Supplementary material

Mom's diet matters: Maternal prebiotic intake in mice reduces anxiety and alters brain gene expression and the fecal microbiome in offspring.

Jenna C Hebert[a], Daniel E Radford-Smith[b], Fay Probert[b], Nicholas Ilott[c], Ka Wai Chan[b], Daniel C Anthony[b*], Philip WJ Burnet[a*†]

*These authors made an equal contribution

[a.] Department of Psychiatry, University of Oxford, Warneford Hospital, Warneford Lane, Oxford OX3 7JX, UK

[b.] Department of Pharmacology, University of Oxford, Mansfield Road, Oxford OX1 3QT, UK

[c.] Oxford Centre for Microbiome Studies, Kennedy Institute of Rheumatology, University of Oxford, Roosevelt Drive, Oxford OX3 7FY, UK

†Address correspondence to: phil.burnet@psych.ox.ac.uk

Philip WJ Burnet, PhD

Department of Psychiatry University of Oxford Warneford Lane

Oxford, OX3 7JX United Kingdom

Keywords: galacto-oligosaccharide; short-chain fatty acids; early life; NMDAR; prefrontal cortex; hippocampus; perinatal; 16S sequencing; gut-brain axis; NMR

Supplementary methods

Breeding

Male and female mice were housed in individually ventilated cages (IVCs). After habituating to the animal facility for at least one week, they were organized into mating pairs. Pairs were checked daily for vaginal plugs. The date a vaginal plug was found was noted as embryonic day [E]0.

B-GOS administration

When a vaginal plug was found, females were randomly assigned to receive either normal drinking water or water supplemented with 1.5 mg/mL B-GOS (approximately 0.25 mg/g per day). This dose was chosen because it was shown to increase *Bifidobacterium* levels in female mice in a small pilot study (**Fig. S1A**). Testing multiple doses on all the outcomes in these experiments would have required a prohibitive number of mice, and we did not have an *a priori* primary outcome for a more suitable pilot test. This is a key limitation of the study. In future studies, however, a dose-response could be possible using select outcomes that changed significantly after maternal B-GOS, such as rearing in the OFT, fecal SCFAs and NMDAR expression. A dose-response curve would also bolster the conclusions from this study.

The B-GOS was weighed in a weigh boat and added to 200 mL water. The bottles were shaken with one finger over the spout to prevent spillage. The weigh boat was then rinsed with 50 mL water to collect any remaining powder into the bottle. The bottles were rinsed out and refilled Monday, Wednesday, and Friday every week.

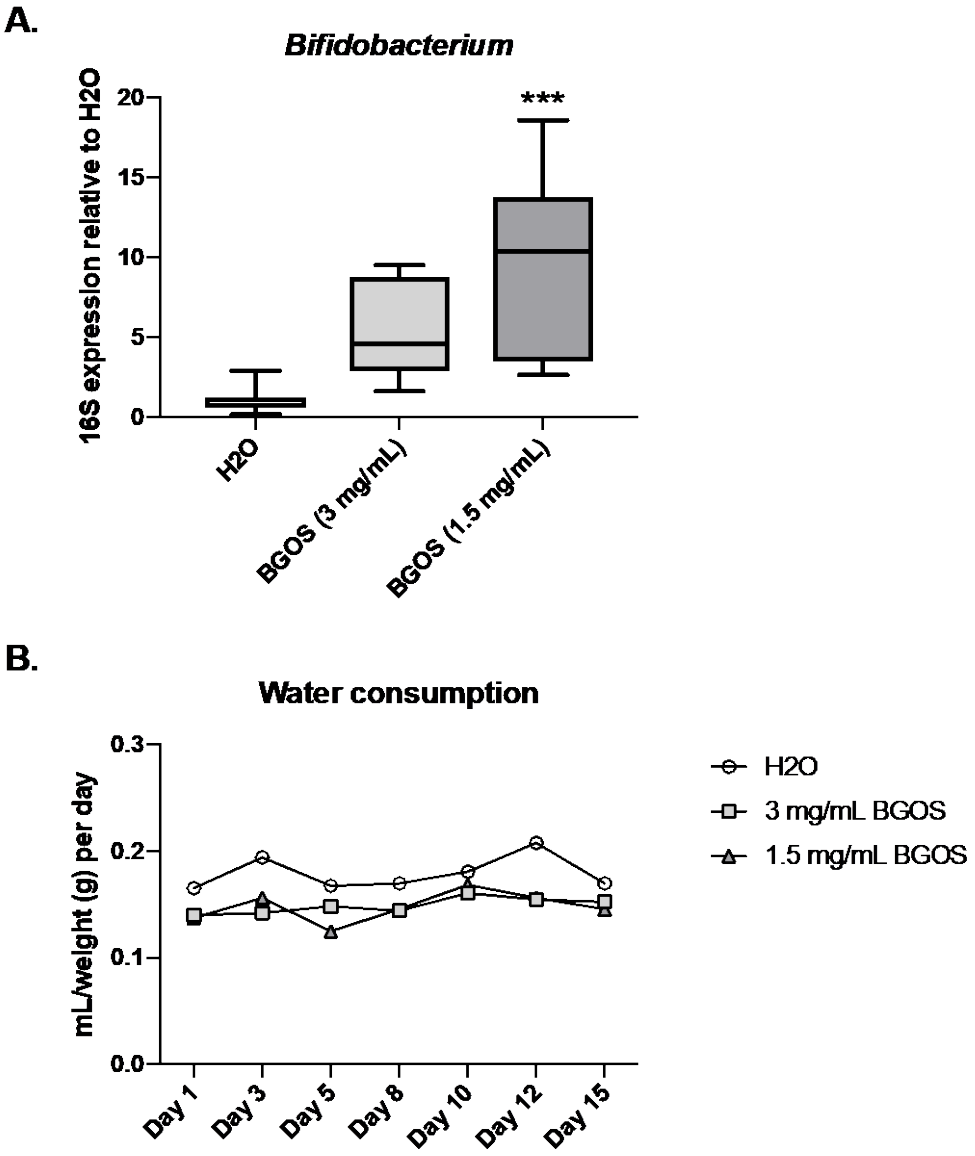

Fig. S1. B-GOS dose pilot study. A small study was performed to determine an appropriate dose of prebiotic for supplementation in the drinking water. The outcome of interest was *Bifidobacterium* levels relative to total bacteria (measured using a universal 16S primer) in fecal samples. C57Bl/6 dams (N = 4/group) were administered either normal drinking water, 3 mg/mL B-GOS, or 1.5 mg/mL B-GOS. Fecal samples were collected, and DNA was extracted for qPCR. *Bifidobacterium* levels were analyzed by normalizing expression of *Bifidobacterium*-specific 16S rRNA to total bacteria. **A)** According to a one-way ANOVA, there were significant differences in relative *Bifidobacterium* levels between groups, $F(2, 20) = 8.71$, $p = 0.002$. Post-hoc comparisons showed that the 1.5 mg/mL dose of B-GOS resulted in a significant increase in *Bifidobacterium* relative to the control group, $p = 0.0009$, while the 3 mg/mL dose was not significant relative to controls, $p = 0.06$. **B)** In addition, the volume of water consumed was monitored for the first two weeks of B-GOS administration to confirm that the prebiotic did not influence water consumption. There were several limitations to this study. Group numbers were low, and all animals in each group were housed in the same cage. In addition, these females were not pregnant/suckling. These aspects of the pilot study design may explain why we did not see a significant increase in *Bifidobacterium* in B-GOS-fed dams. Nevertheless, this dose of B-GOS did influence offspring development and gut microbiome composition.

LPS administration

Lipopolysaccharide (LPS) from *Escherichia coli* O111:B4 (Sigma-Aldrich) was dissolved the morning of injections in sterile 0.9% sodium chloride. For intraperitoneal (I.P.) injections in pups at weaning age, LPS was administered in a single dose of 1 mg/kg. For subcutaneous (S.C.) injections for P9 pups, LPS was delivered at 3 mg/kg. Pups were randomly assigned to receive saline or LPS.

Behavior

Open field test (OFT)

In the open field test (OFT, **Fig. S2A**), when the test is performed for a brief period of time (2–10 minutes), distance traveled and rearing are measures of exploration of a novel environment. A rear occurs when the mouse elevates its forepaws, keeping the hind paws on the floor, to assess the environment. In addition, mice tend to stay close to the walls, a phenomenon known as thigmotaxis. The time spent in the central zone of the open field is therefore a proxy for reduced anxiety-like behavior (Gould, 2009). This test was performed as described previously (Couch et al., 2015) for five minutes. Distance traveled (cm) and time in center zone (s) were measured with ANY-maze software (Stoelting Co.). Total number of rears was scored manually by an independent observer blinded to the treatment groups.

Light-dark box (LDB)

The light-dark box (LDB, **Fig. S2B**), measures the balance between the mouse's tendency to explore and neophobia (avoidance of the unfamiliar). Increases in exploration in the light compartment – time spent, latency to enter, entries to the light side, and rearing in the light – correspond to reduced anxiety-like behavior (Gould, 2009). This test was performed as described

previously (Savignac et al., 2016) for five minutes. Time in the light side (s), latency to enter the light (s), number of entries into the light, and number of rears were scored manually.

Forced swim test (FST)

The forced swim test (FST, **Fig. S2C**) was first described by Porsolt and colleagues as a novel screen for antidepressant activity (Porsolt et al., 1977), but it has poor construct validity as a model of depression (Gould, 2009). For this reason, we refer to floating as passive stress coping or behavioral despair. This test was performed as previously described (Couch et al., 2015) with one modification – each mouse was placed in the apparatus for a total of five minutes instead of six to reduce the distress experienced by the animals. Time floating (s) and latency to float (s) were scored in the last three minutes. Floating behavior was defined as the minimal amount of movement necessary to stay afloat.

Three-chamber social interaction test (SIT)

Mice are social, curious animals, and they tend to approach unfamiliar conspecifics. Genetic modification or maternal immune activation can reduce interaction time, and reduced social preference in mice is a correlate of impaired social approach behavior in autism (Yang et al., 2016)

The social interaction test (SIT, **Fig. S2D**) was adopted from a published protocol (Yang et al., 2016). The apparatus was a transparent red Plexiglas box (65x30x20 cm^3, *Lxwxh*) divided into three equal-sized chambers separated by sliding doors. The lighting of the room was measured to be 2 lux. The test consisted of two parts: a habituation phase and a testing phase. In the morning, during the ten-minute habituation phase, mice were placed in the central chamber with the doors closed. The doors were opened after five minutes to allow the animal to explore the entire apparatus. In the afternoon, for the ten-minute testing phase, a novel C57Bl/6 mouse of the same

sex was placed in a wired cup in one chamber, and an empty wired cup was placed in the other chamber, serving as a novel object. The mouse was placed in the center of the apparatus, and the doors were immediately opened. The mice were tracked with ANY-maze software. The social preference score was calculated as the percentage of time spent in the chamber with the novel mouse compared to the total time exploring either the novel mouse chamber or the novel object chamber. The apparatus and wired cups were cleaned with Anistel, an animal facility disinfectant, and water in between tests.

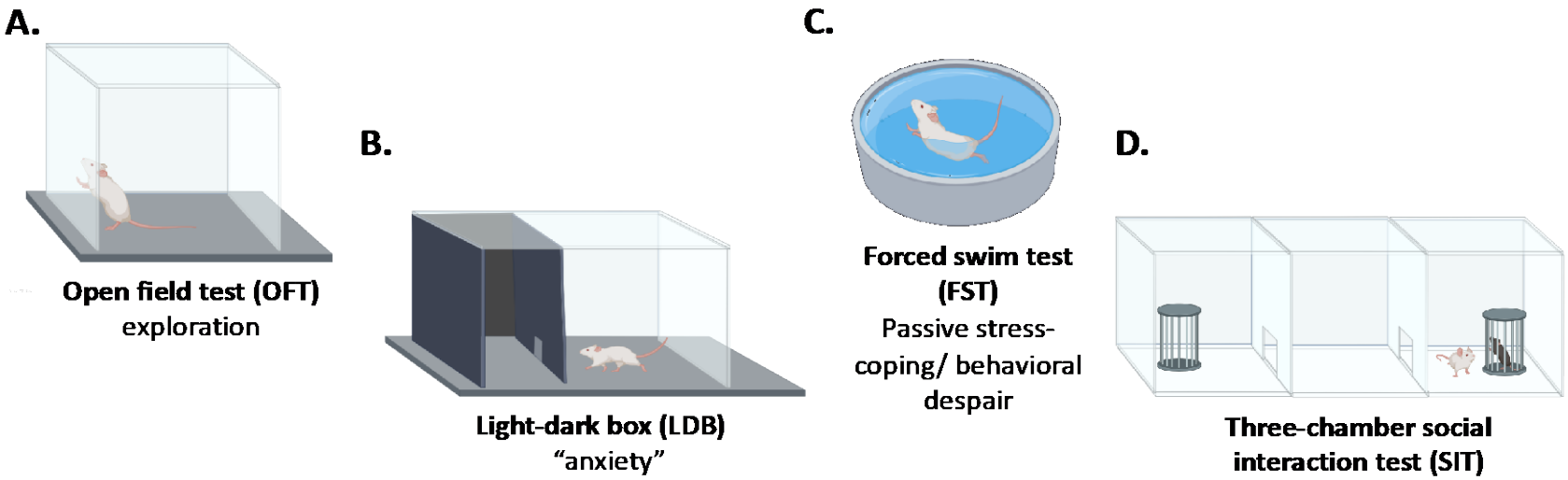

Fig. S2. Schematics of mouse behavioral tests in pups. A) The OFT, **B)** LDB, **C)** FST, and **D)** SIT were performed in offspring.

Maternal care behavior

Assessing maternal care behavior is useful in understanding the influence of a maternal treatment or challenge on offspring. In this study, it was necessary to test the possibility that the prebiotic feeding made the dams more or less attentive to their pups, thus changing behavior and brain gene expression in offspring. Home cage observation is the most useful measure of maternal care because it assesses the dam's behavior in a natural environment without disrupting the nest. It can be challenging for the experimenter, however, to differentiate between similar behaviors (grooming and nursing, for example). Quantifying the latency to retrieve pups that are dispersed from the nest by the experimenter is another common technique. This test is advantageous because it is quicker and less subjective. In this case, however, the disruption to the nest makes the test less ethologically relevant than the home cage observation test (Franks et al., 2011). Given that these paradigms both have advantages and disadvantages, both tests were used to assess maternal care behavior.

Home cage observation

An abbreviated version of a previously published protocol was used (Franks et al., 2011). Mice were transported to the observation room and left to acclimate for 30 minutes. Maternal care behavior was observed on P2 and P4 between 8:00–9:00 and 17:00–18:00. Behavioral observations were recorded every two minutes for a total of 120 observations per dam. The experimenter was blinded to treatment conditions. The following behaviors were scored: nursing, licking/grooming, nest building, and instances of being absent from the nest.

Pup retrieval test

After the home cage observation was complete, the dam was moved to a separate cage. Three pups were removed from the nest, each placed in a different corner of the home cage. The dam was then returned to the home cage. The latency to retrieve the first pup and the time to retrieve all pups were recorded.

Tissue collection

After the last behavioral test, animals were anesthetized with isoflurane gas. Blood samples were taken via cardiac puncture to collect plasma as previously described (Jurynczyk et al., 2017). Mice were then transcardially perfused with saline and heparin, brains were removed, and the prefrontal cortex (PFC) and hippocampus were dissected using a scalpel blade and forceps and immediately frozen on dry ice. In our study, the PFC encompassed the anterior cingulate, prelimbic, infralimbic, and orbital areas (Carlén, 2017).

To collect fecal samples, the intestines were removed, and fecal pellets were dissected from the colon and immediately frozen.

RNA and DNA extraction and quantitative PCR (qPCR)

PFC and hippocampi were homogenized in a Tissue Lyser (Qiagen), and RNA was extracted with Tri Reagent (Sigma-Aldrich) according to the manufacturer's instructions and converted to cDNA using a commercial kit (Life Technologies). For DNA extraction, fecal samples were homogenized in 10 µL/mg 6M guanidine hydrochloride (Sigma-Aldrich) using the Tissue Lyser. Samples were centrifuged for five minutes at room temperature. The supernatant was removed, and 10% silicon dioxide (Sigma-Aldrich) in nuclease-free water was added to isolate the DNA. The pellet was washed with 100 mM

NaCl in 70% ethanol, then re-suspended in nuclease-free H_2O and heated in a water bath at 55 °C. The DNA solution was isolated, and DNA quality and concentration were determined using a Qubit Fluorometer (Thermo Fisher).

SYBR Green (Life Technologies) was used for detection of amplified DNA. qPCR reactions were run using the Applied Biosystems QuantStudio 6 Flex Real-Time PCR System (Thermo Scientific). A melt curve was generated after each run to confirm a single PCR product. Transferrin receptor (TFRC) was used as a housekeeping gene for brain regions. Raw Ct values for reference genes were compared across groups to ensure that there were no effects of treatment on reference gene expression. All samples were run in triplicate, and triplicates labeled as outliers by the QuantStudio software were removed before analysis. $2^{-\Delta Ct}$ was calculated for each sample, where $\Delta Ct = (Ct_{\text{target gene}} - Ct_{\text{reference gene}})$. Fold change in gene expression was then calculated by dividing each group by the average $2^{-\Delta Ct}$ value for the control group. The primer sequences are shown in **Table S1**.

Table S1. qPCR primer sequences.

Gene or bacterial taxa	Forward sequence (5' to 3')	Reverse sequence (5' to 3')	Amplicon size (bp)	Reference
TNF-α	GCCTCCCTCTCATCAGTTCTAT	TTTGCTACGACGTGGGCTA	94	(Couch et al. 2015)
GluNR1	ATCATCCTGCTGGTCAGCGA	AGCAGAGCCGTCACATTCTT	137	Verrall 2008*
GluNR2A	GCTTTCCTTGAACCCTTCAG	GGGGAGCTTTCCCTTTGGCTAAGTT	142	Verrall 2008*
GluNR2B	TTGGTGAGGTGGTCATGAAG	GGCTCTAAGAAGGCAGAAGGT	166	Verrall 2008*
BDNF	AGCTGAGCGTGTGTGACAGT	ACCCATGGGATTACACTTGG	161	(Musazzi et al., 2009)
5-HT2A	CAGGCAAGTCACAGGATAGC	TTAAGCAGAAAGAAAATCCCACAG	93	(Couch et al., 2015)
TFRC	TGGGTCTAAGTCTACAATGGCT	CCCTCATGACGAATCTGTTTG	114	Verrall 2008*
Beta actin	CGTTGACATCCGTAAAGACC	AACAGTCCGCCTAGAAGCAC	281	(Ji et al. 2013)
Bifidobacterium	TCGCGTCYGGTGTGAAAG	CCACATCCAGCRTCCAC	243	(Rinttilä et al., 2004)
16S	CGGCAACGAGCGCAACCC	CCATTGTAGCACGTGTGTAGCC	130	(Denman and McSweeney, 2006)

*Primers were designed using the Primer 3 software (*primer3.ut.ee*, run by Dr Louise Verrall, DPhil Thesis 2008). PCR products spanned exon-exon boundaries such that any contaminating genomic DNA would be unlikely to amplify given its size or, if it were, would be distinguishable by size from mRNA PCR products. Specificity of primers for each transcript was confirmed using the nucleotide-nucleotide alignment search (*http://ncbi.nlm.nih.gov/BLAST/*), and product sequencing (Department of Biochemistry, University of Oxford).

^{1}H Nuclear Magnetic Resonance (NMR) metabolic profiling

^{1}H NMR measures the magnetic field at which protons are at resonance (**Fig. S3A**). Different compounds have characteristics peaks on the NMR spectrum depending on the chemical structure. ^{1}H NMR can be used in an untargeted manner to assess the metabolic "fingerprint" of a sample or group of samples. One-dimensional (1D) ^{1}H NMR is the quickest and most basic technique (**Fig. S3B**), while 2D total correlated spectroscopy

(TOCSY) ^{1}H NMR can be used to differentiate between metabolites with overlapping peaks (**Fig. S3C**).

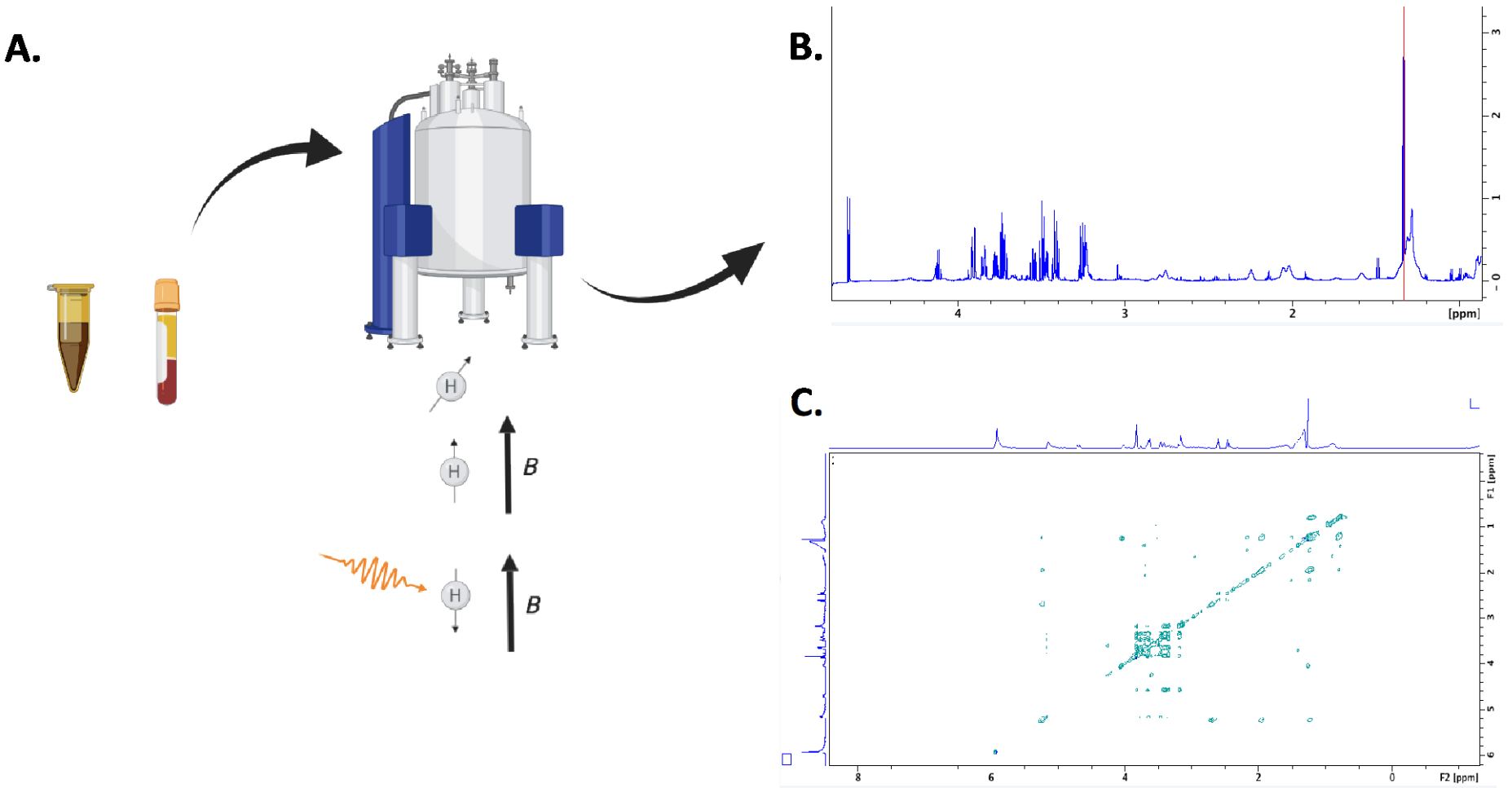

Fig. S3. ^{1}H NMR processing. A) Procedure for characterization of fecal and plasma metabolic profiles with ^{1}H NMR. **B)** Example of a 1D NMR spectrum. **C)** Example of a 2D NMR spectrum.

Sample processing and NMR spectroscopy

Plasma was thawed at room temperature, vortexed, and centrifuged briefly at max speed; 150 µL plasma was added to the 60 µL 75 mM phosphate buffer (5:1 disodium phosphate [Na_2HPO_4] and monosodium phosphate [NaH_2PO_4] in 100% D_2O, pH = 7.4). The solution was transferred to 5 mm Borosilicate NMR tubes (Sigma) with a glass pipette dropper. In addition, 60–70 mg of fecal sample was thawed at room temperature and diluted with 600 µL buffer with 5 µg/mL 3-(trimethylsilyl)propionic-*2,2,3,3-d*$_4$ acid (TSP) as a standard. Samples were homogenized in a Tissue Lyser at 40 Hz for four minutes and centrifuged at 16100 *g* for 15 minutes. The supernatant was transferred to a new Eppendorf tube, and the centrifuging process was repeated. All processing occurred at 4 °C. Samples were then stored at –80 °C. On the day of running the fecal extracts in the NMR, the samples were thawed at room temperature and briefly vortexed and centrifuged at max speed; 400 µL sample was combined with 150 µL buffer and transferred to an NMR tube.

Spectra were processed as described previously (Jurynczyk et al., 2017). ^{1}H spectra were acquired using a one-dimensional Nuclear Overhauser Effect Spectroscopy (NOESY) pre-saturation to reduce the water signal with a 2 s pre-saturation. For plasma, a spin-echo Carr-Purcell-Meiboom-Gill (CPMG) sequence was to suppress signals resulting from large molecular weight compounds such as proteins.

All spectra were preprocessed in Topspin 4.0.5 (Bruker) Processed spectra were exported to ACD/Labs Spectrus Processor Academic Edition 12.01 (Advanced Chemistry Development, Inc., Toronto, Canada) and divided into 0.02 ppm segments called “bins,” and the integral of each bin was calculated. For each sample, the integrals were normalized

to the sum of all the integrals for the sample's spectra, and noise regions were manually removed. The bins, labeled with classes (treatment groups) and animal IDs were then exported to R for analysis. Metabolites were identified by referring to literature values and the human metabolome database (Wishart et al., 2009). Bins with ambiguous signals were either assigned to multiple metabolites or designated with "???" when no candidate metabolites could be assigned. TOCSY spectra were used to separate overlapping signals. In addition, the integrals for regions corresponding to fecal acetate (1.90–1.94 ppm), butyrate (1.54–1.58 ppm), and propionate (1.06–1.08 ppm) were calculated and compared between groups in young and adult offspring using mixed-effects analysis.

Statistical analysis

Univariate analysis

Statistical tests were run in R 3.6.1 (R Foundation for Statistical Computing), and graphs were built in Prism 8.0.0 (Graphpad) and R. Normality of residuals was assessed by examining of linearity of quantile-quantile (Q-Q) plots, and homogeneity of variances was tested with Spearman's rank correlation test for heteroscedasticity (alpha threshold for significance set at $p < 0.01$). Grubbs' test for outliers was applied before statistical analysis, and statistically significant outliers were excluded. Non-Gaussian or heteroscedastic data were square root- or log-transformed. Data transformations are indicated in graph axis labels.

Our data has a nested design with maternal diet, LPS, and sex as fixed effects and litter as a random effect. Therefore, linear mixed-effects analysis applied using the "lmer" function (*lmerTest* package) with restricted maximum likelihood to determine the effects

of maternal diet, LPS, and sex for each outcome (Y) using the following formula: "Y ~ Independent Variable1*Independent Variable2*Sex + (1|Litter)." The "anova" function was then used to calculate *p* values using Satterthwaite's method for approximating degrees of freedom for the *F* tests (Kuznetsova et al., 2017). Because each animal within each litter received the same maternal diet, the denominator degrees of freedom for effects of maternal diet are significantly smaller than what would be expected in a traditional analysis of variance (ANOVA) in which each data point from each animal is independent.

We were not interested in directly comparing outcomes between male and female offspring. Therefore, in the cases of significant two- or three-way interactions with sex, male and female data were analyzed separately to interpret sex-specific effects of maternal diet and LPS treatment. Otherwise, if no significant interactions of any variable with sex were found, male and female data were consolidated using the formula "Y ~ Independent Variable1*Independent Variable2 + (1|Litter)" to simplify the model and minimize the number of comparisons. Tukey post hoc tests were used when a significant interaction was detected.

F0 diet (H_2O vs. B-GOS), acute treatment (saline vs. LPS), and sex were fixed effects, and litter was a random effect in mixed-effects analysis of behavior and brain gene expression. Prenatal diet (H_2O vs. B-GOS), postnatal diet (H_2O vs. B-GOS), and sex were fixed effects, and litter was a random effect in mixed-effects analysis of behavior and fecal SCFAs in cross-fostered offspring (Cohort B). Finally, F0 diet (H_2O vs. B-GOS), P9 treatment, (saline vs. LPS), and sex were fixed effects in analysis of outcomes in adult offspring (Cohort C).

Data were plotted in boxplots, where boxes comprise the median and interquartile range, and the error bars denote minimum and maximum values. Adjusted degrees of freedom (including numerator and denominator degrees of freedom), *F* and *t* scores, and *p* values were reported. Given the large number of statistical tests and results, only significant ($p < 0.05$) outcomes and trends ($p < 0.08$) were included in the results section.

Finally, unpaired Welch's t-tests were used for pairwise comparisons, and one-way ANOVAs were used for the comparison of three or more groups when litter was not a factor.

Multivariate analysis (for ^{1}H NMR data)

Analysis of metabolic data was performed as previously described (Jurynczyk et al., 2017). Sum-normalized bucket integrals were imported into R software for Pareto-scaling and orthogonal partial-least squares discriminant analysis (OPLS-DA) using the *ropls* package in R (Thévenot et al., 2015). OPLS-DA is a supervised multivariate analysis method that attempts to generate models to explain variation in a dataset. These models can also give information about which metabolites drive differences between groups.

OPLS-DA models were built by multiple iterations of randomly separating data into training and test sets. Training sets were used to build the models using a seven-fold internal cross-validation procedure. The models were then tested by forced classification of samples in the independent test sets in a ten-fold external cross-validation scheme to determine the accuracy, sensitivity, and specificity of the models. In addition, goodness of fit was assessed with R^2X (a measure of how well the models explained variation in the observations) and R^2Y (how good the separation between groups was). Finally, predictive

ability of the models was assessed with Q^2. A total of 1000 models were built for each pairwise group comparison. In addition, the same model-building and external cross-validation procedure was used to build models from data randomly sorted into classes – the average accuracy of these models is always expected to be 50%. If the true models performed significantly better than random chance in categorizing the samples in the test sets ($p < 0.001$, Kolmogorov-Smirnov test), a representative OPLS-DA model was generated, and metabolites driving the group separations were identified using the average variable importance in projection (VIP) scores. In the graph of VIP scores vs. ranking, the inflection point was used to determine which variables were important in building the models. The ^{1}H NMR process workflow is summarized in **Fig. S4**.

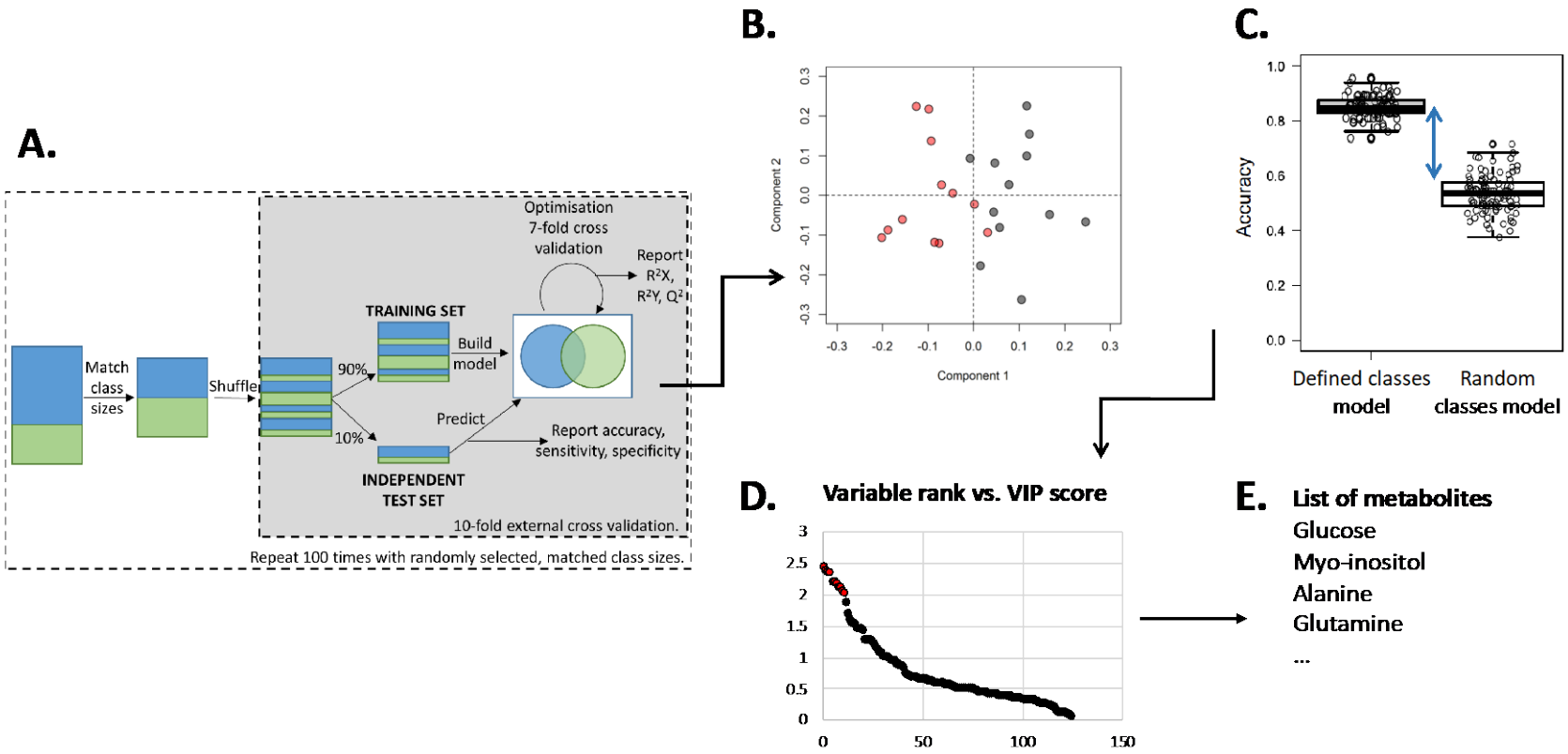

Fig. S4. OPLS-DA analysis procedure. **A)** OPLS-DA procedure for building models and testing models with 10-fold external cross-validation. **B)** Example of one OPLS-DA model. **C)** Accuracy of models in classifying real data is compared to accuracy in classifying randomly labeled data. **D)** Variables corresponding to top VIP scores are selected for metabolite identification. **E)** The last step is to identify the metabolites that correspond to the spectral bins with high VIP scores. These are the metabolites that are significantly different between groups.

16S sequencing

16S rRNA amplicon library preparation and sequencing

Fecal DNA was sent to the Oxford Genomics Centre (OGC) for library preparation and sequencing. Library preparation and sequencing were performed at the Oxford Genomics Center (OGC). DNA samples were normalized to 12.4 ng in 2.5 µl volume. Primers targeting the V3-V4 region of the 16S rRNA gene were used for amplification (Klindworth et al., 2013). The full-length primer sequences, using standard International Union of Pure and Applied Chemistry (IUPAC) nucleotide nomenclature, were: 16S Amplicon PCR Forward Primer = 5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGC CTACGGGNGGCWGCAG, 16S Amplicon PCR Reverse Primer = 5' GTCTCGTGGGCTCG GAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC. Following a post-amplicon PCR clean-up, a second indexing PCR was applied using NexteraXT indices. Following the second PCR clean-up, libraries were then pooled and diluted into a 10nM final multiplex library ready for sequencing on Illumina MiSeq (V3 chemistry).

Data processing

Quality of raw sequencing data was assessed using FastQC (v0.11.07). Following quality assessment, downstream processing was performed using *dada2* (Callahan et al., 2016) which was implemented in pipeline_dada2.py from the NGSKit repository (https://github.com/nickilott/NGSKit). Given the drop in quality at ~200bp, we decided to proceed with the analysis using the first read in the pair (V3 region). The filterAndTrim function in *dada2* was used to truncate the forward and reverse reads. Retained primer sequences were removed from forward reads (17bp) which resulted in final read lengths of 183bp. Error learning, de-replication,

and sample inference were performed using *dada2* with default parameters. Taxonomy was assigned to amplicon sequence variants (ASVs) using the assignTaxonomy function in *dada2* and the NCBI RefSeq training data (https://zenodo.org/record/2541239/files/RefSeq-RDP16S_v2_May2018.fa.gz). ASV count tables were used for downstream analysis.

Alpha and beta diversity

The *phyloseq* package in R (McMurdie and Holmes, 2013) was used for all statistical analysis of microbiome diversity. To investigate global changes in microbiome composition, the Chao1 measure of alpha diversity (number and distribution of ASVs in each sample) was compared between H_2O control and B-GOS animals in the F0 and F1 cohorts using absolute ASV counts. The Bray-Curtis measure of beta diversity (how different each sample is from every other sample based on relative ASV abundances) was used to explore global differences in microbiome composition between groups and plot data in principal coordinate analysis (PCoA) plots. Permutational multivariate analysis of variance (PERMANOVA) was used to determine whether microbiome composition was significantly different between groups based on Bray-Curtis indices.

Identification of differentially abundant taxa

To identify differentially abundant taxa between groups, differential gene expression analysis based on the negative binomial distribution (DESeq) was performed using the *deseq2* package in R. DESeq uses the Wald test to calculate *p* values, which are adjusted for multiple comparisons using the procedure of Benjamini and Hochberg (Love et al., 2014). Because there are strong cage and litter effects in microbiome data (Ubeda et al., 2012), linear mixed-effects analysis with litter as a random effect was performed on log10-transformed relative abundances to determine which ASVs were robust to litter effects. Differentially abundant ASVs that were

detected by both DESeq and mixed-effects analysis ($p < 0.05$) were considered to be significantly different in offspring.

Supplementary results

Effects of perinatal B-GOS intake on PFC gene expression in young offspring and maternal care behavior

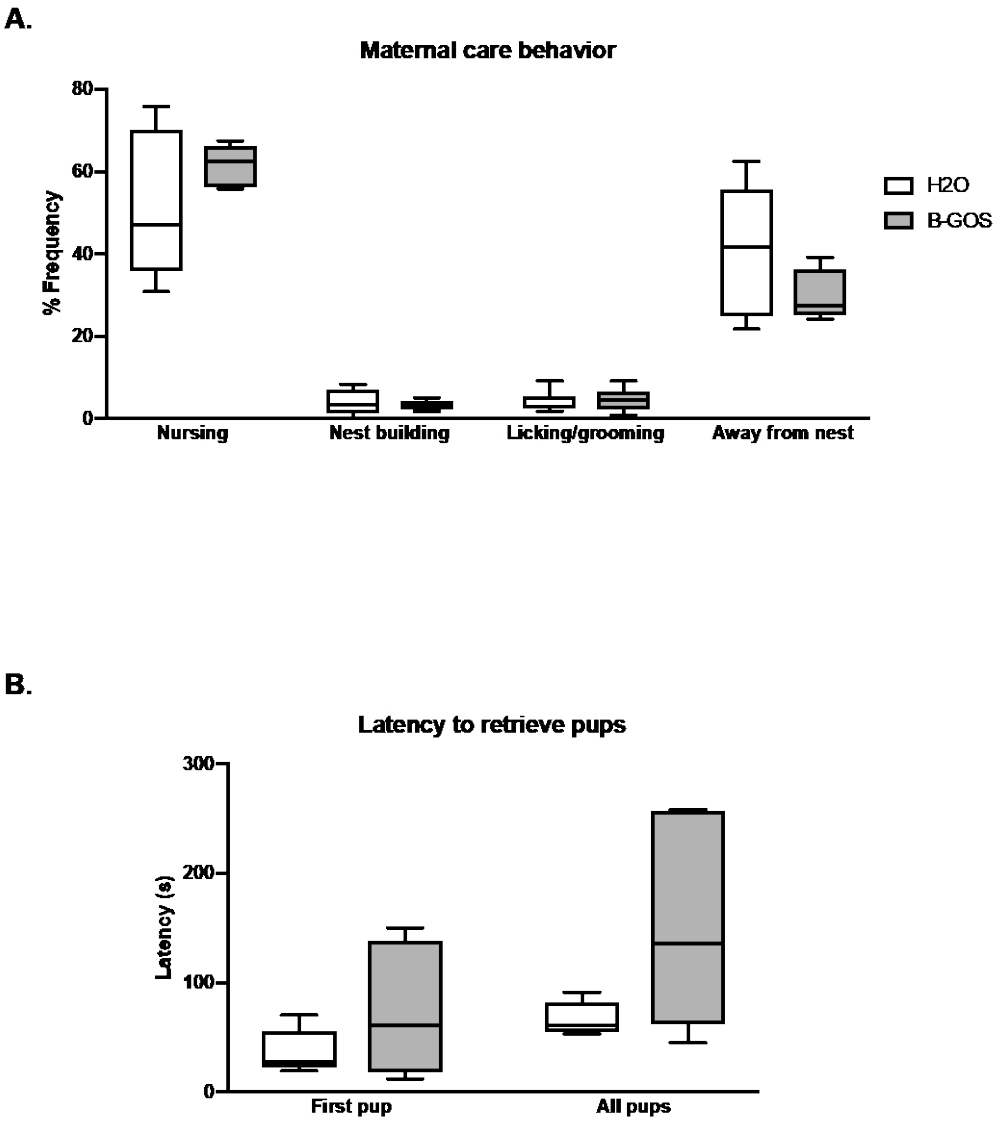

Fig. S5. Prebiotic intake has no effect on maternal care. A) There were no effects of B-GOS on frequency of maternal nursing, nest building, licking/grooming, or being away from the nest in the home cage. **B)** In the pup retrieval test, there were also no differences between B-GOS and plain water on latency to retrieve the first pup or all pups.

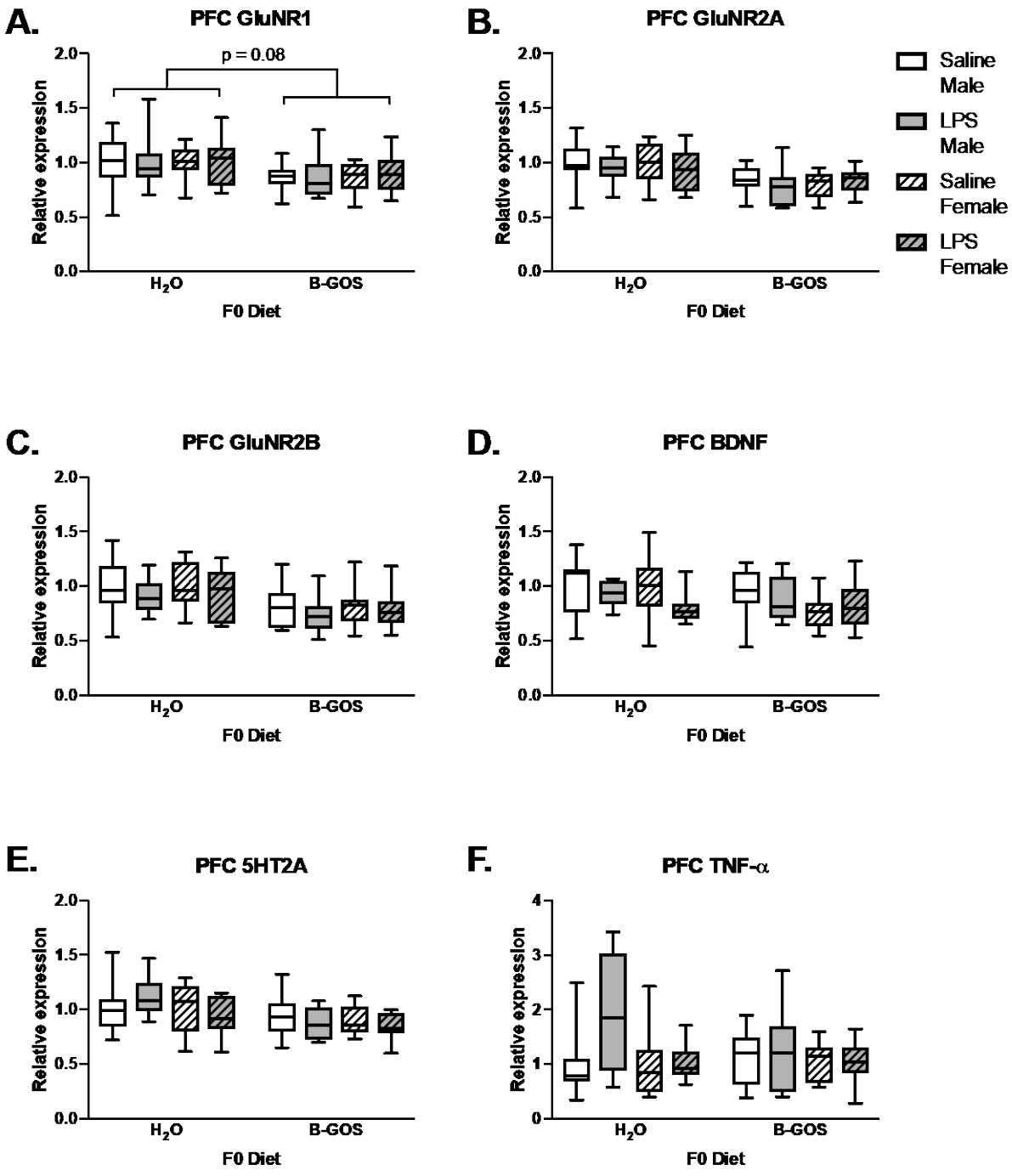

Fig. S6. Maternal B-GOS and acute LPS have no effects on PFC gene expression. There were no significant effects of perinatal prebiotic intake or LPS on **A)** GluNR1, **B)** GluNR2A, **C)** GluNR2B, **D)** BDNF, **E)** 5-HT$_{2A}$, or **F)** TNF-α expression in the PFC in young offspring (Cohort A). Mixed-effects analysis, N = 10–11 males/females; N = 20–22 total.

Additional 16S sequencing results

16S sequencing quality

DNA was extracted from maternal and offspring fecal samples and diluted for sequencing. Average Phred scores (confidence in the assigned base pair [bp]) for each bp position were used to asses sequence quality in every sample. All samples had acceptable quality scores (Phred score > 25) up until read lengths of ~200 bp.

16S sequencing ASVs and taxonomic assignment

Samples had between 455 and 135,372 reads (M = 86,335, SD = 29,895). Two samples with very low read counts (< 600) were removed before analysis. The remaining samples had at least 7259 reads (M = 89,103, SD = 25,957).

Between ~100–350 ASVs were called per sample in the *dada2* pipeline, with the exception of the two samples with very low reads. High-level taxonomic assignment was performed to confirm that the samples contained the expected phyla. As would be anticipated in the mouse gut, Bacteroidetes and Firmicutes made up the vast majority of the microbes present (**Fig. S7**). In addition, most ASVs were able to be assigned at the phylum-level, and a smaller percentage could be assigned at the lower taxonomic levels, with very few assigned at the species level. This is expected for the resolution of 16S sequencing. ASVs that could not be assigned to a phylum were removed before analysis.

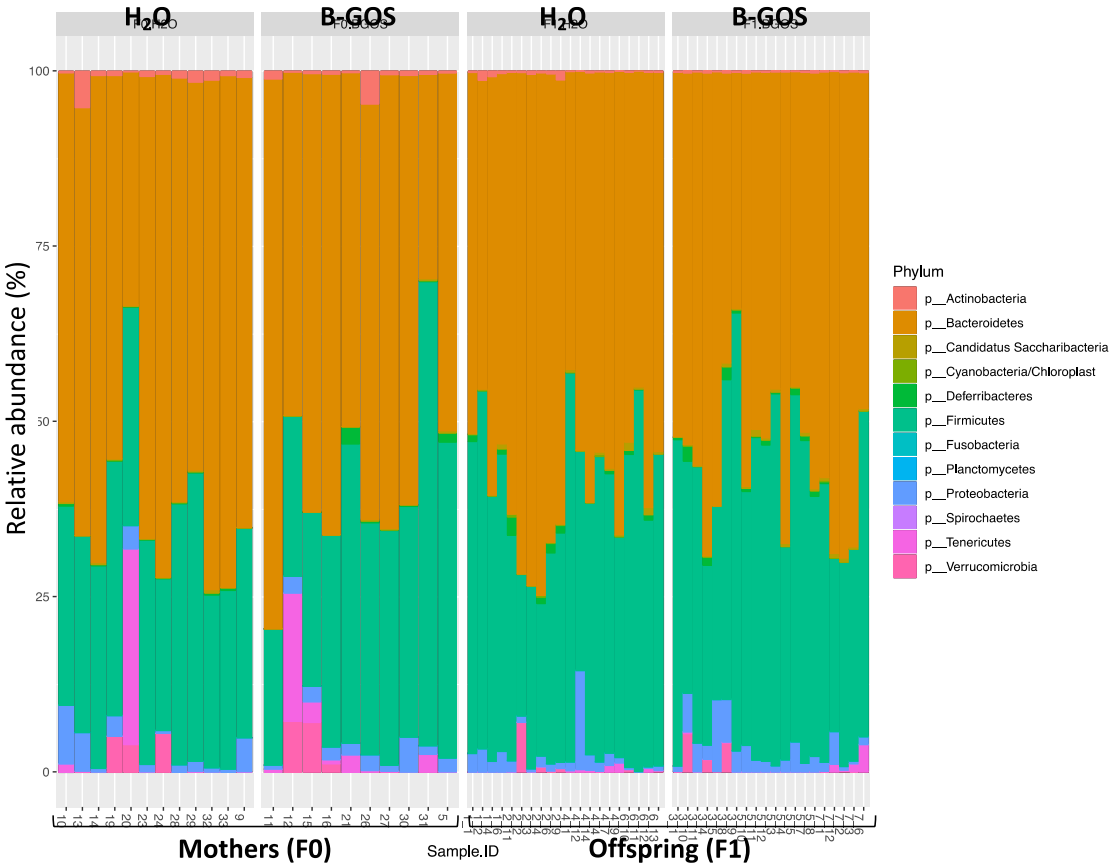

Fig. S7. Taxonomic assignments for dams and offspring (relative abundances at the phylum level). Each multi-colored bar represents fecal microbiome composition of one sample.

Alpha diversity

A Welch's t-test was used to compare alpha diversity (number and distribution of ASVs within samples) between B-GOS-fed and control dams and their offspring. There was no effect of B-GOS on alpha diversity in F0, $t(18.60) = -1.07$, $p = 0.30$ (**Fig. S8A**). Similarly, there was no effect of maternal diet on alpha diversity in offspring, $t(24.10) = -1.50$, $p = 0.15$ (**Fig. S8B**).

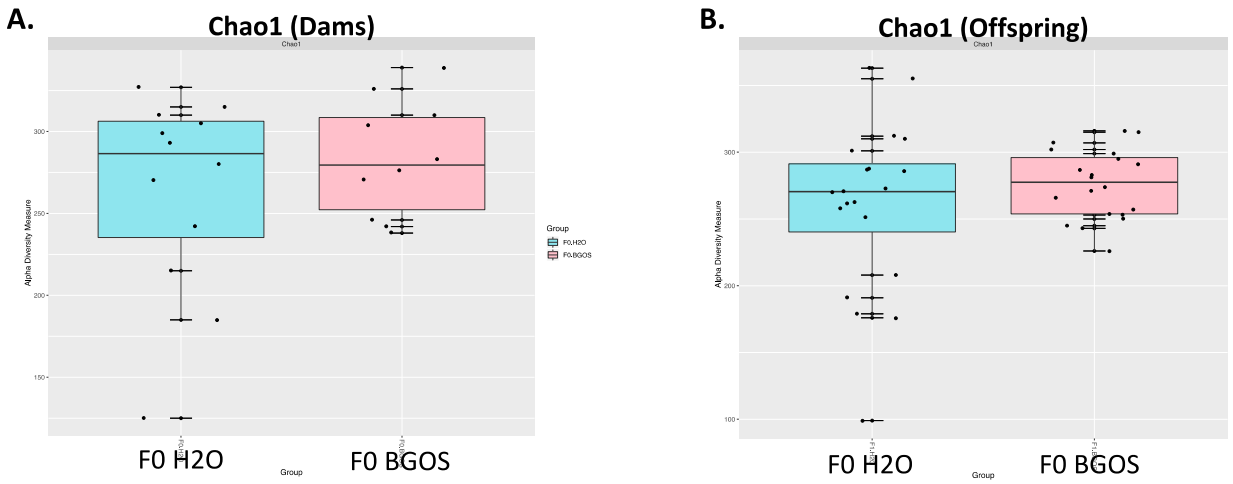

Fig. S8. B-GOS has no effect on alpha diversity in dams or offspring. Absolute numbers of ASVs per sample were used to calculate the Chao1 measure of alpha diversity in **A)** F0 and **B)** F1 animals. There were no significant differences in either cohort. Unpaired Welch's t-test, N = 10–11 dams/group; N = 9–10 male/female offspring, N = 19–20 total/group.

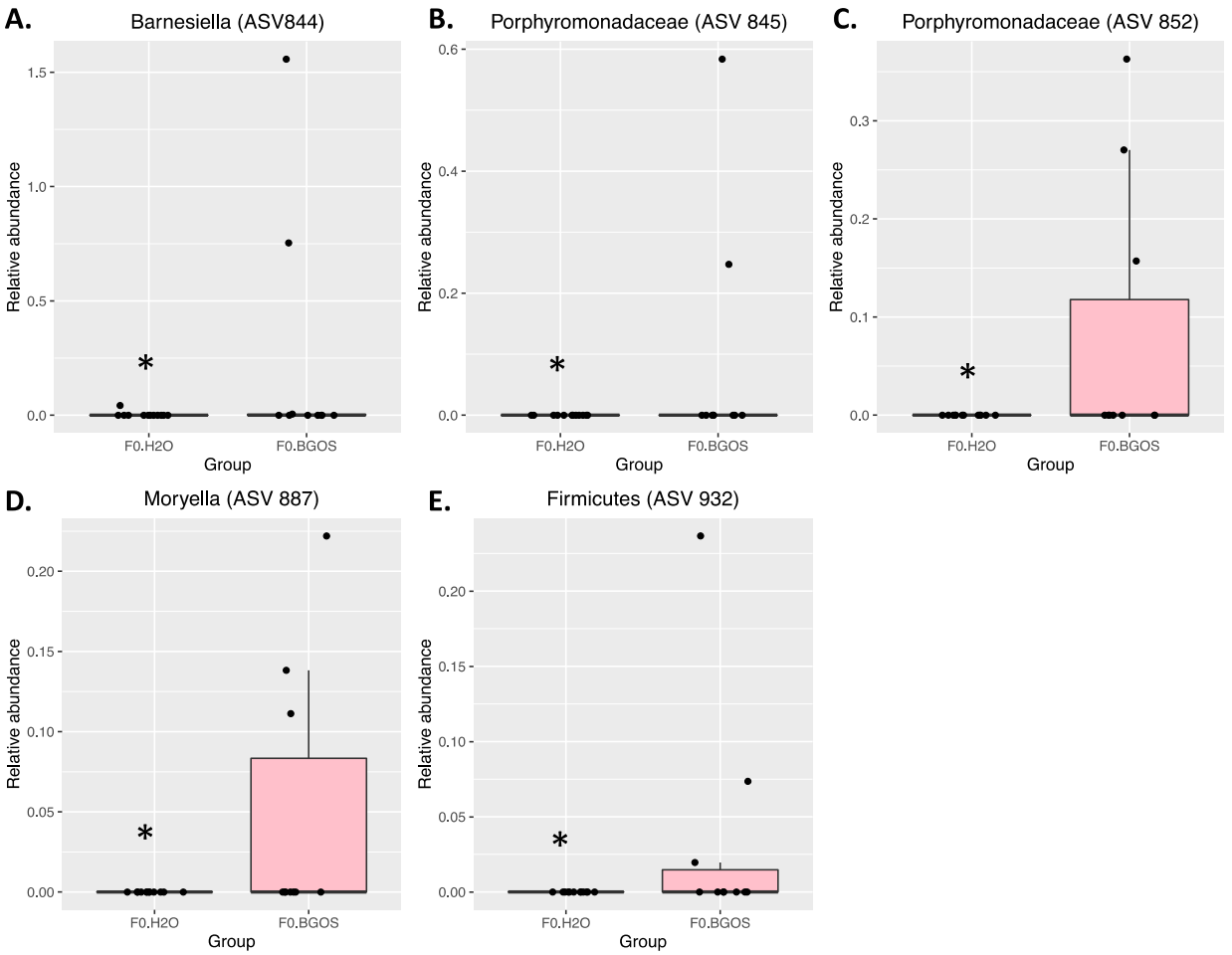

Fig. S9. Relative abundances of taxa detected by DESeq in dams. Each boxplot represents one ASV, and the lowest taxonomic level assigned by *dada2* is indicated in the boxplot titles. DESeq, *$p < 0.05$, N = 10–11/group.

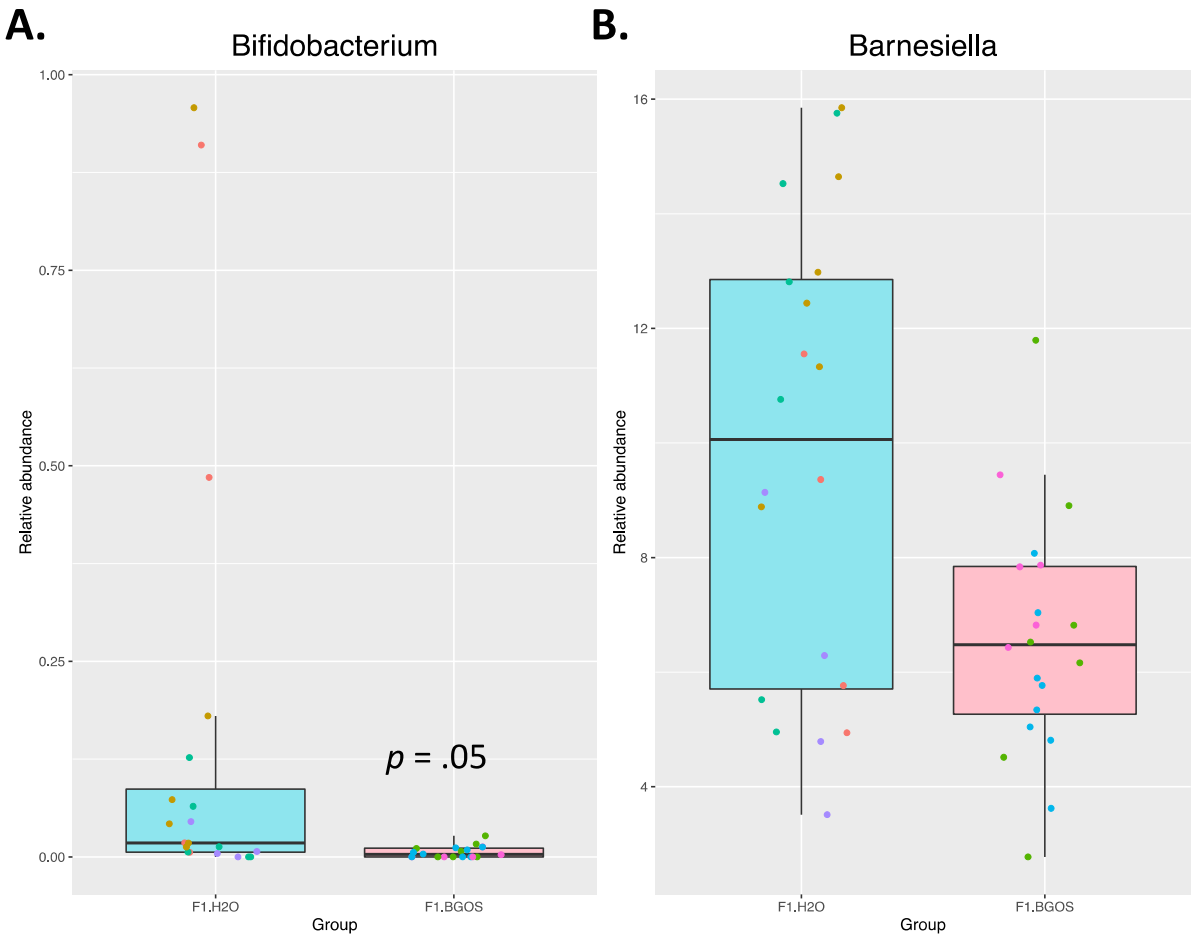

Fig. S10. Relative abundances of total *Bifidobacterium* and *Barnesiella* in offspring. Because ASVs corresponding to unknown species in the *Bifidobacterium* and *Barnesiella* genera were differentially abundant between H_2O and B-GOS pups, ASVs in these genera were pooled and compared between groups. There were no significant differences between groups, although the effect of F0 diet on *Bifidobacterium* just missed the $p = 0.05$ threshold. Mixed-effects analysis, N = 9–10 male/female offspring, N = 19–20 total/group.

Fecal ^{1}H NMR metabolic profiling

Fecal ^{1}H NMR analysis: dams (Cohort B)

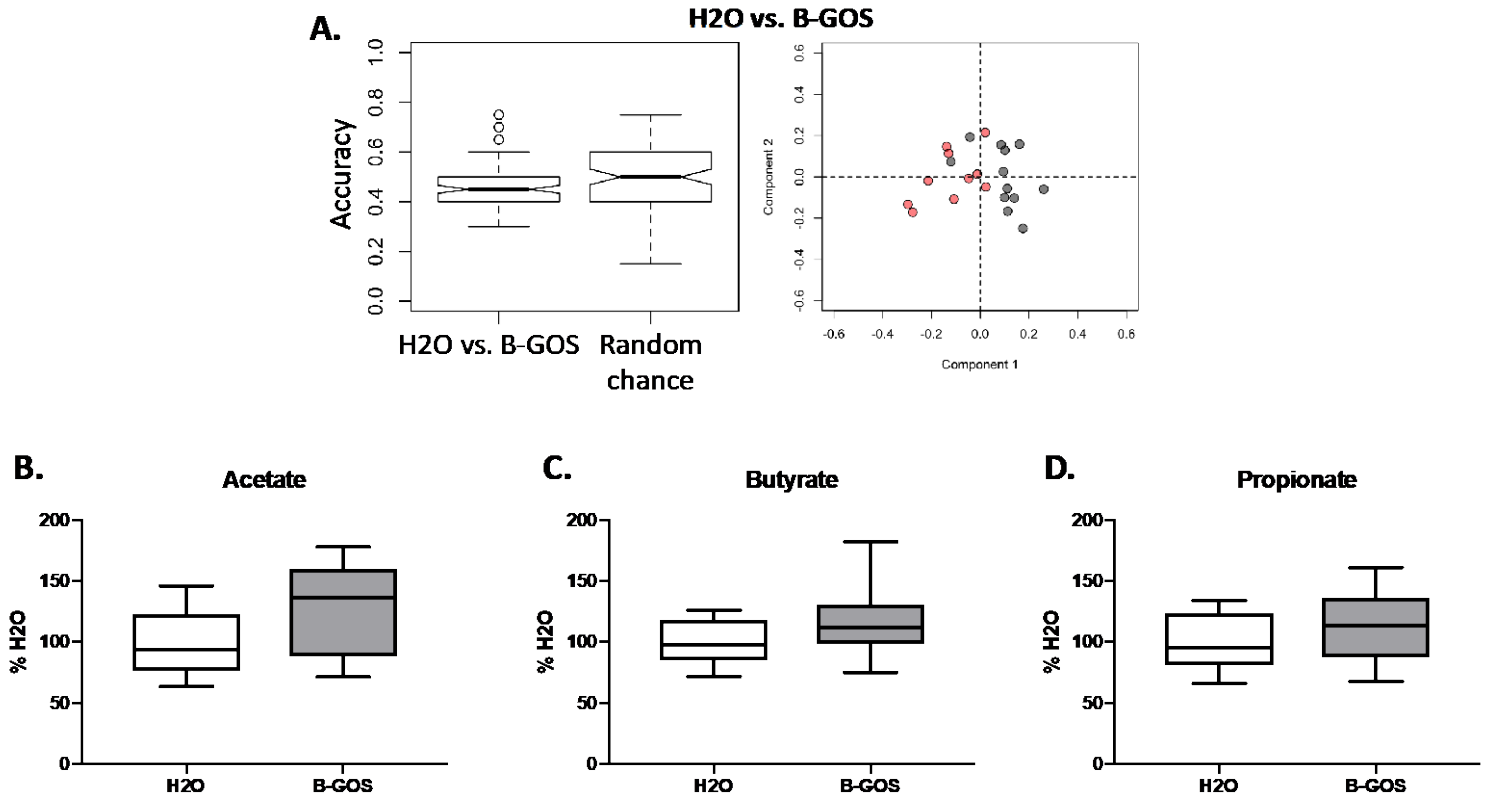

Fig. S11. B-GOS has no effect on fecal metabolite profiles in suckling dams. Fecal samples were collected from dams that received B-GOS or normal water during pregnancy and lactation. Samples were collected when pups were at weaning age. **A)** B-GOS-fed dams could not be distinguished from controls based on OPLS-DA. There were also no differences in fecal levels of the SCFAs **B)** acetate, **C)** butyrate, and **D)** propionate. N = 10–12/group.

Additional fecal ^{1}H NMR analysis: young cross-fostered offspring (Cohort B)

Table S2. Summary of Cohort B fecal ^{1}H NMR OPLS-DA results.

Group A	Group B	Average accuracy of A vs. B models ± SD*	Significantly better performance than random chance?	Average R^2X ± SD*	Average R^2Y ± SD*	Average Q^2 ± SD*
H2O/H2O (n = 13)	B-GOS/B-GOS (n = 13)	78.1 ± 6%	Yes	0.65 ± 0.05	0.88 ± 0.03	0.51 ± 0.05
H2O/H2O (n = 13)	H2O/B-GOS (n = 12)	79.3 ± 6%	Yes	0.69 ± 0.04	0.89 ± 0.03	0.42 ± 0.07
H2O/H2O (n = 13)	B-GOS/H2O (n = 14)	45.1 ± 9%	No	0.71 ± 0.03	0.78 ± 0.03	-0.66 ± 0.22
H2O/B-GOS (n = 12)	B-GOS/H2O (n = 14)	80.1 ± 6%	Yes	0.71 ± 0.03	0.89 ± 0.02	0.49 ± 0.09
H2O (F0) (n = 12)	B-GOS (F0) (n = 10)	47.1 ± 10%	No	0.58 ± 0.03	0.83 ± 0.03	-0.42 ± 0.30

*SD = standard deviation

Effects of perinatal B-GOS on plasma metabolite profiles in young offspring

Plasma samples were collected for ^{1}H NMR metabolic profiling to generate hypotheses about altered host compounds that may be related to the brain and behavioral changes. We were also interested in additional evidence of biological differences resulting from maternal prebiotic intake. Results for young cross-fostered pups (Cohort B) are shown in **Fig. S12**. Differences in metabolic profiles between B-GOS/B-GOS and H_2O/H_2O control pups were driven by glucose signals. Pups born to and suckled by B-GOS-fed dams had reduced blood glucose levels compared to controls. The largest difference in blood metabolites, however, was between pups that were cross-fostered (H_2O/B-GOS and B-GOS/H_2O) and pups that remained with the same dam during both pregnancy and lactation (H_2O/H_2O and B-GOS/B-GOS). This indicates that cross-fostering alone may have effects on blood metabolite levels.

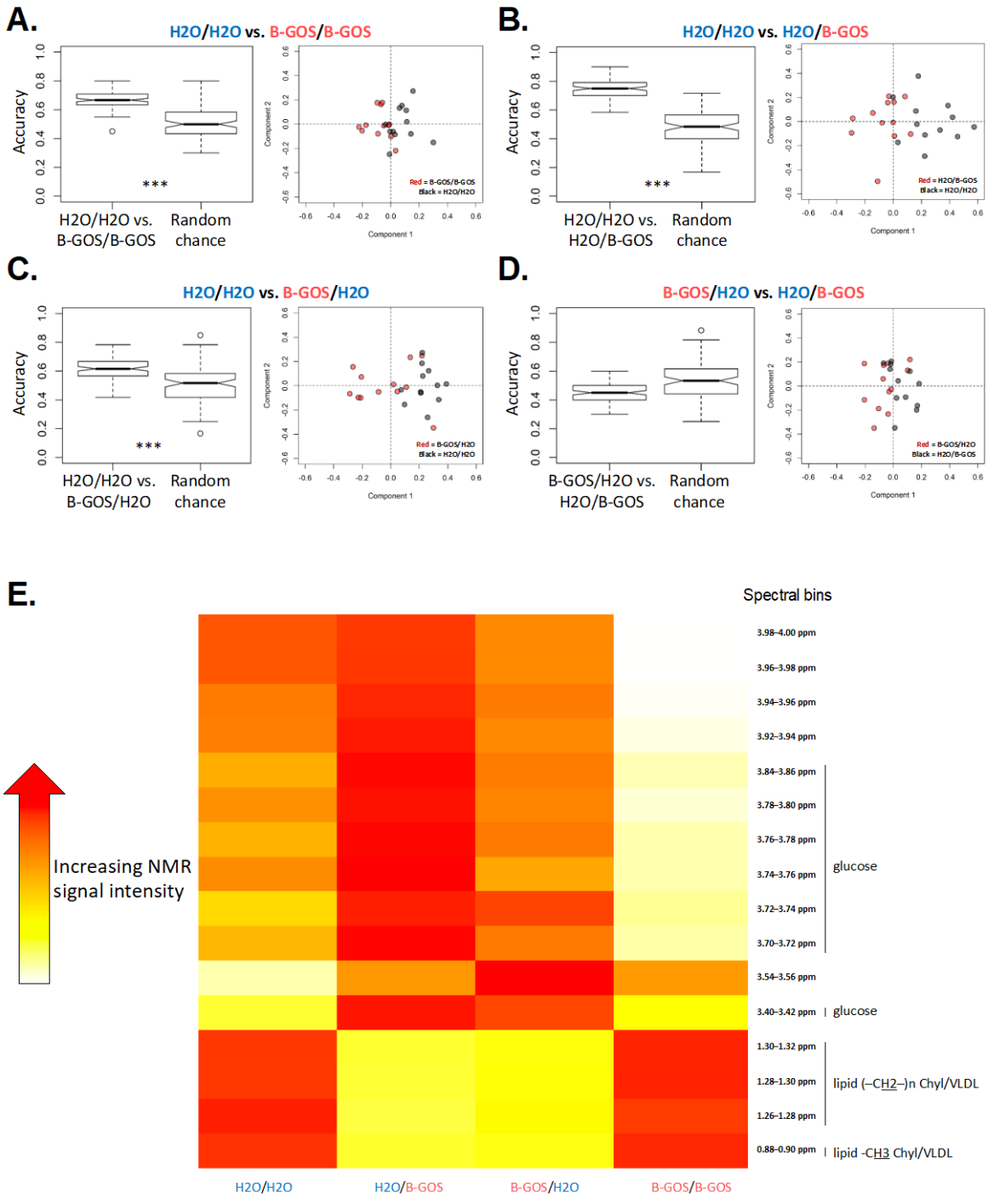

Fig. S12. Effects of maternal B-GOS on plasma metabolite profiles in young offspring. According to OPLS-DA analysis, H_2O/H_2O control offspring had significantly different plasma metabolite profiles than **A)** B-GOS/B-GOS, **B)** H_2O/B-GOS, **C)** and B-GOS/H_2O pups. **D)** B-GOS/H_2O offspring did not differ from H_2O/B-GOS. **E)** A heat map comparing spectral bins and corresponding metabolites between adult offspring groups is shown. Bins were selected based on VIP score. Kolmogorov-Smirnov, ***$p < 0.001$. N = 6–8 males/females; N = 12–14 total/group.

Effects of perinatal B-GOS intake on hippocampal gene expression in adult offspring

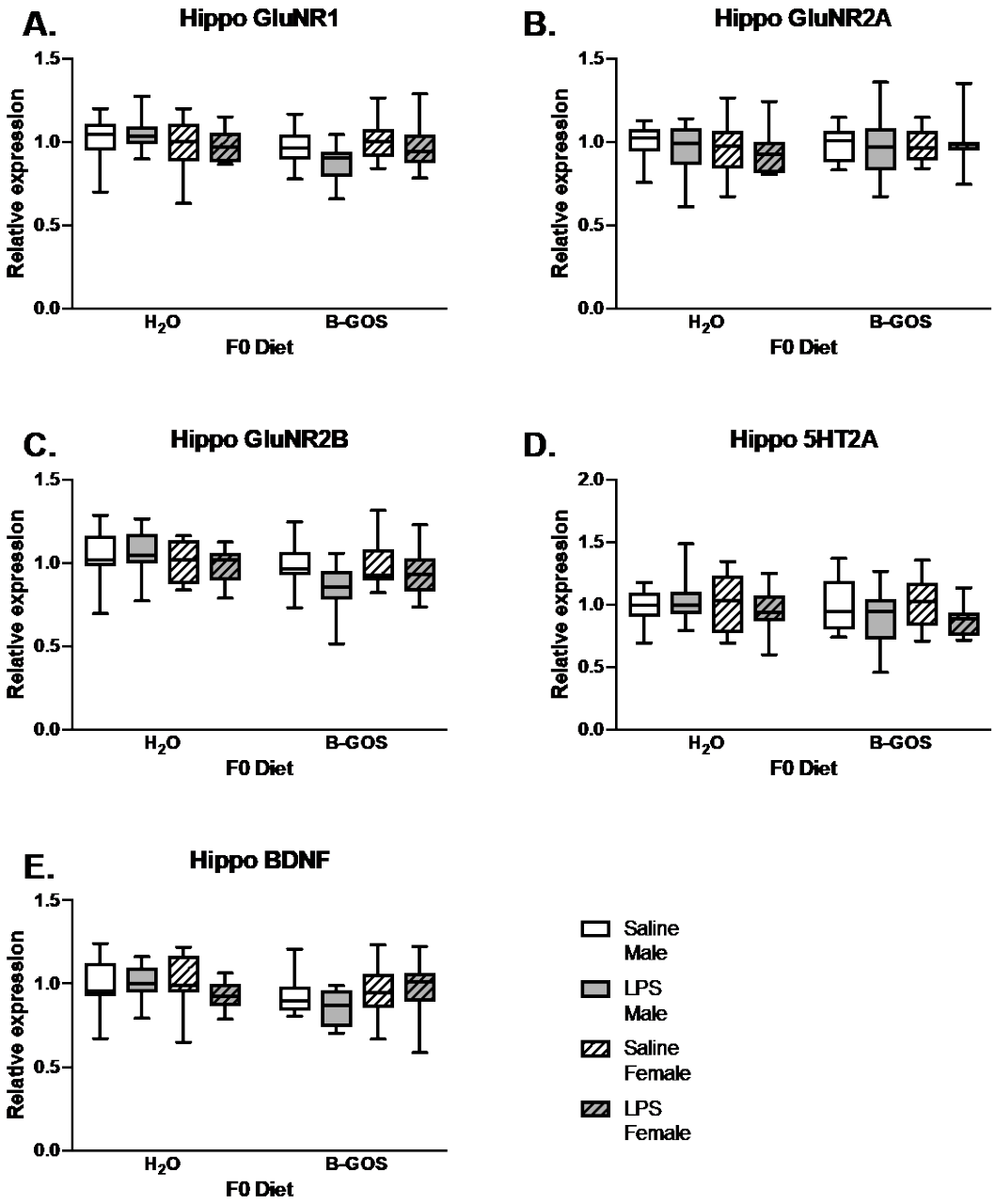

Fig. S13. Effects of maternal B-GOS intake on hippocampal gene expression in adult offspring. There were no significant effects of perinatal prebiotic feeding or P9 LPS treatment on hippocampal gene expression in adult offspring. Mixed-effects analysis, N = 9–10 males/females; N = 19–20 total/group.

Fecal ^{1}H NMR analysis: adult offspring (Cohort C)

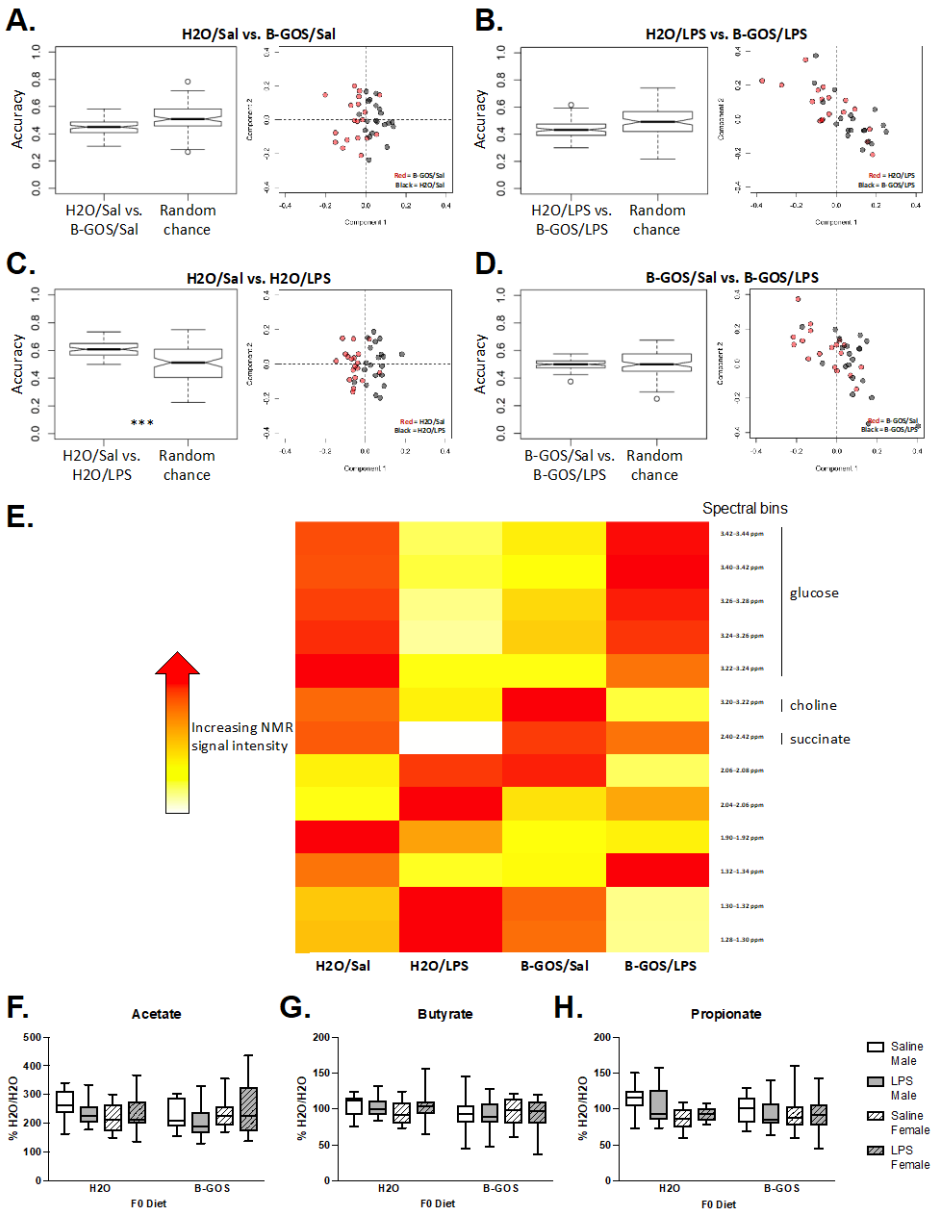

Fig. S14. P9 LPS alters fecal metabolite composition, while maternal prebiotic intake has no effect in adult offspring. A–B) Maternal B-GOS had no effect on fecal metabolites in adult offspring. **C)** P9 LPS significantly altered metabolic profiles in H_2O offspring **D)** but not B-GOS offspring. Kolmogorov-Smirnov, ***$p < 0.001$. **E)** Heat map comparing spectral bins and corresponding metabolites between adult offspring groups. There were no significant effects of maternal B-GOS supplementation on fecal levels of **F)** butyrate, **G)** propionate, or **H)** acetate in adult offspring. Mixed-effects analysis, N = 9–10 males/females; N = 19–20 total/group.

Table S3. Summary of OPLS-DA results for adult offspring fecal ^{1}H NMR.

Group A	Group B	Average accuracy of A vs. B models ± SD*	Significantly better performance than random chance?	Average R^2X± SD*	Average R^2Y ± SD*	Average Q^2 ± SD*
H2O/Saline (n = 20)	B-GOS/Saline (n = 19)	45.1 ± 6%	No	0.48 ± 0.01	0.79 ± 0.01	-0.48 ± 0.13
H2O/LPS (n = 19)	B-GOS/LPS (n = 20)	43.5 ± 6%	No	0.70 ± 0.02	0.69 ± 0.03	-0.76 ± 0.21
H2O/Saline (n = 20)	H2O/LPS (n = 19)	61.1 ± 5%	Yes	0.54 ± 0.02	0.74 ± 0.01	-0.24 ± 0.11
B-GOS/Saline (n = 19)	B-GOS/LPS (n = 20)	50.6 ± 4%	No	0.61 ± 0.01	0.62 ± 0.03	-0.44 ± 0.15

**SD* = standard deviation

Effects of perinatal B-GOS and early-life LPS on body weights and plasma metabolite profiles in adult offspring

Neonatal offspring body weights were recorded at P9 (before saline/LPS injections), P10, and P14. At baseline before administration of LPS (age P9), there were no differences between groups (**Figure S15A**). At P10, however, there was a significant main effect of P9 treatment, $F(1, 125.02) = 70.60$, $p < 0.0001$, showing that P10 LPS-treated pups weighed significantly less than saline-treated animals (**Figure S15B**). Similarly, at P14, there was a significant main effect of P9 treatment on body weight, $F(1, 123.06) = 131.45$, $p < 0.0001$, indicating that LPS-treated animals weighed significantly less than saline-treated mice five days after LPS exposure (**Figure S15C**). There were no effects, however, of perinatal B-GOS or early-life LPS on body weights in adult offspring at ~P56 (data not shown).

In adult offspring (**Fig. S16**), maternal B-GOS altered levels of lipoprotein-associated lipids in adult offspring. Lipoprotein-associated lipid signals in the 0.84–0.90 ppm and 1.26–1.36 ppm regions were roughly divided into lipid signals associated with chylomicron (Chyl)/very low-density lipoprotein (VLDL) and low-density/high-density lipoproteins (HDL/LDL) (Liu, Tang, Nicholson, & Lindon, 2002; Soininen et al., 2009). Signal intensity in peaks corresponding to Chyl/VLDL-associated lipids were lower in offspring of B-GOS-fed dams (B-GOS/Sal and B-GOS/LPS) than H_2O offspring (H_2O/Sal and H_2O/LPS). In contrast, signals corresponding to HDL/LDL were elevated in B-GOS offspring compared to H_2O offspring. Whether these metabolic differences are involved in the brain and behavioral changes should be investigated in future studies. In addition, maternal prebiotic feeding should be investigated as an intervention in animal models of metabolic disorders.

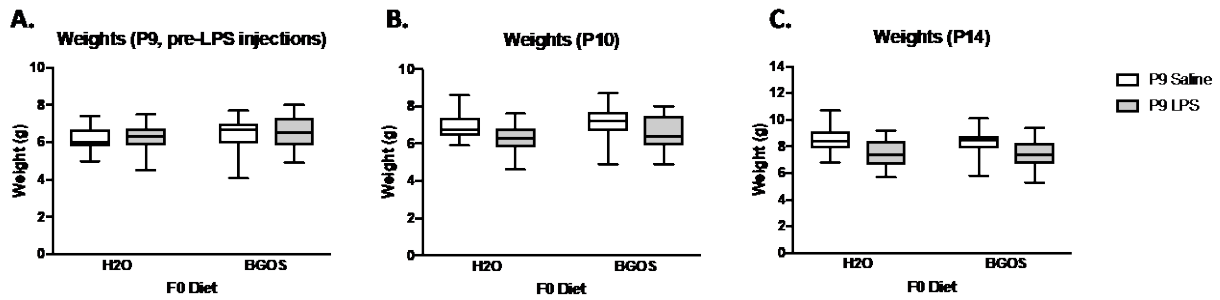

Fig. S15. P9 LPS results in significant weight loss one and five days post injection. A) At baseline, there were no differences between groups. At both B) P10 and C) P14, significant main effects of P9 treatment showed that pups treated with subcutaneous LPS weighed significantly less than those treated with saline ($p < .0001$ for both P10 and P14). Two-way ANOVA. Mixed-effects analysis, N=31–38 total/group.

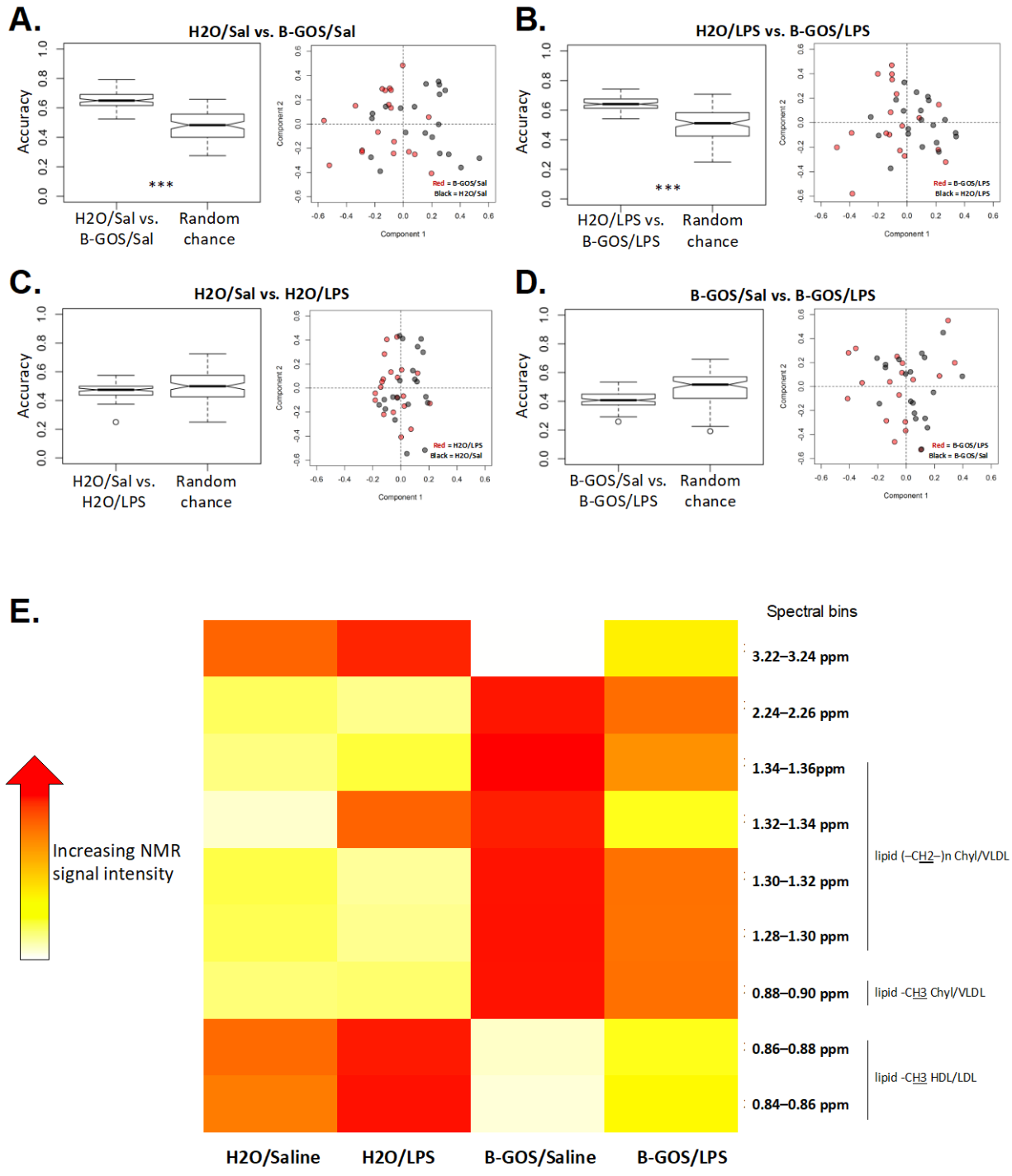

Fig. S16. Effects of maternal B-GOS on plasma metabolite profiles in adult offspring. Maternal B-GOS altered plasma metabolite profiles in all offspring – both those treated with **A)** saline and **B)** LPS at P9. **C–D)** P9 LPS had no effect on plasma metabolites. Kolmogorov-Smirnov, *** $p < 0.001$. N = 8–10 males/females; N = 18–20 total/group.

References

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. Nat. Methods 13, 581–583. https://doi.org/10.1038/nmeth.3869

Carlén, M., 2017. What constitutes the prefrontal cortex? Science (80-.). 358, 478–482. https://doi.org/10.1126/science.aan8868

Couch, Y., Xie, Q., Lundberg, L., Sharp, T., Anthony, D.C., 2015. A Model of Post-Infection Fatigue Is Associated with Increased TNF and 5-HT2A Receptor Expression in Mice. PLoS One 10. https://doi.org/10.1371/journal.pone.0130643

Denman, S.E., McSweeney, C.S., 2006. Development of a real-time PCR assay for monitoring anaerobic fungal and cellulolytic bacterial populations within the rumen. FEMS Microbiol. Ecol. 58, 572–582. https://doi.org/10.1111/j.1574-6941.2006.00190.x

Franks, B., Curley, J.P., Champagne, F.A., 2011. Measuring Variations in Maternal Behavior: Relevance for Studies of Mood and Anxiety, in: Gould, T.D. (Ed.), Mood and Anxiety Related Phenotypes in Mice. Humana Press, pp. 209–224.

Gould, T.D., 2009. Mood and Anxiety Related Phenotypes in Mice: Characterization Using Behavioral Tests.

Ji, J., Hu, S.L., Cui, Z.W., Li, W.F., 2013. Probiotic Bacillus amyloliquefaciens mediate M1 macrophage polarization in mouse bone marrow-derived macrophages. Arch. Microbiol. 195, 349–356. https://doi.org/10.1007/s00203-013-0877-7

Jurynczyk, M., Probert, F., Yeo, T., Tackley, G., Claridge, T.D.W., Cavey, A., Woodhall, M.R., Arora, S., Winkler, T., Schiffer, E., Vincent, A., DeLuca, G., Sibson, N.R., Isabel Leite, M., Waters, P., Anthony, D.C., Palace, J., 2017. Metabolomics reveals distinct, antibody-independent, molecular signatures of MS, AQP4-antibody and MOG-antibody disease. Acta Neuropathol. Commun. 5, 95. https://doi.org/10.1186/s40478-017-0495-8

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., Glöckner, F.O., 2013. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res. 41, 1–11. https://doi.org/10.1093/nar/gks808

Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B., 2017. lmerTest Package: Tests in Linear Mixed Effects Models . J. Stat. Softw. 82. https://doi.org/10.18637/jss.v082.i13

Liu, M., Tang, H., Nicholson, J.K., Lindon, J.C., 2002. Use of 1H NMR-determined diffusion coefficients to characterize lipoprotein fractions in human blood plasma. Magn. Reson. Chem. 40, 83–88. https://doi.org/10.1002/mrc.1121

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 1–21. https://doi.org/10.1186/s13059-014-0550-8

McMurdie, P.J., Holmes, S., 2013. Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS One 8. https://doi.org/10.1371/journal.pone.0061217

Musazzi, L., Cattaneo, A., Tardito, D., Barbon, A., Gennarelli, M., Barlati, S., Racagni, G., Popoli, M., 2009. Early raise of BDNF in hippocampus suggests induction of posttranscriptional

mechanisms by antidepressants. BMC Neurosci. 10, 1–7. https://doi.org/10.1186/1471-2202-10-48

Porsolt, R.D., Pichon, M. Le, Jalfre, M., 1977. Depression: a new animal model sensitive to antidepressant treatments. Nature 266, 731–732.

Rinttilä, T., Kassinen, A., Malinen, E., Krogius, L., Palva, A., 2004. Development of an extensive set of 16S rDNA-targeted primers for quantification of pathogenic and indigenous bacteria in faecal samples by real-time PCR. J. Appl. Microbiol. 97, 1166–1177. https://doi.org/10.1111/j.1365-2672.2004.02409.x

Savignac, H.M., Couch, Y., Stratford, M., Bannerman, D.M., Tzortzis, G., Anthony, D.C., Burnet, P.W.J., 2016. Prebiotic administration normalizes lipopolysaccharide (LPS)-induced anxiety and cortical 5-HT2A receptor and IL1-β levels in male mice. Brain. Behav. Immun. 52, 120–131. https://doi.org/10.1016/j.bbi.2015.10.007

Soininen, P., Kangas, A.J., Würtz, P., Tukiainen, T., Tynkkynen, T., Laatikainen, R., Järvelin, M.R., Kähönen, M., Lehtimäki, T., Viikari, J., Raitakari, O.T., Savolainen, M.J., Ala-Korpela, M., 2009. High-throughput serum NMR metabonomics for cost-effective holistic studies on systemic metabolism. Analyst 134, 1781–1785. https://doi.org/10.1039/b910205a

Thévenot, E.A., Roux, A., Xu, Y., Ezan, E., Junot, C., 2015. Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses. J. Proteome Res. 14, 3322–3335. https://doi.org/10.1021/acs.jproteome.5b00354

Ubeda, C., Lipuma, L., Gobourne, A., Viale, A., Leiner, I., Equinda, M., Khanin, R., Pamer, E.G., 2012. Familial transmission rather than defective innate immunity shapes the distinct intestinal microbiota of TLR-deficient mice. J. Exp. Med. 209, 1445–1456. https://doi.org/10.1084/jem.20120504

Wishart, D.S., Knox, C., Guo, A.C., Eisner, R., Young, N., Gautam, B., Hau, D.D., Psychogios, N., Dong, E., Bouatra, S., Mandal, R., Sinelnikov, I., Xia, J., Jia, L., Cruz, J.A., Lim, E., Sobsey, C.A., Shrivastava, S., Huang, P., Liu, P., Fang, L., Peng, J., Fradette, R., Cheng, D., Tzur, D., Clements, M., Lewis, A., de souza, A., Zuniga, A., Dawe, M., Xiong, Y., Clive, D., Greiner, R., Nazyrova, A., Shaykhutdinov, R., Li, L., Vogel, H.J., Forsythei, I., 2009. HMDB: A knowledgebase for the human metabolome. Nucleic Acids Res. 37, 603–610. https://doi.org/10.1093/nar/gkn810

Yang, M., Silverman, J.L., Crawley, J.N., 2016. Automated Three-Chambered Social Approach Task for Mice. https://doi.org/10.1002/0471142301.ns0826s56.Automated