medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Microbiome profiles are associated with cognitive functioning in 45-month-old children

Fabian Streit[1*+], Emese Prandovszky[3*], Tabea Send[2], Lea Zillich[1], Josef Frank[1], Sarven Sabunciyan[3], Jerome Foo[1], Lea Sirignano[1], Bettina Lange[2], Svenja Bardtke[2], Glen Hatfield[3], Stephanie H Witt[1], Maria Gilles[2], Marcella Rietschel[1], Michael Deuschle[2], Robert Yolken[3]

[1]Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany; [2]Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany; [3] Stanley Division of Developmental Neurovirology, Department of Pediatrics. Johns Hopkins School of Medicine, Baltimore, MD, United States of America

*contributed equally

[+]corresponding author:

Dr. Fabian Streit
Department of Genetic Epidemiology in Psychiatry
Central Institute of Mental Health
J5, 68159 Mannheim
Germany
E-mail: fabian.streit@zi-mannheim.de

Keywords (up to 10): microbiome, cognition, preschoolers, brain, immune-gut axis

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Prenatal, perinatal, and postnatal factors have been shown to shape neurobiological functioning and alter the risk for mental disorders later in life. The gut microbiome is established early in life, and interacts with the brain via the brain-immune-gut axis. However, little is known about how the microbiome relates to early-life cognitive functioning in children. The present study, where the fecal microbiome of 380 children was characterized using 16s rDNA and metagenomic sequencing aimed to investigate the association between the microbiota and cognitive functioning of children at the age of 45 months measured with the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III).

Overall the microbiome profile showed a significant association with cognitive functioning. A strong correlation was found between cognitive functioning and the relative abundance of an unidentified genus of the family *Enterobacteriaceae*. Follow-up mediation analyses revealed significant mediation effects of the level of this genus on the association of maternal smoking during pregnancy and current cigarette smoking with cognitive function. Metagenomic sequencing of a subset of these samples indicated that the identified genus was most closely related to *Enterobacter asburiae*. Analysis of metabolic potential showed a nominally significant association of cognitive functioning with the microbial norspermidine biosynthesis pathway.

Our results indicate that alteration of the gut microflora is associated with cognitive functioning in childhood. Furthermore, they suggest that the altered microflora might interact with other environmental factors such as maternal cigarette smoking. Interventions directed at altering the microbiome should be explored in terms of improving cognitive functioning in young children.

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

The microbiome, a collection of genes of the multitude of microorganisms living in and on the human body, has been shown to play a crucial role in somatic health (1,2). Recently, its role in the etiology of mental disorders has been documented by a large body of research. The gut microflora has been demonstrated to interact with the brain (gut-brain axis), via neural (vagus nerve), endocrine, immune and humoral (cytokines, short-chain and long-chain fatty acids) reciprocal communication routes (3-5). Moreover, it has been implicated in mental health and has been linked to disorders such as schizophrenia, bipolar disorder, attention deficit hyperactivity disorder, and autism spectrum disorders (6-8). Modern sequencing approaches make it possible to investigate these dense microbial populations, such as those prevalent in the human gut (1). It has been found that the microbial composition varies between both individuals (9,10) and different body sites (1). The human microbiome undergoes major changes in the first 1–3 years of life (11,12) but then remains relatively stable on an individual level (9,13,14). However, it has yet to be clarified whether bacterial colonization of the human body starts in utero or during birth (15). During the birth process, infants are exposed to a variety of maternal and environmental microbiota. The microbiome can be influenced by various factors including the mode of delivery (16), breast vs. formula feeding (14,17), and nutrition in general as well as the use of medication, especially antibiotics (18-20).

While the first few years of life are crucial for the establishment of the microbiome, this period is also a critical phase for the development of neurobiological functioning and is likely to represent a critical phase in the etiology of neurodevelopmental psychiatric disorders (21). Cognitive functioning in children highly reflects healthy neurodevelopment and is related to a broad range of mental health outcomes (22,23). Several animal studies have demonstrated an interaction between the gut microbiome and cognitive functioning (24-26). Studies have also linked alterations in the microbiome composition to cognitive functioning in humans (27). It was recently reported that the microbiome of children measured at one year of age

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

correlated with cognitive functioning and brain volume at age one and two (28) an association which has been confirmed in other studies (29-31). However, only a few studies have discussed the relationship between the microbiome and cognitive functioning in older infants and young children, where mental health and cognitive outcomes are already more advanced and easier to evaluate. The pre-school time period between three and five years of age is of particular interest due to its importance in subsequent learning and cognitive development (32).

In young children many factors can contribute to changes in the microbiome and might subsequently have an impact on cognition later in life. For example, exposure to antibiotics —a major modulator of the microflora— in the first six months of life has been linked to lower cognitive abilities at 11 years of age (33). Additional factors that have been shown to affect the microflora during infancy including mode of delivery (Caesarean section vs. vaginal birth), type of feeding (breastfeeding vs. formula), gestational age, diet, and socioeconomic status (34-36). Maternal smoking is another major neurodevelopmental risk factor during pregnancy, and has previously been linked to reduced birth weight in this cohort (37) and in others (38). However, the effect of these factors on the microbiome and cognitive functioning in the pre-school aged children has not been extensively studied (39).

In the present study, we aimed to I) investigate the link between cognitive functioning and the intestinal microbiome at the age of 45 months, II) identify particular bacteria species and microbial metabolic pathways that are associated with cognitive performance, and III) test whether the most strongly associated covariate (smoking during pregnancy) influences theses associations.

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Methods

Study population

Study participants were part of an ongoing longitudinal study on pre-, peri-, and postnatal stress and offspring development and health (Pre-, Peri-, and Postnatal Stress: Epigenetic Impact on Depression; POSEIDON) (40-42). In total, 410 pregnant women were enrolled in the study around four to eight weeks prior to delivery. Inclusion criteria for the mothers were: age 16–45 years, German-speaking, and being the child's primary caregiver. Exclusion criteria were evidence of infection with hepatitis B, or hepatitis C or human immunodeficiency virus, current inpatient treatment due to a diagnosis of a psychiatric disorder, a lifetime diagnosis of schizophrenia or psychotic disorder, or a substance dependency other than nicotine during pregnancy. For children, the exclusion criteria were birth before a gestational age of 30 weeks, birth weight lower than 1,500 grams, or multiple births, as well as any congenital disease, malformation, or chromosomal abnormality. The study was registered in the German Clinical Trials Register (DRKS00006338), approved by the Ethics Committee of the Medical Faculty Mannheim of the University of Heidelberg and conducted in accordance with the Declaration of Helsinki. All families provided written informed consent.

To date, four waves of assessments have been carried out: T1 (during the third trimester of pregnancy), T2 (between day 1–3 after birth), T3 (at the age of 6 months), and T4 (at the age of 45 months). At T4, of the original 410 families, 302 remained, and the dropouts were replaced with 101 newly recruited families. For the newly recruited families, pre-, peri-, and early postnatal factors were coded based on parental reports or medical records. Details of the assessment procedure have been described previously (42-44).

Assessment included the following variables used in the present study: smoking during pregnancy assessed via self-report, Caesarean section (C-section) or natural birth, breastfeeding, age of the mother as well as sex, birth weight, and gestational age of the

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

child. Measures at T4 used in the present study included the child's age and BMI, use of antibiotic medication during the previous 6 months, current maternal smoking, multilingual upbringing, and the years of maternal education as a marker of socioeconomic status (45). Perceived stress ratings of the mothers using Perceived Stress Scale (PSS) (Cohen, Kamarck, & Mermelstein, 1983) were available during pregnancy for the original sample only, and at T4 for the original and the newly recruited sample.

Cognitive functioning was assessed using the Wechsler Preschool and Primary Scale of Intelligence - Third Edition (WPPSI-III) (46) at T4, the same time point at which stool sampling occurred. Five subtests of the WPPSI-III were conducted: the core subtests *Receptive Vocabulary*, *Information*, *Block Design*, and *Object Assembly*, and the supplementary subtest *Picture Naming*. For all subtests, the raw scores were converted to standardized scores based on the child's exact age. Full-scale IQ (FIQ) score was calculated from the four core subtests, verbal IQ (VIQ) from the *Receptive Vocabulary* and *Information* subtests, performance IQ (PIQ) from the *Block Design* and *Object Assembly* subtests, and the general language composite (GLC) from the *Receptive Vocabulary* and *Picture Naming* subtests. Subjects were included in the analysis if values were available for at least one of the WPPSI-III scales, resulting in the exclusion of 21 subjects. Prior to the statistical analyses, the test documentation was reevaluated as an additional quality control step. Data points for each subtest were excluded if, for example, an apparent lack of motivation or lack of understanding the instructions was evident. This reevaluation resulted in the exclusion of additional 14 subjects. Five more samples were excluded due to irregularities in shipping or handling (e.g. delayed shipping, extended storage at home before shipping).

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Microbiome analysis

Sampling and storage

Stool samples were collected by the parents at the age of 45 months (T4). The samples were collected using a standard stool sample kit (Sarstedt, Nuembrecht, Germany) on the days prior to or the day of the T4 appointment and temporarily stored in a home refrigerator until being delivered by the parents at ambient temperature to the lab on the day of the appointment. In circumstances when this was not achievable, stool samples were sent to the laboratory at ambient temperature by mail service, which usually took 1-2 days. The effect of the shipment on the microflora was analyzed and taken into account as noted below. Upon arrival, stool samples were frozen at -80°C and stored in the Biobank of Psychiatric Diseases Mannheim (BioPsy) (47) until further processing.

DNA extraction

Fecal samples were thawed and an aliquot of 100 mg was used to extract genomic DNA using the ZymoBIOMICS™ 96 MagBead DNA Kit according to the manufacturer's protocol. The protocol included a bead-beating step to enable complete homogenization/disruption of the microbial cell walls. ZymoBIOMICS™ Microbial Community Standard (Zymo Research Europe GmbH) was included in the DNA extraction process to ensure an accurate representation of microbial communities in samples (48) and to assess potential environmental contaminants. DNA quantity and quality were determined by Thermo Scientific NanoDrop™ 1000 Spectrophotometer and normalized for further analysis to 5 ng/µL.

Characterization of the Microbiome

Library preparation

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The extracted DNA was amplified using primers directed at conserved bacterial 16S ribosomal sequences in the v3–v4 region, and 380 individually indexed (Nextera XT, Illumina, San Diego, CA) libraries were generated following the Illumina 16S metagenomic library preparation protocol. Each library was quantified using Qubit (Life Technologies) and the concentration was normalized to 4 nM. A 5 µl aliquot of each library was pooled and denatured using 0.2 N NaOH and subsequently diluted to 10 pM using HT1 (hybridization buffer) supplied by Illumina (San Diego, CA). Prior to loading, the library was spiked with a 10% 20 pM PhiX control. Sequencing was performed on the MiSeq Illumina sequencing platform using a 2 × 300 PE sequencing reagent kit (Cat # MS-102-3003 Illumina, San Diego, CA).

Sequence Analysis

Raw paired-end reads were first de-multiplexed using 5′ index information, and adapter sequences were trimmed off prior access. Sequence reads were analyzed separately using the QIIME 2 (v 2018.11) platform (49). In addition to filtering out any remaining PhiX contaminants and chimeric sequences, the DADA-q2 plugin (50) was utilized to remove low-quality reads. All unique sequence reads were collected into a feature table, which was filtered to remove low-abundance features using the feature-table q2 plugin (51). Next, multiple sequence alignment was performed by employing the MAFFT alignment plugin (52) followed by computing a phylogenetic tree using the FastTree program (53). Sampling depth was set to 2263 sequence reads. Phylogenetic diversity analyses utilized this phylogenetic tree in order to calculate the following alpha diversity matrices: number of observed species, Shannon's and Faith's phylogenetic diversity, and Pileou's Evenness. To explore the bacterial composition of the samples, taxonomy was assigned to sequences using a pre-trained naive Bayes classifier (54) and q2-feature-classifier plugin (https://github.com/qiime2/q2-feature-classifier). This classifier was trained on Greengenes v13_8 99% OTUs, where the sequences have been trimmed to only include v3–v4 regions of the 16S rRNA gene bound by Illumina primer pairs. Analyses of taxa focused on the level

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

of genera since that is the taxonomic level that can be determined with the highest degree of certainty with this method (55). Prior to analysis, the number of counts for each genus were log-transformed using natural logarithm after adding 1, to allow for the analysis of samples with 0 counts.

Metagenomic Shotgun Sequencing

A subset of 39 samples and 1 blank were selected for metagenomic sequencing. The selection was based on a broad distribution of WPSSI-III full scale IQ score and *Enterobacteriaceae* genus abundance, including samples with both low and high values. After exclusion for invalid WPSSI values or irregularities in sample shipping or handling, data from 33 subjects was used for statistical analysis. Aliquots of 1ng of extracted fecal DNA from the selected samples were used to generate paired-end metagenomic libraries using Illumina Nextera XT Low Input DNA kit following the manufacturer's instructions (Nextera XT DNA Library Prep Kit Reference Guide (15031942)). The concentration of each library was measured by using Quibit and normalized to 4nM. Libraries were then sequenced using the Illumina HiSeq 4000, using the sequencing core facility at University of Maryland. Sequence reads with homology to human sequences were removed according to the selected downstream pipeline. The remaining sequence reads were then mapped to microbial databases employing Centrifuge (v1.0.3) (56) as previously described. Centrifuge uses a novel indexing scheme based on the Burrows-Wheeler transform (BWT) and the Ferragina-Manzini (FM) index optimized specifically for the metagenomic classification problem. Centrifuge output was visualized using the R-based application Pavian (https://ccb.jhu.edu/software/pavian/) (57). The metabolic potential of the microbiome was characterized using the program humann2 (v1.7) as described elsewhere (58). For association analysis with cognition, we selected a set of pathways, I.) focusing on the pathway abundances at community level, II.) filtering hits present in the blank sample, or present in more than 37 and less than 5 samples and III.) with a variance at least 3 times

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

higher than the average abundance. This resulted in a list of 122 pathways (see Supplementary Table S6).

Statistical Analyses

Restricted maximum likelihood (REML) models were computed to assess the association of the overall microbiome profile with the cognitive measures as implemented in the software package omic-data-based complex trait analysis (OSCA v.0.45) (59). This approach applies REML to assess the overall association of the microbiome profile with the tested variables, expressed as the share of explained variance (59). Corresponding 95% confidence intervals (CI) were calculated with bootstrapping as implemented in FIESTA v1.0 (61). In a second step, potentially confounding covariates for those analyses were identified in separate REML model for each confounding variable: maternal age, sex of the child, infant feeding (formula vs. breast), gestational age, birth weight, antibiotic exposure during the 6 months prior to T4 (no vs. yes), type of birth (vaginal delivery vs. C-section), maternal cigarette smoking during pregnancy (none vs. any), current maternal cigarette smoking at T4 (none vs. any), child age at T4, child BMI at T4, multilingual upbringing at T4, level of maternal education (in years of education), and shipment vs. delivery by parents. Subsequently, the cognition REML models were rerun, adding confounding variables which showed a significant association with the microbiome profile as covariates.

Further analyses were conducted with the R open-source environment for statistical computation and graphics (v3.5.1) (R Development Core Team, 2013). Regression analyses were used to explore bivariate associations between the WPPSI-III full-scale IQ score and the following microbiome measures: alpha diversity measures, relative abundance of specific genera, and predicted metabolic pathways. Standardized betas (*ß)* and their 95% confidence intervals (CI) are given. The resulting *p* values were corrected for multiple comparisons (indicated as *q*-values) using Bonferroni correction for the number of tested diversity measures (q = 0.05/4 = 0.0125) or genera (q = 0.05/82 = 6.1×10^{-4}). Initially,

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

regressions were computed without controlling for covariates. Subsequently, significant (q<.05) bivariate associations of diversity measures or specific genera with the WPPSI-III FIQ score were explored by adding potential confounding variables as covariates in the regression model if they were associated with the respective diversity measure or genus at the nominal level (p<.05, same variables as in the REML models). Additional regression analyses were performed to test the association between the taxa and pathways identified by metagenomic sequencing and the WPPSI-III subscales. Additionally, we explored whether the maternal perceived stress during pregnancy related to the main findings, by testing the association of the PSS with the associated microbiome measures. Subjects with missing data for a WPPSI-III scale or covariates were excluded from the respective analyses (pairwise deletion).

Mediation analysis

Mediation analysis was used to investigate the hypothesis that the occurrence of an identified taxa mediates the effect of the most strongly associated covariates (i.e. smoking during pregnancy or current maternal smoking, see results) on the WPPSI-III full-scale IQ score, using quasi-Bayesian Monte Carlo simulation with 10,000 simulations as implemented in the mediation package.

Results

Study population

Microflora profiling was successful in 364 out of 380 samples (Figure 1). For the subsequent statistical analyses one subject with an outlying gestational age, 35 subjects with a lack of valid measures for any of the WPPSI-III scales, and 5 with irregularities in shipping/storage

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

were excluded, resulting in a total of 323 subjects. For the remaining families, demographics (maternal age, child age, sex), pre-, peri- and postnatal factors (gestational age, birth weight, smoking, C-section, breastfeeding, maternal education, multilingual upbringing, child's BMI, antibiotics use), and cognition (WPSSI-III FIQ, VIQ, PIQ, GLC) are depicted in Table 1.

Variance in cognition explained by the microbiome

After quality and content filtering (removing environmental contaminants and keeping genera that present at least in 10 samples), 82 genera remained and were utilized in subsequent downstream analyzes. The variance explained in the WPPSI-III scales by the overall microbiome was characterized by restricted maximum likelihood (REML) models. These analyses showed that a significant proportion of the variance of WPPSI-III FIQ (variance explained = 13.3%, p=.0010) was association with the overall composition of the microbiome. Furthermore, the microbiome was associated with the subscales WPPSI-III VIQ and WPPSI-III GLC (variance explained = 15.8%; 10.3%, p=.000027; p=.0014 respectively) but not of WPPSI-III PIQ (0.5%; p = .46; Table 2). Of the assessed covariates, smoking during pregnancy, current smoking, breastfeeding, maternal education, multilingual upbringing, and shipment showed a significant association with the microbiome profile in REML models (see Supplementary Table S3). When those variables were jointly added as covariates to the REML cognition models, the explained variance of the WPPSI-III scales by variation in the overall composition of the microbiome variables was substantially reduced, with only the WPPSI-III VIQ scale remaining significant (see Supplementary Table S4). When each of these covariates was added individually to the REML model of WPPSI-III FIQ, the strongest reduction in explained variance was observed for smoking during pregnancy (see Supplementary Table S5).

Significant negative correlations were observed between alpha diversity metrics and cognition. As shown in Table 3, the associations of WPPSI-III FIQ with Faith's phylogenetic diversity index remained significant after correction for multiple testing. We observed

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

significant (p<0.05) negative correlations of antibiotic use with Shannon diversity index and Pileou's Evenness and of gestation age with Shannon diversity index and number of observed species, and a positive correlation of C-section with Faith's phylogenetic diversity index (see Supplementary Figure S1). When we adjusted for these covariates in the association analyses with the respective diversity measures, we observed similar results, however the associations did not remain significant after correction for multiple testing (Supplementary Table S2).

As shown in Figures 2 and 3, relative abundance of one genus - an unidentified genus within the family *Enterobacteriaceae* were significantly associated with lower WPPSI-III FIQ (*ß*=-.20 [-.31; -0.095], *p*=.00031, *q*=.025, Supplementary Figure S2) after correction for the 82 tested genera. Relative abundance of this genus was also associated with lower performance on the WPPSI-III VIQ (*ß*=-19 [-.30; -.083] *p*=. 0.000636, *q*=.052), PIQ (*ß*=-.14 [-.25; -.031], *p*=.012, *q*=1), and GLC (*ß*=-.21 [-.32; -.10], *p*=.00015, *q*=.012) scores (Figure 2 A-D). Explorative analyses revealed an additional genus, Eubacterium genus of the Erysipelotrichaceae family (*ß*=-.20 [-.31; -.094], *p*=.0003047; *q*=.025) being associated with WPPSI-III verbal IQ subscales (*q*<.05).

We subsequently carried out additional analyses with the unidentified genus within the family *Enterobacteriaceae.* It was nominally associated with following covariates (p<.05): maternal age (*ß*=.13 [-.24; -.024], *p*=.018), shipment (*ß*=-.18 [.070; .28], *p*=.0014), and more pronounced with maternal cigarette smoking during pregnancy (*ß*=.22 [.12; .33], *p*=0.0000505, Figure 3B) and smoking at T4 (*ß*=.18 [.077; .29], *p*=0.00088). The relationship between the unidentified genus within the family *Enterobacteriaceae* and WPPSI-III FIQ remained nominally significant when those variables were employed as covariates (*ß*=-.14 [-.25; -.026], *p*=.016, *q* = 1) in a regression model. As the variables most strongly associated with the genus were maternal smoking during pregnancy and current (T4) maternal smoking, those were further explored in mediation models. Mediation analyses (Fig 3.) revealed a significant mediation effect of this genus in the family Enterobacteriaceae for the effects of maternal smoking during pregnancy (proportion mediated = 13.7% [3.4%; 32%]; *p*=.007) and

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

current smoking (proportion mediated = 12.0% [3.1%; 29%]; p=.0034) on WPPSI-III FIQ (controlling for age of the mother and shipment). Maternal smoking during T4 and maternal smoking during pregnancy were strongly associated (r = .64, p < 0.001), and a joint model predicting *Enterobacteriaceae* levels with both smoking variables indicated significant effects of pregnancy smoking (*ß*=-.18 [.039; .33], *p*=.013) but not for current smoking (*ß*=-.071 [-.072; .21], *p*=.33).

We further characterized the microbiome by metagenomic analysis performed in a subset of 33 samples (demographics details of subset are shown in Supplemental Table S1C). This analysis indicated that the organisms which belongs to this unidentified genus of the family *Enterobacteriaceae* were most closely related to *Enterobacter asburiae* (*ß*=.87 [.70; 1.0], $p=3.8x10^{-11}$, Figure 2E), and were also related to *Enterobacter cloacae* (*ß*=.76 [.54; .99], $p=2.4x10^{-7}$) and *Kluyvera intermedia* (*ß*=.51 [.20; .81], *p*=.0026). Of those organism, *Enterobacter cloacae* was significantly associated with the WPPSI-III FIQ, and *Enterobacter cloacae and Enterobacter asburiae* were associated with the WPPSI-III PIQ subscale.

In the subjects for which the data was available (T1: n=240, T4: n=320), we additionally explored whether the associations of WPPSI FIQ with Faith's phylogenetic diversity index and the Enterobacteriaceae genus were affected by perceived stress during pregnancy and current perceived stress. The PSS at T1 was not significantly associated with Faith's phylogenetic diversity index (r=-.11; *p*=.074) but showed a nominal association with the Enterobacteriaceae genus species (r=.15; *p*=.022). The PSS at T4 was neither associated with the genus nor Faith’s phylogenetic diversity metric (p>.31). When PSS T1 was added as an additional predictor to the model with maternal age, shipment, smoking during pregnancy, and smoking at T4, the relationship between the unidentified genus within the family Enterobacteriaceae and WPPSI-III FIQ remained nominally significant (ß=-.16 [-.29; -.032], *p*=.015).

We also employed metabolic profiling to explore the relationship between cognitive functioning and the metabolic potential of the microbiome explored by metagenomic

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

sequencing (details of metagenomics sequencing subset shown in Supplementary Table S1C). We found a nominally significant negative association between the WPPSI-III FIQ and the predicted level of norspermidine after adjusting for the covariates smoking during pregnancy, present cigarette smoking, maternal age, and shipment (*ß*=-.47 [-.82; -.12], *p*=.014; Figure 2F; Details of the components of this pathway are shown in Supplementary Figure S3, all other pathways p > 0.05).

Discussion

The present prospective cohort study revealed an association between the composition of the fecal bacterial microflora and cognitive functioning as measured by the WPPSI-III in healthy children at 45 months of age.

Overall, we found that a significant portion of the difference in cognitive functioning of these children as measured by the WPPSI-III is correlated with variance in the composition of their fecal microflora. Of interest, this correlation was observed for the FIQ and the language-related subscales (VIQ & GLC), but not for the PIQ subscale of the WPPSI-III.

Additionally, we found evidence of a negative association of cognition measures with alpha diversity, with a higher diversity being associated with lower cognitive scores. Again, the significant associations were limited to the FIQ and the language-related subscales (VIQ & GLC). Such associations between the composition of the bacterial microflora and cognitive functioning are consistent with the findings of Carlson and colleagues (28). They measured the fecal microbiome in children at the age of one and assessed cognitive functioning (Mullen scales of early learning) at the age of one and two years. They identified three distinct bacterial clusters, where significant differences between the clusters were associated with cognitive functioning as measured by Mullen scales of early learning by both overall score and by subdomain performance at the age of two years. As observed in the current study, they found negative correlations in alpha diversity measures, with higher diversity

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

being associated with lower cognitive functioning scores, however, no association remained significant after correction for multiple testing.

At the genus level, we found that lower levels of cognitive functioning in preschool-aged children were associated with an increased level of DNA of an unidentified genus of the family Enterobacteriaceae. In previous studies, supportive evidence links exposure to similar organisms to altered cognitive functioning in aging humans (62) and in animal models of aging (63). Notably, antibiotics usage, which has been linked to lower cognitive abilities later in childhood (33) showed no correlation with this genus. The mechanisms of these associations are still unidentified but are likely to involve alterations of immune activity within the brain via alterations of the gut-immune-brain axis (64,65). Metagenomic sequencing of a subset of samples indicated that the associated taxon was closely related to *Enterobacter asburiae* and *Enterobacter cloacae*. These are highly homologous organisms which inhabit the gastrointestinal tract of infants and young children and are occasionally associated with symptomatic infections (66,67). It is yet to be clarified whether this taxon represents a novel species related to both *Enterobacter asburiae* and *Enterobacter cloacae* or a mixture of several related species within the *Enterobacter cloacae* complex (68). Additional metagenomic sequencing will be required to further characterize these organisms and determined mechanisms by which they might be related to cognitive functioning.

The finding of a potential association between lower levels of cognitive functioning and norspermidine synthesis is also of interest. Norspermidine is a polyamine generated by bacteria that is structurally and functionally similar to spermidine, a polyamine synthesized in humans and other animals (69). Spermidine and other polycationic polyamines have been shown to improve cognitive functioning in animals and are being investigated for the prevention of cognitive decline in aging humans (70). The mechanism of action is uncertain but likely to be related to the biological functions of polycationic molecules, including modulation of the activity of ionic channels, protein synthesis, protein kinases, and cell

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

proliferation and apoptosis. (71). Loss of function mutations in spermidine synthase are associated with decreased levels of cognitive functioning in Snyder-Robinson syndrome (72). Animal model studies indicate that the mechanism of the cognitive dysfunction in this syndrome is likely related to lysosomal dysfunction and oxidative stress (73). However, as these analyses were based on a comparatively small subset of the cohort, and this association only reached nominal significance ($p < 0.05$), additional studies will be required to conform this potential association and to define precise mechanisms of association.

Previous studies have indicated that cognition in early childhood is impaired by maternal smoking (74). The results of the mediation analyses in the present study indicate that the effects of both smoking during pregnancy and current smoking on cognition might be (partially) mediated by the microbiome. Maternal smoking has been linked to adverse outcomes in both somatic and mental health of the offspring (75). Both animal and human studies indicate possible effects of smoking on the microbiome. For example, in rats, smoking was associated with reductions in *Bifidobacteria* (76), and side-stream smoking increased the levels of *Clostridium* and reduced the levels of *Firmicutes* (*Lactococci* and *Reminococcus*) and *Enterobacteriacae* families in mice (77). In humans, it has been shown, that smokers and non-smokers differ in regard to their microbiome (e.g. 78,79) and that changes in smoking behavior affect the microbiome (80). As the mother's microbiome has been shown to influence the microbiome of the child (15), it is plausible that cigarette smoking during pregnancy could influence the child's microbiome profile as well. Notably, prenatal smoking showed a stronger link to the *Enterobacteriaceae* genus than current smoking in the present study.

There are several possible caveats to our findings. The significant correlation between altered level of this unidentified genus of the Enterobacteriaceae family and cognitive functioning might be due to a direct connection between the microbiome and brain development. However, it is also possible that, the presence of this genus might be related

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

to unmeasured concomitant factors, such as altered diet, family structure, and social interactions (62). Among the above listed concomitant factors, diet is quite unlikely, since all children except one were attending national daycare where they have been exposed to a very similar diet. Several postnatal factors were not investigated in the present study, including household size, exposure to pets and other animals, travel, urban vs. suburban living, and additional lifestyle factors (e.g. diet or parental alcohol consumption), which might be associated with the microbiome, smoking and cognitive functioning. The present study indicates that such covariates can influence observed associations of outcomes with the microbiome. The potential role of these factors as mediators between the microbiome and cognitive development should be the subject of additional studies. While we did not observe a strong influence of measures of prenatal and current perceived stress on the association of the microbiome with cognition, the direct associations of pre-, peri- and postnatal stress measures with the microbiome should be assessed in more detail in the future. Another limitation of this present study was that in the case of the subset of subjects who were newly recruited at T4, environmental factors from earlier time points such as smoking during pregnancy or breastfeeding could only be assessed retrospectively by self-report. However, these variables were taken into account as binary variable in our analyses, hence the retrospective self-report is unlikely to introduce strong bias. Another limitation of our study is that our microbial data collection is cross-sectional, precluding studies of temporal variation. Additional longitudinal studies are planned to confirm our findings

Our study documents an association between intestinal microflora composition and cognitive functioning, with the microbiome mediating the effect of maternal smoking on cognition in a group of otherwise healthy preschool-aged children. The microbiome of a child can be altered by environmental factors, such as diet, antibiotics, and exposure to cigarette smoke, as well as therapeutically by administration of probiotic or prebiotic medications (81). Follow-up studies need to test the possibility of a direct interaction between the microbiome and cognitive functioning in this age group. The establishment of a cause-and-effect association

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

may lead to the evaluation of interventions directed at improving cognitive functioning in infants and young children.

Acknowledgements

We thank all parents and children for taking part in this study and our student employees and interns for their support with data acquisition and data entry. We thank Regev Schweiger (FIESTA) and Jian Yang (OSCA) for feedback on the application of their software.

Funding

Stanley Medical Research Institute. National Institute of Mental Health (P50) Silvio O Conte Center at Johns Hopkins. ERA-NET NEURON grants “EMBED” (01EW1904), and “SynSchiz—Linking synaptic dysfunction to disease mechanisms in schizophrenia—a multilevel investigation” (01EW1810 to MR). German Research Foundation grants RI908/11-2 and WI3439/3-2.

Financial Disclosures

Michael Deuschle participated as principal investigator in studies of Johnson & Johnson and received speaker fees of AbbVie. The other authors report no biomedical financial interests or potential conflicts of interest.

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

References

1. Cho I, Blaser MJ (2012): The human microbiome: at the interface of health and disease. *Nat Rev Genet*. 13:260-270.
2. Warner BB (2018): The contribution of the gut microbiome to neurodevelopment and neuropsychiatric disorders. *Pediatr Res*.
3. Carabotti M, Scirocco A, Maselli MA, Severi C (2015): The gut-brain axis: interactions between enteric microbiota, central and enteric nervous systems. *Annals of gastroenterology: quarterly publication of the Hellenic Society of Gastroenterology*. 28:203.
4. Cryan JF, O'Riordan KJ, Cowan CS, Sandhu KV, Bastiaanssen TF, Boehme M, et al. (2019): The microbiota-gut-brain axis. *Physiological reviews*. 99:1877-2013.
5. Mayer EA, Tillisch K, Gupta A (2015): Gut/brain axis and the microbiota. *The Journal of clinical investigation*. 125:926-938.
6. Severance EG, Yolken RH (2018): Deciphering microbiome and neuroactive immune gene interactions in schizophrenia. *Neurobiology of disease*.
7. Valles-Colomer M, Falony G, Darzi Y, Tigchelaar EF, Wang J, Tito RY, et al. (2019): The neuroactive potential of the human gut microbiota in quality of life and depression. *Nature microbiology*. 4:623-632.
8. Sharon G, Cruz NJ, Kang DW, Gandal MJ, Wang B, Kim YM, et al. (2019): Human Gut Microbiota from Autism Spectrum Disorder Promote Behavioral Symptoms in Mice. *Cell*. 177:1600-1618.e1617.
9. Bogaert D, Keijser B, Huse S, Rossen J, Veenhoven R, van Gils E, et al. (2011): Variability and diversity of nasopharyngeal microbiota in children: a metagenomic analysis. *PLoS One*. 6:e17035.
10. Eckburg PB (2005): Diversity of the Human Intestinal Microbial Flora. *Science*. 308:1635-1638.
11. van Best N, Hornef MW, Savelkoul PHM, Penders J (2015): On the origin of species: Factors shaping the establishment of infant's gut microbiota. *Birth Defects Res C Embryo Today*. 105:240-251.
12. Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, et al. (2011): Succession of microbial consortia in the developing infant gut microbiome. *Proceedings of the National Academy of Sciences*. 108:4578-4585.
13. Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R (2009): Bacterial community variation in human body habitats across space and time. *Science*. 326:1694-1697.
14. Davis EC, Wang M, Donovan SM (2017): The role of early life nutrition in the establishment of gastrointestinal microbial composition and function. *Gut Mi crobes*. 8:143-171.
15. Mueller NT, Bakacs E, Combellick J, Grigoryan Z, Dominguez-Bello MG (2015): The infant microbiome development: mom matters. *Trends Mol Med*. 21:109-117.
16. Dominguez-Bello MG, Costello EK, Contreras M, Magris M, Hidalgo G, Fierer N, et al. (2010): Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci U S A*. 107:11971-11975.
17. Madan JC, Hoen AG, Lundgren SN, Farzan SF, Cottingham KL, Morrison HG, et al. (2016): Association of Cesarean Delivery and Formula Supplementation With the Intestinal Microbiome of 6-Week-Old Infants. *JAMA Pediatr*. 170:212-219.
18. Imhann F, Bonder MJ, Vich Vila A, Fu J, Mujagic Z, Vork L, et al. (2016): Proton pump inhibitors affect the gut microbiome. *Gut*. 65:740-748.
19. Langdon A, Crook N, Dantas G (2016): The effects of antibiotics on the microbiome throughout development and alternative approaches for therapeutic modulation. *Genome Med*. 8:39.
20. Rogers MAM, Aronoff DM (2016): The influence of non-steroidal anti-inflammatory drugs on the gut microbiome. *Clin Microbiol Infect*. 22:178.e171-178.e179.

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

21. Sharon G, Sampson TR, Geschwind DH, Mazmanian SK (2016): The central nervous system and the gut microbiome. *Cell*. 167:915-932.
22. Dickinson D (2014): Zeroing in on early cognitive development in schizophrenia. Am Psychiatric Assoc.
23. Koenen KC, Moffitt TE, Roberts AL, Martin LT, Kubzansky L, Harrington H, et al. (2009): Childhood IQ and adult mental disorders: a test of the cognitive reserve hypothesis. *American Journal of Psychiatry*. 166:50-57.
24. Fröhlich EE, Farzi A, Mayerhofer R, Reichmann F, Jačan A, Wagner B, et al. (2016): Cognitive impairment by antibiotic-induced gut dysbiosis: Analysis of gut microbiota-brain communication. *Brain Behav Immun*. 56:140-155.
25. Gareau MG, Wine E, Rodrigues DM, Cho JH, Whary MT, Philpott DJ, et al. (2011): Bacterial infection causes stress-induced memory dysfunction in mice. *Gut*. 60:307-317.
26. Savignac HM, Tramullas M, Kiely B, Dinan TG, Cryan JF (2015): Bifidobacteria modulate cognitive processes in an anxious mouse strain. *Behav Brain Res*. 287:59-72.
27. Sarkar A, Harty S, Lehto SM, Moeller AH, Dinan TG, Dunbar RI, et al. (2018): The microbiome in psychology and cognitive neuroscience. *Trends in cognitive sciences*. 22:611-636.
28. Carlson AL, Xia K, Azcarate-Peril MA, Goldman BD, Ahn M, Styner MA, et al. (2018): Infant Gut Microbiome Associated With Cognitive Development. *Biol Psychiatry*. 83:148-159.
29. Brett BE, de Weerth C (2019): The microbiota–gut–brain axis: A promising avenue to foster healthy developmental outcomes. *Developmental psychobiology*. 61:772-782.
30. Gao W, Salzwedel AP, Carlson AL, Xia K, Azcarate-Peril MA, Styner MA, et al. (2019): Gut microbiome and brain functional connectivity in infants-a preliminary study focusing on the amygdala. *Psychopharmacology*.1-11.
31. Sordillo JE, Korrick S, Laranjo N, Carey V, Weinstock GM, Gold DR, et al. (2019): Association of the infant gut microbiome with early childhood neurodevelopmental outcomes: an ancillary study to the VDAART randomized clinical trial. *JAMA network open*. 2:e190905-e190905.
32. Li Y, Grabell AS, Wakschlag LS, Huppert TJ, Perlman SB (2017): The neural substrates of cognitive flexibility are related to individual differences in preschool irritability: a fNIRS investigation. *Developmental cognitive neuroscience*. 25:138-144.
33. Slykerman RF, Coomarasamy C, Wickens K, Thompson JMD, Stanley TV, Barthow C, et al. (2019): Exposure to antibiotics in the first 24 months of life and neurocognitive outcomes at 11 years of age. *Psychopharmacology (Berl)*. 236:1573-1582.
34. Renson A, Herd P, Dowd JB (2020): Sick Individuals and Sick (Microbial) Populations: Challenges in Epidemiology and the Microbiome. *Annual Review of Public Health*. 41:63-80.
35. Kang SS, Jeraldo PR, Kurti A, Miller MEB, Cook MD, Whitlock K, et al. (2014): Diet and exercise orthogonally alter the gut microbiome and reveal independent associations with anxiety and cognition. *Molecular neurodegeneration*. 9:1-12.
36. Penders J, Thijs C, Vink C, Stelma FF, Snijders B, Kummeling I, et al. (2006): Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics*. 118:511-521.
37. Witt SH, Frank J, Gilles M, Lang M, Treutlein J, Streit F, et al. (2018): Impact on birth weight of maternal smoking throughout pregnancy mediated by DNA methylation. *BMC Genomics*. 19:290-290.
38. Banderali G, Martelli A, Landi M, Moretti F, Betti F, Radaelli G, et al. (2015): Short and long term health effects of parental tobacco smoking during pregnancy and lactation: a descriptive review. *Journal of translational medicine*. 13:1-7.
39. DeWeerdt S (2018): How baby's first microbes could be crucial to future health. *Nature*. 555:S18-s19.
40. Dukal H, Frank J, Lang M, Treutlein J, Gilles M, Wolf IA, et al. (2015): New-born females show higher stress- and genotype-independent methylation of SLC6A4 than males. *Borderline Personal Disord Emot Dysregul*. 2:8.

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

41. Nieratschker V, Massart R, Gilles M, Luoni A, Suderman MJ, Krumm B, et al. (2014): MORC1 exhibits cross-species differential methylation in association with early life stress as well as genome-wide association with MDD. *Transl Psychiatry*. 4:e429.
42. Send TS, Gilles M, Codd V, Wolf I, Bardtke S, Streit F, et al. (2017): Telomere Length in Newborns is Related to Maternal Stress During Pregnancy. *Neuropsychopharmacology*. 42:2407-2413.
43. Wolf IA-C, Gilles M, Peus V, Scharnholz B, Seibert J, Jennen-Steinmetz C, et al. (2017): Impact of prenatal stress on the dyadic behavior of mothers and their 6-month-old infants during a play situation: role of different dimensions of stress. *J Neural Transm*. 124:1251-1260.
44. Send T, Bardtke S, Gilles M, Wolf I, Sütterlin M, Kirschbaum C, et al. (2019): Stress reactivity in preschool-aged children: Evaluation of a social stress paradigm and investigation of the impact of prenatal maternal stress. *Psychoneuroendocrinology*. 101:223-231.
45. Dickerson FB, Boronow JJ, Stallings CR, Origoni AE, Cole S, Yolken RH (2004): Association between cognitive functioning and employment status of persons with bipolar disorder. *Psychiatric Services*. 55:54-58.
46. Petermann F, Lipsius M (2009): WPPSI-III. Wechsler Preschool and Primary Scale of Intelligence–Third Edition. Deutsche Version. Frankfurt: Pearson Assessment.
47. Witt S, Dukal H, Hohmeyer C, Radosavljevic-Bjelic S, Schendel D, Frank J, et al. (2016): Biobank of Psychiatric Diseases Mannheim – BioPsy. *Open Journal of Bioresources*. 3.
48. Costea PI, Zeller G, Sunagawa S, Pelletier E, Alberti A, Levenez F, et al. (2017): Towards standards for human fecal sample processing in metagenomic studies. *Nat Biotechnol*. 35:1069-1076.
49. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. (2019): Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature biotechnology*. 37:852-857.
50. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016): DADA2: high-resolution sample inference from Illumina amplicon data. *Nature methods*. 13:581.
51. McDonald D, Clemente JC, Kuczynski J, Rideout JR, Stombaugh J, Wendel D, et al. (2012): The Biological Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome. *Gigascience*. 1:2047-2217X-2041-2047.
52. Katoh K, Standley DM (2013): MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular biology and evolution*. 30:772-780.
53. Price MN, Dehal PS, Arkin AP (2010): FastTree 2–approximately maximum-likelihood trees for large alignments. *PloS one*. 5.
54. Werner JJ, Koren O, Hugenholtz P, DeSantis TZ, Walters WA, Caporaso JG, et al. (2012): Impact of training sets on classification of high-throughput bacterial 16s rRNA gene surveys. *The ISME journal*. 6:94-103.
55. Edgar RC (2018): Accuracy of taxonomy prediction for 16S rRNA and fungal ITS sequences. *PeerJ*. 6:e4652.
56. Kim D, Song L, Breitwieser FP, Salzberg SL (2016): Centrifuge: rapid and sensitive classification of metagenomic sequences. *Genome research*. 26:1721-1729.
57. Breitwieser FP, Salzberg SL (2016): Pavian: Interactive analysis of metagenomics data for microbiomics and pathogen identification. *BioRxiv*.084715.
58. Franzosa EA, McIver LJ, Rahnavard G, Thompson LR, Schirmer M, Weingart G, et al. (2018): Species-level functional profiling of metagenomes and metatranscriptomes. *Nature methods*. 15:962.
59. Zhang F, Chen W, Zhu Z, Zhang Q, Nabais MF, Qi T, et al. (2019): OSCA: a tool for omic-data-based complex trait analysis. *Genome biology*. 20:1-13.
60. Zhang F, Chen W, Zhu Z, Zhang Q, Nabais MF, Qi T, et al. (2019): OSCA: a tool for omic-data-based complex trait analysis. *Genome biology*. 20:107.

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

61. Schweiger R, Kaufman S, Laaksonen R, Kleber ME, Marz W, Eskin E, et al. (2016): Fast and Accurate Construction of Confidence Intervals for Heritability. *American journal of human genetics*. 98:1181-1192.
62. Liu P, Wu L, Peng G, Han Y, Tang R, Ge J, et al. (2019): Altered microbiomes distinguish Alzheimer's disease from amnestic mild cognitive impairment and health in a Chinese cohort. *Brain, behavior, and immunity*.
63. Chen J, Buchanan JB, Sparkman NL, Godbout JP, Freund GG, Johnson RW (2008): Neuroinflammation and disruption in working memory in aged mice after acute stimulation of the peripheral innate immune system. *Brain, behavior, and immunity*. 22:301-311.
64. Schaafsma W, Basterra LB, Jacobs S, Brouwer N, Meerlo P, Schaafsma A, et al. (2017): Maternal inflammation induces immune activation of fetal microglia and leads to disrupted microglia immune responses, behavior, and learning performance in adulthood. *Neurobiology of disease*. 106:291-300.
65. Severance EG, Yolken RH (2019): From infection to the microbiome: an evolving role of microbes in schizophrenia.
66. Mezzatesta ML, Gona F, Stefani S (2012): Enterobacter cloacae complex: clinical impact and emerging antibiotic resistance. *Future microbiology*. 7:887-902.
67. Dalben M, Varkulja G, Basso M, Krebs V, Gibelli M, Van der Heijden I, et al. (2008): Investigation of an outbreak of Enterobacter cloacae in a neonatal unit and review of the literature. *Journal of Hospital Infection*. 70:7-14.
68. Nazir F, Ibrahim M, Zaman G, Hussain A, Yar AM, Bo Z (2018): Genetic Diversity and Functional Analysis of Sigma Factors in Enterobacter cloacae Complex Resourced From Various Niche. *Evolutionary Bioinformatics*. 14:1176934318754878.
69. Pegg AE (2014): The function of spermine. *IUBMB life*. 66:8-18.
70. Schwarz C, Stekovic S, Wirth M, Benson G, Royer P, Sigrist SJ, et al. (2018): Safety and tolerability of spermidine supplementation in mice and older adults with subjective cognitive decline. *Aging (Albany NY)*. 10:19.
71. Guerra GP, Rubin MA, Mello CF (2016): Modulation of learning and memory by natural polyamines. *Pharmacological Research*. 112:99-118.
72. Peron A, Spaccini L, Norris J, Bova SM, Selicorni A, Weber G, et al. (2013): Snyder–Robinson syndrome: A novel nonsense mutation in spermine synthase and expansion of the phenotype. *American Journal of Medical Genetics Part A*. 161:2316-2320.
73. Li C, Brazill JM, Liu S, Bello C, Zhu Y, Morimoto M, et al. (2017): Spermine synthase deficiency causes lysosomal dysfunction and oxidative stress in models of Snyder-Robinson syndrome. *Nature communications*. 8:1257.
74. Tzoumakis S, Carr VJ, Dean K, Laurens KR, Kariuki M, Harris F, et al. (2018): Prenatal maternal smoking, maternal offending, and offspring behavioural and cognitive outcomes in early childhood. *Criminal behaviour and mental health : CBMH*. 28:397-408.
75. Lee SH, Weerasinghe WSP, van der Werf JH (2017): Genotype-environment interaction on human cognitive function conditioned on the status of breastfeeding and maternal smoking around birth. *Scientific reports*. 7:6087.
76. Tomoda K, Kubo K, Asahara T, Andoh A, Nomoto K, Nishii Y, et al. (2011): Cigarette smoke decreases organic acids levels and population of bifidobacterium in the caecum of rats. *J Toxicol Sci*. 36:261-266.
77. Wang H, Zhao J-X, Hu N, Ren J, Du M, Zhu M-J (2012): Side-stream smoking reduces intestinal inflammation and increases expression of tight junction proteins. *World J Gastroenterol*. 18:2180-2187.
78. Allais L, Kerckhof F-M, Verschuere S, Bracke KR, De Smet R, Laukens D, et al. (2016): Chronic cigarette smoke exposure induces microbial and inflammatory shifts and mucin changes in the murine gut. *Environ Microbiol*. 18:1352-1363.
79. Capurso G, Lahner E (2017): The interaction between smoking, alcohol and the gut microbiome. *Best Pract Res Clin Gastroenterol*. 31:579-588.
80. Biedermann L, Zeitz J, Mwinyi J, Sutter-Minder E, Rehman A, Ott SJ, et al. (2013): Smoking cessation induces profound changes in the composition of the intestinal microbiota in humans. *PloS one*. 8:e59260-e59260.

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

81. Robertson RC, Manges AR, Finlay BB, Prendergast AJ (2019): The Human Microbiome and Child Growth - First 1000 Days and Beyond. *Trends in microbiology*. 27:131-147.

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figures and tables

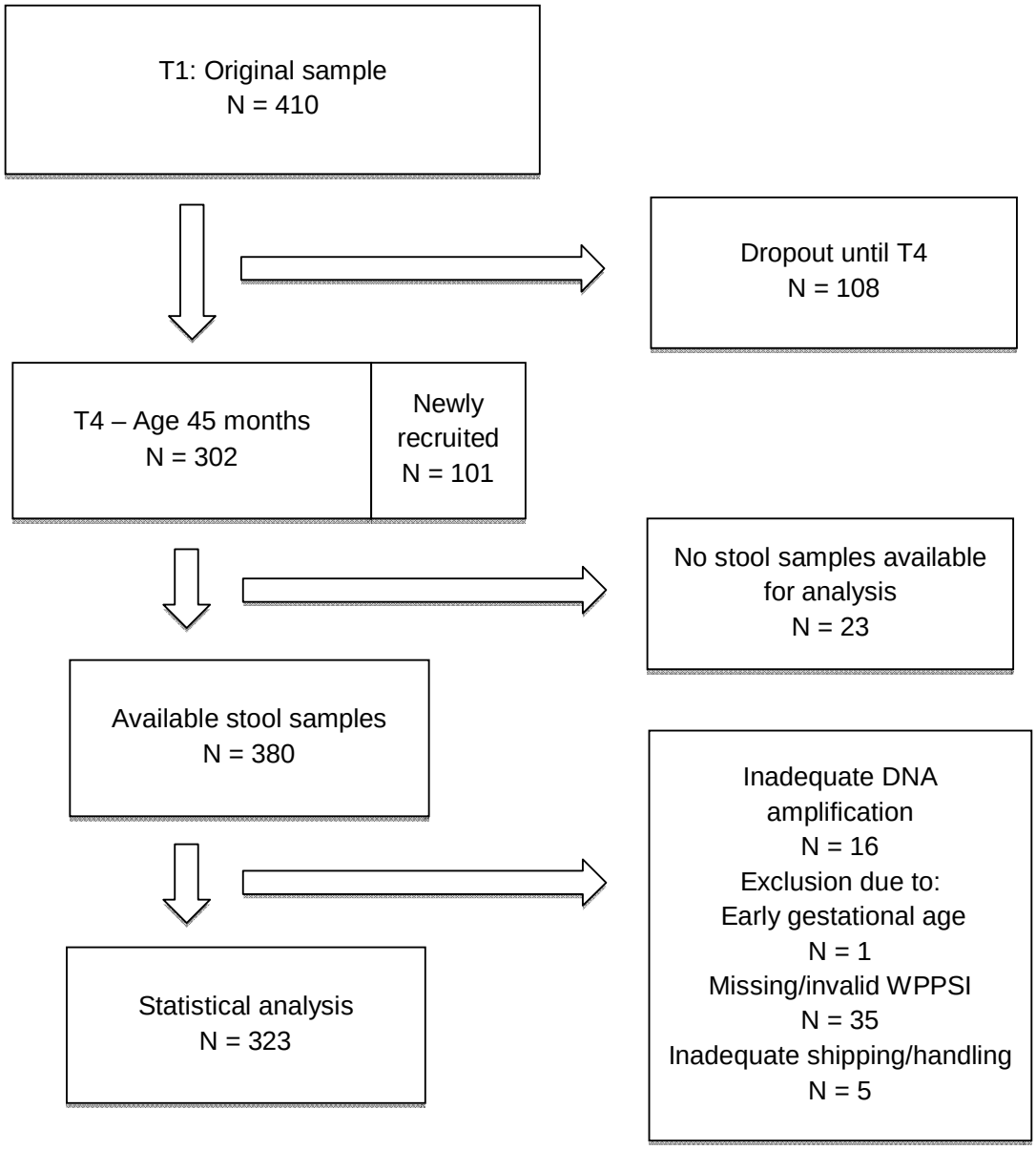

Figure 1: Consort flow diagram indicating the number of subjects participating in the POSEIDON study at the T1 and T4 time points, along with the number of subjects who provided stool samples and the number of samples that were included in the statistical analysis following successful library preparation and microbiome analyses.

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

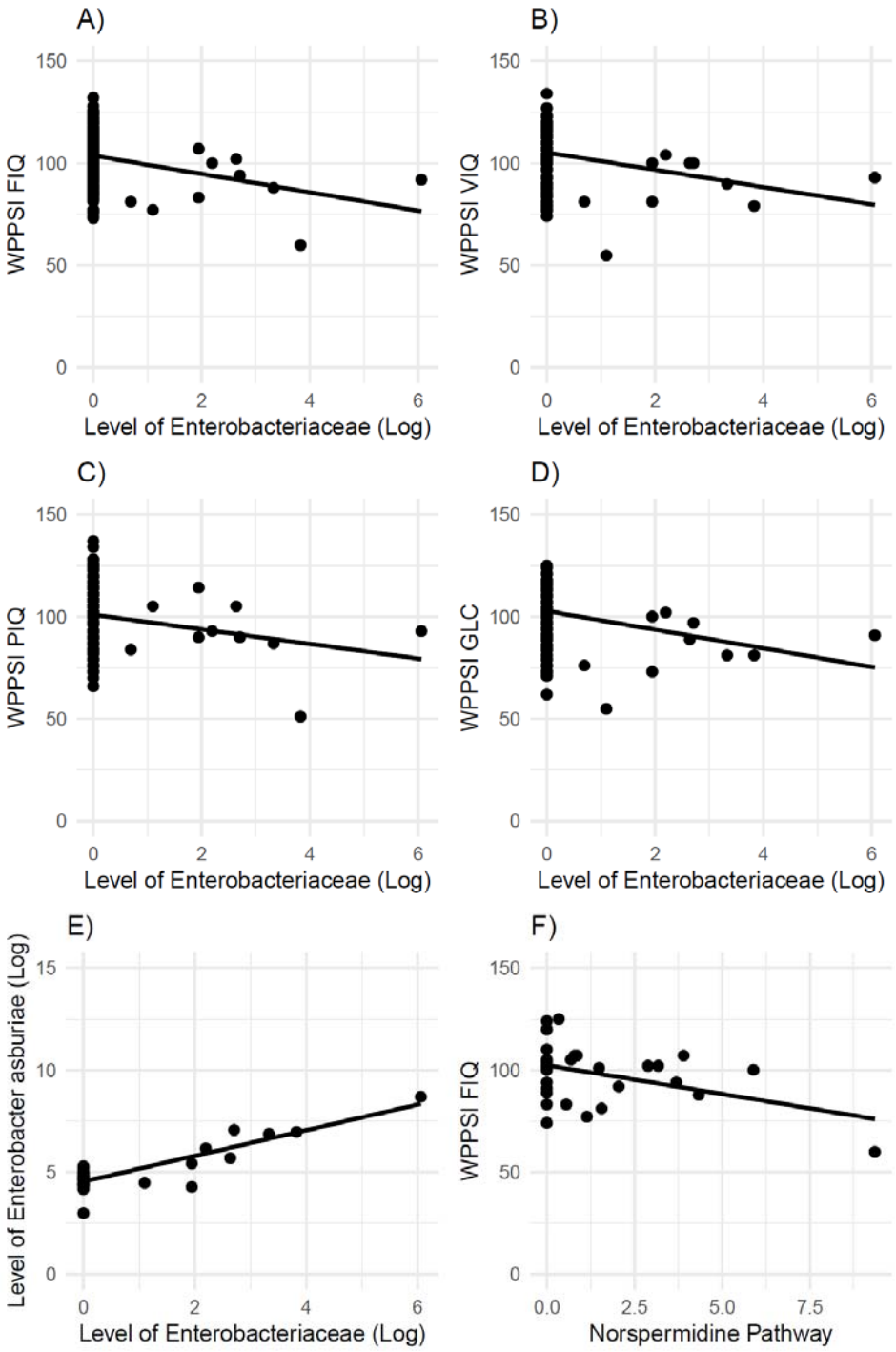

Figure 2: Correlation between the log transformed relative abundance of an unidentified genus within the family *Enterobacteriaceae* and A) the full scale IQ and B-D) subscales of the WPPSI-III, E) log transformed relative abundance of genus within the family *Enterobacteriaceae* with level of *Enterobacter asburiae* and F) WPPSI-III full-scale IQ score and predicted level of norspermidine biosynthesis. WPPSI-III = Wechsler Preschool and Primary Scale of Intelligence; FIQ = full-scale IQ; GLC = general language composite; PIQ = performance IQ; VIQ = verbal IQ

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

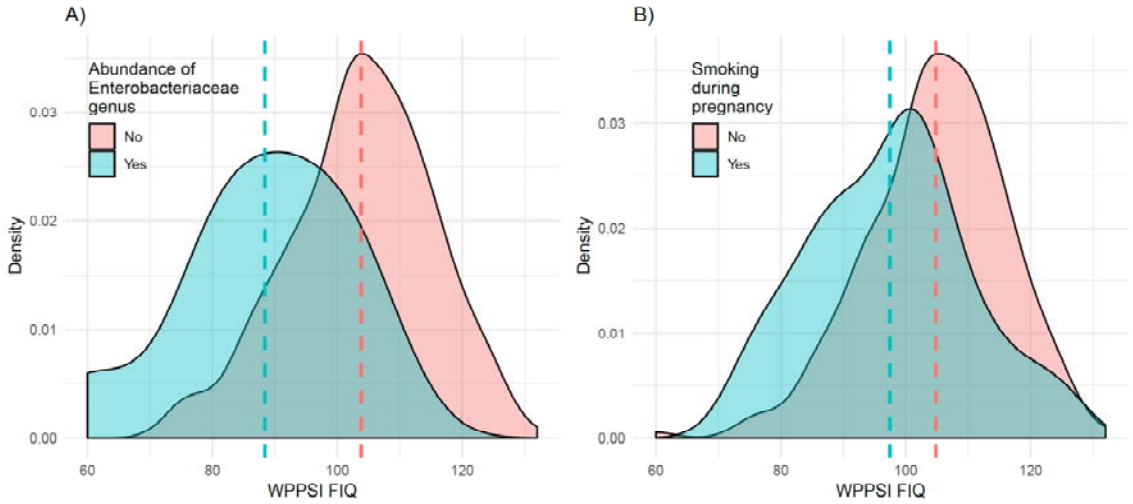

Figure 3: Distribution and mean of the full-scale WPPSI-III score A) by abundance of the genus within the family *Enterobacteriacea*e and B) by smoking during pregnancy. The group means are indicated by the dashed lines. WPPSI-III = Wechsler Preschool and Primary Scale of Intelligence Version III. FIQ = full-scale IQ

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

	Frequency / Mean (SD)	N
Maternal age (years)	31.33 (4.89)	323
Smoking during pregnancy	20.7%	323
Gestational age (weeks)	39.14 (1.25)	318
Birth weight (gram)	3,392 (468.48)	317
Current maternal smoking	24.1%	323
C-section	31.3%	323
Breastfeeding	83.2%	322
Sex (female)	53.2%	323
Antibiotics last 6 months	31.4%	322
Maternal education (years)	14.78 (2.23)	323
Multilingual upbringing	23.4%	321
Child age (months)	45.02 (1.00)	323
Child BMI	15.44 (1.35)	320
Perceived stress pregnancy	21.12 (8.36)	240
Current perceived stress	23.91 (8.02)	320
Shipment	18%	323
WPPSI-III FIQ	103.30 (11.81)	308
WPPSI-III VIQ	104.81 (11.72)	311
WPPSI-III PIQ	100.62 (13.48)	315
WPPSI-III GLC	102.33 (11.39)	320

Table 1: Information on pre-, peri-, and postnatal factors and cognition in the final study sample (n = 323). WPPSI-III = Wechsler Preschool and Primary Scale of Intelligence; SD = standard deviation; FIQ = full-scale IQ; GLC = general language composite; PIQ = performance IQ; VIQ = verbal IQ

medRxiv preprint doi: https://doi.org/10.1101/2021.06.26.21259573; this version posted July 22, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Phenotype	(V/G)Vp (CI)	*P*
WPPSI-III FIQ	13.4% (4.9%; 25.1%)	.0010
WPPSI-III VIQ	15.8% (7.1%; 29.4%)	.0000273
WPPSI-III PIQ	0.5% (0.0%; 12.4%)	.46
WPPSI-III GLC	10.3% (2.7%; 22.2%)	.0014

Table 2: Explained variance of cognition phenotypes by variation in the microbiome estimated using restricted maximum likelihood (REML) models. (V/G)Vp = explained variance; CI = 95% confidence interval; WPPSI-III = Wechsler Preschool and Primary Scale of Intelligence Version III; FIQ = full-scale IQ; GLC = general language composite; PIQ = performance IQ; VIQ = verbal IQ.

Phenotype	Shannon diversity index	*Pielou's evenness index*	Faith phylogenetic diversity index	*Observed OTUs*
WPPSI-III FIQ	-0.103 (-0.215;0.008)	-0.068 (-0.179;0.044)	-0.144** (-0.255;-0.033)	-0.101 (-0.212;0.011)
WPPSI-III VIQ	-0.131* (-0.242;-0.021)	-0.101 (-0.212;0.010)	-0.136* (-0.247;-0.026)	-0.112* (-0.223;-0.002)
WPPSI-III PIQ	-0.044 (-0.154;0.067)	-0.026 (-0.137;0.085)	-0.086 (-0.196;0.024)	-0.046 (-0.157;0.064)
WPPSI-III GLC	-0.129* (-0.238;-0.020)	-0.078 (-0.187;0.032)	-0.133* (-0.242;-0.024)	-0.129* (-0.238;-0.020)

Table 3: Associations of measures of cognition with measures of alpha diversity. Standardized betas and the 95% confidence interval are given; WPPSI-III = Wechsler Preschool and Primary Scale of Intelligence Version III; FIQ = full-scale IQ; GLC = general language composite; PIQ = performance IQ; VIQ = verbal IQ; * = $p<.05$, ** = $p < 0.0125$ (0.05 / 4)