CAMBRIDGE
UNIVERSITY PRESS

Regular Article

A longitudinal study of the gut microbiota during the first three years of life: Links with problem behavior and executive functions at preschool age

Yvonne Willemsen[*1] ⬤, Yangwenshan Ou[*1,2], Clara Belzer[2], Alejandro Arias Vásquez[3], Hauke Smidt[2], Roseriet Beijers[1,4] and Carolina de Weerth[1]

[1]Department of Cognitive Neuroscience, Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Center, Nijmegen, The Netherlands, [2]Laboratory of Microbiology, Wageningen University & Research, Wageningen, The Netherlands, [3]Department of Psychiatry and Human Genetics, Donders Center for Medical Neuroscience, Radboud University Medical Center, Nijmegen, The Netherlands and [4]Behavioral Science Institute, Radboud University, Nijmegen, The Netherlands

Abstract

Early life is a sensitive period when microbiota-gut-brain interactions may have important impact on development. This study investigated the associations of the gut microbiota in the first three years of life (two, six, and 12 weeks, and one and three years) with problem behavior and executive functions in $N = 64$ three-year-old children. Higher relative abundance of *Streptococcus* at the age of two weeks, as well as its trajectory over time (including ages two, six and 12 weeks, and one and three years), was related to worse executive functions. Higher relative abundance of [*Ruminococcus*] *torques* group at the age of three years, as well as its trajectory from one to three years, was associated with less internalizing behavior. Besides, several robust age-specific associations were identified: higher *Bifidobacterium* relative abundance (age three years) was associated with more internalizing and externalizing issues; higher *Blautia* relative abundance (age three years) was linked to less internalizing behavior; and increased relative abundance of an unidentified *Enterobacteriaceae* genus (age two weeks) was related to more externalizing behavior. Our findings provide important longitudinal evidence that early-life gut microbiota may be linked to behavioral and cognitive development in low-risk children.

Keywords: gut microbiota; early life; problem behavior; cognition; executive functions

(Received 6 March 2023; revised 17 October 2023; accepted 20 October 2023)

Introduction

The human gut harbors a great number of microorganisms, of which bacteria are an essential part. These microorganisms are collectively termed the 'gut microbiota' (Thursby & Juge, 2017). Not only has the gut microbiota been involved in many health outcomes, such as obesity, type 2 diabetes, and irritable bowel syndrome (Vos et al., 2022), but it has also been linked to mental health (Cryan et al., 2019). Accumulating evidence from animal and adult human studies has uncovered several key bidirectional communication pathways between the gut microbiota and brain functioning, named the microbiota-gut-brain axis (MGBA) (Cryan et al., 2019). Remarkably, the MGBA is not only functional in adults but starts playing an equally or even more important role at early ages with regard to child behavior and cognition (Cryan

et al., 2019). Both the gut microbiota and the brain develop at a rapid pace during early life, however, only few studies investigated associations between the gut microbiota and behavior in such sensitive periods. Therefore, this study aimed to investigate the relations of the gut microbiota in the first three years of life with child problem behavior and executive functions at the age of three years.

The bidirectional interactions of the MGBA occur through intricately innervated and highly adaptable neuronal pathways, and extremely delicate and difficult-to-measure molecular communication systems (Cryan et al., 2019; de Weerth, 2017). For instance, short-chain fatty acids (SCFAs), mainly produced through dietary fiber fermentation by the gut microbiota, likely affect the brain via the vagus nerve, immunity, and the endocrine system (Dalile et al., 2019). Furthermore, specific microbial taxa can generate γ-aminobutyric acid (GABA), which is the main inhibitory neurotransmitter of the central nervous system and regulates many physiological functions (Mazzoli & Pessione, 2016; Silva et al., 2020). The symporter that mediates the uptake of microbiota-derived GABA is present through the gastrointestinal tract, suggesting that luminal GABA is able to cross the gut barrier and influence extra-gut targets. Although remaining controversial,

recent studies suggest the permeability of the blood-brain barrier to GABA, implying its direct impact on the central nervous system (Mazzoli & Pessione, 2016). Besides, GABA receptors are widely expressed in enteric neurons and immune cells, indicating the role of GABA in regulating the gut-to-brain signaling and neuro-inflammation (Auteri et al., 2015; Hyland & Cryan, 2010). Such pathways along the MGBA may partially explain how the gut microbiota impacts mental health.

The colonization of the gut by microorganisms mostly commences soon after birth and continues in the following years. The general consensus is that the gut microbiota develops into an adult-like configuration around the age of three years (Derrien et al., 2019), while some studies suggested that this step toward maturation may take longer than previously thought (Ou et al., 2022, 2023). Gut microbial disturbances during the early dynamic and sensitive colonization period can result in subsequent health problems, such as developing allergies and obesity (Zhuang et al., 2019). This is explained by the early-life programing theory, which refers to long-lasting changes and disruptions as a consequence of environmental exposures at a young age (Tarry-Adkins & Ozanne, 2011). In early life, the brain experiences numerous quick developments in neuronal proliferation, migration, differentiation, synaptogenesis, myelination, and apoptosis (Rice & Barone, 2000), largely impacting brain functioning, cognition, and behavior (Erus et al., 2015). Simultaneously, the microbiota is becoming established in the gut of infants and young children (de Weerth, 2017; Wang et al., 2018). Thus, alterations of the gut microbiota in early life may exert considerable effects on the development of the brain. Indeed, there is compelling evidence from animal studies supporting such a hypothesis (Clarke et al., 2014; Leclercq et al., 2017; O'Mahony et al., 2014; Stilling et al., 2015). This marks early life as a sensitive time window to obtain and maintain microbiota composition that will promote normal physical and mental development. However, we know little about early-life gut microbiota in association with child behavior and cognition. Specifically, how the gut microbiota and brain functioning, in particular host behavior, are interconnected in low-risk community infants and children (i.e., generally healthy and neurotypically developing) is underexplored. Knowledge of these associations, particularly when uncovered by comprehensive longitudinal studies, can provide insight into the typical early development of the gut microbiota in relation to child behavior and cognition.

First studies have found evidence for associations between the gut microbiota and child behavior and cognition. Regarding behavior, Loughman et al., reported that increased relative abundances of taxa belonging to the genus *Prevotella* at one year of age were associated with less internalizing behavior at age two (i.e., problem behavior affecting internal psychological conditions, characterized by withdrawal, anxiety, and emotional problems (T. M. Achenbach, 1966)) (Loughman et al., 2020). In our previous study, we found that the rise of *Prevotella* 9 in middle childhood was related to more externalizing behavior at age ten (i.e., problem behavior exhibited in the external environment, including features like impulsivity, aggression, and hyperactivity (T. M. Achenbach, 1966)) (Ou et al., 2022). Besides, Laue et al., observed a negative relation between *Streptococcus peroris* and internalizing behavior in girls before school age, and a positive association between *Lachnospiraceae* species and externalizing behavior in both genders (Laue et al., 2021). Furthermore, *Lachnospiraceae* species and *Veillonella* were linked to more internalizing behavior in preschoolers; interestingly, *Veillonella* was positively related to

externalizing behavior as well (Van De Wouw et al., 2022). Additionally, increased alpha diversity was observed in preschool children with less internalizing (Laue et al., 2021; Van De Wouw et al., 2022).

Four other studies have found an underlying link between infant gut microbiota and child cognition (Aatsinki et al., 2020; Carlson et al., 2017; Rothenberg et al., 2021; Streit et al., 2021; Tamana et al., 2021). Cognition is fundamental for the development of executive functions, including higher-level cognitive processes like inhibitory control (Diamond, 2013). Specifically, a cross-sectional study found more *Enterobacteriaceae* species in relation to worse cognition at age 45 months (Streit et al., 2021). Longitudinal research reflected that high relative abundances of *Bacteroides* at age one year were related to better cognition at age two (Carlson et al., 2017; Tamana et al., 2021). Furthermore, *Faecalibacterium* at one year of age was associated with reduced cognitive functions at age two (Carlson et al., 2017). Additionally, a lower relative abundance of *Bifidobacterium* and a higher relative abundance of *Clostridium* at two-and-a-half months were linked to increased attention at eight months (Aatsinki et al., 2020). Moreover, Rothenberg et al., found that children with better cognition showed enriched *Faecalibacterium*, *Sutterella,* and *Clostridium* cluster XIVa at age three years. Finally, high alpha diversity at age one year was reported in two-year-old children with worse cognition (Carlson et al., 2017).

To conclude, a number of associations have been observed between the gut microbiota and problem behavior and cognition in early life, but findings are variable and inconsistent across studies, mainly due to different methodologies used regarding microbiota analyses, genomics, epidemiology, and statistics. Furthermore, most of the previous studies have assessed problem behavior and cognition by using only one questionnaire of a single reporter. In the current longitudinal study in a community sample of children, we investigated the gut microbiota in relation to problem behavior (i.e., internalizing and externalizing behavior) and executive functions (i.e., advanced cognitive abilities, including inhibitory control (Diamond, 2013)) using questionnaires of multiple reporters and behavioral tasks. We had the following two hypotheses: (1) relative abundances and alpha diversity (i.e., Chao1, Shannon, and phylogenetic diversity) of the gut microbiota at age three years are associated with reported problem behavior and executive functions at the same age; (2) relative abundances and alpha diversity of the gut microbiota at early ages (i.e., two, six, and 12 weeks, and one year) are associated with reported problem behavior and executive functions at age three.

We investigated these hypotheses in three ways: (1) as the gut microbiota is highly dynamic in early life, its composition at different ages may be differently associated with problem behavior and executive functions later in life. For this reason, we analyzed the overall gut microbiota composition in relation to preschool-aged cognitive measures in an age-specific manner; (2) for the same reason, relations regarding a single taxon and an alpha diversity index were analyzed in an age-specific manner; (3) based on the age-specific analyses, we explored the trajectories of taxa and alpha diversity parameters in association with mental outcomes over the whole study period. Figure 1 shows the workflow of our analyses. Considering that most published findings were at the genus level, we performed our analyses at the same taxonomic level. However, given the paucity of studies on these relations at such early ages, we did not hypothesize specific associations between microbial taxa and mental outcomes.
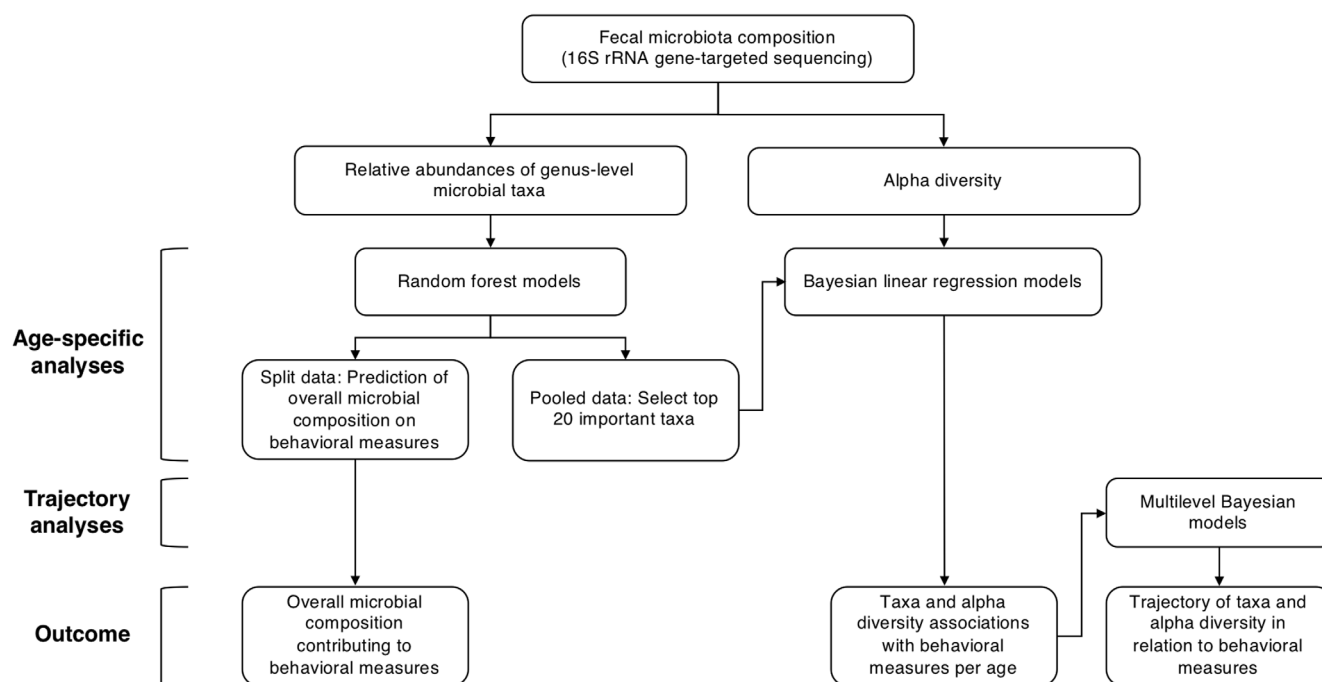
Figure 1. Workflow of the analyses.

Materials and methods

Participants

The current study is part of the longitudinal Dutch study named *BINGO* where early factors associated with child development were investigated. Participants were healthy children and their parents living in the Netherlands. Detailed in- and exclusion criteria are described in a previous publication (Hechler et al., 2018). At baseline, 88 pregnant women were recruited in the Arnhem-Nijmegen region for the *BINGO* study; 96% were born in the Netherlands. Postnatal exclusion criteria included: complications during pregnancy, gestational age at birth < 37 weeks, birth weight < 2500 g, 5-min Apgar score < 7, and congenital malformations. Seventy-seven mothers were followed up after postnatal exclusion. At three years of child age, 76 families were approached (one drop-out occurred during the previous measurement rounds). Among them, two families could not be contacted, six families did not participate due to time constraints, and one family dropped out due to personal reasons. Parental demographics did not differ significantly between participating and nonparticipating families. This resulted in a final sample of 67 families. Of them, 64 families participated in home visits when their children reached age three, and the other three families were unable to join home visits but filled out questionnaires in this assessment round. Both parents participated in 54 families (81%, 54/67), and only mothers participated in 13 families (19%, 13/67).

Ethics

The *BINGO* study was independently reviewed by the Ethics Committee of Social Sciences of Radboud University, and no formal objection to this research was made [ECSW2014–1003–189 and amendment: ECSW–2018–034]. The current study was preregistered on the Open Science Framework: https://osf.io/vwgef with amendment: https://osf.io/nyeb4.

Data collection procedure

Collection of child stool samples was done at the ages of two, six, and 12 weeks, and one and three years. Stool samples were stored in the participant's freezer (−20°C) until they were collected with a portable freezer. The stool samples were stored at -80°C at Radboud University prior to being processed at Laboratory of Microbiology at Wageningen University & Research.

Home visits took place when the child turned three years old. Prior to the home visit, mothers and their partners independently filled in digital questionnaires about their child's problem behavior and executive functions. During the home visit, the child performed inhibitory control tasks. Tasks were video recorded and afterward rated by two trained observers. Observers were trained by use of a coding manual specific to each task.

Measures

Gut microbiota composition

Stool samples were collected with a polystyrene 10 mL stool container. Total DNA was extracted from 0.01 to 0.15 g of stool sample with 300 μL of Stool Transport and Recovery Buffer by double bead-beating steps as previously described (Gu et al., 2018). The variable V4 region of prokaryotic 16S ribosomal RNA (rRNA) genes was then amplified by PCR in duplicate reactions, by using primers 515F-n (5'-GTGCCAGCMGCCGCGGTAA) and 806R-n (5'-GGACTACHVGGGTWTCTAAT) (Gu et al., 2018). The 16S rRNA gene sequencing was completed on the Illumina HiSeq platform by Eurofins Genomics Germany GmbH.

Behavioral measures
Parental questionnaires. Mothers and their partners filled in all questionnaires mentioned below. However, because fewer partners completed the questionnaires, we used partner reports for sensitivity analyses to validate the maternal reports by calculating Kendall correlations between both. The non-

parametric Kendall method was chosen due to its better performance in handling non-normally distributed data and tied values (Kendall, 1945). Maternal reports were used as the final measure of reported problem behavior and executive functions.

To assess child problem behavior, the Child Behavior Checklist for ages of one and a half to five years (CBCL, 103 items) (Thomas M. Achenbach & Ruffle, 2000) and the Strengths and Difficulties questionnaire (SDQ, 25 items) (Goodman, 1997) were used. The CBCL and the SDQ include internalizing and externalizing subscales, consisting of items scored on a three-point Likert scale. The SDQ can detect problem behavior as accurately as the CBCL does (Goodman & Scott, 1999). Raw scores for both questionnaires were used as outcome measure in order to compare and possibly aggregate the measures. However, given that the Kendall correlations on the same subscales of the CBCL and the SDQ were lower than 0.5 (Table S1), we included both instruments separately in the analyses. In both instruments, higher scores on subscales indicate more problem behavior.

To evaluate child executive functions, the Behavior Rating Inventory of Executive Function – preschool version (BRIEF-P, 63 items) questionnaire for preschoolers (Sherman & Brooks, 2010) and the Ratings of Everyday Executive functions (REEF, 77 items) (Nilsen et al., 2017) were used. The BRIEF-P and the REEF are scored on three- and four-point scales, respectively. A higher score on the BRIEF-P indicates worse executive functions, while a higher score on the REEF indicates better executive functions. The BRIEF-P is a commonly used questionnaire that measures general executive functions and does not differentiate between different situations. The REEF rates executive functions in different situations (e.g., executive functions around friends, during grocery shopping, or in the community) and determines an average score. Raw scores for both questionnaires were used as outcome measure in order to compare and possibly aggregate the measures. However, Kendall correlations between the BRIEF-P and the REEF were lower than 0.5 (Table S1), hence both instruments were included in the analyses.

Parental questionnaires were considered acceptable and reliable based on their ω_{total} (ranging between 0.65 and 0.94) or Cronbach's α values (ranging between 0.83 and 0.96) (Table S2) (Revelle & Condon, 2019).

Inhibitory control tasks. Six different behavioral tasks with good reliability (i.e., Flanker, Whisper, Gift Wrap, Gift Delay, Snack Delay, and Bear Dragon) were performed to measure inhibitory control as previously stated in detail (Willemsen et al., 2021). Observer reliability was determined by the Intraclass Correlation Coefficient (ICC) relying on absolute agreement. The ICC's for the inhibitory control tasks ranged from $r = 0.84$ to $r = 0.96$ ($p < 0.001$). Snack Delay and Bear Dragon were excluded from the analyses due to insufficient variation and low number of children that passed the practice trials, respectively. The other four tasks were included in our study. Higher scores on these tasks indicate better inhibitory control.

Statistical analyses

Pre-processing of sequence data
Sequence data were processed via *NG-Tax* 2.0 with default settings (Poncheewin et al., 2020; Ramiro-Garcia et al., 2018), with SILVA SSU 16S rRNA gene reference database (version 132) (Quast et al., 2012). The raw amplicon sequence variant (ASV) count data were

used to calculate alpha diversity by the *ape* (Paradis, 2020) and the *picante* (Kembel, 2020) packages. Then, ASV count data were glommed at the genus level prior to analyses.

Gut microbiota composition and development over the first three years of life
For descriptive purposes, we first delineated gut microbiota composition and development in the first three years after birth (including all samples at the age of two, six, and 12 weeks, and one and three years). We compared differences in alpha diversity indices, including Chao1, Shannon, and phylogenetic diversity, between ages using Wilcoxon rank-sum tests corrected with the False Discovery Rate (FDR) method. Next, we also compared beta diversity between ages by conducting Principal Coordinate Analysis (PCoA) via the *vegan* package (Oksanen et al., 2020). Considering that PCoA can be applied to different dissimilarity and distance metrics that all differ in specific aspects and corresponding interpretation, we included the Bray-Curtis, weighted Jaccard (formula = 2*Bray-Curtis dissimilarity / (1 + Bray-Curtis dissimilarity)), unweighted UniFrac, weighted UniFrac, and Aitchison (the Euclidean distance based on centered-log-transformed ASV count data) methods, to comprehensively describe the compositional differences. Except for the Aitchison distance, we transformed genus-level count data into relative abundances before calculating other dissimilarity and distance metrices. Significance was determined as a p value lower than 0.05 for non-multiple tests and an FDR-adjusted p value lower than 0.05 for multiple tests.

Additionally, we visualized average and individual relative abundances at the genus level over the study period by using a bar plot and a heatmap, respectively. To identify differentially abundant microbial taxa at the genus level between ages, we conducted the Linear Discriminant Analysis Effect Size (LEfSe) method by using the *microbiomeMarker* R package (Segata et al., 2011), with a log-transformed Linear Discriminant Analysis (LDA) score higher than two indicating significance.

Confounding effects
In our original preregistration, we considered child age and diet quality as potential confounders (i.e., variables that influence both the independent variables and the outcome). After reconsideration, both variables were removed as potential confounders due to two major reasons (amendment can be found via https://osf.io/nyeb4): (1) low variation in child age (see Figure 2 for notes regarding ages); (2) our previous study using the same cohort found no significant associations of diet quality with behavior and executive functions (Willemsen et al., 2021). Given these considerations, no confounders were accounted for in the models performed in this study. Note that potential covariates of the independent variables only (i.e., the gut microbiota) were not accounted for in downstream analyses (Cinelli et al., 2020), as they would remove variation in the gut microbiota data, which was not the purpose of this study. These potential confounders and covariates as well as their relations to the gut microbiota and behavioral outcomes are displayed in a directed acyclic graph (Figure S1).

Data imputation and transformation
Missing values (proportion of missing values is shown in Table S3) in problem behavior, executive functions, and inhibitory control were imputed ten times together, by using the predictive mean match method in the R package *mice* (Buuren, 2021). The
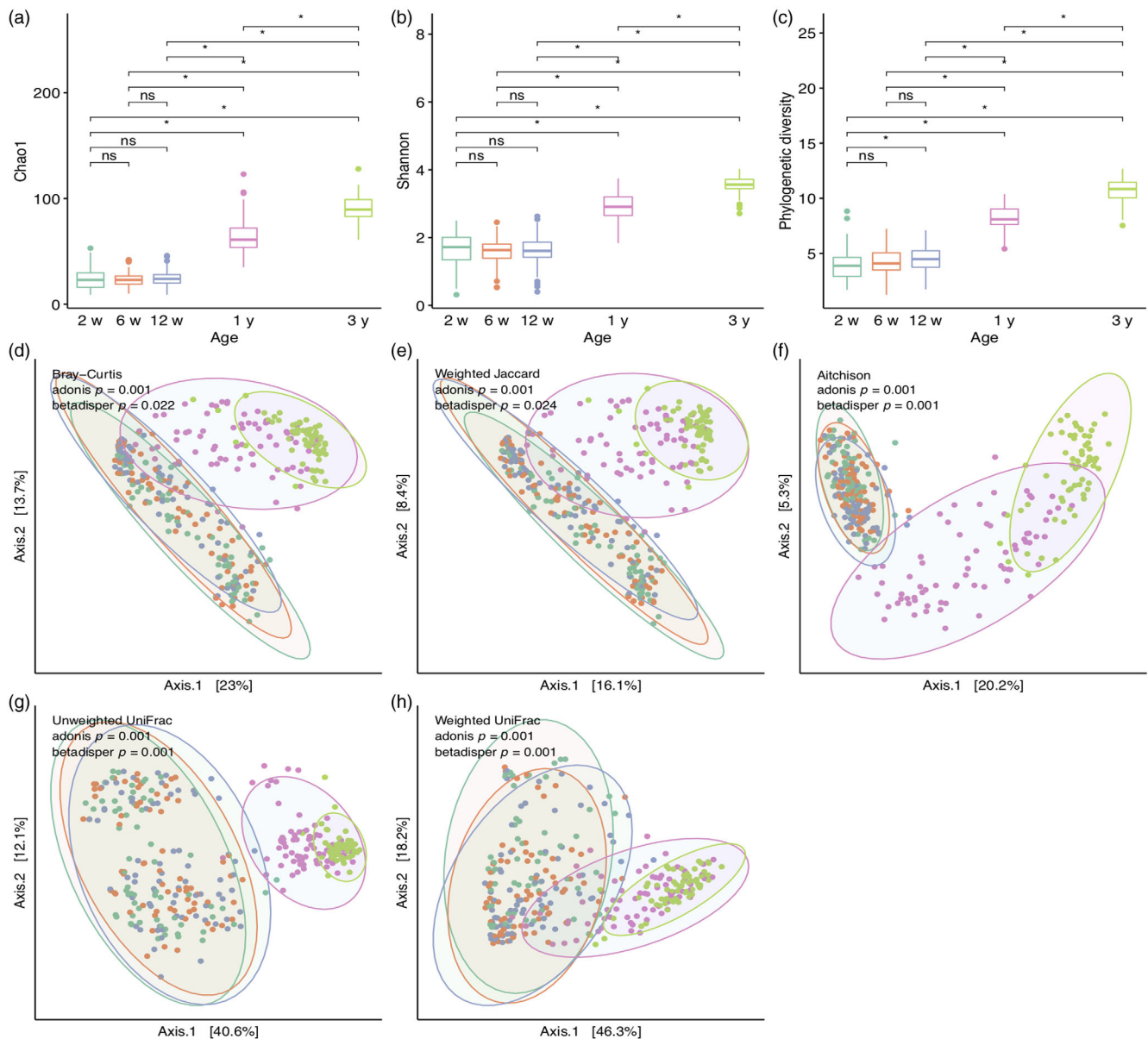
Figure 2. Alpha and beta diversity of the gut microbiota in the first three years of life. (a–c) alpha diversity as measured by Chao1, Shannon, and phylogenetic diversity indices. Wilcoxon rank-sum tests were conducted between ages and corrected with the FDR method (ns, not significant; *, <0.01). Age2w_mean±sd = 2.08 ± 0.28. Age6w_mean ±sd = 6.23 ± 0.55. Age12w_mean±sd = 12.27 ± 0.42. Age1y_mean±sd = 1.04 ± 0.08. Age3y_mean±sd = 3.18 ± 0.10. (d–h) principal coordinate plots of beta diversity, based on different pairwise dissimilarity (Bray-curtis and weighted Jaccard) and distance (UniFrac and Aitchison) matrices, with points and ellipses colored by ages (Lake blue, two weeks; orange, six weeks; purple, 12 weeks; Pink, one year; grass green, three years).

imputation model was conducted separately at each age. For instance, at the age of three, 64 children provided gut microbiota data, and their missing values in aforementioned behavioral measures were imputed jointly in one model. No auxiliary variables (i.e., variables that are not included in analyses, but are correlated with imputed variables) were considered in the imputation.

For both random forest models and the Bayesian linear regression models, genus-level relative abundance data were used. Numeric variables were standardized (i.e., subtracting the mean and dividing by the standard deviation) for the Bayesian models only, as random forest models rely on decision trees for which standardization is considered unnecessary.

Main analyses

To determine whether gut microbiota composition in the first three years of life (i.e., two, six, and 12 weeks, and one and three years) is associated with problem behavior (i.e., internalizing and externalizing behavior) and executive functions (including inhibitory control) at age three, we conducted random forest models and the Bayesian linear regression models (Bürkner, 2017; Kuhn et al., 2020). Random forest is first of all well suited to analyze microbiome data as it is appropriate for high-dimensional data, invariant to scaling of inputs, computationally efficient, and able to uncover nonlinear relationships (Belk et al., 2018; Louppe, 2014; Namkung, 2020). The first random forest model was applied to assess the contribution of the total gut microbiota composition on

our behavioral outcomes. This was done for the purpose of exploring the gut microbial community as a whole to account for the complex interplay between taxa. The following random forest model was applied as a preselection tool, to select possibly important taxa from high-dimensional data, before passing them on to the Bayesian linear regression model. The Bayesian model was first used to determine age-specific relations (i.e., directions and strengths) of a selected taxon and an alpha diversity index with each outcome measure. By looking at the different time points separately, these analyses can help identify periods of development that are sensitive to certain microbial compositions. Although not preregistered, after reconsideration, we decided to perform an additional analysis to optimize the use of our longitudinal data. Based on the age-specific results, we implemented a multilevel Bayesian model to determine whether trajectories of change in the gut microbiota were associated with the outcome measures at age three. Figure 1 shows the workflow of our analyses.

Age-specific analyses - Determining the contribution of the overall gut microbiota to each behavioral measure through random forest models. Data were imputed ten times: data were randomly split into a training dataset (including 80% participants) and a testing dataset (including 20% participants), leading to ten training datasets and ten corresponding testing datasets. The procedure of data splitting was applied to children who provided gut microbiota information at each age separately. To prevent data leakage, the missing values of behavioral measures were imputed in training and testing datasets separately (ten times) as described earlier. Next, we included genus-level relative abundances of overall gut microbiota as independent variables and one behavioral measure as an outcome. This step was performed on each individual behavioral measure separately. To train the model, a ten-repeated ten-fold cross-validation was conducted on each complete training dataset including imputed values via the *caret* package (Kuhn et al., 2020). Afterward, we used the trained model to obtain predicted behavioral outcomes of each corresponding complete testing dataset including imputed values. Similarity between predicted and actual behavioral outcomes of the complete testing dataset was measured by the Pearson correlation with its p value obtained from a permutation test ($N = 1000$). Considering that data splitting and imputation resulted in multiple datasets, we used the median value of the Pearson correlation coefficient from multiple cases to represent the final similarity. The p value corresponding to this median was included. P values were adjusted with FDR methods, with corrected values under 0.05 indicating significance.

Age-specific analyses - Preselecting potentially important gut microbiota contributing to each behavioral measure through random forest models. To identify microbial taxa that contribute to each behavioral outcome, we measured the change in the generalized cross-validation (GCV) value in the random forest model. Larger GCV changes indicate more contribution of the independent variable to the model, in other words, this analysis shows which taxa are potentially more important (Kuhn et al., 2020). Unlike the first random forest model, we did not split the data but used the whole dataset here, because we prioritized the structure of the model and a large sample size can provide more valid information. Missing values of behavioral measures in the whole dataset were imputed as described in the section on data imputation and transformation. Then, relative abundances of all taxa were treated as independent variables with one behavioral

measure as an outcome. This procedure was performed on each behavioral measure separately. Next, we carried out a ten-repeated ten-fold cross-validation on each complete dataset containing imputed data and calculated average GCV values of multiple datasets acquired from data imputation. Based on the size of average GCV values, we selected the largest 20 taxa as the top 20 in importance. These 20 taxa were then passed to the Bayesian linear regression models to confirm their actual associations with behavioral measures.

Age-specific analyses – Associating the gut microbiota with behavioral measures by using Bayesian linear regression models. We implemented Bayesian linear regression models to estimate the relations of relative abundances of the selected top 20 microbial taxa with a prevalence value higher than 10% and alpha diversity with the child behavioral measures. Compared to standard linear regression models, the Bayesian linear regression models compute the probability of different effects rather than simply reporting single estimates of the "true effect" (Bürkner, 2018). We performed the Bayesian models by using the *brms* R package built based on the programing language *Stan* (Bürkner, 2017). The *brm* function within the *brms* package was used with the Gaussian distribution (mean = 0, std = 1) as the prior distribution for all beta coefficients and the Student's t-distribution for error distribution (due to better performance in handling extreme values) (Lange et al., 1989). A list containing multiple complete datasets including imputed data was directly passed to the *brm* function, which in turn generated a single estimate. Other arguments of the *brm* function were set as follows: chains = 4, iter = 2000, and warmup = 1000. Under these settings, chains converged properly with Rhat values lower than 1.01. Regarding the outcomes of the Bayesian models, the less the posterior distribution overlaps with zero, the more likely a relation is positive or negative. In the current study, we defined a relation as positive or negative with confidence when its 95% credible interval (CI) excluded the value zero.

Trajectory analyses – Relating the developmental trajectories of the gut microbiota to behavioral measures through multilevel Bayesian linear regression models. To make maximum use of our longitudinal data, we conducted multilevel models to investigate relations between the developmental trajectories of the gut microbiota and behavioral measures. The multilevel models were performed on microbial taxa and alpha diversity with confident age-specific relations to behavioral measures (i.e., as determined by the Bayesian linear regressions described above). In the multilevel models, microbial and behavioral information as well as the actual age were level 1 variables, and the child was the level 2 variable. Note that missing values in behavioral measures and actual age were not imputed and that in these analyses we used the same distributions and arguments as described earlier. Before performing a testing model, we first checked the intraclass correlation (ICC) of an intercept-only model. When the 95% CI of an ICC excluded the value zero in the intercept-only model, multilevel strategies were used. A trajectory relation was considered with confidence when there was no overlap between its 95% CI and zero.

With respect to taxa, when their prevalence was higher than 10% at five-time points (i.e., two, six, and 12 weeks, and one and three years), multilevel models were performed on the pooled data of all ages. When only the first three time points met the 10% criteria, multilevel models were carried out by pooling samples at these three ages together. When only the prevalences at the last two

ages were higher than 10%, multilevel approaches were done in the pooled aged-one-and-three years samples. Rhat values were used to check chain convergence.

Results

Population characteristics and descriptives

Demographic data and descriptives of study variables are shown in Table 1. Roughly 50% of the children were girls. Mothers were mostly highly educated (86.2%). Scores on the questionnaires measuring child problem behavior and executive functions did not differ significantly between mothers and their partners, and they were significantly positively intercorrelated (Table S1).

Gut microbiota composition and development over the first three years of life

We analyzed microbial composition of 345 fecal samples taken at five-time points. A total of 42,056,591 high-quality reads were obtained after being processed with *NG-Tax* 2.0. Within these reads, 220 microbial taxa were identified at the genus level mainly belonging to the phyla *Firmicutes*, *Actinobacteria*, *Bacteroidetes*, *Proteobacteria*, and *Verrucomicrobia*.

For descriptive purposes, we compared alpha and beta diversity between ages (Fig. 2; beta-diversity comparisons between the first three ages and between the last two ages are displayed in Figures S2 and S3) and delineated a general developmental trajectory of the gut microbiota over time (Fig. 3a). Diversity comparisons reflected profound compositional differences between infancy and preschool age. These differences were visualized by the heatmap showing individual relative abundance data (Figure S4). LEfSe identified a total of 106 differentially abundant microbial taxa between ages (log-transformed LDA scores higher than two; Table S4). Due to the large number of significant taxa, only the taxa with log-transformed LDA scores higher than four are highlighted and displayed in Figure 3b, such as an unidentified genus within *Enterobacteriaceae*, *Lactobacillus*, *Bifidobacterium*, *Faecalibacterium*, and *Blautia*.

Age-specific analyses

Determining the contribution of the overall gut microbiota to each behavioral measure through random forest models

To explore whether the overall microbial composition in the first three years (i.e., at ages two, six, and 12 weeks, and one and three years) contributes to problem behavior and executive functions at age three, we compared the similarity between the actual and the predicted behavioral results. As shown in Table S5, 92% (46/50) of the models showed insignificant absolute correlation coefficients (i.e., lower than 0.3), indicating a low likelihood that the gut microbiota can contribute to behavioral outcomes. Regarding the 8% (4/50) models with correlation coefficients higher than 0.3, the similarity remained insignificant between the actual and predicted data, implying the same low likelihood. The random forest models showed that the overall gut microbiota did not contribute to problem behavior and executive functions in the present study.

Preselecting potentially important gut microbiota contributing to each behavioral measure through random forest models

As planned in our preregistration, we preselected the microbial taxa that may contribute to the behavioral outcomes the most (i.e., top important taxa) based on GCV values, by performing separate random forest models at each age. The top 20 important taxa at the genus level are depicted in Figures 4 and 5, with the following observations:

(1) *Bacteroides* and *Clostridium* sensu stricto 1 were the most frequent contributors to CBCL internalizing behavior;
(2) *Bacteroides* and *Bifidobacterium* were the most frequent contributors to CBCL externalizing behavior, SDQ internalizing and externalizing behavior, and BRIEF-P executive functions;
(3) *Bacteroides* and *Blautia* were the most frequent contributors to REEF executive functions;
(4) Additionally, *Bacteroides* and *Bifidobacterium* were the most frequent contributors to the behavioral measures of inhibitory control (i.e. Flanker, Whisper, Gift Wrap, and Gift Delay) (Fig. 6).

Associating the gut microbiota with behavioral measures by using Bayesian linear regression models

To confirm whether the aforementioned top 20 important microbial taxa were associated with problem behavior and executive functions, we performed the Bayesian linear regression model on each genus-level taxa (relative abundance) and behavior pair. Table 2 shows the strongest observed associations of these pairs (i.e. estimates higher than 0.2 or lower than −0.2).

Remarkably, there were several highly present taxa (i.e. prevalent in more than 80% of the samples, relative abundance higher than 10%) in relation to the outcome measures: *Bifidobacterium* at age three years was associated with more internalizing and externalizing behavior (est. = 0.27 for both), *Blautia* at three years was linked to less internalizing behavior (est. = −0.25), and an unidentified taxa within the *Enterobacteriaceae* family was related to more externalizing behavior (est. = 0.25).

Next, we checked for consensus between the questionnaires assessing the same construct. For internalizing behavior, there was no consensus between the associations found for the CBCL and the SDQ. For externalizing behavior, more *Parabacteroides* at two weeks were associated with less externalizing behavior in both the CBCL (est. = −0.30) and the SDQ (est. = −0.28). An opposite finding was found for *Butyricicoccus* at one year in relation to more externalizing behavior by the CBCL (est. = 0.23), while at three years, it was associated with less externalizing behavior by the SDQ (est. = −0.35).

Within the CBCL results, *Barnesiella* at age three years was associated with more internalizing (est. = 0.31) and externalizing behavior (est. = 0.33). Within the SDQ results, *Bifidobacterium* at age three years was associated with more internalizing and externalizing behavior (est. = 0.27 for both).

Regarding executive functions, *Ruminococcus* 2 at one year and [*Ruminococcus*] *torques* group at age three years, were associated with better executive functions as measured by the BRIEF-P (est. = −0.30, note that higher scores on the BRIEF-P indicate worse executive functions) and worse executive functions measured by the REEF (est. = −0.24), respectively. Lastly, *Halomonas* at six weeks was associated with worse executive functions as measured by the BRIEF-P (est. = 0.24) and the REEF (est. = −0.24).

Different associations were found for the Flanker and the Gift Delay tasks. For the Flanker, relations were identified at the age of six weeks, and one and three years, while for the Gift Wrap, associations were observed at age one year only. This may be due to a highly dynamic gut microbiota ecosystem in early life, of which composition at different ages may be variously linked to executive functions.

Table 1. Descriptives of study subjects

Categorical variable	Characteristics		
	Ratio	Sample size	Completion rate %
Child sex	girl: boy = 34:30	64	100
Delivery mode	vaginal: C-section = 54:7	61	95
Antibiotic treatment	no: yes = 60:3	63	98
Educational level (%)	low: middle: high = 0:12.5:87.5	64	100

Numeric variable	Mean ± SD	Minimum	Lower quartile	Median	Upper quartile	Maximum	Sample size	Completion rate %
Age at age three in years	3.2 ± 0.1	3.1	3.1	3.2	3.2	3.5	63	98
Gestational age in weeks	39.8 ± 1.5	35.6	38.9	40	40.9	42.1	63	98
Birth weight in grams	3556 ± 426.2	2570	3270	3480	3885	4445	61	95
Total breastfeeding duration in months	9.6 ± 8.1	0	4	8	13.2	36	64	100
Total exclusive breastfeeding duration in months	3.9 ± 1.7	0	3	4	5	7	53	82
Age at solid food introduction in months	4.6 ± 1	3	4	4	5	7	59	92
Average diet quality at age three	4 ± 1.2	2	3.3	3.9	4.8	7.2	64	100
CBCL_M_Internalizing	7.1 ± 5.7	0	3	5.5	10.8	24	62	97
CBCL_M_Externalizing	11.8 ± 7.3	0	7	12	15.8	31	62	97
SDQ_M_Internalizing	3.6 ± 2.5	0	2	3	5	11	62	97
SDQ_M_Externalizing	5.5 ± 3	1	3.2	5	7	14	62	97
BRIEF-P_M_TotalScore	94.4 ± 15.4	69	83	92	106.5	146	63	98
REEF_M_TotalScore	151.1 ± 31	74	133.2	153	172.8	215	62	97
Flanker	1.6 ± 0.3	0.9	1.4	1.7	1.9	2	47	73
Whisper	1.9 ± 0.3	0.9	1.8	2	2	2	60	94
Gift Wrap	2.2 ± 0.9	0	1.5	2.5	3	3	60	94
Gift Delay	3.9 ± 0.2	2.9	3.9	4	4	4	61	95
CBCL_P_Internalizing	7.5 ± 5.5	0	4	6	11	22	49	96
CBCL_P_Externalizing	12.2 ± 5.8	1	8	12	17	24	49	96
SDQ_P_Internalizing	3.5 ± 2.4	0	2	3	5	9	44	86
SDQ_P_Externalizing	5.4 ± 2.9	0	3	5	7	12	50	98
BRIEF-P_P_TotalScore	97.2 ± 17.8	69	86.2	96	109.2	137	50	98
REEF_P_TotalScore	147 ± 28.9	78	133	148	164	212	49	96

In the assessment round at age three, 64 children, 64 mothers, and 51 partners participated in the study. In total, 66, 70, 73, 72, and 64 fecal samples were collected at ages two, six, and 12 weeks, and one, and three years, respectively. Completion percentages were based on number of participating individuals (i.e., completion rates for mothers are based on 64 participating mothers, and completion rates for partners are based on 51 participating partners). P = partner; M = mother; CBCL = the child behavioral checklist; SDQ = the strengths and difficulties questionnaire; BRIEF-P = behavior rating inventory of executive functions – preschool; REEF = ratings of everyday executive functioning. Differences were compared between mother and partner reports by Wilcoxon tests. None of them were significant before or after FDR adjustments.

There were some overlapping associations between the questionnaires on problem behavior and executive functions. *Parabacteroides* at two weeks were associated with better executive functions (REEF, est. = 0.25), and less externalizing behavior (CBCL and SDQ, est. = −0.30 and −0.28, respectively). Another consistent result was *Streptococcus* at two weeks in relation to worse executive functions (BRIEF-P, est. = 0.40) and more externalizing behavior (CBCL, est. = 0.26).

We also measured behavioral relations to alpha diversity, including Chao 1, Shannon and phylogenetic diversity by using the Bayesian linear regression models (strongest results are displayed in Table 3). Interestingly, relations were only observed for alpha diversity at age two weeks. Higher Chao 1 values were associated with less internalizing behavior (CBCL) (est. = −0.28).

Furthermore, Chao1 values were in positive relation to better executive function performance (REEF and Gift Wrap, est. = 0.31 and 0.43, respectively). Lastly, higher phylogenetic diversity at age two weeks was also linked to better inhibitory control during the Gift Wrap task (est. = 0.32).

Trajectory analyses

Relating the developmental trajectories of the gut microbiota to behavioral measures through multilevel Bayesian linear regression models

Based on the results of age-specific Bayesian models and the 10% prevalence rule applied to microbial taxa (Table S6), we identified 16 pairs (including 12 pairs of taxa and behavioral
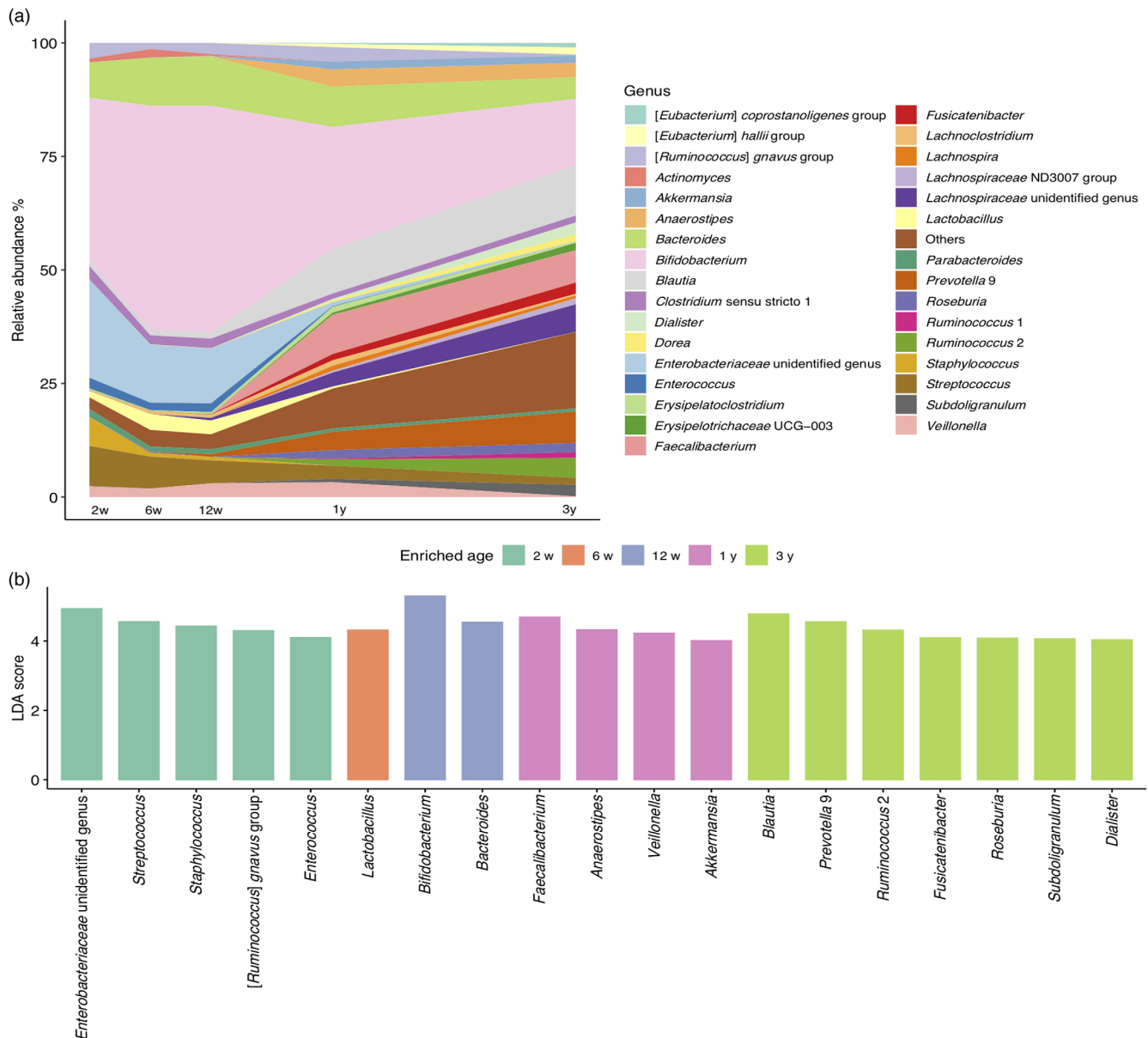
Figure 3. Characteristics of the gut microbiota in the first three years of life. (*a*) Average relative abundances of the gut microbiota at the genus level over time. Others represent genera with relative abundances lower than 1%. (*b*) Differentially abundant genus-level taxa between ages, identified by linear discriminant analysis effect size (LEfSe) with log-transformed linear discriminant analysis (LDA) scores higher than four.

measures, and four pairs of alpha diversity and behavioral measures) available at all five ages (i.e., two, six, and 12 weeks, and one and three years), three at the first three ages, and 12 at the last two ages (Table S7). Higher relative abundances of *Streptococcus* over the first three years of life were weakly related to worse executive functions reported by the BRIEF-P (est. = 0.05; higher scores on the BRIEF-P indicating worse performance), conforming to earlier age-specific findings. We also found that the trajectory of [*Ruminococcus*] *torques* group from age one to three was negatively related to internalizing behavior (SDQ, est. = −0.22), implying that higher relative abundances were associated with fewer internalizing difficulties during this period. No enduring associations were observed with confidence regarding alpha diversity.

Discussion

In this longitudinal study, we investigated associations of the gut microbiota during early life with problem behavior and executive functions, including inhibitory control, at three years of age. Several associations with behavior and cognition were found for relative abundances of microbial taxa and alpha diversity throughout the first three years of life. Table S8 provides an overview of the different microbiota taxa and microbial diversity index at different ages relative to the developmental findings. In addition, Table S8 shows the existing literature relative to our findings. Based on this table, we discuss the most prominent findings below.

We found evidence that increased relative abundance of *Streptococcus*, specifically at the age of two weeks and over the first

Figure 4. Heatmap showing the top 20 important microbial taxa over time and their associations to problem behavior at age three as reported by the mother. The top 20 important genus-level taxa within each age (i.e., 2w, two weeks; 6w, six weeks; 12w, 12 weeks; 1y, one year; 3y, three years) per behavioral measure are shown on the right side of the figure. Behavioral measures include: CBCL-I = internalizing behavior measured by the CBCL; SDQ-I = internalizing behavior measured by the SDQ; CBCL-E, externalizing behavior measured by the CBCL; SDQ-E, externalizing behavior measured by the SDQ. The orange scale indicates the importance of the taxa, with darker color referring to increased importance. The importance was determined by the generalized cross-validation value, with a larger value change indicating more contribution of a taxon to the model, i.e., which taxon is more important. As not all taxa appeared in the top 20 list at each time point, these absent taxa are colored in gray. Numbers on the left side of the figure show how many times a taxon appeared to be in the top 20 list of a behavioral measure over time. The frequently appearing taxa are bolded and colored in orange (five times), yellow (four times), or green (three times).
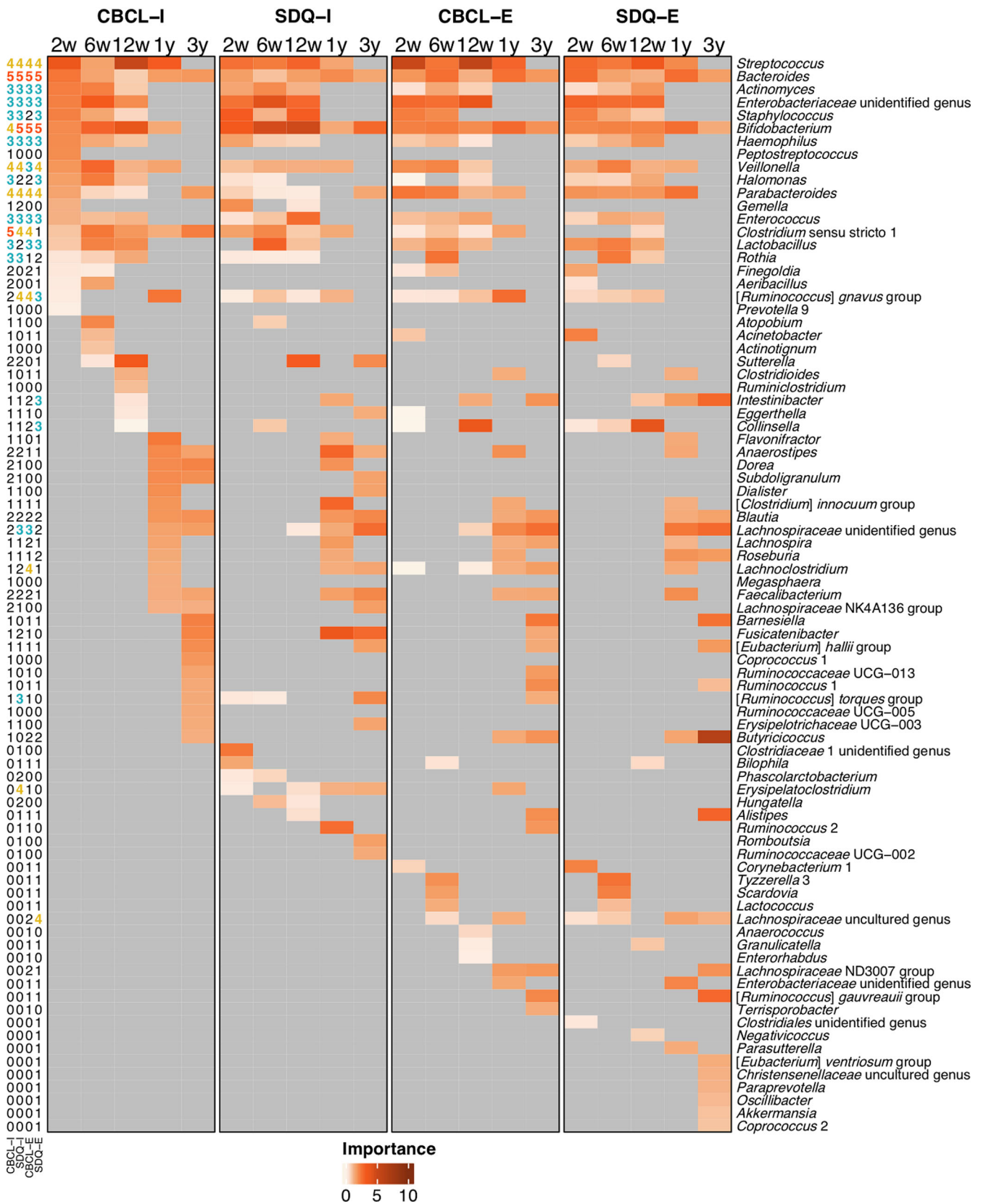
Figure 5. Heatmap showing the top 20 important microbial taxa over time and their associations to executive functions at age three as reported by the mother. The top 20 important genus-level taxa within each age (i.e., 2w, two weeks; 6w, six weeks; 12w, 12 weeks; 1y, one year; 3y, three years) per cognitive measure are shown on the right side of the figure. The measures include: BRIEF-P = executive functions measured by the BRIEF-P; REEF =, executive functions measured by the REEF. The orange scale indicates the importance of the taxa, with darker color referring to increased importance. The importance was determined by the generalized cross-validation value, with a larger value change indicating more contribution of a taxon to the model, i.e., which taxon is more important. As not all taxa appeared in the top 20 list at each time point, these absent taxa are colored in gray. Numbers on the left side of the figure show how many times a taxon appeared to be in the top 20 list of a measure over time. The frequently appearing taxa are bolded and colored in orange (five times), yellow (four times), or green (three times).

three years after birth, was associated with worse executive functions at the age of three years. This result indicates that *Streptococcus* might affect cognitive development throughout early life. Relations between early-life relative abundances of *Streptococcus* and behavior and cognition in typically developing children have not been observed in previous literature. For comparative purposes, we, therefore, examined studies on microbiota composition in children diagnosed with neurodevelopmental disorders as they mostly have comorbid behavioral and cognitive issues (Schoemaker et al., 2014). According to a systematic review, children with ASD (autism spectrum disorder) frequently show an

overgrowth of *Streptococcus* (Bundgaard-Nielsen et al., 2020). Although gut microbiota dysbiosis in ASD was seemingly partially attributed to an altered dietary pattern (Li et al., 2022; Welberg, 2022), diet was not correlated to the mental outcomes of our community samples. In addition to *Streptococcus*, both age-specific and trajectory relations were discerned for the [*Ruminococcus*] *torques* group: higher relative abundances at the age of three years and throughout the period from one to three years of age were associated with fewer child internalizing problems. Previous research showed excessive absolute abundances of fecal [*Ruminococcus*] *torques* group in children with ASD
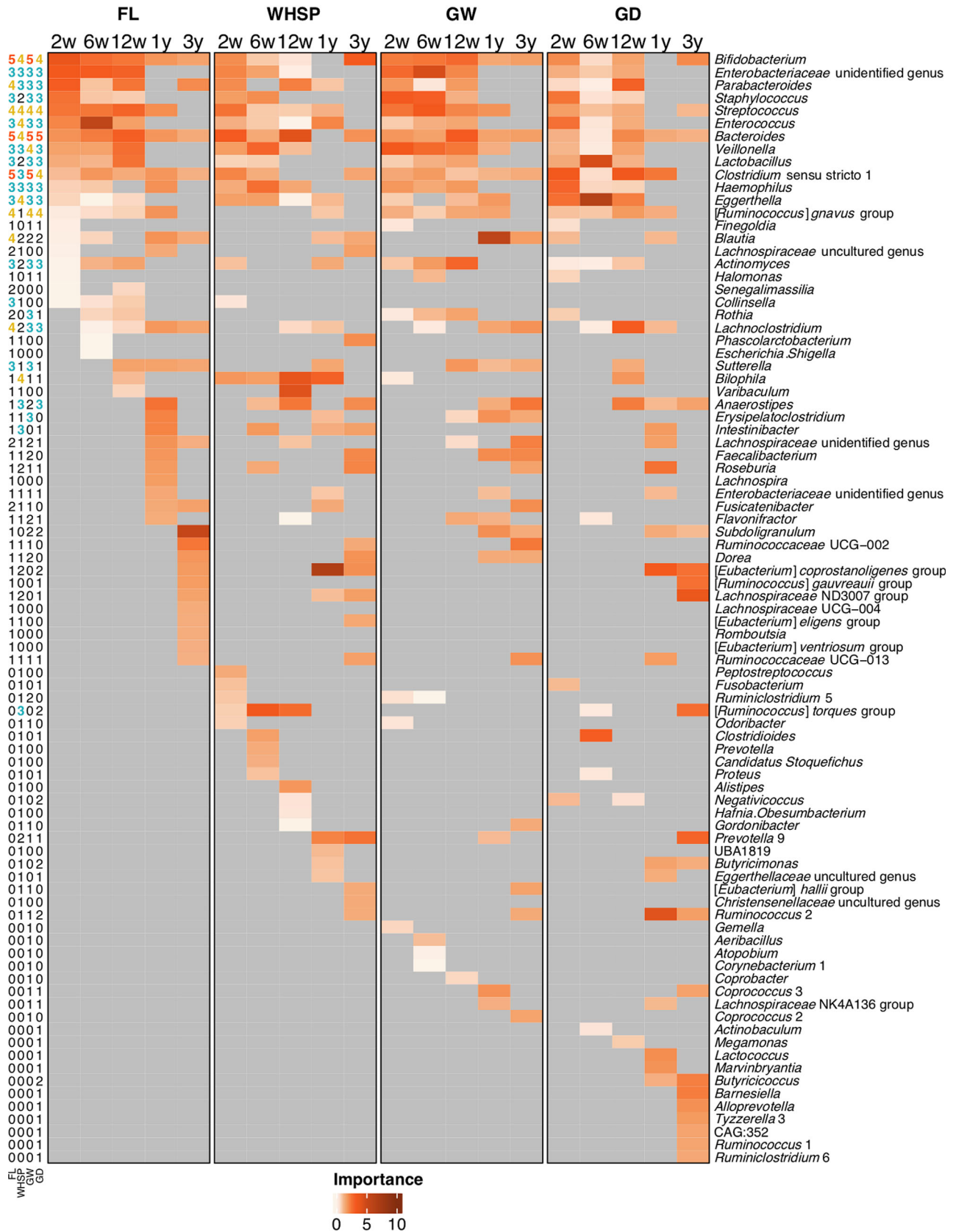
Figure 6. Heatmap showing the top 20 important microbial taxa over time and their associations to observed inhibitory control behavior at age three. The top 20 important genus-level taxa within each age (i.e., 2w, two weeks; 6w, six weeks; 12w, 12 weeks; 1y, one year; 3y, three years) per inhibitory control task are shown on the right side of the figure. The tasks include: FL = flanker; WHSP = whisper; GW = gift wrap; GD = gift delay. The orange scale indicates the importance of the taxa, with darker color referring to increased importance. The importance was determined by the generalized cross-validation value, with a larger value change indicating more contribution of a taxon to the model, i.e., which taxon is more important. As not all taxa appeared in the top 20 list at each time point, these absent taxa are colored in gray. Numbers on the left side of the figure show how many times a taxon appeared to be in the top 20 list of a task over time. The frequently appearing taxa are bolded and colored in orange (five times), yellow (four times), or green (three times).

Table 2. Associations of the gut microbiota in the first three years of life with behavioral measures at age three

Behavior at age three	Age of the gut microbiota	Genus	Estimate	Estimate error	95% CI	Prevalence %	Relative abundance % mean ± SD
Mother−reported							
CBCL_Internalizing	12w	Intestinibacter	0.23	0.12	[0.01, 0.47]	16	0.2 ± 0.6
	3y	Barnesiella	0.31	0.12	[0.08, 0.55]	50	0.4 ± 0.6
CBCL_Externalizing	2w	Streptococcus	0.26	0.12	[0.01, 0.5]	94	8.8 ± 10.1
	2w	Parabacteroides	− 0.3	0.12	[− 0.51, −0.06]	35	1.8 ± 4.9
	1y	Clostridium sensu stricto 1	0.23	0.11	[0.01, 0.46]	62	1.2 ± 3.5
	1y	Butyricicoccus	0.23	0.12	[0.01, 0.48]	56	0.4 ± 0.5
	1y	Parabacteroides	− 0.22	0.12	[− 0.45, −0.01]	44	0.8 ± 1.8
	3y	Barnesiella	0.33	0.12	[0.1, 0.56]	50	0.4 ± 0.6
SDQ_Internalizing	1y	Ruminococcus 2	−0.36	0.11	[− 0.58, −0.14]	39	1.4 ± 2.4
	3y	Bifidobacterium	0.27	0.13	[0.01, 0.53]	100	14.6 ± 11.6
	3y	Blautia	−0.25	0.12	[− 0.48, 0]	100	11.1 ± 4.5
	3y	[Ruminococcus] torques group	−0.25	0.13	[− 0.51, −0.01]	84	0.8 ± 0.7
	3y	Sutterella	0.25	0.13	[0.01, 0.5]	61	0.3 ± 0.3
SDQ_Externalizing	2w	Enterobacteriaceae unidentified genus	0.25	0.12	[0.01, 0.5]	89	21.6 ± 24.1
	2w	Parabacteroides	− 0.28	0.12	[− 0.51, −0.05]	35	1.8 ± 4.9
	6w	Halomonas	0.28	0.11	[0.06, 0.5]	11	0.1 ± 0.2
	3y	Butyricicoccus	− 0.35	0.12	[− 0.57, −0.11]	89	0.4 ± 0.3
	3y	Bifidobacterium	0.27	0.13	[0.01, 0.52]	100	14.6 ± 11.6
	3y	Oscillibacter	0.28	0.12	[0.04, 0.51]	22	0 ± 0.1
BRIEF − P	2w	Streptococcus	0.4	0.12	[0.15, 0.64]	94	8.8 ± 10.1
	6w	Halomonas	0.24	0.12	[0.01, 0.49]	11	0.1 ± 0.2
	12w	Streptococcus	0.31	0.12	[0.07, 0.55]	88	5.1 ± 10.4
	12w	Intestinibacter	0.3	0.11	[0.08, 0.53]	16	0.2 ± 0.6
	1y	Ruminococcus 2	−0.3	0.12	[− 0.54, −0.08]	39	1.4 ± 2.4
	1y	Clostridium sensu stricto 1	0.27	0.12	[0.03, 0.5]	62	1.2 ± 3.5
	3y	Blautia	−0.3	0.13	[− 0.57, −0.05]	100	11.1 ± 4.5
REEF	2w	Parabacteroides	0.25	0.12	[0, 0.47]	35	1.8 ± 4.9
	6w	Halomonas	− 0.24	0.12	[− 0.48, −0.01]	11	0.1 ± 0.2
	1y	Lachnospiraceae unidentified genus	0.28	0.11	[0.06, 0.5]	78	3.1 ± 4.1
	3y	[Ruminococcus] torques group	− 0.24	0.13	[− 0.49, −0.01]	84	0.8 ± 0.7
Inhibitory control tasks							
Flanker	6w	Bacteroides	0.28	0.12	[0.05, 0.51]	59	10.6 ± 16.3
	1y	Anaerostipes	−0.29	0.12	[− 0.51, −0.07]	96	3.8 ± 3.6
	1y	Sutterella	−0.24	0.12	[− 0.48, −0.01]	46	0.3 ± 0.6
	3y	Subdoligranulum	−0.25	0.13	[− 0.5, 0]	95	2.5 ± 1.8
	3y	Ruminococcaceae UCG − 013	0.26	0.12	[0.03, 0.51]	73	0.2 ± 0.2
Gift Wrap	1y	Subdoligranulum	− 0.31	0.12	[− 0.54, −0.08]	31	0.7 ± 1.4
	1y	Coprococcus 3	− 0.26	0.12	[− 0.48, −0.02]	14	0.1 ± 0.3
	1y	Veillonella	0.24	0.12	[0.01, 0.47]	71	3.3 ± 5
	1y	Lachnospiraceae NK4A136 group	− 0.3	0.12	[− 0.55, −0.08]	38	0.4 ± 0.7

Associations of estimates with 95% credible intervals (CIs) excluding 0 are presented. Behavioral measures include: CBCL = the child behavioral checklist; SDQ = the strengths and difficulties questionnaire; BRIEF − P = behavior rating inventory of executive functions − preschool; REEF = ratings of everyday executive functioning.

Table 3. Associations of alpha diversity in the first three years of life with problem behavior, executive functioning, and inhibitory controls at age three

Behavior at age three	Age of the gut microbiota	Alpha diversity	Estimate	Estimate error	95% CI
Mother-reported					
CBCL_Internalizing	2w	chao1	−0.28	0.12	[−0.51, −0.04]
REEF	2w	chao1	0.31	0.13	[0.07, 0.57]
Inhibitory control tasks					
Gift Wrap	2w	chao1	0.43	0.12	[0.19, 0.64]
	2w	PD	0.32	0.12	[0.08, 0.56]

Associations of estimates with 95% credible intervals (CIs) excluding 0 are presented. CBCL = the child behavioral checklist; REEF = ratings of everyday executive functioning.

(Wang et al., 2013) and this taxon was strikingly increased in patients with inflammatory bowel disease (Png et al., 2010).

With respect to microbial taxa that only showed age-specific associations, we observed higher relative abundances of *Bifidobacterium* at age three years related to more internalizing and externalizing behavior at the same age. Interestingly, a systematic review showed decreased *Bifidobacterium* in ASD children compared to neurotypically developing controls (Xu et al., 2019). Besides, supplementing ASD children with a prebiotic galacto-oligosaccharide increased *Bifidobacterium* populations in the gut and alleviated autistic symptoms (Grimaldi et al., 2017). However, opposite roles of *Bifidobacterium* have been described in major depressive disorder (MDD) (Cheung et al., 2019; Knudsen et al., 2021). Such inconsistency also takes place within ADHD studies. Two studies reported that *Bifidobacterium longum* mitigated ADHD (Finegold et al., 2010; Pärtty et al., 2015), while another study found overgrowing *Bifidobacterium* species in ADHD subjects (Aarts et al., 2017). In addition, in the current study, a higher relative abundance of *Blautia* at the age of three years was related to fewer internalizing difficulties (as well as to better executive functions at the same age). *Blautia* is suggested to play an important role in nutrient absorption and digestion (Eren et al., 2015), and in child gut microbiota development towards a normal adult-like configuration (Hsiao et al., 2014). In line with our findings, depleted *Blautia* was seen in ASD populations aged from two to 18 years old by several studies as concluded in a systematic review (Liu et al., 2019). However, elevated levels of *Blautia* were reported in relation to MDD in adults (Cheung et al., 2019) and ADHD in three-year-old children (Laue et al., 2021), indicating that different mechanisms may be involved depending on the psychopathology and chronological age. Lastly, we observed a positive relation between one unidentified *Enterobacteriaceae* genus at the age of two weeks and externalizing problems at the age of three years. Of interest, more *Enterobacteriaceae* species were cross-sectionally related to decreased cognitive functioning at the age of 45 months (Streit et al., 2021).

Another of our findings was that higher alpha diversity at age two weeks was linked to fewer internalizing problems and better executive functions at age three years. In accordance with our internalizing behavior result, Laue et al., observed that higher alpha diversity in the first two months of life was related to less internalizing behavior in three-year-old boys (Laue et al., 2021). Furthermore, van de Wouw et al., found lower alpha diversity in three-to-five-year-old children with clinically relevant CBCL cutoff scores for internalizing behavior (Van De Wouw et al., 2022). Besides, Eckermann et al., observed higher Shannon alpha diversity at the age of one, three, and four months in relation to better cognitive ability as measured by Digit Span forwards test at

the age of ten years, although not for Shannon alpha diversity at the age of six and ten years (Eckermann et al., 2022). Despite the generally weak or absent relations found between child gut microbiota and executive functions, the abovementioned studies may indicate that higher alpha diversity in the first years of life is related to improved subsequent mental outcomes at later ages. On the contrary, Carlson et al., found higher alpha diversity at one year of age related to worse cognition at two years of age in typically developing toddlers (Carlson et al., 2017). Additionally, a recent paper from van de Wouw et al. (2023) found evidence for a weak to modest cross-sectional relation between higher Shannon diversity and worse verbal comprehension in three- to four-year-old children (van de Wouw et al., 2023). Together, these findings illustrate the fact that we are yet to understand the potential impact of alpha diversity levels at different developmental stages. In general, alpha diversity levels of newborns start increasing immediately after birth due to colonization of microorganisms. With time, breastfed-infants tend to form a *Bifidobacterium*-predominated configuration which is often less diverse than formula-fed infants. After the introduction of solid food, child alpha diversity starts to increase, gradually reaching a steady state resembling gut microbiota composition of adults. Given these apparent normative fluctuations in levels of alpha diversity in the first months and years of life, having a comparatively high (or low) alpha diversity will potentially impact a child's development differently depending on the child's specific age.

Our study contributes to the growing body of literature on the gut microbiota, problem behavior, and executive functions. A strength of our study is the longitudinal design, which covered the period from birth to age three years and allowed for the assessment of multiple developmental stages of the gut microbiota. Another advantage was that questionnaires were filled in by both mother and partner. Partner reports were used for sensitivity analyses and because they showed positive correlations with maternal reports, they enhanced the reliability of our study measures. Furthermore, problem behavior and executive functions were assessed with multiple questionnaires (i.e., CBCL and SDQ for problem behavior, and BRIEF-P and REEF for executive functions), allowing us to determine conformity and consistency between various questionnaires. Finally, we used standardized behavioral tasks as a tool to objectively determine child executive functions.

A limitation of our study is the possible overreliance on the compositional features of the gut microbiota using relative abundances instead of absolute abundances. This approach may increase the chance of spurious associations as relative abundances are dependent on each other. Besides, 16S rRNA gene sequence data are limited at species-level resolution and profiling precise gene functions (Durazzi et al., 2021). Hence, although multiple

associations were identified, it is worth noting that these relations do not indicate causality. The results await follow-up studies, preferably preclinical experimental studies, to determine if individual microbes (e.g., aforementioned *Streptococcus* and *Bifidobacterium*) and microbial communities as a whole influence behavior and cognition, e.g., by generating neurotransmitters (e.g., GABA and serotonin) and their precursors (e.g., tryptophan and phenylalanine) in the gut (Altaib et al., 2021; Barandouzi et al., 2022; Biederman & Spencer, 1999; Gizer et al., 2009; Kandel et al., 2000; Kanehisa et al., 2022; Staller & Faraone, 2007). Further research including quantitative PCR, whole-genome shotgun metagenomic sequencing, targeted fecal metabolomics, and experimental studies in animal models, would improve the understanding of current correlational results and provide insight into microbial functions and even causality. An analytic limitation of our study is LEfSe, which was used to identify differentially abundant taxa in this study and has recently been pointed out to have higher sensitivity to false positive rates compared to other microbial composition analyses, such as ALDEx2, ANCOM − II, and DESeq2 (Nearing et al., 2022). Due to such methodological limitations, LEfSe-based significant findings should be carefully validated in future studies. Another limitation of our study is the relatively small sample size and mostly highly educated study population, limiting the generalizability of the findings. The restricted sample size may also hamper deep inference with respect to taxa with low prevalence rates to some degree. Our findings on microbial relations to the mental outcomes need to be confirmed in a larger, more representative cohort. Also, the participants were highly educated, which may hamper translating of our findings to individuals with a lower socioeconomic status. Lastly, recent developments in longitudinal analytical approaches (Kodikara et al., 2022), such as zero-inflated beta regression models, block bootstrap methods, and SplinectomeR, will better facilitate the identification of differentially abundant microbial taxa between groups (e.g., below and above clinical cutoffs) over time.

To conclude, our results provide tentative evidence supporting the idea that in a child's first years of life, the gut microbiota might play a vital role in the development of the brain, in line with the early-life programing theory (Tarry-Adkins & Ozanne, 2011). Potential mechanisms are likely related to microbiota-derived metabolites (Ahmed et al., 2022). As the nature of this study was exploratory and the body of similar research needs to grow to a large extent, it is still premature to translate our correlational findings into clinical implications. Replications in other longitudinal studies on healthy community children are necessary to confirm our findings. Ideally, to avoid inconsistent results caused by different methods used, replication studies should apply the same methodology regarding microbiology, genomics, epidemiology, and statistics (Ou et al., 2024). This will shed more light on key microbial taxa and latent pathways of associations between early gut microbiota and child behavior and cognition.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0954579423001402.

Author contribution. Yvonne Willemsen and Yangwenshan Ou contributed equally.

References

Aarts, E., Ederveen, T. H. A., Naaijen, J., Zwiers, M. P., Boekhorst, J., Timmerman, H. M., Smeekens, S. P., Netea, M. G., Buitelaar, J. K., Franke, B., van Hijum, S. A. F. T., & Arias Vasquez, A. (2017). Gut microbiome in ADHD and its relation to neural reward anticipation. *PLOS ONE*, *12*(9), e0183509. https://doi.org/10.1371/journal.pone.0183509

Aatsinki, A.-K., Kataja, E.-L., Munukka, E., Lahti, L., Keskitalo, A., Korja, R., Nolvi, S., Häikiö, T., Tarro, S., Karlsson, H., & Karlsson, L. (2020). Infant fecal microbiota composition and attention to emotional faces. *Emotion*, *22*(6), 1159–1170. https://doi.org/10.1037/emo0000924

Achenbach, T. M. (1966). The classification of children's psychiatric symptoms: A factor-analytic study. *Psychological Monographs*, *80*(7), 1–37. https://doi.org/10.1037/H0093906

Achenbach, Thomas M., & Ruffle, T. M. (2000). The child behavior checklist and related forms for assessing behavioral/emotional problems and competencies. *Pediatrics in Review*, *21*(8), 265–271. https://doi.org/10.1542/pir.21-8-265

Ahmed, H., Leyrolle, Q., Koistinen, V., Kärkkäinen, O., Layé, S., Delzenne, N., & Hanhineva, K. (2022). Microbiota-derived metabolites as drivers of gut-brain communication. *Gut Microbes*, *14*(1), 2102878. https://doi.org/10.1080/19490976.2022.2102878

Altaib, H., Nakamura, K., Abe, M., Badr, Y., Yanase, E., Nomura, I., & Suzuki, T. (2021). Differences in the concentration of the fecal neurotransmitters GABA and glutamate are associated with microbial composition among healthy human subjects. *Microorganisms*, *9*(2), 378. https://doi.org/10.3390/MICROORGANISMS9020378

Auteri, M., Zizzo, M. G., & Serio, R. (2015). GABA and GABA receptors in the gastrointestinal tract: From motility to inflammation. *Pharmacological Research*, *93*, 11–21. https://doi.org/10.1016/J.PHRS.2014.12.001

Barandouzi, Z. A., Lee, J., del Carmen Rosas, M., Chen, J., Henderson, W. A., Starkweather, A. R., & Cong, X. S. (2022). Associations of neurotransmitters and the gut microbiome with emotional distress in mixed type of irritable bowel syndrome. *Scientific Reports*, *12*(1), 1648. https://doi.org/10.1038/S41598-022-05756-0

Belk, A., Xu, Z. Z., Carter, D. O., Lynne, A., Bucheli, S., Knight, R., & Metcalf, J. L. (2018). Microbiome data accurately predicts the postmortem interval using random forest regression models. *Genes*, *9*(2), 104. https://doi.org/10.3390/GENES9020104

Biederman, J., & Spencer, T. (1999). Attention-deficit/hyperactivity disorder (adhd) as a noradrenergic disorder. *Biological Psychiatry*, *46*(9), 1234–1242. https://doi.org/10.1016/S0006-3223(99)00192-4

Bundgaard-Nielsen, C., Knudsen, J., Leutscher, P. D. C., Lauritsen, M. B., Nyegaard, M., Hagstrøm, S., & Sørensen, S. (2020). Gut microbiota profiles of autism spectrum disorder and attention deficit/hyperactivity disorder: A systematic literature review. *Gut Microbes*, *11*(5), 1172–1187. https://doi.org/10.1080/19490976.2020.1748258

Bürkner, P. C. (2017). Brms: An R package for bayesian multilevel models using stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Bürkner, P. C. (2018). Advanced bayesian multilevel modeling with the R package brms. *R Journal*, *10*(1), 395–411. https://doi.org/10.32614/rj-2018-017

Buuren, S.van (2021). *mice: Multivariate imputation by chained equations.*

Carlson, A. L., Xia, K., Azcarate-Peril, M. A., Goldman, B. D., Ahn, M., Styner, M. A., Thompson, A. L., Geng, X., Gilmore, J. H., & Knickmeyer, R. C. (2017). Infant gut microbiome associated with cognitive development. *Biological Psychiatry*, *83*(2), 148–159. https://doi.org/10.1016/j.biopsych.2017.06.021

Cheung, S. G., Goldenthal, A. R., Uhlemann, A. C., Mann, J. J., Miller, J. M., & Sublette, M. E. (2019). systematic review of gut microbiota and major depression. *Frontiers in Psychiatry*, *10*. https://doi.org/10.3389/FPSYT.2019.00034

Cinelli, C., Forney, A., & Pearl, J. (2020). A crash course in good and bad controls. *SSRN Electronic Journal*, 1–10. https://doi.org/10.2139/ssrn.3689437

Clarke, G., O'Mahony, S. M., Dinan, T. G., & Cryan, J. F. (2014). Priming for health: Gut microbiota acquired in early life regulates physiology, brain and behaviour. *Acta Paediatrica*, *103*(8), 812–819. https://doi.org/10.1111/APA.12674

Cryan, J. F., O'riordan, K. J., Cowan, C. S. M., Sandhu, K. V., Bastiaanssen, T. F. S., Boehme, M., Codagnone, M. G., Cussotto, S., Fulling, C., Golubeva, A. V., Guzzetta, K. E., Jaggar, M., Long-Smith, C. M., Lyte, J. M., Martin, J. A., Molinero-Perez, A., Moloney, G., Morelli, E., Morillas, E., . . . Dinan, T. G. (2019). The microbiota-gut-brain axis. *Physiological Reviews*, *99*(4), 1877–2013. https://doi.org/10.1152/physrev.00018.2018

Dalile, B., Van Oudenhove, L., Vervliet, B., & Verbeke, K. (2019). The role of short-chain fatty acids in microbiota-gut–brain communication. *Nature Reviews Gastroenterology & Hepatology*, *16*(8), 461–478. https://doi.org/10.1038/s41575-019-0157-3

de Weerth, C. (2017). Do bacteria shape our development? Crosstalk between intestinal microbiota and HPA axis. *Neuroscience & Biobehavioral Reviews*, *83*, 1–14. https://doi.org/10.1016/j.neubiorev.2017.09.016

Derrien, M., Alvarez, A.-S., & de Vos, W. M. (2019). The gut microbiota in the first decade of life. *Trends in Microbiology*, *27*(12), 997–1010. https://doi.org/10.1016/j.tim.2019.08.001

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Durazzi, F., Sala, C., Castellani, G., Manfreda, G., Remondini, D., & De Cesare, A. (2021). Comparison between 16S rRNA and shotgun sequencing data for the taxonomic characterization of the gut microbiota. *Scientific Reports*, *11*(1), 3030. https://doi.org/10.1038/S41598-021-82726-Y

Eckermann, H. A., Ou, Y., Lahti, L., & de Weerth, C. (2022). Can gut microbiota throughout the first 10 years of life predict executive functioning in childhood? *Developmental Psychobiology*, *64*(3), 1–14. https://doi.org/10.1002/dev.22226

Eren, A. M., Sogin, M. L., Morrison, H. G., Vineis, J. H., Fisher, J. C., Newton, R. J., & McLellan, S. L. (2015). A single genus in the gut microbiome reflects host preference and specificity. *The ISME Journal*, *9*(1), 90–100. https://doi.org/10.1038/ISMEJ.2014.97

Erus, G., Battapady, H., Satterthwaite, T. D., Hakonarson, H., Gur, R. E., Davatzikos, C., & Gur, R. C. (2015). Imaging patterns of brain development and their relationship to cognition. *Cerebral Cortex (New York, N.Y. : 1991)*, *25*(6), 1676–1684. https://doi.org/10.1093/CERCOR/BHT425

Finegold, S. M., Dowd, S. E., Gontcharova, V., Liu, C., Henley, K. E., Wolcott, R. D., Youn, E., Summanen, P. H., Granpeesheh, D., Dixon, D., Liu, M., Molitoris, D. R., & Green, J. A. (2010). Pyrosequencing study of fecal microflora of autistic and control children. *Anaerobe*, *16*(4), 444–453. https://doi.org/10.1016/J.ANAEROBE.2010.06.008

Gizer, I. R., Ficks, C., & Waldman, I. D. (2009). Candidate gene studies of ADHD: A meta-analytic review. *Human Genetics*, *126*(1), 51–90. https://doi.org/10.1007/s00439-009-0694-x

Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586. https://doi.org/10.1111/j.1469-7610.1997.tb01545.x

Goodman, R., & Scott, S. (1999). Comparing the strengths and difficulties questionnaire and the child behavior checklist: Is small beautiful? *Journal of Abnormal Child Psychology*, *27*(1), 17–24. https://doi.org/10.1023/a:1022658222914

Grimaldi, R., Cela, D., Swann, J. R., Vulevic, J., Gibson, G. R., Tzortzis, G., & Costabile, A. (2017). In vitro fermentation of B-GOS: Impact on faecal bacterial populations and metabolic activity in autistic and non-autistic children. *FEMS Microbiology Ecology*, *93*(2), fiw233. https://doi.org/10.1093/FEMSEC/FIW233

Gu, F., Borewicz, K., Richter, B., van der Zaal, P. H., Smidt, H., Buwalda, P. L., & Schols, H. A. (2018). In vitro fermentation behavior of isomalto/Malto-polysaccharides using human fecal inoculum indicates prebiotic potential. *Molecular Nutrition & Food Research*, *62*(12), 1800232. https://doi.org/10.1002/mnfr.201800232

Hechler, C., Beijers, R., Riksen-Walraven, J. M., & de Weerth, C. (2018). Are cortisol concentrations in human breast milk associated with infant crying? *Developmental Psychobiology*, *60*(6), 639–650. https://doi.org/10.1002/dev.21761

Hsiao, A., Ahmed, A. M. S., Subramanian, S., Griffin, N. W., Drewry, L. L., Petri, W. A., Haque, R., Ahmed, T., & Gordon, J. I. (2014). Members of the human gut microbiota involved in recovery from Vibrio cholerae infection. *Nature*, *515*(7527), 423–426. https://doi.org/10.1038/NATURE13738

Hyland, N. P., & Cryan, J. F. (2010). A gut feeling about GABA: Focus on GABAB receptors. *Frontiers in Pharmacology*, *1*, 124. https://doi.org/10.3389/FPHAR.2010.00124

Kandel, E., Schwartz, J., Jessell, T., & Siegelbaum, S. (2000). *Principles of neural science*. McGraw-Hill companies. https://www.academia.edu/download/30536508/neuroscience_syllabus.pdf

Kanehisa, M., Sato, Y., & Kawashima, M. (2022). KEGG mapping tools for uncovering hidden features in biological data. *Protein Science*, *31*(1), 47–53. https://doi.org/10.1002/PRO.4172

Kembel, S. W. (2020). *picante: Integrating phylogenies and ecology*.

Kendall, M. G. (1945). The treatment of ties in ranking problems. *Biometrika*, *33*(3), 239–251. https://doi.org/10.1093/BIOMET/33.3.239

Knudsen, J. K., Bundgaard-Nielsen, C., Hjerrild, S., Nielsen, R. E., Leutscher, P., & Sørensen, S. (2021). Gut microbiota variations in patients diagnosed with major depressive disorder—A systematic review. *Brain and Behavior*, *11*(7), e02177. https://doi.org/10.1002/BRB3.2177

Kodikara, S., Ellul, S., & Lê Cao, K.-A. (2022). Statistical challenges in longitudinal microbiome data analysis. *Briefings in Bioinformatics*, *23*(4), 1–18. https://doi.org/10.1093/BIB/BBAC273

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., Benesty, M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang, Y., Candan, C., Hunt, T., & Team, R. C. (2020). *caret: Classification and regression training*.

Lange, K. L., Little, R. J. A., & Taylor, J. M. G. (1989). Robust statistical modeling using the t distribution. *Journal of the American Statistical Association*, *84*(408), 881. https://doi.org/10.2307/2290063

Laue, H. E., Karagas, M. R., Coker, M. O., Bellinger, D. C., Baker, E. R., Korrick, S. A., & Madan, J. C. (2021). Sex-specific relationships of the infant microbiome and early-childhood behavioral outcomes. *Pediatric Research*, *92*(2), 1–12. https://doi.org/10.1038/s41390-021-01785-z

Leclercq, S., Mian, F. M., Stanisz, A. M., Bindels, L. B., Cambier, E., Ben-Amram, H., Koren, O., Forsythe, P., & Bienenstock, J. (2017). Low-dose penicillin in early life induces long-term changes in murine gut microbiota, brain cytokines and behavior. *Nature Communications*, *8*(1), 1–12. https://doi.org/10.1038/ncomms15062

Li, G., Song, B., Wang, C., Tang, D., Li, K., He, X., & Cao, Y. (2022). Diet, microbe, and autism: Cause or consequence? *Cell Host & Microbe*, *30*(1), 5–7. https://doi.org/10.1016/J.CHOM.2021.12.018

Liu, F., Li, J., Wu, F., Zheng, H., Peng, Q., & Zhou, H. (2019). Altered composition and function of intestinal microbiota in autism spectrum disorders: A systematic review. *Translational Psychiatry*, *9*(1), 43. https://doi.org/10.1038/s41398-019-0389-6

Loughman, A., Ponsonby, A.-L., O'Hely, M., Symeonides, C., Collier, F., Tang, M. L. K., Carlin, J., Ranganathan, S., Allen, K., Pezic, A., Saffery, R., Jacka, F., Harrison, L. C., Sly, P. D., Vuillermin, P., & Group, B. I. S. I. (2020). Gut microbiota composition during infancy and subsequent behavioural outcomes. *EBioMedicine*, *52*, 102640. https://doi.org/10.1016/j.ebiom.2020.102640

Louppe, G. (2014). *Understanding random forests: From theory to practice*. Cornell University.

Mazzoli, R., & Pessione, E. (2016). The neuro-endocrinological role of microbial glutamate and GABA signaling. *Frontiers in Microbiology*, *7*, 1934. https://doi.org/10.3389/FMICB.2016.01934/XML/NLM

Namkung, J. (2020). Machine learning methods for microbiome studies. *Journal of Microbiology*, *58*(3), 206–216. https://doi.org/10.1007/S12275-020-0066-8

Nearing, J. T., Douglas, G. M., Hayes, M. G., MacDonald, J., Desai, D. K., Allward, N., Jones, C. M. A., Wright, R. J., Dhanani, A. S., Comeau, A. M., & Langille, M. G. I. (2022). Microbiome differential abundance methods produce different results across 38 datasets. *Nature Communications*, *13*(1), 1–16. https://doi.org/10.1038/s41467-022-28034-z

Nilsen, E. S., Huyder, V., McAuley, T., & Liebermann, D. (2017). Ratings of everyday executive functioning (REEF): A parent-report measure of preschoolers' executive functioning skills. *Psychological Assessment*, *29*(1), 50–64. https://doi.org/10.1037/pas0000308

O'Mahony, S. M., Felice, V. D., Nally, K., Savignac, H. M., Claesson, M. J., Scully, P., Woznicki, J., Hyland, N. P., Shanahan, F., Quigley, E. M., Marchesi, J. R., O'Toole, P. W., Dinan, T. G., & Cryan, J. F. (2014). Disturbance of the gut microbiota in early-life selectively affects visceral pain in adulthood without impacting cognitive or anxiety-related behaviors in male rats. *Neuroscience*, *277*, 885–901. https://doi.org/10.1016/J. NEUROSCIENCE.2014.07.054

Oksanen, J. F., Simpson, G. L., Guillaume Blanchet, F., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., . . . Weedon, J. (2020). *vegan: Community ecology package*.

Ou, Y., Belzer, C., Smidt, H., & de Weerth, C. (2022). Development of the gut microbiota in healthy children in the first ten years of life: Associations with internalizing and externalizing behavior. *Gut Microbes*, *14*(1), 2038853. https://doi.org/10.1080/19490976.2022.2038853/SUPPL_FILE/KGMI_ A_2038853_SM3188.ZIP

Ou, Y., Belzer, C., Smidt, H., & de Weerth, C. (2023). Development of the gut microbiota in the first 14 years of life and its relations to internalizing and externalizing difficulties and social anxiety during puberty. *European Child & Adolescent Psychiatry*. https://doi.org/10.1007/s00787-023-02205-9

Ou, Y., Belzer, C., Smidt, H., & de Weerth, C. (2024). Methodological recommendations for human microbiota-gut-brain axis research. *Microbiome Research Reports*, *3*(1), 1. http://dx.doi.org/10.20517/mrr.2023. 33

Paradis, E. (2020). *ape: Analyses of phylogenetics and evolution*.

Pärtty, A., Kalliomäki, M., Wacklin, P., Salminen, S., & Isolauri, E. (2015). A possible link between early probiotic intervention and the risk of neuropsychiatric disorders later in childhood: A randomized trial. *Pediatric Research*, *77*(6), 823–828. https://doi.org/10.1038/PR.2015.51

Png, C. W., Lindén, S. K., Gilshenan, K. S., Zoetendal, E. G., McSweeney, C. S., Sly, L. I., McGuckin, M. A., & Florin, T. H. J. (2010). Mucolytic bacteria with increased prevalence in IBD mucosa augment in vitro utilization of mucin by other bacteria. *The American Journal of Gastroenterology*, *105*(11), 2420–2428. https://doi.org/10.1038/AJG.2010.281

Poncheewin, W., Hermes, G. D. A., van Dam, J. C. J., Koehorst, J. J., Smidt, H., & Schaap, P. J. (2020). NG-tax 2.0: A semantic framework for high-throughput amplicon analysis. *Frontiers in Genetics*, *10*, 1–12. https://doi. org/10.3389/fgene.2019.01366

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1), D590–D596. https://doi.org/10.1093/nar/ gks1219

Ramiro-Garcia, J., Hermes, G. D. A., Giatsis, C., Sipkema, D., Zoetendal, E. G., Schaap, P. J., & Smidt, H. (2018). NG-tax, a highly accurate and validated pipeline for analysis of 16S rRNA amplicons from complex biomes. *F1000Research*, *5*, 1791. https://doi.org/10.12688/ f1000research.9227.2

Revelle, W., & Condon, D. M. (2019). Reliability from α to ω: A tutorial. *Psychological Assessment*, *31*(12), 1395–1411. https://doi.org/10.1037/ pas0000754

Rice, D., & Barone, S. (2000). Critical periods of vulnerability for the developing nervous system: Evidence from humans and animal models. *Environmental Health Perspectives*, *108*(Suppl 3), 511–533. https://doi.org/ 10.1289/EHP.00108S3511

Rothenberg, S. E., Chen, Q., Shen, J., Nong, Y., Nong, H., Trinh, E. P., Biasini, F. J., Liu, J., Zeng, X., Zou, Y., Ouyang, F., & Korrick, S. A. (2021). Neurodevelopment correlates with gut microbiota in a cross-sectional analysis of children at 3 years of age in rural China. *Scientific Reports*, *11*(1), 1–11. https://doi.org/10.1038/s41598-021-86761-7

Schoemaker, K., Bunte, T., Espy, K. A., Deković, M., & Matthys, W. (2014). Executive functions in preschool children with ADHD and DBD: An 18-month longitudinal study. *Developmental Neuropsychology*, *39*(4), 302–315. https://doi.org/10.1080/87565641.2014.911875

Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., & Huttenhower, C. (2011). Metagenomic biomarker discovery and explanation. *Genome Biology*, *12*(6), R60. https://doi.org/10.1186/gb-2011-12-6-r60

Sherman, E. M. S., & Brooks, B. L. (2010). Behavior rating inventory of executive function – preschool version (BRIEF-P): Test review and clinical guidelines for use. *Child Neuropsychology*, *16*(5), 503–519. https://doi.org/10. 1080/09297041003679344

Silva, Y. P., Bernardi, A., & Frozza, R. L. (2020). The role of short-chain fatty acids from gut microbiota in gut-brain communication. *Frontiers in Endocrinology*, *11*, 25. https://doi.org/10.3389/FENDO.2020.00025

Staller, J. A., & Faraone, S. V. (2007). Expert review of neurotherapeutics targeting the dopamine system in the treatment of attention-deficit/ hyperactivity disorder targeting the dopamine system in the treatment of attention-deficit/hyperactivity disorder. *Expert Review of Neurotherapeutics*, *7*(4), 351–362. https://doi.org/10.1586/14737175.7.4.351

Stilling, R. M., Ryan, F. J., Hoban, A. E., Shanahan, F., Clarke, G., Claesson, M. J., Dinan, T. G., & Cryan, J. F. (2015). Microbes & neurodevelopment – absence of microbiota during early life increases activity-related transcriptional pathways in the amygdala. *Brain, Behavior, and Immunity*, *50*, 209–220. https://doi.org/10.1016/J.BBI.2015.07.009

Streit, F., Prandovszky, E., Send, T., Zillich, L., Frank, J., Sabunciyan, S., Foo, J., Sirignano, L., Lange, B., Bardtke, S., Hatfield, G., Witt, S. H., Gilles, M., Rietschel, M., Deuschle, M., & Yolken, R. (2021). Microbiome profiles are associated with cognitive functioning in 45-month-old children. *Brain, Behavior, and Immunity*, *98*, 151–160. https://doi.org/10.1016/J.BBI.2021.08.001

Tamana, S. K., Tun, H. M., Konya, T., Chari, R. S., Field, C. J., Guttman, D. S., Becker, A. B., Moraes, T. J., Turvey, S. E., Subbarao, P., Sears, M. R., Pei, J., Scott, J. A., Mandhane, P. J., & Kozyrskyj, A. L. (2021). Bacteroides-dominant gut microbiome of late infancy is associated with enhanced neurodevelopment. *Gut Microbes*, *13*(1), 1–17. https://doi.org/10.1080/ 19490976.2021.1930875

Tarry-Adkins, J. L., & Ozanne, S. E. (2011). Mechanisms of early life programming: Current knowledge and future directions. *The American Journal of Clinical Nutrition*, *94*(6 Suppl), S1765–S1771. https://doi.org/10. 3945/AJCN.110.000620

Thursby, E., & Juge, N. (2017). Introduction to the human gut microbiota. *Biochemical Journal*, *474*(11), 1823–1836. https://doi.org/10.1042/ BCJ20160510

van de Wouw, M., Rojas, L., Vaghef-Mehrabani, E., Wang, Y., Fichter, C., Workentine, M. L., Dewey, D., Arrieta, M. C., Reimer, R. A., Tomfohr-Madsen, L., & Giesbrecht, G. F. (2023). Exploring associations between the gut microbiota and full-scale intelligence in preschool children. *Neuroscience Letters*, *810*, 137357. https://doi.org/10.1016/J.NEULET.2023.137357

Van De Wouw, M., Wang, Y., Workentine, M. L., Vaghef-Mehrabani, E., Dewey, D., Reimer, R. A., Tomfohr-Madsen, L., & Giesbrecht, G. F. (2022). Associations between the gut microbiota and internalizing behaviors in preschool children. *Psychosomatic Medicine*, *84*(2), 159–169. https://doi. org/10.1097/PSY.0000000000001026

Vos, W. M.de, Tilg, H., Hul, M. Van, & Cani, P. D. (2022). Gut microbiome and health: Mechanistic insights. *Gut*, *71*(5), 1020–1032. https://doi.org/10. 1136/GUTJNL-2021-326789

Wang, L., Christophersen, C. T., Sorich, M. J., Gerber, J. P., Angley, M. T., & Conlon, M. A. (2013). Increased abundance of sutterella spp. and ruminococcus torques in feces of children with autism spectrum disorder. *Molecular Autism*, *4*(1), 42. https://doi.org/10.1186/2040-2392-4-42

Wang, S., Harvey, L., Martin, R., van der Beek, E. M., Knol, J., Cryan, J. F., & Renes, I. B. (2018). Targeting the gut microbiota to influence brain development and function in early life. *Neuroscience and Biobehavioral Reviews*, *95*, 191–201. https://doi.org/10.1016/j.neubiorev.2018.09.002

Welberg, L. (2022). Diet drives altered microbiota in autism. *Nature Neuroscience*, *25*(1), 2–2. https://doi.org/10.1038/s41593-021-00990-7

Willemsen, Y., Beijers, R., Arias Vasquez, A., & de Weerth, C. (2021). Do breastfeeding history and diet quality predict inhibitory control at preschool age? *Nutrients*, *13*(8), 2752. https://doi.org/10.3390/nu13082752

Xu, M., Xu, X., Li, J., & Li, F. (2019). Association between gut microbiota and autism spectrum disorder: a systematic review and meta-analysis. *Frontiers in Psychiatry*, *10*, 473. https://doi.org/10.3389/fpsyt.2019.00473

Zhuang, L., Chen, H., Zhang, S., Zhuang, J., Li, Q., & Feng, Z. (2019). Intestinal microbiota in early life and its implications on childhood health. *Genomics, Proteomics and Bioinformatics*, *17*(1), 13–25. https://doi.org/10. 1016/j.gpb.2018.10.002