 2012. Genome-wide DNA hydroxymethylation changes are associated with neurodevelopmental genes in the developing human cerebellum. *Hum Mol Genet* **21**: 5500-5510.
- Weinstock M. 1997. Does prenatal stress impair coping and regulation of hypothalamic-pituitary-adrenal axis? *Neurosci Biobehav Rev* **21**: 1-10.
- Weiss LA, Arking DE, Daly MJ, Chakravarti A. 2009. A genome-wide linkage and association scan reveals novel loci for autism. *Nature* **461**: 802-808.
- Wong ML, Dong C, Maestre-Mesa J, Licinio J. 2008. Polymorphisms in inflammation-related genes are associated with susceptibility to major depression and antidepressant response. *Mol Psychiatry* **13**: 800-812.
- Wong TH, Chiu WZ, Breedveld GJ, Li KW, Verkerk AJ, Hondius D, Hukema RK, Seelaar H, Frick P, Severijnen LA et al. 2014. PRKAR1B mutation associated with a new neurodegenerative disorder with unique pathology. *Brain : a journal of neurology* **137**: 1361-1373.
- Yang Y, Li N, Qiu J, Ge H, Qin X. 2020. Identification of the Repressive Domain of the Negative Circadian Clock Component CHRONO. *International journal of molecular sciences* **21**.
- Yao B, Jin P. 2014. Cytosine modifications in neurodevelopment and diseases. *Cell Mol Life Sci* **71**: 405-418.
- Zannas AS, Wiechmann T, Gassen NC, Binder EB. 2016. Gene-Stress-Epigenetic Regulation of FKBP5: Clinical and Translational Implications. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology* **41**: 261-274.
- Zeng J, Konopka G, Hunt BG, Preuss TM, Geschwind D, Yi SV. 2012. Divergent whole-genome methylation maps of human and chimpanzee brains reveal epigenetic basis of human regulatory evolution. *Am J Hum Genet* **91**: 455-465.
- Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9**: R137.
- Zhubi A, Chen Y, Dong E, Cook EH, Guidotti A, Grayson DR. 2014. Increased binding of MeCP2 to the GAD1 and RELN promoters may be mediated by an enrichment of 5-hmC in autism spectrum disorder (ASD) cerebellum. *Transl Psychiatry* **4**: e349.
- Zweier C, Peippo MM, Hoyer J, Sousa S, Bottani A, Clayton-Smith J, Reardon W, Saraiva J, Cabral A, Gohring I et al. 2007. Haploinsufficiency of TCF4 causes syndromal mental retardation with intermittent hyperventilation (Pitt-Hopkins syndrome). *Am J Hum Genet* **80**: 994-1001.

Figure Legends

Figure 1: Results from social behavioral tests in female offspring. (A-D) 3-chamber social interaction test: (A-B) time spent in the chamber with either an unfamiliar mouse inside a cup (mouse) or with an empty cup (empty). Data analysis details: (A) two way ANOVA for control-WT versus PS-WT groups: effect of time spent in either chamber: $F_{(1,42)} = 28.5$, P -value < 0.001 . (B) two way ANOVA for control-HET versus PS-HET groups: effect of time spent in either chamber: $F_{(1,32)} = 16.2$, P -value < 0.001 and interaction between time spent in either chamber and treatment: $F_{(1,32)} = 5.3$, P -value = 0.002. (C-D) Time spent interacting with either an unfamiliar mouse inside a cup (mouse) or with an empty cup (empty). Data analysis details: (C) two way ANOVA for control-WT versus PS-WT groups: effect of time spent interacting: $F_{(1,42)} = 29$, P -value < 0.001 . (D) two way ANOVA for control-HET versus PS-HET groups: effect of time spent interacting in either chamber: $F_{(1,32)} = 9.9$, P -value = 0.003). All asterisks denote P -value ≤ 0.05 , two-way ANOVA with Bonferroni multiple comparison test determined post hoc significance. Δ denotes the difference in time spent. (E) Ten-minute reciprocal social interaction test in female offspring. One-way ANOVA detected effect of groups on time spent grooming/digging during a 10-minute reciprocal social interaction test ($F_{(3,16)} = 10.9$, P -value < 0.001). Hashtag (#) denotes significance for time spent grooming/digging between PS-HET in comparison to Control-WT, Control-HET, and PS-WT (P -value < 0.05). $N = 9$ -10 per treatment for each genotype.

Figure 2: Distribution of hippocampal HET DhMRs. (A) A Manhattan plot shows the distribution of 5hmC sequenced data across the genome (x-axis; alternating black/grey for alternating Chromosomes) and the level of significance ($-\log_{10}(P$ -value)) along the y-axis. Dots

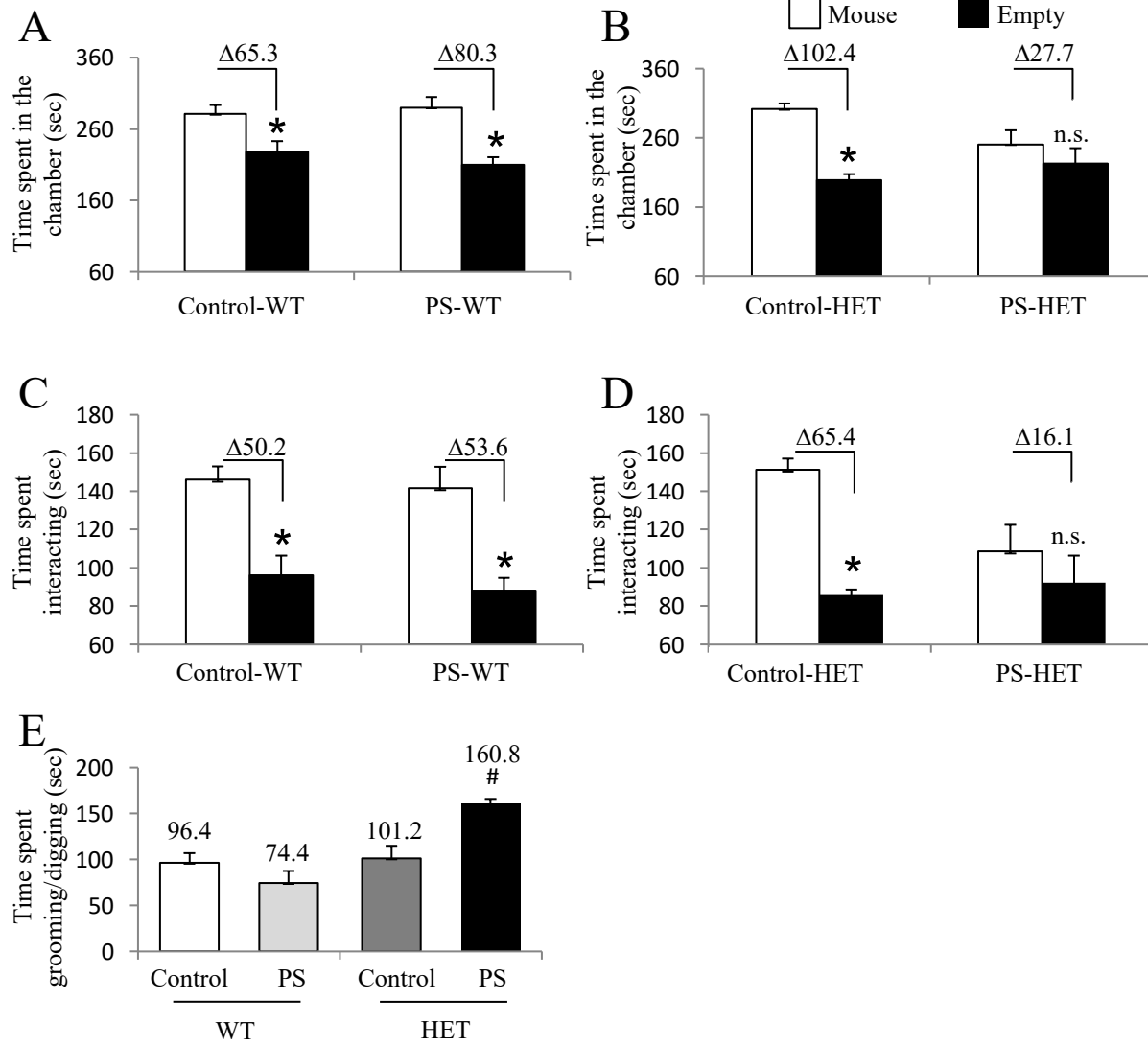
above the top blue line represent PS-HET-specific (hyper)-DhMRs, while dots below the bottom blue line represent Control-HET-specific (hypo)-DhMRs (P -value < 0.05). (B) A pie chart displays the proportion of DhMRs across standard genomic structures. (C-D) Bar plot showing the top ten significantly over-represented ontological terms (y-axis) based on gene ontological analysis of PS-HET-specific DhMR-associated genes (C) and Control-HET-specific DhMR-associated genes (D). The number of DhMR-associated genes linked to the ontological terms are displayed on the x-axis. Bar color is based on P -value.

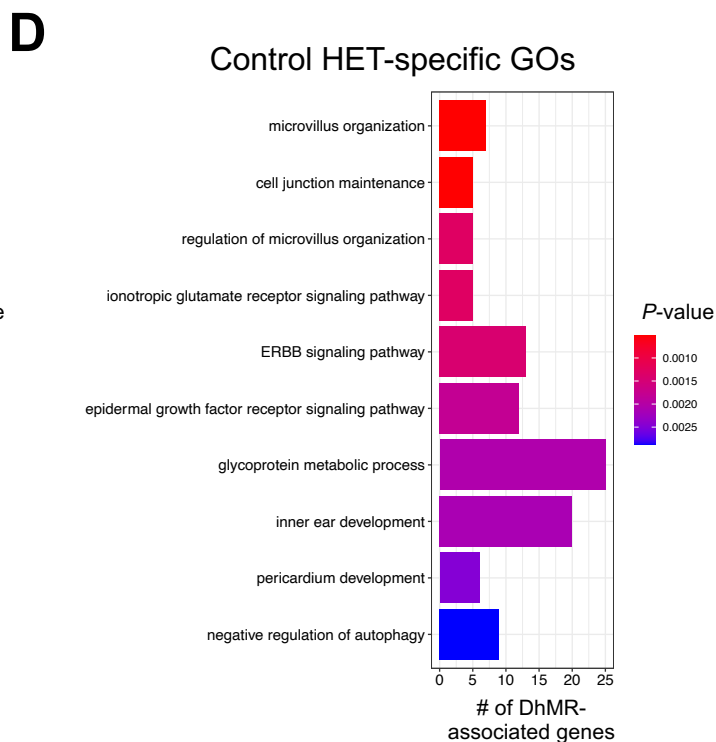
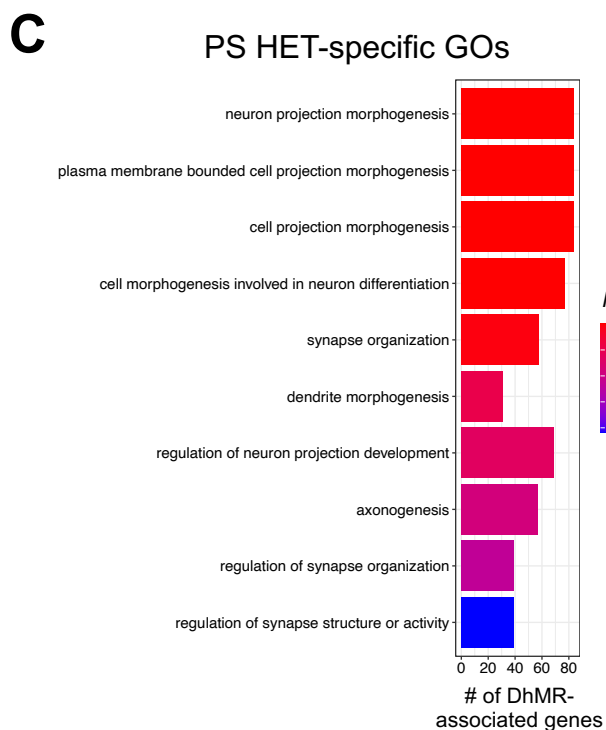
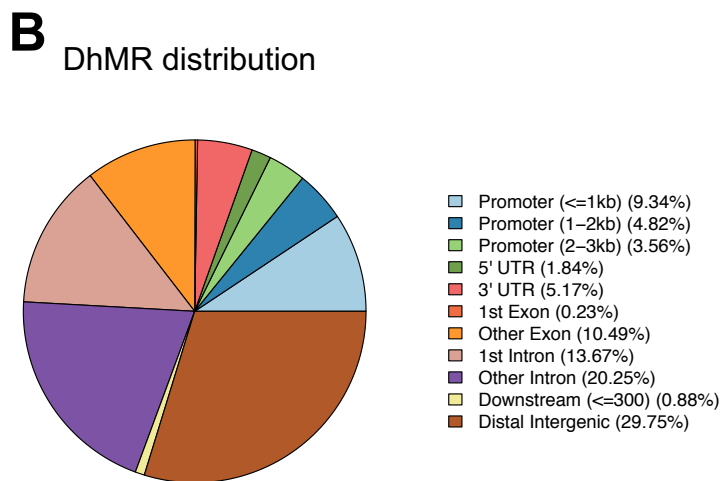
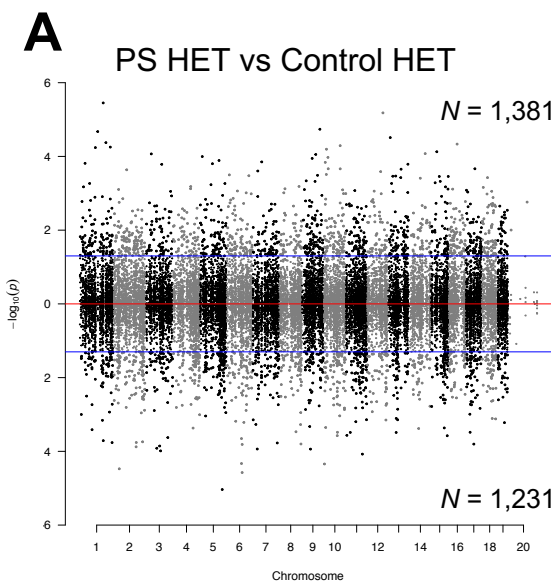
Figure 3: Characterization of differentially expressed genes (DEGs) between PS-HET and Control-HET. (A) A modified volcano plot depicts the \log_2 (posterior fold change; x-axis) versus the posterior probability of differential expression (y-axis). Open circles (red/black) represent each gene examined and differentially expressed genes (red) are shown above the significance line (blue, FDR P -value < 0.1). (B-C) A bar plot showing the top significantly over-represented ontological terms based on gene ontological analysis of PS-HET-specific (B) differentially expressed genes and Control-HET-specific (C) differentially expressed genes. The number of DEGs linked to the ontological terms are displayed on the x-axis. Bar color is based on P -value. (D) A Venn diagram depicts the overlap of DhMR-associated genes (green circle) and differentially expressed genes (orange circle) from the PS-HET versus Control-HET comparison. The asterisk (*) denotes a significant overlap (hypergeometric enrichment P -value < 0.05).

Figure 4: Functional assays reveal that 5hmC represses CLOCK binding. (A) Western blot analysis identified no gross differences in CLOCK protein expression between Control-WT (lanes 1 and 2), PS-WT (lanes 3 and 4), Control-HET (lanes 5 and 6), or PS-HET (lanes 7 and 8)

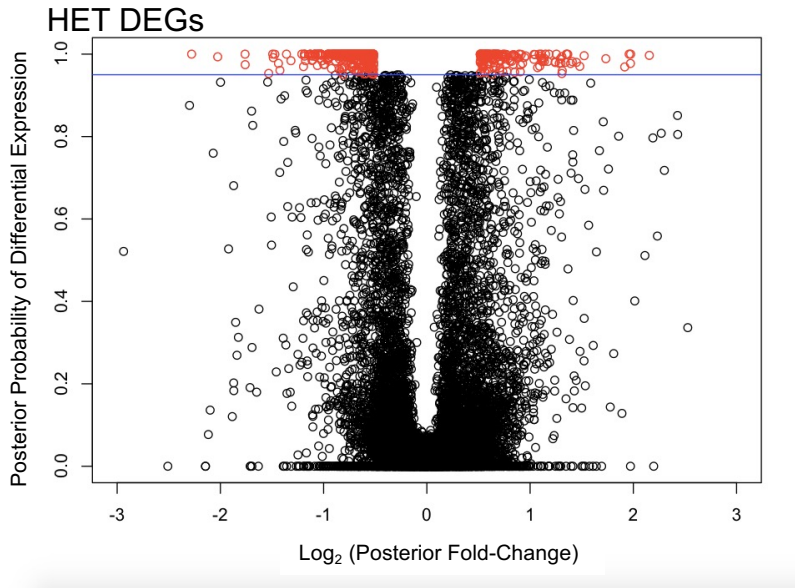
(top panel) or in ACTB (middle panel), used as a loading control in all lanes. A bar plot (bottom panel) depicts the quantified levels of CLOCK proteins normalized to ACTB from each respective lane. Error bars show the standard error of the mean ($N=2$). (B) A volcano plot displays the $-\log_2(\text{fold-change})$ of PS-HET compared to Control-HET (x-axis) and the significance (y-axis) from differential analysis of CHIP-seq data. Open circles (blue/black) and red triangles represent differentially hydroxymethylated regions (DhMRs) containing a canonical CLOCK binding site. DhMRs with significant differential binding of CLOCK (blue open circles and red triangles) are located above the significance line ($P\text{-value} < 0.05$). Red triangle DhMRs also exhibit differential gene expression (*Gigyf1*, *Palld*, and *Fry*). (C) Gene schematics of *Gigyf1* (blue), *Palld* (yellow), and *Fry* (orange) are shown and indicate the location of the DhMRs containing the differential CLOCK binding (green triangle). (D) Electromobility shift assays with hippocampal lysates. Lysate shift (lanes 2, 6, and 10), cold competition (lanes 3, 7, and 11), and supershift (lanes 4, 8, and 12) of DNA probes corresponding to a CLOCK E-box motif in the DhMR of *Palld* that exhibited differential CLOCK binding. Biotinylated DNA probes contained either an unmodified E-box motif (lanes 1-4), a methylated E-box motif (lanes 5-8), or a hydroxymethylated E-box motif (lanes 9-12). Regions containing bands of interest are highlighted (dashed boxes) with DNA::protein complexes that are abrogated with cold competition are indicated (black arrowheads) and supershift complexes are shown (carrot). (E-G) Quantifications of biotinylated DNA probe intensities. The E-box modifications (x-axis) and mean biotinylated DNA probe intensity (y-axis) are depicted \pm SEM for *Gigyf1* (E), *Palld* (F), and *Fry* (G) probes. The asterisk (*) denotes a significant decrease in CLOCK binding box (* $P\text{-value} < 0.05$; *** $P\text{-value} < 0.001$).

Figure 5: A working model of a functional role for 5hmC in developmental brain disorder-like behaviors. Control *Cntnap2*^{+/-} mice (left panel) 5hmC levels (orange lollipops) remain abundant and prevent CLOCK (blue semicircle) from binding to genes implicated in developmental brain disorders (*e.g.*, *Gigyf1*), which results in adequate gene expression levels (black curved lines) and unaffected behaviors. *Cntnap2*^{+/-} mice exposed to prenatal stress (PS; lightning bolt; right panel) exhibit a loss of 5hmC abundance in genes implicated in developmental brain disorders (*e.g.*, *Gigyf1*). This reduction in 5hmC levels allows CLOCK to bind to DNA and repress transcription, which results in developmental brain disorder-like phenotypes.

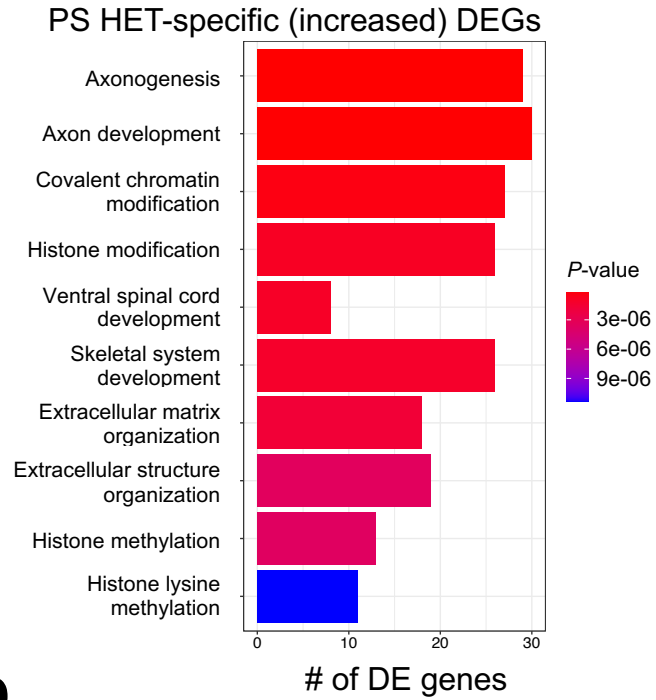




A

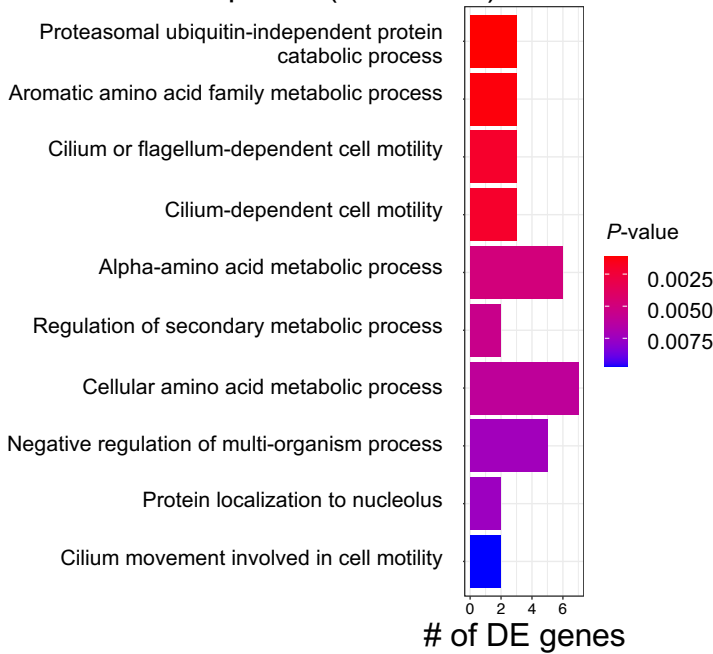


B

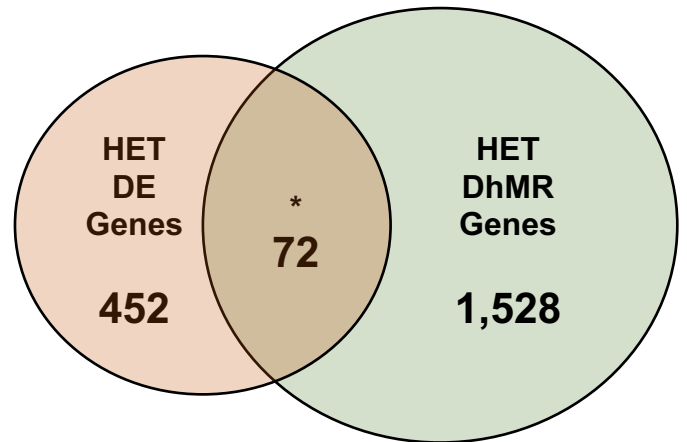


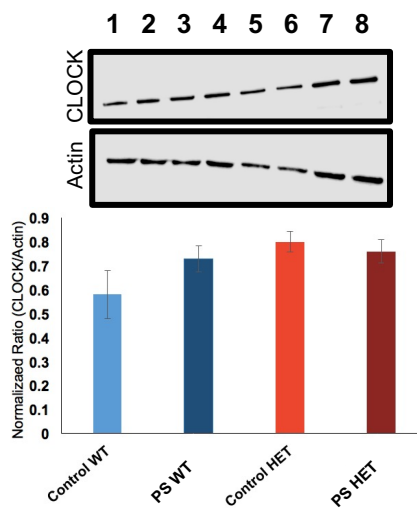
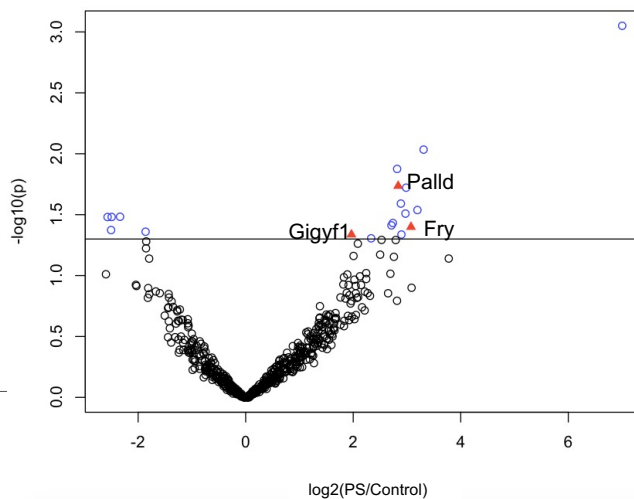
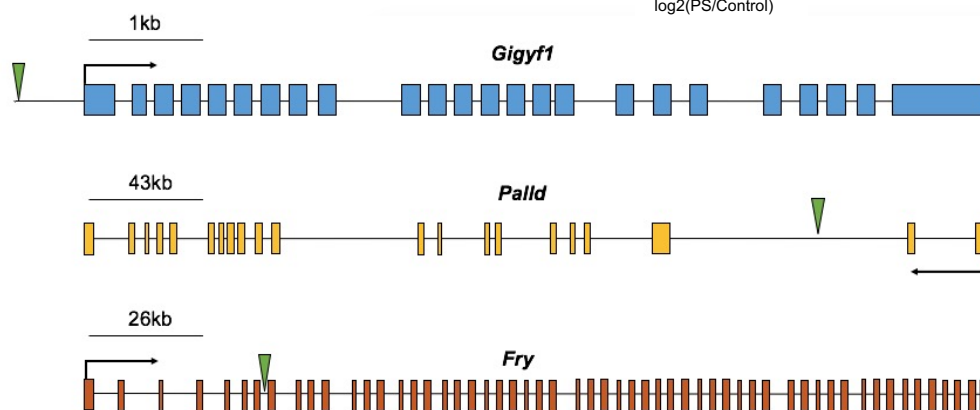
C

Control HET-specific (decreased) DEGs



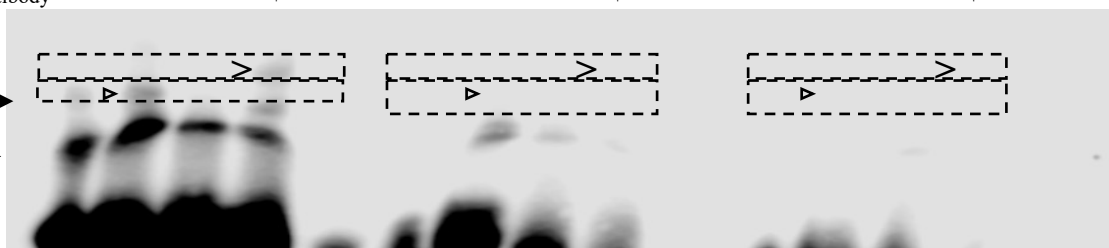
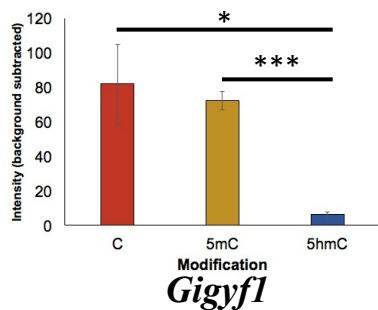
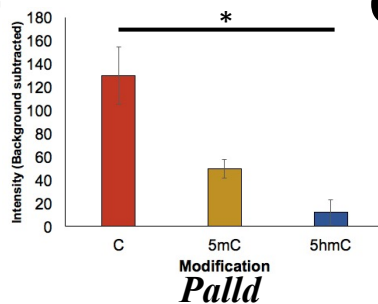
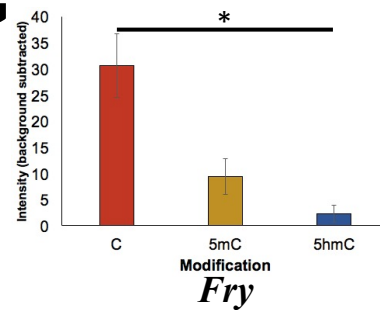
D



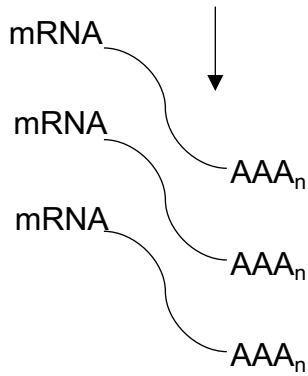
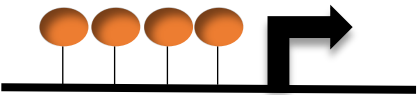
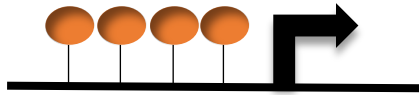
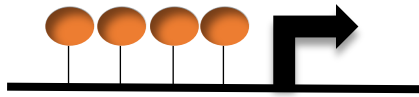
A**B****C****D**

Free <i>Palld</i> probes	+	+	+	+	-	-	-	-	-	-	-	-
5mC <i>Palld</i> probes	-	-	-	-	+	+	+	+	-	-	-	-
5hmC <i>Palld</i> probes	-	-	-	-	-	-	-	-	+	+	+	+
Nuclear extracts	-	+	+	+	-	+	+	+	-	+	+	+
Unlabeled probes	-	-	+	-	-	-	+	-	-	-	+	-
Antibody	-	-	-	+	-	-	-	+	-	-	-	+

Supershift
DNA::protein
complexes
→

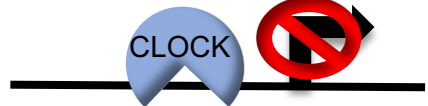
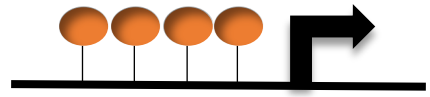
**E****F****G**

Cntnap2^{+/-}



Unaffected Behavior

Cntnap2^{+/-}



No transcripts

Developmental Brain Disorder-like Behavior

Table 1: Summary of behavioral tests in female offspring

	<i>Control</i>		<i>Prenatal Stressed</i>	
	WT	<i>Cntnap2</i> ^{+/-}	WT	<i>Cntnap2</i> ^{+/-}
Open Field				
Time spent in the center (s)	44 ± 10	35 ± 3	42 ± 4	45 ± 3
Number of entries to the center	19 ± 3	17 ± 1	19 ± 1	19 ± 0
Elevated Plus Maze				
Time spent in closed arm (s)	222 ± 9	244 ± 8	240 ± 12	233 ± 9
Time spent in open arm (s)	204 ± 11	166 ± 8	182 ± 13	183 ± 9
Light/Dark Box				
Time spent in light side (s)	352 ± 18	357 ± 19	392 ± 11	376 ± 24
Forced swim test				
Time spent floating	194 ± 7	152 ± 13	172 ± 16	192 ± 11
Ten-minute grooming test				
Time spent grooming	64.5±8.8	52.2±4.8	59.3±7.8	44.2±5

Comparison of different behavioral tests between prenatal stressed (PS) and non-stressed (Control) *Cntnap2*^{+/-} and WT mice. The value are presented as mean ± S.E.M. *p≤ 0.05. N = 6-8. Data analysis: 1) Open Field (time spent in the center: effect of PS $F_{(1,31)} = 0.59$, $P = 0.4$; effect of genotype $F_{(1,31)} = 0.9$, $P = 0.3$; number of entries to the center: effect of PS $F_{(1,31)} = 0.27$, $P = 0.6$; effect of genotype $F_{(1,31)} = 0.02$, $P = 0.8$), 2) Elevated Plus Maze (time spent in closed arm: effect of PS $F_{(1,34)} = 0.08$, $P = 0.7$; effect of genotype $F_{(1,34)} = 0.4$, $P = 0.5$; time spent in open arm: effect of PS $F_{(1,34)} = 0.04$, $P = 0.8$; effect of genotype $F_{(1,34)} = 2.8$, $P = 0.1$), 3) Light/Dark Box (time spent in light side: effect of PS $F_{(1,19)} = 2.4$, $P = 0.13$; effect of genotype $F_{(1,19)} = 0.1$, $P = 0.7$) and 4) Forced Swim Test (time spent floating: effect of PS $F_{(1,26)} = 0.31$, $P = 0.5$; effect of genotype $F_{(1,26)} = 0.49$, $P = 0.4$). 5) Ten-minutes grooming test: effect of PS $F_{(1,34)} = 0.8$, $P = 0.3$; effect of genotype $F_{(1,34)} = 3.6$, $P = 0.06$.

Table 2: A subset of the transcription factors with binding motifs associated with DhMRs in hippocampus PS-HET x Control-HET

Gene	Name	Role	FDR P-value*	DhMR type
<u>Arnt (Hif1b)</u>	Aryl hydrocarbon receptor nuclear translocator	Cellular oxygen sensing	0.00001	PS-HET, Ctrl-HET
<i>Arntl1</i>	Brain and Muscle ARNT-Like 1	Regulates biological functions under circadian control	0.00001	Ctrl-HET
<u>Bhlhe40/41</u>	Basic helix-loop-helix family	Regulation of circadian rhythms	0.00001	PS-HET, Ctrl-HET
<u>Clock</u>	Circadian locomotor output cycles kaput	Maintenance of circadian rhythms	0.00001	PS-HET, Ctrl-HET
<u>Epas1 (Hif2a)</u>	Endothelial PAS domain protein 1	Cellular oxygen sensing	0.00001	PS-HET, Ctrl-HET
<u>Hif1a</u>	Hypoxia-inducible factors	Cellular oxygen sensing	0.00001	PS-HET, Ctrl-HET
<i>Max</i>	Myc-associated factor X	Cell proliferation and apoptosis	0.00001	Ctrl-HET
<i>Mnt</i>	Max-binding protein MNT	Gene-specific transcriptional activation or repression	0.00001	Ctrl-HET
<u>Myc</u>	Cellular Myelocytomatosis	Regulates cell division and proto-oncogenes	0.00001	PS-HET, Ctrl-HET
<u>Nf1</u>	Nuclear factor I	Neuronal regulation	0.00001	PS-Het, Ctrl-HET
<u>Npas2</u>	Neuronal PAS domain protein	Paralogous to Clock, maintenance of circadian rhythms	0.00001	Ctrl-HET
<u>Usf1/2</u>	Upstream stimulator factor	Cell metabolism	0.00001	PS-HET, Ctrl-HET

Underlined transcription factors are also found in hippocampal PS-WT x Control WT

*FDR P-value for the enrichment of each transcription factor binding motif in hippocampal tissue