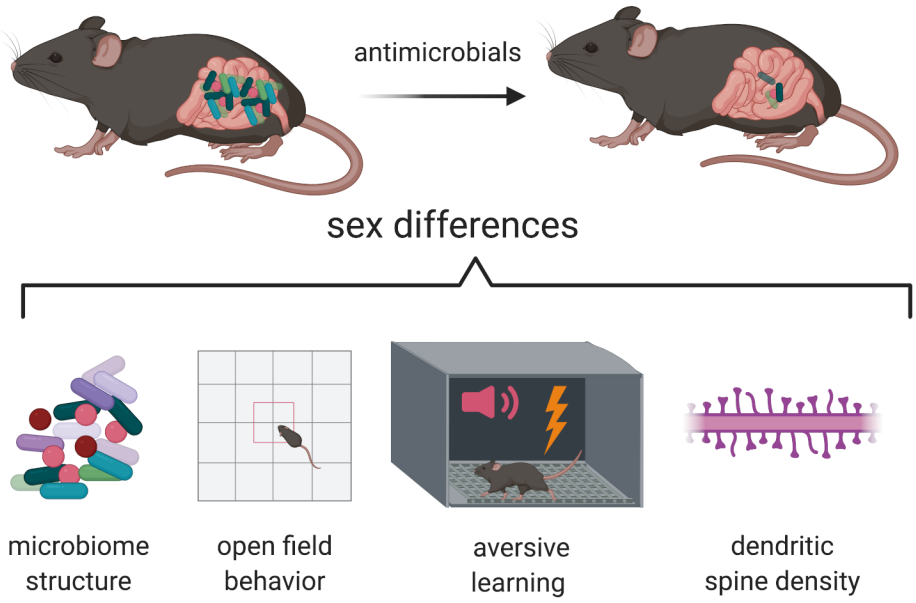

JNR
JOURNAL OF
Neuroscience
Research
June 2021 ■ Volume 99, Number 6
WILEY

Graphical Abstract



Adult female and male C57BL/6N mice voluntarily consumed either a 10-day course of antimicrobials or a control (tap water). Results showed that the antimicrobials exerted sex-specific effects on gut microbiome structure, open field behavior, cued aversive learning, and amygdala dendritic spine density (Figure created in https://biorender.com/).

Received: 21 July 2020 | Revised: 8 March 2021 | Accepted: 28 March 2021

DOI: 10.1002/jnr.24848

RESEARCH ARTICLE

JOURNAL OF Neuroscience Research

Sex differences in gut microbiota modulation of aversive conditioning, open field activity, and basolateral amygdala dendritic spine density

Caroline Grace Geary[1] | Victoria Christina Wilk[2] | Katherine Louise Barton[1] | Parvaneh Ottavia Jefferson[1] | Tea Binder[2] | Vasvi Bhutani[1] | Claire Luisa Baker[1] | Achal James Fernando-Peiris[2] | Alexa Lee Mousley[1] | Stefano Freitas Andrade Rozental[1] | Hannah Mae Thompson[1] | Justin Charles Touchon[2] | David Justin Esteban[2] | Hadley Creighton Bergstrom[1]

[1]Department of Psychological Science, Program in Neuroscience and Behavior, Vassar College, Poughkeepsie, NY, USA

[2]Department of Biology, Vassar College, Poughkeepsie, NY, USA

Correspondence
Hadley Creighton Bergstrom, Department of Psychological Science, Vassar College, 124 Raymond Avenue, Poughkeepsie, NY 12604, USA.
Email: habergstrom@vassar.edu

Funding information
Asprey Center for Collaborative Approaches in Science; Ruth M. Berger Foundation; The Lucy Maynard Salmon Research Fund

Abstract

Gut microbiota influence numerous aspects of host biology, including brain structure and function. Growing evidence implicates gut microbiota in aversive conditioning and anxiety-related behaviors, but research has focused almost exclusively on males. To investigate whether effects of gut dysbiosis on aversive learning and memory differ by sex, adult female and male C57BL/6N mice were orally administered a moderate dose of nonabsorbable antimicrobial medications (ATMs: neomycin, bacitracin, and pimaricin) or a control over 10 days. Changes in gut microbiome composition were analyzed by 16S rRNA sequencing. Open field behavior, cued aversive learning, context recall, and cued recall were assessed. Following behavioral testing, the morphology of basolateral amygdala (BLA) principal neuron dendrites and spines was characterized. Results revealed that ATMs induced gut dysbiosis in both sexes, with stronger effects in females. ATMs also exerted sex-specific effects on behavior and neuroanatomy. Males were more susceptible than females to microbial modulation of locomotor activity and anxiety-like behavior. Females were more susceptible than males to ATM-induced impairments in aversive learning and cued recall. Context recall remained intact, as did dendritic structure of BLA principal neurons. However, ATMs exerted a sex-specific effect on spine density. A second experiment was conducted to isolate the effects of gut perturbation to cued recall. Extinction was also examined. Results revealed no effect of ATMs on cued recall or extinction, suggesting that gut dysbiosis preferentially impacts aversive learning. These data shed new light on how gut microbiota interact with sex to influence aversive conditioning, open field behavior, and BLA dendritic spine architecture.

KEYWORDS
antibiotics, dendritic spines, estrous, fear conditioning, gut-brain axis, gut microbiome

Edited by Karina Alvina, David McArthur, and Junie Warrington. Reviewed by Caitlin Cowan and Karina Alvina.

Significance

Gut microbiota can influence brain function and behavior, including trauma and anxiety-related disorders. Although these disorders disproportionately affect women, preclinical research has focused almost exclusively on male rodent models. We investigated the impact of antimicrobial administration on gut microbiome structure, aversive conditioning, open field behavior, and basolateral amygdala principal neuron morphology in adult female and male mice. Results revealed that treatment exerted numerous effects, many of which were stronger in one sex than the other. Our findings underscore the importance of studying both sexes and support a role for microbial modulation of aversive learning, anxiety-like behavior, and amygdala spine patterning.

1 | INTRODUCTION

The gastrointestinal tract harbors a diverse microbial ecosystem which plays a fundamental role in human health. Although the ability of pathogens to modulate host physiology and behavior is well established (Lyte, 2013), important influences of resident microbiota did not enter the spotlight of neuroscience until Sudo et al. (2004) found that microbiota modulate neural networks involved in stress responsivity. Since then, a surge of evidence has amassed linking resident microbiota with host behavior and brain function, including processes such as synaptogenesis, myelination, tryptophan metabolism, microglial activation, and regulation of neurotransmitters and neurotrophic factors (Abdel-Haq et al., 2019; Bercik et al., 2011; Chu et al., 2019; Cryan & O'mahony, 2011; Desbonnet et al., 2015; Heijtz et al., 2011). Gut microbiota components and metabolites can act both locally and systemically through neural, immune, and endocrine pathways of communication between the gut and the brain—now termed the microbiota–gut–brain (MGB) axis—which comprise a bidirectional communication network (Cryan & Dinan, 2012; El Aidy et al., 2016; Forsythe et al., 2016; Martin et al., 2018).

Gut dysbiosis (broadly defined as a disruption of the homeostatic balance between host and associated intestinal microbiota, and characterized by a substantial shift in bacterial composition or metabolic activities) is associated with many chronic illnesses, particularly functional gastrointestinal disorders like irritable bowel syndrome (IBS; DeGruttola et al., 2016; Kawoos et al., 2017; van de Guchte et al., 2018). The prevalence of anxious and depressive symptoms in patients with functional bowel disorders is high, relative to both the general population as well as patients with other chronic illnesses (Jones et al., 2006; MacQueen et al., 2017). Increasing evidence suggests that the MGB axis may be a therapeutic target for a wide range of psychiatric conditions, including trauma and anxiety-related disorders (Cowan et al., 2018; Jaggar et al., 2019; Malan-Muller et al., 2018; MacQueen et al., 2017; Peirce & Alviña, 2019).

A leading behavioral model for associative "fear" learning and memory is aversive conditioning (Bergstrom, 2016). In classical aversive conditioning, a sensory stimulus, such as an auditory tone, is temporally paired with a naturally aversive stimulus, such as a mild foot shock (the unconditioned stimulus, US). Following learning, the tone by itself (now a conditioned stimulus, CS) is capable of triggering a conditioned defensive response, such as freezing. Studies generally assess subjects' freezing response to a discrete CS ("cued recall") and, separately, to the testing chamber in which learning occurred ("context recall"). Also assessed is the ability to learn when a CS no longer predicts an associated US ("extinction").

Various probiotics (microorganisms beneficial to health) have been found to enhance aspects of aversive conditioning in male rodents, including learning, cued and context recall, and extinction (Bravo et al., 2011; Fox et al., 2017; Savignac et al., 2015). In contrast, germ-free (GF) male mice exhibit impaired cued aversive memory recall and extinction (Chu et al., 2019; Hoban et al., 2018). Treatment with antimicrobial drugs (ATMs) is another well-established method of disrupting gut microbiota homeostasis, and has important clinical relevance given the widespread global use of antibiotics (Bercik et al., 2011; Jaggar et al., 2019; Rousham et al., 2018). Both ATM-treated and gnotobiotic (colonized by a small consortium of known bacteria) male mice have been found to exhibit impaired aversive extinction (Chu et al., 2019), suggesting that, even in developed adult brains, normal aversive conditioning circuitry depends on continuous input from a rich and diverse gut microbiome.

The amygdala is a central structure in emotional processing, motivation, and associative learning (LeDoux, 2000). Considerable research implicates the amygdala in MGB signaling: altered or absent microbial communities have been linked to changes in amygdala transcriptome (Hoban et al., 2016, 2017, 2018), neurochemistry (Bercik et al., 2011; Bravo et al., 2011), and structure (Luczynski et al., 2016; see Cowan et al., 2018 for a review). Morphological changes in amygdala principal neuron dendrites and spines—key indicators of experience-induced plasticity (Chklovskii et al., 2004)—have also been reported in GF male mice (Luczynski et al., 2016). Together, these data support a model in which gut microbiota modulate aversive conditioning, potentially via amygdala neuroplasticity.

Although a role for the MGB axis in aversive conditioning and amygdala plasticity has been established, preclinical research has centered almost exclusively on male rodents. Considering that women are more likely to be diagnosed with a trauma or anxiety-related disorder than men (Altemus et al., 2014), the study of sex differences in aversive conditioning and anxiety-related behaviors may have important implications for psychiatric care. Sex differences in aversive conditioning (Asok et al., 2019; Blume et al., 2017) and gut microbiome structure (Jašarević et al., 2016; Markle et al., 2013) have also been reported, albeit inconsistently (Ding & Schloss, 2014;

Org et al., 2016). Furthermore, functional bowel disorders like IBS, the most common form, disproportionately affect women (Kibune-Nagasako et al., 2016; Lovell & Ford, 2012).

To address a research gap in the characterization of sex differences related to microbial modulation of aversive conditioning, we induced gut dysbiosis in female and male C57BL/6N mice by administering a moderate impact ATM cocktail in drinking water for 10 days. We assessed auditory cued and contextual aversive learning and memory, as well as locomotor activity and anxiety-like behavior in the novel open field test (OFT). 16S rRNA sequencing was used to confirm that ATMs altered gut bacterial community structures. Following behavioral testing, we analyzed changes in architecture of basolateral amygdala (BLA) principal neuron dendrites and dendritic spines. Because ATM treatment impaired both aversive learning and cued recall, we speculated that decreased freezing in cued recall may have been due to deficits in learning or memory consolidation. To investigate whether gut dysbiosis has a specific impact on cued aversive memory retrieval, we conducted a second experiment. In this experiment, ATMs were administered after consolidation of the cued aversive memory, and extinction learning and estrous stage were also examined.

2 | METHODS AND MATERIALS

2.1 | Animals

Female (n = 46; weight = 29.0 g ± 5.4 (mean ± standard deviation, here and throughout)) and male (n = 48; weight = 29.2 g ± 4.8) C57BL/6N (B6) mice were 10 to 16 weeks old (mean = 13 weeks) at the beginning of the first experiment. Female (n = 24; weight = 26.5 g ± 1.8) and male (n = 24; weight = 33.8 g ± 2.5) B6 mice were 10 to 16 weeks old (mean = 14 weeks) at the beginning of the second experiment. Mice were originally derived from Charles River Laboratories (Wilmington, MA, USA) and bred in-house at the Vassar College vivarium (Poughkeepsie, NY, USA) over multiple generations. Each year mice were supplemented into the breeding colony to maintain genetic heterogeneity. Mice were kept on a 12-hr light/dark cycle (lights on at 0600) in standard polycarbonate cages (28 × 17 × 12 cm) in a climate-controlled (21°C, humidity 65%) vivarium. *Ad libitum* food (LabDiet® JL Rat and Mouse/Irr 6F), water, and environmental enrichment (a combination of Nestlets and either an EcoForage pack or a wood gnawing block and EnviroPak) were provided. Cages were changed 2×/week. Mice were housed in groups of two to five until 1 week prior to the start of the experiment, at which point they were single-housed. Mice were pseudo-randomly assigned (based on body weight) to experimental groups before the start of each experiment. A total of six cohorts were tested in the first experiment, and a total of two cohorts were tested in the second experiment. The first four cohorts were of the same sex (either all female or all male). The final four cohorts were of mixed sex (half female and half male). Experiments took place over a 26-month period, from 2017 to 2020. Mice were not handled by experimenters prior to testing. All experimental protocols were approved by the Vassar College Institutional Animal Care and Use Committee. Disclosure of animal housing, husbandry, and experimental procedures followed principles for transparent reporting (Prager et al., 2011) and reproducibility (Prager et al., 2019) in behavioral neuroscience.

2.2 | Antimicrobial regimen

Half to the subjects in each cohort ("ATM mice") were administered ATMs (2 mg/ml neomycin; 2 mg/ml bacitracin; 1.2 μg/ml pimaricin; Sigma-Aldrich, St. Louis, MO) in tap water for 10 days. The other half ("controls") received tap water. The ATM cocktail was freshly prepared every second day. In the first experiment, ATMs were started 5 days prior to behavioral testing (Figure 1a). In the second experiment, ATMs were started the day after context recall (Figure 1b). To our knowledge, three other studies have used this cocktail to study the MGB axis. Based on findings that higher doses of this cocktail may deleteriously affect voluntary consumption (Odeh, 2013), we used a dose that was relatively low compared to the other two studies (Bercik et al., 2011; Verdu et al., 2006). Advantages of this cocktail include that it does not induce intestinal inflammation, and it is nonabsorbable from the gut (Bercik et al., 2011; Odeh, 2013; Van Der Waaij et al., 1974; Van der Waaij & Sturm, 1968; Verdu et al., 2006). This is supported by evidence showing that behavioral changes associated with this cocktail are induced by enteral but not intraperitoneal administration (Bercik et al., 2011). Furthermore, bacitracin and neomycin do not readily cross the blood–brain barrier (Desrochers & Schacht, 1982; Teng & Meleney, 1950).

2.3 | PCR amplification and library preparation

Fecal pellets were collected in the final two cohorts of Experiment 1 on Days 0, 3, 7, and 10 (Figure 1a). Fecal pellets were not collected during Experiment 2. One pellet was obtained per subject by individually placing mice in a plastic box sterilized with 70% ethanol. Immediately following defecation, pellets were transferred to a microfuge tube using sterilized tweezers, then stored at −80°C. To extract DNA, pellets were heated at 65°C for 10 min, then homogenized and processed using a QIAmp Powerfecal DNA kit (Qiagen). The measurement of DNA concentrations in extracts in a subset of samples was approximately 20–40 ng/μl. DNA was stored at −20°C.

The 16s rRNA gene amplicon libraries were constructed in a two-stage polymerase chain reaction (PCR) procedure. In the first stage, the V3 and V4 regions of the 16S rRNA genes were amplified with Illumina V3-V4 forward (TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG) and reverse primers (GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC). Reactions included 12.5 μl of 2× KAPA HiFi Hot Start Ready Mix (Roche Sequencing Systems, Indianapolis, IN, USA), 5 μl each of the forward and reverse primers at 1 μM, and 2.5 μl of the template

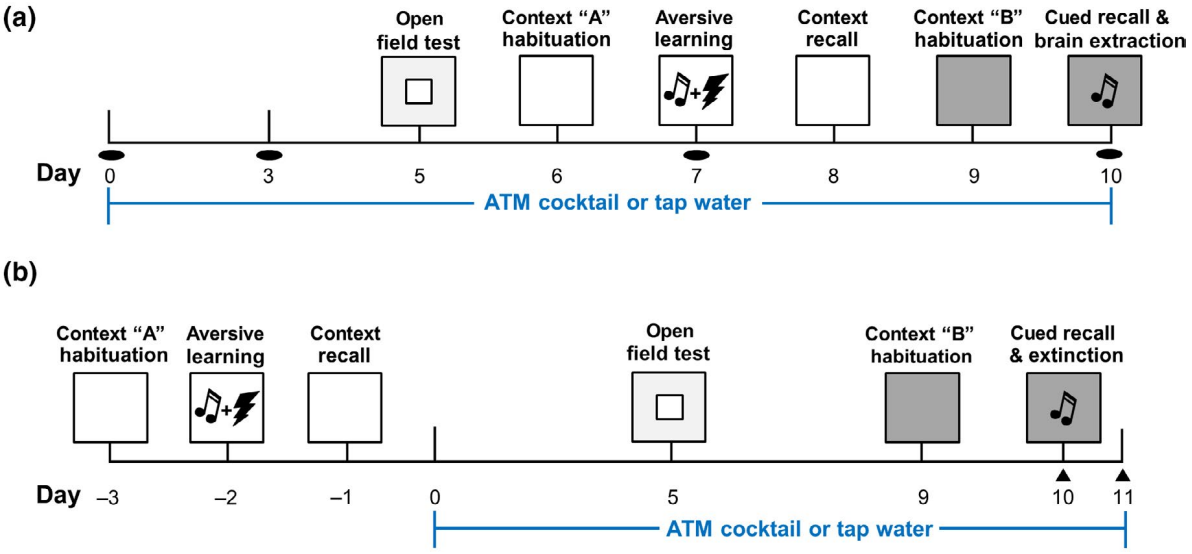


FIGURE 1 Experimental schematics. (a) Experiment 1 schematic. Half of the subjects were administered antimicrobial medications (ATMs) in tap water for 10 days. Controls received tap water. Fecal pellets (black ovals) were collected on Days 0, 3, 7, and 10. (b) Experiment 2 ("ATM Delay") Schematic. To isolate the effects of gut dysbiosis on cued recall, ATMs were started after rather than prior to learning of the aversive cue. Vaginal cytology (black triangles) was performed on Days 10 and 11

DNA. Thermocycling conditions were as follows: 95°C for 3 min, 95°C for 30 s, 55°C for 30 s, 72°C for 30 s, and 72°C for 5 min. Steps 2–4 were repeated 34 times. Gel electrophoresis was used to confirm the presence of approximately 500bp products. Products were quantified using a NanoDrop One (Thermo Scientific, Waltham, MA, USA).

A QIAquick PCR Purification Kit (Qiagen Inc, Germantown, MD, USA) was used to purify the original PCR product, then the second-stage PCR was performed to attach Illumina i7 and i5 indexes (Illumina Nextera XT Library Preparation Kit, Illumina Inc, San Diego, CA, USA) to individual samples. The reaction contained 12.5 μl of 2× KAPA HiFi Hot Start Ready Mix, 5 μl of the forward and reverse primers at 2 μM, and 2.5 μl of the purified first-stage PCR product. Thermocycling conditions were the same as described above, except steps 2–4 were repeated eight times. Gel electrophoresis was used to confirm the presence of product. A QIAquick PCR Purification Kit was again used to purify second-stage PCR products. After determining the concentration using a NanoDrop One, each product was diluted to a concentration of 8 nM prior to combining all samples for sequencing. The final concentration of the combined sample was quantified using Agilent Bioanalyzer. The sample was sent to Cornell Biotechnology Institute, where the DNA was size-selected using a BluePippin (Sage Science, Beverly, MA, USA) before the sample was sequenced on an Illumina MySeq (Illumina pipeline software v2.18) with paired-end 250 base read length.

2.4 | Gut microbiome data analysis

Paired-end reads were processed and analyzed using QIIME v2018.8 (Bolyen et al., 2019). Reads were trimmed at the 5′ end to remove primer sequences, truncated at the 3' end to remove low-quality ends, and denoised using Dada2 (Callahan et al., 2016). Sequences were clustered at 100% identity to generate amplicon sequence variants (ASVs). A rooted phylogenetic tree was generated for phylogenetic diversity analysis using the q2-phylogeny plugin, which uses mafft to generate a multiple sequence alignment (Katoh & Standley, 2013) and fasttree to generate the tree (Price et al., 2010). Diversity as a function of sampling depth was investigated using *qiime diversity alpha-rarefaction* with default settings (including calculation of Faith's phylogenetic diversity (PD; Faith, 1992), Shannon index, and ASV richness) and a maximum depth of 50,000 reads. Then, three metrics of α-diversity—PD, species evenness, and ASV richness—were calculated using *qiime diversity core-metrics-phylogenetic* with a sampling depth of 10,000 reads per sample. Taxonomy was assigned to ASVs using *qiime feature-classifier classify-sklearn* with a SILVA-99 classifier.

All measures of α-diversity were assessed using linear mixed effects models (LMMs). Analyses were performed in R v3.6.1 (R Core Team, 2019). All analyses included experimental cohort as a random effect. Thus, our LMMs were analogous to two-way ANOVAs, but they controlled for nonindependence of mice between cohorts. Fixed effects were treatment, sex, and their interaction. α-diversity analyses were conducted in two stages. First, we analyzed all data from Day 0 to assess pretreatment differences. Next, we analyzed data from Days 3, 7, and 10 to assess the effects of ATM treatment. Where no effect of treatment day was found, data from Days 3, 7, and 10 were collapsed into control and ATM groups. Significance was assessed using likelihood ratio tests of increasingly simplified nested models. Model fit was assessed using quantile–quantile plots of residuals. Box plots were generated in R. The top and bottom of each box plot represent 75% (upper) and 25% (lower) quartiles, respectively. The center

line represents the median, and the whiskers represent the minimum and maximum values.

Whereas α-diversity describes structure within microbial communities, β-diversity describes how communities are differentiated (Lozupone & Knight, 2005). Several complementary measures of β-diversity (weighted UniFrac, unweighted UniFrac, Bray–Curtis, Manhattan, Euclidean, and Canberra; Parks & Beiko, 2013) were calculated using phyloseq v1.24.2 (McMurdie & Holmes, 2013) in R then analyzed using permutational multivariate analysis of variance (PERMANOVA) with 999 permutations. Main effects were sex and treatment. Analyses were again conducted for Day 0 data, then for treatment Days 3, 7, and 10 together. Significant PERMANOVA results were further analyzed using betadisper from the package vegan (Oksanen et al., 2019). One-way ANOVA was used to determine if results could be explained by differences in dispersion. Principal coordinate (PC) plots were generated in R to visualize the data. Distances were calculated in phyloseq, and ordination was performed using betadisper.

Differential abundance analysis was performed on samples from Days 3, 7, and 10 using the q2-gneiss plugin for QIIME 2 (Morton et al., 2017). ASVs with fewer than 50 sequence counts, or which appeared in under five samples, were removed; any counts of 0 were replaced with a pseudocount of 1 to allow log transformation. Within q2-gneiss, correlation clustering using Ward's hierarchical clustering by treatment day was used to define partitions of co-occurring ASVs, which were then log-transformed and followed by a multivariate response linear regression using the ordinary least squares function to identify meaningful covariates (sex, treatment, day) that impacted the model. For subsequent determination of significant balances, q2-gneiss was performed on female and male samples separately using treatment and treatment day for the linear regression. The largest significant balances in both sexes were selected for inspection of taxa populating the balance. Plots were generated in R, as described above. Significance for all data was set at $p < 0.05$.

2.5 | Animal health

General health and weight of mice were monitored and recorded daily. Percent body weight change was calculated from baseline (Day 0). In one cohort in the first experiment, ATM dose was dropped to 50% on Day 6 in six female subjects who were approaching >15% weight loss; dosing was returned to 100% the following day, when body weight increased. All subjects in both experiments displayed normal behaviors and signs of good health (active, alert, well-groomed, clear eyes) throughout ATM administration. Water consumption was assessed only in the second experiment. Water bottles were weighed daily, including prior to placement in or collection from cages. Daily water consumption for each subject was determined by subtracting final from initial bottle weight, then converting the result to ml. To compare water intake across subjects of varying body weights, each subject's daily water intake (ml) was divided by its daily body weight (g).

2.6 | Behavioral testing

Visual depictions of Experiments 1 and 2 are shown in Figure 1. Behavioral testing was conducted during the light cycle. Unless otherwise indicated, all testing equipment was wiped with 70% ethanol before and after each test. Subjects were counterbalanced by treatment and age to minimize potential order effects. The days prior to both fear conditioning (Context A) and cued recall (Context B), subjects were allowed to freely explore the respective context chamber configuration for 30 min ("habituation").

2.6.1 | Open field test

OFT was used to assess baseline locomotor activity (determined by the sum of total distance traveled) and anxiety-like behavior (determined by % time spent in the arena's center, which is considered a more anxiogenic environment than the enshadowed periphery; Bailey & Crawley, 2009). Mice were placed in the bottom-right corner of the novel, dimly lit rectangular arena (3 lux; 54 × 39 × 38 cm; center: 13 × 27 cm (~17% of the total area)) and allowed to explore freely for 10 min.

2.6.2 | Aversive learning (Context A)

The aversive conditioning chamber measured 18 × 18 × 45 cm and was enclosed in a sound-attenuating chamber (58 × 45 × 61 cm; Coulbourn Instruments, Whitehall, PA, USA). Programming was delivered via Graphic State software (Coulbourn Instruments). Following a 180-s habituation period, mice were presented with three 20-s auditory cues (CS; 72–75 dB; 5 kHz), each of which co-terminated with a 0.6 mA, 0.5 s foot shock (US). Intertrial intervals (ITIs) varied in length (20, 80, and 60 s), with ITI3 immediately following the presentation of the third CS. The total training time was 6 min 40 s.

2.6.3 | Context recall (Context A)

Mice were placed back in the original training context (Context A; commercial chamber) and allowed to explore freely for 20 min.

2.6.4 | Auditory cued recall (Context B)

Auditory cued recall was conducted in the same chamber as Context A, but the chamber was disguised (hereafter referred to as Context B) to circumvent the recall of Context A. Disguises to Context A included the modification of habituation conditions (dim lights, cages hand-carried, and white noise, as opposed to bright lights, cages wheeled on cart, and quiet), the aversive conditioning apparatus (1% acetic acid, plastic floor covered in clean bedding, and two white

stripes added to each plexiglass wall, instead of ethanol, steel bar flooring, and transparent walls), and the experimental protocol (ITI1 and ITI2 were switched, and no foot shocks were delivered). The total test time was 6 min 40 s.

2.6.5 | Auditory cued extinction (Context B)

Auditory cued extinction was conducted only in the second experiment, and took place in Context B, as an extension of the cued recall test. Mice were presented with thirty 20-s CSs (72–75 dB; 5-kHz). ITIs were created using a random number generator, and varied in length between 5 and 60 s. The total test time was 23 min 20 s.

2.7 | Behavioral quantification

Cameras positioned directly above the OFT arena and aversive conditioning chambers recorded video footage. Footage was analyzed using the video-tracking system SMART v3.0 (Panlab, Harvard Apparatus, Barcelona, Spain). For OFT, SMART quantified total distance traveled and % time spent in the arena's center. For aversive conditioning assays, "freezing" was used as a behavioral measure of a conditioned defensive response. Freezing was defined as immobility, except for respiration, that lasted >1 s. For each time bin (habituation, CSs, and ITIs), freezing duration was averaged and converted into a percentage. Freezing in Experiment 1 was quantified using SMART. Because our laboratory subsequently upgraded video-tracking software to Freezeframe (Actimetrics, Wilmette, IL, USA), this program was used to quantify freezing in Experiment 2. To verify tracking software accuracy and comparability of freezing quantification, a subset of the digitally-analyzed footage from both experiments was hand-scored by two blind experimenters to 95% inter-rater reliability. Comparison of these results with those of the automated tracking software revealed a remarkably high inter-rater reliability ($R^2 = 0.99$).

2.8 | Behavioral statistics

Prior to analyses, outliers were determined as values greater than or equal to 1.5 times the interquartile range. Locomotor activity, anxiety-like behavior, and the first three aversive conditioning tests (aversive learning, context recall, and cued recall) were each analyzed in SPSS v. 26 (IBM, Armonk, NY, USA) using a two-way ANOVA. Dependent variables were distance traveled (meters), % time spent in the center of the arena, and % freezing, respectively. Independent variables were sex and treatment (ATM vs. Control). Because sex differences were hypothesized a priori, all two-way ANOVAs were followed up by within-sex ANOVAs. Sphericity was assessed using Mauchly's test; if significant, the Greenhouse–Geisser adjustment was applied. Homogeneity of variance tests were conducted using Levene's test; if significant, the Welch *F*-ratio was used. Bonferonni's post hoc tests were conducted as necessary. For relevant comparisons between sex, lower and upper bound 95% confidence intervals (CIs) were reported, as well as Cohen's *d* values. Cued expression (cued recall and extinction in the second experiment), relative body weight change, and water consumption were each analyzed in R v3.6.1 using LMMs in the package lme4 (Bates et al., 2014). Dependent variables were % freezing, % body weight change, and relative volume of liquid consumed, respectively. Fixed effects were sex, treatment, and CS or Day (as appropriate), as well as all two-way and three-way interactions. We included mouse ID nested within testing cohort as a random intercept and CS or Day as a random slope to account for the nonindependence of individuals measured repeatedly over time. Significance of fixed effects was assessed with nested likelihood ratio tests. Model fit was checked visually by assessing quantile–quantile plots of residuals. In order to explicitly compare the effect of ATMs within each sex, we ran a priori contrasts using the emmeans package (Lenth, 2020). We compared the difference in estimated marginal means between ATM and control groups within each sex, as well as calculated the *p* value from Tukey post hoc comparisons to assess if ATMs had a stronger effect on males or females. All subjects underwent each behavioral test and were excluded from analyses only when subjects were determined to be outliers, or when data were unavailable (footage lost due to malfunctioning tracking software). In the first experiment, four outliers were excluded from both OFT distance traveled and % center time. Three outliers were excluded from both aversive learning and cued recall. There were no outliers in context recall, but data from 35 mice were lost due to digital recording error. No outliers were detected in the second experiment. Statistical significance for all data was set at $p < 0.05$. All tests were two-tailed. Box plots were generated as described above, except in OriginPro v. 2019b. All other plots were generated in R. Gray-shaded areas in % body weight change and liquid consumption plots represent smoothed 95% confidence intervals.

2.9 | Golgi-Cox staining

Six mice per group were randomly selected for brain extraction from the final two cohorts of the first experiment. Immediately following the conclusion of auditory cued recall, these mice were anesthetized with a ketamine/xylazine cocktail (100:10 mg/ml) and intracardially perfused with saline. The whole brain was submerged in a Golgi-Cox solution composed of 5% w/v mercuric chloride, 5% w/v potassium chromate, and 5% w/v potassium dichromate. Brains were stored in the dark, at room temperature, for 5 days. The solution was refreshed after the second day. Next, brains were transferred to a cryoprotectant (Zaqout & Kaindl, 2016) and stored at 4°C until slicing. Coronal sections of 200-μm thick were cut in cryoprotectant using a vibratome (VT1200, Leica Biosystems Inc., Buffalo Grove, IL, USA). Sections were transferred to 3% gelatinized slides and left to dry in the dark at room temperature for not more than 2 days.

To develop the sections, they were first rinsed in dH20, dehydrated in 50% ethanol, and alkalinized in a 33% ammonia hydroxide solution. Following another rinse in dH20 and immersion in 5% sodium thiosulfate (in the dark), slides were dehydrated in a graded series of ethanol dilutions, cleared using xylenes, and cover slipped using a mounting medium (Permount, Fisher Scientific, Hampton, NH, USA). These procedures were adapted from several previous protocols (Bergstrom et al., 2010; Gibb & Kolb, 1998; Zaqout & Kaindl, 2016).

2.10 | Imaging and 3D dendrite reconstruction

2.10.1 | Dendrite visualization

Dendrites from Golgi-Cox-stained BLA principal neurons were visualized using brightfield microscopy (Axio Imager M2, Zeiss, Thornwood, NY, USA) under a 63× (0.75 N.A.) air objective, and manually reconstructed in 3D using Neurolucida (MBF Biosciences, Williston, VT, USA). All investigators were blind to treatment condition. Principal neurons in the BLA complex (defined as the lateral, basolateral, and basomedial subnuclei) were sampled randomly, and from both hemispheres evenly. Neurons chosen for reconstruction were well-stained (fully impregnated dendritic trees), with unobstructed dendritic arbors that could be followed from soma to terminal tip without interruption. Morphometric analysis was restricted to cells located between bregma −1.00 and −2.30 mm. The BLA was identified by the contour of the two branching fiber tracts that delineate its medial and lateral borders, which become readily visible using the Golgi stain. Principal neurons were differentiated from stellate neurons and interneurons based on morphometric criteria that included spines on later branch orders, an "apical-like" dendritic tree, and biconical dendritic radiation (Bergstrom et al., 2010).

2.10.2 | Dendritic spine visualization

Dendritic spines were visualized under a 100× (1.3 N.A.) oil-immersion objective. Spines were characterized on a randomly chosen dendritic tree (soma to terminal tip) from a previously reconstructed neuron. Spines were defined as small (<3 μm in length) protrusions emanating from the dendritic shaft. Spines were characterized as one of four types: stubby (protrusions that lack a neck), mushroom (neck with a larger head), thin (long, skinny neck and small, bulbous head), or unknown (Yuste, 2010).

2.11 | Dendrite and dendritic spine morphometric statistics

The sample size for morphometric analyses was based on previous reports (Bergstrom et al., 2008, 2010). Quantitative morphological measurements were extracted using Neuroexplorer software (MBF Biosciences, Williston, VT, USA). All data were analyzed in R v3.6.1. Apical and basilar Sholl data were restricted to radii between 30 and 250 μm because there were too few data points to analyze above or below these values. All data were analyzed using generalized LMMs in the glmmTMB package (Brooks et al., 2017), with sex and treatment as fixed effects, radius as a random slope, and individual ID as a random intercept (Wilson et al., 2017). The random slope controls for variation in repeated measurements within a single individual (akin to a repeated measures design), while the random intercept controls for multiple cells measured within a single individual. The error distribution for different measures of Sholl data was determined by comparing the fit of different possible distributions (i.e., normal, lognormal, Poisson, and negative binomial) with the fitdistr function in the MASS package (Venables & Ripley, 2002). Dendritic intersections and overall length were analyzed with a negative binomial distribution. Given the large number of zeros in the data (radii without nodes), we first analyzed node data as a binomial variable, investigating the presence/absence of nodes at each radius. We then analyzed radii where nodes were present using a negative binomial distribution to determine if the number of nodes varied by sex or treatment.

Spine data was analyzed in terms of spine density per branch order (spines per μm). Since spine density values were very small (mean = 0.048) and contained a large number of zeros (branches without spines), we analyzed data as generalized LMMs with a Tweedie distribution using the glmmTMB package. Tweedie distributions (Jørgensen, 1987) are increasingly used in neuroscience (Moshitch & Nelken, 2014), and describe a probability distribution for continuous positive data that include a large portion of the data at zero, and where the values are generally very small. We began by analyzing the three-way interaction between branch order, sex, and treatment. Random effects included branches and cells nested within individuals. Because this three-way interaction was significant, we then analyzed sex x treatment interactions within each branch order separately, with individual ID as a random effect. Significance was set at $p < 0.05$ and assessed using likelihood ratio tests of increasingly simplified nested models (Crawley, 2012), essentially equivalent to a two-tailed significance test. Box plots were generated in R, as described above. Error bars on line graphs represent standard deviation.

2.12 | Vaginal cytology and histology

Because the influence of ovarian hormones on aversive extinction (but not cued aversive learning or recall) has been established (Milad et al., 2009; Velasco et al., 2019), we conducted a preliminary investigation of potential ATM × estrous stage interactions during the second experiment. Vaginal cells were collected on Day 10, after the cued recall/extinction test, then again on Day 11 to confirm estrous stage identification accuracy. Each mouse was held by its tail, paws resting on the cage lid, while 0.2 ml of sterile water was pipetted in and out of the vaginal opening. The pipette did not penetrate the orifice. Water-containing cells was released onto a glass slide and left

to dry. Once dry, estrous smears were submerged in a 0.1% crystal violet solution for 1 min, then exposed to two 1-min rinses in distilled water prior to coverslipping. Slides were examined using brightfield microscopy. Cornified squamous epithelial cells, leukocytes, and nucleated epithelial cells were identified at magnifications of 40×, and their relative proportions were used to determine estrous stage as previously described (Byers et al., 2012).

2.13 | Estrous stage statistics

We created a binary analysis by combining diestrus and proestrus (stages in which levels of ovarian hormones are relatively high) as well as estrus and metestrus (stages in which levels of ovarian hormones are relatively low). To assess the effects of ovarian hormones on cued recall and extinction, mean freezing during the 33 CSs of the cued expression paradigm was averaged in blocks of 3. We compared freezing across these 11 CS bins using LMMs in the lme4 package of R v3.6.1 (R Core Team, 2019). Fixed effects were estrous stage, treatment, and CS, as well as all two-way and three-way interactions. We included mouse ID nested within testing cohort as a random intercept and CS as a random slope to account for the nonindependence of individuals measured repeatedly across time. Significance of fixed effects and model fit were assessed as above. Mean freezing during habituation trials was not included in our analyses. In order to explicitly compare the effect of ATMs within each estrous stage, we ran a priori contrasts using the emmeans package. We compared the difference in estimated marginal means between ATM and control groups within each estrous stage, as well as calculated the *p* value from Tukey post hoc comparisons to assess if ATMs exerted differential effects on females in diestrus/proestrus versus estrus/metestrus. Plots were generated as described above for % body weight change and liquid consumption.

3 | RESULTS

3.1 | Experiment 1: ATMs on aversive conditioning

3.1.1 | 16S rRNA sequencing

We obtained over 14 million reads from 53 samples, with >5,000 reads each. After denoising with Dada2, there were over 10 million reads and 52 samples with >5,000 reads. Calculation of the Shannon index diversity metric on rarefied samples revealed that 5,000 reads were sufficient to capture sample diversity, although observed ASVs and PD continued to rise steadily until 20,000 reads (Figure S1). All further analyses were performed using 52 samples containing at least 5,000 reads to balance sequencing depth and sample retention. These included 21 samples from 12 different female mice and 31 samples from 12 different male mice. Samples were collected longitudinally (from the same mice on Days 0, 3, 7, and 10).

3.1.2 | Gut microbiome structure

Microbiomes were characterized using multiple measures of diversity. PD describes the phylogenetic diversity of a community, ASV richness reports the number of ASVs (unique sequences), and evenness describes the similarity of ASV relative abundances of members of the community. LMMs revealed no differences between groups for any measure of α-diversity on Day 0 (Figure 2a). These data indicate that, prior to ATM administration, gut microbiomes were similar between females and males as well as between mice assigned to control and treatment groups. Initial analyses of α-diversity measures for treatment days (Days 3, 7, and 10) revealed no effect of day, so these days were collapsed. LMMs for treatment days revealed a main effect of treatment on PD ($\chi^2 = 10.5$, $p = 0.001$), richness ($\chi^2 = 21.8$, $p < 0.001$), and evenness ($\chi^2 = 23.8$, $p < 0.001$), with ATM mice showing reductions in all measures relative to controls. There was also a main effect of sex on PD ($\chi^2 = 9.6$, $p = 0.002$), richness ($\chi^2 = 4.9$, $p = 0.027$), and evenness ($\chi^2 = 4.8$, $p = 0.029$). The interaction of sex x treatment did not significantly affect evenness or richness, but it affected PD ($\chi^2 = 3.9$, $p = 0.047$) such that PD decreased more in ATM females than in ATM males. Together, these results suggest that female mice are more susceptible than males to the effects of ATM treatment on PD; otherwise, ATMs generally impact gut microbiomes similarly across sex.

β-diversity was assessed using several complementary measures. Prior to treatment (Day 0), PERMANOVAs of unweighted UniFrac, Bray–Curtis, Manhattan, Euclidean, and Canberra distances revealed no differences between groups (see Figure 2b for unweighted UniFrac and Bray–Curtis plots). Unexpectedly, PERMANOVA of weighted UniFrac distances revealed a difference between ATM and control microbiomes prior to treatment (pseudo-$F = 6.6$, $p = 0.006$), which was partly due to differences in dispersion ($F_{1,13} = 7.4$, $n = 15$, $p = 0.017$). Given the absence of differences in all other measures on Day 0, our small sample size at this time point, and the randomization of previously group-housed subjects into ATM and control groups, it is unlikely that this difference is meaningful.

Initial analyses of unweighted UniFrac and Bray–Curtis distances revealed no effect of day throughout ATM administration (Days 3, 7, and 10), suggesting that gut microbial communities in both sexes showed relatively rapid and persistent changes in response to ATM administration. Thus, treatment days were collapsed. PERMANOVA of unweighted UniFrac distances revealed an effect of treatment (pseudo-$F = 3.57$, $p = 0.001$) and sex (pseudo-$F = 2.61$, $p = 0.004$), with no interaction. Differences in dispersion (spread of the data) partially accounted for the effect of treatment ($F_{1,35} = 8.12$, $n = 37$, $p = 0.008$) but not sex, suggesting that differences between females and males were due entirely to varying positions in centroids (arithmetic mean position of the points). PERMANOVA of Bray–Curtis distances revealed an effect of treatment (pseudo-$F = 9.37$, $p = 0.01$) and sex (pseudo-$F = 3.44$, $p = 0.009$), with no interaction and no differences in dispersion. Figure 2b,c shows PC plots with distinct clusters by treatment in both females and males. In these plots, centroids (the larger of the two shape symbols) and the surrounding

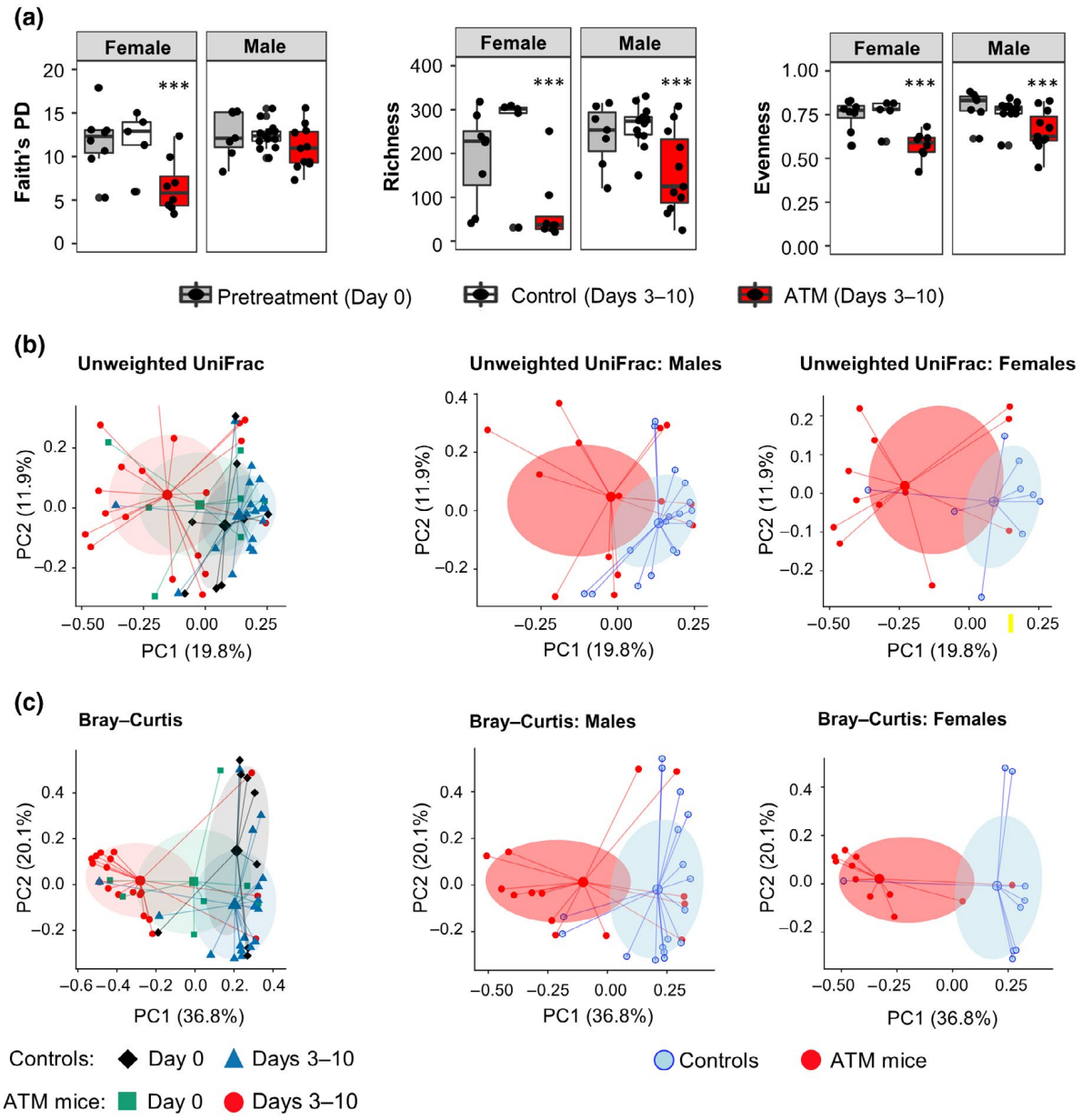


FIGURE 2 Antimicrobial medications (ATMs) induce gut dysbiosis in both sexes, with stronger effects in females. Microbiome structure was evaluated only in Experiment 1. (a) α-diversity metrics did not differ between ATM mice and controls on Day 0, so groups were collapsed as "Pretreatment." Effects did not differ within treatment groups across Days 3, 7, and 10, so these days were collapsed as "Control" or "ATM." Following ATM administration, only females showed reductions in phylogenetic diversity (PD). Both sexes showed reductions in ASV richness and evenness. Principal coordinate (PC) plots of (b) unweighted UniFrac and (c) Bray–Curtis distances revealed that microbiomes of ATM mice became dissimilar from controls following ATM administration. The left shows pretreatment (Day 0) and treatment (Days 3, 7, and 10) data for all control and ATM subjects. The center and right show the same data separated sex. PERMANOVAs of UniFrac and Bray–Curtis distances showed no significant differences prior to treatment. Following treatment, both showed the effects of sex and treatment ($p < 0.01$), with no interaction. Large shapes represent centroids, and small shapes represent samples. Shaded ovals indicate 1 standard deviation from the centroid. $n = 5$–13 samples/group, $^{***}p < 0.001$

transparent oval (which represents standard deviation) show the arithmetic mean position of the cluster. Dispersion can be visualized by the spread of the lines connecting samples (the smaller of the two shape symbols) to the centroid. A change in centroid position suggests a change in community structure. The plots clearly show a change in centroids in ATM mice of both sexes. Together, these results confirm that ATMs were ingested and induced states of gut dysbiosis in both females and males. They further suggest that states of dysbiosis were partially sex specific.

We next sought to characterize gut microbiome taxonomic shifts (Figure 3a). Microbiomes of controls of both sexes were dominated by Bacteroidetes and Clostridia. Microbiomes of ATM mice of both sexes showed reductions in Clostridia, as well as increases in Verrucomicrobia and Gammaproteobacteria.

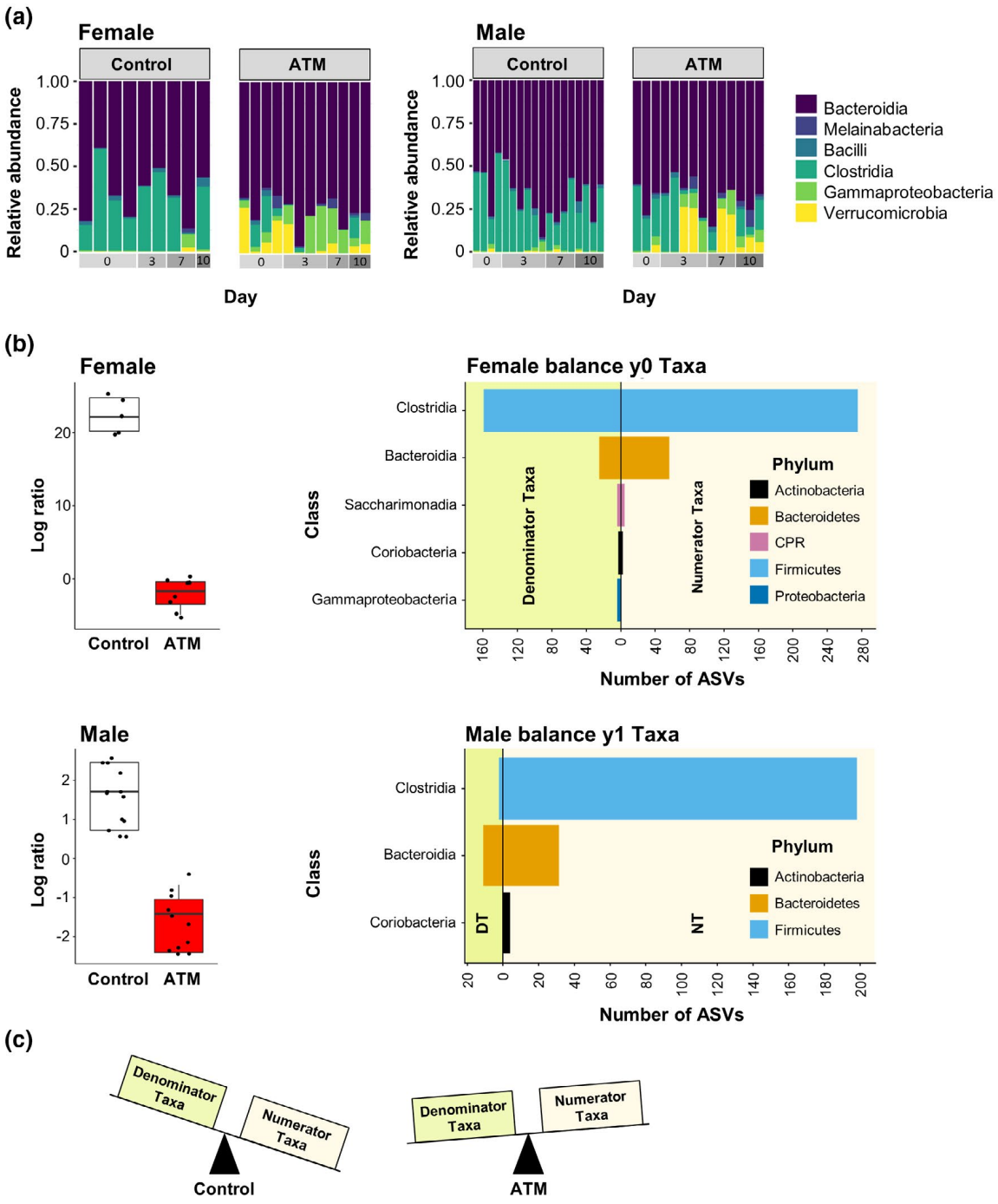


FIGURE 3 Antimicrobial medications (ATMs) alter gut microbiome taxonomic structure. (a) Relative abundance of major phyla (bacterial classes present in ≥1% total abundance) across the experiment. (b) Gneiss was used to calculate the log ratios of numerator:denominator taxa in both sexes to reveal differences in microbial subcommunities (called balances) between ATM mice and controls during treatment days. Shown for each sex are the log ratio of the largest significant balance (*left)* and the taxa comprising the balance (*right*). Taxa represented by four or less ASVs were excluded for clarity. Note that these values represent the taxonomic composition of the subcommunities, not the relative abundance of the ASVs. Both sexes showed a downward shift in the log ratio for ATM mice relative to controls, (c) visually depicted by the two scales. Given that we expect ATMs to deplete microbial populations, this shift most likely indicates that treatment either decreased relative abundance of numerator taxa, or it decreased relative abundance of both numerator and denominator taxa, with a larger decrease in the former. Females: $n = 13$/group, Males: $n = 25$/group

Differential abundance analysis was performed using Gneiss, which uses the concept of balance trees to determine shifts in microbial subcommunities described using proportional data (Morton et al., 2017). Because microbiome sequencing data are proportional—meaning that a change in one microbe will result in a change in the proportion of other microorganisms—it is difficult to infer which taxa are changing. The analysis begins with construction of a bifurcating tree using Ward's hierarchical clustering,

relating taxa based on the similarity of abundance across samples. The tips of the tree therefore represent the relative abundance of a taxon, and closely correlated taxa appear close together in the tree. An isometric log transformation is then performed to compute the log ratios between the groups of taxa, and a multivariate response linear regression is performed on these balances. A balance describes the subtrees (i.e. subcommunities) from a particular bifurcating node of the tree. A change in the structure of a microbial community can be detected by shifts in the balances (i.e. a community dominated by taxa in one subcommunity becomes dominated by members of another subcommunity). Linear modeling identified treatment as the main covariate explaining variation in the communities (10.9%). Sex explained 5.4% and treatment day explained 3.0%. Subsequent within-sex linear regressions revealed that treatment explained 22.0% of the variation among females and 11.4% of the variation among males. Balances comprising the subcommunities that show the largest shift in response to ATM treatment are shown in Figure 3b. A log ratio of 0 would indicate that the relative abundances of taxa in the two subcommunities were equal. Compared to controls, ATM females showed a downward shift in the log ratio of this balance, indicating a transition in dominance toward the subcommunity placed in the denominator and away from the subcommunity placed in the numerator. The composition of the subcommunities is shown in the panel on the right. We note that, while both subcommunities were primarily comprised of Clostridia and Bacteroidia, our data are presented at a high taxonomic rank. Thus, the species of Clostridia and Bacteroidia in one subcommunity are different from those in the other subcommunity. Compared to controls, ATM males showed a similar shift in the largest significant balance. The shift in dominance away from the subcommunity in the numerator and toward the subcommunity in the denominator is depicted in the scales in Figure 3c. Given that we expect ATMs to deplete microbial diversity, the most likely explanation for the shift is either a decreased abundance of taxa in the numerator subcommunity, or a decreased numerator and denominator taxa, with a greater decrease in numerator taxa.

3.1.3 | Open field test

Two-way ANOVA on total distance traveled revealed main effects of sex ($F_{1,76} = 5.1$; $p = 0.027$, $d = 0.61$) and treatment ($F_{1,76} = 7.0$; $p = 0.009$, $d = 0.74$), with no interaction (Figure 4a). Females ($n = 41$, 95% CI [37.9, 42.5]) traveled further than males ($n = 39$, 95% CI [33.2, 38.9]), and treated mice ($n = 39$, 95% CI [38.5, 42.6]) traveled further than controls ($n = 41$, 95% CI [32.9, 38.8]. Within-sex ANOVAs showed that ATMs increased distance traveled in males (treated: $n = 18$, 95% CI [37.7, 42.7]; controls: $n = 21$, 95% CI [27.9, 36.9]) (Welch $F_{1,30.4} = 9.9$, $p = 0.004$, $d = 0.84$), and did not significantly affect distance traveled in females ($d = 0.09$). These results suggest that ATMs increased locomotor activity in both sexes, with stronger effects in males.

Two-way ANOVA on % time spent in the field's center revealed a sex x treatment interaction ($F_{1,75} = 4.4$; $p = 0.039$; Figure 4a). Within males, % center time was increased in treated subjects ($n = 18$, 95% CI [5.7, 10.5]) relative to controls ($n = 21$, 95% CI [3.3, 6.1]) (Welch $F_{1,27.7} = 6.7$, $p = 0.015$, $d = 0.74$). We detected no effect of ATMs on % center time in females ($d = 0.10$). These results suggest that males are more susceptible than females to microbial modulation of anxiety-like behavior.

3.1.4 | Aversive learning

Two-way ANOVA revealed effects of sex ($F_{1,82} = 9.95$, $p = 0.002$, $d = 0.87$) and treatment ($F_{1,82} = 11.1$; $p < 0.001$, $d = 0.91$; Figure 4b), with no interaction. These data suggest that ATMs impaired learning of the CS/US association. Within-sex ANOVAs showed that treatment reduced freezing in ATM females ($n = 23$, 95% CI [32.5, 37.9]) relative to controls ($n = 22$, 95% CI [36.6, 48.9]) ($F_{1,43} = 6.3$; $p = 0.016$, $d = 0.79$). Treatment did not significantly affect freezing in males ($p = 0.098$, $d = 0.38$). There was no difference in freezing between female and male controls. Importantly, there were also no pre-CS/US differences in freezing. Overall, these results suggest that females are more susceptible than males to the effects of gut dysbiosis on aversive learning.

3.1.5 | Context recall

Two-way ANOVA detected no significant sex, treatment, or interaction effects (Figure 4c).

3.1.6 | Cued recall

Two-way ANOVA revealed an effect of treatment ($F_{1,70} = 9.2$; $p = 0.003$, $d = 0.82$), such that ATM mice froze less than controls (Figure 4d). This finding suggests that ATMs impaired memory retrieval of the CS/US association. There was no effect of sex, or interaction of sex × treatment. Within-sex ANOVAs showed that treated females froze less ($n = 19$, 95% CI [54.9, 75.7]) than controls ($n = 19$, 95% CI [73.4, 86.6]) ($F_{1,36} = 6.3$, $p = 0.017$, $d = 0.68$). Treatment did not significantly affect freezing in males ($p = 0.088$, $d = 0.40$). We detected no differences in freezing between female and male controls or between groups in the pre-CS (habituation) period. These results could indicate that females are more sensitive than males to microbial modulation of cued memory retrieval.

3.1.7 | Dendritic and spine morphology of BLA principal neurons

The effects of gut dysbiosis on aversive conditioning led us to investigate the potential changes in BLA principal neuron

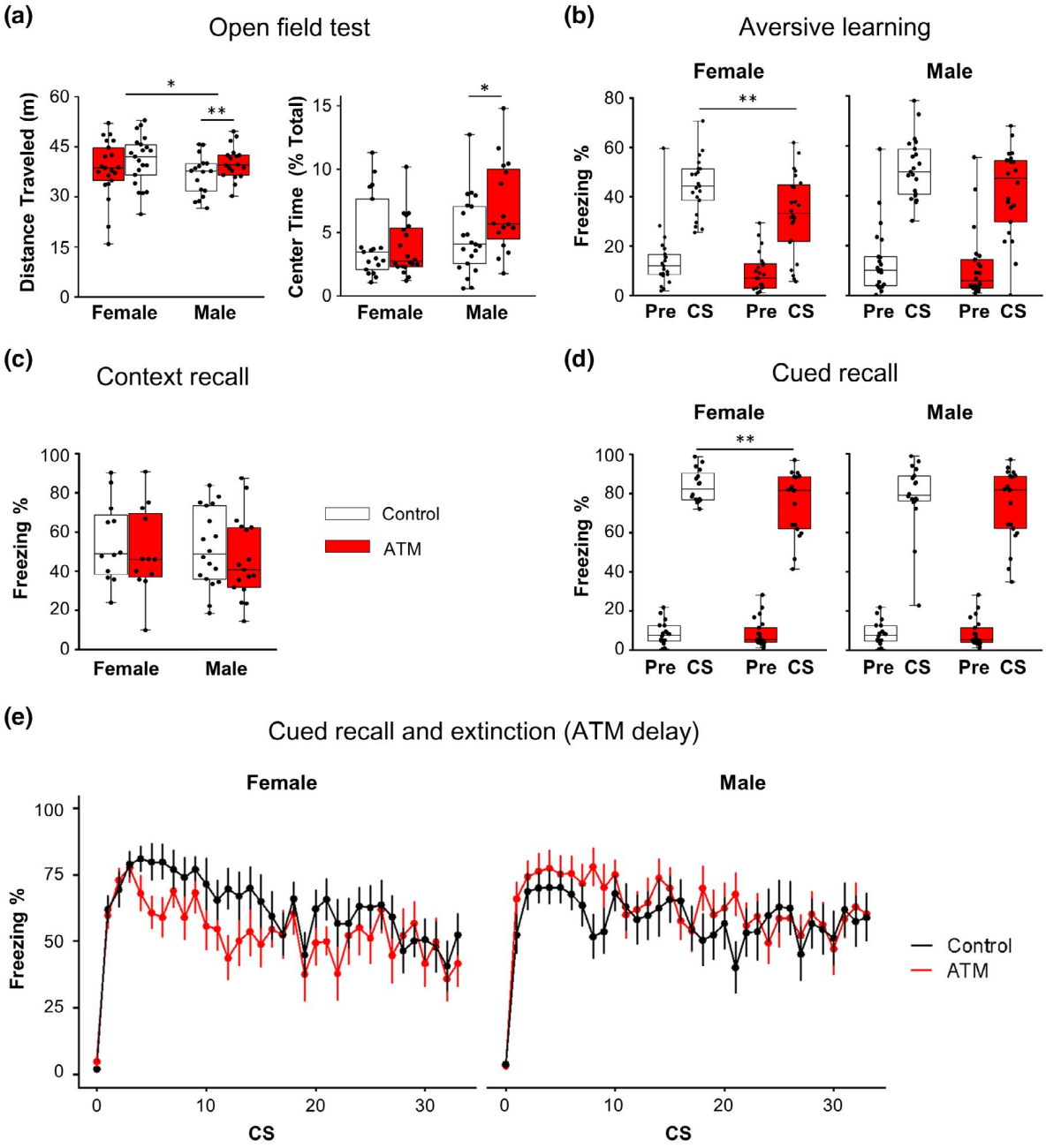


FIGURE 4 Female C57BL/6N mice are more susceptible than males to gut dysbiosis-driven impairment of cued aversive learning. The first four panels show data from Experiment 1. (a) Antimicrobial medications (ATMs) increased distance traveled in males more than females. Females traveled further than males overall ($p < 0.05$). A sex x treatment interaction indicated that ATMs increased % time spent in the arena's center only in males ($p < 0.05$; $n = 18–22$/group). (b) ATMs reduced freezing in both sexes during aversive learning, with stronger effects in females ($n = 21–23$/group). (c) No differences were detected in context recall ($n = 12–18$/group). (d) ATMs reduced freezing in both sexes in cued recall ($n = 19–21$/group), again with stronger effects in females. (e) In the "ATM Delay" experiment, treatment did not alter cued recall (CSs 1–3) or extinction learning in either sex ($n = 12$/group). CS 0 represents the habituation period. *$p < 0.05$

morphology. Generalized linear mixed effects modeling did not reveal the effects of ATMs on apical or basilar dendritic structure (Figure 5c), but effects were detected for dendritic spines. The analysis of spine density revealed a three-way interaction between branch order, sex, and treatment, indicating that spine density was affected by treatment and sex, but effects differed across branch order ($\chi^2 = 9.9$, $p = 0.019$; Figure 5e). Branch order analyses revealed effects at first- and third-order dendrites. For the first-order tree, there was a sex x treatment interaction ($\chi^2 = 4.4$, $p = 0.035$) which indicated that ATMs decreased spine density in females relative to males (post hoc comparison; $\chi^2 = 5.1$, $p = 0.025$). There were no differences between female and male controls. At the third-order tree, there were effects of sex ($\chi^2 = 4.0$, $p = 0.046$) and treatment ($\chi^2 = 4.0$, $p = 0.046$), but no

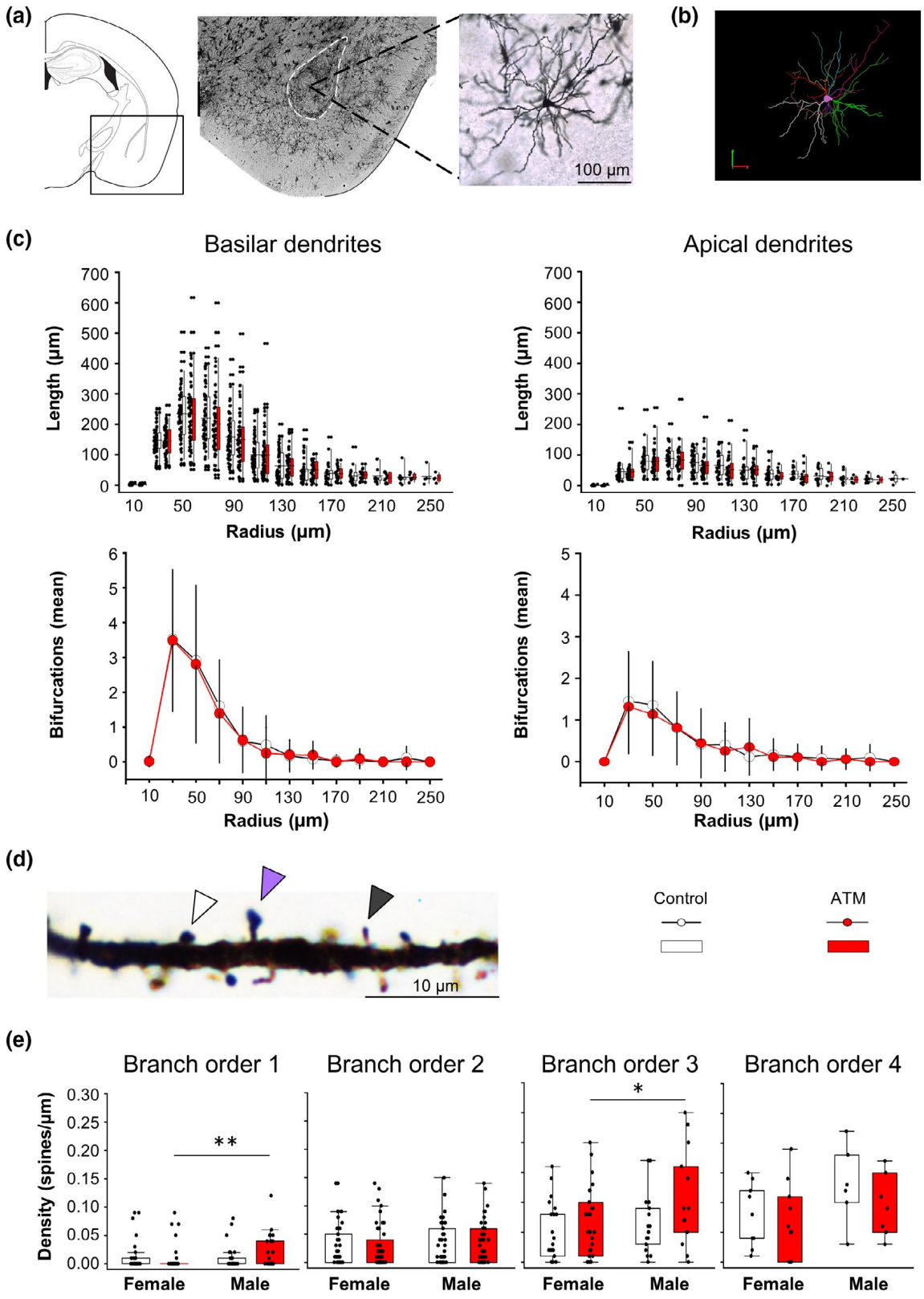


FIGURE 5 Morphometrics of Golgi-Cox-stained BLA principal neuron dendrites and spines. Brains were extracted only in Experiment 1. (a) The location of the BLA is shown in the atlas (*left*), and enclosed within the dashed lines in the micrograph (20×, *center*). Randomly sampled principal neurons (40×, *right*) were (b) reconstructed in 3D using Neurolucida. Mixed model analyses revealed no effects of sex or treatment on (c) average dendritic lengths or (d) bifurcations. (e) Spines (100×) were classified into three morphological types: stubby, mushroom, and thin (white, purple, and gray arrows, respectively). No overall differences were detected between types (data not shown). (f) No overall differences in spine density were detected, but ATMs exerted an opposing effect between sex at first-order branches, with reduced spine density in females relative to males. At third-order branches, overall spine density was higher in ATM mice relative to controls ($p < 0.05$), and higher in males than females. $n = 4–6$ mice/group and 35–42 reconstructed cells/group. $^{*}p < 0.05$, $^{**}p < 0.01$

interaction. Males had a higher density of spines than females, and treated mice had a higher density of spines than controls. No effects were detected at second- and fourth-order trees.

The analysis of spine morphology indicated that over 50% of spines on first-order branches in treated mice were "thin" type spines, compared with fewer than 10% in controls. There were no differences between groups across "stubby" and "mushroom" spine types.

3.1.8 | Relative body weight

LMMs revealed that relative body weight was affected by treatment ($\chi^2 = 140.1$, $p < 0.001$) and day ($\chi^2 = 8.82$, $p = 0.003$), as well as interactions between sex x day ($\chi^2 = 15.6$, $p < 0.001$) and treatment x day ($\chi^2 = 39.2$, $p < 0.001$; Figure S2a). Both sexes showed an initial drop in body weight in response to ATMs (Days 1–6). The effect of treatment on body weight appeared to be relatively similar in both sexes (estimated mean differences between ATM and control, females: 8.72, males: 9.71; Tukey post hoc comparisons, females: $t = -12.82$, $p < 0.001$, males: $t = -14.60$, $p < 0.001$).

3.2 | Experiment 2: ATMs on cued aversive recall and extinction ("ATM delay")

3.2.1 | Open field test

Two-way ANOVA on total distance traveled detected no significant effects, indicating that locomotor activity was not affected by treatment, and did not differ between sex (Figure S3). Two-way ANOVA on % center time revealed an increase in the ATM group relative to controls ($F_{1,44} = 4.8$; $p = 0.034$, $d = 0.57$), suggesting that gut dysbiosis reduced anxiety-like behavior in both sexes (Figure S3). There was no effect of sex, or interaction of sex × treatment. Within-sex analyses revealed that % center time between ATM mice and controls did not significantly differ in females ($p = 0.247$, $d = 0.21$), but there was a trend toward significance in males ($p = 0.075$, $d = 0.43$).

3.2.2 | Aversive learning and context recall

Two-way ANOVA detected no significant sex, treatment, or interaction effects in either aversive learning or context recall. These results indicate that there were no baseline differences between sex or treatment groups prior to ATM administration.

3.2.3 | Cued expression (recall and extinction)

Cued recall and extinction learning were each assessed using three-way mixed effects models. No significant sex, treatment, or interaction effects were detected (Figure 4e), although we did detect a trend toward a significant sex × treatment interaction ($\chi^2 = 3.5$, $p = 0.060$).

3.2.4 | Estrous cycle stage and cued extinction

LMMs revealed a main effect of CS on cued expression ($\chi^2 = 14.7$, $p < 0.001$). This result confirmed that freezing behavior declined across extinction learning. Importantly, we detected a CS × estrous stage interaction ($\chi^2 = 10.1$, $p = 0.001$; Figure S4b). The examination of the slopes in the trend lines of Figure S4b revealed that mice in estrus/metestrus froze more over time relative to mice in diestrus/proestrus. Post hoc analysis revealed a significant effect of ATMs on females in diestrus/proestrus relative to females in estrus/metestrus (Tukey post hoc comparisons: diestrus/proestrus, $t = -2.19$, $p = 0.041$; estrus/metestrus, $t = 0.34$, $p = 0.739$).

3.2.5 | Relative body weight

LMMs revealed that relative body weight was affected by sex ($\chi^2 = 7.5$, $p = 0.006$) and day ($\chi^2 = 11.2$, $p < 0.001$), as well as interactions between sex × day ($\chi^2 = 6.4$, $p = 0.011$) and treatment × day ($\chi^2 = 27.1$, $p < 0.001$; Figure S2b). These results indicate that body weight in both sexes was affected in a complex manner by ATMs and sex, and that these effects changed across time. The effect of ATMs on body weight was somewhat stronger in males than females (estimated mean differences between ATM and control, females: 3.29, males: 4.40; Tukey post hoc comparisons, females: $t = -2.55$, $p = 0.014$, males: $t = -3.41$, $p = 0.001$). For both sexes, differences in body weight between ATM and control groups were most pronounced between Days 4–6.

3.2.6 | Water consumption

LMMs revealed main effects of treatment ($\chi^2 = 33.8$, $p < 0.001$) and sex ($\chi^2 = 21.7$, $p < 0.001$), as well as a treatment × day interaction ($\chi^2 = 3.9$, $p = 0.047$; Figure S5). Due to the separate significant effects of treatment and sex, we were curious about how ATMs impacted treatment groups within each sex. We found that ATMs significantly reduced water consumption in both females and males, but the effect was almost twice as large in the former (estimated mean differences between ATM and control, females: 0.060, males: 0.033; Tukey post hoc comparisons, females: $t = -3.92$, $p < 0.001$, males: $t = -2.17$, $p = 0.035$).

4 | DISCUSSION

The past two decades have witnessed remarkable strides in understanding the role that gut microbiota play in human health, including in trauma and anxiety-related disorders. A growing body of research has given rise to a framework for how gut microbiota influence aversive conditioning and associated neurobiological mechanisms of plasticity (Cowan et al., 2018), but our understanding of these processes remains nascent. As is the case in much neuroscience research, sex

differences remain largely unexplored (Shansky & Woolley, 2016). In this study, we investigated the effects of ATM treatment on gut microbiome structure, aversive conditioning, anxiety-like behavior, and BLA principal neuron dendritic architecture in female and male B6 mice. We found that ATMs altered gut bacterial communities, open field behavior, auditory cued aversive learning and memory recall, and BLA dendritic spine patterning; many of these effects were influenced by sex. We detected no differences in context recall. Because impairments in cued recall could have been a result of impairments in aversive learning or memory consolidation, we conducted a second experiment to determine whether gut dysbiosis specifically impacts aversive memory retrieval. In this "ATM Delay" experiment, treated subjects showed no behavioral differences relative to controls, indicating that perturbations in gut microbiota were specific to aversive learning mechanisms.

4.1 | Gut microbiome and aversive conditioning

In the first experiment, ATM mice froze less than controls during both aversive acquisition and cued recall. Because no differences in cued recall were detected in the ATM Delay experiment, we conclude that deficits in memory retrieval were due to ATM-induced alterations in learning performance or memory consolidation. In support of this conclusion, ATM-induced changes in freezing in Experiment 1 were nearly identical during aversive learning (18.6%) and cued recall (14.5%). Interestingly, dysbiosis-related impairment of aversive learning was stronger in females than males. This finding, to our best knowledge, is novel.

We are also the first study to report reductions in aversive learning using an antimicrobial cocktail. Previous studies have demonstrated that other methods of altering the gut microbiome influence aversive conditioning in male mice. The probiotics *Bifidobacterium breve* and *B. longum* have been shown to augment aversive learning performance (Savignac et al., 2014), and both the latter strain and *Lactobacillus rhamnosus* have been shown to augment cued recall (Bravo et al., 2011). Two studies in B6 males—one examining GF status (Hoban et al., 2018) and another examining the effects of a fecal transplant from mice fed a high-fat diet (Bruce-Keller et al., 2015)—reported impairments in cued aversive recall, but not aversive learning. Another study in GF and gnotobiotic mice found no difference in either aversive learning or cued recall (Chu et al., 2019). Together, these data demonstrate that different microbial strains can exert diverse effects on cued aversive conditioning.

Whereas auditory cued aversive conditioning depends on the amygdala, context recall depends on the amygdala and hippocampus (Chaaya et al., 2019). Previous work has shown that microbiota modulate hippocampal structure and function, both in GF subjects (Clarke et al., 2013; Heijtz et al., 2011; Luczynski et al., 2016) as well as models of gut dysbiosis in developmentally normal adult rodents (Bercik et al., 2011; Bravo et al., 2011; Provensi et al., 2019; Savignac et al., 2015). However, we found no effects of ATMs on context recall, either in the pre-CS (habituation) period of cued recall or in the context recall test. Our data are consistent with previous work in male mice in GF (Hoban et al., 2018) and fecal transplant studies (Bruce-Keller et al., 2015), suggesting that hippocampal-dependent memory is less sensitive than amygdala-dependent conditioning to perturbations of the gut microbiome.

We assessed extinction in the second experiment. We found that females in diestrus/proestrus (high estrogen/high progesterone) exhibited facilitated extinction relative to females in estrus/metestrus (low estrogen/low progesterone). This finding supports previous work indicating that high levels of ovarian hormones augment aversive extinction (Milad et al., 2009; Velasco et al., 2019). We also found that ATM mice in diestrus/proestrus, but not estrus/metestrus, froze less than controls throughout cued expression. This result suggests that gut dysbiosis may exert distinct effects on cued aversive recall and extinction when ovarian hormone levels are high. In contrast with previous reports that the MGB axis modulates extinction in male rodents (Bravo et al., 2011; Chu et al., 2019; Fox et al., 2017; Hoban et al., 2018; Savignac et al., 2015), we observed no effect of ATMs on extinction in male mice. Inconsistent results could be due to differences in methodology or microbiome composition. Overall, our data indicate that adult females are more sensitive than males to the effects of gut dysbiosis on cued aversive conditioning and extinction.

4.2 | Gut microbiome and locomotor activity

Results from studies investigating microbial modulation of locomotor activity have been inconsistent. Many report no change in locomotor activity, regardless of methodological approach (Bruce-Keller et al., 2015; Burokas et al., 2017; De Palma et al., 2017; Neufeld et al., 2011a; Savignac et al., 2014; Zheng et al., 2016). This includes a study using a higher dose of the same ATM cocktail in BALB/c males (Bercik et al., 2011). Although ATM mice in our second experiment showed no difference in locomotor activity in the novel OFT, ATM mice in our first experiment showed enhanced locomotor activity. This effect was stronger in males than females. Our results support previous work in male GF mice (Heijtz et al., 2011; Nishino et al., 2013) as well as in female and male B6 mice (Bridgewater et al., 2017). Together, these data suggest that males may be more vulnerable than females to microbial modulation of locomotor activity.

Our OFT findings from the first experiment add to a large body of literature across rodent models, including B6 mice, showing that females tend to travel further than males (Bridgewater et al., 2017; Foldi et al., 2011; McCormick et al., 2005; Ramos et al., 2003). However, sex differences in murine locomotor activity are not ubiquitously reported (Davis et al., 2017; De Palma et al., 2017), and we detected no sex differences in our second experiment. Disparities between our experiments may be due to differences in methodology. Importantly, we questioned whether impairments in cued learning and recall in the first experiment could have been due to alterations in locomotor activity. This possibility—though it cannot be ruled out—is unlikely, as freezing did not differ between treated and control mice in the pre-CS (habituation) periods of conditioning tests.

4.3 | Gut microbiome and anxiety-like behavior

In the first experiment, ATMs increased OFT % center time only in males, suggesting that gut dysbiosis induced a sex-specific reduction in anxiety-like behavior. Results from the second experiment, while limited in comparability due to exposure to the fear learning paradigm, revealed increased OFT % center time in both sexes. Together, these results suggest that males are more susceptible than females to microbial modulation of anxiety-like behavior.

The anxiolytic phenotype we observed is consistent with numerous studies in GF and ATM male mice (Bercik et al., 2011; Clarke et al., 2013; Desbonnet et al., 2015; Heijtz et al., 2011; Zheng et al., 2016), suggesting that dysbiosis or absence of resident microbiota decreases anxiety-like behavior. Anxiety-like behavior in male rodents has also been shown to decrease in response to prebiotics and probiotics (Bravo et al., 2011; Burokas et al., 2017; Savignac et al., 2014), and to increase in response to fecal transplants from mice fed a high-fat diet (Bridgewater et al., 2017; Bruce-Keller et al., 2015) or from human patients with major depressive disorder (Kelly et al., 2016; Zheng et al., 2016). Collectively, these data suggest that the mechanisms underlying microbial modulation of trait anxiety are diverse, and specific to particular bacterial strains and interacting microbial communities.

Our finding that females are less vulnerable to microbial modulation of anxiety-like behavior than males is consistent with a small but growing body of preclinical evidence. One study reported an anxiolytic phenotype in female GF Swiss Webster mice, but colonization with resident microbiota in adulthood did not reverse this phenotype (Neufeld et al., 2011b) as it did in males of the same strain (Clarke et al., 2013). Since the latter study conventionalized subjects at 3 weeks of age, it is possible that this discrepancy reflects a critical developmental window, rather than a sex-specific effect. However, two studies in adult B6 mice—one assessing the effects of supplementation with an omega-3 polyunsaturated fatty acid (Davis et al., 2017) and another assessing the effects of a high-fat diet (Bridgewater et al., 2017)—also support a model of increased male vulnerability. Because anxiety is a multidimensional trait, future work should examine whether this effect of sex persists across multiple behavioral tests. We also note that indices of anxiety that are well-defined in males may be differentially expressed in females, and there remains a need to investigate potential sex-specific behavioral response profiles (Shansky, 2015).

4.4 | Gut microbiome and BLA principal neuron dendritic and spine morphology

The amygdala is a well-established network node for encoding, storage, and retrieval of aversive CS/US associations, and affiliated defensive behavioral responses (Bergstrom, 2016). The BLA plays a particularly crucial role in aversive conditioning, and also influences innate anxiety-like behavior (Cowan et al., 2018; Felix-Ortiz et al., 2013; Simpson et al., 2010). We found that ATMs increased BLA principal neuron dendritic spine density on third-order branches (49%) and exerted sex-specific effects on first-order branches, with females showing decreased spine density relative to males (162% difference). This finding is meaningful because branch orders may receive distinct inputs and possess varying densities and classes of receptors (Klenowski et al., 2015). Our data are consistent with a recent study indicating that BLA principal neurons in GF males show increased spine density (33%) across morphological types (Luczynski et al., 2016). The same study also observed elongated BLA principal neuron dendrites, with no differences in bifurcations. We did not detect changes in either of these measures, but it would be useful to explore whether a longer ATM regimen might yield different results. Overall, our results suggest that morphological alterations in BLA principal neuron spine density may underlie changes in aversive memory processing. Dendritic structure reflects function (Chklovskii et al., 2004), and the % spine density changes we observed after ATM treatment fall within and exceed ranges observed after fear conditioning (30%–35% increase) (Heinrichs et al., 2013; Lai et al., 2012; Ostroff et al., 2010; Radley et al., 2006). Together, our findings add to a growing number of studies indicating that microbiota influence amygdala structure and function (Bercik et al., 2011; Bravo et al., 2011; Chu et al., 2019; Heijtz et al., 2011; Hoban et al., 2016, 2017, 2018; Luczynski et al., 2016; Neufeld et al., 2011a).

4.5 | ATM-induced shifts in gut microbiome composition

We found that ATMs reduced phylogenetic diversity only in females, and perturbed ASV richness and evenness to a similar degree in both sexes. β-diversity analyses and PC plots provided further evidence that ATMs altered gut microbiome structure and introduced significant variation within microbial communities; effects did not differ between sex, except ATM females showed greater reductions in rare species. Taxonomic differences between ATM mice and controls were apparent by Day 3 of treatment, and did not differ across time in either sex. There were also no differences in any measure of alpha or β-diversity across Days 3–10. These data suggest that states of ATM-induced gut dysbiosis remained relatively stable across the experiment. Because controls also showed no differences in taxonomic or diversity measures across Days 3–10, our findings further suggest that exposure to our behavioral tests alone did not induce significant dysbiosis. We note that, while our approach did not encompass nonbacterial members of the gut microbiome (i.e., viruses and eukaryotes), bacteria are the most likely source of neuroactive components and metabolites (Glowacki & Martens, 2020). Overall, our findings confirmed that our ATM regimen induced gut dysbiosis in both sexes. We also found evidence that dysbiotic states associated with our cocktail are partially sex specific, with stronger effects in females.

4.6 | Limitations of ATM treatment

While ATM models of gut dysbiosis are plentiful, treatment regimens (specific ATMs used, doses, and lengths of administration) vary widely, and strain-dependent dosing remains to be established (Bercik et al., 2011; Jang et al., 2018; Kennedy et al., 2018; Odeh, 2013). Although our ATM cocktail was conservatively dosed relative to other studies using the same mixture (Bercik et al., 2011; Verdu et al., 2006), we found that ATMs mildly reduced body weight and water consumption in both sexes (Figures S1 and S2). Although the possibility that these changes affected our results cannot be ruled out, there are several reasons that they did not likely play a significant role. First, 25% water deprivation has not been shown to result in corticosterone response and physiological adaptation over 8 days (Bekkevold et al., 2013). Second, water deprivation has not been shown to affect aversive learning or cued recall in studies with comparable methodologies (three CS/US pairings; Maren et al., 1994; Pouzet et al., 2001). Third, our ATM mice showed no consistent changes in locomotor behavior, as might be expected if treatment exerted a toxic effect (Bercik et al., 2011) or if water deprivation was more severe (Tsunematsu et al., 2008). Finally, our subjects' weight loss remained under 15%, we observed no visible declines in animal health or behavior, and reductions in water consumption were relatively mild.

4.7 | Clinical implications

Trauma and anxiety-related disorders are major health burdens worldwide, with increased incidence in women relative to men (Altemus et al., 2014). Associated symptoms are highly prevalent in patients with dysbiosis-related syndromes like IBS, which also disproportionately affect women (Kibune-Nagasako et al., 2016; Lovell & Ford, 2012). Our data add to growing evidence for a role of gut microbiota in psychiatric conditions (Cowan et al., 2018; Cryan & Dinan, 2012; Kelly et al., 2016; Malan-Muller et al., 2018; MacQueen et al., 2017), and highlight the need to consider sex as a biological variable (Shansky & Woolley, 2016). While animal models can provide powerful insight into the mechanisms underlying disease states, caution should be exercised in translating results. MGB research is still in its infancy, and "dysbiotic patterns" remain poorly defined, and differ between species as well as conspecifics (Walter et al., 2020). Nonetheless, elucidation of the effects observed in this study could have important clinical utility moving into the future.

5 | CONCLUSIONS

We found novel evidence that female mice are more susceptible than males to the microbial modulation of aversive learning. This was supported by partially sex-specific effects of ATMs on gut microbiome composition and BLA dendritic spine density. Conversely, we found that females may be more resilient than males to the effects of gut dysbiosis on anxiety-like behavior. These findings advance our understanding of the influences that gut microbiota and sex can exert on various aspects of aversive conditioning and anxiety-related behavior during adulthood.

DECLARATION OF TRANSPARENCY

The authors, reviewers, and editors affirm that in accordance to the policies set by the *Journal of Neuroscience Research*, this manuscript presents an accurate and transparent account of the study being reported and that all critical details describing the methods and results are present.

ACKNOWLEDGMENTS

The authors are grateful to Kelli Duncan and Bojana Zupan (Vassar College) for their guidance on estrous stage determination and brain tissue processing, respectively. They are also grateful to Julie Williams and Yi Yuan (Vassar College) for their excellent care of the mouse colony. They thank Rebecca Shansky (Northeastern University) for helpful discussions on interpreting sex differences. This study was supported by grants from the Asprey Center for Collaborative Approaches in Science, the Lucy Maynard Salmon Research Fund at Vassar College, and the Ruth M. Berger Foundation.

CONFLICT OF INTEREST

The authors declare no competing interests.

AUTHOR CONTRIBUTIONS

All of the authors read and approved the final manuscript. *Conceptualization*, C.G.G.; *Methodology*, C.G.G., H.C.B., D.J.E., and V.C.W.; *Investigation*, V.C.W., K.L.B., P.O.J., T.B., V.B., C.G.G., C.B.B., A.F.P., A.M., H.M.T., and S.R.; *Formal Analysis*, J.C.T, H.C.B., D.J.E., V.C.W., and C.G.G.; *Resources*, H.C.B. and D.J.E.; *Writing – Original Draft*, C.G.G., H.C.B., D.J.E., J.C.T., and V.C.W.; *Writing – Review & Editing*, C.G.G., H.C.B., D.J.E., J.C.T., and V.C.W.; *Visualization*, H.C.B., D.J.E., V.C.W., and C.G.G.; *Supervision*, H.C.B. and D.J.E.; *Funding Acquisition*, H.C.B., D.J.E., and C.G.G.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1002/jnr.24848.

DATA AVAILABILITY STATEMENT

The microbiome sequence data in this study will be made available at the Sequence Read Archive (SRA) under BioProject ID: PRJNA646617. The BLA principal neuron dendrite reconstructions will be made available at NeuroMorpho.org. Other data that support our findings are available from the corresponding author upon reasonable request.

ORCID

Hannah Mae Thompson https://orcid.org/0000-0002-4587-5084

Hadley Creighton Bergstrom https://orcid.org/0000-0003-3677-1619

REFERENCES

Abdel-Haq, R., Schlachetzki, J. C., Glass, C. K., & Mazmanian, S. K. (2019). Microbiome–microglia connections via the gut–brain axis. *Journal of Experimental Medicine*, *216*(1), 41–59. https://doi.org/10.1084/jem.20180794

Altemus, M., Sarvaiya, N., & Epperson, C. N. (2014). Sex differences in anxiety and depression clinical perspectives. *Frontiers in Neuroendocrinology*, *35*(3), 320–330. https://doi.org/10.1016/j.yfrne.2014.05.004

Asok, A., Kandel, E. R., & Rayman, J. B. (2019). The neurobiology of fear generalization. *Frontiers in Behavioral Neuroscience*, *12*, 329. https://doi.org/10.3389/fnbeh.2018.00329

Bailey, K. R., & Crawley, J. N. (2009). Anxiety-related behaviors in mice. In J. J. Buccafusco (Ed.), *Methods of Behavior Analysis in Neuroscience*. Boca Raton: CRC Press/Taylor & Francis.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bekkevold, C. M., Robertson, K. L., Reinhard, M. K., Battles, A. H., & Rowland, N. E. (2013). Dehydration parameters and standards for laboratory mice. *Journal of the American Association for Laboratory Animal Science*, *52*(3), 233–239.

Bercik, P., Denou, E., Collins, J., Jackson, W., Lu, J., Jury, J., Deng, Y., Blennerhassett, P., Macri, J., McCoy, K. D., Verdu, E. F., & Collins, S. M. (2011). The intestinal microbiota affect central levels of brain-derived neurotropic factor and behavior in mice. *Gastroenterology*, *141*(2), 599–609.e3. https://doi.org/10.1053/j.gastro.2011.04.052

Bergstrom, H. C. (2016). The neurocircuitry of remote cued fear memory. *Neuroscience & Biobehavioral Reviews*, *71*, 409–417. https://doi.org/10.1016/j.neubiorev.2016.09.028

Bergstrom, H. C., McDonald, C. G., French, H. T., & Smith, R. F. (2008). Continuous nicotine administration produces selective, age-dependent structural alteration of pyramidal neurons from prelimbic cortex. *Synapse*, *62*(1), 31–39. https://doi.org/10.1002/syn.20467

Bergstrom, H. C., Smith, R. F., Mollinedo, N. S., & McDonald, C. G. (2010). Chronic nicotine exposure produces lateralized, age-dependent dendritic remodeling in the rodent basolateral amygdala. *Synapse*, *64*(10), 754–764. https://doi.org/10.1002/syn.20783

Blume, S. R., Freedberg, M., Vantrease, J. E., Chan, R., Padival, M., Record, M. J., DeJoseph, M. R., Urban, J. H., & Rosenkranz, J. A. (2017). Sex-and estrus-dependent differences in rat basolateral amygdala. *Journal of Neuroscience*, *37*(44), 10567–10586. https://doi.org/10.1523/JNEUROSCI.0758-17.2017

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, *37*(8), 852–857. https://doi.org/10.1038/s41587-019-0209-9

Bravo, J. A., Forsythe, P., Chew, M. V., Escaravage, E., Savignac, H. M., Dinan, T. G., Bienenstock, J., & Cryan, J. F. (2011). Ingestion of lactobacillus strain regulates emotional behavior and central GABA receptor expression in a mouse via the vagus nerve. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(38), 16050–16055. https://doi.org/10.1073/pnas.1102999108

Bridgewater, L. C., Zhang, C., Wu, Y., Hu, W., Zhang, Q., Wang, J., Li, S., & Zhao, L. (2017). Gender-based differences in host behavior and gut microbiota composition in response to high fat diet and stress in a mouse model. *Scientific Reports*, *7*(1), 1–12. https://doi.org/10.1038/s41598-017-11069-4

Brooks, M. E., Kristensen, K., Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Mächler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, *9*(2), 378–400. https://doi.org/10.32614/RJ-2017-066

Bruce-Keller, A. J., Salbaum, J. M., Luo, M., Blanchard, E. IV, Taylor, C. M., Welsh, D. A., & Berthoud, H. (2015). Obese-type gut microbiota induce neurobehavioral changes in the absence of obesity. *Biological Psychiatry*, *77*(7), 607–615. https://doi.org/10.1016/j.biopsych.2014.07.012

Burokas, A., Arboleya, S., Moloney, R. D., Peterson, V. L., Murphy, K., Clarke, G., Stanton, C., Dinan, T. G., & Cryan, J. F. (2017). Targeting the microbiota-gut-brain axis: Prebiotics have anxiolytic and antidepressant-like effects and reverse the impact of chronic stress in mice. *Biological Psychiatry*, *82*(7), 472–487. https://doi.org/10.1016/j.biopsych.2016.12.031

Byers, S. L., Wiles, M. V., Dunn, S. L., & Taft, R. A. (2012). Mouse estrous cycle identification tool and images. *PLoS ONE*, *7*(4), e35538. https://doi.org/10.1371/journal.pone.0035538

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from illumina amplicon data. *Nature Methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Chaaya, N., Jacques, A., Belmer, A., Beecher, K., Ali, S. A., Chehrehasa, F., Battle, A. R., Johnson, L. R., & Bartlett, S. E. (2019). Contextual fear conditioning alter microglia number and morphology in the rat dorsal hippocampus. *Frontiers in Cellular Neuroscience*, *13*, 214. https://doi.org/10.3389/fncel.2019.00214

Chklovskii, D. B., Mel, B. W., & Svoboda, K. (2004). Cortical rewiring and information storage. *Nature*, *431*(7010), 782–788.

Chu, C., Murdock, M. H., Jing, D., Won, T. H., Chung, H., Kressel, A. M., & Zhou, L. (2019). The microbiota regulate neuronal function and fear extinction learning. *Nature*, *574*(7779), 543–548.

Clarke, G., Grenham, S., Scully, P., Fitzgerald, P., Moloney, R. D., Shanahan, F., Dinan, T. G., & Cryan, J. F. (2013). The microbiome-gut-brain axis during early life regulates the hippocampal serotonergic system in a sex-dependent manner. *Molecular Psychiatry*, *18*(6), 666–673. https://doi.org/10.1038/mp.2012.77

Cowan, C. S., Hoban, A. E., Ventura-Silva, A. P., Dinan, T. G., Clarke, G., & Cryan, J. F. (2018). Gutsy moves: The amygdala as a critical node in microbiota to brain signaling. *BioEssays*, *40*(1), 1700172. https://doi.org/10.1002/bies.201700172

Crawley, M. J. (2012). *The R book*. John Wiley & Sons.

Cryan, J. F., & Dinan, T. G. (2012). Mind-altering microorganisms: The impact of the gut microbiota on brain and behaviour. *Nature Reviews Neuroscience*, *13*(10), 701–712. https://doi.org/10.1038/nrn3346

Cryan, J. F., & O'Mahony, S. M. (2011). The microbiome-gut-brain axis: From bowel to behavior. *Neurogastroenterology & Motility*, *23*(3), 187–192. https://doi.org/10.1111/j.1365-2982.2010.01664.x

Davis, D. J., Hecht, P. M., Jasarevic, E., Beversdorf, D. Q., Will, M. J., Fritsche, K., & Gillespie, C. H. (2017). Sex-specific effects of docosahexaenoic acid (DHA) on the microbiome and behavior of socially-isolated mice. *Brain, Behavior, and Immunity*, *59*, 38–48. https://doi.org/10.1016/j.bbi.2016.09.003

De Palma, G., Lynch, M. D. J., Lu, J., Dang, V. T., Deng, Y., Jury, J., Umeh, G., Miranda, P. M., Pigrau Pastor, M., Sidani, S., Pinto-Sanchez, M. I., Philip, V., McLean, P. G., Hagelsieb, M.-G., Surette, M. G., Bergonzelli, G. E., Verdu, E. F., Britz-McKibbin, P., Neufeld, J. D., … Bercik, P. (2017). Transplantation of fecal microbiota from patients with irritable bowel syndrome alters gut function and behavior in recipient

mice. *Science Translational Medicine*, *9*(379), eaaf6397. https://doi.org/10.1126/scitranslmed.aaf6397

DeGruttola, A. K., Low, D., Mizoguchi, A., & Mizoguchi, E. (2016). Current understanding of dysbiosis in disease in human and animal models. *Inflammatory Bowel Diseases*, *22*(5), 1137–1150. https://doi.org/10.1097/MIB.0000000000000750

Desbonnet, L., Clarke, G., Traplin, A., O'Sullivan, O., Crispie, F., Moloney, R. D., Cotter, P. D., Dinan, T. G., & Cryan, J. F. (2015). Gut microbiota depletion from early adolescence in mice: Implications for brain and behaviour. *Brain, Behavior, and Immunity*, *48*, 165–173. https://doi.org/10.1016/j.bbi.2015.04.004

Desrochers, C. S., & Schacht, J. (1982). Neomycin concentrations in inner ear tissues and other organs of the guinea pig after chronic drug administration. *Acta Oto-Laryngologica*, *93*(1–6), 233–236. https://doi.org/10.3109/00016488209130877

Ding, T., & Schloss, P. D. (2014). Dynamics and associations of microbial community types across the human body. *Nature*, *509*(7500), 357–360.

El Aidy, S., Stilling, R., Dinan, T. G., & Cryan, J. F. (2016). Microbiome to brain: Unravelling the multidirectional axes of communication. In *Microbial endocrinology: Interkingdom signaling in infectious disease and health* (pp. 301–336). Springer. https://doi.org/10.1007/978-3-319-20215-0_15

Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, *61*(1), 1–10. https://doi.org/10.1016/0006-3207(92)91201-3

Felix-Ortiz, A. C., Beyeler, A., Seo, C., Leppla, C. A., Wildes, C. P., & Tye, K. M. (2013). BLA to vHPC inputs modulate anxiety-related behaviors. *Neuron*, *79*(4), 658–664. https://doi.org/10.1016/j.neuron.2013.06.016

Foldi, C. J., Eyles, D. W., McGrath, J. J., & Burne, T. H. (2011). The effects of breeding protocol in C57BL/6J mice on adult offspring behaviour. *PLoS ONE*, *6*(3), e18152. https://doi.org/10.1371/journal.pone.0018152

Forsythe, P., Kunze, W., & Bienenstock, J. (2016). Moody microbes or fecal phrenology: What do we know about the microbiota-gut-brain axis? *BMC Medicine*, *14*(1), 58. https://doi.org/10.1186/s12916-016-0604-8

Fox, J. H., Hassell, J. E., Siebler, P. H., Arnold, M. R., Lamb, A. K., Smith, D. G., Day, H. E. W., Smith, T. M., Simmerman, E. M., Outzen, A. A., Holmes, K. S., Brazell, C. J., & Lowry, C. A. (2017). Preimmunization with a heat-killed preparation of mycobacterium vaccae enhances fear extinction in the fear-potentiated startle paradigm. *Brain, Behavior, and Immunity*, *66*, 70–84. https://doi.org/10.1016/j.bbi.2017.08.014

Gibb, R., & Kolb, B. (1998). A method for vibratome sectioning of Golgi-Cox stained whole rat brain. *Journal of Neuroscience Methods*, *79*(1), 1–4. https://doi.org/10.1016/S0165-0270(97)00163-5

Glowacki, R. W., & Martens, E. C. (2020). In sickness and health: Effects of gut microbial metabolites on human physiology. *PLOS Pathogens*, *16*(4), e1008370. https://doi.org/10.1371/journal.ppat.1008370

Heijtz, R. D., Wang, S., Anuar, F., Qian, Y., Bjorkholm, B., Samuelsson, A., Hibberd, M. L., Forssberg, H., & Pettersson, S. (2011). Normal gut microbiota modulates brain development and behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(7), 3047–3052. https://doi.org/10.1073/pnas.1010529108

Heinrichs, S. C., Leite-Morris, K. A., Guy, M. D., Goldberg, L. R., Young, A. J., & Kaplan, G. B. (2013). Dendritic structural plasticity in the basolateral amygdala after fear conditioning and its extinction in mice. *Behavioural Brain Research*, *248*, 80–84. https://doi.org/10.1016/j.bbr.2013.03.048

Hoban, A. E., Moloney, R. D., Golubeva, A. V., McVey Neufeld, K. A., O'Sullivan, O., Patterson, E., Stanton, C., Dinan, T. G., Clarke, G., & Cryan, J. F. (2016). Behavioural and neurochemical consequences of chronic gut microbiota depletion during adulthood in the rat. *Neuroscience*, *339*, 463–477. https://doi.org/10.1016/j.neuroscience.2016.10.003

Hoban, A. E., Stilling, R. M., M. Moloney, G., Moloney, R. D., Shanahan, F., Dinan, T. G., Cryan, J. F., & Clarke, G. (2017). Microbial regulation of microRNA expression in the amygdala and prefrontal cortex. *Microbiome*, *5*(1), 102. https://doi.org/10.1186/s40168-017-0321-3

Hoban, A. E., Stilling, R. M., Moloney, G., Shanahan, F., Dinan, T. G., Clarke, G., & Cryan, J. F. (2018). The microbiome regulates amygdala-dependent fear recall. *Molecular Psychiatry*, *23*(5), 1134–1144. https://doi.org/10.1038/mp.2017.100

Jaggar, M., Rea, K., Spicak, S., Dinan, T. G., & Cryan, J. F. (2019). You've got male: Sex and the microbiota-gut brain axis across the lifespan. *Frontiers in Neuroendocrinology*, *56*, 100815.

Jang, H., Lee, H., Jang, S., Han, M. J., & Kim, D. (2018). Evidence for interplay among antibacterial-induced gut microbiota disturbance, neuroinflammation, and anxiety in mice. *Mucosal Immunology*, *11*(5), 1386–1397. https://doi.org/10.1038/s41385-018-0042-3

Jašarević, E., Morrison, K. E., & Bale, T. L. (2016). Sex differences in the gut microbiome–brain axis across the lifespan. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1688), 20150122. https://doi.org/10.1098/rstb.2015.0122

Jones, R., Latinovic, R., Charlton, J., & Gulliford, M. (2006). Physical and psychological co-morbidity in irritable bowel syndrome: A matched cohort study using the general practice research database. *Alimentary Pharmacology & Therapeutics*, *24*(5), 879–886. https://doi.org/10.1111/j.1365-2036.2006.03044.x

Jørgensen, B. (1987). Exponential dispersion models. *Journal of the Royal Statistical Society: Series B (Methodological)*, *49*(2), 127–145.

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4), 772–780. https://doi.org/10.1093/molbev/mst010

Kawoos, Y., Wani, Z. A., Kadla, S. A., Shah, I. A., Hussain, A., Dar, M. M., Margoob, M. A., & Sideeq, K. (2017). Psychiatric co-morbidity in patients with irritable bowel syndrome at a tertiary care center in Northern India. *Journal of Neurogastroenterology and Motility*, *23*(4), 555–560. https://doi.org/10.5056/jnm16166

Kelly, J. R., Borre, Y., O' Brien, C., Patterson, E., El Aidy, S., Deane, J., Kennedy, P. J., Beers, S., Scott, K., Moloney, G., Hoban, A. E., Scott, L., Fitzgerald, P., Ross, P., Stanton, C., Clarke, G., Cryan, J. F., & Dinan, T. G. (2016). Transferring the blues: Depression-associated gut microbiota induces neurobehavioural changes in the rat. *Journal of Psychiatric Research*, *82*, 109–118. https://doi.org/10.1016/j.jpsychires.2016.07.019

Kennedy, E. A., King, K. Y., & Baldridge, M. T. (2018). Mouse microbiota models: Comparing germ-free mice and antibiotics treatment as tools for modifying gut bacteria. *Frontiers in Physiology*, *9*, 1534. https://doi.org/10.3389/fphys.2018.01534

Kibune-Nagasako, C., García-Montes, C., Silva-Lorena, S. L., & Aparecida-Mesquita, M. (2016). Irritable bowel syndrome subtypes: Clinical and psychological features, body mass index and comorbidities. *Revista Española De Enfermedades Digestivas*, *108*(2), 59–64.

Klenowski, P. M., Fogarty, M. J., Belmer, A., Noakes, P. G., Bellingham, M. C., & Bartlett, S. E. (2015). Structural and functional characterization of dendritic arbors and GABAergic synaptic inputs on interneurons and principal cells in the rat basolateral amygdala. *Journal of Neurophysiology*, *114*(2), 942–957. https://doi.org/10.1152/jn.00824.2014

Lai, C. S. W., Franke, T. F., & Gan, W. B. (2012). Opposite effects of fear conditioning and extinction on dendritic spine remodelling. *Nature*, *483*(7387), 87–91.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*(1), 155–184. https://doi.org/10.1146/annurev.neuro.23.1.155

Lenth, R. (2020). *Emmeans: Estimated marginal means, aka least-squares means*. https://CRAN.R-project.org/package=emmeans

Lovell, R. M., & Ford, A. C. (2012). Effect of gender on prevalence of irritable bowel syndrome in the community: Systematic review and meta-analysis. *American Journal of Gastroenterology*, *107*(7), 991–1000. https://doi.org/10.1038/ajg.2012.131

Lozupone, C., & Knight, R. (2005). UniFrac: A new phylogenetic method for comparing microbial communities. *Applied and Environment Microbiology*, *71*(12), 8228–8235. https://doi.org/10.1128/AEM.71.12.8228-8235.2005

Luczynski, P., Whelan, S. O., O'Sullivan, C., Clarke, G., Shanahan, F., Dinan, T. G., & Cryan, J. F. (2016). Adult microbiota-deficient mice have distinct dendritic morphological changes: Differential effects in the amygdala and hippocampus. *European Journal of Neuroscience*, *44*(9), 2654–2666. https://doi.org/10.1111/ejn.13291

Lyte, M. (2013). Microbial endocrinology in the microbiome-gut-brain axis: How bacterial production and utilization of neurochemicals influence behavior. *PLoS Pathogens*, *9*(11), e1003726. https://doi.org/10.1371/journal.ppat.1003726

MacQueen, G., Surette, M., & Moayyedi, P. (2017). The gut microbiota and psychiatric illness. *Journal of Psychiatry & Neuroscience*, *42*(2), 75–77. https://doi.org/10.1503/jpn.170028

Malan-Muller, S., Valles-Colomer, M., Raes, J., Lowry, C. A., Seedat, S., & Hemmings, S. M. (2018). The gut microbiome and mental health: Implications for anxiety-and trauma-related disorders. *OMICS: A Journal of Integrative Biology*, *22*(2), 90–107. https://doi.org/10.1089/omi.2017.0077

Maren, S., DeCola, J. P., & Fanselow, M. S. (1994). Water deprivation enhances fear conditioning to contextual, but not discrete, conditional stimuli in rats. *Behavioral Neuroscience*, *108*(3), 645–649. https://doi.org/10.1037/0735-7044.108.3.645

Markle, J. G., Frank, D. N., Mortin-Toth, S., Robertson, C. E., Feazel, L. M., Rolle-Kampczyk, U., & Danska, J. S. (2013). Sex differences in the gut microbiome drive hormone-dependent regulation of autoimmunity. *Science*, *339*(6123), 1084–1088.

Martin, C. R., Osadchiy, V., Kalani, A., & Mayer, E. A. (2018). The brain-gut-microbiome axis. *Cellular and Molecular Gastroenterology and Hepatology*, *6*(2), 133–148. https://doi.org/10.1016/j.jcmgh.2018.04.003

McCormick, C. M., Robarts, D., Kopeikina, K., & Kelsey, J. E. (2005). Long-lasting, sex-and age-specific effects of social stressors on corticosterone responses to restraint and on locomotor responses to psychostimulants in rats. *Hormones and Behavior*, *48*(1), 64–74. https://doi.org/10.1016/j.yhbeh.2005.01.008

McMurdie, P. J., & Holmes, S. (2013). Phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE*, *8*(4), e61217. https://doi.org/10.1371/journal.pone.0061217

Milad, M. R., Igoe, S. A., Lebron-Milad, K., & Novales, J. E. (2009). Estrous cycle phase and gonadal hormones influence conditioned fear extinction. *Neuroscience*, *164*(3), 887–895. https://doi.org/10.1016/j.neuroscience.2009.09.011

Morton, J. T., Sanders, J., Quinn, R. A., McDonald, D., Gonzalez, A., Vázquez-Baeza, Y., Navas-Molina, J. A., Song, S. J., Metcalf, J. L., Hyde, E. R., Lladser, M., Dorrestein, P. C., & Knight, R. (2017). Balance trees reveal microbial niche differentiation. *mSystems*, *2*(1), 162. https://doi.org/10.1128/mSystems.00162-16

Moshitch, D., & Nelken, I. (2014). Using tweedie distributions for fitting spike count data. *Journal of Neuroscience Methods*, *225*, 13–28. https://doi.org/10.1016/j.jneumeth.2014.01.004

Neufeld, K. M., Kang, N., Bienenstock, J., & Foster, J. A. (2011a). Reduced anxiety-like behavior and central neurochemical change in germ-free mice. *Neurogastroenterology & Motility*, *23*(3), 255-e119. https://doi.org/10.1111/j.1365-2982.2010.01620.x

Neufeld, K. M., Kang, N., Bienenstock, J., & Foster, J. A. (2011b). Effects of intestinal microbiota on anxiety-like behavior. *Communicative & Integrative Biology*, *4*(4), 492–494. https://doi.org/10.4161/cib.15702

Nishino, R., Mikami, K., Takahashi, H., Tomonaga, S., Furuse, M., Hiramoto, T., Aiba, Y., Koga, Y., & Sudo, N. (2013). Commensal microbiota modulate murine behaviors in a strictly contamination-free environment confirmed by culture-based methods. *Neurogastroenterology & Motility*, *25*(6), 521-e371. https://doi.org/10.1111/nmo.12110

Odeh, S. (2013). *The impact of antibiotics on the gut-brain axis* (Doctoral Dissertation). McMaster University.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., & Wagner, H. (2019). *Community ecology package*. R Package Version. https://CRAN.R-project.org/package=vegan

Org, E., Mehrabian, M., Parks, B. W., Shipkova, P., Liu, X., Drake, T. A., & Lusis, A. J. (2016). Sex differences and hormonal effects on gut microbiota composition in mice. *Gut Microbes*, *7*(4), 313–322. https://doi.org/10.1080/19490976.2016.1203502

Ostroff, L. E., Cain, C. K., Bedont, J., Monfils, M. H., & LeDoux, J. E. (2010). Fear and safety learning differentially affect synapse size and dendritic translation in the lateral amygdala. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(20), 9418–9423. https://doi.org/10.1073/pnas.0913384107

Parks, D. H., & Beiko, R. G. (2013). Measures of phylogenetic differentiation provide robust and complementary insights into microbial communities. *ISME Journal*, *7*(1), 173–183. https://doi.org/10.1038/ismej.2012.88

Peirce, J. M., & Alviña, K. (2019). The role of inflammation and the gut microbiome in depression and anxiety. *Journal of Neuroscience Research*, *97*(10), 1223–1241. https://doi.org/10.1002/jnr.24476

Pouzet, B., Zhang, W., Richmond, M. A., Rawlins, J. N. P., & Feldon, J. (2001). The effects of water deprivation on conditioned freezing to contextual cues and to a tone in rats. *Behavioural Brain Research*, *119*(1), 49–59. https://doi.org/10.1016/S0166-4328(00)00337-5

Prager, E. M., Bergstrom, H. C., Grunberg, N. E., & Johnson, L. R. (2011). The importance of reporting housing and husbandry in rat research. *Frontiers in Behavioral Neuroscience*, *5*, 38. https://doi.org/10.3389/fnbeh.2011.00038

Prager, E. M., Chambers, K. E., Plotkin, J. L., McArthur, D. L., Bandrowski, A. E., Bansal, N., Martone, M. E., Bergstrom, H. C., Bespalov, A., & Graf, C. (2019). Improving transparency and scientific rigor in academic publishing. *Journal of Neuroscience Research*, *97*(4), 377–390. https://doi.org/10.1002/jnr.24340

Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2–approximately maximum-likelihood trees for large alignments. *PLoS ONE*, *5*(3), e9490. https://doi.org/10.1371/journal.pone.0009490

Provensi, G., Schmidt, S. D., Boehme, M., Bastiaanssen, T. F. S., Rani, B., Costa, A., Busca, K., Fouhy, F., Strain, C., Stanton, C., Blandina, P., Izquierdo, I., Cryan, J. F., & Passani, M. B. (2019). Preventing adolescent stress-induced cognitive and microbiome changes by diet. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(19), 9644–9651. https://doi.org/10.1073/pnas.1820832116

R Core Team. (2019). *A language and environment for statistical computing*. R Foundation for Statistical Computing.

Radley, J. J., Johnson, L. R., Janssen, W. G. M., Martino, J., Lamprecht, R., Hof, P. R., LeDoux, J. E., & Morrison, J. H. (2006). Associative Pavlovian conditioning leads to an increase in spinophilin-immunoreactive dendritic spines in the lateral amygdala. *European Journal of Neuroscience*, *24*(3), 876–884. https://doi.org/10.1111/j.1460-9568.2006.04962.x

Ramos, A., Correia, E. C., Izídio, G. S., & Brüske, G. R. (2003). Genetic selection of two new rat lines displaying different levels of anxiety-related behaviors. *Behavior Genetics*, *33*(6), 657–668.

Rousham, E. K., Unicomb, L., & Islam, M. A. (2018). Human, animal and environmental contributors to antibiotic resistance in low-resource settings: Integrating behavioural, epidemiological and one health approaches. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1876), 20180332.

Savignac, H. M., Kiely, B., Dinan, T. G., & Cryan, J. F. (2014). Bifidobacteria exert strain-specific effects on stress-related behavior and physiology in BALB/c mice. *Neurogastroenterology & Motility*, *26*(11), 1615–1627.

Savignac, H. M., Tramullas, M., Kiely, B., Dinan, T. G., & Cryan, J. F. (2015). Bifidobacteria modulate cognitive processes in an anxious mouse strain. *Behavioural Brain Research*, *287*, 59–72. https://doi.org/10.1016/j.bbr.2015.02.044

Shansky, R. M. (2015). Sex differences in PTSD resilience and susceptibility: Challenges for animal models of fear learning. *Neurobiology of Stress*, *1*, 60–65. https://doi.org/10.1016/j.ynstr.2014.09.005

Shansky, R. M., & Woolley, C. S. (2016). Considering sex as a biological variable will be valuable for neuroscience research. *Journal of Neuroscience*, *36*(47), 11817–11822. https://doi.org/10.1523/JNEUROSCI.1390-16.2016

Simpson, H. B., Neria, Y., Lewis-Fernández, R., & Schneier, F. (2010). *Anxiety disorders: Theory, research and clinical perspectives*. Cambridge University Press.

Sudo, N., Chida, Y., Aiba, Y., Sonoda, J., Oyama, N., Yu, X.-N., Kubo, C., & Koga, Y. (2004). Postnatal microbial colonization programs the hypothalamic–pituitary–adrenal system for stress response in mice. *Journal of Physiology*, *558*(1), 263–275. https://doi.org/10.1113/jphysiol.2004.063388

Teng, P., & Meleney, F. L. (1950). Bacitracin levels in the cerebrospinal fluid after parenteral injections: Bacitracin therapy of experimental staphylococcal meningitis in the dog. *Surgery*, *27*(3), 403–413.

Tsunematsu, T., Fu, L., Yamanaka, A., Ichiki, K., Tanoue, A., Sakurai, T., & van den Pol, A. N. (2008). Vasopressin increases locomotion through a V1a receptor in orexin/hypocretin neurons: Implications for water homeostasis. *Journal of Neuroscience*, *28*(1), 228–238. https://doi.org/10.1523/JNEUROSCI.3490-07.2008

van de Guchte, M., Blottière, H. M., & Doré, J. (2018). Humans as holobionts: Implications for prevention and therapy. *Microbiome*, *6*(1), 81. https://doi.org/10.1186/s40168-018-0466-8

Van Der Waaij, D., Berghuis-de Vries, J. M., & Altes, C. K. (1974). Oral dose and faecal concentration of antibiotics during antibiotic decontamination in mice and in a patient. *Epidemiology & Infection*, *73*(2), 197–203.

Van der Waaij, D., & Sturm, C. A. (1968). Antibiotic decontamination of the digestive tract of mice. Technical procedures. *Laboratory Animal Care*, *18*(1), 1.

Velasco, E. R., Florido, A., Milad, M. R., & Andero, R. (2019). Sex differences in fear extinction. *Neuroscience & Biobehavioral Reviews*, *103*, 81–108. https://doi.org/10.1016/j.neubiorev.2019.05.020

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics* (4th ed.). Springer.

Verdu, E. F., Bercik, P., Verma-Gandhu, M., Huang, X., Blennerhassett, P., Jackson, W., & Collins, S. M. (2006). Specific probiotic therapy attenuates antibiotic induced visceral hypersensitivity in mice. *Gut*, *55*(2), 182–190. https://doi.org/10.1136/gut.2005.066100

Walter, J., Armet, A. M., Finlay, B. B., & Shanahan, F. (2020). Establishing or exaggerating causality for the gut microbiome: Lessons from human microbiota-associated rodents. *Cell*, *180*(2), 221–232. https://doi.org/10.1016/j.cell.2019.12.025

Wilson, M. D., Sethi, S., Lein, P. J., & Keil, K. P. (2017). Valid statistical approaches for analyzing sholl data: Mixed effects versus simple linear models. *Journal of Neuroscience Methods*, *279*, 33–43. https://doi.org/10.1016/j.jneumeth.2017.01.003

Yuste, R. (2010). *Dendritic spines*. MIT Press.

Zaqout, S., & Kaindl, A. M. (2016). Golgi-cox staining step by step. *Frontiers in Neuroanatomy*, *10*, 38. https://doi.org/10.3389/fnana.2016.00038

Zheng, P., Zeng, B., Zhou, C., Liu, M., Fang, Z., Xu, X., Zeng, L., Chen, J., Fan, S., Du, X., Zhang, X., Yang, D., Yang, Y., Meng, H., Li, W., Melgiri, N. D., Licinio, J., Wei, H., & Xie, P. (2016). Gut microbiome remodeling induces depressive-like behaviors through a pathway mediated by the host's metabolism. *Molecular Psychiatry*, *21*(6), 786–796. https://doi.org/10.1038/mp.2016.44

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the Supporting Information section.

Transparent Science Questionnaire for Authors

Transparent Peer Review Report

FIGURE S1 α-diversity rarefaction curves confirm that bacterial communities were sufficiently sampled. Rarefaction plots of α-diversity metrics were created to evaluate sufficiency of sampling depth. Samples were rarefied in QIIME2 with 10 replicates at each depth. As depth increases, samples with fewer total sequences are dropped from the analysis. A plateau suggests sufficient sampling depth. Faith's phylogenetic diversity and richness are more sensitive to rare species and sequencing errors, and are not expected to fully plateau. (a) Shannon Index, (b) Faith's Phylogenetic Diversity, and (c) observed ASVs (richness). Dark Blue: ATM-treated, Light Blue: Controls

FIGURE S2 ATMs decrease body weight in both sexes. (a) Relative to controls, both female (n = 23/group) and male (n = 24/group) ATM mice in Experiment 1 showed reductions in body weight, particularly across Days 1–6. Effects were relatively similar across sex. (b) Treatment also reduced body weight in females (n = 12/group) and males (n = 12/group) in Experiment 2, particularly across Days 4–6. In this "ATM Delay" experiment, effects appeared to be stronger in males

FIGURE S3 ATMs alter time spent in the center of the open field in the ATM Delay experiment. Locomotor activity did not significantly differ between treatment groups or sex. ATMs increased % center time in both sexes. There was a trend toward a significant increase in ATM males relative to male controls. n = 12/group, $^{*}p < 0.05$

FIGURE S4 Ovarian hormones modulate aversive extinction and interact with ATMs to influence cued recall. Estrous stage and aversive extinction were evaluated only in Experiment 2. (a) Micrographs (40×) of 0.1% crystal violet-stained vaginal cells during (i) proestrus, (ii) estrus, (iii) metestrus, and (iv) diestrus. (b) Bin 0 represents the habituation period, which was not included in our analyses. Bins 1–11 each represent the mean of three CS trials. The slopes of the trend lines in each panel reveal that mice in diestrus/proestrus froze less than mice in estrus/metestrus during the latter CS presentations, suggesting that high levels of ovarian hormones facilitate aversive extinction. We also found that ATM mice in diestrus/proestrus, but not estrus/metestrus, froze less than controls throughout cued expression ($p < 0.05$ and $p > 0.05$,

respectively). This suggests that high levels of ovarian hormones may interact with gut dysbiosis to attenuate cued recall or augment aversive extinction. *n* = 6/group

FIGURE S5 ATMs attenuate voluntary water consumption in females more than males. Voluntary water consumption was assessed only in Experiment 2. To determine whether treatment influenced water consumption, we measured each mouse's daily water bottle volume loss. Since fluid intake can differ across body size, we normalized data by dividing water loss (ml) by the mouse's daily weight (g). Both sexes showed mild reductions in drinking behavior, though this effect was nearly twice as pronounced in females as in males. Females: *n* = 12/group, Males: *n* = 12/group

How to cite this article: Geary CG, Wilk VC, Barton KL, et al. Sex differences in gut microbiota modulation of aversive conditioning, open field activity, and basolateral amygdala dendritic spine density. *J Neurosci Res*. 2021;00:1–22. https://doi.org/10.1002/jnr.24848