Transcription and DNA methylation signatures of paternal care in hippocampal dentate gyrus of prairie voles

Nicholas J. Waddell[1], Yan Liu[2,3], Javed M. Chitaman[1,3], Graham J. Kaplan[1,3], Zuoxin Wang[2,3,#], Jian Feng[1,3,#]

1. *Department of Biological Science, Florida State University, Tallahassee, FL 32306, USA.*
2. *Department of Psychology, Florida State University, Tallahassee, FL 32306, USA.*
3. *Program in Neuroscience, Florida State University, Tallahassee, FL 32306, USA.*

#Correspondence:

Zuoxin Wang, Ph.D.

Email: zwang@psy.fsu.edu

Phone: 850 644 5057

Jian Feng, Ph.D.

Email: feng@bio.fsu.edu

Phone: 850 644 3706

ABSTRACT

In socially monogamous prairie voles (*Microtus ochrogaster*), parental behaviors not only occur in mothers and fathers, but also can exist in virgin males (“parental”). However, some virgin males display aggressive behaviors towards conspecific pups (“attacker”). Although this behavioral dichotomy in response to pup exposure has been well documented, little is known about the gene expression changes underlie the parental behavioral differences and their regulatory mechanisms. To address this, we profiled the transcriptome and DNA methylome of hippocampal dentate gyrus of four prairie vole groups, attacker virgin males, parental virgin males, fathers, and mothers. We found a concordant pattern of gene transcription in parental virgin males and fathers, when comparing to the attacker group. The methylome analysis also revealed signaling pathways enriched for epigenetic changes involving both receptor-mediated and secondary messenger signaling across both behavioral phenotypes and sexual experiences. Furthermore, we found correlations between gene expression changes and DNA methylation differences between attacker and parental virgin males, which suggests a canonical gene expression regulatory role of DNA methylation in paternal care. Therefore, our study presents an integrated view of transcriptome and epigenome that provides a DNA epigenetic based molecular insight of paternal behavior.

INTRODUCTION

In mothers, gene expression changes across brain regions underlie the physiological and behavioral adaptations for pup-rearing[1]. Though still few in number, accumulating evidence suggests that neural gene expression changes are associated with sexual experience in fathers of biparental species, where both parents participate in pup-rearing and paternal care contributes to pup development, maturation, and reproductive success[2]. However, the mechanism of gene expression regulation responsible for these processes remains elusive.

The prairie vole, *Microtus orchrogaster*, has become a valuable organism to model social bonding where vole individuals mate exclusively, share nests, and exhibit biparental care of newborn pups[3-5]. Both mother and father voles nearly equally participate in parental behaviors, such as pup grooming, huddling, retrieving, and nest building [6]. Though these parental behaviors normally present after a litter is born, they may spontaneously occur in some sexually naïve voles when exposed to conspecific pups ("parental")[7, 8]. While about 60% virgin males are spontaneously parental, other virgin males display aggression towards conspecific pups ("attacker")[9-11]. Therefore, to study the behavioral dichotomy in male prairie voles may provide molecular insights of paternal care.

Environmental stimuli play a large role in parental behavior manifestation and modulation during post-partum periods, which can also affect gene expression programs through epigenetic regulation of gene expression[12, 13]. For example, it was shown that pup exposure facilitates parental behaviors, and this response can be altered by using compounds that affect epigenetic states, such as histone deacetylase inhibitors[14]. However, the potential role of DNA methylation, a major epigenetic mechanism, in paternal behaviors remains largely unknown. Though DNA methyltransferases (DNMTs) catalyze the covalent addition of a methyl group to the cytosine nucleotide, DNA methylation can also be altered through a series of oxidation reactions by the Ten-eleven translocation (TET) methylcytosine oxidases, that may lead to unmethylated cytosines[15-17]. DNA methylation has been implicated in various basic brain functions and diseases[18-20]. Notably, differential DNA methylation in the hippocampus has been related to altered maternal care[21].

Mating and social interaction which lead to pair bond formation has been found to modulate cell proliferation and differentiation to neuronal precursor cells in the dentate gyrus (DG) of hippocampus [22, 23]. While pup exposure elicited cell proliferation in the DG[24], fatherhood decreased cell survival in the DG[25]. In addition, exposure to psychostimulant drugs, such as amphetamine, not only diminished pair bonding[26, 27], but also impaired social recognition and decreased neuronal and neurochemical

activation in the DG[28]. In prairie voles, DG contains receptors for oxytocin[29] a key molecule in pair bonding and parental behaviors. It was found that oxytocin receptors are subjected to DNA methylation mediated gene expression regulation[30], whose density in DG is associated with mating tactics and reproductive success in male voles[31]. Although these indicate a potential role of DNA methylation in DG's function in mediating social behaviors in prairie voles, it remains unknown how gene expression and DNA methylation changes occur across the whole genome and their interplay in parental behaviors. Here, we aim to address this question by profiling transcriptome and DNA methylome in DG with a focus on the paternal behavioral dichotomy in prairie voles.

METHODS

Animal Subjects

Subjects were sexually-naïve male and female prairie voles, *Microtus ochrogaster*, from a laboratory breeding colony. Subjects were weaned at 21 days of age and housed in same-sex sibling pairs in plastic cages (12 W x 28 L x 16 H cm) containing cedar chip bedding with water and food provided ad libitum. All cages were maintained under a 14:10 light:dark cycle, and the temperature was kept at 20°C. Adult subjects (at 90-120 days of age) were randomly assigned into experimental groups in which males and females were paired or virgin males were continuously housed in same-sex sibling pairs. Females gave birth following 21-23 days of pairing with a male, and the mother and father voles were continuously housed with their offspring. At three days postpartum, Mothers and fathers were tested for their parental behaviors towards a conspecific pup. Age-matched virgin males were also tested for their spontaneous parental behaviors towards a conspecific pup. As virgin males either display parental behaviors towards pups or attack pups[10, 32, 33]), these males were further classified into "Parental" and "Attacker" groups, respectively. All experimental procedures were approved by the Florida State University Institutional Animal Care and Use Committee and were in accordance with the U.S. National Institutes of Health Guide for the Care and Use of Laboratory Animals[34].

Parental Behavior Test

The parental behavior test was conducted as previously described[10, 32, 33]. Briefly, all subjects were tested in a Plexiglas cage (20 W x 45 L x 25 H cm) with a thin layer of cedar chip bedding and ad lib food and water as described for the housing cages. The subject was placed in the testing cage and allowed for a 15-min habituation. Afterwards, an unfamiliar stimulus pup (at 3-day age) was introduced into the testing cage at the opposite corner from the subject, and the subject's behaviors were digitally recorded for 60 minutes. If the virgin male subject showed aggression towards the pup, the experimenter tapped the cage to immediately stop the behavioral testing, the pup was removed, and the subject was classified as "Attacker".

All behavioral videos were scored by a trained experimenter blind to the treatment groups using JWatcher software programv1.0[35]. The duration and frequency in the first 10-min of the subject's interactions with the stimulus pup were quantified. The scored behaviors included both parental behaviors (pup huddling, retrieving, licking and grooming, and nest building) and non-parental behaviors (autogrooming, locomotion, sniffing, and resting)[8, 32, 36]. For virgin male attackers, only the latency of the first attack was scored.

Behavior Data Analysis

Group differences in all behavioral measurements were analyzed using a one-way ANOVA. Post-hoc analyses were conducted using Tukey's HSD tests ($p < 0.05$). Plots generated from ggplot2 v3.3.3[37].

Brain Tissue Collection

After the parental behavioral test, subjects were decapitated. Brains were extracted and immediately frozen on dry-ice. Brains were sliced into 200 µm sections on a cryostat and thaw mounted on slides. Thereafter, 1 mm-diameter punches from 4 consecutive sections were taken bilaterally from the dentate gyrus of the hippocampus[38]. Tissue punches were stored at -80ºC until further processing.

Next Generation Sequencing Library Preparation

DNA and RNA were isolated from the same tissue using the Qiagen AllPrep DNA/RNA Micro Kit (Qiagen, #80284) according to the manufacturer's protocol and then quantified by Qubit fluorometry.

150 ng of total RNA from one single animal was applied to the NEBNext rRNA Depletion Kit (New England Biolabs, #E6310L) according to the manufacturer's protocol. Ribo-depleted RNA sequencing (RNA-seq) libraries were constructed using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs, #E7765S). RNA samples were fragmented, and cDNA was synthesized. The ends of cDNA fragments were ligated with universal Illumina adapter sequences. RNAseq libraries were individually indexed with NEBNext Multiplex Oligos for Illumina (New England Biolabs, #E7335S) and amplified for 11 cycles of PCR amplification. All clean-up steps were accomplished by using the supplied purification beads within the NEBNext Ultra II Directional RNA Library Prep Kit. RNAseq libraries were then sequenced 50-bp paired-ended on an Illumina NovaSeq 6000 Sequencer with a 5% PhiX spike-in control.

Reduced representation bisulfite sequencing (RRBS)[39] libraries were prepared using a Premium RRBS Kit (Diagenode, #C02030032) according to the manufacturer's protocol. Briefly, 100 ng high-quality genomic DNA from one single animal was digested with MspI,

end-repaired, ligated to adapters, and then size selected using AMPure XP beads (Beckman Coulter, #A63881). All size-selected samples were treated with sodium bisulfite conversion. Spike-in control DNA was included for the monitoring of bisulfite conversion efficiency. Libraries were then purified after 14 cycles of PCR amplification. An Agilent Bioanalyzer and KAPA Library Quantification Kit were run to assess library quality and quantity. RRBS libraries were sequenced 100-bp single-ended on an Illumina NovaSeq 6000 Sequencer with a 5% PhiX spike-in control.

Sequencing Data Pre-processing

Raw sequencing reads were evaluated for quality and sequencing artifacts using FastQC v0.11.9[40]. To address a positional sequencing error within the RNAseq library, where an entire tile within the sequencing chip had extremely low quality, FilterByTile[41] was applied without incurring biases on the rest of the dataset. Afterwards, sequencing reads were trimmed of adapters and to a minimum of 20 quality score and 20 read length using TrimGalore v0.6.4[42]. RRBS sequencing reads were trimmed with TrimGalore's RRBS mode which eliminates synthetic cytosine signals from the ends of reads that was incorporated during the end-repair process. Reads were checked for quality again after trimming.

RNAseq Analysis

Alignment, Assignment, and Differential Expression

Alignment of RNAseq reads was performed using the splice-junction aware alignment software STAR v2.5.4b[43] with the MicOch1.0[44] annotation set as the reference genome. Aligned reads were assigned and counted to gene-level features using FeatureCounts v2.0.0[45].

Gene-level counts from RNAseq reads were imported to R and analyzed using EdgeR v3.28.1[46]. Genes with low counts and those missing counts from most of the samples per group were removed from the analysis, resulting in about 70% of annotated genes to be considered in downstream analyses. Subsequently, we estimated dispersion, biological coefficients of variation (BCV) and normalization factors for the dataset. RNAseq samples were evaluated in two-dimensional space using multi-dimensional scaling to determine whether certain principal components of the dataset are driving the variation among and between groups. Then, the RNAseq design matrix and generalized linear models were created contrasting the four groups of voles (Father, Mother, Parental, Attacker) in the experiment. Mother voles were only compared to father voles as these groups have experienced pair-bonding while the sexually naïve male vole groups have not, in order to reduce confounding variation in the analysis. Hypothesis testing was performed through the likelihood ratio test and any genes with a Log_2 fold change less

than -0.5 or greater than 0.5 with a significant p-value were claimed differentially expressed. These results were represented using volcano plots, Log_2 fold change on the x-axis with $-Log_{10}$ p-values on the y-axis, generated using ggplots2 v3.3.3[37].

Gene Ontology Enrichment Analysis

As Kyoto Encylcopedia of Gene and Genomes (KEGG)[47] pathways do not have an annotation for the prairie vole, differentially expressed prairie vole genes were annotated to orthologous gene IDs from the mouse genome using Ensembl's biomart[44] annotation database and SQL manipulations. Gene IDs with at least 80% homology between mouse and prairie vole genes were considered for the enrichment analysis. The analysis was performed using a web-based tool, WebGestalt[48], a hypergeometric overlaps test. Significant pathways were considered by an FDR value < 0.05. For biological process gene ontology pathway enrichment testing, differentially expressed prairie vole gene names were passed to gProfiler[49] for gene ontology overrepresentation analysis using Ensembl's[44] prairie vole annotation. Enrichment ratio is defined by the following formula:

$$\frac{Intersection\ Size}{\left(\frac{Term\ Size}{Effective\ Domain\ Size}\right) * Query\ Size}$$

Rank-Rank Hypergeometric Overlaps

Rank-rank hypergeometric overlaps (RRHO) analysis identifies overlapping transcriptome expression profiles without pre-set thresholds, and determines the degree and the direction of overlapping genes[50]. An improved version is applied to allow discordant signatures to be assessed as robustly as concordant signatures. With this, visualization of each quadrant is separated, where the lengths of each side representing the relative length of each input gene list[51, 52]. Each expression list was ranked by multiplying the $-Log_{10}$(p-value) and the sign of the Log_2 fold expression change. RRHO difference maps were generated by representing the p-value from the hypergeometric test, and their significance adjusted through the Benjamini and Yekutieli method[53] and $-Log_{10}$ transformed.

Clustering Analysis

Differential Expression Clustering: All genes that had a significant differential expression between experimental groups were collected, and normalized gene counts were formatted into a table in R. Z-scores were calculated for each gene and passed to the pheatmap v1.0.12[54] package for hierarchical clustering using k-means clustering with Euclidean distance.

Gene Ontology Clustering: Genes that were found in cluster 6 of the Differential Expression clustering heatmap were evaluated for overrepresented ontologies as shown above. However, there were many ontologies found enriched in this dataset. To simplify the interpretation, the gene ontology pathways were clustered, and a network was constructed using Gene Ontology Markov Clustering (GOMCL)[55]. The resulting network and annotation table were passed to Cytoscape v3.9.1[56] for network visualization.

RRBS Analysis

Alignment, Methylation Calling, Differential Methylation Analysis

Bismark v22.3[57] genome preparation tool was used to create appropriate reference genomes for bisulfite sequencing alignment. Subsequently, quality-checked RRBS sequencing reads were aligned using the bowtie2 based methylation alignment algorithm in the Bismark suite with increased seed extension effort (parameters: -N 1, -L 20, -D 20). To extract methylation status from the alignment data, a methylation extractor tool of Bismark is applied to create a sample-wise list of CpG positions with the number of reads at that location with methylated calls and unmethylated calls. The resulting files were formatted and imported to R for differential methylation analysis. Differential methylation was calculated using a similar design structure to the differential expression analysis, which was processed using DSS general v2.34.0[58] that implements a Bayesian hierarchical model for dispersion estimation of the beta binomial distribution. Significantly methylated CpG sites were those from the Wald testing procedure with p-value thresholding < 0.05 and an absolute methylation difference of 15 percent. This data was represented by a volcano plot created using ggplot2[37].

Gene Annotation and KEGG Pathway Analysis

Differentially Methylated CpGs were annotated to an imported prairie vole genome using a Homer suite tool[59]. CpGs were annotated to the closest transcription starting site (TSS) and their genomic feature context was evaluated. The associated gene names for these differential CpGs were evaluated for overrepresented KEGG pathways as mentioned above using the orthologous mouse gene annotations, GRMc38[44], from Ensembl's Biomart database.

RNAseq and RRBS Correlation Analysis

Differentially expressed genes with differential methylation loci located within their respective gene boundary were filtered to be evaluated. When examining within the entire gene boundary, all genes with both expression and DNA methylation changes were separated into deciles according to the values of Log_2 fold change from differential gene expression and gene-boundary associated differential methylation values were assigned to their gene. In addition, the effects of promoter DNA methylation on gene expression

were also evaluated, in which all differentially expressed genes were classified into four quartiles based on the value of Log2 differential gene expression fold changes as long as they contained differential methylation in their promoter. Gene promoters were assigned by using default parameters in Homer as mentioned above. Finally, spearman's correlation is applied to check both the deciles and quartiles for significant correlations.

RESULTS

Behavior

During the pup exposure test, both mother and father voles displayed high levels of parental behaviors. Some virgin males also displayed spontaneous parental behaviors including huddling over, retrieving, licking and grooming the conspecific pup as well as nest building (Parental). Although no differences were noticed in overall parental behavior frequency and duration, a significant difference was found in pup huddling between mothers and parental voles, revealed through a one-way ANOVA ($F(2,18) = 6.812$, $p = 0.006$, Tukey's HSD $p = 0.005$, Fig 1C). There was also a significant difference in huddle duration in the same direction ($F(2,18) = 4.453$, $p = 0.027$, Tukey's HSD $p = 0.023$, Fig. 1E). In contrast, some other virgin males displayed aggressive behavior toward the conspecific pup (Attacker) as indicated by their short latency to attack (Fig. 1B). Among the non-parental behaviors, a group difference was found in rest frequency, a self-serving behavior, between parental males and mothers. ($F(2,18) = 6.191$, $p = 0.009$, Tukey's HSD $p = 0.008$ Fig. 1D).

Transcriptome

We found that the number of differentially expressed genes (DEGs) in the Attacker versus Parental comparison were higher compared to the Attacker versus Father comparison, yet the pattern of expression between the two comparisons is quite similar. Both datasets had more upregulated genes (Attacker versus Parental: N=518, Attacker versus Father: N=305. $Log_2FC<0.5$ and p-value < 0.05. Fig. 2A and Supplemental Table 2) than down-regulated genes (Attacker versus Parental: N=35, Attacker versus Father: N=47; $Log_2FC<0.5$ and p-value < 0.05. Fig. 2A and Supplemental Table 2), and both comparisons also had more genes with low fold expression changes (Attacker versus Parental: N=760, Attacker versus Father: N=711; $-0.5 < Log_2FC < 0.5$ and p-value < 0.05; Fig. 2A and Supplemental Table 2). Using only upregulated and downregulated genes from each comparison, we investigated KEGG pathways that were overrepresented. First, when compare Attacker and Parental virgin males, we found that many pathways were overrepresented with a significant FDR and strong enrichment ratio. These pathways included extracellular matrix (ECM)-receptor interaction (FDR = 1.43×10^{-13}, Enrichment = 10.26, Fig. 2C and Supplemental Table 3), Wnt signaling pathway (FDR = 2.58×10^{-5}, Enrichment = 4.39, Fig. 2C and Supplemental Table 3), and Hippo signaling

pathway (FDR = $1.12x10^{-5}$, Enrichment = 4.48, Fig. 2C and Supplemental Table 3). When we evaluated the Attacker versus Father, we found the enrichment of the same pathways, such as ECM-receptor interaction (FDR = 0.0182, Enrichment = 5.91, Fig. 2D and Supplemental Table 3), and Hippo signaling pathway (FDR = 0.04998, enrichment = 3.64, Fig. 2D and Supplemental Table 3). To compare these two analyses, we looked for overlapping gene IDs within significantly enriched KEGG pathways. We found that 32 genes were shared between the two analyses using just the overlapping KEGG pathways (Fig. 2E). Within this overlapping dataset, many genes are selectively enriched in certain pathways, such as Wnt signaling. For example, *Tcf7* and *Tcf7L2*, transcription factor genes that are strongly associated as a positive regulator of adult neurogenesis, are both differentially expressed within the Attacker versus Parental (*Tcf7L2*: Log_2FC = 1.06, p-value = 6.81x10-5; *Tcf7*: Log_2FC = 0.77, p-value = 0.0146, Fig. 2F, Supplemental Table 2) and Attacker versus Father comparisons (*Tcf7L2*: Log_2FC = 0.635, p-value = 0.0239; *Tcf7*: Log_2FC = 0.712, p-value = 0.037, Fig. 2F,Supplemental Table 2). We also found the main target for the canonical Wnt and BMP signaling *Lef1* is also upregulated in both conditions (*Lef1*: Log_2FC = 0.563, p-value = 0.011, Attacker versus Parental; *Lef1*: Log_2FC = 0.489, p-value = 0.037, Attacker versus Father, Fig. 2F, Supplemental Table 2). Together these results point to the molecular signaling differences in the aggressive animals compared to the pup care phenotypes.

To further investigate the potential overlap between these two datasets and determine a biological consequence of the behavioral phenotypes, we utilized an unfiltered transcriptome approach to elucidate the similarity between the datasets. As shown in Fig 3A, there is a large degree of concordant transcriptome signals (3,723 genes upregulated, 4,392 genes downregulated, Fig 3A and Supplemental Table 4) with very few discordant signals (1 gene down in Attacker versus Parental and up in Attacker versus Father, Fig 3A and Supplemental Table 4). While these results are promising in that the pattern shows concordant signals, a more nuanced approach is necessary to determine the biological consequence. To achieve this, we constructed a clustered heatmap of all DEGs from any of the four comparisons in the dataset, by plotting their inter-group gene expression z-scores. Of the 8 clusters generated through k-means clustering, cluster 6 exhibits our previously discovered pattern in the transcriptome dataset where the Attacker group had upregulated genes compared to both the Parental and father voles as seen in Figure 3B.

As this cluster 6 gene list was large and produced a big biological process gene ontology enrichment result (240 significantly overrepresented pathways with 213 pathways in the top 10 clusters, Supplemental Table 6), we opted to simplify the interpretation by constructing a clustered similarity network of the annotated biological process gene ontology pathways. Of note, clusters 1 and 2 contained the largest set of both pathways

and genes (Cluster 1 = 81 pathways, 281 genes; Cluster 2 = 88 pathways, 259 genes, Fig 3C, Supplemental Table 7). Cluster 1 is represented by the parent GO term, “Regulation of Response to Stimulus”, while cluster 2 is represented by “Cellular Developmental Process”. Within cluster 1, there are similarities between the previous KEGG pathway analysis and the gene ontology results shown here. We also found several overrepresented pathways such as, Wnt signaling pathway (Enrichment = 3.42, adjusted p-value = 0.00128, Fig 3D and Supplemental Table 7). For cluster 2 of the GO similarity network analysis, nervous system development (Enrichment = 1.851, adjusted p-value = 0.000349, Supplemental Table 7), both positive and negative regulation of cell-differentiation (positive regulation: Enrichment = 2.259, adjusted p-value = 0.0123; negative regulation: Enrichment = 2.85, adjusted p-value = 4.55×10^{-5}, Supplemental Table 7), and many other developmental pathways are present.

In other clusters generated from the heatmap, clusters 2 and 3 show a similar but opposite pattern of gene expression as cluster 6. From cluster 3, there were many gene ontology terms that were overrepresented, such as synaptic signaling (Enrichment = 5.65, adjusted p-value = 7.48×10^{-10}, Supplemental table 5.5), neurogenesis (Enrichment = 2.67, adjusted p-value = 9.02×10^{-5}), and numerous others involving various morphogenesis processes. For example, *Fezf2* is downregulated in aggressive voles compared to parental voles (Log_2FC = -.404, p-value = 0.004, Supplemental Table 2) suggesting that *Fezf2* is implicated in differences seen between these behavioral phenotypes. The clustering analysis reveals the strong upregulation in genes from heatmap cluster 6 is associated with signaling mechanisms and cellular developmental processes. Together, they suggest that unique signaling pathway signatures and an up-regulation of developmental processes are occurring in the dentate gyrus of Attacker voles.

DNA Methylome

We looked for an explanation on why the Attacker and Parental voles had different transcriptomes that may contribute to behavioral phenotypes, so we profiled the DNA methylome for any implications in this relationship. We found differentially methylated CpGs when comparing the reduced representative methylome between the Attacker and Parental voles, and in the Attacker versus Father comparison, with hypermethylated CpGs (13,717 and 16,970 CpGs, respectively, Fig. 4A and Supplemental Table 8) and hypomethylated CpGs (10,495 and 11,579 CpGs, respectively, Fig. 4A and Supplemental Table 8). To learn more about the potential role these differentially methylated CpGs have in this model, we characterized their genomic context using the homer annotation software. We found that most of the differentially methylated CpGs were found in intergenic and intronic genomic regions (about 50% and 32%, for both comparisons, respectively, Fig. 4B and Supplemental Table 8). When looking at the number of

promoters associated differentially methylated CpGs, these account for 5% of the dataset in each comparison (Fig. 4B).

To obtain a better understanding of the functional consequence of these differentially methylated CpGs, we took CpGs that were found within prairie vole gene boundaries and looked for overrepresented pathways in a KEGG pathway analysis. First, we looked at all overlapping pathways between all four comparisons in the study, and found that, Rap1 signaling pathway (Attacker versus Parental - Enrichment = 2.04, FDR = 0.0036; Attacker versus Father – Enrichment = 2.03, FDR = 5.81×10^{-4};Parental versus Father – Enrichment = 2.08, FDR = 7.03×10^{-5}; Father versus Mother – Enrichment = 2.16, FDR = 9.83×10^{-11}; Fig 3C), calcium signaling pathway (Attacker versus Parental - Enrichment = 1.9, FDR = 0.016; Attacker versus Father – Enrichment = 2.19, FDR = 2.43×10^{-4}; Parental versus Father – Enrichment = 1.97, FDR = 0.00105; Father versus Mother – Enrichment = 1.45, FDR = 0.021; Fig 3C), and axon guidance (Attacker versus Parental - Enrichment = 1.99, FDR = 0.012; Attacker versus Father – Enrichment = 2.55 , FDR = 5.70×10^{-6}; Parental versus Father – Enrichment = 2.28, FDR = 2.78×10^{-5}; Father versus Mother – Enrichment = 2.29, FDR = 9.83×10^{-11}; Fig 3C), which have pathway interactions leading to proliferation, were overrepresented for genes with differential methylation. The oxytocin signaling pathway was found enriched for all comparisons except for Attacker versus Parental, which are both sexually-naïve vole groups (Fig. 3C, Supplemental Table 10). For the Attacker versus Parental comparison, the unique KEGG pathways overrepresented by differentially methylated genes have implications in energy homeostasis involving AMPK signaling pathway (Enrichment = 1.95, FDR = 0.0365) and nicotinate and nicotinamide metabolism (Enrichment = 2.88, FDR = 0.046).

Transcriptome-Methylome Correlation

To determine if there is an association between the methylome and transcriptome datasets, we separated the differentially expressed genes into Log_2FC quantiles. We looked at these quartiles or deciles with two different genomic feature scopes, either promoter only, or gene-body methylation, respectively. We then tested each of these quantiles for correlations with the associated DNA methylation data. Across the different groupings of quantiles and DNA methylation, we found that in most cases the extreme quantiles had a significantly negative correlation. In the Attacker versus Parental group and while looking at the entire gene body, the highest Log_2FC quantile was significantly negatively correlated with the DNA methylation signals (cor = -.17, p-value = 0.028, Fig. 5E left and Supplemental Table 11). While looking only at promoters the lowest Log_2FC quantile was significantly negatively correlated (cor = -.511, p-value = 0.0075, Fig. 5E right and Supplemental Table 11). For the Aggressive versus Father comparison, only the promoter-focused analysis was significantly negatively correlated. However, it was for the highest Log_2FC quantile (cor = -0.599, p-value = 0.0105, Fig. 5F right and

Supplemental Table 11). Of the significant Log_2FC quantile correlations with DNA methylation, one gene *Fzd10* was found to have a strong negative correlation of gene expression with DNA methylation in its promoter region (Log_2FC = 1.28, methylation difference = -0.61, Aggressive versus Parental, Supplemental Table 12). Given the relative uniformity of cellular composition in DG, the enrichment of DNA methylation changes at the extreme gene expression deciles may suggest DNA methylation as a canonical gene expression regulatory mechanism exists in the majority of cells in DG to mediate paternal behavioral care.

DISCUSSION

In this study, we examined the dentate gyrus transcriptome and methylome of prairie voles following a pup-exposure paradigm where they displayed either parental behaviors or latency to attack conspecific pups. As noted earlier, there was a separation of behavioral phenotypes among sexually naïve males, where some behaved parentally and the rest displayed aggression. However, it was noted that the virgin parental voles did not act just like the fathers and mothers in their behavior, but they showed a preference for resting. This makes them less parentally focused as pair-bonded individuals, which might be explained by changes associated with sexual experience.

Though the prairie vole genome is still under construction where large portions of chromosomes are isolated on independent genome scaffolds, and the RRBS sequencing only surveys a small portion of the DNA methylome[60], our exploratory approach on the epigenetic underpinnings of paternal care appears to be promising. We found that spontaneously aggressive male prairie voles exhibited both transcriptome and methylome changes compared against spontaneously parental male voles. Though variable gene transcription patterns have been reported in other brain regions[61], in our study of the DG, we found many genes to be differentially expressed between the attacker virgin voles and the parental virgin males and fathers. Overlapping genes from these datasets were involved in several meaningful pathways including extracellular matrix-receptor interactions and components of the Wnt and Hippo signaling pathways. These include the up-regulated genes *Tcf7* and *Tcf7l2*, which are main effectors of the Wnt signaling pathway and these transcription factors have been reported by other studies to be involved in adult hippocampal neurogenesis[62]. With increased canonical Wnt signaling, beta-catenin stabilizes and can be transported to the nucleus where it complexes with TCF/LEF transcription factors to positively regulate neurogenesis[63].

In order to further profile the degree of similarity between the attacker and parental voles in a whole transcriptome manner, we utilized a modification of a hypergeometric overlaps test named RRHO. In this analysis we found that there was a large portion of the

transcriptome showing the same pattern as the baseline differential expression analysis. To quantify the transcriptome patterns within the four experimental groups we clustered the dataset using k-means clustering on 8 clusters. There was one gene cluster that was predominantly upregulated and once again matched our previous transcriptome pattern, where the Attacker group had increased expression compared to both parentally behaving groups, Parental and Fathers. Within these clusters neurodevelopmental factors were found to be over-represented. *Fezf2* is a transcriptional repressor whose activation limits the diversity differentiating neuron populations[64, 65] and mark glutamatergic cell fate delineations by activating gene programs that include *Vglut1* and suppressing the expression of *Gad1*[66]. Furthermore, many members of the Wnt signaling pathway were over-represented, including multiple frizzled receptors, *Fzd2*, *Fzd7*, and *Fzd10*, downstream effectors of Wnt signaling, *Tcf7, Tcf7l2, Lef1*, and two different Wnt-receptors, *Wnt6* and *Wnt9a*. Both *Tcf7l2* and *Lef1* have previously been shown that they are necessary for proper neurogenesis function and hippocampal formation[67]. These findings further support that neurodevelopmental processes are disrupted in DG between aggressive and alloparental voles.

Looking at the aforementioned Wnt-signaling effectors, there were many differentially methylated CpG sites associated with *Tcf7l2*, *Tcf7*, and *Lef1*. The promoter of *Lef1* has decreased methylation in the attackers which coincides with increased gene expression, a canonical mechanism of DNA methylation's effect on gene transcription. In the case of increased *Lef1* expression using a lentiviral overexpression construct, there was an increase in the number of newly born neurons, highlighted by *TuJ1* expression[68]. Our results are supported by a few earlier experiments, where attacker male virgin voles show increased newly born immature neurons in the dentate gyrus[24], yet the survival rate of these adult-generated neurons is reduced[25]. There is evidence that increased survivability of newly born neurons are predictive of stress susceptibility in rodents[69]. However, it has been shown that fatherhood in prairie voles is associated with decreased proliferating neuron survival in the dentate gyrus[25], that may affect neural plasticity. Therefore, pair-bonding experience or fatherhood provides a resilience compensation of parental care through neuroplasticity changes in prairie vole fathers. In sum, these reports alongside our study support that the paternal care may be mediated through DNA methylation dependent key signaling pathway changes (e.g. Wnt-signaling) that contribute to DG neural-developmental plasticity alterations.

Data Availability

Next generation sequencing data is being deposited in GEO databases and will be available upon publication.

Author contributions

The study was conceived by Z.W., N.J.W., and J.F. Y.L. performed animal experiments. Y.L., G.J.K., N.J.W., performed behavioral data analysis. J.M.C. prepared sequencing libraries. N.J.W. performed bioinformatic sequencing data analysis with feedback from other authors. N.J.W., Z.W., and J.F. prepared the manuscript. All authors reviewed and approved the manuscript.

Acknowledgements

This work was supported by National Institutes of Health grants (R01MH108527 and R21MH111998 to ZW, DP1DA046587 and R01DA046720 to J.F.). J.M.C was a recipient of FSU Neuroscience Fellowship and FSU Legacy Fellowship.

Figures

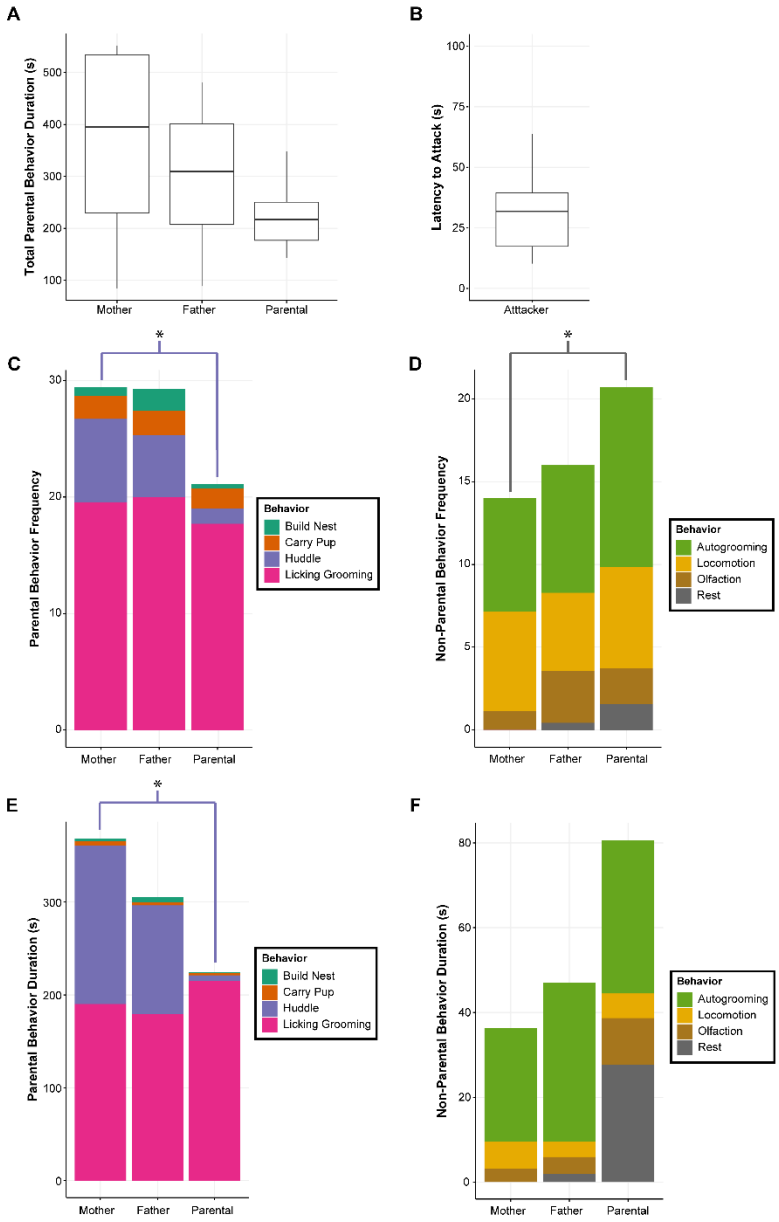


Fig. 1 - Paternal behavioral dichotomy in virgin male prairie voles. A – Boxplot summarizing the duration of total parental behavior in seconds for each of experimental groups. Whiskers of boxplot represent the ranges of each group, the upper and lower bounds of the boxes represent inter-quartile range (IQR) from quartile 1 and quartile 3, and the horizontal line represents the median. **B** – Boxplot showing the latency to attack in seconds by the attacker virgin males. The upper and lower bounds of the boxes represent IQR from quartile 1 and quartile 3, and the horizontal line represents the median of the dataset. **C** – Stacked barplots represent the total summed average of each parental behavior in frequency. Significance bars are color coded according to the legend. Data

was analyzed with a one-way ANOVA with TukeyHSD post-hoc correction. (*) - p-value < 0.05. **D** – Stacked barplots represent the total summed average of each non-parental behavior in frequency. Significance bars are color coded according to the legend. Data was analyzed with a one-way ANOVA with TukeyHSD post-hoc correction. (*) - p-value < 0.05. **E** – Stacked barplots represent the total summed average of each parental behavior in duration. Significance bars are color coded according to the legend. Data was analyzed with a one-way ANOVA with TukeyHSD post-hoc correction. (*) - p-value < 0.05. **F** – Stacked barplots represent the total summed average of each non-parental behavior in duration.

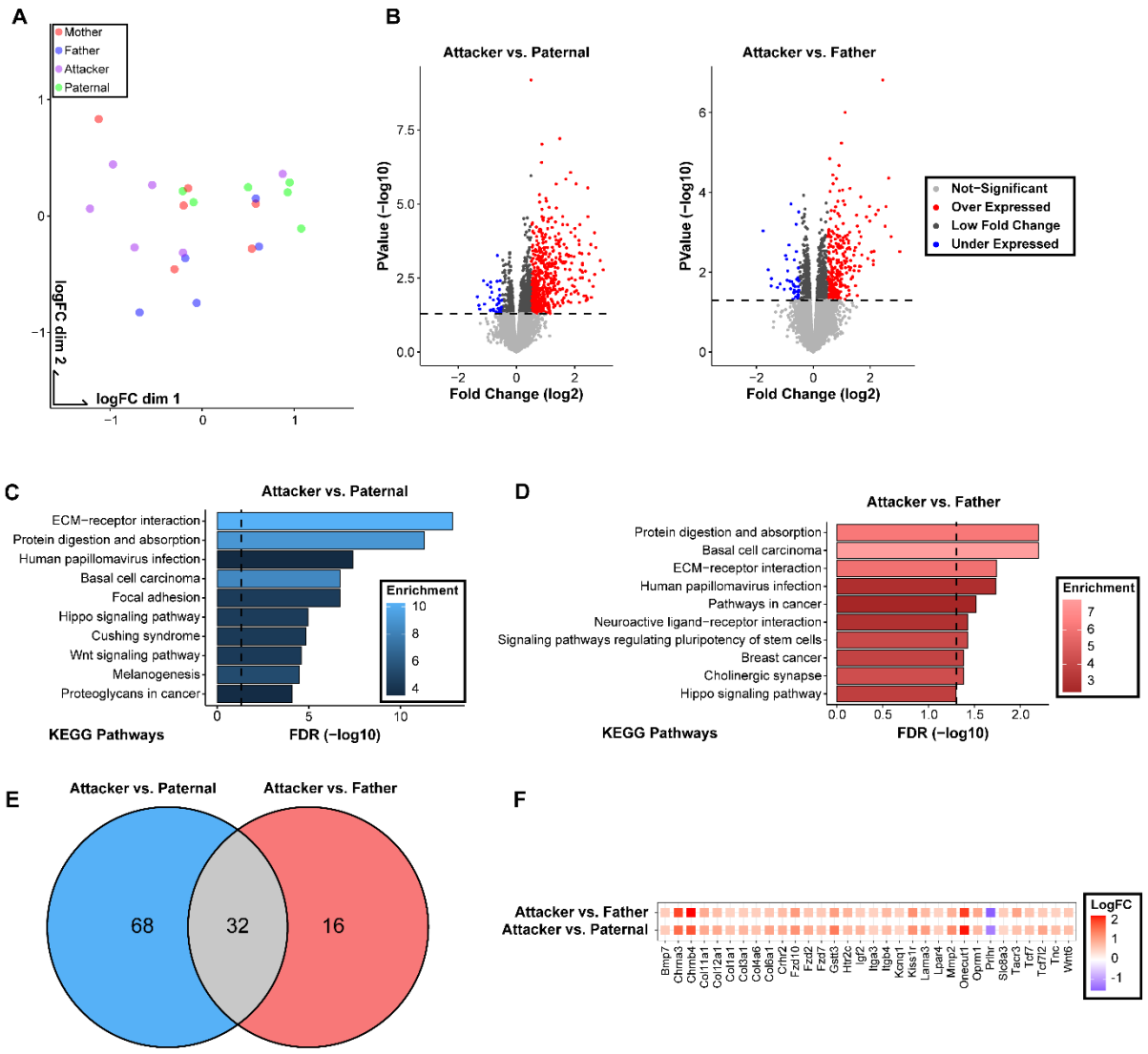
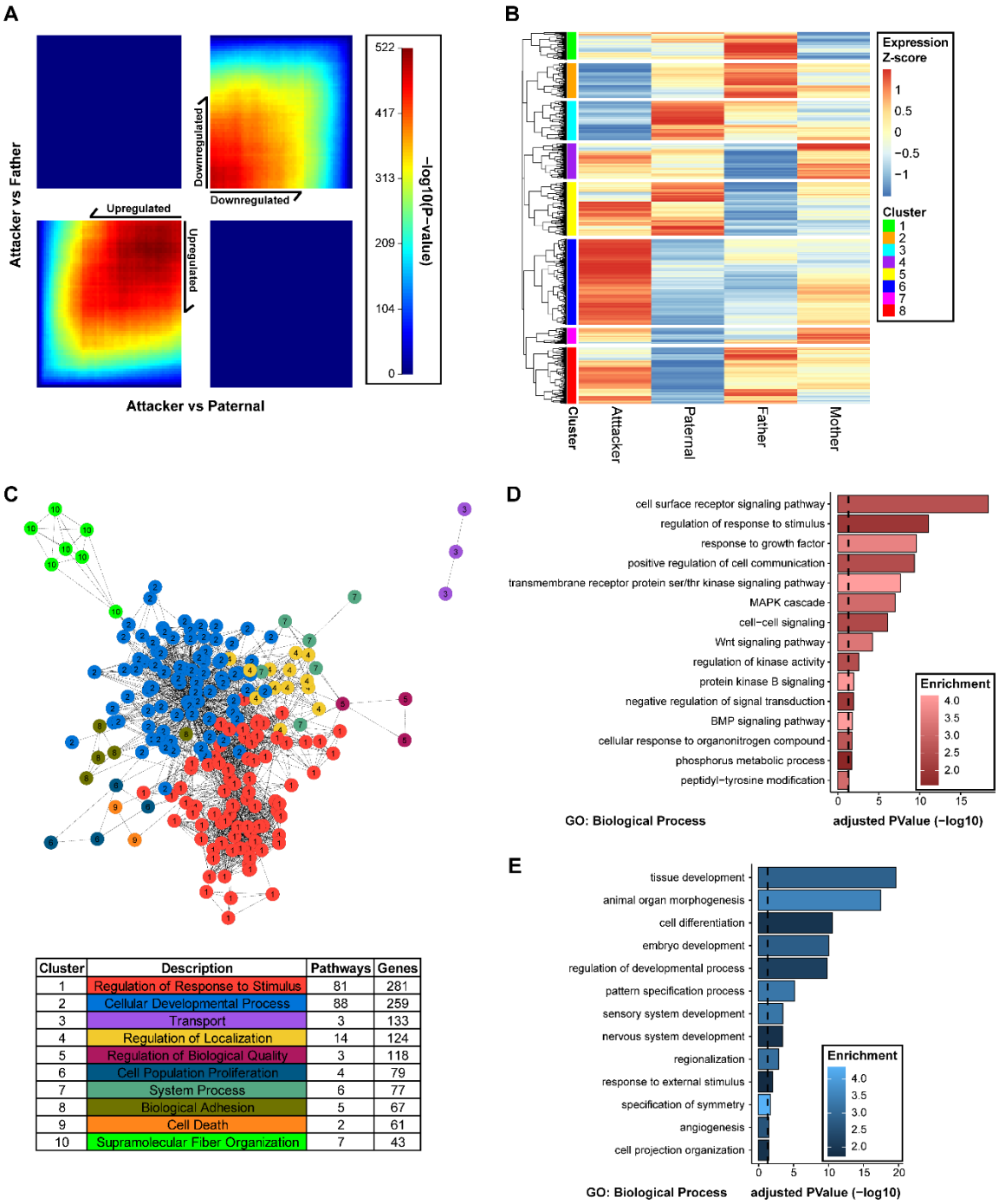


Fig. 2 - Overlapping pathways in attackers compared to both parental and father voles. A – Plot showing the first two principal components that explain variability within the RNA-seq dataset. Points are color coded according to experimental group and labels were dodged to increase clarity. **B** – Volcano plot visualizing differentially expressed genes, in Log_2 transformed Fold Change (Log_2FC) on the x-axis, with $-Log_{10}$ adjusted p-values on the y-axis. Genes considered to be overexpressed have greater than 0.5 Log_2FC, while underexpressed genes have less than -0.5 in Log_2FC. The left plot shows the genes from the Attacker versus Paternal comparison, and the right plot shows the differentially expressed genes between the Attacker versus Fathers comparison. Genes labeled in dark gray correspond to significantly expressed genes with Log_2FC values between the previous two thresholds. The dashed line represents $-Log_{10}(0.05)$ to show the boundary of significance with data points below the line representing non-significant data and colored in light gray. **C** & **D**– Bar plot representing significantly enriched KEGG

Pathways resulting from (C) the Attacker versus Paternal comparison and (D) shows pathways from the Attacker versus Fathers comparison. The overrepresented pathways were modeled with the hypergeometric test using the whole annotated transcriptome as background. The color gradient for each bar represented the enrichment ratio (observed/expected) with the degree of blue color showing increasing enrichment. The x-axis represents -Log_{10} transformed FDR values. **E** – Venn diagram showing the overlapping relationship between the genes of enriched pathways found in the KEGG pathway overrepresentation test. Blue represents the genes unique to the Attacker vs Paternal comparison, while the red represents the unique Attacker vs Fathers comparison, with the grey showing the union of the two datasets. **F** – This dotplot represents the Log_2FC values from the shared genes from the venn diagram in panel E. The color assigned to each gene represents the Log_2FC from the associated RNAseq comparison.



Cluster	Description	Pathways	Genes
1	Regulation of Response to Stimulus	81	281
2	Cellular Developmental Process	88	259
3	Transport	3	133
4	Regulation of Localization	14	124
5	Regulation of Biological Quality	3	118
6	Cell Population Proliferation	4	79
7	System Process	6	77
8	Biological Adhesion	5	67
9	Cell Death	2	61
10	Supramolecular Fiber Organization	7	43

Fig. 3 - Overlapping transcriptome profiles of attackers compared to parental and father voles. A – This RRHO map compares all expressed genes between the Attacker versus Paternal and Attacker versus Fathers comparisons. The overlap is made in a whole-transcriptome and threshold-free manner, where each pixel represents

overlapping genes. The color of each pixel represents the Benjamini-Yekutieli adjusted -Log_{10}(p-value) of a hypergeometric test, with warmer colors representing more significant results indicating either discordant or concordant expression profiles depending on the quadrant. **B** – Clustered Heatmap generated by the “pheatmap” software. Rows represent z-scores of normalized gene expression among the four groups in the analysis. Clustering was performed using k means clustering to generate 8 clusters based on the normalized z-score of genes. **C** – Cluster 2 significantly over-represented biological process gene ontologies were clustered with GOMCL to reduce redundancy and simplify the pathway analysis. The legend shows the simplified pathway labels with the number of pathways within this cluster, and the number of unique genes within these simplified pathways to the right. **D** & **E** – Representative GO Biological Process Pathways shown from GOMCL cluster 1 (D) and cluster 2 (E). The overrepresented pathways were modeled using the hypergeometric test with the whole annotated transcriptome as background. The color gradient for each bar represented the enrichment ratio (observed/expected). The x-axis represents -Log_{10} transformed FDR values.

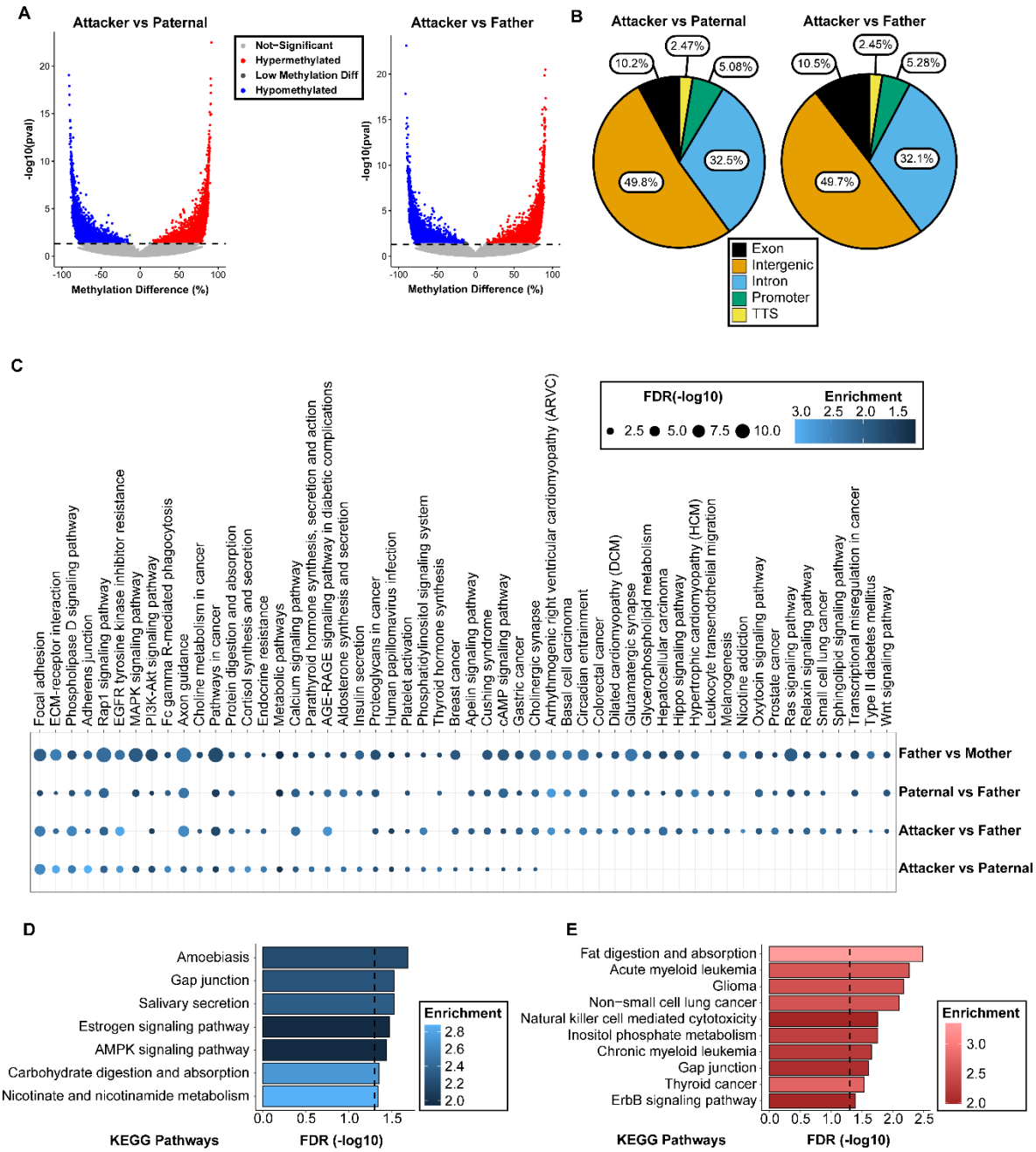


Fig. 4 - Differential methylation analysis. A – Volcano plot visualizing differentially methylated CpG sites. Methylation % is represented on the x-axis, while the -Log_{10}(p-value) is represented on the y-axis. Hypermethylated CpGs were characterized by an increase in 15% of methylated CpG calls compared to control, while hypomethylated CpGs were a decrease in 15% of methylated CpG calls compared to control. And significantly methylated CpGs between these cutoffs was considered low methylation difference. The horizontal dashed line corresponds to -Log_{10}(0.05) to represent a threshold for significance with anything below being considered non-significant. The left

plot represents the comparison Attacker versus Paternal, while the right represents the Attacker versus Fathers comparison. **B** – These pie charts represent the genomic feature distribution of the differentially methylated CpGs from the comparisons listed above each plot. Labels within each section represent the % of differentially methylated CpGs that fall within these features. **C** – This plot represents the overrepresented KEGG pathways of genes that contain differentially methylated CpG sites within their gene boundary (-2000 bp upstream of transcription start site -< 500 bp downstream for the transcription termination site). This plot shows the pathways that are shared between any comparison from this dataset. Size is reflective of the enrichment ratio from the hypergeometric distribution, while the color represents the significance values. **D** & **E** – This plot represents the overrepresented KEGG pathways of genes that are not shared between any two comparisons in the analysis. The genes used for this test contain differentially methylated CpG sites within their gene boundary (-2000 bp upstream of transcription start site -< 500 bp downstream for the transcription termination site) of Attacker versus Paternal (D) and Attacker versus Fathers (E). The dashed line represents $-Log_{10}(0.05)$ as a threshold of significant results. Y-axis represents FDR, while color of each bar represents the degree of enrichment from the hypergeometric overlaps test.

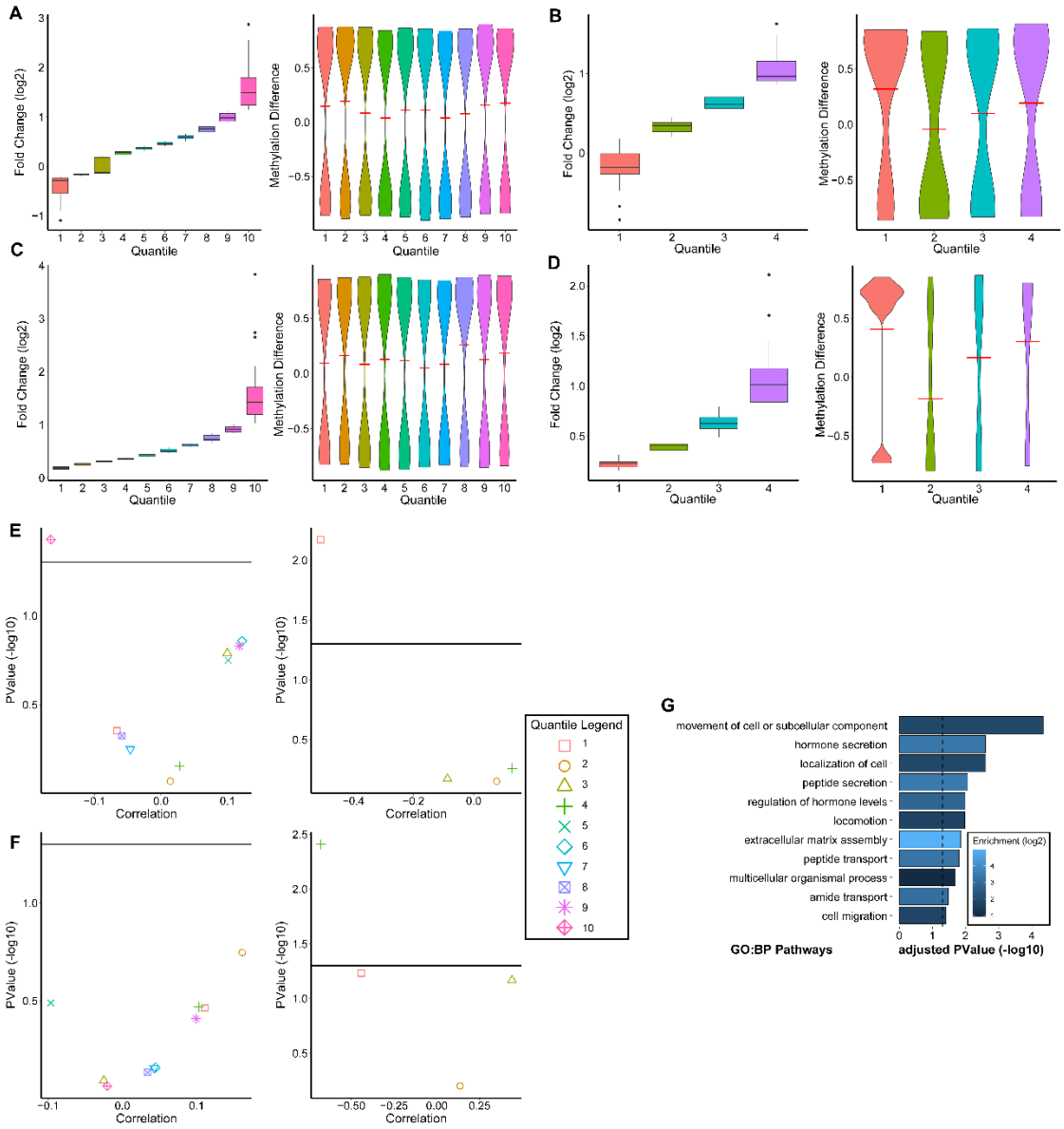


Fig. 5 - Correlations between gene expression and DNA methylation. A – Data from Attacker versus Paternal comparison – (Left) – Genes that have both differential expression and differentially methylated CpGs within their gene boundary were segmented into 10 deciles. The left plot represents the Fold Changes of the differentially expressed genes after Log_2 transformation. The horizontal line represents the median of each group. (Right) – From the 10 deciles created on fold change, methylation of each quantile is represented as a violin plot. The width of each plot indicates the number of datapoints in that portion. The red line represents the mean of the methylation differences from the points in each violin plot. **B** – Data from Attacker versus Paternal comparison, and data is represented in the same manner as (A), but for genes with differentially methylated CpGs in their promoter. The data was separated into 4 quartiles instead of 10 deciles. **C** – Data from Attacker versus Fathers comparison and using genes with

differentially methylated CpGs within their gene boundary, where the data was segmented into 10 deciles in the same manner as in (A). **D** – Data from Attacker versus Fathers comparison, and data is represented the same manner as (B), using only genes with differentially methylated CpGs in their promoter. **E** – Represents a scatter plot of spearman correlation coefficients and the associated p-values from the correlation analysis of quantiles from the Attacker versus Paternal comparison. The horizontal line represents $-Log_{10}(0.05)$ as a threshold of significance. (Left) – showing the 10 deciles from panel (A). (Right) – showing the 4 quartiles from panel (B). **F** – Represents a similar scatterplot as shown in (E), but data comes from the Attacker versus Fathers comparison, where data for the left and right plots correspond to panel (C) and (D) respectively. **G** – Genes from the significantly correlated quantiles from panels (E) and (F) were combined and tested for overrepresented categories from the biological process gene ontology annotations. The dashed line represents $-Log_{10}(0.05)$ as a threshold of significance, color of each bar represents Log_2 transformed enrichment values from the hypergeometric overlaps test.

List of Supplementary Materials

Supplemental Table 1: Behavioral Analysis

Supplemental Table 2: Differentially Expressed Genes

Supplemental Table 3: Overrepresented KEGG Pathways from DEGs

Supplemental Table 4: RRHO Gene List

Supplemental Table 5: Normalized Gene Counts for Clustering

Supplemental Table 6: Cluster-based Gene Ontology Pathways

Supplemental Table 7: GOMCL Ontology Network Analysis

Supplemental Table 8: Differentially Methylated Loci

Supplemental Table 9: Differentially Methylated Regions

Supplemental Table 10: Overrepresented KEGG Pathways from DMLs

Supplemental Table 11: Quantile Correlation Results

Supplemental Table 12: Gene Expression and Methylation Tables from Correlation

Supplemental Table 13: Overrepresented Biological Process Pathways from all Significantly Correlated Quantiles

Literature Cited

1. Ray, S., et al., *An Examination of Dynamic Gene Expression Changes in the Mouse Brain During Pregnancy and the Postpartum Period.* G3 (Bethesda), 2015. **6**(1): p. 221-33.
2. Perea-Rodriguez, J.P., et al., *Effects of reproductive experience on central expression of progesterone, oestrogen alpha, oxytocin and vasopressin receptor mRNA in male California mice (Peromyscus californicus).* J Neuroendocrinol, 2015. **27**(4): p. 245-52.
3. Kleiman, D.G., *Monogamy in mammals.* Q Rev Biol, 1977. **52**(1): p. 39-69.
4. Young, L.J. and Z. Wang, *The neurobiology of pair bonding.* Nat Neurosci, 2004. **7**(10): p. 1048-54.
5. Young, K.A., et al., *The neurobiology of pair bonding: insights from a socially monogamous rodent.* Front Neuroendocrinol, 2011. **32**(1): p. 53-69.
6. Thomas, J.A. and E.C. Birney, *Parental Care and Mating System of the Prairie Vole, Microtus ochrogaster.* Behavioral Ecology and Sociobiology, 1979. **5**(2): p. 171-186.
7. Wang, Z. and M.A. Novak, *Alloparental care and the influence of father presence on juvenile prairie voles, Microtus ochrogaster.* Animal Behaviour, 1994. **47**(2): p. 281-288.
8. Lonstein, J.S. and G.J. De Vries, *Comparison of the Parental Behavior of Pair-Bonded Female and Male Prairie Voles (Microtus ochrogaster).* Physiology & Behavior, 1999. **66**(1): p. 33-40.
9. Lonstein, J.S. and G.J. De Vries, *Sex differences in the parental behavior of rodents.* Neuroscience & Biobehavioral Reviews, 2000. **24**(6): p. 669-686.
10. Liu, Y., et al., *The ventromedial hypothalamic circuitry and male alloparental behaviour in a socially monogamous rodent species.* Eur J Neurosci, 2019. **50**(11): p. 3689-3701.
11. Kenkel, W.M., A.M. Perkeybile, and C.S. Carter, *The neurobiological causes and effects of alloparenting.* Dev Neurobiol, 2017. **77**(2): p. 214-232.
12. Stolzenberg, D.S., J.S. Stevens, and E.F. Rissman, *Experience-facilitated improvements in pup retrieval; evidence for an epigenetic effect.* Horm Behav, 2012. **62**(2): p. 128-35.
13. Zocher, S., et al., *Environmental enrichment preserves a young DNA methylation landscape in the aged mouse hippocampus.* Nat Commun, 2021. **12**(1): p. 3892.
14. Mayer, H.S., et al., *Histone deacetylase inhibitor treatment promotes spontaneous caregiving behaviour in non-aggressive virgin male mice.* J Neuroendocrinol, 2019. **31**(9): p. e12734.
15. Xu, G.L. and M. Bochtler, *Reversal of nucleobase methylation by dioxygenases.* Nat Chem Biol, 2020. **16**(11): p. 1160-1169.
16. Wu, H. and Y. Zhang, *Reversing DNA methylation: mechanisms, genomics, and biological functions.* Cell, 2014. **156**(1-2): p. 45-68.
17. Bestor, T.H., *The DNA methyltransferases of mammals.* Hum Mol Genet, 2000. **9**(16): p. 2395-402.

18. Brown, A.N. and J. Feng, *Drug Addiction and DNA Modifications.* Adv Exp Med Biol, 2017. **978**: p. 105-125.
19. Smith, Z.D. and A. Meissner, *DNA methylation: roles in mammalian development.* Nat Rev Genet, 2013. **14**(3): p. 204-20.
20. Day, J.J. and J.D. Sweatt, *Epigenetic mechanisms in cognition.* Neuron, 2011. **70**(5): p. 813-29.
21. Turecki, G. and M.J. Meaney, *Effects of the Social Environment and Stress on Glucocorticoid Receptor Gene Methylation: A Systematic Review.* Biol Psychiatry, 2016. **79**(2): p. 87-96.
22. Fowler, C.D., et al., *The effects of social environment on adult neurogenesis in the female prairie vole.* J Neurobiol, 2002. **51**(2): p. 115-28.
23. Castro, A.E., et al., *Effects of Mating and Social Exposure on Cell Proliferation in the Adult Male Prairie Vole (Microtus ochrogaster).* Neural Plast, 2020. **2020**: p. 8869669.
24. Ruscio, M.G., et al., *Pup exposure elicits hippocampal cell proliferation in the prairie vole.* Behav Brain Res, 2008. **187**(1): p. 9-16.
25. Lieberwirth, C., et al., *Fatherhood reduces the survival of adult-generated cells and affects various types of behavior in the prairie vole (Microtus ochrogaster).* Eur J Neurosci, 2013. **38**(9): p. 3345-55.
26. Liu, Y., et al., *Nucleus accumbens dopamine mediates amphetamine-induced impairment of social bonding in a monogamous rodent species.* Proc Natl Acad Sci U S A, 2010. **107**(3): p. 1217-22.
27. Young, K.A., et al., *Oxytocin reverses amphetamine-induced deficits in social bonding: evidence for an interaction with nucleus accumbens dopamine.* J Neurosci, 2014. **34**(25): p. 8499-506.
28. Liu, Y., et al., *Amphetamine exposure alters behaviors, and neuronal and neurochemical activation in the brain of female prairie voles.* Neuroscience, 2022. **498**: p. 73-84.
29. Inoue, K., et al., *Oxytocin receptors are widely distributed in the prairie vole (Microtus ochrogaster) brain: Relation to social behavior, genetic polymorphisms, and the dopamine system.* J Comp Neurol, 2022.
30. Danoff, J.S., et al., *Genetic, epigenetic, and environmental factors controlling oxytocin receptor gene expression.* Clin Epigenetics, 2021. **13**(1): p. 23.
31. Ophir, A.G., et al., *Oxytocin receptor density is associated with male mating tactics and social monogamy.* Horm Behav, 2012. **61**(3): p. 445-53.
32. Lei, K., et al., *Effects of pair bonding on parental behavior and dopamine activity in the nucleus accumbens in male prairie voles.* Eur J Neurosci, 2017. **46**(7): p. 2276-2284.
33. Tabbaa, M., et al., *Paternal deprivation affects social behaviors and neurochemical systems in the offspring of socially monogamous prairie voles.* Neuroscience, 2017. **343**: p. 284-297.
34. Council, N.R., *Guide for the Care and Use of Laboratory Animals, 8th edition*. 2011, Washington, DC: The National Academies Press.
35. Blumstein, D.T., J.C. Daniel, and C.S. Evans. *JWatcher*. 2000; Available from: https://www.jwatcher.ucla.edu/.

36. Lonstein, J.S., B.D. Rood, and G.J. De Vries, *Parental responsiveness is feminized after neonatal castration in virgin male prairie voles, but is not masculinized by perinatal testosterone in virgin females.* Horm Behav, 2002. **41**(1): p. 80-7.
37. Wickham, H., *ggplot2: Elegant Graphics for Data Analysis*. 2016: Springer-Verlag New York.
38. Paxinos, G. and C. Watson, *The Rat Brain in Stereotaxic Coordinates*. 7 ed. 2013: Academic Press.
39. Meissner, A., et al., *Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis.* Nucleic Acids Res, 2005. **33**(18): p. 5868-77.
40. Andrews, S. *FastQC: a quality control tool for high throughput sequence data*. 2010; Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.
41. Bushnell, B. *BBMap*. 2014; Available from: sourceforge.net/projects/bbmap/.
42. Krueger, F. *Trim Galore*. 2012; Available from: https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.
43. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.
44. Cunningham, F., et al., *Ensembl 2022.* Nucleic Acids Res, 2022. **50**(D1): p. D988-D995.
45. Liao, Y., G.K. Smyth, and W. Shi, *featureCounts: an efficient general purpose program for assigning sequence reads to genomic features.* Bioinformatics, 2014. **30**(7): p. 923-30.
46. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.* Bioinformatics, 2010. **26**(1): p. 139-40.
47. Kanehisa, M. and S. Goto, *KEGG: kyoto encyclopedia of genes and genomes.* Nucleic Acids Res, 2000. **28**(1): p. 27-30.
48. Liao, Y., et al., *WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs.* Nucleic Acids Res, 2019. **47**(W1): p. W199-W205.
49. Raudvere, U., et al., *g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update).* Nucleic Acids Res, 2019. **47**(W1): p. W191-W198.
50. Plaisier, S.B., et al., *Rank-rank hypergeometric overlap: identification of statistically significant overlap between gene-expression signatures.* Nucleic Acids Res, 2010. **38**(17): p. e169.
51. Cahill, K.M., et al., *Improved identification of concordant and discordant gene expression signatures using an updated rank-rank hypergeometric overlap approach.* Sci Rep, 2018. **8**(1): p. 9588.
52. Seney, M.L., et al., *Opposite Molecular Signatures of Depression in Men and Women.* Biol Psychiatry, 2018. **84**(1): p. 18-27.
53. Benjamini, Y. and D. Yekutieli, *The control of the false discovery rate in multiple testing under dependency.* The Annals of Statistics, 2001. **29**(4).
54. Kolde, R., *Pheatmap: pretty heatmaps.* R package version, 2012. **1**(2): p. 726.

55. Wang, G., D.H. Oh, and M. Dassanayake, *GOMCL: a toolkit to cluster, evaluate, and extract non-redundant associations of Gene Ontology-based functions.* BMC Bioinformatics, 2020. **21**(1): p. 139.
56. Shannon, P., et al., *Cytoscape: a software environment for integrated models of biomolecular interaction networks.* Genome Res, 2003. **13**(11): p. 2498-504.
57. Krueger, F. and S.R. Andrews, *Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications.* Bioinformatics, 2011. **27**(11): p. 1571-2.
58. Park, Y. and H. Wu, *Differential methylation analysis for BS-seq data under general experimental design.* Bioinformatics, 2016. **32**(10): p. 1446-53.
59. Heinz, S., et al., *Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities.* Mol Cell, 2010. **38**(4): p. 576-89.
60. Veillard, A.-C., et al., *Diagenode® Premium RRBS technology: cost-effective DNA methylation mapping with superior coverage.* Nature Methods, 2016. **13**(2): p. i-ii.
61. Duclot, F., et al., *Transcriptomic Regulations Underlying Pair-bond Formation and Maintenance in the Socially Monogamous Male and Female Prairie Vole.* Biol Psychiatry, 2022. **91**(1): p. 141-151.
62. Lie, D.C., et al., *Wnt signalling regulates adult hippocampal neurogenesis.* Nature, 2005. **437**(7063): p. 1370-5.
63. Katoh, M. and M. Katoh, *WNT signaling pathway and stem cell signaling network.* Clin Cancer Res, 2007. **13**(14): p. 4042-5.
64. Tsyporin, J., et al., *Transcriptional repression by FEZF2 restricts alternative identities of cortical projection neurons.* Cell Rep, 2021. **35**(12): p. 109269.
65. Guo, H., et al., *Single-cell methylome landscapes of mouse embryonic stem cells and early embryos analyzed using reduced representation bisulfite sequencing.* Genome Res, 2013. **23**(12): p. 2126-35.
66. Lodato, S., et al., *Gene co-regulation by Fezf2 selects neurotransmitter identity and connectivity of corticospinal neurons.* Nat Neurosci, 2014. **17**(8): p. 1046-54.
67. Chodelkova, O., et al., *Tcf7L2 is essential for neurogenesis in the developing mouse neocortex.* Neural Dev, 2018. **13**(1): p. 8.
68. Armenteros, T., et al., *BMP and WNT signalling cooperate through LEF1 in the neuronal specification of adult hippocampal neural stem and progenitor cells.* Sci Rep, 2018. **8**(1): p. 9241.
69. Lagace, D.C., et al., *Adult hippocampal neurogenesis is functionally important for stress-induced social avoidance.* Proc Natl Acad Sci U S A, 2010. **107**(9): p. 4436-41.