Title: Fetal programming by the parental microbiome of offspring behavior, and DNA methylation and gene expression within the hippocampus

Authors:

Kevin L. Gustafson, Department of Veterinary Pathobiology, College of Veterinary Medicine, University of Missouri, Columbia, MO, 65201, USA, klgbkt@umsystem.edu

Susheel Bhanu Busi, UK Centre for Ecology and Hydrology, Wallingford, Oxfordshire, OX10 8BB, UK, SusBus@ceh.ac.uk

Zachary L. McAdams, Department of Veterinary Pathobiology, College of Veterinary Medicine, University of Missouri, Columbia, MO, 65201, USA, zlmg2b@missouri.edu

Rachael E. McCorkle, College of Veterinary Medicine, University of Missouri, Columbia, MO, 65211, USA, remccorkle@gmail.com

Pavlo Khodakivskyi, Department of Chemistry, College of Arts and Science, University of Missouri, Columbia, MO, 65211, USA, p.khodakivsky@missouri.edu

Nathan J. Bivens, University of Missouri Genomics Technology Core, University of Missouri, Columbia, MO, 65211, USA, bivensn@missouri.edu

Daniel J. Davis, Department of Veterinary Pathobiology, College of Veterinary Medicine, University of Missouri, Columbia, MO, 65201, USA, davisdaniel@missouri.edu

Murugesan Raju, University of Missouri Bioinformatics and Analytics Core, University of Missouri, Columbia, MO, 65211, USA, rajum@missouri.edu

Lyndon M. Coghill, University of Missouri Bioinformatics and Analytics Core, University of Missouri, Columbia, MO, 65211, USA, lcoghill@missouri.edu

Elena A. Goun, Department of Chemistry, College of Arts and Science, University of Missouri, Columbia, MO, 65211, USA, elena.goun@missouri.edu

James Amos-Landgraf, Department of Veterinary Pathobiology, College of Veterinary Medicine, University of Missouri, Columbia, MO, 65201, USA, amoslandgrafj@missouri.edu

Craig L. Franklin, Department of Veterinary Pathobiology, College of Veterinary Medicine, University of Missouri, Columbia, MO, 65201, USA, franklinc@missouri.edu

Paul Wilmes, Department of Life Sciences and Medicine, Faculty of Science, Technology and Medicine, University of Luxembourg, L-4362 Esch-sur-Alzette, Luxembourg, paul.wilmes@uni.lu

Rene Cortese, Department of Child Health & Obstetrics, Gynecology, and Women's Health, School of Medicine, University of Missouri, Columbia, MO, 65212, USA, rene.cortese@health.missouri.edu

Aaron C. Ericsson, Department of Veterinary Pathobiology, College of Veterinary Medicine, University of Missouri, Columbia, MO, 65201, USA, ericssona@missouri.edu

Abstract

Background

The microorganisms colonizing the gastrointestinal tract of animals, collectively referred to as the gut microbiome, affect numerous host behaviors dependent on the central nervous system (CNS). Studies comparing germ-free mice to normally colonized mice have demonstrated influences of the microbiome on anxiety-related behaviors, voluntary activity, and gene expression in the CNS. Additionally, there is epidemiologic evidence supporting an intergenerational influence of the maternal microbiome on neurodevelopment of offspring and behavior later in life. There is limited experimental evidence however directly linking the maternal microbiome to long-term neurodevelopmental outcomes, or knowledge regarding mechanisms responsible for such effects.

Results

Here we show that that the maternal microbiome has a dominant influence on several offspring phenotypes including anxiety-related behavior, voluntary activity, and body weight. Adverse outcomes in offspring were associated with features of the maternal microbiome including bile salt hydrolase activity gene expression (*bsh*), abundance of certain bile acids, and hepatic expression of *Slc10a1*. In cross-foster experiments, offspring resembled their birth dam phenotypically, despite faithful colonization in the postnatal period with the surrogate dam microbiome. Genome-wide methylation analysis of hippocampal DNA identified microbiome-associated differences in methylation of 196 loci in total, 176 of which show conserved profiles between mother and offspring. Further, single-cell transcriptional analysis revealed accompanying differences in expression of several differentially methylated genes within certain hippocampal cell clusters, and vascular expression of genes associated with bile acid transport.

Inferred cell-to-cell communication in the hippocampus based on coordinated ligand-receptor expression revealed differences in expression of neuropeptides associated with satiety.

Conclusions

Collectively, these data provide proof-of-principle that the maternal gut microbiome has a dominant influence on the neurodevelopment underlying certain offspring behaviors and activities, and selectively affects genome methylation and gene expression in the offspring CNS in conjunction with that neurodevelopment.

Keywords: Fetal programming, Maternal microbiome, Gut-brain axis, DNA methylation, Hippocampus, Gene expression, Bile acids, Neurodevelopment

Background

Neurodevelopmental and behavioral disorders are a growing concern worldwide. According to the World Health Organization, nearly one billion people worldwide live with a mental disorder[1]. Behavioral disorders are commonly associated with social impairments, decreased productivity, financial losses, and general maladjustment[2]. Previous studies have found that such disorders contribute substantially to global nonfatal health loss[1]. Anxiety disorders (AD) are increasing in prevalence, affecting close to 1 in 10 children and adolescents between the ages of 3 and 17[3]. Like most mood disorders, AD are multifactorial and often result from a combination of genetic, environmental, and experiential factors. Similarly, one in five children in the U.S. are obese or overweight[4], reflecting the combined influence of western diet, increasingly sedentary lifestyles, and other factors. Moreover, AD and obesity/overweight (OO) are reciprocal risk factors, frequently occurring as co-morbidities[5-9].

A growing body of research has linked the gut microbiome (GM) to neurodevelopment and behavior[10-15], and growth rate or weight gain[16-19]. Work with germ-free mice shows the importance of the GM in normative behavior and metabolism[20-22], and transfer of an anxiety-related phenotype or increased energy harvest via fecal microbiome transfer indicates that certain features within naturally occurring microbiomes influence these phenotypes[23-25]. There is also evidence that indicates that the effects of the GM can go beyond simply influencing the host. Research in rodents has revealed that the GM of a pregnant dam can influence the fetus and phenotype of the offspring following birth. Eloquent studies in mice have shown that effects of diet and exercise on the maternal GM can be transferred to the offspring, relieving negative metabolic phenotypes [26, 27]. There are also developmental components to both AD and OO,

raising the question of how the maternal microbiome during pregnancy affects fetal development and subsequent behavior and energy metabolism in the adult offspring. The maternal gut microbiome during pregnancy produces metabolites which reach peripheral circulation and the fetal CNS[28], and maternal proteins and peptides produced by enteroendocrine cells in response to the microbiome likely also cross the placenta and reach fetal circulation[29-31]. Disruption of the maternal GM can affect these processes as demonstrated by increased anxiety in the offspring of mice with antibiotic- or diet-induced dysbiosis[32-35].

There are still major gaps in our knowledge however regarding the mechanisms through which the maternal microbiome during pregnancy programs long-lasting changes in offspring behavior and metabolism. These intergenerational effects suggest fetal imprinting by an unknown mechanism, while differences in anxiety-related behavior (and other complex behaviors) indicate a neurodevelopmental basis. Owing to the genetic, dietary, and environmental heterogeneity, analysis of these processes in a human population requires very large sample sizes and long-term tracking of mother-child pairs. To circumvent these factors, here we use two groups of genotype-, age-, and sex-matched outbred CD-1 mice consuming the same diet. To be clear, all mice in these two colonies are of the same genetic background, and only differ in the two microbiomes they harbor. These microbiomes, originally derived from Jackson Laboratory and Envigo (now known as Inotiv), are characterized by low and high alpha diversity relative to each other and distinct beta diversity. These two colonies were developed at MU Mutant Mouse Resource and Research Center (MMRRC) by initially transferring CD-1 embryos into respective C57BL/6 dams and allowing the dams to transfer their GMs to offspring via natural postnatal transmission. These CD-1 pups became the founders of these two colonies which have been maintained and continually monitored for GM stability within our facility for over 35 generations. Additionally, a rotational breeding scheme and routine introduction of CD-1 genetics via embryo transfer from CD-1 mice purchased from Charles River allows for the maintenance of allelic heterozygosity within each colony and ensures these colonies do not become genetically distinct from each other. Since CD-1 mice that harbor a Jackson Laboratory origin GM have a GM with low phylogenetic richness and diversity, the GM of these mice was designated GM^{Low}. The CD-1 colony with an Envigo origin GM has relatively high phyologenetic diversity and is thus designated GM^{High}. Phenotypic assessments of these two colonies revealed differences in anxiety-related behavior, voluntary activity, fetal growth, food intake, and adult growth[36, 37].

We hypothesized that the maternal GM would influence the neurodevelopment of the offspring via fetal programming while *in utero* by GM-derived metabolites. Taking advantage of the

phenotypic differences in these two colonies, we utilized cross-foster studies to determine the relative influence of the parental (i.e., prenatal) and offspring (i.e., postnatal) microbiome on offspring phenotypes. Cross-fostering between dams of the reciprocal GM (e.g., pups born to GM^{Low} cross-fostered to GM^{High}, and vice versa) allows offspring to develop under the influence of the birth dam GM *in utero*, and then acquire the surrogate dam GM during postnatal life. Microbiome-associated differential phenotypes in which cross-fostered offspring match the phenotypes observed in surrogate dams suggest a postnatal influence, while similarities between cross-fostered offspring and their birth dams suggest a dominant prenatal influence of the parental microbiome. Here, we expand on previous behavioral phenotyping to include control and cross-fostered (CF) offspring, demonstrating a dominant influence of the birth dam GM on offspring development and behavior at seven weeks of age. This work was complemented by microbial and metabolic profiling of mice in each colony, genome-wide methylome analysis of hippocampal DNA from dams and control and CF offspring, and single nuclei transcriptome analysis of RNA from control and CF offspring. Previously identified phylogenetic differences are now complemented by differences in certain metabolites, including bile acids (BA), and differential ileal and hepatic expression of BA receptors and transporters. Analysis of hippocampal DNA revealed dominant effects of the maternal microbiome on CpG methylation, maintained in offspring independent of the postnatal microbiome. Single-nuclei RNA sequencing (snRNA-seq) of hippocampal RNA confirmed fetal programming of several cell-specific differentially methylated genes, including genes involved in G protein-coupled receptor and orexigenic signaling pathways.

Methods

Mice

All mice tested in the current study were outbred CD-1 mice (Crl:CD1(ICR)) generated from breeders obtained from the Mutant Mouse Resource and Research Center at the University of Missouri (MU MMRRC). Multiple different cohorts of mice were utilized for various outcomes. CD-1 mice were from two colonies in which the founders were originally purchased from Charles River (Frederick, MD), and were generated via rederivation to harbor either a high richness Envigo (now Inotiv, Indianapolis, IN) origin GM (GM^{High}), or a low richness Jackson Laboratory origin GM (GM^{Low}) as previously described[24]. All donor mice were reared at the MU MMRRC and the two colonies have been maintained and continually monitored for GM stability within our facility for over 35 generations. Additionally, a rotational breeding scheme and routine

introduction of CD-1 genetics via embryo transfer from CD-1 mice purchased from Charles River allows for the maintenance of allelic heterozygosity within each colony and ensures these colonies do not become genetically distinct from each other. Since CD-1 mice that harbor a Jackson Laboratory-origin GM were found to have a GM with low phylogenetic richness and diversity, the GM of these mice was designated GM^{Low}. Similarly, since CD-1 mice that harbored an Envigo-origin GM were found to have a GM with high phylogenetic richness and diversity relative to GM^{Low}, the GM of these mice was designated GM^{High}. Colonies of mice were housed under barrier conditions in microisolator cages with compressed pelleted paper bedding and nestlets, on ventilated racks with *ad libitum* access to irradiated chow and acidified, autoclaved water, under a 14:10 light/dark cycle. Mice were determined to be free of all bacterial pathogens including *Bordetella bronchiseptica*, *Filobacterium rodentium*, *Citrobacter rodentium, Clostridium piliforme, Corynebacterium bovis, Corynebacterium kutscheri, Helicobacter* spp., *Mycoplasma* spp., *Rodentibacter* spp., *Pneumocystis carinii, Salmonella* spp., *Streptobacillus moniliformis, Streptococcus pneumoniae*; adventitious viruses including H1, Hantaan, KRV, LCMV, MAD1, MNV, PVM, RCV/SDAV, REO3, RMV, RPV, RTV, and Sendai viruses; intestinal protozoa including *Spironucleus muris, Giardia muris, Entamoeba muris,* trichomonads, and other intestinal flagellates and amoebae; intestinal parasites including helminths; and external parasites including all species of lice and mites, via quarterly sentinel testing performed by IDEXX BioAnalytics (Columbia, MO). Fecal samples were collected from pregnant dams at 19 days of gestation, and from mouse pups at time of weaning (21 days of age) using previously described methods[38]. Briefly, mice were placed in an empty autoclaved cage within a biological safety cabinet and allowed to defecate. Freshly evacuated samples feces were immediately collected into a sterile collection tube using autoclaved wooden toothpicks discarded after each single usage. All samples were promptly placed on ice. Following fecal sample collection, samples were stored in a -80°C freezer until DNA extraction was performed. Samples were collected from all experimental mice at 50 days of age, at time of necropsy. All dams were mated with sires of the same GM and were housed together until approximately day 14 of gestation, at which time sires were removed. All dams were singly housed for the last week of gestation to ensure pups were correctly assigned to their birth dams. All dams were handled minimally during gestation, and only handled for routine cage changes by vivarium care staff and once for pre-parturition fecal sample collection. During the one week of anxiety-related behavior testing, only the investigators handled and entered the home cages to avoid unknown and excessive disturbances to the mice.

Gut microbiome analysis

DNA extraction. Fecal DNA was extracted using QIAamp PowerFecal Pro DNA kits (Qiagen), according to the manufacturer's instructions, with the exception that the initial sample disaggregation was performed using a TissueLyser II (Qiagen), rather than a vortex and adaptor as described in the protocol.

16S rRNA amplicon library preparation and sequencing. Extracted fecal DNA was processed at the University of Missouri DNA Core Facility. Bacterial 16S rRNA amplicons were constructed via amplification of the V4 region of the 16S rRNA gene using previously developed universal primers (U515F/806R), flanked by Illumina standard adapter sequences[39, 40]. Oligonucleotide sequences are available at proBase[41]. Dual-indexed forward and reverse primers were used in all reactions. PCR was performed in 50 µL reactions containing 100 ng metagenomic DNA, primers (0.2 µM each), dNTPs (200 µM each), and Phusion high-fidelity DNA polymerase (1U, Thermo Fisher). Amplification parameters were $98°C^{(3\ min)}$ + [$98°C^{(15\ sec)}$ + $50°C^{(30\ sec)}$ + $72°C^{(30\ sec)}$] × 25 cycles + $72°C^{(7\ min)}$. Amplicon pools (5 µL/reaction) were combined, thoroughly mixed, and then purified by addition of Axygen Axyprep MagPCR clean-up beads to an equal volume of 50 µL of amplicons and incubated for 15 minutes at room temperature. Products were washed multiple times with 80% ethanol and the dried pellet was resuspended in 32.5 µL EB buffer (Qiagen), incubated for two minutes at room temperature, and then placed on a magnetic stand for five minutes. The final amplicon pool was evaluated using the Advanced Analytical Fragment Analyzer automated electrophoresis system, quantified using quant-iT HS dsDNA reagent kits, and diluted according to Illumina's standard protocol for sequencing on the MiSeq instrument.

*Bioinformatics***.** DNA sequences were assembled and annotated at the MU Informatics Research Core Facility. Primers were designed to match the 5' ends of the forward and reverse reads. Cutadapt[42] (version 2.6) was used to remove the primer from the 5' end of the forward read. If found, the reverse complement of the primer to the reverse read was then removed from the forward read as were all bases downstream. Thus, a forward read could be trimmed at both ends if the insert was shorter than the amplicon length. The same approach was used on the reverse read, but with the primers in the opposite roles. Read pairs were rejected if one read or the other did not match a 5' primer, and an error-rate of 0.1 was allowed. Two passes were made over each read to ensure removal of the second primer. A minimal overlap of three bp with the 3' end of the primer sequence was required for removal. The QIIME2[43] DADA2[44] plugin (version 1.10.0) was used to denoise, de-replicate, and count ASVs (amplicon sequence variants), incorporating the following parameters: 1) forward and reverse reads were truncated

to 150 bases, 2) forward and reverse reads with number of expected errors higher than 2.0 were discarded, and 3) Chimeras were detected using the "consensus" method and removed. R version 3.5.1 and Biom version 2.1.7 were used in QIIME2. Taxonomies were assigned to final sequences using the Silva.v132[45] database, using the classify-sklearn procedure. The cladogram was constructed with GraPhlAn using genus-level taxonomic classifications[46]. Branch color depicts phylum-level classification. The outer ring denotes Benjamini-Hochberg-corrected *p* values from Wilcox Rank-Sum tests comparing the relative abundance of each genus between GMs. The color of the outer ring indicates the GM with the greater average relative abundance of that genus.

Real-time reverse transcription-polymerase chain reaction (qRT-PCR)

Total RNA was isolated from tissues using the Qiagen RNeasy Mini Kit per manufacturer's instructions. RT-PCR was performed using the BioRad iTaq Universal SYBR Green One-Step Kit following the manufacturer instructions. Briefly, each reaction consisted of 5 µL of SYBR Green Supermix, 0.125 µL iScript reverse transcriptase, 0.45 µL of forward and reverse primers, 1.475 µL of water, and 2.5 µL of template RNA. The reaction was run on a BioRad C1000 Touch thermal cycler with a BioRad CFX384 Real-Time System with the following parameters: 50˚C for 10 min for reverse transcription, 95˚C for 1 min for DNA Polymerase activation and DNA denaturation, and 40 cycles of 95˚C for 10 sec and 60˚C for 30 sec. Melt-curve analysis was performed using the following parameters: 65-95˚C with 0.5˚C increments for 5 sec/step. Primers used can be found in **Table S1**.

Bile salt hydrolase metagenomic and metatranscriptomic analysis of mouse feces

Metagenomic and metatranscriptomic *bsh* (K01442) read counts were acquired from previous multi-omic analysis of GM^{Low} (GM1) and GM^{High} (GM4)[47]. Expression counts from metatranscriptomic analysis were normalized to *bsh* metagenomic reads.

Measurement of BSH activity in mouse feces

BSH activity in the mouse feces was measured using previously reported bioluminescent bile acid activatable luciferin probes (BAL) protocol[48] with major modification by replacing whole-cell bioluminescence readout with the recombinant luciferase enzymatic assay[49]. Mouse fecal samples were soaked in PBS (pH 7.4, Gibco, ref# 10010-023) supplemented with 2-mercaptoethanol (Acros Organics, 20 mM) at a concentration of 10 mg/mL on ice for 30 min. The mixtures were homogenized by sonication in ultrasound cleaner (Elmasonic Easy 40 H, 340

W) at 0°C for 30 min, stirring every 10 min. Resulting mixtures along with blank buffer (3 replicates by 50 µL) were mixed with working solutions of BAL probes (50 µL, 20 µM in PBS) along with a solution of luciferin (50 µL, 2 µM in PBS) in a 96-well assay plate (Corning, ref# 3595) and incubated at 37°C for 1 h. After incubation, the mixtures were diluted with 2% Triton X-100 in PBS (100 µL) to stop the reaction. In a separate 96-well flat bottom black plate (Corning, ref# 3650), resulting mixtures (5 µL) were diluted with PBS (50 µ). A luciferase solution containing recombinant luciferase from Photinus pyralis (Sigma-Aldrich, 20 µg/mL), ATP disodium trihydrate (Fisher Scientific, 2 mM), and magnesium sulfate heptahydrate (Fisher Scientific, 2 mM) in PBS (50 µL) was added to each well simultaneously. Bioluminescence was measured immediately in an IVIS Spectrum (Xenogen) imaging system for 20 min with 1 min intervals using automatic settings. Raw data were processed using Living Image 4.2 software (Caliper LifeSciences), further data processing was carried out in Excel (Microsoft 365), and finally visualization and statistical calculations were performed in Prism 9 (GraphPad software). Deconjugation potentials or percentage of probe hydrolysis were calculated as the ratio of the signal from the BAL probe to the signal from luciferin in the corresponding fecal extract and reported as the mean value of 3 replicates. The signals from incubation of BAL probes in blank buffer provided a background result of nonspecific hydrolysis of the probes.

Metabolite analyses

GC-MS. Fecal and serum samples were diluted in 18 volumes of ice-cold 2:2:1 methanol/acetonitrile/water containing a mixture of internal standards (D4-citric acid, D4-succinic acid, D8-valine, and U13C-labeled glutamine, glutamic acid, lysine, methionine, serine, and tryptophan; Cambridge Isotope Laboratories), where the 1-part water was composed of sample volume + water. Sample extraction mixtures were vortexed for 10 minutes at RT and rotated for 1 hour at -20°C. Mixtures were centrifuged for 10 minutes at 21,000 × g, and 150 µL of the cleared metabolite extracts were transferred to autosampler vials and dried using a SpeedVac vacuum concentrator (Thermo). Dried metabolite extracts were reconstituted in 30 µL of 11.4 mg/mL methoxyamine (MOX) in anhydrous pyridine, vortexed for 5 minutes, and heated for 1 hour at 60°C. Next, to each sample 20 µL of N,O-Bis(trimethylsilyl)trifluoroacetamide (TMS) was added, samples were vortexed for 1 minute, and heated for 30 minutes at 60°C. Derivatized samples were analyzed by GC-MS. One µL of derivatized sample was injected into a Trace 1300 GC (Thermo) fitted with a TraceGold TG-5SilMS column (Thermo) operating under the following conditions: split ratio = 20:1, split flow = 24 µL/minute, purge flow = 5

mL/minute, carrier mode = Constant Flow, and carrier flow rate = 1.2 mL/minute. The GC oven temperature gradient was as follows: 80°C for 3 minutes, increasing at a rate of 20°C/minute to 280°C, and holding at a temperature at 280°C for 8 minutes. Ion detection was performed by an ISQ 7000 mass spectrometer (Thermo) operated from 3.90 to 21.00 minutes in EI mode (-70eV) using select ion monitoring (SIM).

LC-MS SCFA analysis. 18-fold (w/v) extraction solvent (Acetonitrile:Methanol:Water (2:2:1)) containing deuterated SCFA standards (D3-acetate, D7-butyrate, and D5-propionate) was added to each sample and rotated at -20°C for 1 hr and then centrifuged at 21,000 × g for 10 min. Supernatant was used for LC-MS SCFA analysis. LC-MS data was acquired on a Thermo Q Exactive hybrid quadrupole Orbitrap mass spectrometer with a Vanquish Flex UHPLC system or Vanquish Horizon UHPLC system. The LC column used was a ZIC-pHILIC guard column (20 × 2.1 mm). The injection volume was 2 µL. For the Mobile phase, Solvent A consisted of 20 mM ammonium carbonate [(NH4)2CO3] and 0.1% ammonium hydroxide (v/v) [NH4OH] at pH ~9.1] and Solvent B consisted of Acetonitrile. This method was run at a flow rate of 0.1 mL/min, and the injection volume was 2 µL. Linear gradient was used at 70% solvent B with a 5 min elution time. The mass spectrometer was operated in targeted selected ion-monitoring (tSIM) mode from 1 to 5 minutes. An inclusion list for the three short chain fatty acids and their deuterated versions were used in tSIM method.

LC-MS bile acid analysis. Extraction solvent (methanol:acetonitrile:water, 2:2:1) was spiked with (5 µLl/mL) deuterated bile acids MaxSpec Mixture (Cayman Chemicals Item no. 33506). 18-fold volume extraction buffer was added to each sample. The samples were placed in a -20°C freezer for 1 hour while rotating. The samples were then centrifuged at 21,000 × g for 10 minutes. Supernatant was transferred to LC-MS autosampler vials for analysis. LC-MS data was acquired on a Thermo Q Exactive hybrid quadrupole Orbitrap mass spectrometer with a Vanquish Flex UHPLC system or Vanquish Horizon UHPLC system. A Thermo Hypersil GOLD (2.1 × 150 mm, 1.9 µm) UHPLC column was used with a column Temperature of 30°C. For the Mobile Phase, solvent A consisted of 1% acetonitrile in water with 0.1% formic acid, and solvent B is 99% acetonitrile with 0.1% formic acid. The gradient started at 50% Solvent B and was held for 2.5 minutes; then increased to 100% B at 10 minutes and held for 0.5 minutes before re-equilibration to 50% solvent B for 5.5 min. Flow Rate was 0.4 mL/min, and injection volume was 3 µL. The mass spectrometer was operated in full-scan negative mode, with the spray voltage set to 3.0 kV, the heated capillary held at 275°C, and the HESI probe held at 350°C. The sheath gas flow was set to 40 units, the auxiliary gas flow was set to 15 units, and the sweep gas flow

was set to 1 unit. MS data acquisition was performed in a range of m/z 70–1,000, with the resolution set at 70,000, the AGC target at 1×10^6, and the maximum injection time at 200 ms[50].

Metabolomic Data Analysis. GC-MS Raw data were analyzed using TraceFinder 5.1 (Thermo). Metabolite identification and annotation required at least two ions (target + confirming) and a unique retention time that corresponded to the ions and retention time of a reference standard previously determined in-house. A pooled-sample generated prior to derivatization was analyzed at the beginning, at a set interval during, and the end the analytical run to correct peak intensities using the NOREVA tool[51]. NOREVA corrected data were then normalized to the total signal per sample to control for extraction, derivatization, and/or loading effects. Acquired LC-MS data were processed by Thermo Scientific TraceFinder 4.1 software, and metabolites were identified based on the University of Iowa Metabolomics Core facility standard-confirmed, inhouse library. NOREVA was used for signal drift correction[51]. For bile acids, data were normalized to one of the d4-bile acid standards. For SCFA, analyte signal was corrected by normalizing to the deuterated analyte signal and the signal obtained from processing blank was subtracted.

Cross-fostering

Mice in cross-foster (CF) groups were cross-fostered at less than 24 hours of age to a surrogate dam of the reciprocal GM. Following identification of recently birthed litters from both GMs, cages were moved to a biosafety cabinet. Litters were removed from the cage of the biological dam and placed onto clean paper towels. Bedding from the cage of the surrogate dam was gently mixed with the pups to transfer the surrogate dam scent to the pups and reduce the possibility of cannibalism. The pups were then placed into the surrogate dam cage, and cages were returned to the appropriate housing rack.

Behavior testing

Open Field Exploration. The open field exploration test was used to evaluate anxiety-related behavior and locomotor function. Environmental control chambers (Omnitech Electronics, Inc., Columbus, OH) consisting of 4 separate environmental isolation chambers containing a plexiglass box (41cm × 41cm × 30cm) placed onto an infrared grid (41cm × 41cm) to track locomotion. Lighting for each isolation chamber was set to 159 lux. Mice were allowed to acclimate to the behavior room for 1 hour prior to testing. Before starting each test, the plexiglass was cleaned with 0.25% bleach, followed by 70% ethanol to remove any residual

olfactory cues. Each mouse was placed into the middle of the open field exploration test and recorded for 30 minutes by the Fusion behavior monitoring software (Omnitech Electronics, Inc., Columbus, OH). The first 20 minutes of the test were considered acclimation time, and the final 10 minutes were analyzed following completion of the test. Total distance traveled (cm), time spent in the center zone (seconds), distance traveled in the center zone, and vertical activity (rearing) were measured.

Light/Dark Transition. The light/dark transition test was performed within the environmental control chambers (Omnitech Electronics, Inc., Columbus, OH, USA). The apparatus consisted of a plexiglass box within the environmental chambers (41cm × 41cm × 30cm) that were partitioned into two equal sections by a black plexiglass insert with a door that allowed one half to be a dark zone, and a second half to be a light zone. The light zone was illumination was set to 200 lux. Mice were allowed to acclimate to the behavior room for 1 hour prior to testing. Before the start of each test, the plexiglass box and insert were cleaned with 0.25% bleach, followed by 70% ethanol to remove any residual olfactory cues. Mice were placed into the light zone facing away from the dark zone and monitored by Fusion behavior monitoring software (Omnitech Electronics, Inc., Columbus, OH, USA) for 15 minutes. Time spent in the light zone (sec), distance travelled in the light zone (cm), and number of transitions between light and dark zones were measured.

Elevated Plus Maze. The elevated plus maze test consisted of an apparatus with two open arms (32.5 × 5 cm, with 2-mm ledges) and two closed arms (32.5 × 5 cm, with 14.5 cm high walls). The open arms were arranged perpendicular to the closed arms so the apparatus formed the shape of a plus sign with a center square (5 × 5 cm). The entire apparatus was raised 50 cm above the floor. The center zone of the apparatus was illuminated to 50 lux. Mice were allowed to acclimate to the behavior room for 1 hour prior to testing. Prior to testing, the apparatus was cleaned with 0.25% bleach followed by 70% ethanol to remove olfactory cues. Each mouse was placed in the center square facing an open arm and was recorded and monitored for 5 minutes. Distance Traveled in the open arms (cm), time spent in the open arms (sec), and number of entries into the open arms was calculated from distance measurements and entry counts obtained by Any-Maze monitoring software (Stoelting Co., Wood Dale, IL, USA).

Voluntary running. New litters of CD-1 GM^{Low} and GM^{High} mice were generated to evaluate voluntary wheel running assays (the mice used in the behavior assays did not undergo wheel running evaluation). Litters for running wheel experiments were culled to six pups per litter (3 male, 3 female) at birth, and then weaned into cages of same-sex trios at weaning. During wheel set-up at seven weeks of age, mice were transferred from their home cage to a new static

microisolator cage containing a wireless running wheel (Med Associates, ENV-047) connected to a wireless hub and laptop computer in the animal room. Only investigators entered the behavior room to check mice and equipment daily during the 12 days of testing to avoid excessive disturbances to the mice. Mice were singly housed during the experiment, assigned to running wheel cages using a random number generator, and were placed in alternating order on the shelf such that microbiome group and sex were consistently alternated. Following five days of acclimation, data were collected continuously for seven consecutive days using Wheel Manager software, v2.04.00 (Med Associates, SOF-860). Data were analyzed using Wheel Analysis software, v2.02.01 (Med Associates, SOF-861). No other mice were housed in the room containing running wheel cages, traffic was limited to once daily checks at the same time of day by one laboratory staff, and no cage changes were performed during the acclimation and testing period.

Necropsy

Mice were transported to the necropsy room at 50 days of age and allowed to acclimate to the room for 1 hour. Mice were then euthanized one at a time by CO_2 asphyxiation out of sight of other mice. The euthanasia chamber was cleaned with 70% ethanol between mice to eliminate olfactory cues. Following loss of paw pinch and righting reflexes, blood was collected by cardiac puncture and placed into serum separator tubes. The brain was then removed, and the hippocampus was gently dissected out, placed in a 2 mL tube, and promptly plunged into liquid nitrogen to flash freeze. Liver and ileum were isolated, placed in a 2 mL tube, and promptly plunged into liquid nitrogen. Two fecal pellets were collected from the colon for 16S rRNA amplicon sequencing. Blood was allowed to clot for 30 minutes at room temperature and was then centrifuged at 4,000 RPM for 15 minutes, and serum placed into a 1.5 mL microcentrifuge tube. Hippocampus, feces, and serum were promptly placed into -80°C freezer for storage.

Methylome analysis

Due to constraints on resources, methylome analysis were performed in only one sex. To assess intergenerational effects on DNA methylation, we analyzed dams and female offspring in both control and CF mice. Dams from both colonies were time mated, and following birth, litters were culled to six female pups. Three of the pups from each litter remained with the birth dam, and three were cross-fostered onto a dam of the opposite GM so that every dam from both GMs had three of their birth pups and three cross-foster surrogate pups. Hippocampi were collected from all dams following weaning, flash frozen in liquid nitrogen, and stored at -80°C. At seven

weeks of age, hippocampi from the offspring were collected, flash frozen in liquid nitrogen, and stored at -80°C. Hippocampal genomic DNA was isolated from adult female CF and control offspring hippocampi using the DNeasy kit (Qiagen, Valencia, CA) following manufacturer instructions. For studying genome-wide DNA methylation profiles, 1 µg of genomic DNA was treated with sodium bisulfite (Zymo Research, Irvine, CA). Converted DNA was analyzed using Infinium Mouse Methylation BeadChip assay (Illumina, San Diego, CA). This array includes over 285,000 CpG sites covering all RefSeq genes, including CpG islands, translation start sites, enhancers, imprinted loci, and other regions[52]. All data analyses were conducted using the R environment version 4.2.0. Microarray data was processed using the *ENmix* version.1.34.02[53] and *minfi* v.1.44.0[54] packages. Quantile normalization of U or M intensities for Infinium I or II probes were performed, respectively. A model-based correction was performed, and a dye-bias correction was conducted using *RELIC* [55]. β-values representing the averaged total intensity value per CG position was calculated as unmethylated intensity (U) + methylated intensity (M) [M / (U + M + 100)]. Probes with a detection $p < 1 \times 10^{-6}$ and less than 3 beads were defined as low quality. Samples with low quality methylation measurements > 5% or low intensity bisulfite conversion probes were removed from further analysis. Differentially methylated regions (DMRs) between the experimental groups were determined using the *ENmix* version.1.34.02[53] package. For each position, the magnitude of the DNA methylation difference was expressed as Fold Changes in the logarithmic scale (logFC) and the significance of the difference as a FDR-corrected p value (q value).

Isolation of hippocampal nuclei

Single nuclei were isolated from mouse hippocampal tissue samples as follows. Briefly, Nuclei Lysis Buffer was prepared by adding 12 mL of Nuclei EZ Prep Lysis Buffer (Sigma-Aldrich, St. Louis, MO, USA) to a 15mL tube and adding 1 cOmplete Ultra tablet (Sigma-Aldrich, St. Louis, MO, USA) and allowing tablet to dissolve. Two 4-mL aliquots of the Nuclei EZ Prep Lysis Buffer + cOmplete Ultra tablets were then placed in 15 mL tubes. Twenty (20) µL of Protector RNase inhibitor (Sigma-Aldrich, St. Louis, MO, USA) and 20 uL of Superase-In (MilliporeSigma, Burlington, MA, USA) were added to one 4 mL aliquot to make Nuclei Lysis Buffer 1 (NLB1). Four (4) µL of Protector RNase inhibitor (Sigma-Aldrich, St. Louis, MO, USA) and 4 µL of Superase-In (MilliporeSigma, Burlington, MA, USA) were added to the second -mL aliquot to make Nuclei Lysis Buffer 2 (NLB2). Suspension Buffer (SB) was prepared by adding 1 mL of fetal bovine serum (Sigma-Aldrich, St. Lous, MO, USA) to 9 mL of 1× phosphate-buffered saline with 4 µL of Protector RNase inhibitor (Sigma-Aldrich, St. Louis, MO, USA). Eight (8) hippocampi halves from eight individual mice were pooled and homogenized to a single cell

suspension in a 2 mL dounce homogenizer with 2 mL of NBL1. The sample was then strained through a 200 µm strainer (pluriSelect Life Science, El Cajon, CA, USA) and the strained cell suspension returned to the 2 mL dounce and homogenized to a nuclei suspension. The nuclei were strained through a 40 µm strainer (pluriSelect Life Science, El Cajon, CA, USA) and centrifuged at 500 RCF at 4°C for 5 minutes. Supernatant was removed, and pellet was resuspended with NLB2 and incubated at 4°C for 5 minutes. Nuclei were then centrifuged at 500 RCF at 4°C for 5 minutes, supernatant was removed, and nuclei were resuspended in suspension buffer.

10x Genomics single cell 3' RNA-Seq library preparation

Libraries were constructed by following the manufacturer's protocol with reagents supplied in 10x Genomics Chromium Next GEM Single Cell 3′ Kit v3.1. Briefly, nuclei suspension concentration was measured with an Invitrogen Countess II automated cell counter. Nuclei suspension (1,200 nuclei per microliter), reverse transcription master mix, and partitioning oil were loaded on a Chromium Next GEM G chip with a capture target of 8,000 nuclei per library. Post-Chromium controller GEMs were transferred to a PCR strip tube and reverse transcription performed on an Applied Biosystems Veriti thermal cycler at 53°C for 45 minutes. cDNA was amplified for 13 cycles and purified using Axygen AxyPrep MagPCR Clean-up beads. Fragmentation of cDNA, end-repair, A-tailing and ligation of sequencing adaptors was performed according to manufacturer specifications. The final library was quantified with the Qubit HS DNA kit and the fragment size determined using an Agilent Fragment Analyzer system. Libraries were pooled and sequenced on an Illumina NovaSeq 6000 to generate 50,000 reads per nuclei.

Single cell bioinformatics

FASTQ files were generated from the raw base call outputs using Cell Ranger (10x Genomics) pipeline, *mkfastq* v3.0. Using default parameters, a UMI (gene-barcode) count matrix per sample was obtained using the built-in Cell Ranger count pipeline. To reduce noise, we only kept genes that were detected in at least three barcodes, and subsequently removed ribosomal-encoded genes from the count matrix. Scrublet[56] was then used to identify potential multiplet-barcodes and only those with a doublet score of less than 0.15 were used for downstream analyses. The files were then combined in a single embedding for the control and CF groups separately, following the Seurat v3 integration workflow[57]. SCTransform was used to normalize each sample, followed by the identification of integration anchors and variable features using

the Seurat workflow. Dimension reduction was performed scaled data after 4000 highly variable genes across the samples were identified (*SelectIntegrationFeatures* function). The *IntegrateData* function was then used to obtain a combined and centered matrix, where the top 30 components were used to carry out ordination analyses. These components were used to build a SNN (shared nearest neighbour) graph which was subsequently clustered using the *Louvain* algorithm for speed and computational efficiency. The principal components were then mapped into two dimensions using the default uniform manifold approximation and projection (UMAP) algorithm, where the n = 30 neighbours was set, with a minimum distance of 0.4. Finally, the *FindAllMarkers* function was used to identify marker genes for each cluster. The top marker genes were used manually based on literature searches to assign cell type annotations for each cluster. This was further corroborated by cluster annotations using the Azimuth mouse reference datasets[57]. The cell types across samples and groups were combined with their pseudo bulk profiles, and the resulting gene-cell type matrix was normalized by estimating transcripts per million and transformed (log2) for downstream analyses. To obtain statistically enriched differential gene expression, we used generalized additive regression models, where in the control or CF variables, alongside the GM^{Low} or GM^{High} status were encoded as independent variables. The models were analyzed for each cluster independent of the other, where per gene log2 fold-change was determined. Significance was identified as those genes with an adjusted p value of less than 0.05, following Benjamini-Hochberg correction. All figures and statistical analyses were performed using R v4.1[58].

Cell-to-cell communication was inferred using log10-transformed gene expression data collected from snRNA-seq of the mouse hippocampus using CellChat[59, 60]. Cell clusters as identified above were grouped into the following cell types based on Azimuth classification: glutamatergic and GABAergic neurons, periendothelial cells, microglia, astrocytes, and oligodendrocytes. Ligand-receptor interactions were inferred using the mouse reference database provided by CellChat (Accessed May 25, 2023). Overexpressed genes and interactions were determined using default settings. Cell-cell communication probability was inferred using default settings. Information flow was determined by the summation of communication probability for each pathway.

Statistics

Two-way Analysis of Variance (ANOVA) followed by Tukey's *post hoc* analysis was used to test for main effects of GM and sex in OFE, LDT, and EPM behavior tests for all behavior testing parameters. Due to lack of normality, CF LDT parameter distance travelled in the light zone and

CF EPM parameters time spent in open arms and distance travelled in open arms were normalized by square root transformation. Two-way ANOVA followed by Tukey's *post hoc* was used to test for main effects of GM and Time (day) for voluntary running distance data. Two-way ANOVA followed by Tukey's *post hoc* was used to test for main effects of GM and sex in the body weight data. Two-way Permutational analysis of variance (PERMANOVA) was used to test group beta-diversity for main effects of GM and sex. Two-way ANOVA followed by *Tukey's post hoc* was used to test Chao-1 richness for main effects of GM and treatment. Since it was not possible to include a male dam group, the main effect of sex was excluded from Chao-1 analysis. Chao-1 richness was calculated using PAST 4.03 software[61]. Differences in family- and genus-level relative abundance between GMs was assessed using Wilcoxon Rank-Sum tests with a Benjamini-Hochberg correction for multiple comparisons. Due to uniform lack of normality across metabolites, a Mann-Whitney U test was used to test for differences in metabolite concentrations between GM^{Low} and GM^{High} treatment groups. Spearman's rank correlation was used to test correlations between genus-level abundance and statistically significant metabolites. Two-way ANOVA followed by Tukey's *post hoc* was used to test for main effects of GM and sex in the gene expression data. A student's t-test was used to test for differences in the GM^{Low} and GM^{High} groups in the *bsh* metagenomic and metatranscriptomic read counts. A two-way ANOVA followed by Tukey's *post hoc* was used to test for main effects of GM and sex in the BSH activity data. All univariant data analysis was performed using SigmaPlot 14.0 (Systat Software, Inc, San Jose, CA). Shapiro-Wilk test was used to test for normality using SigmaPlot 14.0. Two-way PERMANOVA testing was based on Bray-Curtis dissimilarities using the *adonis2* library from the *vegan* library[62].

Results

Taxonomic differences are associated with different levels of biologically relevant metabolites

Compositional differences between GM^{Low} and GM^{High} have been described previously[36, 37, 63], including greater richness in GM^{High} compared to GM^{Low} (**Fig. 1A**). Families enriched in GM^{Low} included *Erysipelatoclostridiaceae*, *Erysipelotrichaceae*, *Sutterellaceae*, *Saccharimonadaceae*, *Acholeplasmataceae*, and unresolved members of orders *Tissierellales* and RF39; while families enriched in GM^{High} included *Prevotellaceae*, *Marinifilaceae*, *Clostridiaceae*, *Desulfovibrionaceae*, *Deferribacteraceae*, and an unresolved family within the order *Rhodospirillales* (**Table S2**). Genera that were enriched in GM^{Low} included *Anaeroplasma,*

Lachnoclostridium, and *Oscillospira*; while genera that were enriched in GM^{High} included *Odoribacter, Alloprevotella, Rikenella, Bilophila*, and *Desulfovibrio* (**Fig. S1**; **Table S3**). To determine whether these phylogenetic differences were associated with different metabolite profiles, fecal samples were collected from both male and female GM^{Low} and GM^{High} mice, and a combination of mass spectrometry (MS)-based platforms were used to measure short-chain fatty acids; branched chain fatty acids, branched chain keto-acids, and other lipids; unconjugated primary and secondary bile acids; all proteinogenic amino acids and several nonproteinogenic and acetylated amino acids; tryptophan derivatives including kynurenine, serotonin, and several indoles; B vitamins; dicarboxylic acids; glucose, fructose, and other compounds within glycolysis; all ribonucleosides and nitrogenous bases; compounds within the pentose phosphate pathway; compounds within the classic TCA cycle; and other biologically relevant microbial metabolites. Analysis of fecal metabolites revealed differences in sugar molecules involved in glycolysis, multiple amino acids, and primary bile acids (**Fig. 1B**; **Table S4**). Specifically, glucose-6-phosphate, fructose-6-phophate, ribulose-5-phosphate, and β-alanine were enriched in GM^{Low} feces, while cysteine, succinate, lactate, chenodeoxycholic acid (CDCA) and deoxycholic acid (DCA) were enriched in GM^{High} feces. These differences were particularly apparent in the feces of male mice (**Fig. S2A**), while female mice also showed a single secondary bile acid (lithocholic acid, LCA) enriched in the feces of GM^{Low} (**Fig. S2B**). When genus-level taxonomic abundances were compared to each of the significant metabolites, numerous significant correlations were identified indicating that taxonomic features in the GM were strongly associated with the differentially abundant metabolites (**Fig. S3; Table S5**). Metabolites were also measured in serum, collected at the same time as fecal samples. Remarkably, the only significant difference in the serum was a primary bile acid (CDCA) (**Fig. 1C**; **Table S6**). There were also several bile acids, both primary and secondary, that while not reaching statistical significance, had greater levels within the serum of female GM^{Low} and male GM^{High} mice (**Fig. S2C-D**).

To determine whether these differences in bile acids are associated with differences in bile acid cellular transport and receptor signaling, quantitative RT-PCR was performed with mRNA from both ileal and hepatic tissue. *Asbt*, a gene involved in transporting bile acid from the ileal lumen into the ileal epithelium did not show a difference in gene expression (**Fig. 1D**). However, *Ost-β*, a gene involved in transport of bile acids from the ileal epithelium into vascular circulation, showed greater expression in ileum of GM^{High} mice (**Fig. 1E**). *Gpbar1*, a G protein-coupled receptor (GPCR) also known as TGR5, was found to have greater expression in the ileum of GM^{Low} mice (**Fig. 1F**). *Fxr*, a gene critically involved in regulation of hepatic bile acid synthesis,

showed no difference in ileal expression (**Fig. 1G**). Accordingly, *Cyp7a1*, a gene downstream of ileal Fxr signaling that encodes the rate-limiting protein in bile acid synthesis, also showed no difference in hepatic expression (**Fig. 1H**). Expression of *Slc10a1,* a bile acid transport protein, was higher in the liver of GM^{Low} mice (**Fig. 1I**). *S1pr2*, a GPCR activated by bile acids, was also expressed at a greater level in the liver of GM^{Low} mice (**Fig 1J**). These data indicated that the differential abundance of BAs detected in feces and serum of GM^{Low}- and GM^{High}-colonized mice are also associated with differential ileal and hepatic expression of bile acid receptors and transporters.

We next examined whether bile salt hydrolase (BSH), the bacterial enzyme used to deconjugate bile acids, levels showed differential expression or activity between GM^{Low}- and GM^{High}-colonized female mice. While no difference was detected in total gene read count (**Fig. S2E**), microbial expression of *bsh* was significantly higher in GM^{Low} (**Fig. S2F)**. The BSH family of enzymes is widely expressed among gut bacteria, and prone to significant variations in structure with some isoforms exhibiting different deconjugation activity toward various bile acids and possessing various levels of enzymatic activities[64, 65]. Using a panel of highly sensitive bioluminescent assays[48], we compared BSH enzymatic activities of both microbiomes toward various bile acids. Our results demonstrated significantly greater BSH activity specific for cholic acid in GM^{Low} (**Fig. S2G**), but no other differences were detected in BSH activities towards other bile acids examined (**Fig. S2H-K**). Collectively, the differences in bacterial *bsh* expression and fecal and serum bile acid levels suggest greater uptake by GM^{Low} mice and greater fecal loss by GM^{High} mice. This is supported by greater hepatic expression of bile acid receptors and transporters.

Complex microbiome-dependent behaviors determined by the parental microbiome

Prior work has revealed reproducible differences in behavior and growth between GM^{Low}- and GM^{High}-colonized mice[36, 37]. To determine the developmental period in which these phenotypic differences are established, we used an experimental design with four groups, comprising mice born to dams harboring GM^{Low} or GM^{High} and remaining with their birth dams until weaning (control), or cross-fostered at birth to nursing dams harboring the reciprocal microbiome (**Fig. 2A**). These groups are denoted as CF^{Low} and CF^{High}, with the microbiome designation indicating the postnatal offspring composition acquired via cross-foster (CF). Comparisons were then made between the two control groups, and between the two CF groups in behavior tests associated with anxiety-related behavior and voluntary activity between five and seven weeks of age, and body weight (BW) at three and seven weeks of age.

In agreement with previous studies[37], age-, sex-, and genotype-matched CD-1 mice colonized with GM^{Low} spent less time and traveled less distance in the light portion of a light-dark transition (LDT) test (**Fig. 2B, Fig. S4A**), and the open arms of an elevated plus maze (EPM) test (**Fig. 2C, Fig. S4B**), relative to mice colonized with GM^{High}, indicating differential effects on anxiety-related behavior of the two microbiomes. No behavioral differences were observed in the open-field exploration test between GM^{Low} and GM^{High} mice (**Fig. S4C-D**). To assess voluntary physical activity, mice were singly housed with bluetooth wireless running wheels for a five-day acclimation period followed by a seven-day test period. Both male (**Fig. 2D**) and female (**Fig. 2E**) mice colonized with GM^{High} ran significantly more than mice with GM^{Low} ($p < 0.0001$, F = 29.2, and $p = 0.0002$, F = 14.4, respectively). Previously reported differences in body weight (BW) at weaning and adulthood were also reproducible in the GM^{Low} and GM^{High} groups[37] (**Fig. 2F**, **Fig. S4E**). Collectively, these data confirmed microbiome-associated differences in anxiety-related behavior, voluntary physical activity, and body weight, in sex-, age-, and genotype-matched mice.

Microbial 16S rRNA amplicon sequencing was used to confirm that CF mice harbored a microbiome of comparable richness and beta-diversity to that present in surrogate dams, in both directions of CF (**Fig. S5A-C**). Behavioral analysis showed the GM-associated differences in the LDT, independent of sex, were reversed in the comparison between CF^{Low} and CF^{High} mice, with the CF^{High} mice demonstrating behavior suggestive of greater anxiety (**Fig. 2G, Fig. S4F**). Similarly, the robust GM-dependent differences in the EPM were reversed in the CF^{Low} and CF^{High} mice (**Fig. 2H, Fig. S4G**) indicating that the birth dam microbiome has a substantial, if not dominant, influence on neurodevelopmental events contributing to anxiety-related behavior in the offspring. While a significant difference of total distance traveled in the OFE test was observed, distance traveled in the center was not found to be significant in the cross-foster groups **(Fig. S4H-I)**. The significant GM-associated differences in voluntary physical activity were absent in male and female CF mice (**Fig. 2I-J**). Comparison of BW revealed no difference at weaning (**Fig. S4J**) and reversal in adulthood (**Fig. 4K**), such that CF^{High} mice weighed more than age-matched CF^{Low} mice. Collectively, these data supported an equivalent or dominant effect of the birth dam GM on offspring behavioral phenotypes and body weight.

Fetal programming of gene methylation in hippocampus by the parental microbiome

Reasoning that an influence on offspring behavior by the parental microbiome must have a biological foundation in the CNS, and a mechanism by which cellular function and gene expression are established during fetal development, we next examined gene expression and

its epigenetic regulation in offspring hippocampus, given its central role in anxiety-related behavior[66-71]. To identify effects of the parental microbiome on fetal programming of DNA methylation in the hippocampus, we performed genome-wide DNA methylation profiling in female mice. Analysis of methylation across the entire array, and parallel comparisons between samples from GM^{Low} or GM^{High} mice, and from CF^{Low} or CF^{High} mice, identified only 196 differentially methylated regions (DMRs) with beta values differing by $\log_2 FC > 1$ (**Table S7**). Remarkably, 176 of those 196 sites (89.8%) showed a difference in methylation at the same site in the reciprocal contrast, such that offspring methylation matched that of their birth dam (**Fig. 3A, $R^2 = -0.644$, $p = 2 \times 10^{-7}$**). DMRs identified in offspring hippocampi were distributed across the genome, occurring most often early in gene bodies or enhancer/promoter regions and roughly correlating in frequency to chromosomal gene content, with the exception of an apparent enrichment for DMRs on chromosome 14 (**Fig. 3B, $p = 0.003$**). Differential methylation was detected at five contiguous markers mapped to promoters, enhancers, and exon 1 of the *Pde1c* gene, encoding phosphodiesterase 1c, a regulator of Ca^{2+} and cGMP-dependent intracellular signaling. This methylation pattern differed between hippocampus from GM^{Low} and GM^{High} mice, and was conserved between birth dam and offspring, regardless of postnatal colonization (**Fig. 3C**). Similarly, microbiome-associated differences in methylation were identified at multiple closely spaced markers on both strands of chromosome 14, including markers associated with the genes *Ang5* and *Ang6* (**Fig. 3D**), encoding members 5 and 6, respectively, of the angiogenin, ribonuclease A family.

To identify shared pathways or commonalities among the functional products (i.e., proteins) encoded by protein-coding genes among the DMRs, a STRING analysis was performed using a final input of 144 gene names[72]. Interaction analysis resulted in assembly of one large network with 32 nodes, a smaller network with six nodes, and four dyads (**Fig. 3E**). While enrichment analysis failed to detect greater network connectivity than would occur at random, stratified analysis of the 32 nodes in the large network revealed 10 significantly enriched Gene Ontology (GO) Biological Processes; five GO Molecular Functions including TGF-β receptor binding (GO:0005160, strength 1.94, FDR-adjusted $p = 0.015$) and GTPase activity (GO:0003924, strength = 1.09, FDR-adjusted $p = 0.042$); five KEGG pathways including Gap junction (mmu04540, strength = 1.38, FDR-adjusted $p = 0.033$) and TGF-β signaling pathway (mmu04350, strength = 1.32, FDR-adjusted $p = 0.033$); and four Reactome pathways including axon guidance (MMU-422475, strength = 1.08, FDR-adjusted $p = 0.048$). Network STRING analysis results provided in **Table S8**.

Fetal programming of gene expression in hippocampus by parental microbiome

We next performed single nuclei RNAseq on hippocampus collected from control mice and cross-fostered mice. Based on expression of cell-specific markers, 11 cell clusters were identified in the control mice (**Fig. S6A-B**) and 24 cell clusters were identified in the cross-fostered mice (**Fig. S6C-D**). Gene expression in GM^{Low} and GM^{High} control mice was compared using generalized additive regression models, and differentially expressed genes (DEGs) within each treatment group were identified (**Table S9**). Similarly, mice from CF^{Low} and CF^{High} were examined for DEGs within each cell cluster (**Table S10**). After DEGs from each cell cluster had been determined using a cutoff magnitude difference of $Log_2FC > 1.5$, STRING analysis was used to determine the protein-protein interactions of the DEGs from each treatment[72]. The number of DEGs from each cell cluster were compared to the mean node degree received from the STRING analysis results to determine which cell clusters contained a high number of DEGs that were most likely to interact within protein pathways. Interestingly the cell cluster with the highest number of DEGs, as well as the highest number of mean node degree of protein-protein interactions, in both control and cross-foster mice was identified as hippocampal endothelial cells (**Fig. S7A-B**). STRING analysis was used to generate interaction networks using DEGs identified in the endothelial cells of control mice as well as the cross-fostered mice (**Fig. 4A-B**). Of these interactions, six DEGs (*Fas, Fzd6, Gja1, Ntng1, Pik3r3, Sox17*) showed a pattern of fetal programming. Of note, numerous DEGs (including *Pde1c*, *Dock1, and Pdzrn3*) identified were also found to be differentially methylated or closely related to differentially methylated genes. When L5 IT glutamatergic neurons and astrocytes were examined using interaction networks, they also contained multiple DEGs that showed a pattern of fetal programming as well as DEGs that were identified as differentially methylated (**Fig. S8A-B; Fig. S9A-B**). When we examined the Log_2FC of *Dock1* expression of CF^{Low} and CF^{High}, *Dock1* was shown to be upregulated in the clusters of endothelial cells, oligodendrocyte, microglial cells, and subsets of astrocytes and glutamatergic neurons of CF^{High} mice compared with CF^{Low} (**Fig. 4C**). Similarly, when we examined expression of *Pde1c*, it too was found to be increased in the endothelial cell cluster of CF^{High} mice, though not consistent with the expression patterns of other cells seen with *Dock1* (**Fig. 4D**). We next used CellChat software[59] to impute cell-cell communication via patterns of coordinated ligand-receptor expression in each control and cross-foster group, to identify cell signaling pathways that show a pattern of fetal programming. There were 42 cell-cell communication pathways that were shared among the control mice and the cross-foster mice (**Fig. 4E**). Of the 42 pathways identified, 15 showed a pattern of fetal programming including VEGF, IGF, IL2, TGFβ, WNT, and NPY (**Fig. 4F-G**). Interestingly, three cell-cell communication

pathways that were identified in only GM^{Low} and CF^{High} were appetite-stimulating orexigenic neuropeptide pathways.

Discussion

Studies comparing germ-free and SPF mice demonstrate that the parental microbiome can affect offspring phenotypes associated with neurodevelopment[73], metabolic diseases including obesity[74], and organogenesis in the CNS and intestines[75]. While challenging to study in human cohorts, recent retrospective analyses suggest a dominant influence of the maternal microbiome on offspring phenotypes related to asthma[76], neurodevelopment[77], and metabolic diseases including obesity and diabetes[78]. The current data demonstrate that differences among native parental microbiomes can influence neurodevelopment and behavioral outcomes in the offspring. The current findings and prior studies[36, 37] show reproducible effects of these native SPF microbiomes on certain phenotypes. GM^{Low}-colonized CD-1 mice are consistently heavier than age- and sex-matched mice colonized with GM^{High}, and the same effect is observed in inbred C57BL/6J and BTBR T^{+} $Itpr3^{tf}$/J mice[79]. Behavior and BW data from the cross-foster (CF) mice provide strong evidence of a dominant effect of the parental GM on these behavioral phenotypes in offspring. While we did not measure food intake in the current study, previous work showed that the heavier GM^{Low}-colonized mice consume more food (normalized to BW) than age- and sex-matched GM^{High}-colonized at all timepoints examined[36]. The differences in BW between CF and control offspring would suggest that these feeding behaviors are similarly programmed during fetal, embryonic, or even pre-fertilization events. As such, these findings raise the possibility of a connection between the anxiety-related behaviors and the behaviors underlying the difference in BW and voluntary physical activity, representing a constellation of behavioral phenotypes influenced by common features within the parental microbiomes.

The GM can influence host physiology through microbially derived metabolites in peripheral circulation[73], interactions with the immune system[80], and stimulation of the vagus nerve or enteric nervous system by microbially derived neurotransmitters[81] and other molecules. Gut metabolites have been implicated as a means by which the parental microbiome can influence fetal development[73, 82]. The present data provide evidence of a functional difference between these native SPF microbiomes, including differential abundance of several bile acids. Bile acids stored in the gallbladder are released into the duodenum following food intake, and the observed differences in fecal bile acids may reflect differences in *bsh* expression, food intake, or other factors. Regardless, the observed differences in ileal and hepatic bile acid transporters

and receptors indicate that the differences in bile acid levels are physiologically relevant to the host. The observed difference in expression of *Slc10a1 (Ntcp)* may reflect a mechanism to regulate reabsorption of bile acids. GM^{Low} mice also demonstrated greater hepatic expression of *S1pr2*, a GPCR that when bound to primary conjugated bile acids is involved in the regulation of hepatic lipid metabolism[83]. There is considerable interest in the role of bile acids in anxiety and depressive disorders[84, 85], and causative links have been shown between bile acids, bile acid receptor signaling, and these outcomes[86-89]. CDCA, present at greater levels in the serum of GM^{Low}-colonized females, has been shown to readily cross the blood brain barrier[90] and influence the expression of transcription factors CREB and BDNF through FXR activation, which when down-regulated can lead to decreased neuroplasticity and mood disorders including anxiety[91]. The current findings suggest that differences in the native microbiome, independent of dietary challenge or host insult, can have intergenerational effects on these outcomes.

Analyses of hippocampal DNA methylation and gene expression were performed to document and compare the effects of the parental or offspring GM on those processes, and identify specific genes, pathways, and cell types associated with the observed phenotypic differences. It is well-established that the GM can influence the epigenome of the host[92-94]. The distribution and relationship of genes affected by differential methylation reflects a semi-stochastic effect across the genome, with enrichment of genes and pathways associated with TGF-β signaling and GTPase activity, both of which were identified in the single-cell transcriptome data as well. Four of the 144 protein-coding genes (*Dlx5*, *Drd1*, *Zfp64*, and *BC034090*) identified as DMRs here are included among a comprehensive list of 384 genes known to undergo fetal imprinting[95]. This suggests that the GM-associated effects on methylation of those and perhaps other DMRs occurred in germline cells pre-fertilization. Indeed, several recent studies have revealed the influence of the paternal microbiome on germline methylation and offspring outcomes[96-98]. As all matings in the current study were between mice sharing the same microbiome, it is unclear whether the effects of the GM on offspring DNA methylation occurred pre- or post-fertilization and whether the maternal lor paternal microbiome had a dominant or selective influence. As even transient co-housing to breed mice results in sharing of the GM, *in vitro* fertilization or similar methods would be needed to investigate those questions. There was an incredibly high degree of conservation across all mice in the degree of methylation at the vast majority of CpG sites included in the BeadChip array. However, we also observed a high degree of conservation between dams and offspring in the specific DMRs affected by the GM, suggesting the affected loci are not the result of random DNA methyltransferase (DNMT) activity, but rather an outcome with a teleological explanation. While speculative, the methylation and gene expression profiles

following a pattern of fetal programming may represent an intergenerational feedback mechanism wherein nutrient availability in the parent may program the trafficking of, or receptor response to, microbial metabolites as a way of fine-tuning offspring metabolism.

It is worth noting that the number and connectivity of DEGs were greatest in endothelial cells, in both control and CF mice. These cells supply blood to tissues within the CNS and comprise the blood-brain barrier. The greater size of GM^{Low}-colonized mice would necessitate greater amount of peripheral vasculature to adequately perfuse tissues. Indeed, prior work found the total cardiac weight of GM^{Low}-colonized mice to be significantly greater than age- and sex-matched GM^{High}-colonized mice, and no difference in cardiac weight when normalized to total BW[37], indicating differential growth of the circulatory system commensurate with the difference in BW. Moreover, the same study found no difference in body composition based on DEXA scanning, and a significant correlation between BW and crown-to-rump length, further indicating that the observed phenotypic difference is associated with somatic growth rather than adiposity. Several of the pathways identified in the CellChat analysis showing patterns of fetal programming represented growth factors including TGF-β, vascular endothelial growth factor (VEGF) and insulin-like growth factor (IGF).

Two of the DEGs identified in the endothelial cells included *Pde1c* and *Dock1*, two genes that were also found to be differentially methylated. PDE1C is a member of the phosphodiesterase family of enzymes involved in the production of cyclic guanosine monophosphate (cGMP) and cyclic adenosine monophosphate (cAMP). The production of cAMP is necessary to maintain the integrity of the blood-brain barrier (BBB)[99], and excessive levels of cGMP are associated with anxiety and depression[100]. DOCK1 is a protein belonging to the dedicator of cytokinesis (DOCK) family of guanine exchange factors (GEFs) involved in activation of G proteins. DOCK1 is involved in neuronal development and angiogenesis[101]. The difference in expression of *Pde1c* and *Dock1* in the endothelial cells may indicate a difference in permeability of the BBB within the hippocampus. It is also worth noting that a number of other phosphodiesterase and DOCK genes were identified as DEGs in endothelial cells, with a largely uniform direction of difference. This suggests a consistent differential effect of these microbiomes on two major mechanisms of regulating intracellular signal transduction (i.e., cyclic nucleotide generation and GEF activity) across a sizeable range of surface receptors, including GPCRs. GPCRs are among the largest classes of receptors and common drug targets[102], responding to neurotransmitters, hormones, and a wide range of sensory cues. With widespread and strong expression in the CNS[103],

GPCRs are also broadly expressed by enteroendocrine cells[104], vagal efferents[105], and other cells throughout the gut[106].

Another gene of interest found to be differentially expressed in astrocytes is the gene encoding for brain-derived neurotrophic factor (BDNF). BDNF over-expression leads to a decrease in anxiety-related behavior in mice[107]. However, we observed relatively greater *Bdnf* expression in GM^{Low} and CF^{High} mice which were found to have relatively increased anxiety-related behavior, and these findings therefore need to be explored further.

Interestingly, CellChat analysis detected coordinated expression of genes involved in orexigenic pathways including *Ghrelin*, *Npr*, and *Qrfp* only in GM^{Low} and CF^{High} mice, the two groups with greater BW. These differences in intercellular signaling may provide an explanation for the observed differences in feed intake. These findings highlight cell type specific differences in hippocampal gene expression in genes identified within DMRs and genes in pathways associated with growth and feeding behavior, giving a possible reason for the increased weight noted in these two groups.

We recognize a number of limitations within this study. For example, methylome analysis was restricted to samples from female offspring and their dams due to resource constraints. Similarly, methylome and transcriptome analyses were limited to hippocampal tissue. While these data provide strong proof-of-principle and demonstrate the utility of the experimental model, additional work is needed to determine whether the observed differences in hippocampal methylation and gene expression are conserved in other regions of the CNS or even other tissues. Lastly, the parental generation was represented by the dams in all analyses presented here. There has been growing evidence for paternal programming of offspring through the epigenome[96, 98], and additional studies are needed to determine whether the observed effects are due to pre- or post-fertilization events..

Conclusion

In total, the findings presented here demonstrate that features within healthy native GMs exert an intergenerational effect on offspring behavior, growth, DNA methylation, and gene expression within the central nervous system, and strongly suggest a relationship between these factors during fetal development. Moreover, these findings implicate bile acids as potential mediators of these effects, including changes in GPCR signal transduction and pathways involved in feeding behavior.

List of abbreviations

AD – Anxiety disorders
BSH – Bile salt hydrolase enzyme
bsh – Bile salt hydrolase gene
CF – Cross-fostered offspring
CNS – Central Nervous System
DEGs – Differentially expressed genes
DMRs – Differentially methylated regions
EPM – Elevated plus maze test
GM – Gut microbiome
LDT – Light/dark transition test
OFE – Open-field exploration test
OO – Obesity/Overweight
snRNA-seq – Single nuclei RNA sequencing

Declarations

Ethics Approval and consent to participate

All activities described here were performed in accordance with the guidelines put forth in the Guide for the Care and Use of Laboratory Animals and were approved by the Institutional Animal Care and Use Committee (IACUC) of the University of Missouri, an AAALAC accredited institution.

Consent for publication

Not applicable

Availability of data and material

All data supporting our analyses and reported conclusions have been deposited in the appropriate data repositories and are publicly available. Metagenomic, metatranscriptomic, 16S rRNA amplicon sequencing, and single nuclei RNAseq data are available at the National Center for Biotechnology and Informatics (NCBI) Sequence Read Archive (SRA) under the BioProject accession number PRJNA885816. Mouse Methylation BeadChip array data are available at the Gene Expression Omnibus (GEO) under accession GSE239371.

Competing Interests

The authors declare no competing interests.

Funding

This project was funded by NIH R03 OD028259-01 (Ericsson). Generation of all CD-1 mice used in this study was supported by NIH U42 OD010918. KG was also supported by NIH T32 OD011126 and the Joseph Wagner Fellowship Endowment in Laboratory Animal Medicine. ZM was also supported by NIH T32 GM008396.

Authors and Affiliations

Department of Veterinary Pathobiology, College of Veterinary Medicine, University of Missouri, Columbia, MO, 65201, USA

Kevin L. Gustafson, Zachary L. McAdams, Daniel J. Davis, Lyndon M. Coghill, James Amos-Landgraf, Craig L. Franklin, and Aaron C. Ericsson

University of Missouri College of Veterinary Medicine, University of Missouri, Columbia, MO, 65211, USA

Kevin L. Gustafson, Rachael E. McCorkle, James Amos-Landgraf, Craig L. Franklin, and Aaron C. Ericsson

Department of Life Sciences and Medicine, Faculty of Science, Technology and Medicine, University of Luxembourg, L-4362 Esch-sur-Alzette, Luxembourg

Susheel Bhanu Busi and Paul Wilmes

UK Centre for Ecology and Hydrology, Wallingford, Oxfordshire, OX10 8BB, UK

Susheel Bhanu Busi

Department of Chemistry, College of Arts and Science, University of Missouri, Columbia, MO, 65211, USA

Pavlo Khodakivskyi and Elena A. Goun

University of Missouri Genomics Technology Core, University of Missouri, Columbia, MO, 65211, USA

Nathan J. Bivens

University of Missouri Animal Modeling Core, University of Missouri, Columbia, MO, 65201, USA

Daniel J. Davis

University of Missouri Bioinformatics and Analytics Core, University of Missouri, Columbia, MO, 65211, USA

Murugesan Raju and Lyndon M. Coghill

University of Missouri Mutant Mouse Resource and Research Center (MMRRC), University of Missouri, Columbia, MO, 65201, USA

Kevin L. Gustafson, Zachary L. McAdams, James Amos-Landgraf, Craig L. Franklin, and Aaron C. Ericsson

University of Missouri Metagenomics Center, University of Missouri, Columbia, MO, 65201, USA

Aaron C. Ericsson

Department of Child Health & Obstetrics, Gynecology, and Women's Health, School of Medicine, University of Missouri, Columbia, MO, 65212, USA

Rene Cortese

Author contributions

K.L.G and A.C.E conceived the project and wrote the manuscript. K.L.G., Z.L.M., R.E.M., P.K., E.A.G. performed experiments and analyzed data. Z.L.M. assisted in microbiome bioinformatic analysis, and figure generation and editing. S.B.B., Z.L.M, M.R, and L.M.C performed the snRNAseq bioinformatic analysis. R.C. assisted in designing and collecting data for the methylome analysis experiment, and performed the methylome bioinformatic analysis. N.J.B. assisted in designing and collecting data for the snRNAseq experiment. P.W. contributed resources and assisted with snRNAseq analysis. D.J.D. assisted in design and performance of qRT-PCR assays. C.L.F. and J.A.L. were responsible for generating and maintaining the GM^{Low} and GM^{High} colonies of CD-1 mice used in the experiments. K.L.G., S.B.B., Z.L.M., R.E.M., P.K.,

N.J.B., D.J.D., M.R., L.M.C., E.A.G., J.A.L., C.L.F., P.W., R.C., and A.C.E. all assisted in editing the manuscript.

Corresponding Author

ACE at ericssona@missouri.edu

Acknowledgements

The authors would like to thank the University of Iowa metabolomics core for assisting with metabolomic data acquisition and analysis. The authors would also like to thank the University of Missouri Mutant Mouse Resource and Research Center for donating the CD-1 mice used in this study. The authors would also like to thank Rebecca Dorfmeyer for prepping fecal DNA samples for 16S rRNA amplicon analysis.

References

1. Organization WH. World mental health report: Transforming mental health for all: World Health Organization, 2022.
2. Freidl EK, Stroeh OM, Elkins RM, et al. Assessment and Treatment of Anxiety Among Children and Adolescents. Focus (Am Psychiatr Publ) 2017;15:144-156.
3. Prevention CfDCa. Anxiety and depression in children: Get the facts. Children's Mental Health. Volume 2023: cdc.gov, 2023.
4. Prevention CfDCa. Prevalence of Childhood Obesity in the United States. Overweight and Obesity. Volume 2023. cdc.gov, 2023.
5. Gariepy G, Nitka D, Schmitz N. The association between obesity and anxiety disorders in the population: a systematic review and meta-analysis. Int J Obes (Lond) 2010;34:407-19.
6. Grundy A, Cotterchio M, Kirsh VA, et al. Associations between anxiety, depression, antidepressant medication, obesity and weight gain among Canadian women. PLoS One 2014;9:e99780.
7. Lindberg L, Hagman E, Danielsson P, et al. Anxiety and depression in children and adolescents with obesity: a nationwide study in Sweden. BMC Med 2020;18:30.
8. Sharafi SE, Garmaroudi G, Ghafouri M, et al. Prevalence of anxiety and depression in patients with overweight and obesity. Obesity Medicine 2020;17:1-5.
9. Wang S, Sun Q, Zhai L, et al. The Prevalence of Depression and Anxiety Symptoms among Overweight/Obese and Non-Overweight/Non-Obese Children/Adolescents in China: A Systematic Review and Meta-Analysis. Int J Environ Res Public Health 2019;16.
10. Bercik P, Denou E, Collins J, et al. The intestinal microbiota affect central levels of brain-derived neurotropic factor and behavior in mice. Gastroenterology 2011;141:599-609, 609 e1-3.
11. Bravo JA, Forsythe P, Chew MV, et al. Ingestion of Lactobacillus strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. Proc Natl Acad Sci U S A 2011;108:16050-5.

1010 12. Buffington SA, Di Prisco GV, Auchtung TA, et al. Microbial Reconstitution Reverses Maternal Diet-
1011 Induced Social and Synaptic Deficits in Offspring. Cell 2016;165:1762-1775.
1012 13. Hsiao EY, McBride SW, Hsien S, et al. Microbiota modulate behavioral and physiological
1013 abnormalities associated with neurodevelopmental disorders. Cell 2013;155:1451-63.
1014 14. Catron TR, Swank A, Wehmas LC, et al. Microbiota alter metabolism and mediate
1015 neurodevelopmental toxicity of 17beta-estradiol. Sci Rep 2019;9:7064.
1016 15. Kim E, Paik D, Ramirez RN, et al. Maternal gut bacteria drive intestinal inflammation in offspring
1017 with neurodevelopmental disorders by altering the chromatin landscape of CD4(+) T cells.
1018 Immunity 2022;55:145-158 e7.
1019 16. Cho HM, Gonzalez-Ortiz G, Melo-Duran D, et al. Stimbiotic supplementation improved
1020 performance and reduced inflammatory response via stimulating fiber fermenting microbiome
1021 in weaner pigs housed in a poor sanitary environment and fed an antibiotic-free low zinc oxide
1022 diet. PLoS One 2020;15:e0240264.
1023 17. Everard A, Belzer C, Geurts L, et al. Cross-talk between Akkermansia muciniphila and intestinal
1024 epithelium controls diet-induced obesity. Proc Natl Acad Sci U S A 2013;110:9066-71.
1025 18. Lee H, Lee Y, Kim J, et al. Modulation of the gut microbiota by metformin improves metabolic
1026 profiles in aged obese mice. Gut Microbes 2018;9:155-165.
1027 19. Rodrigues RR, Greer RL, Dong X, et al. Antibiotic-Induced Alterations in Gut Microbiota Are
1028 Associated with Changes in Glucose Metabolism in Healthy Mice. Front Microbiol 2017;8:2306.
1029 20. Desbonnet L, Clarke G, Shanahan F, et al. Microbiota is essential for social development in the
1030 mouse. Mol Psychiatry 2014;19:146-8.
1031 21. Diaz Heijtz R, Wang S, Anuar F, et al. Normal gut microbiota modulates brain development and
1032 behavior. Proc Natl Acad Sci U S A 2011;108:3047-52.
1033 22. Rabot S, Membrez M, Bruneau A, et al. Germ-free C57BL/6J mice are resistant to high-fat-diet-
1034 induced insulin resistance and have altered cholesterol metabolism. FASEB J 2010;24:4948-59.
1035 23. Collins SM, Kassam Z, Bercik P. The adoptive transfer of behavioral phenotype via the intestinal
1036 microbiota: experimental evidence and clinical implications. Curr Opin Microbiol 2013;16:240-5.
1037 24. De Palma G, Lynch MD, Lu J, et al. Transplantation of fecal microbiota from patients with irritable
1038 bowel syndrome alters gut function and behavior in recipient mice. Sci Transl Med 2017;9.
1039 25. Turnbaugh PJ, Ley RE, Mahowald MA, et al. An obesity-associated gut microbiome with
1040 increased capacity for energy harvest. Nature 2006;444:1027-31.
1041 26. Bhagavata Srinivasan SP, Raipuria M, Bahari H, et al. Impacts of Diet and Exercise on Maternal
1042 Gut Microbiota Are Transferred to Offspring. Front Endocrinol (Lausanne) 2018;9:716.
1043 27. Bruce-Keller AJ, Fernandez-Kim S-O, Townsend RL, et al. Maternal obese-type gut microbiota
1044 differentially impact cognition, anxiety and compulsive behavior in male and female offspring in
1045 mice. PLOS ONE 2017;12:e0175577.
1046 28. Pessa-Morikawa T, Husso A, Karkkainen O, et al. Maternal microbiota-derived metabolic profile
1047 in fetal murine intestine, brain and placenta. BMC Microbiol 2022;22:46.
1048 29. Deng X, Guarita DR, Pedroso MR, et al. PYY inhibits CCK-stimulated pancreatic secretion through
1049 the area postrema in unanesthetized rats. Am J Physiol Regul Integr Comp Physiol
1050 2001;281:R645-53.
1051 30. Cabral A, Cornejo MP, Fernandez G, et al. Circulating Ghrelin Acts on GABA Neurons of the Area
1052 Postrema and Mediates Gastric Emptying in Male Mice. Endocrinology 2017;158:1436-1449.
1053 31. Yamamoto H, Kishi T, Lee CE, et al. Glucagon-like peptide-1-responsive catecholamine neurons in
1054 the area postrema link peripheral glucagon-like peptide-1 with central autonomic control sites. J
1055 Neurosci 2003;23:2939-46.

1056 32. Arslanova A, Tarasova A, Alexandrova A, et al. Protective Effects of Probiotics on Cognitive and
1057 Motor Functions, Anxiety Level, Visceral Sensitivity, Oxidative Stress and Microbiota in Mice with
1058 Antibiotic-Induced Dysbiosis. Life (Basel) 2021;11.
1059 33. O'Connor R, Moloney GM, Fulling C, et al. Maternal antibiotic administration during a critical
1060 developmental window has enduring neurobehavioural effects in offspring mice. Behav Brain
1061 Res 2021;404:113156.
1062 34. Sun Y, Zhu H, Cheng R, et al. Outer membrane protein Amuc_1100 of Akkermansia muciniphila
1063 alleviates antibiotic-induced anxiety and depression-like behavior in mice. Physiol Behav
1064 2023;258:114023.
1065 35. Bruce-Keller AJ, Fernandez-Kim SO, Townsend RL, et al. Maternal obese-type gut microbiota
1066 differentially impact cognition, anxiety and compulsive behavior in male and female offspring in
1067 mice. PLoS One 2017;12:e0175577.
1068 36. Cheatham CN, Gustafson KL, McAdams ZL, et al. Standardized Complex Gut Microbiomes
1069 Influence Fetal Growth, Food Intake, and Adult Body Weight in Outbred Mice. Microorganisms
1070 2023;11.
1071 37. Ericsson AC, Hart ML, Kwan J, et al. Supplier-origin mouse microbiomes significantly influence
1072 locomotor and anxiety-related behavior, body morphology, and metabolism. Commun Biol
1073 2021;4:716.
1074 38. Wolfe AE, Moskowitz JE, Franklin CL, et al. Interactions of Segmented Filamentous Bacteria
1075 (Candidatus Savagella) and bacterial drivers in colitis-associated colorectal cancer development.
1076 PLoS One 2020;15:e0236595.
1077 39. Walters WA, Caporaso JG, Lauber CL, et al. PrimerProspector: de novo design and taxonomic
1078 analysis of barcoded polymerase chain reaction primers. Bioinformatics 2011;27:1159-61.
1079 40. Caporaso JG, Lauber CL, Walters WA, et al. Global patterns of 16S rRNA diversity at a depth of
1080 millions of sequences per sample. Proc Natl Acad Sci U S A 2011;108 Suppl 1:4516-22.
1081 41. Loy A, Maixner F, Wagner M, et al. probeBase--an online resource for rRNA-targeted
1082 oligonucleotide probes: new features 2007. Nucleic Acids Res 2007;35:D800-4.
1083 42. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads.
1084 EMBnet.journal 2011;17:10-12.
1085 43. Bolyen E, Rideout JR, Dillon MR, et al. Reproducible, interactive, scalable and extensible
1086 microbiome data science using QIIME 2. Nat Biotechnol 2019;37:852-857.
1087 44. Callahan BJ, McMurdie PJ, Rosen MJ, et al. DADA2: High-resolution sample inference from
1088 Illumina amplicon data. Nat Methods 2016;13:581-3.
1089 45. Pruesse E, Quast C, Knittel K, et al. SILVA: a comprehensive online resource for quality checked
1090 and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res
1091 2007;35:7188-96.
1092 46. Asnicar F, Weingart G, Tickle TL, et al. Compact graphical representation of phylogenetic data
1093 and metadata with GraPhlAn. PeerJ 2015;3:e1029.
1094 47. Zachary L McAdams SBB, Kevin L Gustafson, Nathan Bivens, Craig L Franklin, Paul Wilmes, Aaron
1095 C Ericsson. Multi-omics analysis of mouse fecal microbiome reveals supplier-dependent
1096 functional differences and novel metagenome-assembled genomes. bioRxiv 2022.
1097 48. Khodakivskyi PV, Lauber CL, Yevtodiyenko A, et al. Noninvasive imaging and quantification of bile
1098 salt hydrolase activity: From bacteria to humans. Sci Adv 2021;7.
1099 49. Vorobyeva AG, Stanton M, Godinat A, et al. Development of a Bioluminescent Nitroreductase
1100 Probe for Preclinical Imaging. PLoS One 2015;10:e0131037.
1101 50. Cantor JR, Abu-Remaileh M, Kanarek N, et al. Physiologic Medium Rewires Cellular Metabolism
1102 and Reveals Uric Acid as an Endogenous Inhibitor of UMP Synthase. Cell 2017;169:258-272.e17.

1103 51. Li B, Tang J, Yang Q, et al. NOREVA: normalization and evaluation of MS-based metabolomics
1104 data. Nucleic Acids Res 2017;45:W162-w170.
1105 52. Zhou W, Hinoue T, Barnes B, et al. DNA methylation dynamics and dysregulation delineated by
1106 high-throughput profiling in the mouse. Cell Genom 2022;2.
1107 53. Xu Z, Niu L, Li L, et al. ENmix: a novel background correction method for Illumina
1108 HumanMethylation450 BeadChip. Nucleic Acids Res 2016;44:e20.
1109 54. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive Bioconductor
1110 package for the analysis of Infinium DNA methylation microarrays. Bioinformatics 2014;30:1363-
1111 9.
1112 55. Xu Z, Langie SA, De Boever P, et al. RELIC: a novel dye-bias correction method for Illumina
1113 Methylation BeadChip. BMC Genomics 2017;18:4.
1114 56. Wolock SL, Lopez R, Klein AM. Scrublet: Computational Identification of Cell Doublets in Single-
1115 Cell Transcriptomic Data. Cell Syst 2019;8:281-291 e9.
1116 57. Stuart T, Butler A, Hoffman P, et al. Comprehensive Integration of Single-Cell Data. Cell
1117 2019;177:1888-1902 e21.
1118 58. Team RDC. R: A Language and Environment for Statistical Computing. 2010.
1119 59. Jin S, Guerrero-Juarez CF, Zhang L, et al. Inference and analysis of cell-cell communication using
1120 CellChat. Nat Commun 2021;12:1088.
1121 60. Suoqin J, Maksim VP, Qing N. CellChat for systematic analysis of cell-cell communication from
1122 single-cell and spatially resolved transcriptomics. bioRxiv 2023:2023.11.05.565674.
1123 61. Hammer Ø, Harper DA. Past: paleontological statistics software package for educaton and data
1124 anlysis. Palaeontologia electronica 2001;4:1.
1125 62. Dixon P. VEGAN, A Package of R Functions for Community Ecology. Journal of Vegetation Science
1126 2003;14:927-930.
1127 63. Ericsson AC, Davis JW, Spollen W, et al. Effects of vendor and genetic background on the
1128 composition of the fecal microbiota of inbred mice. PLoS One 2015;10:e0116704.
1129 64. Song Z, Cai Y, Lao X, et al. Taxonomic profiling and populational patterns of bacterial bile salt
1130 hydrolase (BSH) genes based on worldwide human gut microbiome. Microbiome 2019;7:9.
1131 65. Abdulrab S, Al-Maweri S, Halboub E. Ursodeoxycholic acid as a candidate therapeutic to alleviate
1132 and/or prevent COVID-19-associated cytokine storm. Med Hypotheses 2020;143:109897.
1133 66. Adhikari A, Topiwala MA, Gordon JA. Synchronized activity between the ventral hippocampus
1134 and the medial prefrontal cortex during anxiety. Neuron 2010;65:257-69.
1135 67. Baksh RA, Ritchie CW, Terrera GM, et al. The association between anxiety disorders and
1136 hippocampal volume in older adults. Psychol Aging 2021;36:288-297.
1137 68. Jimenez JC, Su K, Goldberg AR, et al. Anxiety Cells in a Hippocampal-Hypothalamic Circuit.
1138 Neuron 2018;97:670-683 e6.
1139 69. Mineur YS, Obayemi A, Wigestrand MB, et al. Cholinergic signaling in the hippocampus regulates
1140 social stress resilience and anxiety- and depression-like behavior. Proc Natl Acad Sci U S A
1141 2013;110:3573-8.
1142 70. Revest JM, Dupret D, Koehl M, et al. Adult hippocampal neurogenesis is involved in anxiety-
1143 related behaviors. Mol Psychiatry 2009;14:959-67.
1144 71. Satpute AB, Mumford JA, Naliboff BD, et al. Human anterior and posterior hippocampus respond
1145 distinctly to state and trait anxiety. Emotion 2012;12:58-68.
1146 72. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with
1147 increased coverage, supporting functional discovery in genome-wide experimental datasets.
1148 Nucleic Acids Res 2019;47:D607-D613.
1149 73. Vuong HE, Pronovost GN, Williams DW, et al. The maternal microbiome modulates fetal
1150 neurodevelopment in mice. Nature 2020;586:281-286.

1151 74. Kimura I, Miyamoto J, Ohue-Kitano R, et al. Maternal gut microbiota in pregnancy influences
1152 offspring metabolic phenotype in mice. Science 2020;367.
1153 75. Husso A, Pessa-Morikawa T, Koistinen VM, et al. Impacts of maternal microbiota and microbial
1154 metabolites on fetal intestine, brain, and placenta. BMC Biol 2023;21:207.
1155 76. Lee-Sarwar KA, Chen YC, Chen YY, et al. The maternal prenatal and offspring early-life gut
1156 microbiome of childhood asthma phenotypes. Allergy 2023;78:418-428.
1157 77. Sun Z, Lee-Sarwar K, Kelly RS, et al. Revealing the importance of prenatal gut microbiome in
1158 offspring neurodevelopment in humans. EBioMedicine 2023;90:104491.
1159 78. Seneviratne SN, Rajindrajith S. Fetal programming of obesity and type 2 diabetes. World J
1160 Diabetes 2022;13:482-497.
1161 79. McAdams ZL, Gustafson KL, Russell AL, et al. Supplier-origin gut microbiomes affect host body
1162 weight and select autism-related behaviors. Gut Microbes 2024;16.
1163 80. Martini E, Krug SM, Siegmund B, et al. Mend Your Fences: The Epithelial Barrier and its
1164 Relationship With Mucosal Immunity in Inflammatory Bowel Disease. Cell Mol Gastroenterol
1165 Hepatol 2017;4:33-46.
1166 81. Kaelberer MM, Rupprecht LE, Liu WW, et al. Neuropod Cells: The Emerging Biology of Gut-Brain
1167 Sensory Transduction. Annu Rev Neurosci 2020;43:337-353.
1168 82. Hellmuth C, Lindsay KL, Uhl O, et al. Maternal Metabolomic Profile and Fetal Programming of
1169 Offspring Adiposity: Identification of Potentially Protective Lipid Metabolites. Mol Nutr Food Res
1170 2019;63:e1700889.
1171 83. Nagahashi M, Takabe K, Liu R, et al. Conjugated bile acid-activated S1P receptor 2 is a key
1172 regulator of sphingosine kinase 2 and hepatic gene expression. Hepatology 2015;61:1216-26.
1173 84. MahmoudianDehkordi S, Bhattacharyya S, Brydges CR, et al. Gut Microbiome-Linked Metabolites
1174 in the Pathobiology of Major Depression With or Without Anxiety-A Role for Bile Acids. Front
1175 Neurosci 2022;16:937906.
1176 85. Chen S, Shao Q, Chen J, et al. Bile acid signalling and its role in anxiety disorders. Front
1177 Endocrinol (Lausanne) 2023;14:1268865.
1178 86. Weng H, Deng L, Wang T, et al. Humid heat environment causes anxiety-like disorder via
1179 impairing gut microbiota and bile acid metabolism in mice. Nat Commun 2024;15:5697.
1180 87. Wu L, Han Y, Zheng Z, et al. Obeticholic Acid Inhibits Anxiety via Alleviating Gut Microbiota-
1181 Mediated Microglia Accumulation in the Brain of High-Fat High-Sugar Diet Mice. Nutrients
1182 2021;13.
1183 88. Tao Y, Zhou H, Li Z, et al. TGR5 deficiency-induced anxiety and depression-like behaviors: The role
1184 of gut microbiota dysbiosis. J Affect Disord 2024;344:219-232.
1185 89. Wang H, Tan YZ, Mu RH, et al. Takeda G Protein-Coupled Receptor 5 Modulates Depression-like
1186 Behaviors via Hippocampal CA3 Pyramidal Neurons Afferent to Dorsolateral Septum. Biol
1187 Psychiatry 2021;89:1084-1095.
1188 90. Kamp F, Hamilton JA, Kamp F, et al. Movement of fatty acids, fatty acid analogues, and bile acids
1189 across phospholipid bilayers. Biochemistry 1993;32:11074-86.
1190 91. Abdo Qaid EY, Abdullah Z, Zakaria R, et al. Minocycline Attenuates Lipopolysaccharide-Induced
1191 Locomotor Deficit and Anxiety-like Behavior and Related Expression of the BDNF/CREB Protein in
1192 the Rat Medial Prefrontal Cortex (mPFC). Int J Mol Sci 2022;23.
1193 92. Obata Y, Furusawa Y, Endo TA, et al. The epigenetic regulator Uhrf1 facilitates the proliferation
1194 and maturation of colonic regulatory T cells. Nat Immunol 2014;15:571-9.
1195 93. Ansari I, Raddatz G, Gutekunst J, et al. The microbiota programs DNA methylation to control
1196 intestinal homeostasis and inflammation. Nat Microbiol 2020;5:610-619.
1197 94. Ramos-Molina B, Sanchez-Alcoholado L, Cabrera-Mulero A, et al. Gut Microbiota Composition Is
1198 Associated With the Global DNA Methylation Pattern in Obesity. Front Genet 2019;10:613.

95. Santini L, Halbritter F, Titz-Teixeira F, et al. Genomic imprinting in mouse blastocysts is predominantly associated with H3K27me3. Nat Commun 2021;12:3804.
96. Argaw-Denboba A, Schmidt TSB, Di Giacomo M, et al. Paternal microbiome perturbations impact offspring fitness. Nature 2024;629:652-659.
97. Kilama J, Dahlen CR, Reynolds LP, et al. Contribution of the seminal microbiome to paternal programming. Biol Reprod 2024;111:242-268.
98. Masson BA, Kiridena P, Lu D, et al. Depletion of the paternal gut microbiome alters sperm small RNAs and impacts offspring physiology and behavior in mice. Brain Behav Immun 2024;123:290-305.
99. Radeva MY, Waschke J. Mind the gap: mechanisms regulating the endothelial barrier. Acta Physiol (Oxf) 2018;222.
100. Gutierrez-Rodelo C, Martinez-Tolibia SE, Morales-Figueroa GE, et al. Modulating cyclic nucleotides pathways by bioactive compounds in combatting anxiety and depression disorders. Mol Biol Rep 2023;50:7797-7814.
101. Hernandez-Vasquez MN, Adame-Garcia SR, Hamoud N, et al. Cell adhesion controlled by adhesion G protein-coupled receptor GPR124/ADGRA2 is mediated by a protein complex comprising intersectins and Elmo-Dock. J Biol Chem 2017;292:12178-12191.
102. Insel PA, Snead A, Murray F, et al. GPCR expression in tissues and cells: are the optimal receptors being used as drug targets? Br J Pharmacol 2012;165:1613-1616.
103. Hakak Y, Shrestha D, Goegel MC, et al. Global analysis of G-protein-coupled receptor signaling in human tissues. FEBS Lett 2003;550:11-7.
104. Smith CA, O'Flaherty EAA, Guccio N, et al. Single-cell transcriptomic atlas of enteroendocrine cells along the murine gastrointestinal tract. PLoS One 2024;19:e0308942.
105. Egerod KL, Petersen N, Timshel PN, et al. Profiling of G protein-coupled receptors in vagal afferents reveals novel gut-to-brain sensing mechanisms. Mol Metab 2018;12:62-75.
106. Ito J, Ito M, Nambu H, et al. Anatomical and histological profiling of orphan G-protein-coupled receptor expression in gastrointestinal tract of C57BL/6J mice. Cell Tissue Res 2009;338:257-69.
107. Quesseveur G, David D, Gaillard M, et al. BDNF overexpression in mouse hippocampal astrocytes promotes local neurogenesis and elicits anxiolytic-like activities. Translational psychiatry 2013;3:e253-e253.

Figure legends

Figure 1. Microbiomes linked to fetal programming of complex behaviors differ in synthesis or catabolism of several metabolites including bile acids. (A) Hierarchical clustering of data from adult GM^{Low}- or GM^{High}-colonized mice, demonstrating segregation of GMs by richness and composition (legend at right). Volcano plots showing metabolites enriched in GM^{Low} (red dots) or GM^{High} (blue dots) **(B)** feces or **(C)** serum. Horizontal dashed line indicates significance of $p<0.05$ between GM^{Low} and GM^{High} mice by Wilcox rank sum test. Normalized ileal expression of *Asbt* (**D**), *Ostβ* (**E**), *Gpbar1* (**F**), and *Fxr* (**G**) male (M) and female (F) mice colonized with GM^{Low} (red dots) or GM^{High} (blue dots). Normalized hepatic expression of *Cyp7a1* (**H**), *Slc10a1* (**I**), and *S1pr2* (**J**) in the same groups of mice. Gene expression

normalized to *β-Actin* expression. Results indicate main effect of GM in one- or two-way ANOVA, following correction for multiple tests. * $p<0.05$, ** $p<0.01$, **** $p<0.0001$. Abbreviations: 2-HBA (2-hydroxybutyrate), α-MCA (alpha-Muricholic Acid), β-ALA (beta-Alanine), BUTY (Butyrate), CA (Cholic Acid), CDCA (Chendeoxycholic Acid), CYS (Cysteine), DCA (Deoxycholic Acid), F6P (Fructose-6-Phosphate), G6P (Glucose-6-Phosphate), GCA (Glycocholic Acid), GLY (Glycerate), INO (Inositol), LAC (Lactate), LEU (Leucine), MVA (Mevalonic Acid), R5P (Ribulose-5-Phosphate), SUC (Succinate), TCA (Taurocholic Acid), TCDCA (Taurochenodeoxycholic Acid), TDCA (Taurodeoxycholic Acid), TLCA (Taurolithocholic Acid), UDCA (Ursodeoxycholic Acid).

Figure 2. Anxiety-related behavior and other outcomes in offspring are influenced by the maternal gut microbiome. (A) Schematic showing the experimental groups and timing of outcome measures. Behavior outcomes in male (M) and female (F) mice colonized with GM^{Low} or GM^{High}, including (**B**) time spent in the light portion of a light/dark test, (**C**) time spent in the open arms of an elevated plus maze, and distance run per day (Km) by (**D**) male and (**E**) female mice. (**F)** Body weight per mouse (grams) at day 50 in control mice. (**G-J)** Behavior outcomes in CF^{Low} and CF^{High} mice as those shown in panels **B-E**. **(K)** Body weight per mouse at day 50 in cross-fostered mice. p and F values represent effect of GM in two-way ANOVA with Tukey post hoc. Significant sex-dependent differences were detected in distance on running wheels and body weight in both control and cross-fostered mice, with no significant interactions between GM and sex in any tests. * $p<0.05$, ** $p<0.01$, **** $p<0.0001$

Figure 3. Maternal microbiome is associated with gene methylation in offspring hippocampus. (A) Dot plot showing normalized difference [Log_2(FC)] in mean methylation between GM^{Low} and GM^{High}, and between CF^{Low} and CF^{High}, of all CpG markers achieving a $Log_2(FC) > 1$ in either comparison. **(B)** Ratio of observed to predicted differentially methylated regions (DMRs), and locations of DMRs relative to gene bodies, on each chromosome. **(C)** Mean beta values of 27 CpG sites spanning the enhancers, promoters, introns and exons across the *Pde1c* gene in all four groups, with specific sites indicated above (upper); and log_2(FC) between control and cross-fostered groups (lower). Arrows on X-axis indicate direction of transcription. Numbers on the X-axis indicated number of CpG sites analyzed. **(D)** Mean beta values (upper) and Log_2(FC) between groups (lower) across adjacent regions of the positive (left) and negative (right) strand of a region of chromosome 14 containing genes for *Ang5*, *Ang6*, and other genes. **(E)** Protein interaction networks among products of DMRs, with node color indicating Log_2(FC) in methylation between GM^{Low} and GM^{High}.

Figure 4. Interaction networks constructed using differentially expressed genes (DEGs) identified in hippocampal endothelial and perivascular cells of control (**A**) and cross-fostered (**B**) mice. Highlighted nodes include DEGs identified in both comparisons and showing a pattern of fetal imprinting, and genes that are also differentially methylated (*Pde1c*, *Dock1*), or closely related (PDE cluster, *Dcc*, *Elmo1*, *Dock4*, *Dock5*). **c, d.** Log_2FC in expression of (**C**) Pde1c and (**D**) Dock1 in hippocampal cell clusters of male and female CF^{Low} and CF^{High} mice, as determined via real-time snRNAseq. **(E)** Venn diagram showing number of inferred cell-to-cell communication pathways identified in hippocampus of GM^{Low}, GM^{High}, CF^{Low}, and CF^{High} mice. Pathways listed above and below the diagram were selectively identified in the indicated groups. Bar charts showing the relative degree of cell-to-cell communication between (**F**) GM^{Low} and GM^{High}, and (**G**) CF^{Low}, and CF^{High} mice in the indicated pathways.

Supplementary Figures

Figure S1. Cladogram representing genera taxa identified within each phyla**.** Arrows indicate genus with statistically significant abundance differences**.** Wilcox rank sum test with Benjamini-Hochberg corrected p values. List of genera identified provided in **Table S3**.

Figure S2. Volcano plots showing the metabolites enriched in feces of GM^{Low}(red dots) and GM^{High} (blue dots) **(A)** males, and **(B)** females. Volcano plots showing the metabolites enriched in serum of GM^{Low}(red dots) and GM^{High} (blue dots) **(C)** males, and **(D)** females and Dot plots representing **(E)** gene and **(F)** the normalized (transcript/gene) expression in the feces of adult female mice colonized with GM^{Low} or GM^{High} (n = 3/GM). Box plots and individual data showing hydrolytic activity of bile salt hydrolase (BSH) in adult male and female mice colonized with GM^{Low} or GM^{High}, determined using five different bile acid-conjugated bioluminescent probes specific for **(G)** cholic acid (CA**)**, **(H)** deoxycholic acid (DCA), **(I)** chenodeoxycholic acid (CDCA), **(J)** ursodeoxycholic acid (UDCA), and (**K**) lithocholic acid (LCA). *$p<0.05$, **** $p<0.0001$. Abbreviations: 2-HGA (2-hydroxygluterate), 3-HPA (3-Hydroxyproprionate), α-KGA (alpha-Ketogluterate), α-KIC (alpha-Ketoisocaproate), ASP (Asparagine), β-ALA (beta-Alanine), CDCA (Chendeoxycholic Acid), CYS (Cysteine), DCA (Deoxycholic Acid), F6P (Fructose-6-Phosphate), G6P (Glucose-6-Phosphate), GLY (Glycerate), IAA (Indoleacetic Acid), ISO (Isocitrate), LAC (Lactate), LCA (Lithocholic Acid), LEU (Leucine), MAL (Malate), MET (Methionine), ORN (Ornithine), PHEN (Phenylalanine), PRO (Proline), PTA (Pantothenic Acid), R5P (Ribulose-5-Phosphate), SUC (Succinate), TDCA (Taurodeoxycholic Acid), TRY (Tryptamine), VAL (Valine).

Figure S3. Spearman correlation coefficient plots comparing relative abundance at the genera level and metabolite concentrations to the p value of the Spearman correlation. Spearman correlation coefficient is plotted along the x-axis for each metabolite, and -Log_{10}(p value) of the correlation between relative taxa abundance and metabolite concentration is plotted on the y-axis. Red dots represent genera with an increased relative abundance in GM^{Low}, and blue dots represent genera with an increased relative abundance in GM^{High}. Dotted lines in each plot represent statistical significance of $p<0.05$.

Figure S4. Dot and bar plots showing other results of behavior tests in adult male (M) and female (F) GM^{Low} and GM^{High} mice including (**A**) distance traveled in the light portion of the light/dark test and (**B**) open arms of the elevated plus maze. (**C**) Total distance traveled and (**D**) time spent in the center zone during OFE. (**E**) Body weight at 21 days of age of GM^{Low} and GM^{High} mice. (**F-J**) Same outcomes in CF^{Low} and CF^{High} mice as those shown in panels **A-D. (J)** Body weight at 21 days of age of CFLow and CFHigh mice**.** * $p<0.05$, ** $p<0.01$, **** $p<0.0001$.

Figure S5. Dot and bar plots showing differences in (**A**) richness and principal coordinate analysis (PCoA) plots showing differences in beta-diversity based on (**B**) Jaccard and (**C**) Bray-Curtis distances, between adult male (M) and female (F) mice colonized with GM^{Low} or GM^{High}, and similarities in all of the above metrics between CF mice at seven weeks of age and their cognate birth dams.

Figure S6. UMAP projections of hippocampal cell clusters in **(A)** control GM^{Low} and **(B)** GM^{High}, and cross-foster **(C)** CF^{Low} and **(D)** CF^{High} mice.

Figure S7. Dot plot correlation between the number of differentially expressed genes (DEGs) and the mean node degree determined by protein interaction analysis of those DEGs, in comparisons of **(A)** GM^{Low} and GM^{High} offspring and in **(B)** CF^{Low} and CF^{High} offspring. Marker shapes denote cell type including endothelial/perivascular cells (Endo, red squares); astrocytes and oligodendrocytes (Glial, grey diamonds); GABAergic neurons (GABA, dark blue triangles); and glutamatergic neurons (Gluta, brown triangles).

Figure S8. Interaction networks constructed using differentially expressed genes (DEGs) identified in hippocampal level five intratelencephalon (L5 IT)-projecting glutamatergic neurons from **(A)** control and **(B)** cross-fostered mice. Labeled nodes include DEGs identified in both comparisons and showing a pattern of fetal imprinting, and genes that are also differentially methylated (*Dock1*), or closely related (PDE cluster, *Adcy8*, *Dcc*, *Ntn1*).

Figure S9. Interaction networks constructed using differentially expressed genes (DEGs) identified in hippocampal astrocyte subsets from **(A)** control and **(B)** cross-fostered mice. Labeled nodes include DEGs identified in both comparisons and showing a pattern of fetal imprinting, and genes that are also differentially methylated (*Dock1*), or closely related (PDE cluster, *Adcy8*, *Dcc*, *Ntn1*).

Supplementary Tables

Table S1. Primer pairs used for qRT-PCR analysis**.**

Table S2. Relative abundance of family taxa harbored by either GM^{Low} or GM^{High}.

Table S3. Fecal metabolite concentrations from GM^{Low} and GM^{High} mice.

Table S4. Serum metabolite concentrations from GM^{Low} and GM^{High} mice.

Table S5. Relative abundance of identified genera in GM^{Low} and GM^{High} mice.

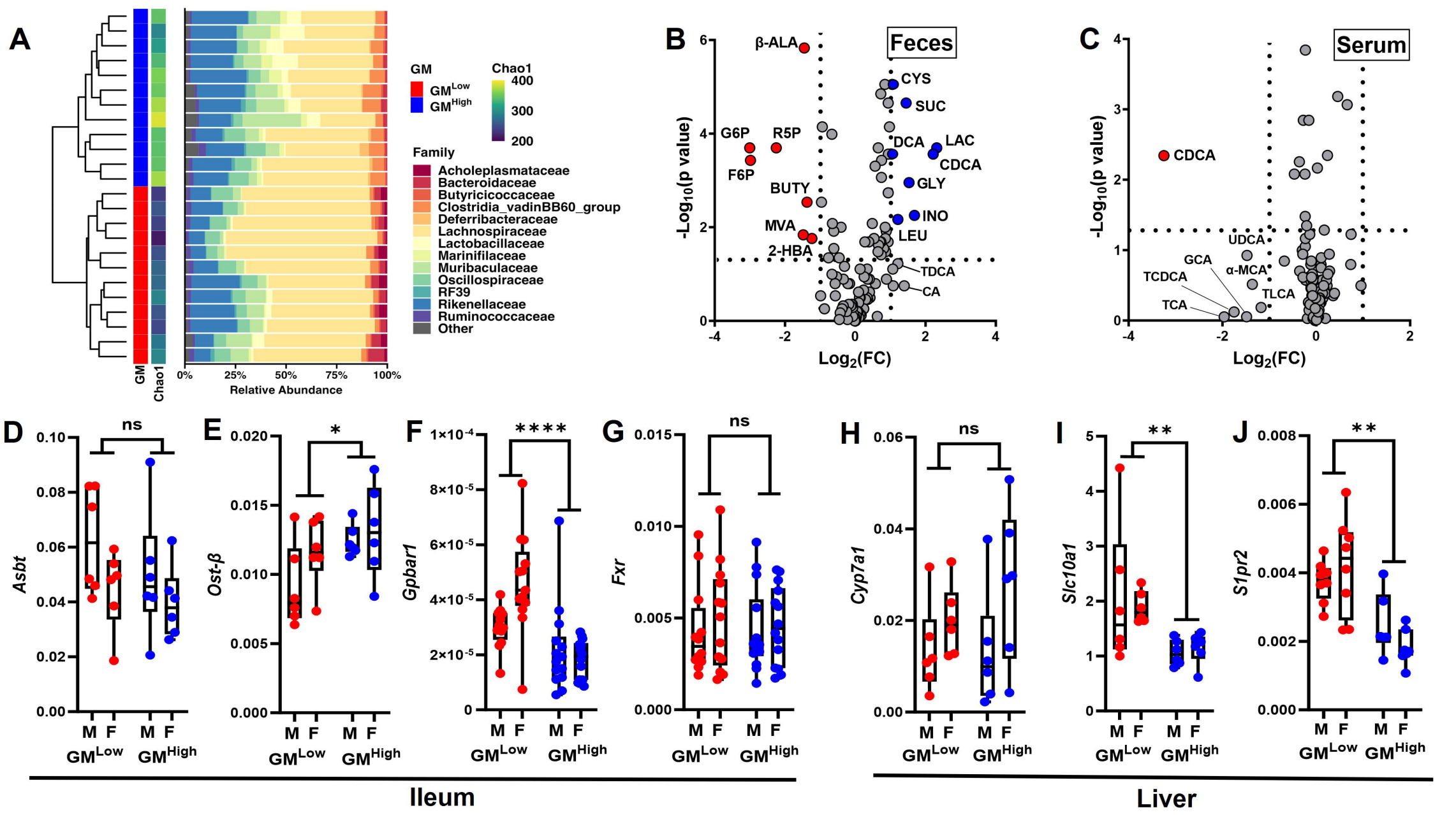
Table S6. Comparison of identified bacterial genera abundance to statistically significant fecal metabolite concentrations between GM^{Low} and GM^{High} mice.

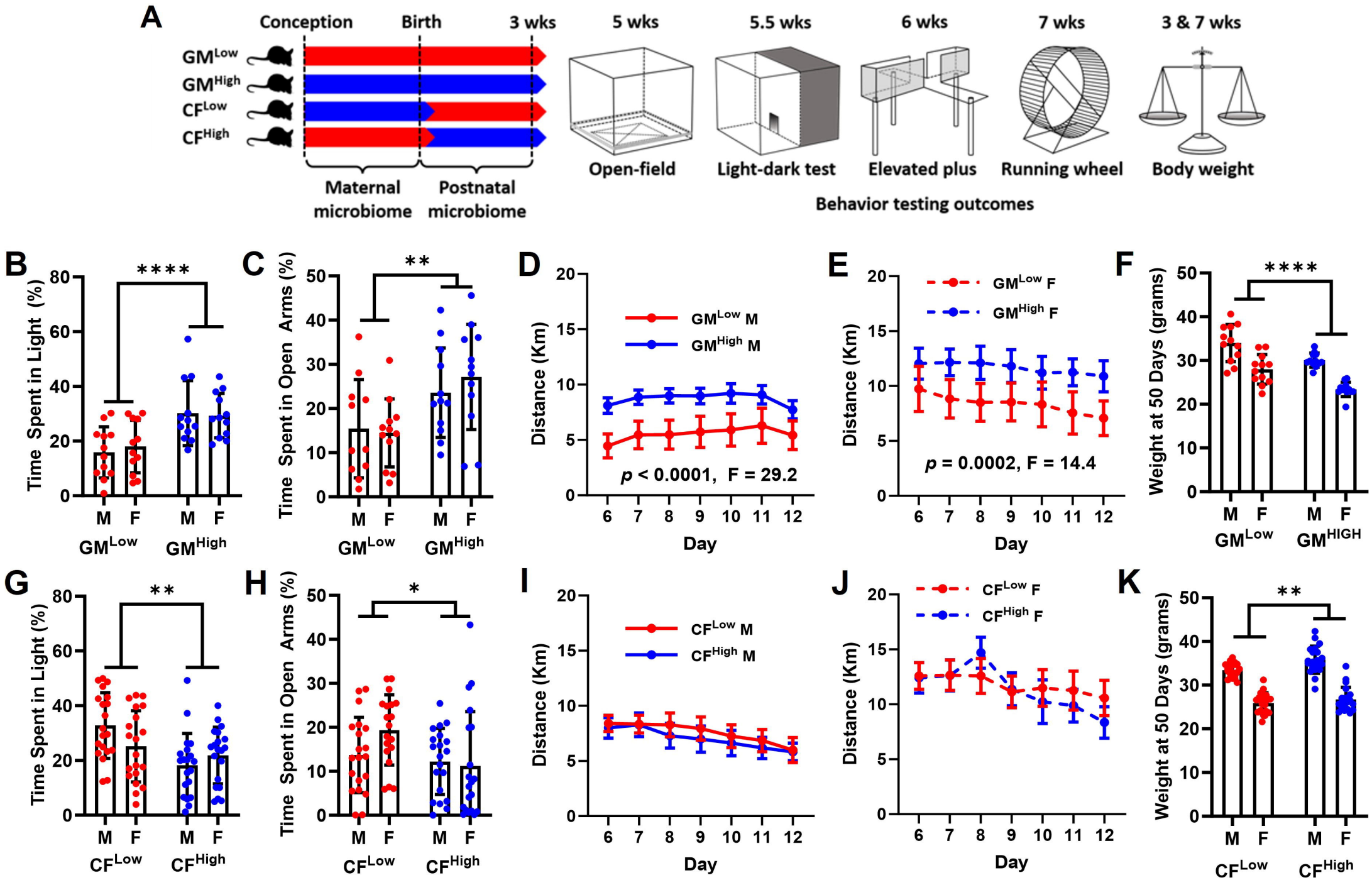
Table S7. Differentially methylated regions identified within the genomes of female GM^{Low} and GM^{High} mouse hippocampi.

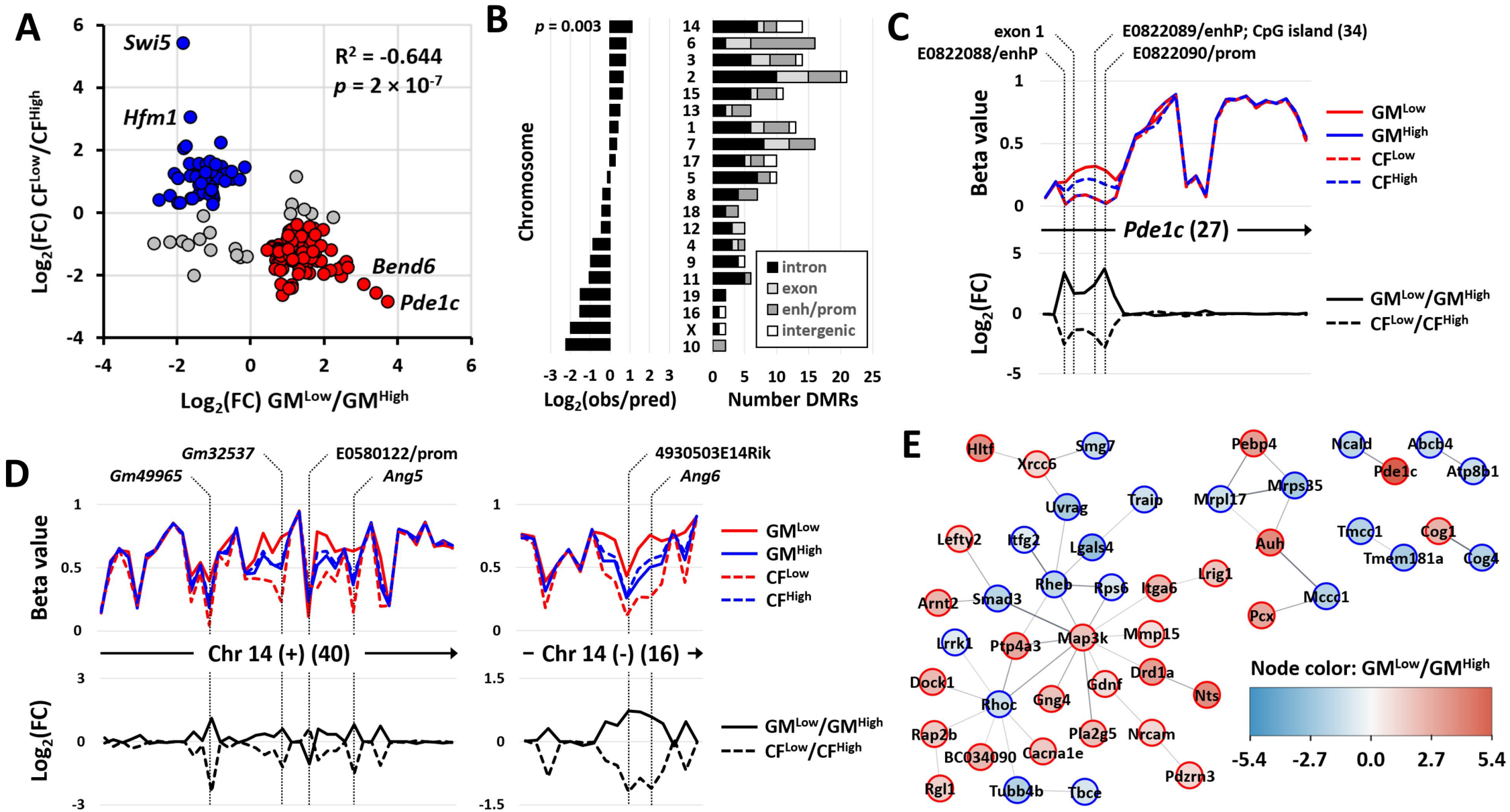
Table S8. Network STRING analysis results utilizing the 196 DMRs identified in the methylome analysis

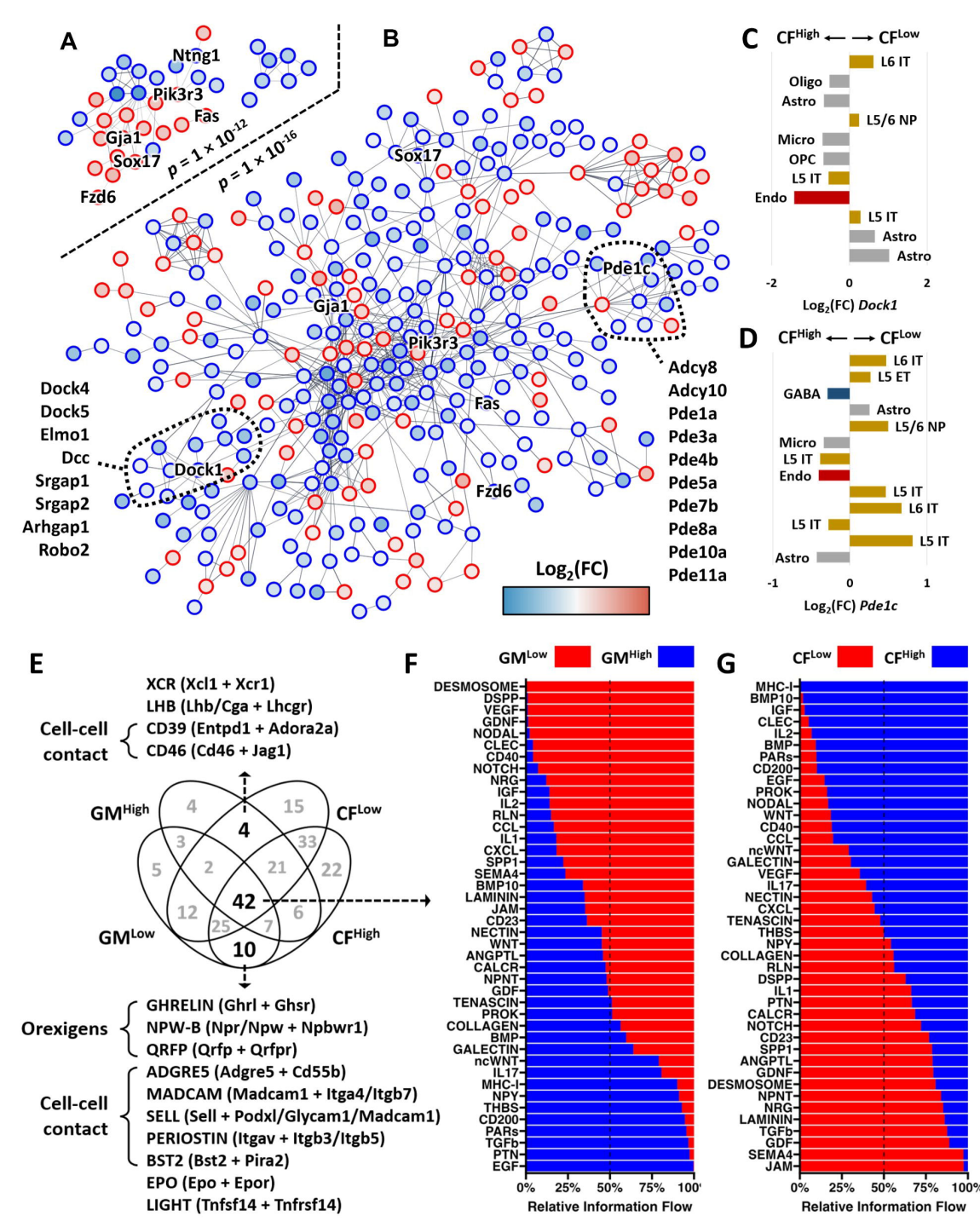
Table S9. Differentially expressed genes identified within the hippocampal cells of GM^{Low} and GM^{High} control male and female mice.

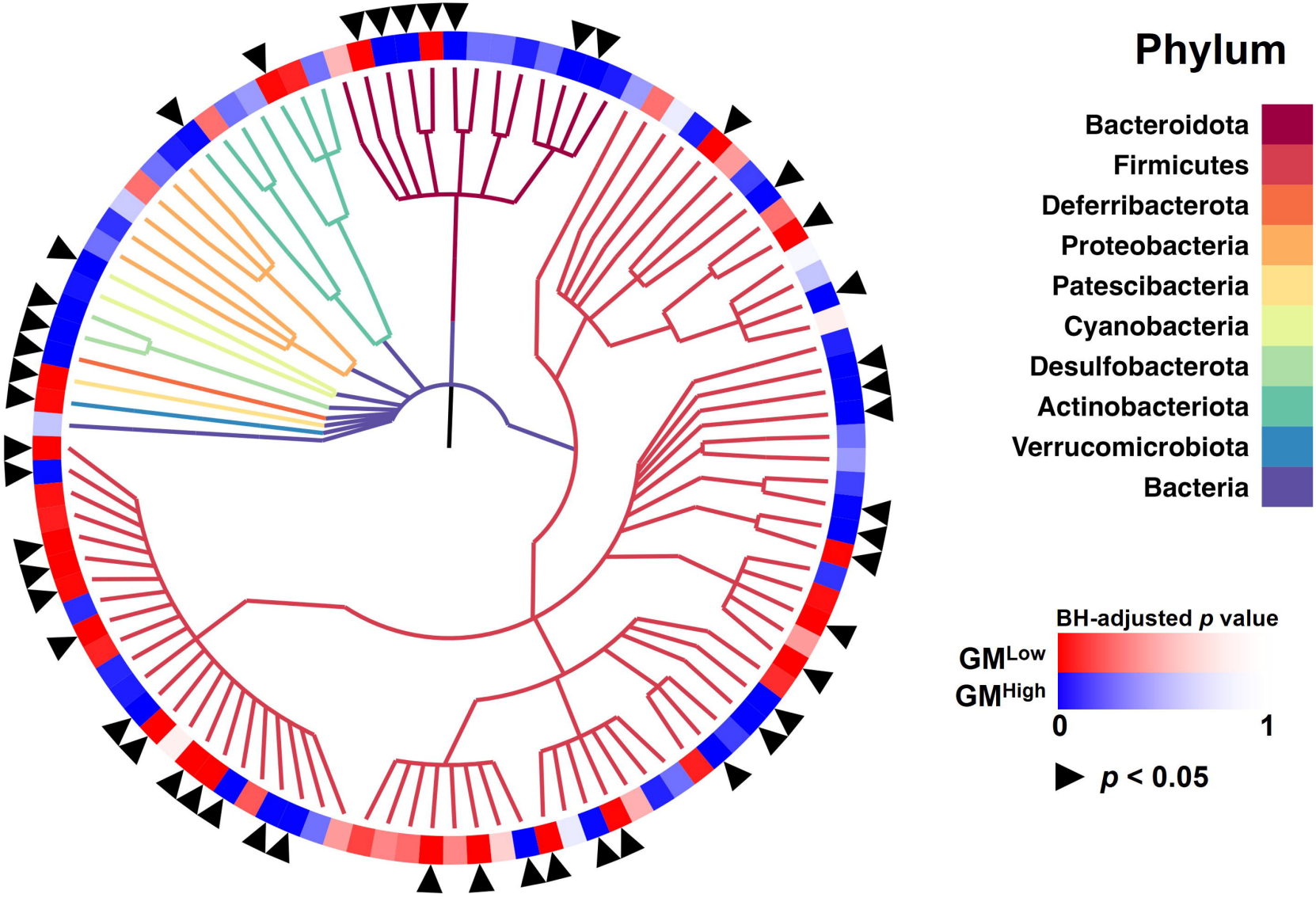
Table S10. Differentially expressed genes identified within the hippocampal cells of CF^{Low} and CF^{High} male and female mice.

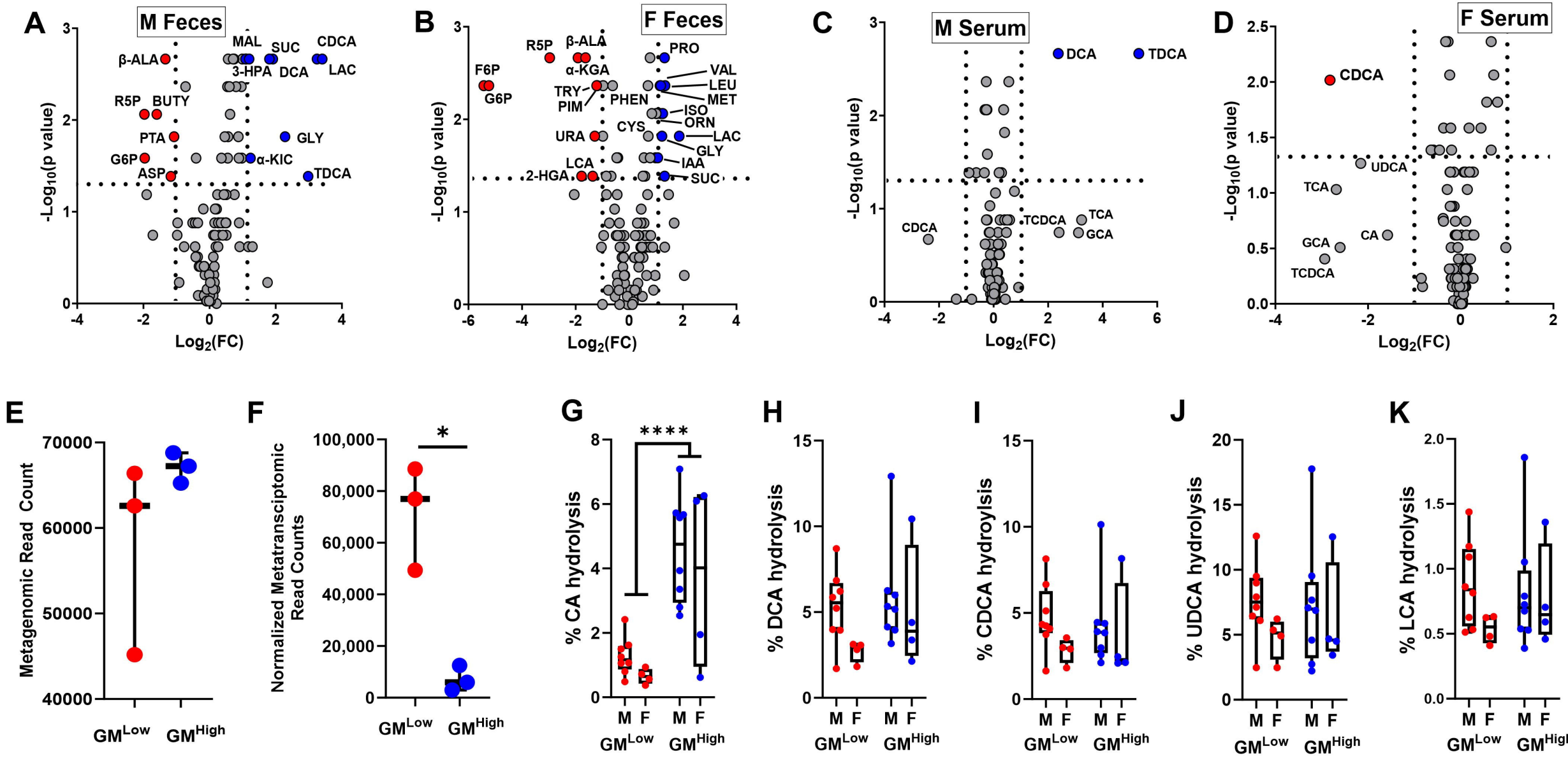

A
GM
GM^Low
GM^High
Chao1
400
300
200
Family
Acholeplasmataceae
Bacteroidaceae
Butyricicoccaceae
Clostridia_vadinBB60_group
Deferribacteraceae
Lachnospiraceae
Lactobacillaceae
Marinifilaceae
Muribaculaceae
Oscillospiraceae
RF39
Rikenellaceae
Ruminococcaceae
Other
0%
25%
50%
75%
100%
Relative Abundance
B
Feces
β-ALA
CYS
SUC
G6P
R5P
F6P
DCA
LAC
CDCA
BUTY
GLY
INO
MVA
LEU
2-HBA
TDCA
CA
-Log10(p value)
Log2(FC)
C
Serum
CDCA
UDCA
GCA
α-MCA
TCDCA
TLCA
TCA
-Log10(p value)
Log2(FC)
D
Asbt
ns
E
Ost-β
*
F
Gpbar1
****
G
Fxr
ns
H
Cyp7a1
ns
I
Slc10a1
**
J
S1pr2
**
M F M F
GM^Low GM^High
Ileum
Liver

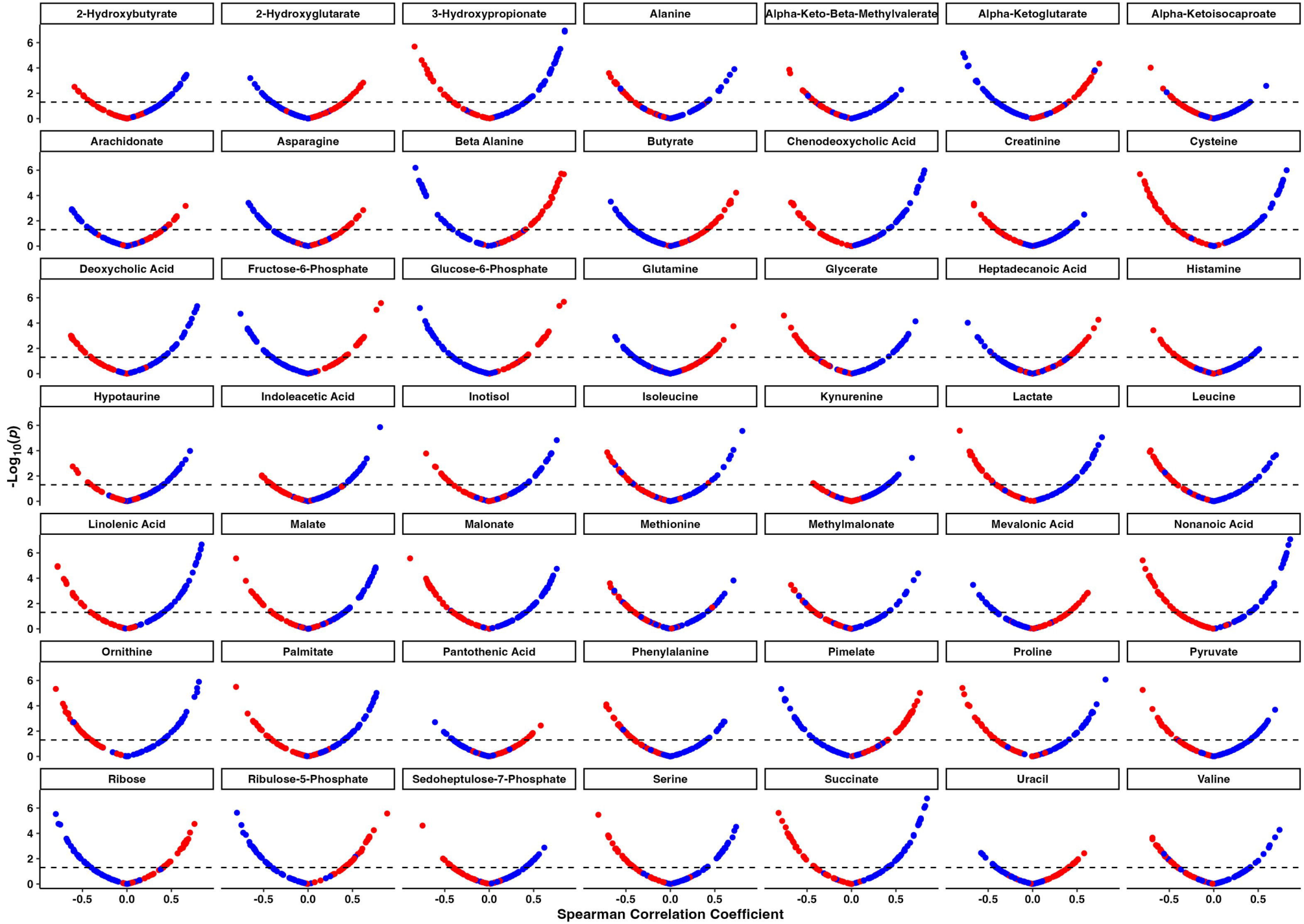

A
Conception
Birth
3 wks
5 wks
5.5 wks
6 wks
7 wks
3 & 7 wks
GM^Low
GM^High
CF^Low
CF^High
Maternal microbiome
Postnatal microbiome
Open-field
Light-dark test
Elevated plus
Running wheel
Body weight
Behavior testing outcomes
B
Time Spent in Light (%)
****
M F M F
GM^Low GM^High
C
Time Spent in Open Arms (%)
**
M F M F
GM^Low GM^High
D
Distance (Km)
GM^Low M
GM^High M
p < 0.0001, F = 29.2
Day
E
Distance (Km)
GM^Low F
GM^High F
p = 0.0002, F = 14.4
Day
F
Weight at 50 Days (grams)
****
M F M F
GM^Low GM^HIGH
G
Time Spent in Light (%)
**
M F M F
CF^Low CF^High
H
Time Spent in Open Arms (%)
*
M F M F
CF^Low CF^High
I
Distance (Km)
CF^Low M
CF^High M
Day
J
Distance (Km)
CF^Low F
CF^High F
Day
K
Weight at 50 Days (grams)
**
M F M F
CF^Low CF^High

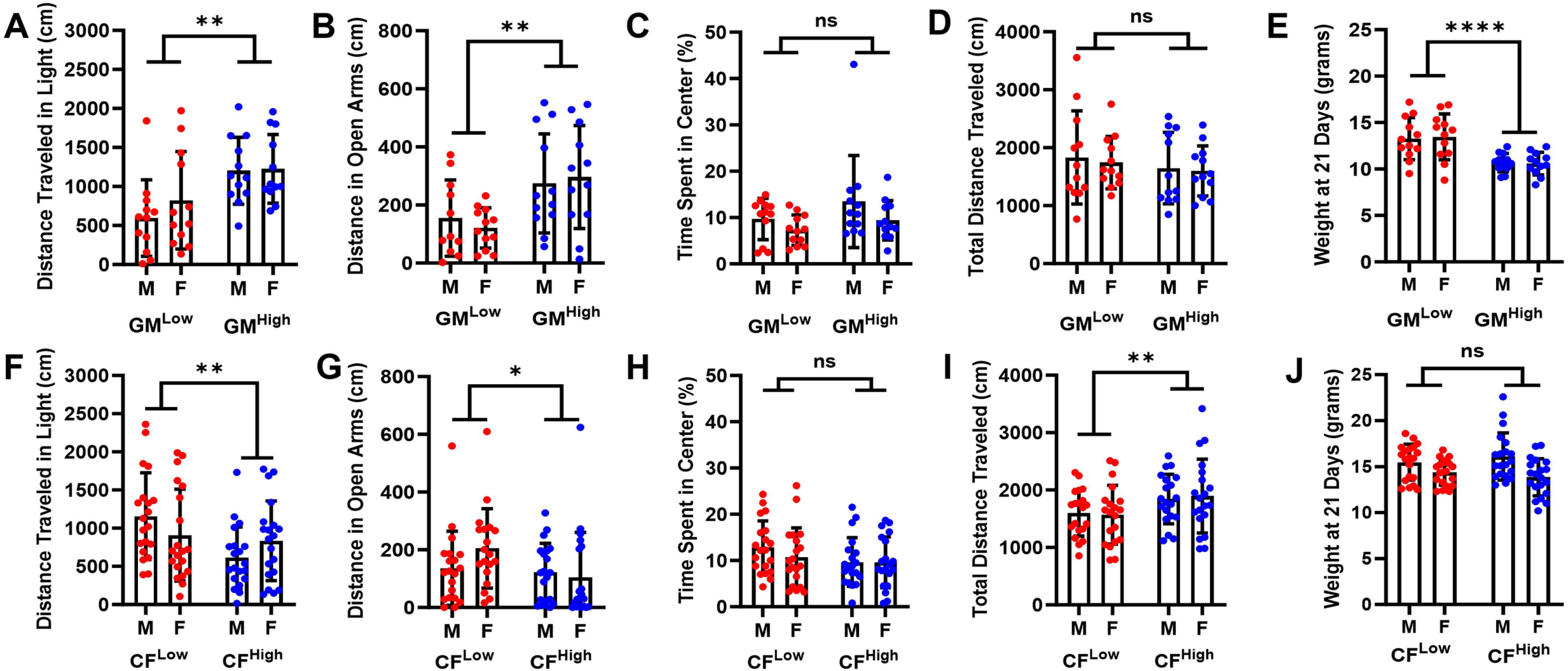

A
Swi5
Hfm1
Bend6
Pde1c
R² = -0.644
p = 2 × 10⁻⁷
Log2(FC) CFLow/CFHigh
Log2(FC) GMLow/GMHigh
B
p = 0.003
Chromosome
14
6
3
2
15
13
1
7
17
5
8
18
12
4
9
11
19
16
X
10
intron
exon
enh/prom
intergenic
Log2(obs/pred)
Number DMRs
C
exon 1
E0822088/enhP
E0822089/enhP; CpG island (34)
E0822090/prom
Beta value
GMLow
GMHigh
CFLow
CFHigh
Pde1c (27)
Log2(FC)
GMLow/GMHigh
CFLow/CFHigh
D
Gm32537
Gm49965
E0580122/prom
Ang5
4930503E14Rik
Ang6
Beta value
GMLow
GMHigh
CFLow
CFHigh
Chr 14 (+) (40)
Chr 14 (-) (16)
Log2(FC)
GMLow/GMHigh
CFLow/CFHigh
E
Hltf
Smg7
Xrcc6
Uvrag
Traip
Lefty2
Itfg2
Lgals4
Rheb
Rps6
Itga6
Arnt2
Smad3
Map3k
Mmp15
Lrrk1
Ptp4a3
Drd1a
Dock1
Gdnf
Rhoc
Gng4
Nts
Rap2b
Pla2g5
Nrcam
BC034090
Cacna1e
Pdzrn3
Rgl1
Tubb4b
Tbce
Pebp4
Mrps35
Mrpl17
Auh
Lrig1
Mccc1
Pcx
Ncald
Abcb4
Pde1c
Atp8b1
Tmcc1
Cog1
Tmem181a
Cog4
Node color: GMLow/GMHigh
-5.4
-2.7
0.0
2.7
5.4

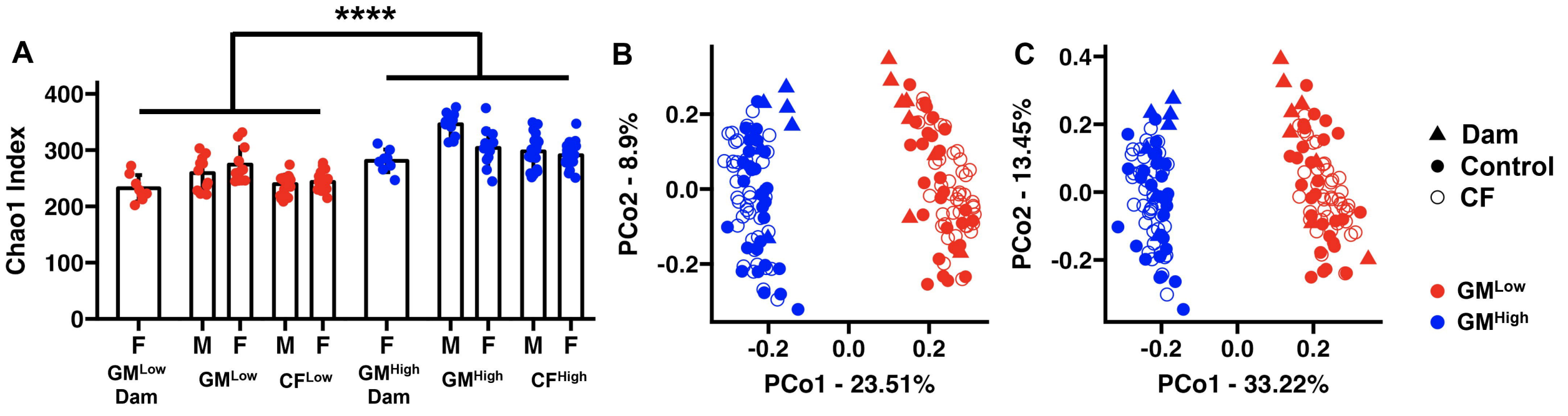

A
Ntng1
Pik3r3
Fas
Gja1
Sox17
Fzd6
p = 1 × 10-12
p = 1 × 10-16
B
Sox17
Pde1c
Gja1
Pik3r3
Fas
Fzd6
Dock1
Dock4
Dock5
Elmo1
Dcc
Srgap1
Srgap2
Arhgap1
Robo2
Adcy8
Adcy10
Pde1a
Pde3a
Pde4b
Pde5a
Pde7b
Pde8a
Pde10a
Pde11a
Log2(FC)
C
CFHigh ← → CFLow
L6 IT
Oligo
Astro
L5/6 NP
Micro
OPC
L5 IT
Endo
L5 IT
Astro
Astro
-2
0
2
Log2(FC) Dock1
D
CFHigh ← → CFLow
L6 IT
L5 ET
GABA
Astro
L5/6 NP
Micro
L5 IT
Endo
L5 IT
L6 IT
L5 IT
L5 IT
Astro
-1
0
1
Log2(FC) Pde1c
E
Cell-cell contact
XCR (Xcl1 + Xcr1)
LHB (Lhb/Cga + Lhcgr)
CD39 (Entpd1 + Adora2a)
CD46 (Cd46 + Jag1)
GMHigh
CFLow
GMLow
CFHigh
4
15
3
33
5
2
21
22
4
42
12
25
7
6
10
Orexigens
GHRELIN (Ghrl + Ghsr)
NPW-B (Npr/Npw + Npbwr1)
QRFP (Qrfp + Qrfpr)
Cell-cell contact
ADGRE5 (Adgre5 + Cd55b)
MADCAM (Madcam1 + Itga4/Itgb7)
SELL (Sell + Podxl/Glycam1/Madcam1)
PERIOSTIN (Itgav + Itgb3/Itgb5)
BST2 (Bst2 + Pira2)
EPO (Epo + Epor)
LIGHT (Tnfsf14 + Tnfrsf14)
F
GMLow
GMHigh
DESMOSOME
DSPP
VEGF
GDNF
NODAL
CLEC
CD40
NOTCH
NRG
IGF
IL2
RLN
CCL
IL1
CXCL
SPP1
SEMA4
BMP10
LAMININ
JAM
CD23
NECTIN
WNT
ANGPTL
CALCR
NPNT
GDF
TENASCIN
PROK
COLLAGEN
BMP
GALECTIN
ncWNT
IL17
MHC-I
NPY
THBS
CD200
PARs
TGFb
PTN
EGF
0%
25%
50%
75%
100%
Relative Information Flow
G
CFLow
CFHigh
MHC-I
BMP10
IGF
CLEC
IL2
BMP
PARs
CD200
EGF
PROK
NODAL
WNT
CD40
CCL
ncWNT
GALECTIN
VEGF
IL17
NECTIN
CXCL
TENASCIN
THBS
NPY
COLLAGEN
RLN
DSPP
IL1
PTN
CALCR
NOTCH
CD23
SPP1
ANGPTL
GDNF
DESMOSOME
NPNT
NRG
LAMININ
TGFb
GDF
SEMA4
JAM
0%
25%
50%
75%
100%
Relative Information Flow

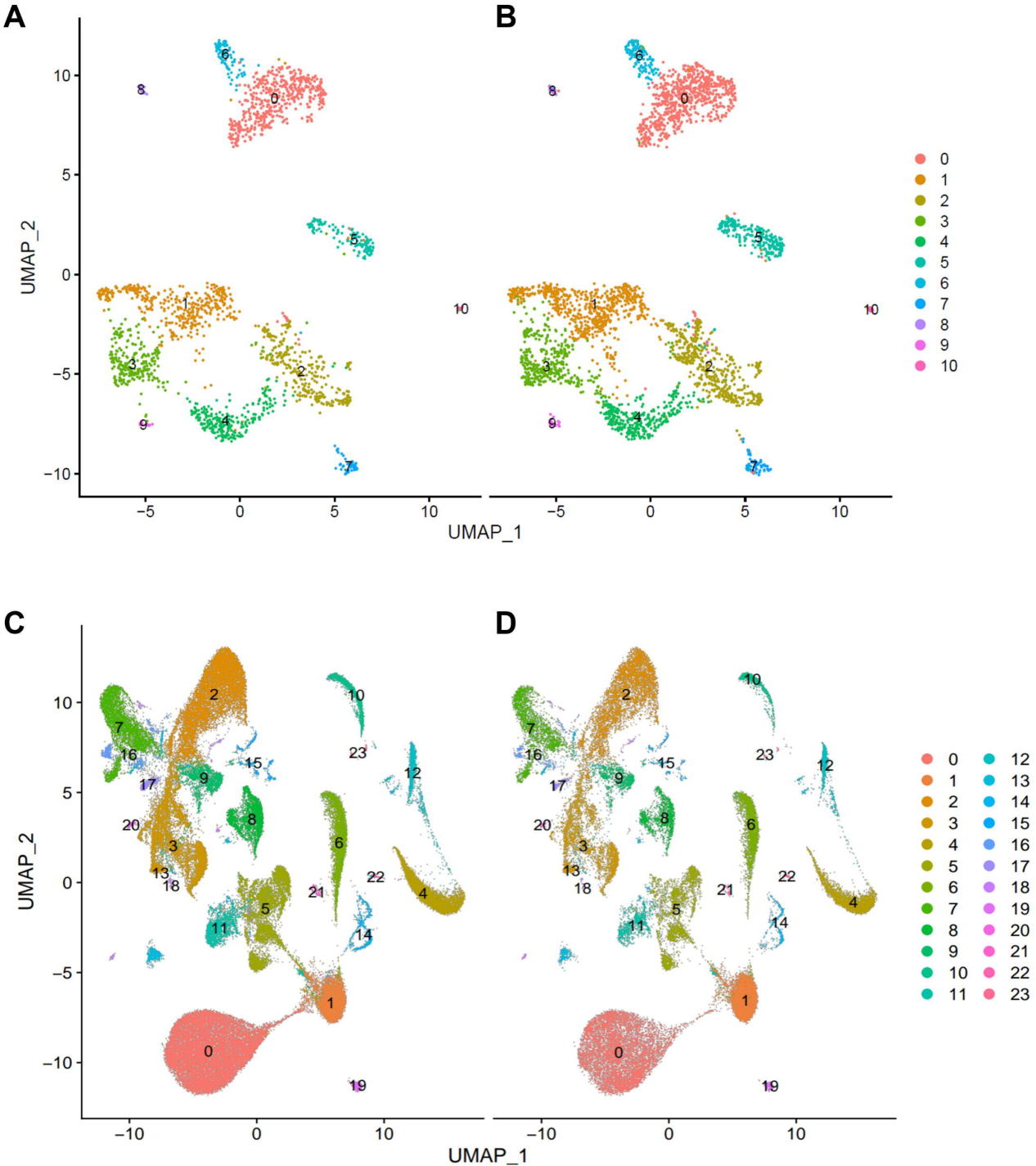

Phylum
Bacteroidota
Firmicutes
Deferribacterota
Proteobacteria
Patescibacteria
Cyanobacteria
Desulfobacterota
Actinobacteriota
Verrucomicrobiota
Bacteria
BH-adjusted p value
GMLow
GMHigh
0
1
p < 0.05

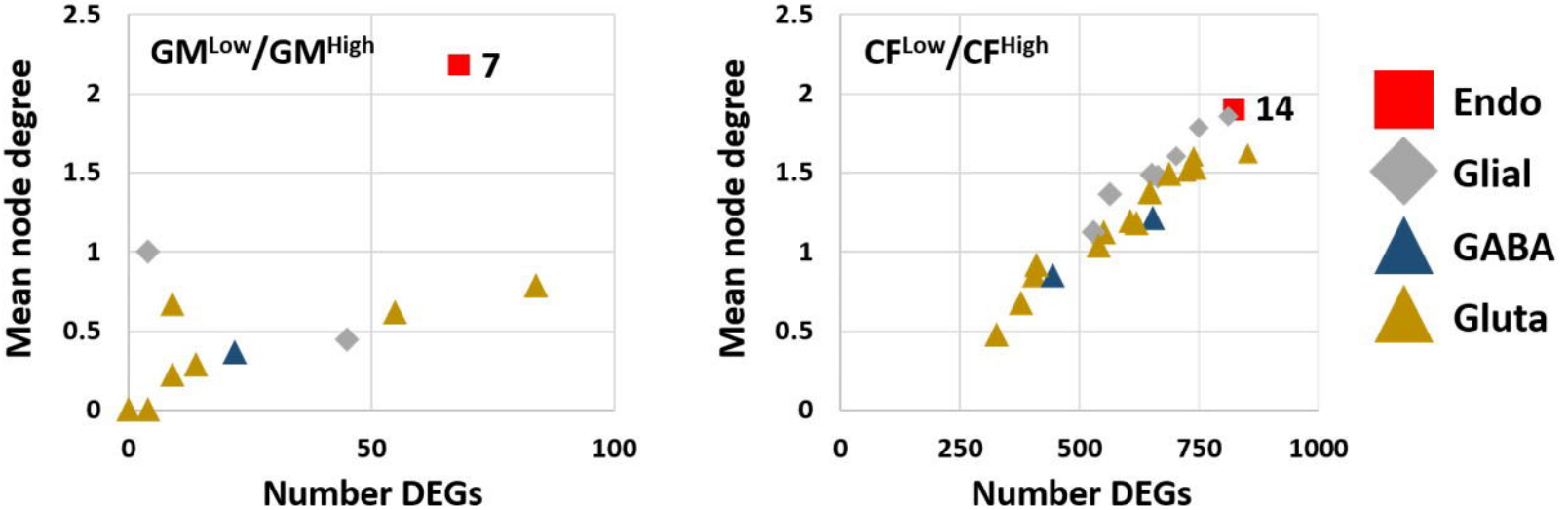

A
M Feces
β-ALA
MAL
SUC
CDCA
3-HPA
DCA
LAC
R5P
BUTY
PTA
GLY
G6P
α-KIC
ASP
TDCA
-Log10(p value)
Log2(FC)
B
F Feces
R5P
β-ALA
PRO
F6P
α-KGA
VAL
LEU
G6P
TRY
PIM
PHEN
MET
ISO
ORN
CYS
URA
LAC
GLY
LCA
IAA
2-HGA
SUC
C
M Serum
DCA
TDCA
CDCA
TCDCA
TCA
GCA
D
F Serum
CDCA
UDCA
TCA
GCA
CA
TCDCA
E
Metagenomic Read Count
GMLow
GMHigh
F
Normalized Metatranscriptomic Read Counts
*
G
****
% CA hydrolysis
M F M F
GMLow GMHigh
H
% DCA hydrolysis
I
% CDCA hydroylsis
J
% UDCA hydrolysis
K
% LCA hydrolysis

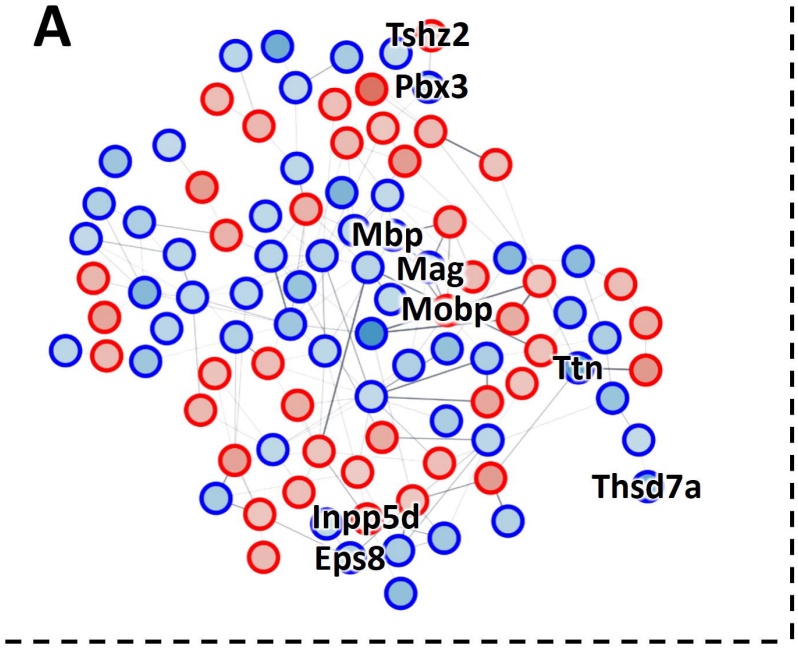

2-Hydroxybutyrate
2-Hydroxyglutarate
3-Hydroxypropionate
Alanine
Alpha-Keto-Beta-Methylvalerate
Alpha-Ketoglutarate
Alpha-Ketoisocaproate
Arachidonate
Asparagine
Beta Alanine
Butyrate
Chenodeoxycholic Acid
Creatinine
Cysteine
Deoxycholic Acid
Fructose-6-Phosphate
Glucose-6-Phosphate
Glutamine
Glycerate
Heptadecanoic Acid
Histamine
Hypotaurine
Indoleacetic Acid
Inotisol
Isoleucine
Kynurenine
Lactate
Leucine
Linolenic Acid
Malate
Malonate
Methionine
Methylmalonate
Mevalonic Acid
Nonanoic Acid
Ornithine
Palmitate
Pantothenic Acid
Phenylalanine
Pimelate
Proline
Pyruvate
Ribose
Ribulose-5-Phosphate
Sedoheptulose-7-Phosphate
Serine
Succinate
Uracil
Valine
-Log10(p)
0
2
4
6
-0.5
0.0
0.5
Spearman Correlation Coefficient

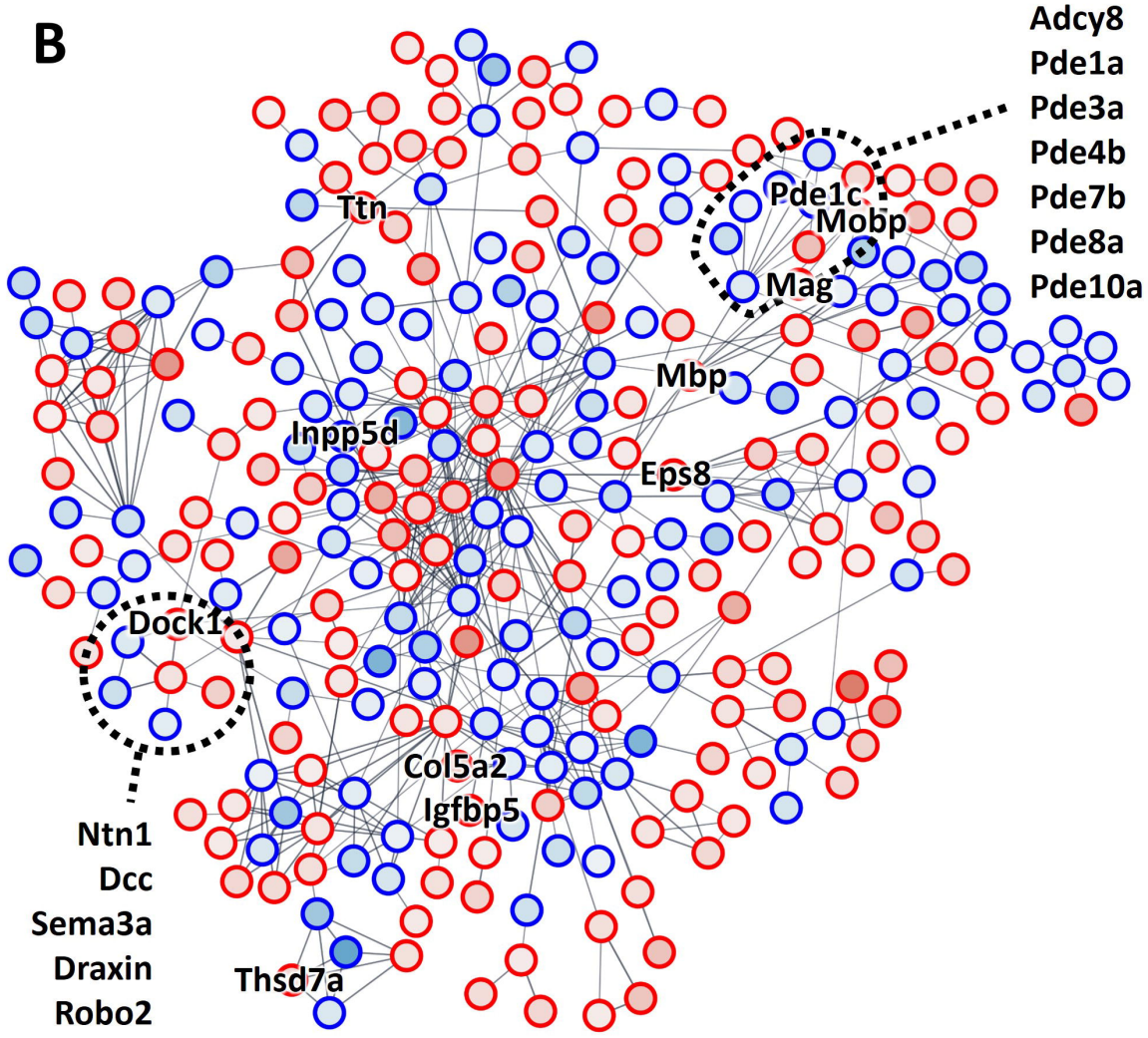

A
Distance Traveled in Light (cm)
**
M F M F
GM^Low GM^High
B
Distance in Open Arms (cm)
**
M F M F
GM^Low GM^High
C
Time Spent in Center (%)
ns
M F M F
GM^Low GM^High
D
Total Distance Traveled (cm)
ns
M F M F
GM^Low GM^High
E
Weight at 21 Days (grams)
****
M F M F
GM^Low GM^High
F
Distance Traveled in Light (cm)
**
M F M F
CF^Low CF^High
G
Distance in Open Arms (cm)
*
M F M F
CF^Low CF^High
H
Time Spent in Center (%)
ns
M F M F
CF^Low CF^High
I
Total Distance Traveled (cm)
**
M F M F
CF^Low CF^High
J
Weight at 21 Days (grams)
ns
M F M F
CF^Low CF^High

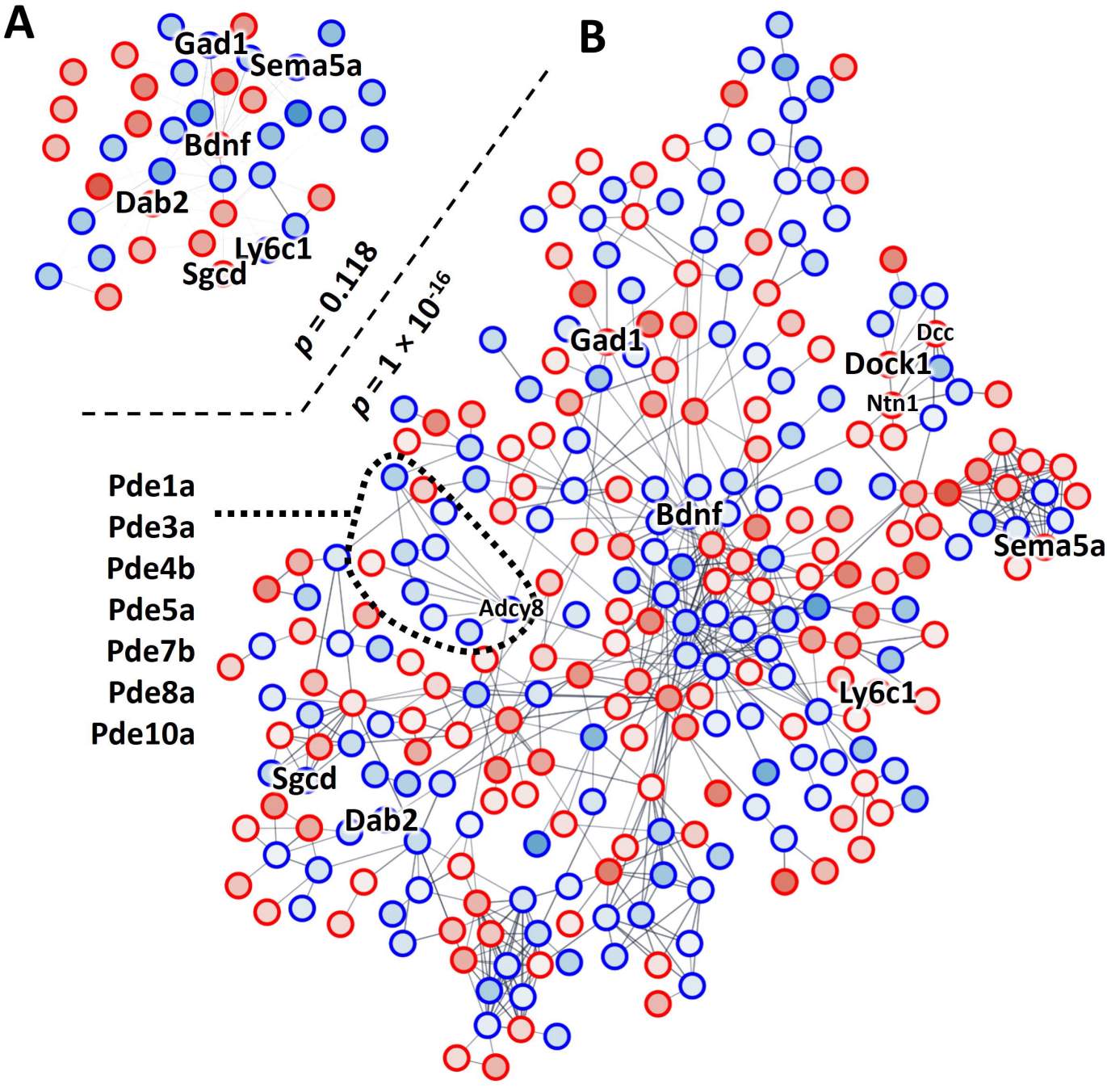

A
****
Chao1 Index
400
300
200
100
0
F
M
F
M
F
F
M
F
M
F
GMLow Dam
GMLow
CFLow
GMHigh Dam
GMHigh
CFHigh
B
PCo2 - 8.9%
0.2
0.0
-0.2
-0.2
0.0
0.2
PCo1 - 23.51%
C
PCo2 - 13.45%
0.4
0.2
0.0
-0.2
-0.2
0.0
0.2
PCo1 - 33.22%
Dam
Control
CF
GMLow
GMHigh


A
B
C
D
UMAP_1
UMAP_2
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23


GMLow/GMHigh
7
Mean node degree
Number DEGs
2.5
2
1.5
1
0.5
0
0
50
100
CFLow/CFHigh
14
Mean node degree
Number DEGs
0
250
500
750
1000
Endo
Glial
GABA
Gluta


A
Tshz2
Pbx3
Mbp
Mag
Mobp
Ttn
Thsd7a
Inpp5d
Eps8


B
Adcy8
Pde1a
Pde3a
Pde4b
Pde7b
Pde8a
Pde10a
Pde1c
Mobp
Mag
Ttn
Mbp
Inpp5d
Eps8
Dock1
Col5a2
Igfbp5
Ntn1
Dcc
Sema3a
Draxin
Robo2
Thsd7a


A
Gad1
Sema5a
Bdnf
Dab2
Ly6c1
Sgcd
p = 0.118
p = 1 × 10-16
B
Gad1
Dcc
Dock1
Ntn1
Pde1a
Pde3a
Pde4b
Pde5a
Pde7b
Pde8a
Pde10a
Bdnf
Sema5a
Adcy8
Ly6c1
Sgcd
Dab2