

RESEARCH

Open Access



# Prospective association of the infant gut microbiome with social behaviors in the ECHO consortium

Hannah E. Laue<sup>1†</sup>, Kevin S. Bonham<sup>2†</sup>, Modupe O. Coker<sup>3</sup>, Yuka Moroishi<sup>1</sup>, Wimal Pathmasiri<sup>4</sup>, Susan McRitchie<sup>4</sup>, Susan Sumner<sup>4</sup>, Anne G. Hoen<sup>1</sup>, Margaret R. Karagas<sup>1</sup>, Vanja Klepac-Ceraj<sup>2\*†</sup> and Juliette C. Madan<sup>1,5\*†</sup> on behalf of program collaborators for Environmental influences on Child Health Outcomes

## Abstract

**Background** Identifying modifiable risk factors of autism spectrum disorders (ASDs) may inform interventions to reduce financial burden. The infant/toddler gut microbiome is one such feature that has been associated with social behaviors, but results vary between cohorts. We aimed to identify consistent overall and sex-specific associations between the early-life gut microbiome and autism-related behaviors.

**Methods** Utilizing the Environmental influences on Children Health Outcomes (ECHO) consortium of United States (U.S.) pediatric cohorts, we gathered data on 304 participants with fecal metagenomic sequencing between 6-weeks to 2-years postpartum (481 samples). ASD-related social development was assessed with the Social Responsiveness Scale (SRS-2). Linear regression, PERMANOVA, and Microbiome Multivariable Association with Linear Models (MaAsLin2) were adjusted for sociodemographic factors. Stratified models estimated sex-specific effects.

**Results** Genes encoding pathways for synthesis of short-chain fatty acids were associated with higher SRS-2 scores, indicative of ASDs. Fecal concentrations of butyrate were also positively associated with ASD-related SRS-2 scores, some of which may be explained by formula use.

**Limitations** The distribution of age at outcome assessment differed in the cohorts included, potentially limiting comparability between cohorts. Stool sample collection methods also differed between cohorts. Our study population reflects the general U.S. population, and thus includes few participants who met the criteria for being at high risk of developing ASD.

**Conclusions** Our study is among the first multicenter studies in the U.S. to describe prospective microbiome development from infancy in relation to neurodevelopment associated with ASDs. Our work contributes to clarifying

<sup>†</sup>Hannah E. Laue and Kevin S. Bonham have contributed equally to this work.

<sup>†</sup>Vanja Klepac-Ceraj and Juliette C. Madan have equally contributed to this work.

\*Correspondence:

Vanja Klepac-Ceraj  
vklepacc@wellesley.edu

Juliette C. Madan  
Juliette.Madan@dartmouth.edu

Full list of author information is available at the end of the article



which microbial features associate with subsequent diagnosis of neuropsychiatric outcomes. This will allow for future interventional research targeting the microbiome to change neurodevelopmental trajectories.

**Keywords** Microbiome, Autism, Social behavior, Gut metagenome, Gene set enrichment analysis

## Background

The connection between the gut microbiome and the brain has important foundational implications in the critical window following birth until early childhood [1, 2]. During the neonatal period through 3 years of age, microbe-microbe and microbe-human interactions are required for the developing innate and adaptive immune systems to achieve competency [3], and the intestinal microbiome plays a critical role in neurodevelopment during this important window [4–6]. For example, brain volume differences between specific-pathogen-free mice and germ-free mice, even when the microbiomes of the germ-free mice are conventionalized after weaning [7]. Signals from the gut microbiome influence neurophysiology and neurodevelopment via vagus nerve activation, endocrine signaling, and impacts on systemic inflammation and neuroimmune cells [8]. Cross-talk between the gut and developing brain is likely bidirectional and primarily driven by microbially-derived metabolites such as short-chain fatty acids (SCFAs) produced by bacteria as byproducts of fermentation of indigestible polysaccharides, which rapidly change during this critical period of microbiome acquisition and neurodevelopment [9, 10]. Microbial metabolites, including SCFAs, are required for proper functioning and maturation of neuroimmune cells, such as microglia, the innate immune cells critical for neurodevelopment and modulation of neuroinflammation [11].

We recently described associations of early-life gut bacteria and their functions with social behavior at preschool age in a single large United States (U.S.) birth cohort study, complementing research from other groups [12–15]. The gut microbiome is a promising modifiable target for interventions aiming to reduce adverse autism spectrum disorder (ASD) behaviors, which impose a significant financial burden due to increased healthcare utilization, special education costs, and loss of income [16–20]. Modification of the intestinal microbiome results in shifts in key metabolites implicated in social behaviors [13] and alteration of the microbiome can result in reduction of ASD-associated behaviors in both animal and human studies of antibiotics and fecal transplants [21–25]. These and other studies have identified gut microbial differences after ASD diagnosis, when amending the microbiome may be less impactful. Furthermore, a recent case-control study suggested that the mechanism of microbiome-ASD associations is via early

dietary preferences in children with ASD [26], highlighting the need for research with microbiome characterization prior to the manifestation of dietary preferences. Such prospective studies beginning in the first weeks of life can clarify opportunities for altering the microbiome to optimize neurodevelopmental outcomes.

We aimed to prospectively characterize the association between the microbiome beginning at birth and ASD-related social behaviors during childhood when rapid neurodevelopment lays the foundation for lifelong optimal health outcomes. Leveraging the resources of the Environmental influences on Child Health Outcomes (ECHO) consortium funded by the National Institutes of Health, we performed a multi-center investigation to identify prospective associations between the gut microbiome and autism-related social behaviors and to determine whether these associations were consistent across cohorts and ages.

## Methods

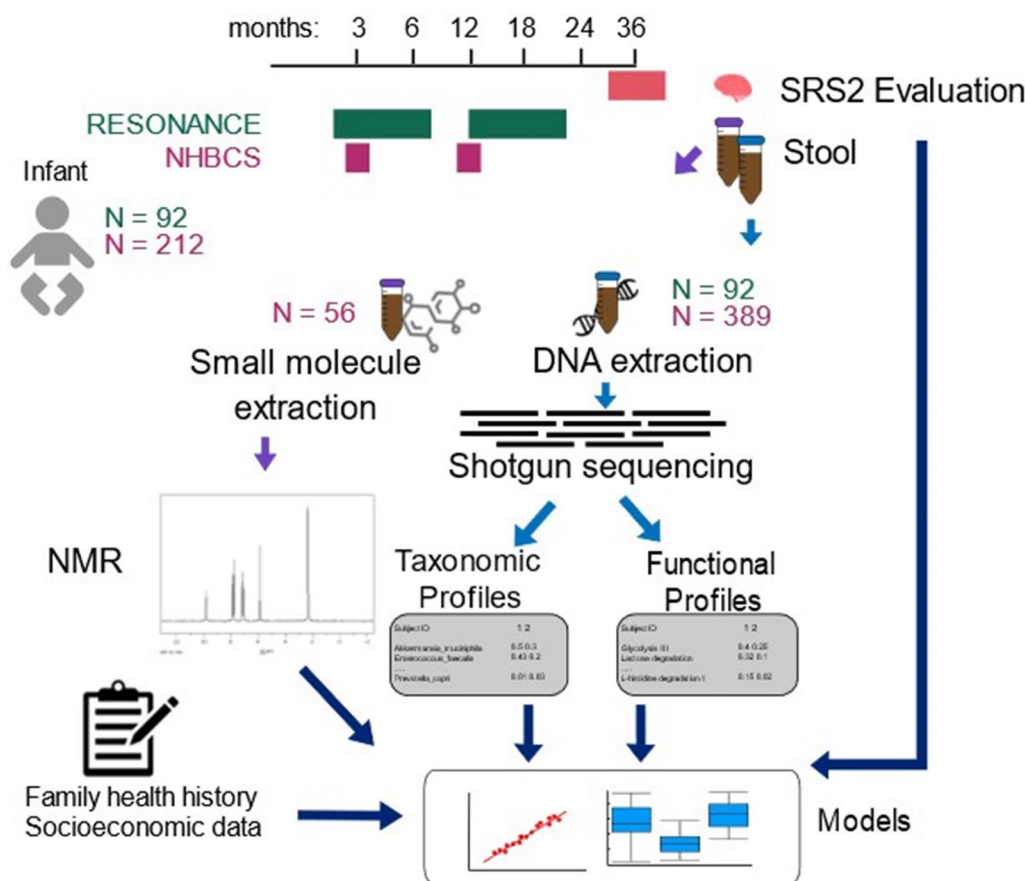
### Cohort descriptions

An overview of the methods is presented in Fig. 1. The RESONANCE cohort and NHBCS are part of ECHO, a consortium that harmonizes data across cohorts to probe scientific questions regarding children's environmental health that are unanswerable by a single cohort [27, 28]. Background on each of the cohorts, including eligibility criteria, is provided in the Supplemental Methods.

### Shotgun metagenomic sequencing

Fecal samples were collected in tubes containing an RNA stabilizer from a subset of child participants for microbiome sequencing and analysis. In the NHBCS, fecal samples ( $n=389$ ) were collected from 212 children at approximately 6 weeks and 12 months of age, stored in a polyethylene bag, and frozen in home freezers or directly transported on ice to maternal postnatal follow-up visits. Within 24 h of receipt, the samples were aliquoted and frozen at 80 °C. We restricted this analysis to samples collected before four years of age ( $n=92$ ). DNA was extracted from the stool using the Zymo DNA extraction kit (Zymo Research, Irvine, California) and prepared for metagenomic sequencing using Illumina NextSeq at Marine Biological Laboratory in Woods Hole, Massachusetts. In the RESONANCE cohort, stool samples were collected by parents in OMR-200 tubes (OMNIgene

# Cohort and Methods



**Fig. 1** Diagram of Study Methods. Fecal samples were collected from RESONANCE participants between six weeks and two years postpartum and from New Hampshire Birth Cohort Study (NHBCS) participants at six weeks and/or one year. In both cohorts, DNA was extracted from stools and underwent shotgun metagenomic sequencing. In both cohorts, assessment of autism-spectrum disorder behaviors with the Social Responsiveness Scale, 2nd edition, was completed after age 3 years. Additionally, a subset of NHBCS fecal samples underwent nuclear magnetic resonance spectroscopy to quantify small molecules. These data were integrated with sociodemographic data into statistical models

GUT, DNA Genotek, Ottawa, Ontario, Canada), immediately stored on ice, and brought within 24 h to the lab in Rhode Island where they were immediately frozen at  $-80^{\circ}\text{C}$ . Stool samples were not collected if the subject had taken antibiotics within the last 2 weeks. DNA extraction was performed at Wellesley College (Wellesley, Massachusetts). Nucleic acids were extracted from stool samples using the Rneasy PowerMicrobiome kit automated on the QIAcube (Qiagen, Germantown, Maryland), excluding the DNA degradation steps. Shotgun metagenomic sequencing was performed on extracted DNA at the Integrated Microbiome Resource (IMR, Dalhousie University, Nova Scotia, Canada).

For the NHBCS, sequencing libraries were prepared using Nugen’s Ovation Ultralow V2 protocol. The

extracted DNA samples were sheared to a mean insert size of 400 bp using a Covaris S220 focused ultrasonicator, resulting in an average of 30 million reads per sample. For the RESONANCE cohort, a pooled library (max 96 samples per run) was prepared using the Illumina Nextera Flex Kit for MiSeq and NextSeq from 1 ng of each sample. Samples were then pooled onto a plate and sequenced on the Illumina NextSeq 550 platform using 150+150 bp paired-end “high-output” chemistry, generating ~400 million raw reads and ~120 Gb of sequence per plate. For both cohorts, the BioBakery v3 pipeline was used to process the raw sequences [29]. KneadData (v0.7.7) was used to trim adapters from raw sequence reads and remove reads matching a human genome reference. Metagenomic Phylogenetic

Analysis (MetaPhlan v3.0.7) was used to generate taxonomic profiles by aligning reads to a reference database of marker genes (mpa\_v30\_CHOCOPHlan\_201901) [29]. Finally, the Human Microbiome Project Unified Metabolic Analysis Network (HUMANn v3.0.0a4 with databases at v201901b) was used to functionally profile metagenomes [29].

### Fecal metabolomics

In the NHBCS, stool samples were collected for metabolomics analysis in trace element free tubes and sent to the Eastern Regional Comprehensive Metabolomics Research Core at the Research Triangle Institute for nuclear magnetic resonance (NMR) analysis. NMR spectra of fecal samples were acquired on a Bruker 700 MHz NMR spectrometer (Billerica, Massachusetts). The relative concentrations of library-matched metabolites associated with host and gut microbe co-metabolism [30–32] were obtained using the Chenomx NMR Suite 8.1 Professional (Chenomx Inc., Edmonton, Alberta, Canada) [33, 34]. Relative concentrations were  $\log_2$ -transformed for analysis.

### Outcome assessment

Caregivers assessed the participant's social skills using the Social Responsiveness Scale, 2nd edition (SRS-2), which captures ASD-related social behaviors in multiple domains, in both cohorts [35]. Standardized total T-scores have a mean (standard deviation, SD) of 50 (10), with higher scores indicative of social behaviors associated with ASD. In the NHBCS, caregivers completed the preschool version of the SRS-2 when participants were ~3 years of age and the school-age version when the participants were 5 years of age. If a preschool assessment was not available, the participant was included with their school-age score. In the RESONANCE cohort, caregivers completed the SRS-2 preschool version for participants 2.5–4 years of age and the school-age version for participants 4–19 years of age.

### Covariates

Covariates were selected a priori based on their potential to confound the association between the microbiome and SRS-2 scores or as predictors of the outcome. At recruitment, study staff implemented questionnaires to collect data on maternal smoking during pregnancy (any vs. none), parity (parous vs. nulliparous), parental age (continuous), and sociodemographic characteristics, including maternal education (any higher education vs. none) and marital status. Characteristics of the birth (delivery mode: C-section vs. vaginal, peripartum antibiotic exposure: any vs. none, gestational age at delivery: continuous, and child's sex) were abstracted from medical records

(NHBCS) or questionnaires (RESONANCE). Repeated questionnaires throughout childhood captured duration of breastfeeding (continuous) and breastfeeding status at the time of stool sample collection (exclusive vs. any formula). Exact age at SRS-2 implementation was calculated as the difference between the date of birth and the date of survey completion (continuous). To reduce selection bias and improve power, we imputed missing covariates using the *mice* R package [36–38]. Details are available in the Supplemental Methods.

### Statistical analysis

Because many NHBCS participants provided two stool samples, we analyzed early (<6 month) and late (6 month–2 year) samples separately. Additionally, we examined NHBCS and RESONANCE samples separately. We tested the overall association between bacterial species or functions and SRS-2 scores with *adonis2*, a PERMANOVA method that allows for quantification of variable influence ( $R^2$ ) and significance (marginal  $p$ -value) in community structure (Bray–Curtis distances) [39–41]. A  $p$ -value less than 0.05 was considered significant. Bacterial diversity was quantified with Shannon and Inverse Simpson indices. We linearly regressed diversity against SRS-2 scores, adjusting for the previously mentioned covariates, considering a  $p$ -value less than 0.05 significant. We tested whether models built on data from the NHBCS could predict SRS-2 scores in the RESONANCE cohort using the caret package in R [42]. Briefly, we used the predict function in the stats R package to predict SRS-2 scores for RESONANCE observations based on their microbiome and covariates and the model constructed using data from the NHBCS. Functions from the caret package were then used to calculate correlations and errors between actual and predicted scores. To screen for bacterial species associated with SRS-2 scores, we used microbiome multivariable association with linear models (MaAsLin2), which models bacterial relative abundance as the outcome [43]. Associations with a false discovery rate (FDR) of  $q < 0.1$  were modeled in linear regression with bacterial relative abundance predicting SRS-2 scores, adjusting for covariates. Given the sex-ratio differential of ASD presentation, we explored whether relationships between the gut microbiome and behavior were different in biological males and females with interaction models [44].

Potentially neuroactive microbial gene sets were acquired from Supplementary Dataset 1 from Valles-Colomer et al. [45] Kyoto Encyclopedia of Genes and Genomes (KEGG) Orthologs (KOs) from this dataset were mapped to UniRef90s using the mapping files included with HUMANn3. Pearson correlations for each measure (SRS-2 scores or residuals from a complete

linear model regressing SRS-2 scores against all covariates) and each UniRef90 gene family found in  $\geq 4$  subjects were calculated. We performed Mann–Whitney U tests (implemented in the HypothesisTests.jl package) on the correlations of each gene family in a neuroactive gene set against the correlations of all gene families not found in the gene set [46]. FDR  $q < 0.1$  was the threshold for significance [47].

To test if the capacity of the gut microbiome to have neuroactive genes associated with SRS-2 scores related to the metabolic profiles of infants, we regressed SRS-2 scores against relative concentrations of fecal metabolites identified as significant from the Mann–Whitney U test that were also annotated in our dataset. Contribution of specific taxa to the relative abundance of neuroactive gene pathways was determined using stratified MetaPhlAn tables. All models were run with R (4.1.1), RStudio (v1.4.1717), and Julia (v1.6.1). Code and package versions are available at <https://github.com/HEL548/ECHOSRS2>.

## Results

### Population characteristics

ECHO participants contributing to the analysis from the New Hampshire Birth Cohort Study (NHBCS) and Rhode Island RESONANCE cohorts had similar distributions of most characteristics (Table 1). As expected, RESONANCE participants were older at SRS-2 completion ( $8 \pm 2.7$  years compared with  $3.2 \pm 0.4$  in NHBCS). As previously described [12], social behaviors of NHBCS participants were rated as having behavior less indicative of ASD ( $43.6 \pm 4.8$ ) than the normative population ( $50 \pm 10$ ), whereas RESONANCE participants were rated similar to the normative population ( $49.9 \pm 8.5$ ). RESONANCE stool samples, which were collected from older participants (6 months–2 years) than the NHBCS stool samples (6 weeks–2 years), had higher Shannon diversity ( $2.9 \pm 0.5$ ) than the NHBCS samples ( $1.9 \pm 0.7$ ).

### Gut microbiome communities and SRS-2 scores

Bray–Curtis differences in 6-week-old infants’ microbiota in the NHBCS contributed significantly to variability in SRS-2 scores (Additional file 1: Table S1,  $R^2 = 0.01$ ,  $p = 0.016$ ). This association had a similar magnitude among males and females and was of borderline statistical significance in males ( $R^2 = 0.008$ ,  $p = 0.08$ ), but not females ( $R^2 = 0.006$ ,  $p = 0.337$ ). In contrast, microbial species in one-year-old infants did not significantly explain variability in SRS-2 scores ( $R^2 = 0.006$ ,  $p = 0.221$ ). Microbial community structure in the RESONANCE cohort was marginally related to SRS-2 scores, explaining approximately 1.7% of the variability ( $p = 0.068$ ). No association was observed between alpha diversity and SRS-2

**Table 1** Population characteristics of New Hampshire birth cohort study (NHBCS) and RESONANCE participants included in any analysis [Mean  $\pm$  Standard Deviation (% Missing If Any) or n (%)]

	NHBCS participants (n = 212)	Resonance participants (n = 92)
Maternal age (years)	32.4 $\pm$ 4.2	31.3 $\pm$ 5.1
Paternal age (years)	33.9 $\pm$ 5.8 (2.4)	34 $\pm$ 6.3 (10.9)
Maternal smoking during pregnancy		
None	200 (94.3)	85 (92.4)
Any	9 (4.2)	5 (5.4)
Missing	3 (1.4)	2 (2.2)
Maternal education		
Any graduate	93 (43.9)	67 (72.8)
Less than graduate	113 (53.3)	25 (27.2)
Missing	6 (2.8)	0 (0)
Marital status		
Married	188 (88.7)	75 (81.5)
Unmarried	18 (8.5)	17 (18.5)
Missing	6 (2.8)	0 (0)
Parity		
Parous	104 (49.1)	69 (75)
Nulliparous	105 (49.5)	23 (25)
Missing	3 (1.4)	0 (0)
Missing	3 (1.4)	0 (0)
Delivery mode		
Vaginal	149 (70.3)	63 (68.5)
C-section	58 (27.4)	28 (30.4)
Missing	5 (2.4)	1 (1.1)
Breastfeeding at stool collection		
No	78 (36.8)	3 (3.3)
Yes	112 (52.8)	71 (77.2)
Missing	22 (10.4)	18 (19.6)
Peripartum antibiotics		
Any	112 (52.8)	0 (0)
None	95 (44.8)	92 (100)
Missing	5 (2.4)	0 (0)
Gestational age (weeks)	39.2 $\pm$ 1.7	38.5 $\pm$ 1.8
Child’s sex		
Male	119 (56.1)	52 (56.5)
Female	93 (43.9)	40 (43.5)
Child age at follow-up (years)	3.2 $\pm$ 0.4	8 $\pm$ 2.7
Social responsiveness scale T-score	43.6 $\pm$ 4.8	49.9 $\pm$ 8.5
Shannon Index	1.9 $\pm$ 0.7	2.9 $\pm$ 0.5

scores, although increased diversity trended toward being associated with social behavior more indicative of ASD among girls (Additional file 1: Table S2). The models built on NHBCS data did not predict SRS-2 scores in the RESONANCE cohort (Additional file 1: Table S3). Most associations between bacterial species and SRS-2

scores did not reach statistical significance after adjusting for multiple comparisons (Fig. 2, Additional file 1: Table S4). Among females, higher relative abundance of *Enterococcus faecalis* at one year was associated with ASD-related social behaviors in the complete case analysis [ $\beta=0.41$  difference in SRS-2 score per doubling species relative abundance, 95% CI (0.17–0.65),  $p<0.001$ ,  $q=0.002$ ] whereas we observed no association among males [ $\beta=-0.07$ , 95% CI (-0.31–0.18),  $p=0.6$ ,  $q=0.71$ ,  $p_{\text{interaction}}=0.006$ ]. While this association was nominally significant in the multiple imputation analysis, the effect size was smaller among females and did not meet the threshold after adjusting for multiple comparisons [ $\beta=0.23$ , 95% CI (0.01–0.14),  $q=0.34$ ]. Higher relative abundance of *Anaerostipes caccae* was associated with worse social behavior among males [ $\beta=0.31$ , 95% CI (0.16–0.47),  $p<0.001$ ,  $q=0.06$ ] and nominally associated with better social behavior among females [ $\beta=-0.2$ , 95% CI (-0.37 to -0.03),  $p=0.02$ ,  $q=0.73$ ,  $p_{\text{interaction}}<0.001$ ] in the complete case analysis. The association among males was nominally significant in the multiple imputation analysis, but the effect size was smaller [ $\beta=0.27$ , 95% CI (0.11–0.42),  $p<0.001$ ] and did not meet the FDR threshold ( $q=0.27$ ). These associations were not replicable in the RESONANCE cohort due to low prevalence of these species (Fig. 2c).

### Functional potential of gut microorganisms

Differences in the functional capacity of gut bacteria in fecal samples collected at six weeks from NHBCS infants explained 1% of the variability in SRS-2 scores (Additional file 1: Table S1,  $p=0.048$ ) and the contribution of 1-year-old participants was similar (0.9%,  $p=0.067$ ). Feature set enrichment analysis (FSEA) demonstrated a strong association between genes encoding for SCFA synthesis at 6 weeks of age and ASD-related SRS-2 scores after multiple hypothesis correction, for both raw SRS-2 scores and linear model residuals (Fig. 3, Additional file 1: Table S5; butyrate synthesis SRS-2 scores, med=0.064, model residuals: med=0.050, both  $q=0.066$ ; acetate synthesis and SRS-2 scores: med=0.025,  $q=0.066$ , and model residuals: med=0.012,  $q=0.18$ ). *Escherichia coli* was the largest contributor of butyrate synthesis genes

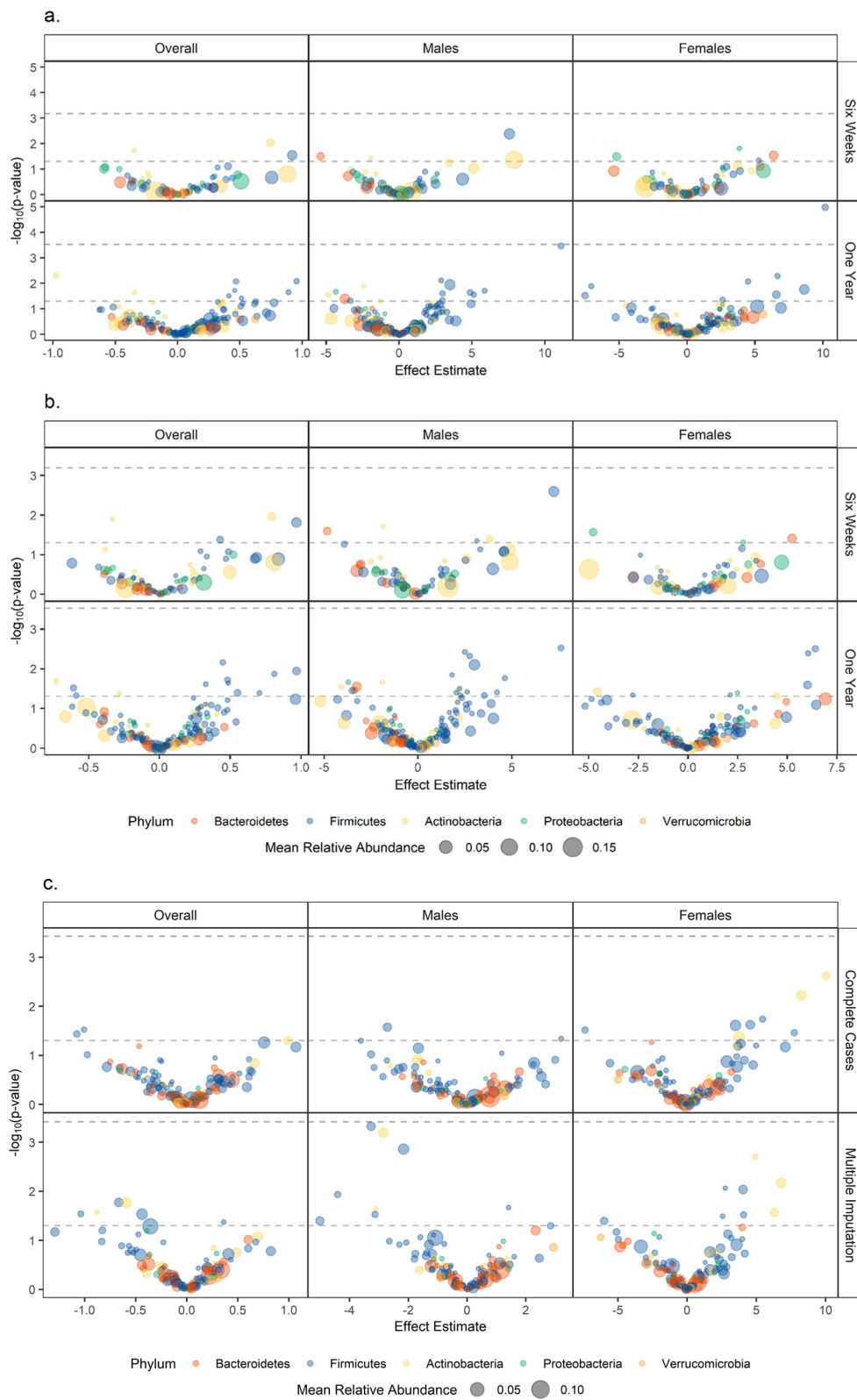
in the NHBCS, determined from the species-stratified table of KO abundance. Interestingly, both synthesis and degradation of the 3-carbon SCFA propionate were significantly associated with ASD-related scores (synthesis, med=0.107,  $q=0.066$ ; degradation, med=0.161,  $q=0.0047$ ). Strikingly, this association was not observed in microbiome samples from older children (Additional file 1: Table S5). Synthesis of the branched SCFA isovaleric acid at 1 year of age was associated with more optimal social behaviors as measured by the SRS-2, but we did not observe any associations with the metabolism of butyrate, propionate, or acetate captured at this age. Rather, we observed a small, but very significant, association between the synthesis of menaquinone, or vitamin K2 (med=-0.048,  $q<0.001$ ), and more optimal social behaviors. We utilized data from the RESONANCE cohort to validate these findings and observed the same association of propionate synthesis in infants (birth to 4 months of age) with autism-related behaviors in mid-childhood (med=0.096,  $p=0.037$ ) and between menaquinone synthesis in older participants (age>9 months) and more optimal social behaviors (med=-0.070,  $p=0.033$ ). As expected, correlations between all gene families and each measure were approximately normally distributed (Fig. 3c).

### Fecal metabolites associated with SRS-2 scores

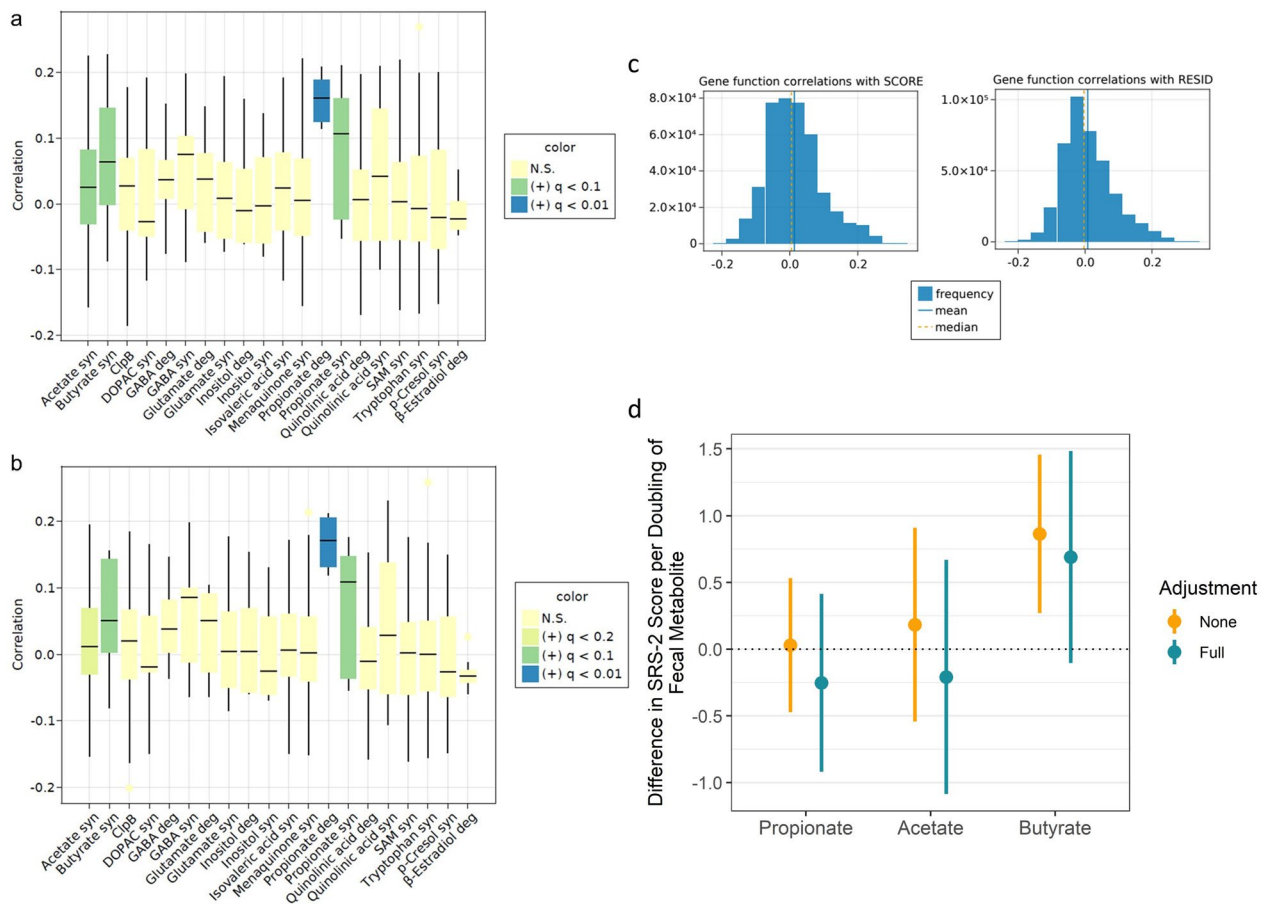
In the NHBCS, the relative concentrations of microbial-related metabolites, including propionate, acetate, and butyrate, were quantified in 6-week-old stool samples from 56 participants. Menaquinone and isovaleric acid were not annotated. Concentrations of acetate and propionate were not significantly associated with SRS-2 scores in unadjusted or adjusted models, whereas higher butyrate concentrations related to higher SRS-2 scores in unadjusted models [ $\beta=0.86$  point increase per doubling of butyrate, 95% CI (0.27–1.46),  $p=0.006$ ; Fig. 3d]. This association was attenuated and no longer significant in adjusted models [ $\beta=0.69$  point increase per doubling of butyrate, 95% CI (-0.1–1.48),  $p=0.097$ ], likely due to a strong association between formula use and increased butyrate, including in NHBCS [34].

(See figure on next page.)

**Fig. 2** Associations Between Bacterial Species Relative Abundance and Social Responsiveness Scale, 2nd Edition (SRS-2), Scores. **A** Complete cases in the New Hampshire Birth Cohort Study (NHBCS). **B** Subjects with missing covariate data imputed in the NHBCS. **C** Among RESONANCE participants. Models adjusted for gestational age, child's sex, peripartum antibiotic exposure, delivery mode, exclusive breastfeeding at the time of stool sample collection, maternal smoking during pregnancy, maternal and paternal age, maternal marital status, parity, maternal education, and the child's age at SRS-2 completion. Each point reflects the association between a bacterial species and SRS-2 scores. Points are colored by phylum. Dashed gray lines indicate  $p=.05$  and the Bonferroni corrected  $p$ -value for each analysis



**Fig. 2** (See legend on previous page.)



**Fig. 3** Associations of Microbial Functions and Social Responsiveness Scale, 2nd Edition (SRS-2) Scores. Feature Set Enrichment Analysis of **A** Raw SRS-2 scores and **B** regression model residuals with microbial gene sets. **C** Distribution of correlations between microbial gene families and SRS-2 scores (SCORE), and regression model residuals (RESID). **D** Association (point estimate and 95% confidence interval) between concentrations of fecal short-chain fatty acids and SRS-2 Scores in the Unadjusted and Fully Adjusted Models. Fully adjusted models include gestational age, child's sex, peripartum antibiotic exposure, delivery mode, exclusive breastfeeding at the time of stool sample collection, maternal smoking during pregnancy, maternal and paternal age, maternal marital status, parity, maternal education, and the child's age at SRS-2 completion

### Discussion

In this longitudinal analysis of the infant and toddler microbiome and social behaviors associated with ASD measured at preschool age, we found suggestive evidence that bacterial species, neuroactive bacterial functions, and metabolite patterns were prospectively associated with ASD-related traits. Our study adds to the existing body of research by examining the prospective development of the intestinal microbiome in early life in relation to subsequent autism-associated behaviors. The findings of this study highlight opportunities for identifying biomarkers for early detection and developing therapies for intervention and prevention of ASD.

We identified a significant association between the genes involved in the synthesis of specific SCFAs, particularly butyrate and acetate, in early life and ASD-related social behavior at approximately age 3. This signal was not evident at older ages, highlighting the

importance of early-life microbiome factors in influencing later behavioral outcomes. Butyrate, in particular, has been linked to immune regulation and reduced inflammation, which are relevant to ASD [48]. It is important to note that we cannot directly infer gut concentrations of molecules based on the differences in gene abundances. Genes abundances reflect selective pressures encountered in the gut ecosystem and do not necessarily mean those genes are expressed. Further, low concentrations of SCFA in the gut may select for microbes that produce them, leading to a higher presence of SCFA synthesis genes.

To address this limitation, we augmented our analysis of the functional capacity of the gut microbiome with fecal metabolomic data in a subset of participants. These data supported our finding of an adverse association between butyrate and SRS-2 scores, but not acetate or propionate. Higher concentrations of butyrate, an



essential energy source for colonic epithelial cells and gut barrier functions that modulate systemic inflammation [49], have been observed in children with clinical ASD < 5 years of age [50]. In one study, a paucity of *Bifidobacterium longum* in 2- to 4-year-olds with ASD, and an increase in *Faecalibacterium prausnitzii*, a known butyrate producer, were highlighted [50], suggesting that the substrates produced during the fermentation of nondigestible oligosaccharides by *Bifidobacteria* and Ruminococcaceae are then utilized by butyrate producers such as *Faecalibacterium*, *Eubacterium*, *Roseburia*, and *Anaerostipes* [51]. Although *Anaerostipes hadrus* was a major contributor of butyrate synthesizing genes, *E. coli* provided the most butyrate synthesis genes in the NHBCS. This is likely due to the excellent characterization of the *E. coli* genome and its high prevalence in NHBCS infants. It is plausible that butyrate influences social behaviors through modulation of the excitatory/inhibitory balance, but the exact mechanism remains uncharacterized. Because both propionate synthesis and degradation pathways were adversely associated with SRS-2 scores, it is possible that the turnover of propionate rather than its concentration is significant to neurobehavior, or that SCFA concentrations must be in a “Goldilocks zone” requiring regulation of both synthesis and degradation [52]. Additionally, the concentration of microbial metabolites in feces, while proximal to their source, may not reflect circulating bioactive molecules in the host. Further research correlating the serum metabolome data with fecal metagenomes and metabolomes in very young children may provide additional insights into this association.

Increased capacity for menaquinone synthesis in older infants was associated with behavioral scores less indicative of ASD in both cohorts. Menaquinone may have anti-inflammatory and antioxidant effects in the brain [53], although it is unclear what proportion of host circulating menaquinone is bacterially-derived [54]. A successful probiotic intervention to relieve anxiety was shown to increase bacterial capacity for menaquinone production [55], among other metabolic changes, suggesting that bacterially-derived menaquinone may relate to better neurobehavioral health [56]. Additional metabolomic data could clarify our understanding of the importance of bacterial production of menaquinone in the human host.

### Limitations

While this study advances our understanding of the microbiome-autism relationship, it is not without limitations. Most NHBCS participants’ social behavior was rated with the preschool-age form, whereas most RESONANCE participants were rated with the school-age form, potentially weakening the validation analysis.

However, the use of T-scores reduces variation resulting from differences in the form. Furthermore, SRS-2 scores have high consistency over time, minimizing the effect of age at which the surveys were administered in the two cohorts [57]. Another limitation is the variation in stool sample collection methods between the two studies, including the timing of sample collection, which may have led to differential detection of certain species; much of the concern about collection and processing methods may be alleviated with a rigorous comparison study of methods identifying data alignment with variable methods [58].

Our study, while leveraging the resources and population of the ECHO consortium, had a limited sample size, including few participants who met the criteria for high risk of ASD. However, it reflects a population-based evaluation of neurodevelopment, which, when combined with case-control and mechanistic studies, adds valuable insight into features that alter subclinical symptoms of ASD. Our sex-specific analysis was likely underpowered, but given the known sex-specific differences in ASD prevalence, these findings should be replicated in another cohort. We incorporated exclusive human milk consumption and duration of human milk consumption in our models, but did not account for dietary patterns in old infants/toddlers. Further research involving clinical populations and animal models is needed to validate our findings and strengthen their applicability and translatability to clinical care.

### Conclusion

In conclusion, this study employed metagenomic sequencing data and advanced analysis methods to investigate the relationship between the infant microbiome and ASD-related behaviors. The use of FSEA allowed us to examine the metabolism of neuroactive molecules for associations with SRS-2 scores with more power than an untargeted pathway analysis. The validation set approach is novel in microbiome epidemiology and is crucial as the field advances. Leveraging the substantial resources of the ECHO consortium enabled this multicenter evaluation of two large US birth cohorts, providing a robust representation of the general population; this enabled an evaluation of the prospective risk of typical versus atypical neurodevelopment from the perspective of the gut-brain axis. This study demonstrated associations between SCFA synthesis, menaquinone production, and social behaviors in early life, suggesting potential avenues for early intervention and prevention. Our research contributes to the burgeoning research by beginning to clarify the features of the microbiome that relate to subsequent behaviors prior to neuropsychiatric diagnoses. In this future, this will allow for research into interventions

## targeting the microbiome and changing neurodevelopmental trajectories.

### Abbreviations

ASD	Autism spectrum disorder
ECHO	Environmental influences on children health outcomes
KEGG	Kyoto encyclopedia of genes and genomes
NHBCS	New Hampshire Birth Cohort Study
NMR	Nuclear magnetic resonance
SCFA	Short-chain fatty acid
SRS-2	Social Responsiveness Scale, 2nd edition

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13229-024-00597-2>.

**Additional file 1:** Supplementary material contains the information about cohort characteristics and exclusion criteria. The following supplementary tables are included: Contribution of Bacterial Species and Functions to SRS-2 Scores determined with Bray-Curtis distances; Difference in SRS-2 Score per Standard Deviation Increase in Bacterial Species Diversity in the New Hampshire Birth Cohort Study; Validation of Diversity Models in RESONANCE Cohort with Multiple Imputation of Missing Covariates; Family Set Enrichment Analysis Results Among New Hampshire Birth Cohort Infants.

### Acknowledgements

The authors wish to thank our ECHO colleagues, the medical, nursing, and program staff; and the children and families participating in the ECHO cohorts. We also acknowledge the contribution of the following ECHO program collaborators: *ECHO Components*—Coordinating Center: Duke Clinical Research Institute, Durham, North Carolina: Smith PB, Newby KL; Data Analysis Center: Johns Hopkins University Bloomberg School of Public Health, Baltimore, Maryland: Jacobson LP; Research Triangle Institute, Durham, North Carolina: Catellier DJ. *Person-Reported Outcomes Core*: Northwestern University, Evanston Illinois: Gershon R, Cella D. *ECHO Awardees and Cohorts*—Rhode Island Hospital, Providence, Rhode Island: Koinis Mitchell D; Rhode Island Hospital, Providence, Rhode Island: Deoni S, D'Sa V.

### Author contributions

Conceptualization: JCM, VKC, MRK Data curation: KSB, YM, AGH, WP, SM, SS Formal Analysis: HEL, KSB, Funding acquisition: JCM, VKC, MRK, AGH Investigation: MRK, Methodology: HEL, KSB, MOC, YM Project administration: JCM, VKC, Resources: HEL, KSB, YM Supervision: JCM, VKC, Validation: Visualization: HEL, KSB, Writing—original draft: HEL, KSB, JCM, VKC, MRK, MOC Writing – review & editing: HEL, KSB, JCM, VKC, MRK, MOC, AGH, YM, WP, SM, SS.

### Funding

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Research reported in this publication was supported by the Environmental influences on Child Health Outcomes (ECHO) program, Office of The Director, National Institutes of Health under Award Numbers U2COD023375 (Coordinating Center), U24OD023382 (Data Analysis Center), U24OD023319 with co-funding from the Office of Behavioral and Social Science Research (PRO Core), UH3OD023275 (Karagas), and UG3OD023313 (Deoni, D'Sa, Mueller); National Institute of Environmental Health (P01ES022832, P20ES018175, P42ES007373), the National Institute of General Medical Sciences (P20GM104416), the National Institute of Diabetes and Digestive and Kidney Diseases (U24DK097193), and the U.S. Environmental Protection Agency (RD-83544201).

### Availability of data and materials

The sequencing data used in this study are available through the National Center for Biotechnology Information (NCBI) Sequence Read Archive: <https://ncbi.nlm.nih.gov/sra> under accession numbers PRJNA296814 and PRJNA695570. Epidemiologic data are not publicly available due to their

sensitive and identifiable nature, but are available from the corresponding author on reasonable request.

### Declarations

#### Ethics approval and consent to participate

NHBCS: All study protocols were reviewed and approved by the Center for the Protection of Human Subjects at Dartmouth, and all methods were carried out in accordance with relevant guidelines and regulations. Caregivers provided written and informed consent. RESONANCE: All procedures were approved by the local institutional review board at Rhode Island Hospital, and all experiments adhered to the regulation of the review board. Written informed consent was obtained from all parents or legal guardians of enrolled participants.

#### Competing interests

The authors declare no potential, perceived, or real conflict of interest. The funder had no role in (1) the study design, (2) data collection, analysis, or interpretation, (3) writing the report, (4) the decision to submit the manuscript for publication. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. HEL, KSB, MOC, YM, MRK, VKC, and JCM wrote the first draft of the manuscript. No authors received an honorarium, grant, or other form of payment to produce the manuscript.

#### Author details

<sup>1</sup>Department of Epidemiology, Geisel School of Medicine, Dartmouth College, Hanover, NH, USA. <sup>2</sup>Department of Biological Sciences, Wellesley College, 106 Central Street, Wellesley, MA 02481, USA. <sup>3</sup>School of Dental Medicine, Rutgers University, Newark, NJ, USA. <sup>4</sup>Department of Nutrition, Nutrition Research Institute, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA. <sup>5</sup>Departments of Pediatrics and Psychiatry, One Medical Center Drive, Dartmouth Hitchcock Medical Center, Lebanon, NH 03756, USA.

Received: 20 March 2024 Accepted: 11 April 2024

Published online: 17 May 2024

### References

- Jena A, et al. Gut-brain axis in the early postnatal years of life: a developmental perspective. *Front Integr Neurosci*. 2020;14:44–44. <https://doi.org/10.3389/fnint.2020.00044>.
- Laue HE, Coker MO, Madan JC. The developing microbiome from birth to 3 years: the gut-brain axis and neurodevelopmental outcomes. *Front Pediatr*. 2022;10:815885. <https://doi.org/10.3389/fped.2022.815885>.
- Belkaid Y, Hand TW. Role of the microbiota in immunity and inflammation. *Cell*. 2014;157:121–41. <https://doi.org/10.1016/j.cell.2014.03.011>.
- Carlson AL, et al. Infant gut microbiome associated with cognitive development. *Biol Psychiatr*. 2018;83:148–59. <https://doi.org/10.1016/j.biopsych.2017.06.021>.
- Coley EJJ, Hsiao EY. Malnutrition and the microbiome as modifiers of early neurodevelopment. *Trends Neurosci*. 2021;44:753–64. <https://doi.org/10.1016/j.tins.2021.06.004>.
- Kau AL, Ahern PP, Griffin NW, Goodman AL, Gordon JI. Human nutrition, the gut microbiome and the immune system. *Nature*. 2011;474:327–36. <https://doi.org/10.1038/nature10213>.
- Yeo XY, et al. Nuanced contribution of gut microbiome in the early brain development of mice. *Gut Microbes*. 2023;15:2283911. <https://doi.org/10.1080/19490976.2023.2283911>.
- Osadchiv V, Martin CR, Mayer EA. The gut-brain axis and the microbiome: mechanisms and clinical implications. *Clin Gastroenterol Hepatol*. 2019;17:322–32. <https://doi.org/10.1016/j.cgh.2018.10.002>.
- Tolhurst G, et al. Short-chain fatty acids stimulate glucagon-like peptide-1 secretion via the G-protein-coupled receptor FFAR2. *Diabetes*. 2012;61:364–71. <https://doi.org/10.2337/db11-1019>.
- Yano JM, et al. Indigenous bacteria from the gut microbiota regulate host serotonin biosynthesis. *Cell*. 2015;161:264–76. <https://doi.org/10.1016/j.cell.2015.02.047>.

11. Erny D, et al. Host microbiota constantly control maturation and function of microglia in the CNS. *Nat Neurosci*. 2015;18:965–77. <https://doi.org/10.1038/nn.4030>.
12. Laue HE, Korrick SA, Baker ER, Karagas MR, Madan JC. Prospective associations of the infant gut microbiome and microbial function with social behaviors related to autism at age 3 years. *Sci Rep*. 2020;10:15515. <https://doi.org/10.1038/s41598-020-72386-9>.
13. Sharon G, et al. Human gut microbiota from autism spectrum disorder promote behavioral symptoms in mice. *Cell*. 2019;177:1600–1618.e1617. <https://doi.org/10.1016/j.cell.2019.05.004>.
14. Cryan JF, O'Riordan KJ, Sandhu K, Peterson V, Dinan TG. The gut microbiome in neurological disorders. *Lancet Neurol*. 2020;19:179–94. [https://doi.org/10.1016/s1474-4422\(19\)30356-4](https://doi.org/10.1016/s1474-4422(19)30356-4).
15. Li Q, Han Y, Dy ABC, Hagerman RJ. The gut microbiota and autism spectrum disorders. *Front Cell Neurosci*. 2017;11:120. <https://doi.org/10.3389/fncel.2017.00120>.
16. Kogan MD, et al. A national profile of the health care experiences and family impact of autism spectrum disorder among children in the United States, 2005–2006. *Pediatrics*. 2008;122:e1149–1158. <https://doi.org/10.1542/peds.2008-1057>.
17. Cidav Z, Marcus SC, Mandell DS. Implications of childhood autism for parental employment and earnings. *Pediatrics*. 2012;129:617–23. <https://doi.org/10.1542/peds.2011-2700>.
18. Montes G, Halterman JS. Association of childhood autism spectrum disorders and loss of family income. *Pediatrics*. 2008;121:e821–826. <https://doi.org/10.1542/peds.2007-1594>.
19. Liptak GS, Stuart T, Auinger P. Health care utilization and expenditures for children with autism: data from U.S. national samples. *J Autism Dev Disord*. 2006;36:871–9. <https://doi.org/10.1007/s10803-006-0119-9>.
20. Shimabukuro TT, Grosse SD, Rice C. Medical expenditures for children with an autism spectrum disorder in a privately insured population. *J Autism Dev Disord*. 2008;38:546–52. <https://doi.org/10.1007/s10803-007-0424-y>.
21. Zhao H, et al. Mo1667 Fecal microbiota transplantation for children with autism spectrum disorder. *Gastrointest Endosc*. 2019;89:AB512–3. <https://doi.org/10.1016/j.gie.2019.03.857>.
22. Tan Q, et al. Probiotics, prebiotics, synbiotics, and fecal microbiota transplantation in the treatment of behavioral symptoms of autism spectrum disorder: a systematic review. *Autism Res*. 2021;14:1820–36. <https://doi.org/10.1002/aur.2560>.
23. Abuaiash S, et al. Fecal transplant and bifidobacterium treatments modulate gut clostridium bacteria and rescue social impairment and hippocampal BDNF expression in a rodent model of autism. *Brain Sci*. 2021. <https://doi.org/10.3390/brainsci11081038>.
24. Xiao L, et al. Fecal microbiome transplantation from children with autism spectrum disorder modulates tryptophan and serotonergic synapse metabolism and induces altered behaviors in germ-free mice. *mSystems*. 2021;6:e01343–e1320. <https://doi.org/10.1128/mSystems.01343-20>.
25. Kang D-W, et al. Distinct fecal and plasma metabolites in children with autism spectrum disorders and their modulation after microbiota transfer therapy. *mSphere*. 2020;5:e00314–00320. <https://doi.org/10.1128/mSphere.00314-20>.
26. Yap CX, et al. Autism-related dietary preferences mediate autism-gut microbiome associations. *Cell*. 2021;184:5916–5931.e5917. <https://doi.org/10.1016/j.cell.2021.10.015>.
27. Gillman MW, Blaisdell CJ. Environmental influences on child health outcomes, a research program of the national institutes of health. *Curr Opin Pediatr*. 2018;30:260–2. <https://doi.org/10.1097/MOP.0000000000000600>.
28. Forrest CB, Blackwell CK, Camargo CA Jr. Advancing the Science of Children's Positive Health in the National Institutes of Health Environmental Influences on Child Health Outcomes (ECHO) Research Program. *J Pediatr*. 2018;196:298–300. <https://doi.org/10.1016/j.jpeds.2018.02.004>.
29. Beghini F, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3. *Elife*. 2021;10:e65088. <https://doi.org/10.7554/eLife.65088>.
30. Li JV, et al. Global metabolic responses of NMRI mice to an experimental *Plasmodium berghei* infection. *J Proteome Res*. 2008;7:3948–56. <https://doi.org/10.1021/pr800209d>.
31. Zheng X, et al. The footprints of gut microbial-mammalian co-metabolism. *J Proteome Res*. 2011;10:5512–22. <https://doi.org/10.1021/pr2007945>.
32. Nicholson JK, et al. Metabolic phenotyping in clinical and surgical environments. *Nature*. 2012;491:384–92. <https://doi.org/10.1038/nature11708>.
33. Weljie AM, Newton J, Mercier P, Carlson E, Slupsky CM. Targeted profiling: quantitative analysis of 1H NMR metabolomics data. *Anal Chem*. 2006;78:4430–42. <https://doi.org/10.1021/ac060209g>.
34. Hoen AG, et al. Association of cesarean delivery and formula supplementation with the stool metabolome of 6-week-old infants. *Metabolites*. 2021. <https://doi.org/10.3390/metabo11100702>.
35. Constantino JN, Gruber CP. Social Responsiveness Scale Second edition (SRS-2) Manual. Western Psychological Services; 2012.
36. Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? *Int J Methods Psychiatr Res*. 2011;20:40–9. <https://doi.org/10.1002/mpr.329>.
37. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Softw*. 2011;45:1–67. <https://doi.org/10.18637/jss.v045.i03>.
38. Harel O, et al. Multiple imputation for incomplete data in epidemiologic studies. *Am J Epidemiol*. 2018;187:576–84. <https://doi.org/10.1093/aje/kwx349>.
39. Oksanen J, et al. The vegan package. *Commun Ecol Packag*. 2007;10:719.
40. Stevens MHH *adonis: Permutational Multivariate Analysis of Variance*, <<https://rdrr.io/rforge/vegan/man/adonis.html>> (2020).
41. Bray JR, Curtis JT. An ordination of the upland forest communities of southern Wisconsin. *Ecol Monogr*. 1957;27:325–49. <https://doi.org/10.2307/1942268>.
42. Kuhn M. Building predictive models in R using the caret package. *J Stat Softw*. 2008;28:1–26. <https://doi.org/10.18637/jss.v028.i05>.
43. Mallick H et al. Multivariable association in population-scale meta-omics studies, <<https://huttenhower.sph.harvard.edu/maaslin/>> (2020).
44. Laue HE, et al. Sex-specific relationships of the infant microbiome and early-childhood behavioral outcomes. *Pediatr Res*. 2021;92(2):580–91. <https://doi.org/10.1038/s41390-021-01785-z>.
45. Valles-Colomer M, et al. The neuroactive potential of the human gut microbiota in quality of life and depression. *Nat Microbiol*. 2019;4:623–32. <https://doi.org/10.1038/s41564-018-0337-x>.
46. HypothesisTests.jl (GitHub, 2021).
47. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B (Methodological)*. 1995;57:289–300.
48. Koenig JE, et al. Succession of microbial consortia in the developing infant gut microbiome. *Proc Natl Acad Sci*. 2011;108:4578–85. <https://doi.org/10.1073/pnas.1000081107>.
49. Rivièrè A, Selak M, Lantin D, Leroy F, De Vuyst L. Bifidobacteria and butyrate-producing colon bacteria: importance and strategies for their stimulation in the human gut. *Front Microbiol*. 2016;7:979. <https://doi.org/10.3389/fmicb.2016.00979>.
50. Coretti L, et al. Gut microbiota features in young children with autism spectrum disorders. *Front Microbiol*. 2018. <https://doi.org/10.3389/fmicb.2018.03146>.
51. Baxter NT, et al. Dynamics of human gut microbiota and short-chain fatty acids in response to dietary interventions with three fermentable fibers. *MBio*. 2019;10:e02566–e2518. <https://doi.org/10.1128/mBio.02566-18>.
52. Kratsman N, Getselter D, Elliott E. Sodium butyrate attenuates social behavior deficits and modifies the transcription of inhibitory/excitatory genes in the frontal cortex of an autism model. *Neuropharmacology*. 2016;102:136–45. <https://doi.org/10.1016/j.neuropharm.2015.11.003>.
53. Farhadi Moghadam B, Fereidoni M. Neuroprotective effect of menaquinone-4 (MK-4) on transient global cerebral ischemia/reperfusion injury in rat. *PLoS ONE*. 2020;15:e0229769. <https://doi.org/10.1371/journal.pone.0229769>.
54. Beulens JWJ, et al. The role of menaquinones (vitamin K2) in human health. *Br J Nutr*. 2013;110:1357–68. <https://doi.org/10.1017/S0007114513001013>.
55. Lew L-C, et al. Probiotic *Lactobacillus plantarum* P8 alleviated stress and anxiety while enhancing memory and cognition in stressed adults: a randomised, double-blind, placebo-controlled study. *Clin Nutr*. 2019;38:2053–64. <https://doi.org/10.1016/j.clnu.2018.09.010>.
56. Ma T, et al. Probiotic consumption relieved human stress and anxiety symptoms possibly via modulating the neuroactive potential of the gut

microbiota. *Neurobiol Stress*. 2021;14:100294. <https://doi.org/10.1016/j.ynstr.2021.100294>.

57. Haraguchi H, Stickley A, Saito A, Takahashi H, Kamio Y. Stability of autistic traits from 5 to 8 years of age among children in the general population. *J Autism Dev Disord*. 2019;49:324–34. <https://doi.org/10.1007/s10803-018-3770-z>.
58. Antosca K, et al. Reliability of stool microbiome methods for DNA yields and sequencing among infants and young children. *Microbiol Open*. 2020;9:e1018–e1018. <https://doi.org/10.1002/mbo3.1018>.

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.